Documenta Mathematica, Journal der Deutschen Mathematiker-Vereinigung, veröffentlicht Forschungsarbeiten aus allen mathematischen Gebieten und wird in traditioneller Weise referiert. Es wird indiziert durch Mathematical Reviews, Science Citation Index Expanded, Zentralblatt für Mathematik.

Artikel können als T\TeX-Dateien per E-Mail bei einem der Herausgeber eingereicht werden. Hinweise für die Vorbereitung der Artikel können unter der unten angegebenen WWW-Adresse gefunden werden.

Documenta Mathematica, Journal der Deutschen Mathematiker-Vereinigung, publishes research manuscripts out of all mathematical fields and is refereed in the traditional manner. It is indexed in Mathematical Reviews, Science Citation Index Expanded, Zentralblatt für Mathematik.

Manuscripts should be submitted as \TeX-files by e-mail to one of the editors. Hints for manuscript preparation can be found under the following web address.

http://www.math.uni-bielefeld.de/documenta

Geschäftsführende Herausgeber / Managing Editors:

Alfred K. Louis, Saarbrücken  
Ulf Rehmann (techn.), Bielefeld  
Peter Schneider, Münster

Herausgeber / Editors:

Christian Bär, Potsdam  
Don Blasius, Los Angeles  
Joachim Cuntz, Münster  
Patrick Delorme, Marseille  
Gavril Farkas, Berlin (HU)  
Edward Frenkel, Berkeley  
Friedrich Götz, Bielefeld  
Ursula Hamenstädt, Bonn  
Lars Hesselholt, Cambridge, MA (MIT)  
Max Karoubi, Paris  
Stephen Lichtenbaum  
Eckhard Meinrenken, Toronto  
Alexander S. Merkurjev, Los Angeles  
Anil Nerode, Ithaca  
Thomas Peternell, Bayreuth  
Eric Todd Quinto, Medford  
Takeshi Saito, Tokyo  
Stefan Schwede, Bonn  
Heinz Siedentop, München (LMU)  
Wolfgang Soergel, Freiburg

ISSN 1431-0635 (Print), ISSN 1431-0643 (Internet)
Documenta Mathematica
Band 16, 2011

Richard Hill and David Loeffler
Emerton’s Jacquet Functors for Non-Borel Parabolic Subgroups 1–31

Chiarellotto, B. and Tsuzuki, N.
Log-Growth Filtration and Frobenius Slope Filtration of $F$-Isocrystals at the Generic and Special Points 33–69

Jürgen Hausen, Elaine Herppich, Hendrik Süß
Multigraded Factorial Rings and Fano Varieties with Torus Action 71–109

Rune Johansen
On the Structure of Covers of Sofic Shifts 111–131

Kengo Matsumoto
Ergodic Properties and KMS Conditions on $C^*$-Symbolic Dynamical Systems 133–175

M. Koenenberg
An Infinite Level Atom coupled to a Heat Bath 177–208

Markus Perling
Divisorial Cohomology Vanishing on Toric Varieties 209–251

Richard Pink, Torsten Wedhorn, and Paul Ziegler
Algebraic Zip Data 253–300

Thomas M. Fiore, Wolfgang Lück, and Roman Sauer
Euler Characteristics of Categories and Homotopy Colimits 301–354

L. Battisti
Surfaces de Stein Associées aux Surfaces de Kato Intermédiaires 355–371

G. A. Noskov
Asymptotic Behavior of Word Metrics on Coxeter Groups 373–398

Ugo Bruzzo and Dimitri Markushevish
Moduli of Framed Sheaves on Projective Surfaces 399–410

F. Déglise
Modules Homotopiques 411–455

Daniel R. Grayson
The Additivity Theorem in Algebraic $K$-Theory 457–464

Marcus Zibrowius
Witt Groups of Complex Cellular Varieties 465–511
| Title                                                                 | Authors                                                                 | Page(s)   |
|----------------------------------------------------------------------|-------------------------------------------------------------------------|-----------|
| The Hilbert-Chow Morphism and the Incidence Divisor                  | Joseph Ross                                                            | 513–543   |
| Presentation of an Iwasawa Algebra: The Case of $\Gamma_1 SL(2, \mathbb{Z}_p)$ | Laurent Clozel                                                          | 545–559   |
| On the Existence of Stationary Solutions for Some Non-Fredholm Integro-Differential Equations | Vitali Vougalter, Vitaly Volpert                                      | 561–580   |
| Fundamental Group of Schurian Categories and the Hurewicz Isomorphism | Claude Cibils, María Julia Redondo, and Andrea Solotar                 | 581–595   |
| Characterising Weak-Operator Continuous Linear Functionals on $\mathcal{B}(H)$ constructively | Douglas S. Bridges                                                      | 597–617   |
| The Classification of Real Purely Infinite Simple $C^*$-Algebras      | Jeffrey L. Boersema, Efren Ruiz, P. J. Stacey                         | 619–655   |
| A Simple Criterion for Extending Natural Transformations to Higher $K$-Theory | Gonçalo Tabuada                                                        | 657–668   |
| Ordinarity of Configuration Spaces and of Wonderful Compactifications | Kirti Joshi                                                            | 669–676   |
| Galois Algebras, Hasse Principle and Induction–Restriction Methods   | E. Bayer-Fluckiger and R. Parimala                                     | 677–707   |
| Irreducible Modules over the Virasoro Algebra                         | Rencai Lü, Xiangqian Guo, and Kaiming Zhao                            | 709–721   |
| Theta Series and Function Field Analogue of Gross Formula            | Fu-Tsun Wei and Jing Yu                                                | 723–765   |
| K-Homology Class of the Dirac Operator on a Compact Quantum Group    | Sergey Neshveyev and Lars Tuset                                       | 767–780   |
| Dilation Theory, Commutant Lifting and Semicrossed Products          | K.R. Davidson and E.G. Katsoulis                                      | 781–868   |
| Green Functions Via Hyperbolic Localization                           | Pramod N. Achar                                                        | 869–884   |
TREVOR ARNOLD AND KOOPA TAK-LUN KOO
CHANGE OF SELMER GROUP FOR BIG GALOIS
REPRESENTATIONS AND APPLICATION TO NORMALIZATION 885–899

GUNTER MALLE AND GABRIEL NAVARRO
EXTENDING CHARACTERS FROM HALL SUBGROUPS 901–919

XIAO-WU CHEN
THE SINGULARITY CATEGORY OF AN
ALGEBRA WITH RADICAL SQUARE ZERO 921–936

LARS HALVARD HALLE AND JOHANNES NICAISE
JUMPS AND MONODROMY OF ABELIAN VARIETIES 937–968

EUGEN HELLMANN
ON FAMILIES OF WEAKLY ADMISSIBLE FILTERED
\( \phi \)-MODULES AND THE ADJOINT QUOTIENT OF GL\(_d\) 969–991
EMERTON’S JACquet Functors for Non-Borel Parabolic Subgroups

RICHARD HILL AND DAVID LoeffLER

Received: May 30, 2010
Revised: October 18, 2010
Communicated by Peter Schneider

Abstract. This paper studies Emerton’s Jacquet module functor for locally analytic representations of \( p \)-adic reductive groups, introduced in [Eme06a]. When \( P \) is a parabolic subgroup whose Levi factor \( M \) is not commutative, we show that passing to an isotypical subspace for the derived subgroup of \( M \) gives rise to essentially admissible locally analytic representations of the torus \( \mathbb{Z}(M) \), which have a natural interpretation in terms of rigid geometry. We use this to extend the construction in of eigenvarieties in [Eme06b] by constructing eigenvarieties interpolating automorphic representations whose local components at \( p \) are not necessarily principal series.

2010 Mathematics Subject Classification: 11F75, 22E50, 11F70
Keywords and Phrases: Eigenvarieties, \( p \)-adic automorphic forms, completed cohomology

1 Introduction

1.1 Background

Let \( \mathcal{G} \) be a reductive group over a number field \( F \). The automorphic representations of the group \( \mathcal{G}(\mathbb{A}) \), where \( \mathbb{A} \) is the adèles of \( F \), are central objects of study in number theory. In many cases, it is expected that the set \( \Pi(\mathcal{G}) \) of automorphic representations contains a distinguished subset \( \Pi(\mathcal{G})^{\text{arith}} \) of representations which are (in some sense) “definable over \( \mathbb{Q} \)”. The subject of this paper is the \( p \)-adic interpolation properties of these representations (and their

The second author is grateful for the support of EPSRC Postdoctoral Fellowship EP/F04304X/2.
associated Hecke eigenvalues). Following the pioneering work of Coleman and Coleman-Mazur [Col06, Col07, CM98] for the automorphic representations attached to modular forms with nonzero Hecke eigenvalue at \( p \), it is expected that these Hecke eigenvalues should be parametrised by \( p \)-adic rigid spaces (eigenvarieties).

A very general construction of eigenvarieties is provided by the work of Emerton [Eme06b], using the cohomology of arithmetic quotients of \( \mathfrak{G} \). For any fixed open compact subgroup \( K_f \subseteq \mathfrak{G}(\mathbb{A}_f) \) (where \( \mathbb{A}_f \) is the finite adèles of \( F \)), and \( \mathfrak{K}_\infty \), the identity component of a maximal compact subgroup of \( \mathfrak{G}(F \otimes \mathbb{R}) \), the quotients \( Y(K_f) = \mathfrak{G}(F)/\mathfrak{G}(\mathbb{A}_f)/\mathfrak{K}_f\mathfrak{K}_\infty \) are real manifolds, equipped with natural local systems \( \mathcal{V}_X \) for each algebraic representation \( X \) of \( \mathfrak{G} \). The cohomology groups \( H^i(Y(K_f), \mathcal{V}_X) \) are finite-dimensional, and passing to the direct limit over \( K_f \) gives an admissible smooth representation \( H^i(\mathcal{V}_X) \) of \( \mathfrak{G}(\mathbb{A}_f) \).

Every irreducible subquotient of \( H^i(\mathcal{V}_X) \) is the finite part of an automorphic representation; we say that the representations arising in this way are cohomological (in degree \( i \)).

Emerton’s construction proceeds in two major steps. Fix a prime \( p \) above \( p \) and an open compact subgroup \( K^{(p)} \subseteq \mathfrak{G}(\mathbb{A}_f) \) (a “tame level”). Firstly, from the spaces \( H^i(Y(K^{(p)}), \mathcal{V}_X) \) for various open compact subgroups \( K^{(p)} \subseteq G = \mathfrak{G}(F_p) \), Emerton constructs Banach space representations \( \hat{H}^i(K^{(p)}) \) of \( G \). For any complete subfield \( L \) of \( F_p \), the spaces \( \hat{H}^i(K^{(p)})_{la} \) of locally \( L \)-analytic vectors are locally \( L \)-analytic representations of \( G \), and there are natural maps

\[
H^i(\mathcal{V}_X)^{K^{(p)}} \to \operatorname{Hom}_\mathfrak{g}(X', \hat{H}^i(K^{(p)})_{la}) \tag{1.1}
\]

where \( \mathfrak{g} = \text{Lie} \ G \). In many cases, these maps are known to be isomorphisms; if this holds, the automorphic representations which are cohomological in degree \( i \) are exactly those which appear as subquotients of \( \operatorname{Hom}_\mathfrak{g}(X', \hat{H}^i(K^{(p)})_{la}) \) for some \( X \) and tame level \( K^{(p)} \).

The second step in the construction is to extract the desired information from the space \( \hat{H}^i(K^{(p)})_{la} \). This is carried out by applying the Jacquet module functor of [Eme06a], for a Borel subgroup \( B \subseteq G \). This then produces an essentially admissible locally analytic representation of the Levi factor \( M \) of \( B \), which is a torus. There is an anti-equivalence of categories between essentially admissible locally analytic representations of \( M \) and coherent sheaves on the rigid-analytic space \( \hat{M} \) parametrising characters of \( M \). The eigenvariety \( E(i, K^{(p)}) \) is then constructed from this sheaf by passing to the relative spectrum of the unramified Hecke algebra \( \mathcal{H}^{\text{ph}} \) of \( K^{(p)} \); points of this variety correspond to characters \( (\kappa, \lambda) \in \hat{M} \times \text{Spec} \mathcal{H}^{\text{ph}} \) such that the \( (M = \kappa, \mathcal{H}^{\text{ph}} = \lambda) \)-eigenspace of \( J_B(\hat{H}^i(K^{(p)})_{la}) \) is nonzero. Hence if the map (1.1) above is an isomorphism, there is a point of \( E(i, K^{(p)}) \) for each automorphic representation \( \pi = \bigotimes_v \pi_v \) which is cohomological in degree \( i \) with \( (\pi_f^{(p)})^{K^{(p)}} \otimes J_B(\pi_p) \neq 0 \).
1.2 Statement of the main result

In this paper, we consider the situation where $B$ is replaced by a general parabolic subgroup $P$ of $G$. This extends the scope of the theory in two ways: firstly, it may happen that no Borel subgroup exists ($G$ may not be quasi-split); and even if a Borel subgroup exists, there will usually be automorphic representations for which $J_B(\pi_p) = 0$, which do not appear in Emerton’s eigenvariety. As above, we choose a number field $F$, a connected reductive group $\mathcal{G}$ over $F$, and a prime $p$ of $F$ above the rational prime $p$. Let $G = \mathcal{G} \times_F F_p$, a reductive group over $F_p$, and $G = \mathcal{G}(F_p)$. We choose a parabolic subgroup $P$ of $G$ (not necessarily arising from a parabolic subgroup of $\mathcal{G}$), with unipotent radical $N$; and let $\mathcal{M}$ be a Levi factor of $P$, with centre $Z$ and derived subgroup $D$. We write $G = \mathcal{G}(F_p)$, and similarly for $P, M, D, Z$. We choose a complete extension $L$ of $Q_p$, contained in $F_p$, so $G, P, M, D, Z$ are locally $L$-analytic groups.

Let $\Gamma = D \times \mathcal{G}(\mathbb{A}_F^p) \times \pi_0$, where $\pi_0$ is the component group of $\mathcal{G}(F \otimes \mathbb{R})$. Let us choose an open compact subgroup $U \subseteq \Gamma$ (this is the most natural notion of a “tame level” in this context), and a finite-dimensional irreducible algebraic representation $W$ of $\mathcal{M}$. As we will explain below, the Hecke algebra $\mathcal{H}(\Gamma//U)$ can be written as a tensor product $\mathcal{H}^{\text{ram}} \otimes \mathcal{H}^{\text{ph}}$, where $\mathcal{H}^{\text{ph}}$ is commutative, and $\mathcal{H}^{\text{ram}}$ is finitely-generated (and supported at a finite set of places $S$).

**Theorem (Theorem 6.3).** There exists a rigid-analytic subvariety $\mathcal{E}(i, P, W, U)$ of $\tilde{Z} \times \text{Spec} \mathcal{H}^{\text{ph}}$, endowed with a coherent sheaf $\mathcal{F}(i, P, W, U)$ with a right action of $\mathcal{H}^{\text{ram}}$, such that:

1. The natural projection $\mathcal{E}(i, P, W, U) \to \mathcal{G}$ has discrete fibres. In particular, the dimension of $\mathcal{E}(i, P, W, U)$ is at most equal to the dimension of $\tilde{Z}$.

2. The point $(\chi, \lambda) \in \tilde{Z} \times \text{Spec} \mathcal{H}^{\text{ph}}$ lies in $\mathcal{E}(i, P, W, U)$ if and only if the $(Z = \chi, \mathcal{H}^{\text{ph}} = \lambda)$-eigenspace of $\text{Hom}_U \left(W, J_P(\tilde{H})_{\mathbb{A}_F}\right)$ is nonzero. If this is so, the fibre of $\mathcal{F}(i, P, W, U)$ at $(\chi, \lambda)$ is isomorphic as a right $\mathcal{H}^{\text{ram}}$-module to the dual of that eigenspace.

3. If there is a compact open subgroup $G_0 \subseteq G$ such that $(\tilde{H}_{\mathbb{A}_F})_{U(\mathfrak{p})}$ is isomorphic as a $G_0$-representation to a finite direct sum of copies of $C^\text{cl}(G_0)$ (where $U(\mathfrak{p}) = U \cap \mathcal{G}(\mathbb{A}_F^p)$), then $\mathcal{E}(i, P, W, U)$ is equidimensional, of dimension equal to the rank of $\tilde{Z}$.

Now let us suppose that $W$ is absolutely irreducible, and write $\Pi(i, P, W, U)$ for the set of irreducible smooth $\mathcal{G}(\mathbb{A}_F) \times \pi_0$-representations $\pi_f$ such that $J_P(\pi_f)^{\mathcal{G}} \neq 0$, and $\pi_f$ appears as a subquotient of the cohomology space $H^1(\mathcal{V}_X)$ for some irreducible algebraic representation $X$ of $G$ such that $(X)^N \cong W \otimes \chi$ for a character $\chi$. To any such $\pi_f$, we may associate the point $(\theta \chi, \lambda) \in \tilde{Z} \times \text{Spec} \mathcal{H}^{\text{ph}}$, where $\theta$ is the smooth character by which $Z$ acts on $J_P(\pi_f)$, and $\lambda$ the character by which $\mathcal{H}^{\text{ph}}$ acts on $J_P(\pi_f)^{\mathcal{G}}$. Let $E(i, P, W, U)_{\text{cl}}$ denote
the set of points of $\hat{Z} \times \text{Spec } \mathcal{H}^{\text{sph}}$ obtained in this way from representations $\pi_f \in \Pi(i, P, W, U)$.

**Corollary (Corollary 6.4).** If the map (1.1) is an isomorphism in degree $i$ for all irreducible algebraic representations $X$ such that $(X')^N$ is a twist of $W$, then $E(i, P, W, U)_{cl} \subset \mathcal{E}(i, P, W, U)$. In particular, the Zariski closure of $E(i, P, W, U)_{cl}$ has dimension at most $\dim Z$.

In the special case when $\mathfrak{S}(F \otimes \mathbb{R})$ is compact modulo centre, a related statement has been proved (by very different methods) by the second author [Loe11]. If $P_1$ and $P_2$ are two different choices of parabolic, with $P_1 \supseteq P_2$, we have a relation between the eigenvarieties attached to $P_1$ and $P_2$ under a mild additional hypothesis, namely that the tame level be of the form $U^{(p)} \times U_p$, with $U^{(p)}$ an open compact subgroup away from $p$ and $U_p$ an open compact subgroup of $D_1 = [M_1, M_1]$ which admits a certain decomposition with respect to the parabolic $P_2 \cap D_1$ (see §5.2 below). In this situation, we have the following:

**Theorem (Theorem 6.5).** If $U$ is of the above type, then the space $\mathcal{E}(i, P_1, W, U)$ is equal to the union of two subvarieties $\mathcal{E}(i, P_1, W, U)_{P_2 - \text{null}}$ and $\mathcal{E}(i, P_1, W, U)_{P_2 - \text{null}}$, which are respectively endowed with sheaves of $\mathcal{H}^{\text{am}}$-modules $\mathcal{T}(i, P, W, U)_{P_2 - \text{null}}$ and $\mathcal{T}(i, P, W, U)_{P_2 - \text{null}}$ whose direct sum is $\mathcal{T}(i, P, W, U)$.

If $\pi_f \in \Pi(i, P, W, U)$ and $\pi_f$ is not annihilated by the map (1.1), then the point of $\mathcal{E}(i, P_1, W, U)$ corresponding to $\pi_f$ lies in the former subvariety if $J_{P_2}(\pi_p) \neq 0$, and in the latter if $J_{P_2}(\pi_p) = 0$. Moreover, there is a closed subvariety of $\mathcal{E}(i, P_2, W^{N_1}, U \cap D_2)$ whose image in $\hat{Z} \times \text{Spec } \mathcal{H}^{\text{sph}}$ is $\mathcal{E}(i, P_1, W, U)_{P_2 - \text{null}}$.

2 Preliminaries

2.1 Notation and definitions

Let $p$ be a prime. Let $K \supseteq \mathbb{Q}_p$ be a complete discretely valued field, which will be the coefficient field for all the representations we consider, and $L$ a finite extension of $\mathbb{Q}_p$ contained in $K$. If $V$ is a locally convex $K$-vector space, we let $V'$ denote the continuous dual of $V$. We write $V'_b$ for $V'$ endowed with the strong topology (which is the only topology on $V'$ we shall consider).

Let $S$ be an abstract semigroup. A topological representation of $S$ is a locally convex Hausdorff topological $K$-vector space $V$ endowed with a left action of $S$ by continuous operators. If $S$ has a topology, we say that the representation is separately continuous if the orbit map of each $v \in V$ is a continuous map $S \to V$, and continuous if the map $S \times V \to V$ is continuous. In particular, this applies when $S$ is a topological $K$-algebra and $V$ is an $S$-module, in which case we shall refer to $V$ as a separately continuous or continuous topological $S$-module.

If $G$ is a locally compact topological group and $V$ is a continuous representation of $G$, then $V'$ is a module over the algebra $D(G)$ of measures on $G$ [Eme04,
5.1.7], defined as $C(G)'$ where $C(G)$ is the space of continuous $K$-valued functions on $G$. If $G$ is a locally $p$-adic analytic group, then for any open compact subgroup $H \subseteq G$, the subalgebra $D(H)$ is Noetherian, and we say $V$ is admissible continuous [ST02a, Lemma 3.4] if $V$ is a Banach space and $V'$ is finitely generated over $D(H)$ for one (and hence every) open compact $H$.

If $G$ is a locally $L$-analytic group, in the sense of [ST02b], then we say the representation $V$ is locally analytic if it is a continuous $G$-representation on a space of compact type, and the orbit maps are locally $L$-analytic functions $G \to V$. This implies [Eme04, 5.1.9] that $V'_b$ is a separately continuous topological module over the topological $K$-algebra $D^{la}(G)$ of distributions on $G$, defined as $C^{la}(G)'_b$ where $C^{la}(G)$ is the space of locally $L$-analytic $K$-valued functions on $G$. For $H$ an open compact subgroup, the subalgebra $D^{la}(H)$ is a Fréchet-Stein algebra [ST03, 5.1], so the category of coadmissible $D^{la}(H)$-modules is defined [ST03, §3]; we say $V$ is admissibly locally analytic if $V'_b$ is coadmissible as a module over $D^{la}(H)$ for one (and hence every) open compact $H$.

Finally, if $G$ is a locally $L$-analytic group for which $Z = Z(G)$ is topologically finitely generated, we say the representation $V$ is $Z$-tempered if it is locally analytic and can be written as an increasing union of $Z$-invariant $BH$-spaces. This implies that for any open compact subgroup $H \subseteq G$, $V'_b$ is a jointly continuous topological module over the algebra $D^{ess}(H, Z(G)) = D^{la}(H) \widehat{\otimes} D^{la}(Z \cap H) C^{can}(\hat{Z})$, where $\hat{Z}$ is the rigid space\footnote{The space $\hat{Z}$ is in fact defined over $L$, but we shall always consider it as a rigid space over $K$ by base extension.} parametrising characters of $Z$. The algebra $D^{ess}(H, Z(G))$ is also a Fréchet-Stein algebra [Eme04, 5.3.22], and we say $V$ is essentially admissibly locally analytic if $V'_b$ is coadmissible as a module over $D^{ess}(H, Z(G))$ for one (and hence every) open compact $H$.

We write $\text{Rep}_{top}(G)$ for the category of topological representations of $G$, with morphisms being $G$-equivariant continuous linear maps. We consider the following full subcategories:

- $\text{Rep}_{cts}(G)$: continuous representations
- $\text{Rep}_{cts, ad}(G)$: admissible continuous representations
- $\text{Rep}_{top, c}(G)$: topological representations on compact type spaces
- $\text{Rep}_{la, c}(G)$: locally analytic representations
- $\text{Rep}_{la, c}(G)$: $Z$-tempered representations
- $\text{Rep}_{la, ad}(G)$: admissible locally analytic representations
- $\text{Rep}_{ess}(G)$: essentially admissible locally analytic representations
- $\text{Rep}_{cts, fd}(G)$: finite-dimensional continuous representations
- $\text{Rep}_{la, fd}(G)$: finite-dimensional locally analytic representations

Documenta Mathematica 16 (2011) 1–31
Each of these categories is stable under passing to closed $G$-invariant submodules. The categories $\text{Rep}_{\text{cts}, \text{ad}}(G)$, $\text{Rep}_{\text{la}, \text{ad}}(G)$ and $\text{Rep}_{\text{ess}}(G)$ have the additional property that all morphisms are strict, with closed image. The definition of $\text{Rep}_{\text{top}}$ and $\text{Rep}_{\text{top}, \text{c}}$ makes sense if $G$ is only assumed to be a semigroup. We will need one more category of representations of semigroups: if $S$ is a semigroup which contains a locally $L$-analytic subgroup $S_0$, we define $\text{Rep}_{\text{la}, \text{c}}^*(S)$ to be the full subcategory of $\text{Rep}_{\text{top}, \text{c}}(S)$ of representations which are locally analytic as representations of $S_0$, and can be written as an increasing union of $Z(S)$-invariant $BH$-subspaces. We will, in fact, only use this when either $S$ is a group (in which case the definition reduces to the definition of $\text{Rep}_{\text{la}, \text{c}}^*$ above) or $S$ is commutative.

Remark. If $V \in \text{Rep}_{\text{top}}(G)$, $V'$ naturally carries a right action of $G$. Hence we follow the conventions of [Eme04, §5.1] by defining the algebra structures on $D(G)$ and its cousins in such a way that the Dirac distributions satisfy $\delta_g \ast \delta_h = \delta_{gh}$, so all of our modules are left modules. The alternative is to consider the contragredient action on $V'$, which is the convention followed in [ST02b, ST03]; we do not adopt this approach here as we will occasionally wish to consider semigroups rather than groups.

2.2 Smooth and locally isotypical vectors

We now present a slight generalisation of the theory of [Eme04, §7]. Let $G$ be a locally compact topological group and $H \leq G$ closed. We suppose that $G$ admits a countable basis of neighbourhoods of the identity consisting of open compact subgroups; this is automatic if $G$ is locally $p$-adic analytic, for instance. The action of any $g \in G$ on $H$ by conjugation gives a homeomorphism from $H$ to itself, so the conjugation action of $G$ preserves the set of open compact subgroups of $H$.

**Definition 2.1.** Let $V$ be an (abstract) $K$-vector space with an action of $G$. We say a vector $v \in V$ is $H$-smooth if there is an open compact subgroup $U$ of $H$ such that $Uv = v$.

Our assumptions imply that the space $V_{H-\text{sm}}$ of $H$-smooth vectors is $G$-invariant.

**Definition 2.2** ([Eme04, 7.1.1]). Suppose $V \in \text{Rep}_{\text{top}}(G)$. We define

$$V_{H-\text{st, sm}} = \lim_{\underset{U \text{ open}}{\longrightarrow}} V^U,$$

equipped with the locally convex inductive limit topology.

Clearly $V_{H-\text{st, sm}}$ can be identified with $V_{H-\text{sm}}$ as an abstract $K$-vector space, but the inductive limit topology on the former is generally finer than the subspace topology on the latter. It is clear that the action of $G$ on $V$ induces a
topological action on $V_{H-\text{st.sm}}$, so $(-)_{H-\text{st.sm}}$ is a functor from $\text{Rep}_{\text{top}}(G)$ to itself, and the natural injection $V_{H-\text{st.sm}} \hookrightarrow V$ is $G$-equivariant. We say $V$ is strictly $H$-smooth if this map is a topological isomorphism.

**Proposition 2.3.**

(i) If $V \in \text{Rep}_{\text{cts}}(G)$, then $V_{H-\text{st.sm}} \in \text{Rep}_{\text{cts}}(G)$.

(ii) If $V \in \text{Rep}_{\text{top,c}}(G)$, then $V_{H-\text{st.sm}}$ is of compact type and the natural map $V_{H-\text{st.sm}} \to V$ is a closed embedding.

**Proof.** To show (i), we argue as in [Eme04, 7.1.10]. We let $G_0$ be an open compact subgroup of $G$ and $(H_i)_{i \geq 0}$ a decreasing sequence of open compact subgroups of $H$ satisfying $\bigcap_i H_i = \{1\}$ and with each $H_i$ normal in $G_0$; it is clear that we may do this, by our assumption on $G$. We set $H_i = G_i \cap H$. Then $V^{H_i}$ is a $G_0$-invariant closed subspace of $V$, and letting $V_i$ denote the kernel of the “averaging” map $V^{H_i} \to V^{H_{i-1}}$, we have $V^{H-\text{st.sm}} = \bigoplus_i V_i$. Since each $V_i$ is in $\text{Rep}_{\text{cts}}(G_0)$, $V_{H-\text{st.sm}} \in \text{Rep}_{\text{cts}}(G_0)$, which implies it is in $\text{Rep}_{\text{cts}}(G)$.

Statement (ii) depends only on $V$ as an $H$-representation, so we are reduced to the case of [Eme04, 7.1.3].

It follows from (ii) that for $V \in \text{Rep}_{\text{top,c}}(G)$ we do not need to distinguish between $V_{H-\text{st.sm}}$ and $V_{H-\text{cts}}$. Moreover, we see that if $V \in \text{Rep}_{\text{top,c}}(G)$ or any of the subcategories of admissible representations introduced above, $V_{H-\text{st.sm}}$ has the same property.

**Definition 2.4.** Let $V, W$ be abstract $K$-vector spaces with an action of $G$. We say a vector $v \in V$ is locally $(H,W)$-isotypical if there is an integer $n$, an open compact subgroup $U$ of $H$, and a $U$-equivariant linear map $W^n \to V$ whose image contains $v$.

The locally $(H,W)$-isotypical vectors clearly form a $G$-invariant subspace of $V$, since $H$ is normal in $G$. By construction, this is the image of the evaluation map $\text{Hom}_{H-\text{sm}}(W,V) \otimes_K W \to V$, where $\text{Hom}_{H-\text{sm}}(W,V)$ denotes the subspace of $H$-smooth vectors in $\text{Hom}_K(W,V) = W' \otimes_K V$ with its diagonal $G$-action.

If $V$ and $W$ are in $\text{Rep}_{\text{top}}(G)$, with $W$ finite-dimensional, then $\text{Hom}_K(W,V)$ has a natural topology (as a direct sum of finitely many copies of $V$) and we write $\text{Hom}_{H-\text{st.sm}}(W,V)$ for $\text{Hom}_K(W,V)_{H-\text{st.sm}}$, with its inductive limit topology as above. Then $\text{Hom}_{H-\text{st.sm}}(W,V) \otimes_K W$ is an object of $\text{Rep}_{\text{top}}(G)$ with a natural morphism to $V$.

We let $V_{(H,W)-\text{iso}}$ denote the image of $\text{Hom}_{H-\text{st.sm}}(W,V) \otimes_K W$ in $V$, endowed with the quotient topology from the source (which is generally finer than the subspace topology on the target). We say $V$ is strictly locally $(H,W)$-isotypical if the map $V_{(H,W)-\text{iso}} \to V$ is a topological isomorphism.

**Definition 2.5.** We say $W$ is $H$-good if $W$ is finite-dimensional, and for any open compact subgroup $U \subseteq H$, $\text{End}_U(W) = \text{End}_H(W) = \text{End}_G(W)$. 

*Documenta Mathematica 16 (2011) 1–31*
Proposition 2.6. Suppose $W$ is $H$-good, with $B = \text{End}_G(W)$. Then for any representation $V$ of $G$ on an abstract $K$-vector space, the natural map

$$\text{Hom}_K(W,V)_{H\text{-sm}} \otimes_B W \to V$$

is a $G$-equivariant injection. Dually, for any abstract right $B$-module $X$ with a $B$-linear $G$-action which is smooth restricted to $H$, the natural map

$$X \to \text{Hom}_K(W,X \otimes_B W)_{H\text{-sm}}$$

is an isomorphism.

Proof. If $G = H$, the first statement is [Eme04, 4.2.4] (the assumption in op.cit. that $W$ be algebraic is only used to show that $W$ is $H$-good). For the general case, the map exists and is injective at the level of $H$-representations, so it suffices to note that the assumption on $W$ implies that the left-hand side has a well-defined $G$-action, for which the map is $G$-equivariant.

For the second part, it suffices to show that the map restricts to a $G$-invariant $U'$, for any open $U' \subseteq H$. Since $W$ is faithful as a $B$-module by construction, the natural map is an injection. Since $X$ is smooth as an $H$-representation, any vector in the left-hand side is in $\text{Hom}_U(W,X \otimes_B W)$ for some $U'$, which we may assume to be normal in $U$. However, we have

$$\text{Hom}_U(W,X \otimes_B W) \subseteq \text{Hom}_{U'}(W,X \otimes_B W) = X \otimes_B \text{Hom}_{U'}(W,W),$$

and since $W$ is $H$-good, we have $\text{Hom}_{U'}(W,W) = B$, so $\text{Hom}_{U'}(W,X \otimes_B W) = X \otimes_B B$. Passing to $U/U'$-invariants gives the result.

Combining the preceding results shows that for $W$ an $H$-good representation, the two functors

$$\text{Hom}_{H\text{-st,sm}}(W,-) \quad \text{and} \quad - \otimes_B W$$

are mutually inverse equivalences between the categories of strictly locally $(H,W)$-isotypical representations of $G$ and strictly $H$-smooth $G$-representations on right $B$-modules.

Proposition 2.7. If $H$ is a locally $L$-analytic group, and $V$ is in $\text{Rep}_{\text{top}}(G) \cap \text{Rep}_{\text{loc}}(H)$, then there is a topological isomorphism $V_{H\text{-st,sm}} \cong V^\mathfrak{h}$, where $\mathfrak{h}$ is the Lie algebra of $H$. More generally, if $W$ is an $H$-good locally analytic representation of $G$, $V_{(H,W)-\text{loc}} \cong \text{Hom}_\mathfrak{h}(W,V) \otimes_B W$.

Proof. Clear from proposition 2.3(i), since a vector $v \in V$ is in $V_{H\text{-sm}}$ if and only if it is $\mathfrak{h}$-invariant.

3 Preservation of admissibility

3.1 Spaces of invariants

In this section we consider a group $G$ and a normal subgroup $H$, and consider the functor of $H$-invariants $V \mapsto V^H : \text{Rep}_{\text{top}}(G) \to \text{Rep}_{\text{top}}(G/H)$. Our aim is
to show that this preserves the various subcategories of admissible representations introduced in the previous section.

**Proposition 3.1.** If $V$ is an admissible Banach representation of a locally $p$-adic analytic group $G$, and $H \leq G$ is a closed normal subgroup, then $V^H$ is an admissible Banach representation of $G/H$.

**Proof.** Suppose first $G$ is compact, so $D(G)$ is Noetherian. Since $H$ is normal and acts continuously on $V$, $V^H$ is a $G$-invariant closed subspace; so $(V^H)'$ is a $D(G)$-module quotient of a finitely-generated $D(G)$-module, and hence is a finitely-generated $D(G)$-module. However, the closed embedding $C(G/H) \hookrightarrow C(G)$ dualises to a surjection $D(G) \to D(G/H)$, and it is clear that the $D(G)$-action on $(V^H)'$ factors through this surjection. Hence $(V^H)'$ is finitely-generated over $D(G/H)$. In the general case, let $G_0$ be a compact open subgroup of $G$ and $H_0 = G_0 \cap H$. Then $G_0/H_0$ is an open compact subgroup of $G/H$. By the above, $V^{H_0}$ is an admissible continuous $G_0/H_0$-representation. Since $V^H$ is a closed $G_0/H_0$-invariant subspace of $V^{H_0}$ it is also admissible continuous as a representation of $G_0/H_0$ and hence of $G/H$. 

We now suppose $G$ is a locally $L$-analytic group. We write $H \leq_L G$ to mean that $H$ is a closed normal subgroup of $G$ and the $\mathbb{Q}_p$-subspace $\text{Lie}(H) \subseteq \text{Lie}(G)$ is in fact an $L$-subspace, so $H$ and $G/H$ also inherit locally $L$-analytic structures.

**Proposition 3.2.** If $V$ is an admissible locally analytic representation of $G$, and $H \leq_L G$. Then $V^H$ is an admissible locally analytic representation of $G/H$.

**Proof.** As above, we may assume $G$ is compact. As in the Banach case, we note that $V^H$ is a closed $G$-invariant subspace of $V$, so it is an admissible locally analytic $G$-representation [ST03, 6.4(ii)] on which the action of $G$ factors through $G/H$. Hence the action of $D^{\text{ad}}(G)$ on $(V^H)'$ factors through $D^{\text{ad}}(G/H)$. Since the natural map $C^{\text{ad}}(G/H) \to C^{\text{ad}}(G)$ is a closed embedding, $D^{\text{ad}}(G/H)$ is a Hausdorff quotient of $D^{\text{ad}}(G)$ and hence a coadmissible $D^{\text{ad}}(G)$-module, and so by [ST03, 3.8] we see that $(V^H)'_0$ is coadmissible as a $D(G/H)$-module as required.

We now assume that $G$ is a locally $L$-analytic group with $Z(G)$ topologically finitely generated, and $H \leq_L G$. In this case $Z(G)/H$ may be much larger than $Z(G)/(Z(G) \cap H)$, as in the case of $\mathbb{Q}_p^{\times} \times \mathbb{Q}_p$; so an element of $\text{Rep}^{\text{sa,c}}_{\text{ad}}(G)$ on which $H$ acts trivially need not lie in $\text{Rep}^{\text{sa,c}}_{\text{ad}}(G/H)$. Moreover, it is not obvious that $Z(G)/H$ need be topologically finitely generated if $Z(G)$ is so. We shall therefore assume that $G$ is a direct product $H \times J$, with $H, J \leq_L G$, and $Z(H)$ and $Z(J)$ are both topologically finitely generated.

**Proposition 3.3.** In the above situation, for any essentially admissible locally analytic $G$-representation $V$, the space $V^H$ is an essentially admissible locally analytic representation of $J$. 

*Documenta Mathematica* 16 (2011) 1–31
Proof. By [Eme04, 6.4.11], any closed invariant subspace of an essentially admissible representation is essentially admissible; so it suffices to assume that $V = V^H$. Let $J_0 \subseteq J$ and $H_0 \subseteq H$ be open compact subgroups. Then $G_0 = J_0 \times H_0$ is an open compact subgroup of $G$. We have $\widehat{Z}(G) = \widehat{Z}(H) \times \widehat{Z}(J)$, and hence $\widehat{Z}(G) = \widehat{Z}(H) \times \widehat{Z}(J)$.

We now unravel the tensor products to find that the algebra

$$D^{\text{ess}}(G_0, Z(G)) = D^{\text{ra}}(G_0)_K \widehat{\otimes} D^{\text{ra}}(J_0)_K C^{\text{can}}(\widehat{Z}(G))$$

decomposes as

$$\left( D^{\text{ra}}(H_0)_K \otimes D^{\text{ra}}(J_0)_K \right) \left( C^{\text{can}}(\widehat{Z}(H))_K \otimes C^{\text{can}}(\widehat{Z}(J))_K \right)$$

$$= \left( D^{\text{ra}}(H_0)_{D^{\text{ra}}(H_0) \cap Z(H)} \otimes C^{\text{can}}(\widehat{Z}(H))_K \right) \left( D^{\text{ra}}(J_0)_{D^{\text{ra}}(J_0) \cap Z(J)} \otimes C^{\text{can}}(\widehat{Z}(J))_K \right)$$

$$= D^{\text{ess}}(H_0, Z(H))_K \otimes D^{\text{ess}}(J_0, Z(J))_K.$$

By assumption, the action of $D^{\text{ess}}(H_0, Z(H))$ on $V_0'$ factors through the augmentation map to $K$; so the action of $D^{\text{ess}}(G_0, Z(G))$ factors through $D^{\text{ess}}(J_0, Z(J))$. Since $D^{\text{ess}}(J_0, Z(J))$ is a Hausdorff quotient of $D^{\text{ess}}(G_0, Z(G))$, it is a coadmissible $D^{\text{ess}}(G_0, Z(G))$-algebra, and thus $V_0'$ is a coadmissible $D^{\text{ess}}(J_0, Z(J))$-module as required.

3.2 Admissible representations of product groups

In this section, we’ll recall the theory presented in [Eme04, §7] of representations of groups of the form $G \times \Gamma$, where $G$ is a locally $L$-analytic group and $\Gamma$ an arbitrary locally profinite (locally compact and totally disconnected) topological group. This will allow us to give more “global” formulations of the results of the previous section.

Let * denote one of the set {“admissible Banach”, “admissible locally analytic”, “essentially admissible locally analytic”}, so we shall speak of “*-admissible representations”. Whenever we consider essentially admissible representations we will assume that the groups concerned have topologically finitely generated centre, so the concept is well-defined.

Definition 3.4 ([Eme04, 7.2.1]). A *-admissible representation of $(G, \Gamma)$ is a locally convex $K$-vector space $V$ with an action of $G \times \Gamma$ such that

- For each open compact subgroup $U \subseteq \Gamma$, $V^U$ has property * as a representation of $G$ (in the subspace topology);

- $V$ is a strictly smooth $\Gamma$-representation in the sense of definition 2.1.
Remark. Our terminology is slightly different from that of [Eme04], where such representations are described as \(\ast\)-admissible representations of \(G \times \Gamma\). We adopt the formulation above in order to avoid ambiguity when \(\Gamma\) is also a locally analytic group.

The results of the preceding section can be combined to prove:

**Proposition 3.5.** If \(G\) and \(H\) are locally \(L\)-analytic groups, \(V\) is a \(\ast\)-representation of \(G \times H\), and \(Z(H)\) is compact if \(\ast = \text{"essentially admissible locally analytic"}, then the space

\[ V_{H\text{-st.sm}} = \lim_{U \subseteq H \text{ open compact}} V^U \]

is a \(\ast\)-admissible representation of \((G, H)\).

**Proof.** Since the natural maps \(V^U \hookrightarrow V^{U'}\) for \(U' \subseteq U\) are closed embeddings, the map \(V^U \hookrightarrow V_{H\text{-st.sm}}\) is also a closed embedding [Bou87, page II.32]; and its image is clearly \((V_{H\text{-st.sm}})^U\), so it suffices to check that \(V^U\) has property \(\ast\) for each \(U\).

In the admissible Banach case, this is clear from proposition 3.1. In the admissible locally analytic case, it likewise follows from proposition 3.2. In the essentially admissible case, it suffices to note that the assumption on \(Z(H)\) implies that \(V\) is essentially admissible as a representation of \(G \times H\) if and only if it is essentially admissible as a representation of \(G \times U\) for any open compact \(U \subseteq H\); so we are in the situation of proposition 3.3. \(\square\)

A slightly more general version of this applies to groups of the form \(G \times H \times J\), where \(G\) and \(H\) are locally \(L\)-analytic and \(J\) is an arbitrary locally compact topological group.

**Theorem 3.6.** Let \(V\) be a \(\ast\)-admissible representation of \((G \times H, J)\), where \(Z(H)\) is compact in the essentially admissible case. Then \(V_{H\text{-st.sm}}\) is a \(\ast\)-admissible representation of \((G, H \times J)\).

**Proof.** We have

\[ V_{H\text{-st.sm}} = (V_{J\text{-st.sm}})_{H\text{-st.sm}} = \lim_{U \subseteq H \text{ open compact}, U' \subseteq J} V^{U \times U'} \]

which is clearly a strict inductive limit; and \(V^{U \times U'}\) is the \(U\)-invariants in the \(\ast\)-admissible \(G \times H\)-representation \(V^{U'}\), and hence an admissible \(G\)-representation. The open compact subgroups of \(H \times J\) of the form \(U \times U'\) are cofinal in the family of all open compact subgroups, so \(V_{H\text{-st.sm}}\) is a \(\ast\)-admissible \((G, H \times J)\)-representation as required. \(\square\)

We write \(\text{Rep}_{\text{cts.ad}}(G, \Gamma)\) for the category of admissible continuous \((G, \Gamma)\)-representations, and similarly for the other admissibility conditions.
3.3 Ordinary parts and Jacquet modules

Let \( G \) be a connected reductive algebraic group over \( L \), and \( P \) a parabolic subgroup of \( G \) with Levi factor \( M \). We write \( Z = Z(M), D = M^\text{ss} \). We use Roman letters \( G, P, M, Z, D \) for the \( L \)-points of these, which are locally \( L \)-analytic groups. Note that the multiplication map \( Z \times D \to M \) has finite kernel and cokernel, and hence a representation of \( M \) has property \( * \) if and only if it has the corresponding property as a representation of \( Z \times D \).

Suppose that \( V \in \text{Rep}_{\text{cts,adm}}(G) \). We say \( V \) is unitary if the topology of \( V \) can be defined by a \( G \)-invariant norm (or equivalently if \( V \) contains a \( G \)-invariant separated open lattice); this is automatic if \( G \) is compact, but not otherwise.

The category \( \text{Rep}_{\text{u,adm}}(G) \) of unitary admissible Banach representations of \( G \) over \( K \) is equivalent to \( \text{Mod}^{\pi-\text{adm}}_{\text{G}}(O_K) \), where \( \text{Mod}^{\pi-\text{adm}}(O_K) \) is the category considered in [Eme10, 2.4.5] and the subscript \( Q \) denotes the category with the same objects but all Hom-spaces tensored with \( Q \).

In [Eme10, §3], Emerton constructs the ordinary part functor \( \text{Ord}_P : \text{Mod}^{\pi-\text{adm}}_{\text{G}}(O_K) \to \text{Mod}^{\pi-\text{adm}}_{\text{M}}(O_K) \).

This functor is additive, so it extends to a functor

\[
\text{Ord}_P : \text{Rep}_{\text{u,adm}}(G) \to \text{Rep}_{\text{u,adm}}(M).
\]

It is easy to extend this to representations of product groups of the type considered above. Let \( \Gamma \) be a locally profinite topological group, and \( V \) a unitary admissible Banach \((G, \Gamma)\)-representation (i.e. admitting a \( G \times \Gamma \)-invariant norm). We define

\[
\text{Ord}_P(V) = \lim_{\text{open}} \text{Ord}_P(V_U).
\]

Given any subgroups \( U' \subseteq U \), there is an “averaging” map \( \pi : V^{U'} \to V^U \); and we may write \( V^{U'} \) as a locally convex direct sum \( V^{U'} = V^U \oplus V^\pi \), where \( V^\pi \) denotes the kernel of \( \pi \). Since the ordinary part functor commutes with direct sums, we find that \( \text{Ord}_P(V^{U'}) = \text{Ord}_P(V^U) \oplus \text{Ord}_P(V^\pi) \); thus the natural map \( \text{Ord}_P(V^U) \to \text{Ord}_P(V^{U'}) \) is a closed embedding, and if \( U' \subseteq U \), we have \( \text{Ord}_P(V^{U'})^U = \text{Ord}_P(V^U) \). Passing to the direct limit, we have \( \text{Ord}_P(V^U) = \text{Ord}_P(V) \) is an admissible Banach \((M, \Gamma)\)-representation.

An identical argument applies to the Jacquet module functor \( J_P : \text{Rep}_{\text{cts}}(G) \to \text{Rep}_{\text{cts}}(M) \) of [Eme06a] (and indeed to any functor which preserves direct sums). Combining this with theorem 3.6 above, we have:

**Proposition 3.7.**

(i) If \( V \in \text{Rep}_{\text{u,adm}}(G, \Gamma) \) and \( W \in \text{Rep}_{\text{cts,fd}}(M) \), then

\[
\text{Hom}_{\text{D-st.sm}}(W, \text{Ord}_P V) \in \text{Rep}_{\text{cts,adm}}(Z, D \times \Gamma).
\]

Moreover, \( \text{Hom}_{\text{D-st.sm}}(W, \text{Ord}_P V) \) is unitary if \( W \) is.
(ii) If \( V \in \text{Rep}_{\text{ess}}(G, \Gamma) \) and \( W \in \text{Rep}_{\text{la, fd}}(M) \), and \( \mathfrak{d} = \text{Lie } D \), then

\[
\text{Hom}_{\text{D-std., sm}}(W, J_P V) = \text{Hom}_{\mathfrak{d}}(W, J_P V) \in \text{Rep}_{\text{ess}}(Z, D \times \Gamma).
\]

4 Jacquet modules of admissible representations

As in section 3.3 above, let \( G \) be the \( L \)-points of a connected reductive algebraic group over \( L \), and \( P \) a parabolic subgroup with Levi subgroup \( M \). Proposition 3.7(ii) gives us a copious supply of essentially admissible locally analytic representations of the torus \( Z = Z(M) \): for any \( V \in \text{Rep}_{\text{ess}}(G) \), any open compact \( U \subseteq D = M^{ss} \), and any finite-dimensional \( M \)-representation \( W \), \( \text{Hom}_U(W, J_P V) = (W' \otimes_K J_P V)^U \in \text{Rep}_{\text{ess}}(Z) \). These correspond, by the equivalence of categories of [Eme06b, 2.3.2], to coherent sheaves on the rigid space \( \hat{Z} \). For \( V \in \text{Rep}_{\text{ess}}(Z) \), we will write \( \text{Exp} V \) for the support of the sheaf corresponding to \( V \), a reduced rigid subspace of \( \hat{Z} \).

In this section, we’ll prove two results describing the geometry of the rigid spaces \( \text{Exp} \text{Hom}_U(W, J_P V) \), for \( U \subseteq D \) open compact, under additional assumptions on \( V \). These generalise the corresponding results in [Eme06a] when \( P \) is a Borel subgroup.

4.1 Compact maps

We begin by generalising some results from [Eme06a, §2.3] on compact endomorphisms of topological modules. Recall that a topological \( K \)-algebra is said to be of compact type if it can be written as an inductive limit of Banach algebras, with injective transition maps that are both algebra homomorphisms and compact as maps of topological \( K \)-vector spaces. If \( A \) is such an algebra, then a topological \( A \)-module is said to be of compact type if it is of compact type as a topological \( K \)-vector space.

In this situation, we have the following definition of a compact morphism (op.cit., def. 2.3.3):

**Definition 4.1.** A continuous \( A \)-linear morphism \( \phi : M \to N \) between compact type topological \( A \)-modules is said to be \( A \)-COMPACT if there is a commutative diagram

\[
\begin{array}{ccc}
M & \xrightarrow{\phi} & N \\
\downarrow{\alpha} & & \downarrow{\beta} \\
N_1 & \xrightarrow{\gamma} & V
\end{array}
\]

(4.1)

where \( N_1 \) is a compact type topological \( A \)-module, \( \alpha \) and \( \beta \) are continuous \( A \)-linear maps, \( V \) is a compact type \( K \)-vector space, and \( \gamma \) is a continuous \( K \)-
linear map for which $A \hat{\otimes}_{K} V \to N_{1}$ is surjective, and the composite dashed arrow is compact as a map of compact type $K$-vector spaces.

**Lemma 4.2.** If $M$ is a compact type module over a compact type topological $K$-algebra $A$; $\phi : M \to M$ is an $A$-compact map; $N$ is a finitely-generated module over a finite-dimensional $K$-algebra $B$; and $\psi : N \to N$ is $K$-linear, then the map $\phi \otimes \psi : M \otimes_{K} N \to M \otimes_{K} N$ is $(A \otimes_{K} B)$-compact.

**Proof.** We may assume without loss of generality that $\psi$ is the identity, by [Eme06a, 2.3.4(i)]. This case follows immediately by tensoring each of the spaces in the diagram with $N$.

**Lemma 4.3.** Let $\sigma : A \to A'$ be a finite morphism of compact type topological $K$-algebras, and $\phi : M \to N$ a morphism of topological $A'$-modules which is $A'$-compact. Then $\phi$ is $A$-compact.

**Proof.** By assumption, we have a diagram as in lemma 4.1, where the map $A' \hat{\otimes}_{K} V \to N_{1}$ is surjective. Let $a_{1}, \ldots, a_{k}$ be a set of elements generating $A'$ as an $A$-module, let $V' = V^{k}$, and define the map $\gamma' : V' \to N_{1}$ by $(v_{1}, \ldots, v_{k}) \mapsto \sum a_{i} \gamma(v_{i})$. Then it is clear that $1 \hat{\otimes} \gamma'$ gives a surjection $A \hat{\otimes}_{K} V^{k} \to N_{1}$. Furthermore, the composite map $\phi \circ \gamma' : V' \to N$ is the map $(v_{1}, \ldots, v_{k}) \mapsto \sum \beta(a_{i} \gamma(v_{i}))$. As $\beta$ is a morphism of $A'$-modules, this equals $\sum a_{i} (\beta \circ \gamma)(v)$, which is clearly compact (since $\beta \circ \gamma$ is). So the map $\gamma' : V' \to N_{1}$ witnesses $\phi$ as an $A$-compact map.

### 4.2 Twisted distribution algebras

Let $L$ be a finite extension of $\mathbb{Q}_{p}$, and $G$ a locally $L$-analytic group. Let $(H_{n})_{n \geq 0}$ be a decreasing sequence of good $L$-analytic open subgroups of $G$, in the sense of [Eme04, §5.2], such that

- the subgroups $H_{n}$ form a basis of neighbourhoods of the identity in $G$;
- $H_{n}$ is normal in $H_{0}$ for all $n$;
- the inclusion $H_{n+1} \hookrightarrow H_{n}$ extends to a morphism of rigid spaces between the underlying affinoid rigid analytic groups $\mathbb{H}_{n+1} \hookrightarrow \mathbb{H}_{n}$, which is relatively compact.

Such a sequence certainly always exists, since the choice of $H_{0}$ determines a Lie $O_{L}$-lattice $\mathfrak{h}$ in the Lie algebra of $G$, and we may take $H_{n}$ to be the subgroup attached to the sublattice $\pi^{n}\mathfrak{h}$. We may use this sequence to write the topological $K$-algebra $A := D^{\text{in}}(H_{0}) = C^{\text{la}}(H_{0})'$ as an inverse limit of the spaces $A_{n} := D(\mathbb{H}_{n}^{0}, H_{0}) = [C(H_{0})_{\mathbb{R} - \text{an}}]'$. For all $n$, $A_{n}$ is a compact type topological $K$-algebra, and the sequence $(A_{n})_{n \geq 0}$ is a weak Fréchet-Stein structure on $A$.

We begin with a construction related to the “untwisting isomorphism” of [Eme04, 3.2.4]. Let $(\rho, W)$ be any finite-dimensional $K$-representation of $H_{0}$,
and let $E = \text{End}_K W$. We consider the following commutative diagram of $K$-vector spaces:

$$
\begin{array}{ccc}
K[H_0] \otimes_K E & \xrightarrow{\gamma} & K[H_0] \otimes_K E \\
\downarrow{g \mapsto g \otimes 1} & & \downarrow{g \mapsto g \otimes \rho(g)m} \\
K[H_0] & \xrightarrow{\beta} & K[H_0] \otimes_K E
\end{array}
$$

(4.2)

Here $\alpha$ and $\beta$ are ring homomorphisms, and although $\gamma$ is not a ring homomorphism, it satisfies the relation $\gamma(\alpha(x)y) = \beta(x)\gamma(y)$, so it intertwines the two $K[H_0]$-module structures on $K[H_0] \otimes_K E$ given by $\alpha$ and $\beta$. Furthermore $\gamma$ is clearly invertible.

We now assume that $(\rho, W)$ is locally analytic (when $W$ is equipped with its unique Hausdorff locally convex topology).\footnote{If $L = \mathbb{Q}_p$, this is equivalent to the (a priori weaker) assumption that $(\rho, W)$ is continuous. This follows from the $p$-adic analogue of Cartan’s theorem, which states that any continuous homomorphism between two $\mathbb{Q}_p$-analytic groups is locally analytic; see [Ser92, Part II, §V.9].}

Hence there is an integer $n(\rho)$ such that $W^{\mathbb{R}_n-an} = W$ for all $n \geq n(\rho)$.

**Proposition 4.4.** Let $n \geq n(\rho)$. Then there exist unique continuous maps $\alpha_n, \beta_n : A_n \to A_n \otimes_K \text{End}(W)$ and $\gamma_n : A_n \otimes_K \text{End}(W) \xrightarrow{\sim} A_n \otimes_K \text{End} W$ extending the maps $\alpha, \beta, \gamma$ above.

**Proof.** Taking the (algebraic) $K$-dual of the diagram (4.2), we have a diagram

$$
\begin{array}{ccc}
\mathcal{F}(H_0, E') & \xrightarrow{\alpha'} & \mathcal{F}(H_0, K) \\
\downarrow{\gamma'} & & \downarrow{\beta'} \\
\mathcal{F}(H_0, E') & \xrightarrow{\beta'} & \mathcal{F}(H_0, E')
\end{array}
$$

where for $K$-vector space $V$, $\mathcal{F}(H_0, V)$ indicates the $K$-vector space of arbitrary functions $H_0 \to V$. One finds that for a function $f : H_0 \to E'$, we have $\alpha'(f)(m) = f(m)(1)$ and $\beta'(f)(m) = f(m)(\rho(m))$, while $\gamma'(f)(m) = x \mapsto f(\rho(m)x)$. All of these maps manifestly preserve the subspaces of $\mathbb{H}_n$-analytic functions for $n \geq n(\rho)$, and are continuous for the natural topologies of these subspaces; so there are corresponding maps between the duals of these subspaces, as required.

**Corollary 4.5.** For each $n \geq n(\rho)$, the map $\beta_n$ makes $B_n = A_n \otimes_K \text{End} W$ a finitely-generated topological $A_n$-module, and the natural map $B_{n+1} \to B_n$ induces an isomorphism $A_n \hat{\otimes}_{A_{n+1}} B_{n+1} \xrightarrow{\sim} B_n$. 
Proof. This is clearly true for the $A_n$-module structure on $B_n$ given by $\alpha_n$, so it follows for the $\beta_n$-structure (since the untwisting isomorphisms $\gamma_n$ and $\gamma_{n+1}$ are compatible with the map $B_{n+1} \to B_n$).

4.3 Twisted Jacquet modules

We now return to the situation considered above, so $G$ is the group of $L$-points of a reductive algebraic group $G$ over $L$ as above, with $P$ a parabolic subgroup, $M$ a Levi subgroup of $P$, $N$ the unipotent radical, and $Z = Z(M)$. We choose a sequence $(H_n)_{n \geq 0}$ of good $L$-analytic open subgroups of $G$ admitting rigid analytic Iwahori decompositions $H_n = \overline{N}_n \times M_n \times N_n$, as in [Eme06a, 4.1.6]. We also impose the additional condition that $M_n = Z_n \times D_n$ where $Z_n$ and $D_n$ are the affinoid subgroups underlying good analytic open subgroups of $Z$ and of $D = M^\circ$; it is clear that we can always do this (by exactly the same method as in Emerton’s case). We let $Z^+$ be the submonoid $\{ z \in Z(M) : zN_0z^{-1} \subseteq N_0 \}$ of $Z$.

Our starting point is the following, which is part of the proof of [Eme06a, 4.2.23]:

**Proposition 4.7.** Let $V$ be an admissible locally analytic representation of $G$. Then for all $n \geq 0$, the action of $M_0 \times Z^+$ on the space

$$U_n = \left( D(\overline{H}_n, H_0) \overset{\otimes}{\longrightarrow} V^\theta_{D_0(H_0)} \right)_{N_0}$$


extends to an $A_n[Z^+]$-module structure. Moreover, the transition map
$A_n \hat{\otimes}_{A_n} U_{n+1} \to U_n$ is $A_n$-compact and $Z^+$-equivariant, and there is
some $z \in Z^+$ (independent of $n$) such that there exists a map $\alpha : U_n \to A_n \hat{\otimes}_{A_n} U_{n+1}$ making the following diagram commute:

\begin{equation}
\begin{array}{ccc}
A_n \hat{\otimes}_{A_n} U_{n+1} & \to & U_n \\
\downarrow \text{id} \otimes z & & \downarrow \text{id} \\
A_n \hat{\otimes}_{A_n} U_{n+1} & \to & U_n.
\end{array}
\end{equation}

We now let $\tilde{U}_n = U_n \otimes_K W$, where $(W, \rho)$ is a fixed, finite-dimensional, con-
tinuous representation of $M$. By the last proposition of the preceding section
(taking the groups there denoted by $G$ and $H_i$ to be those we are now calling
$M$ and $M_i$), we have a diagonal $A_n$-module structure on $\tilde{U}_n$, and there is also
a diagonal action of $Z^+$ on $\tilde{U}_n$ commuting with the $M_0$-action.

**Proposition 4.8.** For any $n \geq n(\rho)$ the following holds:

- $\tilde{U}_n$ is a compact type topological $A_n$-module, and the action of $Z^+$ is
  $A_n$-linear.

- There is an $A_{n+1}[Z^+]$-linear map $U_{n+1} \to U_n$ such that the induced map
  $A_n \hat{\otimes}_{A_n} \tilde{U}_{n+1} \to \tilde{U}_n$ is $A_n$-compact.

- For any good $z \in Z^+$, we can find a map $\tilde{\alpha} : U_n \to A_n \hat{\otimes}_{A_n} \tilde{U}_{n+1}$ such
  that the diagram corresponding to (4.3) commutes.

Also, the direct limit $\varinjlim U_n$ (with respect to the transition maps above) is iso-
morphic as a topological $A[Z^+]$-module to $(V_N \otimes W')'_b$.

**Proof.** Since $\tilde{U}_n$ is isomorphic to $(U_n)^{\otimes \dim W}$ as a topological $K$-vector space,
it is certainly of compact type, and we have already observed that it is a
topological $A_n$-module for all $n \geq n(\rho)$. Furthermore the $Z^+$-action commutes
with the $M_0$-action, and thus it must be $A_n$-linear by continuity.

Moreover, we have an $A_n$-compact map $A_n \hat{\otimes}_{A_n} U_{n+1} \to U_n$. Ten-
soring with the identity map gives a morphism of $A_n \otimes \text{End} W$-modules
$(A_n \hat{\otimes}_{A_n} U_{n+1}) \otimes_K W \to U_n \otimes_K W$, which is $A_n \otimes_K \text{End} W$-compact by
lemma 4.2. But the map $\beta : A_n \to A_n \otimes_K \text{End} W$ is a finite morphism, so by
lemma 4.3, this map is $A_n$-compact.

Finally, we know that there exists a map $\alpha : U_n \to A_n \hat{\otimes}_{A_n} U_{n+1}$ through
which $z$ factors, and it is clear that if we define $\tilde{\alpha}$ to be the map $\alpha \otimes \rho(z)$ then
the diagram corresponding to (4.3) commutes.

The preceding proposition asserts precisely that the hypotheses of [Eme06a,
3.2.24] are satisfied, and that proposition (and its proof) give us the following:
Moreover, if $(Y_n)_{n \geq 0}$ is any increasing sequence of affinoid subdomains of $\hat{Z}$ whose union is the entire space, then for any $n \geq n(\rho)$ we have

$$\left( \bigcap_{k \in \mathbb{Z}} \bigoplus K \mathcal{A}_n \right) \left( \bigcap_{\mathbb{Z}} G \right) X = \bigcap_{K[Z^+]} \left( \bigcap_{k \in \mathbb{Z}} \bigoplus K \mathcal{A}_n \right) \left( \bigcap_{\mathbb{Z}} G \right) \hat{U}_n.$$ 

By [Eme04, 3.2.9] we have $X = [(V^N \otimes_K W')_\ast]_b = [(V^N \otimes_K W')_b]_b = [J_P(V) \otimes_K W']_b$, so the above corollary gives us a description of the strong dual of the $W$-twisted Jacquet module.

We can now prove the first of the two main theorems of this section. Proposition 3.7(ii) above shows that for any $V \in \text{Rep}_{\text{an}}(G)$, $(J_P(V) \otimes_K W') \in \text{Rep}_{\text{an}}(Z, D)$. Equivalently, for any open compact subgroup $\Gamma \subseteq D$, the space $(J_P(V) \otimes_K W')^G$ is an essentially admissible locally analytic $Z$-representation, and hence corresponds to a coherent sheaf on $\hat{Z}$. The previous corollary allows us to describe the support of this sheaf when $V$ is admissible:

**Theorem 4.10.** Suppose $V$ is an admissible locally analytic $G$-representation, $W$ is a finite-dimensional locally analytic representation of $M$, and $\Gamma$ is an open compact subgroup of $D$. Let $E \subseteq \hat{Z}$ be the support of the coherent sheaf on $\hat{Z}$ corresponding to $(J_P(V) \otimes_K W')^G$. Then the natural map $E \to (\text{Lie} \hat{Z})'$ (induced by the differentiation map $\hat{Z} \to (\text{Lie} \hat{Z})'$) has discrete fibres.

**Proof.** Since we are free to replace the sequence $(H_n)$ of subgroups of $G$ with a cofinal subsequence, we may assume that $\Gamma \supseteq D_0$. So it suffices to prove the result for $\Gamma = D_0$. Furthermore, since the differentiation map $\hat{Z}_0 \to (\text{Lie} \hat{Z})'$ has discrete fibres, it suffices to show that for any character $\chi$ of $Z_0$, the rigid space

$$\text{Exp} (J_P(V) \otimes_K W')_{D_0, Z_0=\chi} \subseteq \hat{Z}$$

is discrete. If $\chi$ does not extend to a character of $M$, then this space is clearly empty, so there is nothing to prove; otherwise, let us fix such an extension, which gives us an isomorphism $(J_P(V) \otimes_K W')_{D_0, Z_0=\chi} = [J_P(V) \otimes_K (W \otimes_K \chi)]^{M_0}$. So we may assume without loss of generality that $\chi$ is the trivial character, and it suffices to show that

$$\bigcap_{K[Z^+]} \left( \bigcap_{k \in \mathbb{Z}} \bigoplus K \mathcal{A}_n \right) \left( \bigcap_{\mathbb{Z}} G \right) \hat{U}_n,$$

is finite-dimensional over $K$ for all $n$, or (equivalently) all sufficiently large $n$. If we take the completed tensor product of both sides of the formula in corollary 4.9 with $\bigcap_{K[Z^+]} \left( \bigcap_{k \in \mathbb{Z}} \bigoplus K \mathcal{A}_n \right)$, regarded as a $\bigcap_{K[Z^+]} \left( \bigcap_{k \in \mathbb{Z}} \bigoplus K \mathcal{A}_n \right)$-algebra via the augmenta-
does exist, by hypothesis. Hence some open compact $H$ with a cofinal subsequence if necessary) that are done.

Theorem. As in [Eme06a], we may assume (by replacing the sequence $(G_n)_{n>0}$ with a cofinal subsequence if necessary) that $H = H_\emptyset$ and $\Gamma \supseteq D_0$. But then we can identify $(J_\Gamma(V) \otimes_K W')^{\Gamma}$ with a direct summand of $(J_\Gamma(V) \otimes_K W')^{D_0}$; this identifies $\text{Exp}(J_\Gamma(V) \otimes_K W')^{\Gamma}$ with a union of irreducible components.
of $\text{Exp}(J_P(V) \otimes_K W)^{D_0}$. We may therefore assume that in fact $\Gamma = D_0$. As a final reduction, letting $U_n = \left(D(\mathbb{H}_n^0, H_0) \otimes_{D(n, H_0)} V_n^r\right)_{N_0}$ as before, we note that the untwisting isomorphism $U_n \sim D(\mathbb{N}_n, \mathbb{N}_0)^r \otimes_K A_n$ (equation 4.2.39 in [Eme04]) can be extended to an isomorphism $U_n \otimes_K W \rightarrow D(\mathbb{N}_n, \mathbb{N}_0)^r \dim W \otimes_K A_n$. We thus assume that $W$ is the trivial representation.

Following Emerton, we choose Banach spaces $W_n$ such that the map $D(\mathbb{N}_n+1, \mathbb{N}_0)^r \rightarrow D(\mathbb{N}_n, \mathbb{N}_0)^r$ factors through $W_n$, and (exactly as in the Borel case) for a suitable $z \in \mathbb{Z}^+$ we have

$$J_P(V)_b^r \sim \lim_{\nu} K \{\{\nu, \nu^{-1}\}\} \otimes_K (W_n \otimes_k A_n),$$

for some $A_n$-linear action of $z$ on $W_n \otimes_K A_n$ which factors through $D(\mathbb{N}_n+1, \mathbb{N}_0)^r \otimes_K A_n$. Taking the completed tensor product with the map $A_n \rightarrow D(\mathbb{Z}_n^0, Z_0)$ given by the augmentation map of $D_0$, we have

$$[J_P(V)_b^r]_b \sim \lim_{\nu} K \{\{\nu, \nu^{-1}\}\} \otimes_K W_n \otimes_K D(\mathbb{Z}_n^0, Z_0).$$

Let us write $\mathring{\mathbb{Z}}_0$ as an increasing union of affinoid subdomains $(X_n)_{n \geq 0}$, such that the natural map $D(\mathring{\mathbb{Z}}_0) \sim C^an(\mathring{\mathbb{Z}}_0) \rightarrow C^an(X_n)$ factors through $D(\mathbb{Z}_n^0, Z_0)$. Extending scalars from $D(\mathbb{Z}_n^0, Z_0)$ to $C^an(\mathring{\mathbb{Z}}_0)$ via this map, the above formula becomes

$$[J_P(V)_b^r]_b = \lim_{\nu} K \{\{\nu, \nu^{-1}\}\} \otimes_K W_n \otimes C^an(X_n).$$

The action of $z$ on $W_n \otimes_K C^an(X_n)$ is a $C^an(X_n)$-compact morphism of an orthonormalizable $C^an(X_n)$-Banach module, so the result follows by the methods of [Buz07].

5 Change of parabolic

We now consider the problem of relating the geometric objects arising from the above construction for two distinct parabolic subgroups.

5.1 Transitivity of Jacquet functors

Let us recall the definition of the finite-slope-part functor, which we have already seen in the previous section. We let $Z$ be a topologically finitely generated abelian locally $L$-analytic group, and $Z^+$ an open submonoid of $Z$ which generates $Z$ as a group. Then we have the following functor $\text{Rep}_{\text{top}, c}(Z^+) \rightarrow \text{Rep}_{\text{top}, c}(Z)$:

**Definition 5.1** ([Eme06a, 3.2.1]). For any object $V \in \text{Rep}_{\text{top}, c}(Z^+)$, we define

$$V_{b} = \mathcal{L}_{b, Z^+}(C^an(\mathring{Z}), V),$$

Documenta Mathematica 16 (2011) 1–31
endowed with the action of $Z$ on the first factor.

**Lemma 5.2.** Let $Z$ be a topologically finitely generated abelian group and $Y$ a closed subgroup, and suppose $Y^+$ and $Z^+$ are submonoids of $Y$ and $Z$ satisfying the conditions above, with $Y^+ \subseteq Y \cap Z^+$. Then for all $V \in \text{Rep}_{\text{top}}(Z^+)$, the natural map $V_{Y^+\text{-fs}} \to V$ induces an isomorphism

$$(V_{Y^+\text{-fs}})_Z \text{-fs} \cong V_{Z^+\text{-fs}}.$$ 

*Proof.* Consider the canonical $Z^+$-equivariant map $V_{Z^+\text{-fs}} \to V$. We note that $V_{Z^+\text{-fs}}$ is in $\text{Rep}_{\text{la}}(Z)$, and hence a fortiori in $\text{Rep}_{\text{la}}(Y)$. Hence by the universal property of $[\text{Eme06a}, 3.2.4(ii)]$, the above map factors through $V_{Y^+\text{-fs}}$. The factored map is still $Z^+$-equivariant, so by a second application of the universal property it factors through $(V_{Y^+\text{-fs}})_Z \text{-fs}$. This gives a continuous $Z^+$-equivariant map $V_{Z^+\text{-fs}} \to (V_{Y^+\text{-fs}})_Z \text{-fs}$, which is clearly inverse to the map in the statement of the proposition.

Now suppose $P_1$ and $P_2$ are parabolic subgroups of the reductive group $G$ over $L$, with $P_1 \supseteq P_2$. We let $N_1, N_2$ be their unipotent radicals, so we have a chain of inclusions $G \supseteq P_1 \supseteq P_2 \supseteq N_2 \supseteq N_1$.

Let us choose a Levi subgroup $M_2$ of $P_2$, so $P_2 = N_2 \rtimes M_2$. There is a unique Levi subgroup $M_1$ of $P_1$ containing $M_2$; and $P_{12} = P_2 \cap M_1$ is a parabolic subgroup of $M_1$ of which $M_2$ is also a Levi factor. We write $Z_1, Z_2$ for the centres of $M_1$ and $M_2$.

All of the above are algebraic groups over $L$, so their $L$-points are locally $L$-analytic groups; we denote these groups of points by the corresponding Roman letters.

**Theorem 5.3.**

1. For any unitary admissible continuous $G$-representation $V$, there is a unique isomorphism of admissible continuous $M_2$-representations

$$\text{Ord}_{P_{12}} (\text{Ord}_{P_1} V) = \text{Ord}_{P_2} V$$

commuting with the canonical lifting maps from both sides into $V^{N_2}$.

2. For any essentially admissible locally analytic $G$-representation $V$, there is a unique isomorphism of essentially admissible locally analytic $M_2$-representations

$$J_{P_{12}} (J_{P_1} V) = J_{P_2} V$$

commuting with the canonical lifting maps.

*Proof.* We begin by proving the second statement. We have $N_2 = N_1 \rtimes N_{12}$, where $N_{12} = N_2 \cap M_1$ is the unipotent radical of $P_{12}$. Let $N_{2,0}$ be an open compact subgroup of $N_2$ which has the form $N_{1,0} \rtimes N_{12,0}$, for open compact subgroups of the two factors; such subgroups certainly exist, since the conjugation action of $N_1$ on $N_{12}$ is continuous.
For $i = 1, 2$ we write $M_i^+$ for the submonoid of elements $m \in M_i$ for which $mN_{1,0}m^{-1} \subseteq N_{1,0}$ and $m^{-1}N_{i,0}m \subseteq N_{i,0}$, and $Z_i = M_i^+ \cap Z_i$. Then we have $M_i^+ \subseteq M_i^+$, and in particular $Z_i^+ \subseteq Z_2^+$. We have

$$J_{P_i}V = L_{b,Z_i^+} \left( C^{\text{an}}(\hat{Z}_1), V^{N_{1,0}} \right)$$

endowed with the action of $M_1 = Z_1 \times_{Z_1^+} M_1^+$ determined by the actions of $Z_1$ on $C^{\text{an}}(\hat{Z}_1)$ and $M_1^+$ on $V^{N_{1,0}}$. The restriction of this action to $N_{12,0}$ is simply the action on the right factor (since $N_{12,0} \subseteq M_{I,0} \subseteq M_i^+$) and hence

$$(J_{P_i}V)^{N_{12,0}} = L_{b,Z_i^+} \left( C^{\text{an}}(\hat{Z}_1), (V^{N_{1,0}})^{N_{12,0}} \right) = L_{b,Z_1^+} \left( C^{\text{an}}(\hat{Z}_1), V^{N_{12,0}} \right).$$

The Hecke operator construction of [Eme06a, §3.4] gives two actions of $M_1^+$ on $V^{N_{12,0}}$, given respectively by $m \circ \circ v = \pi_{N_{12,0}}mv$ and $m \circ \circ v = \pi_{N_{12,0}}\pi_{N_{12,0}}mv$, where the operators $\pi_{N_{12,0}}$ are the averaging operators with respect to Haar measure on the subgroups $N_{12,0}$. Since $N_{12,0} = N_{12,0} \times N_{12,0}$, and the Haar measure on the product is the product of the Haar measures on the factors, these two actions coincide. Applying the preceding lemma with $Z = Z_2$ and $Y = Z_1$ gives the result.

The statement for the ordinary part functor can be proved along similar lines, but it is easier to note that the functor $\text{Ord}_P$ is right adjoint to the parabolic induction functor $\text{Ind}_P^G$ [Eme10, 4.4.6], for $P$ an opposite parabolic to $P$. Since a composition of adjunctions is an adjunction, it suffices to check instead that $\text{Ind}_P^G \text{Ind}_P^{M_1} U = \text{Ind}_P^{M_1} U$ for any $U \in \text{Rep}_{\text{adm}}(M_2)$. We may identify $C(G, C(M_1, U))$ with $C(G \times M_1, U)$. Evaluating at $1 \in M_1$ gives a map to $C(G, U)$, and it is easy to check that this restricts to an isomorphism between the subspaces realising the two induced representations.

5.2 HECKE ALGEBRAS AND THE CANONICAL LIFTING

We now turn to studying the Jacquet functor in a special case; later we will combine this with the transitivity result above to deduce a general statement. As before, let $G$ be a reductive algebraic group over $L$, and let $H = [G, G]$, a semisimple group. There is a bijection between parabolics of $G$ and $H$, given by $P \mapsto P' = P \cap H$ and $P' \mapsto P = N_G(P')$. We also choose an opposite parabolic $\overline{P}$, determining a Levi subgroup $M$ of $P$, and also a Levi $M'$ of $P'$ in the obvious way. Write $Z_M$, $Z_{M'}$ and $Z_G$ for the centres of these subgroups, so $Z_M$ is isogenous to $Z_{M'} \times Z_G$. As before, we use Roman letters for the $L$-points of these algebraic groups.

Let $H_0$ be an open compact subgroup of $H$. We say $H_0$ is decomposed with respect to $P'$ and $\overline{P}$ if the product of the subgroups $M_0' = H_0 \cap M'$, $N_0 = H_0 \cap N$ and $N_0 = H_0 \cap N$ is $H_0$, for any ordering of the factors.

We say an element $m \in M$ is positive (for $H_0$) if $mN_0m^{-1} \subseteq N_0$ and $m^{-1}N_0m \subseteq N_0$ (see [Bus01, §3.1]). Let $M_{\circ} \subseteq M$ be the monoid of positive elements, and $Z_{M_{\circ}}$ its intersection with $Z_M$; and let $H_{\circ}(M_0')$ denote the
subalgebra of the Hecke algebra $\mathcal{H}(M'_0)$ supported on $M'^+ = M^+ \cap M'$. Note that $M^\oplus$ is contained in the monoid $M^+$ of the previous section, and clearly has finite index therein.

We say an element $z \in Z_M$ is strongly positive if the sequences $z^n N_0 z^{-n}$ and $z^{-n} N_0 z^n$ tend monotonically to $\{1\}$; if this holds, then $z^{-1}$ and $M^\oplus$ together generate $M$. Such elements exist in abundance; any element whose pairing with the roots corresponding to $P$ has sufficiently large valuation will suffice. In particular, there exist strongly positive elements in $Z_M$.

**Proposition 5.4.** For any essentially admissible $G$-representation $V$, we have $J_P(V) = (V^N_0)_{Y-\text{fs}}$, where $Y$ is any closed subgroup of $M$ that contains a strongly positive element. In particular, $J_P(V) = J_P'(V)$.

**Proof.** For any open compact $N_0 \subset N$, [Eme06a, lemma 3.2.29] and the discussion following it shows that $V^N_0$ is in the category denoted therein by $\text{Rep}_{\text{la},c}(Z^+_G)$; thus the hypotheses of [Eme06a, prop 3.2.28] are satisfied for the subgroup $Y = Z^{M'}_G$. The conclusion of that proposition then states that $J_P(V) = (V^N_0)_{Z^+_G - \text{fs}} = (V^N_0)_{Y-\text{fs}}$.

We now lighten the notation somewhat by writing superscript + for $\oplus$, since the proposition shows that the distinction between $M^+$ and $M^\oplus$ is unimportant from the perspective of Jacquet modules.

**Proposition 5.5.** Let $j$ be the morphism $H^+(M'_0) \to \mathcal{H}(H_0)$ constructed in [Bus01, §3.3]. Then the natural inclusion $V^H_0 \hookrightarrow V^{M'_0 N_0}$ is $H^+(M'_0)$-equivariant, where $H^+(M'_0)$ acts via $j$ on the first space and via its inclusion into $\mathcal{H}(M'_0)$ on the second.

**Proof.** Easy check. □

**Proposition 5.6.** For any essentially admissible locally analytic $G$-representation $V$ which is smooth as an $H$-representation, the above inclusion induces an isomorphism

$$(V^H_0)_{Z^+_M - \text{fs}} \cong (V^{M'_0 N_0})_{Z^+_M - \text{fs}} = J_P(V)^{M'_0}.$$

Moreover, there exists a direct sum decomposition

$$V^H_0 = (V^H_0)_{Z^+_M - \text{fs}} \oplus (V^H_0)_{Z^+_M - \text{null}}$$

where the summands are closed subspaces of $V^H_0$, stable under the action of $Z_G$ and $\mathcal{H}(M'_0)$.

**Proof.** Let $Q = V^{M'_0 N_0}/V^H_0$. By the left-exactness of the finite slope part functor [Eme06a, 3.2.6(ii)], there is a closed embedding

$$(V^{M'_0 N_0})_{Z^+_M - \text{fs}}/(V^H_0)_{Z^+_M - \text{fs}} \hookrightarrow Q_{Z^+_M - \text{fs}}.$$
But since $V$ is smooth as an $H$-representation, every element $v \in V M_0 N_0$ is in fact in $V M_U^* N_0$ for some open $U \subseteq \overline{N}$; any such $U$ contains a $Z_{M_0}^+$-conjugate of $N_0$, so there is some $z \in Z^+$ such that $z v \in V N_0^* M_0^*$. Our hypothesis that $H_0$ is decomposed implies that the averaging operator $\pi_{N_0} : V^n \to V N_0$ preserves $V \pi_{N_0} M_0^*$, so $z v = \pi_{N_0}(z v) \in V H_0$. Therefore $Q$ is $Z_{M_0}^+$-torsion, and thus clearly $Q_{Z_{M_0}^+} = 0$.

For the second statement, let $z$ be any strongly positive element of $Z_{M_0}$. By [Bus01, Theorem 1], there exists an integer $n$ (depending only on $P$, $H_0$ and $z$) such that for any smooth $H$-representation $V$, the action of $z$ on $V H_0$ via $j$ satisfies

$$V H_0 = z^n V H_0 \oplus \ker(z^n | V H_0),$$

with $z$ invertible on the subspace $z^n V H_0$. For representations $V$ as in the statement, the subspace $\ker(z^n | V H_0)$ is clearly closed, and moreover $z^n$ gives a continuous map from the essentially admissible $Z_G$-representation $V H_0$ to itself, so its image is also closed.

In particular, since $V H_0$ is an essentially admissible $Z_G$-representation, $J_P(V) M_0^*$ is essentially admissible as a $Z_G$-representation, not just as a representation of the larger group $Z_G \times Z_{M_0} / (Z_{M_0} \cap H_0)$.

Remark. If $H_0$ satisfies the stronger conditions of [Bus01, §1.2], we obtain a finer decomposition of $V H_0$ into a direct sum of closed $Z_G$-subrepresentations corresponding to Bernstein components of $H$.

5.3 Jacquet modules of locally isotypical representations

We now extend the results on $H$-smooth representations above to certain locally $H$-isotypical representations.

Proposition 5.7. If $W$ is a twist of an absolutely irreducible algebraic representation of $G$, and $P = M N$ is a parabolic subgroup of $G$ with $[M, M] = D$, then $\text{End}_D(W N) = K$, so in particular the $M$-representation $W N$ is $D$-good.

Proof. The twist is of no consequence, so suppose that $W$ is algebraic. Let us choose a maximal torus $T$ in $M$, and a field $K' \supset K$ over which $M$ is split; then there is a Borel subgroup $B \subseteq P$ defined over $K'$ with Levi factor $T$. The theory of highest weights then shows that $W$ is absolutely irreducible if and only if the highest weight space of $W$ is 1-dimensional; applying this condition to $W$ and to the $M$-representation $W N$, we deduce that $W N$ is absolutely irreducible as an $M$-representation. Since $M$ is isogenous to $D \times Z(M)$ and all absolutely irreducible representations of $Z(M)$ are one-dimensional, it follows that $W N$ is in fact absolutely irreducible as a $D$-representation.

Proposition 5.8. If $W \in \text{Rep}_{p a c}(G)$ is $H$-good, with $B = \text{End}_B(W) = \text{End}_G W$, and furthermore $W^n = W N$, then for any $X \in \text{Rep}_{p a c}(G)$ which is smooth as an $H$-representation and has a right action of $B$, we have

$$J_P(X \otimes_B W) = J_P(X) \otimes_B W N.$$
Proof. Compare [Eme06a, 4.3.4]. Since $X$ is smooth as an $H$-representation it is certainly smooth as an $N$-representation. Arguing as in the proof of proposition 2.6, we have $(X \otimes_B W)^{N_0} = X^{N_0} \otimes_B W^{N_0}$, which by assumption equals $X^{N_0} \otimes_B W^N$. Passing to finite-slope parts now yields the result.

The condition $W^n = W^N$ is certainly satisfied for any $W$ that is algebraic as a representation of $N$ (since any open subgroup of $N$ is Zariski-dense).

**Proposition 5.9.** Let $W$ be a twist of an absolutely irreducible algebraic representation of $G$, and let $V \in \text{Rep}_{la,c}(P)$ be locally $(H,W)$-isotypical. Then $J_P(V)$ is locally $(D,W^N)$-isotypical, and

$$\text{Hom}_D(W^N, J_P(V)) = J_P(\text{Hom}_H(W, V)).$$

Proof. Let $X = \text{Hom}_H(W, V)$. By proposition 2.6, we have $V = X \otimes_K W$; so by proposition 5.8 and the remark following, $J_P(V) = J_P(X) \otimes_K W^N$. Since $W^N$ is $D$-good, we can apply the converse implication of proposition 2.6 to deduce that $J_P(X) = \text{Hom}_D(W^N, J_P(V))$ as required.

5.4 Combining the constructions

We now summarize the results of the above analysis.

**Theorem 5.10.** For any $V \in \text{Rep}_{\text{ess}}(G)$, we have:

1. For any parabolic subgroup $P \subseteq G$ with Levi subgroup $M$, any finite-dimensional $W \in \text{Rep}_{la,c}(M)$, and any open compact subgroup $U \subseteq D = [M, M]$, there is a coherent sheaf $\mathcal{F}(V, P, W, U)$ on $\hat{Z}(M)$ with a right action of $\mathcal{H}(U)$, whose fibre at a character $\chi \in \hat{Z}(M)$ is isomorphic (as a right $\mathcal{H}(U)$-module) to the dual of the space $\text{Hom}_D(W, J_P(V))^{\hat{Z}(M) = \chi}$. In particular, a character $\chi$ lies in the subvariety $S(V, P, W, U)$ if and only if this eigenspace is nonzero.

2. If $V \in \text{Rep}_{\text{la,ad}}(G)$, then the projection $S(V, P, W, U) \to \langle \text{Lie } Z \rangle'$ has discrete fibres.

3. If $V$ is isomorphic as an $H$-representation to $C^{la}(H)^m$ for some $m$ and some open compact $H \subseteq G$, then $S(V, P, W, U)$ is equidimensional of dimension $\text{dim } Z$.

4. If $P_1, P_2$ are parabolics with $P_1 \supseteq P_2$ as above, $W$ is an absolutely irreducible algebraic representation of $M_1$, and $U$ is an open compact subgroup of $D_1$ which is decomposed with respect to the parabolic $P_2 \cap D_1$, then there is a decomposition

$$\mathcal{F}(V, P_1, U, W) = \mathcal{F}(V, P_1, U, W)_{Z_2 - \text{null}} \oplus \mathcal{F}(V, P_1, U, W)_{Z_2 - \text{fs}}.$$
where the latter factor is isomorphic to a quotient of the pushforward to $Z_1$ of the sheaf $\mathcal{F}(V, P_2, W^N, U \cap D_2)$ on $Z_2$.

Proof. The only statement still requiring proof is the last one. Let $Y = (J_P V)_{D_1, w-\text{iso}}$. The closed embedding $Y \hookrightarrow J_{P_1}(V)$ induces by functoriality a closed embedding $J_{P_1} Y \hookrightarrow J_{P_1}(J_P V)$. The right-hand side is simply $J_{P_2} V$, by theorem 5.3. Thus we have a closed embedding

$$\text{Hom}_{Z_2}(W^N, J_{P_1} Y) \hookrightarrow \text{Hom}_{Z_2}(W^N, J_{P_2} V).$$

The left-hand side is isomorphic, by proposition 5.9, to $J_{P_2}(\text{Hom}_{M_2}(W, Y))$. We may now apply proposition 5.6 to the $M_1$-representation $\text{Hom}_{M_1}(W, Y) = \text{Hom}_{M_1}(W, J_{P_2} V)$, to deduce that there is a direct sum decomposition

$$\text{Hom}_{U \cap D_2}(W^N, J_{P_2} Y) \subseteq \text{Hom}_{U \cap D_2}(W^N, J_{P_2} V).$$

Dualising, we obtain the stated relation between the sheaves $\mathcal{F}(\ldots)$.

6 Application to completed cohomology

6.1 Construction of eigenvarieties

Let us now fix a number field $F$, a connected reductive group $\mathcal{G}$ over $F$, and a prime $p$ of $F$ above $p$. Let $\mathcal{G} = \mathcal{G} \times_F F_p$, a reductive group over $F_p$, and $G = \mathcal{G}(F_p)$. Let us choose a parabolic subgroup $P$ of $\mathcal{G}$ (not necessarily arising from a parabolic subgroup of $\mathcal{G}$), and set $P = P(F_p)$, and similarly for $M, N, D, Z$ as above. We suppose our base field $L$ is a subfield of $F_p$, so $G, P, M, N, D, Z$ are locally $L$-analytic groups.

We recall from [Eme04, 2.2.16] the construction of the completed cohomology spaces $\hat{H}^i$ for each cohomological degree $i \geq 0$, which are unitary admissible Banach representations of $(G, \mathcal{G}(A_p^\mathbb{R}) \times \pi_0)$, where $\pi_0$ is the group of components of $\mathcal{G}(F \otimes \mathbb{Q})$. The following is immediate from the above:

**Proposition 6.1.** Let $\Gamma = D \times \mathcal{G}(A_p^\mathbb{R}) \times \pi_0$. For any $i \geq 0$, we have:

1. For any $W \in \text{Rep}_{cts, \text{fd}}(M)$, the space

$$\text{Hom}_{D-\text{st}, \text{sm}}(W, \text{Ord}_P \hat{H}^i)$$

is an admissible continuous $(Z, \Gamma)$-representation.
2. For any $W \in \text{Rep}_{\text{la}}(M)$, the space

$$\text{Hom}_{\mathfrak{a}}(W, J_{p} \tilde{H}^{i}_{\text{la}})$$

is an essentially admissible locally $L$-analytic $(Z, \Gamma)$-representation.

Let us fix an open compact subgroup $U \subseteq \Gamma$ (this is the most natural notion of a “tame level” in this context). Then we can use the above result to define an eigenvariety of tame level $U$, closely following [Eme06b, §2.3].

Let $v$ be a (finite or infinite) prime of $S$. We set

$$\Gamma_v = \begin{cases} \mathfrak{G}(F_v) & \text{if } v \nmid \infty \text{ and } v \neq p \\ D & \text{if } v = p \\ \pi_0(\mathfrak{G}(F_v)) & \text{if } v \mid \infty. \end{cases}$$

Then $\Gamma = \prod_v \Gamma_v$. Let us set $U_v = U \cap \Gamma_v$. We say $v$ is unramified (for $U$) if $v$ is finite, $v \neq p$, and $U_v$ is a hyperspecial maximal compact subgroup of $\Gamma_v$. Let $S$ be the (clearly finite) set of ramified primes, and $\Gamma_S = \prod_{v \in S} \Gamma_v$.

It is easy to see that $U = U_S \times U^S$, where $U^S = U \cap \Gamma^S$ and similarly $U_S = U \cap \Gamma_S$, and hence we have a tensor product decomposition of Hecke algebras

$$\mathcal{H}(\Gamma//U) = \mathcal{H}(\Gamma_S//U_S) \otimes \mathcal{H}(\Gamma^S//U^S) =: \mathcal{H}^{\text{ram}} \otimes \mathcal{H}^{\text{sph}}.$$ 

As is well known, the algebra $\mathcal{H}^{\text{sph}}$ is commutative (but not finitely generated over $K$), while $\mathcal{H}^{\text{ram}}$ is finitely generated (but not commutative in general).

By construction, $\mathcal{H}(\Gamma//U)$ acts on the essentially admissible $Z$-representation $\text{Hom}_{\mathfrak{a}}(W, J_{p} \tilde{H}^{i}_{\text{la}})$, and hence it also acts on the corresponding sheaf $\mathcal{F}(i,P,W,U)$ on $\hat{Z}$.

**Definition 6.2.** Let $\mathcal{E}(i, P, W, U)$ be the relative spectrum $\text{Spec} \mathcal{A}$, where $\mathcal{A}$ is the $O_{\hat{Z}}$-subsheaf of $\text{End} \mathcal{F}(i, P, W, U)$ generated by the image of $\mathcal{H}^{\text{sph}}$.

For the definition of the relative spectrum, see [Con06, Thm 2.2.5]. By definition $\mathcal{E}(i, P, W, U)$ is a rigid space over $K$, endowed with a finite morphism $\pi : \mathcal{E}(i, P, W, U) \to \hat{Z}$ and an isomorphism of sheaves of $O_{\hat{Z}}$-algebras $\mathcal{A} \cong \pi_* O_{\mathcal{E}(i, P, W, U)}$. Consequently, $\mathcal{F}(i, P, W, U)$ lifts to a sheaf $\mathcal{F}(i, P, W, U)$ on $\mathcal{E}(i, P, W, U)$.

We can regard $\mathcal{E}(i, P, W, U)$ as a subvariety of $\hat{Z}_K \times \text{Spec} \mathcal{H}^{\text{sph}}$ (although the latter will not be a rigid space if $\mathfrak{G}$ is not the trivial group); in particular, a $K$-point of $\mathcal{E}(i, P, W, U)$ gives rise to a homomorphism $\lambda : \mathcal{H}^{\text{sph}} \to K$.

We record the following properties of this construction, which are precisely analogous to [Eme06b, 2.3.3]:

**Theorem 6.3.**
1. The natural projection $E(i, P, W, U) \to \mathfrak{g}'$ has discrete fibres. In particular, the dimension of $E(i, P, W, U)$ is at most equal to the dimension of $Z$.

2. The action of $\mathcal{H}_{\text{ram}}$ on $F(i, P, W, U)$ lifts to an action on $\mathcal{F}(i, P, W, U)$, and the fibre of $\mathcal{F}(i, P, W, U)$ at a point $(\chi, \lambda) \in \hat{Z} \times \text{Spec} \mathcal{H}_{\text{ram}}$ is isomorphic as a right $\mathcal{H}_{\text{ram}}$-module to the dual of the $(Z = \chi, \mathcal{H}_{\text{ram}} = \lambda)$-eigenspace of $\text{Hom}(W, J_P \mathcal{H}_{\text{la}})$. In particular, the point $(\chi, \lambda)$ lies in $E(i, P, W, U)$ if and only if this eigenspace is non-zero.

3. If there is a compact open subgroup $G_0 \subseteq G$ such that $(\mathcal{H}_{\text{la}})_{U^p}$ is isomorphic as a $G_0$-representation to a finite direct sum of copies of $C^{\text{la}}(G_0)$ (where $U^p = U \cap \mathcal{G}(\mathbb{A}_F^p)$), then $E(i, P, W, U)$ is equidimensional, of dimension equal to the rank of $Z$.

Remark. The hypothesis in the last point above is always satisfied when $i = 0$ and $\mathcal{G}(F \otimes \mathbb{R})$ is compact, since for any open compact subgroup $U^p \subseteq \mathcal{G}(\mathbb{A}_F^p)$, the image of $G(F) \cap U^p$ in $G$ is a discrete cocompact subgroup $\Lambda$, and the $U^p$-invariants $\mathcal{H}_{\text{ram}}^p(U^p)$ are isomorphic as a representation of $G$ and as a $\mathcal{H}(U^p)$-module to $C(\Lambda \setminus G)$. This case is considered extensively in an earlier publication of the second author [Loe11].

Now let us suppose $G$ is split over $K$, and fix an irreducible (and therefore absolutely irreducible) algebraic representation $W$ of $M$. We let $\Pi(P, W, U)$ denote the set of irreducible smooth $G(\mathbb{A}_F^p) \times \pi_0$-representations $\pi_f$ such that $J_P(\pi_f)^U \neq 0$, and $\pi_f$ appears as a subquotient of the cohomology space $H^1(\mathcal{V}_X)$ of [Eme06b, §2.2] for some irreducible algebraic representation $X$ of $G$ with $(X')^N \equiv W \otimes \chi$ for some character $\chi$. To any such $\pi_f$, we may associate the point $(\theta_\chi, \lambda) \in \hat{Z} \times \text{Spec} \mathcal{H}_{\text{ram}}$, where $\theta$ is the smooth character by which $Z$ acts on $J_P(\pi_f)$, and $\lambda$ the character by which $\mathcal{H}_{\text{ram}}$ acts on $J_P(\pi_f)^U$. Let $E(i, P, W, U)_{\text{cl}}$ denote the set of points of $\hat{Z} \times \text{Spec} \mathcal{H}_{\text{ram}}$ obtained in this way from representations $\pi_f \in \Pi(i, P, W, U)$.

COROLLARY 6.4. If the map (1.1) is an isomorphism for all irreducible algebraic representations $X$ such that $(X')^N$ is a twist of $W$, then $E(i, P, W, U)_{\text{cl}} \subseteq E(i, P, W, U)$. In particular, the Zariski closure of $E(i, P, W, U)_{\text{cl}}$ has dimension at most $\dim Z$.

Proof. Let $\pi_f \in \Pi(i, P, W, U)$. Then the locally algebraic $(G, \mathcal{G}(\mathbb{A}_F^p) \times \pi_0)$-representation $\pi_f \otimes X'$ appears in $H^1(\mathcal{V}_X) \otimes_K X'$. By assumption, the latter embeds as a closed subrepresentation of $\mathcal{H}_{\text{la}}^i$. The Jacquet functor is exact restricted to locally $X'$-algebraic representations (since this is so for smooth representations). Moreover, the functor $\text{Hom}_3(W, -)$ is exact restricted to locally $W$-algebraic representations, so $\text{Hom}_3[W, J_P(\pi_f) \otimes_K (X')^N]$ appears as a subquotient of $\text{Hom}_3[W, J_P(\mathcal{H}_{\text{la}})]$. Since $(X')^N = W \otimes \chi$, the former space is simply $J_P(\pi_f) \otimes_K \chi$, so the point $(\theta_\chi, \lambda)$ appears in $E(i, P, W, U)$ as required.

\hfill $\Box$
Remarks. 1. The entire construction can also be carried out with the spaces $\hat{H}^i$ replaced by the compactly supported versions $\hat{H}^i_{\text{par}}$ or the parabolic versions $\hat{H}^i_{\text{par}}$; we then obtain analogues of the above proposition for the compactly supported or parabolic cohomology of the arithmetic quotients.

2. It suffices to check that the map (1.1) is an isomorphism for $L = \mathbb{Q}_p$. This is known to hold in many cases, e.g. in degree $i = 0$ for any $\mathfrak{G}$, and in degree 1 for $\text{GL}_2(\mathbb{Q})$ (as shown in [Eme06b]) or for a semisimple and simply connected group (as shown by the first author in [Hil07]). The “weak Emerton criterion” of [Hil07, defn. 2] suffices to prove corollary 6.4 when $W$ is not a character; this is known in many more cases, e.g. when $i = 2$ and the congruence kernel of $\mathfrak{G}$ is finite. When $W$ is a character $\chi : M \to \mathbb{G}_m$, the weak Emerton criterion implies that the points $E(i, P, W, U)_\mathfrak{G}$ are contained in the union of $E(i, P, W, U)$ and the single point $(\chi^{-1}, 1)$.

Theorem 6.5. Suppose $P_1 \supset P_2$ are two parabolics, and $U = U^{(p)} \times U_p$, where $U^{(p)} \subseteq \mathfrak{G}(\mathbb{A}_f^\mathfrak{G}) \times \pi_0$ and $U_p \subseteq D_1$ is decomposed with respect to $P_2 \cap D_1$. Then $E(i, P_1, W, U)$ is equal to a union of two closed subvarieties

$$E(i, P_1, W, U)_{P_2 \setminus \text{fs}} \cap E(i, P_1, W, U)_{P_2 \setminus \text{null}},$$

which are respectively equipped with sheaves of $\mathcal{H}^{\text{ram}}$-modules $\mathcal{F}(i, P, W, U)_{P_2 \setminus \text{fs}}$ and $\mathcal{F}(i, P, W, U)_{P_2 \setminus \text{null}}$ whose direct sum is $\mathcal{F}(i, P, W, U)$.

The element of $\mathcal{H}^{\text{ram}}$ corresponding to any strictly positive element of $Z_2$ acts invertibly on $\mathcal{F}(i, P, W, U)_{P_2 \setminus \text{fs}}$ and nilpotently on $\mathcal{F}(i, P, W, U)_{P_2 \setminus \text{null}}$; and there is a subvariety of $E(i, P_2, W^{N_{12}}, U \cap D_2)$ whose image in $\tilde{Z}_1 \times \text{Spec} \mathcal{H}^{\text{par}}$ coincides with $E(i, P_1, W, U)_{P_2 \setminus \text{fs}}$.

Proof. By theorem 5.10, we may decompose $\mathcal{F}(i, P_1, W, U)$ as a direct sum of a null part and a finite slope part; this decomposition is clearly functorial, and hence it is preserved by the action of the Hecke algebra $\mathcal{H}^{\text{par}}$, so we may define the spaces $E(i, P_1, W, U)_{P_2 \setminus \text{fs}}$ and $E(i, P_1, W, U)_{P_2 \setminus \text{null}}$ to be the relative spectra of the Hecke algebra acting on the two summands.

For the final statement, we note that there is a quotient $Q$ of $\mathcal{F}(i, P_2, W^{N_{12}}, U \cap D_2)$, corresponding to the $Z_2$-subrepresentation

$$J_{P_2}(\text{Hom}_{Z_2}(W, J_{P_1}^t \hat{H}^i_{\text{par}}))^{\cup M_2} \subseteq \text{Hom}_{Z_2}(W^{N_{12}}, J_{P_2}^t \hat{H}^i_{\text{par}})^{\cup D_2}$$

such that the pushforward of $Q$ to $\tilde{Z}_1$ is isomorphic to $\mathcal{F}(i, P_1, W, U)_{P_2 \setminus \text{fs}}$.

This isomorphism clearly commutes with the action of $\mathcal{H}^{\text{par}}$ on both sides, from which the result follows. \qed

7 Acknowledgements

The second author would like to thank Kevin Buzzard and Matthew Emerton for illuminating conversations; the participants in the autumn 2009 $^cP$-adic
Richard Hill and David Loeffler

Analysis” study group at DPMMS, University of Cambridge, for their assistance in learning the necessary theory; and Sarah for her love and support.

References

[Bou87] Nicolas Bourbaki, Topological vector spaces. Chapters 1–5, Elements of Mathematics (Berlin), Springer-Verlag, Berlin, 1987, Translated from the French by H. G. Eggleston and S. Madan. MR 910295.

[Bus01] Colin J. Bushnell, Representations of reductive $p$-adic groups: localization of Hecke algebras and applications, J. London Math. Soc. (2) 63 (2001), no. 2, 364–386. MR 1810135.

[Buz07] Kevin Buzzard, Eigenvarieties, $L$-functions and Galois representations (Durham, 2004), London Math. Soc. Lecture Notes, vol. 320, Cambridge Univ. Press, 2007, pp. 59–120. MR 2392353.

[Col96] Robert F. Coleman, Classical and overconvergent modular forms, Invent. Math. 124 (1996), no. 1–3, 215–241. MR 1369416.

[Col97] ______, $p$-adic Banach spaces and families of modular forms, Invent. Math. 127 (1997), no. 3, 417–479. MR 1431135.

[CM98] Robert F. Coleman and Barry Mazur, The eigencurve, Galois representations in arithmetic algebraic geometry (Durham, 1996), London Math. Soc. Lecture Notes, vol. 254, Cambridge Univ. Press, 1998, pp. 1–113. MR 1696469.

[Con06] Brian Conrad, Relative ampleness in rigid geometry, Ann. Inst. Fourier (Grenoble) 56 (2006), no. 4, 1049–1126. MR 2266885.

[Eme04] Matthew Emerton, Locally analytic vectors in representations of locally $p$-adic analytic groups, Mem. Am. Math. Soc. (to appear), 2004.

[Eme06a] ______, Jacquet modules of locally analytic representations of $p$-adic reductive groups. I. Construction and first properties, Ann. Sci. École Norm. Sup. (4) 39 (2006), no. 5, 775–839. MR 2292633.

[Eme06b] ______, On the interpolation of systems of eigenvalues attached to automorphic Hecke eigenforms, Invent. Math. 164 (2006), no. 1, 1–84. MR 2207783.

[Eme10] ______, Ordinary parts of admissible representations of $p$-adic reductive groups. I. Construction and first properties, Astérisque 331 (2010), 355–402. MR 2667882.

[Hil07] Richard Hill, Construction of eigenvarieties in small cohomological dimensions for semi-simple, simply connected groups, Doc. Math. 12 (2007), 363–397. MR 2365907.

[Loe11] David Loeffler, Overconvergent algebraic automorphic forms, Proc. London Math. Soc. 102 (2011), no. 2, 193–228.
[ST02a] Peter Schneider and Jeremy Teitelbaum, *Banach space representations and Iwasawa theory*, Israel J. Math. 127 (2002), 359–380. MR 1900706.

[ST02b] ———, *Locally analytic distributions and p-adic representation theory, with applications to GL2*, J. Am. Math. Soc. 15 (2002), no. 2, 443–468. MR 1887640.

[ST03] ———, *Algebras of p-adic distributions and admissible representations*, Invent. Math. 153 (2003), no. 1, 145–196. MR 1990669.

[Ser92] Jean-Pierre Serre, *Lie algebras and Lie groups (1964 lectures given at Harvard University)*, second ed., Lecture Notes in Mathematics, vol. 1500, Springer-Verlag, Berlin, 1992. MR 1176100.

Richard Hill
Department of Mathematics
University College London
Gower Street
London WC1E 6BT, UK
rih@math.ucl.ac.uk

David Loeffler
Warwick Mathematics Institute
Zeeman Building
University of Warwick
Coventry CV4 7AL, UK
d.loeffler.01@cantab.net
LOG-GROWTH FILTRATION
AND FROBENIUS SLOPE FILTRATION OF $F$-ISOCRYSTALS
AT THE GENERIC AND SPECIAL POINTS

BRUNO CHIARELLOTTO$^1$ AND NOBUO TSUZUKI$^2$

Received: October 9, 2009
Revised: February 25, 2010
Communicated by Peter Schneider

Abstract. We study, locally on a curve of characteristic $p > 0$, the relation between the log-growth filtration and the Frobenius slope filtration for $F$-isocrystals, which we will indicate as $\varphi$-$\nabla$-modules, both at the generic point and at the special point. We prove that a bounded $\varphi$-$\nabla$-module at the generic point is a direct sum of pure $\varphi$-$\nabla$-modules. By this splitting of Frobenius slope filtration for bounded modules we will introduce a filtration for $\varphi$-$\nabla$-modules (PBQ filtration). We solve our conjectures of comparison of the log-growth filtration and the Frobenius slope filtration at the special point for particular $\varphi$-$\nabla$-modules (HPBQ modules). Moreover we prove the analogous comparison conjecture for PBQ modules at the generic point. These comparison conjectures were stated in our previous work [CT09]. Using PBQ filtrations for $\varphi$-$\nabla$-modules, we conclude that our conjecture of comparison of the log-growth filtration and the Frobenius slope filtration at the special point implies Dwork’s conjecture, that is, the special log-growth polygon is above the generic log-growth polygon including the coincidence of both end points.

2010 Mathematics Subject Classification: Primary 12H25; Secondary 11S80.
Keywords and Phrases: $p$-adic differential equations, logarithmic growth, Frobenius slopes, Newton polygon.

$^1$The author is supported by Progetto di Ateneo Padova “Rapresentations and Motives” and Progetto Eccellenza “Differential methods in arithmetic geometry and algebra” (Cariparo Found.).
$^2$The author is supported by Grant-in-Aid for Exploratory Research and Core-to-Core Program 18005 of JSPS.
1 Introduction

The local behavior of p-adic linear differential equations is, in one sense, very easy. If the equation has a geometric origin (i.e., if it is furnished with a Frobenius structure), then the radius of convergence of solutions at any nonsingular
point is at least 1. In general, the $p$-adic norm of the coefficients $a_n$ in the Taylor series of a solution is an increasing function on $n$. However, one knows that some solutions are $p$-adically integral power series. B. Dwork discovered these mysterious phenomena and introduced a measure, called logarithmic growth (or log-growth, for simplicity), for power series in order to investigate this delicate difference systematically (see [Dw73] and [Ka73, Section 7]). He studied the log-growth of solutions of $p$-adic linear differential equations both at the generic point and at special points (see [Ro75], [Ch83]), and asked whether the behaviors are similar to those of Frobenius slopes or not. He conjectured that the Newton polygon of log-growth of solutions at a special point is above the Newton polygon of log-growth of solutions at the generic point [Dw73, Conjecture 2]. We refer to it as Conjecture $\text{LG}_{Dw}$ when there are not Frobenius structures, and as Conjecture $\text{LGF}_{Dw}$ where there are Frobenius structures (see Conjecture 2.7). He also proved that the Newton polygon of log-growth of solutions at the generic (resp. special) point coincides with the Newton polygon of Frobenius slopes in the case of hypergeometric Frobenius-differential systems if the systems are nonconstant, thus establishing the conjecture in these cases [Dw82, 9.6, 9.7, 16.9].

On the other hand P. Robba studied the generic log-growth of differential modules defined over the completion of $\mathbb{Q}$($x$) under the $p$-adic Gauss norm by introducing a filtration on them via $p$-adic functional analysis [Ro75] (see Theorem 2.2). His theory works on more general $p$-adically complete fields, for example our field $E$.

Let $k$ be a field of characteristic $p > 0$, let $\mathcal{V}$ be a discrete valuation ring with residue field $k$, and let $K$ be the field of fractions of $\mathcal{V}$ such that the characteristic of $K$ is 0. In [CT09] we studied Dwork’s problem on the log-growth for $\varphi$-$\nabla$-modules over $E$ or $K[x]_0$ which should be seen as localizations of $F$-isocrystals on a curve over $k$ with coefficients in $K$. Here $K[x]_0$ is the ring of bounded functions on the unit disk around $x = 0$, $E$ is the $p$-adically complete field which is the field of fractions of $K$, and $\varphi$ (resp. $\nabla$) indicates the Frobenius structure (resp. the connection) (See the notation and terminology introduced in Section 2). We gave careful attention to Dwork’s result on the comparison between the log-growth and the Frobenius slopes of hypergeometric Frobenius-differential equations. We compared the log-growth and the Frobenius slopes at the level of filtrations.

Let $M$ be a $\varphi$-$\nabla$-module over $K[x]_0$. Let $M_{\eta} = M \otimes_{K[x]_0} E$ be a $\varphi$-$\nabla$-module over $E$ which is the generic fiber of $M$ and let $V(M)$ be the $\varphi$-module over $K$ consisting of horizontal sections on the open unit disk. Denote by $M_{\eta}^\lambda$ the log-growth filtration on $M_{\eta}$ at the generic point indexed by $\lambda \in \mathbb{R}$, and by $V(M)^\lambda$ be the log-growth filtration with real indices on the $\varphi$-module $V(M)$. Furthermore, let $S_\lambda(\cdot)$ be the Frobenius slope filtration such that $S_\lambda(\cdot)/S_{\lambda^-}(\cdot)$ is pure of slope $\lambda$.

We proved that the log-growth filtration is included in the orthogonal part of the Frobenius slope filtration of the dual module under the natural perfect pairing $M_{\eta} \otimes E M_{\eta}^\gamma \rightarrow E$ (resp. $V(M) \otimes K V(M^\gamma) \rightarrow K$) at the generic point.
Chiarellotto, B. and Tsuzuki, N. 

(resp. the special point) [CT09, Theorem 6.17] (see the precise form in Theorem 2.3):

\[ M_\eta \subset (S_{\lambda-\lambda_{\text{max}}}(M_\eta'))^\perp \quad \text{(resp. } V(M)^\lambda \subset (S_{\lambda-\lambda_{\text{max}}}(V(M')))^\perp) \]

for any \( \lambda \in \mathbb{R} \) if \( \lambda_{\text{max}} \) is the highest Frobenius slope of \( M_\eta \). We then conjectured: (a) the rationality of log-breaks \( \lambda \) (both at the generic and special fibers) and (b) if the bounded quotient \( M_\eta/M_0^\eta \) is pure as a \( \phi \)-module then the inclusion relation becomes equality both at the generic and special points [CT09, Conjectures 6.8, 6.9]. Note that there are examples with irrational breaks, and that both \( M_\lambda \supset \bigcirc \bigcirc \) \( M_\lambda \) and \( M_\lambda \supset \bigcirc \bigcirc \) \( M_\lambda \) can indeed occur for log-growth filtrations in absence of Frobenius structures [CT09, Examples 5.3, 5.4]. We state the precise forms of our conjectures in Conjecture 2.4 on \( E \) and Conjecture 2.5 on \( K[[x]]_0 \), and denote the conjectures by \( \text{LGF}_E \) and \( \text{LGF}_{K[[x]]_0} \), respectively. We have proved our conjectures \( \text{LGF}_E \) and \( \text{LGF}_{K[[x]]_0} \) if \( M \) is of rank \( \leq 2 \) [CT09, Theorem 7.1, Corollary 7.2], and then we established Dwork’s conjecture \( \text{LGF}_{\text{Dw}} \) if \( M \) is of rank \( \leq 2 \) [CT09, Corollary 7.3].

Let us now explain the results in the present paper. First we characterize bounded \( \phi\nabla \)-modules over \( E \) by using Frobenius structures (Theorem 4.1):

1. A bounded \( \phi\nabla \)-module \( M \) over \( E \) (i.e., \( M^0 = 0 \), which means that all the solutions on the generic disk are bounded) is isomorphic to a direct sum of several pure \( \phi\nabla \)-modules if the residue field \( k \) of \( V \) is perfect.

Note that the assertion corresponding to (1) is trivial for a \( \phi\nabla \)-module \( M \) over \( K[[x]]_0 \) such that \( M_\eta \) is bounded by Christol’s transfer theorem (see [CT09, Proposition 4.3]). This characterization implies the existence of a unique increasing filtration \( \{P_i(M)\} \) of \( \phi\nabla \)-modules \( M \) over \( E \) such that \( P_i(M)/P_{i-1}(M) \) is the maximally PBQ submodule of \( M/P_{i-1}(M) \) (Corollary 5.5). This filtration is called the PBQ filtration. When we start with a \( \phi\nabla \)-module \( M \) over \( K[[x]]_0 \), we can introduce a similar PBQ filtration for \( M \), i.e., a filtration consisting of \( \phi\nabla \)-submodules over \( K[[x]]_0 \) whose generic fibers will induce the PBQ filtration of the generic fiber \( M_\eta \) (Corollary 5.10). To this end we use an argument of A.J. de Jong in [dJ98] establishing the full faithfulness of the forgetful functor from the category of \( \phi\nabla \)-modules over \( K[[x]]_0 \) to the category of \( \phi\nabla \)-modules over \( E \).

The need to study horizontality behavior for the PBQ condition with respect to the special and generic points leads us to introduce a new condition for \( \phi\nabla \)-modules over \( K[[x]]_0 \), namely, the property of being “horizontally pure of bounded quotient” (which, for simplicity, we abbreviate as HPBQ, cf. Definition 6.1). Then in Theorem 6.5 we prove that

2. our conjecture \( \text{LGF}_{K[[x]]_0} \) (see 2.5) on the comparison between the log-growth filtration and the Frobenius slope filtration at the special point holds for a HPBQ module.
A HPBQ module should be understood as a $\varphi-\nabla$-module for which the bounded quotient is horizontal and pure with respect to the Frobenius. Our method of proof will lead us to introduce the related definition of equislope $\varphi-\nabla$-modules over $K[[x]]_0$ (Definition 6.7): they admit a filtration as $\varphi-\nabla$-modules over $K[[x]]_0$ which induces the Frobenius slope filtration at the generic point. Note that a PBQ equislope object is HPBQ. Using this result, we prove in Theorem 7.1 that

(3) our conjecture $\text{LG}_E$ (see 2.4) on comparison between the log-growth filtration and the Frobenius slope filtration at the generic point holds for PBQ modules over $E$.

Indeed, for a $\varphi-\nabla$-module $M$ over $E$, the induced $\varphi-\nabla$-module $M_\tau = M \otimes_{E_t} E_t[[X-\tau]]_0$ (where $E_t[[X-\tau]]_0$ is the ring of bounded functions on the open unit disk at generic point $\tau$) is equislope. For the proof of comparison for HPBQ modules, we use an explicit calculation of log-growth for solutions of certain Frobenius equations (Lemma 4.8) and a technical induction argument.

For a submodule $L$ of a $\varphi-\nabla$-module $M$ over $E$ with $N = M/L$, the induced right exact sequence

$$L/L^\lambda \to M/M^\lambda \to N/N^\lambda \to 0$$

is also left exact for any $\lambda$ if $L$ is a maximally PBQ submodule of $M$ by Proposition 2.6. Since there do exist PBQ filtrations, the comparison between the log-growth filtrations and the Frobenius slope filtrations for PBQ modules both at the generic point and at the special point implies the rationality of breaks (Theorem 7.2 and Proposition 7.3) as well as Dwork’s conjecture (Theorem 8.1) that the special log-growth polygon lies above the generic log-growth polygon (including the coincidence of both end points):

(4) Our conjecture of comparison between the log-growth filtration and the Frobenius slope filtration at the special point (Conjecture $\text{LG}_K[[\xi]]_0$, 2.5) implies Dwork’s conjecture (Conjecture $\text{LG}_{\text{Dw}}, 2.7$).

As an application, we have the following theorem (Theorem 8.8) without any assumptions.

(5) The coincidence of both log-growth polygons at the generic and special points is equivalent to the coincidence of both Frobenius slope polygons at the generic and special points.

Let us also mention some recent work on log-growth. Y. André ([An08]) proved the conjecture $\text{LG}_{\text{Dw}}$ of Dwork without Frobenius structures, that is, the log-growth polygon at the special point is above the log-growth filtration at the generic point for $\nabla$-modules, but without coincidence of both end points. (Note that his convention on the Newton polygon is different from ours, see Remark 2.8). He used semi-continuity of log-growth on Berkovich spaces. K. Kedlaya
defined the log-growth at the special point for regular singular connections and studied the properties of log-growth [Ke09, Chapter 18].

This paper is organized in the following manner. In Section 2 we recall our notation and results from [CT09]. In Section 3 we establish the independence of the category of $\varphi\nabla$-modules over $E$ (resp. $K[[x]]_0$) of the choices of Frobenius on $E$ (resp. $K[[x]]_0$). In Section 4 we study when the Frobenius slope filtration of $\varphi\nabla$-modules over $E$ is split and prove (1) above. In Section 5 we introduce the notion of PBQ and prove the existence of PBQ filtrations. In Section 6 we study the log-growth filtration for HPBQ $\varphi\nabla$-modules over $E$ is stable under $\sigma$. We also denote by $\sigma$ the unique extension of $\sigma$ on $E$, which is a Frobenius on $E$. In the case we only discuss $\varphi\nabla$-modules over $E$, one can take a Frobenius $\sigma$ on $K$ such that $\sigma(x) \in E$ with $|\sigma(x) - x^q|_0 < 1$.

2 Preliminaries

We fix notation and recall the terminology in [CT09]. We also review Dwork’s conjecture and our conjectures.

2.1 Notation

Let us fix the basic notation which follows from [CT09].

$p$ : a prime number.

$K$ : a complete discrete valuation field of mixed characteristic $(0, p)$.

$\mathcal{V}$ : the ring of integers of $K$.

$k$ : the residue field of $\mathcal{V}$.

$m$ : the maximal ideal of $\mathcal{V}$.

$|\cdot|$ : a $p$-adically absolute value on $K$ and its extension as a valuation field, which is normalized by $|p| = p^{-1}$.

$q$ : a positive power of $p$.

$\sigma$ : $(p)$-Frobenius on $K$, i.e., a continuous lift of $q$-Frobenius endomorphism ($a \mapsto a^q$ on $k$). We suppose the existence of Frobenius on $K$. We also denote by $\sigma$ a $K$-algebra endomorphism on $A_K(0, 1^{-})$, which is an extension of Frobenius on $K$, such that $\sigma(x)$ is bounded and $|\sigma(x) - x^q|_0 < 1$. Then $K[[x]]_0$ is stable under $\sigma$. We also denote by $\sigma$ the unique extension of $\sigma$ on $E$, which is a Frobenius on $E$. In the case we only discuss $\varphi\nabla$-modules over $E$, one can take a Frobenius $\sigma$ on $K$ such that $\sigma(x) \in E$ with $|\sigma(x) - x^q|_0 < 1$. 
$K_{\text{perf}}$: the $p$-adic completion of the inductive limit $K_{\text{perf}}$ of $K_{\text{perf}}^c, K_{\text{perf}}^r, \ldots$. Then $K_{\text{perf}}$ is a complete discrete valuation field such that the residue field of the ring of integers of $K_{\text{perf}}$ is the perfection of $k$ and that the value group of $K_{\text{perf}}$ coincides with the value group of $K$. The Frobenius $\sigma$ uniquely extend to $K_{\text{perf}}$. Moreover, taking the $p$-adic completion $\overline{K^{\text{al}}}$ of the maximally unramified extension $K^{\text{al}}$ of $K_{\text{perf}}$, we have a canonical extension of $K$ as a discrete valuation field with the same value group such that the residue field of the ring of integers is algebraically closed and the Frobenius extends on it. We use the same symbol $\sigma$ for Frobenius on the extension.

$q^\lambda$: an element of $K$ with $\log_q|q^\lambda| = -\lambda$ for a rational number $\lambda$ such that $\sigma(q^\lambda) = q^{\lambda}$. Such a $q^\lambda$ always exists if the residue field $k$ is algebraically closed and $\lambda \in \log_q[K^*]$. In particular, if $k$ is algebraically closed, then there exists an extension $L$ of $K$ as a discrete valuation field with an extension of Frobenius such that $q^\lambda$ is contained in $L$ for a fix $\lambda$. In this paper we freely extend $K$ as above.

$A_K(c, r^-)$: the $K$-algebra of analytic functions on the open disk of radius $r$ at the center $c$, i.e.,

$$A_K(c, r^-) = \left\{ \sum_{n=0}^{\infty} a_n(x-c)^n \in K[[x-c]] \left| \frac{|a_n|}{\gamma} \rightarrow 0 \text{ as } n \rightarrow \infty \right. \right\}.$$ 

$K[x]_0$: the ring of bounded power series over $K$, i.e.,

$$K[x]_0 = \left\{ \sum_{n=0}^{\infty} a_n x^n \in A_K(0,1^-) \left| \sup_n |a_n| < \infty \right. \right\}.$$ 

An element of $K[[x]]_0$ is said to be a bounded function.

$K[x]_\lambda$: the Banach $K$-module of power series of log-growth $\lambda$ in $A_K(0,1^-)$ for a nonnegative real number $\lambda \in \mathbb{R}_{\geq 0}$, i.e.,

$$K[[x]]_\lambda = \left\{ \sum_{n=0}^{\infty} a_n x^n \in A_K(0,1^-) \left| \sup_n |a_n|/(n+1)^\lambda < \infty \right. \right\},$$

with a norm $|\sum_{n=0}^{\infty} a_n x^n|_\lambda = \sup_n |a_n|/(n+1)^\lambda$. $K[[x]]_\lambda$ is a $K[[x]]_0$-modules. $K[x]_\lambda = 0$ for $\lambda < 0$ for the convenient. An element $f \in K[[x]]_\lambda$ which is not contained in $K[x]_\gamma$ for $\gamma < \lambda$ is said to be exactly of log-growth $\lambda$.

$E$: the $p$-adic completion of the field of fractions of $K[x]_0$ under the Gauss norm $|0|$, i.e.,

$$E = \left\{ \sum_{n=-\infty}^{\infty} a_n x^n \left| a_n \in K, \sup_n |a_n| < \infty, |a_n| \rightarrow 0 \text{ as } n \rightarrow -\infty \right. \right\}. $$
$\mathcal{E}$ is a complete discrete valuation field under the Gauss norm $|\cdot|_0$ in fact $K$ is discrete valued. The residue field of the ring $\mathcal{O}_\mathcal{E}$ of integers of $\mathcal{E}$ is $k((x))$.

$t$: a generic point of radius 1.

$\mathcal{E}_t$: the valuation field corresponding to the generic point $t$, i.e., the same field as $\mathcal{E}$ in which $x$ is replaced by $t$: we emphasize $t$ in the notation with the respect to [CT09]. We regard the Frobenius $\sigma$ as a Frobenius on $\mathcal{E}_t$.

$\mathcal{E}_t[X - t]_0$: the ring of bounded functions in $\mathcal{A}_{\mathcal{E}_t}(t, 1^-)$. Then

$$\tau: \mathcal{E} \rightarrow \mathcal{E}_t[X - t]_0 \quad \tau(f) = \sum_{n=0}^{\infty} \frac{1}{n!} \left. \frac{d^n}{dx^n} f \right|_{x-t}(X - t)^n$$

is a $K$-algebra homomorphism which is equivariant under the derivations $\frac{d}{dx}$ and $\frac{d}{dt}$. The Frobenius $\sigma$ on $\mathcal{E}_t[X - t]_0$ is defined by $\sigma$ on $\mathcal{E}_t$ and $\sigma(X - t) = \tau(\sigma(x)) - \sigma(x)|_{x=t}$. $\tau$ is again $\sigma$-equivariant.

For a function $f$ on $R$ and for a matrix $A = (a_{ij})$ with entries in $R$, we define $f(A) = (f(a_{ij}))$. In case where $f$ is a norm $|\cdot|$, then $|A| = \sup_{i,j} |a_{ij}|$. We use 1 (resp. 1$_r$) to denote the unit matrix of suitable degree (resp. of degree $r$). For a decreasing filtration $\{V^\lambda\}$ indexed by the set $\mathbb{R}$ of real numbers, we put

$$V^{\lambda -} = \cap_{\mu<\lambda} V^\mu, \quad V^{\lambda +} = \cup_{\mu>\lambda} V^\mu.$$

We denote by $W^{\lambda -} = \cup_{\mu<\lambda} W^\mu$ and $W^{\lambda +} = \cap_{\mu>\lambda} W^\mu$ the analogous objects for an increasing filtration $\{W^\lambda\}$, respectively.

### 2.2 Terminology

We recall some terminology and results from [CT09].

Let $R$ be either $K$ ($K$ might be $\mathcal{E}$) or $K[x]_0$. A $\varphi$-module over $R$ consists of a free $R$-module $M$ of finite rank and an $R$-linear isomorphism $\varphi: \sigma^* M \rightarrow M$. For a $\varphi$-module over $K$, there is an increasing filtration $\{S_\lambda(M)\}_{\lambda \in \mathbb{R}}$ which is called the Frobenius slope filtration. Then there is a sequence $\lambda_1 < \cdots < \lambda_r$ of real numbers, called the Frobenius slopes of $M$, such that $S_{\lambda_i}(M)/S_{\lambda_i -}(M)$ is pure of slope $\lambda_i$ and $M \otimes \overline{K} \cong \oplus_i S_{\lambda_i}(M) \otimes_K \overline{K}/S_{\lambda_i -}(M) \otimes_K \overline{K}$ is the Dieudonné-Manin decomposition as $\varphi$-modules over $\overline{K}$. We call $\lambda_1$ the first Frobenius slope and $\lambda_r$ the highest Frobenius slope, respectively.

Let $R$ be either $\mathcal{E}$ or $K[x]_0$. A $\varphi$-$\nabla$-module over $R$ consists of a $\varphi$-module $(M, \varphi)$ over $R$ and a $K$-connection $\nabla: M \rightarrow M \otimes_R \Omega_R$, where $\Omega_R = Rdx$, such that $\varphi \circ \sigma^*(\nabla) = \nabla \circ \varphi$. For a basis $(e_1, \cdots, e_r)$, the matrices $A$ and $G$ with entries $R$,

$$\varphi(1 \otimes e_1, \cdots, 1 \otimes e_r) = (e_1, \cdots, e_r)A, \quad \nabla(e_1, \cdots, e_r) = (e_1, \cdots, e_r)Gdx.$$
are called the Frobenius matrix and the connection matrix of \( R \), respectively. Then one has
\[
\frac{d}{dx} A + GA = \left( \frac{d}{dx} \sigma(x) \right) A \sigma(G)
\]
by the horizontality of \( \varphi \). We denote the dual of \( M \) by \( M^\vee \).

Let \( M \) be a \( \varphi - \nabla \)-module over \( K[x]_0 \). We define the \( K \)-space
\[
V(M) = \{ s \in M \otimes_{K[x]} A_K(0, 1^-) | \nabla(s) = 0 \}
\]
of horizontal sections and the \( K \)-space of solutions,
\[
\text{Sol}(M) = \text{Hom}_{K[x]}(M, A_K(0, 1^-)),
\]
on the unit disk. Both \( \dim_K V(M) \) and \( \dim_K \text{Sol}(M) \) equal to \( \text{rank}_{K[x]} M \) by the solvability. If one fixes a basis of \( M \), the solution \( Y \) of the equations
\[
\begin{cases}
A(0) Y = Y A \\
\frac{d}{dx} Y = Y G \\
Y(0) = 1
\end{cases}
\]
in \( A_K(0, 1^-) \) is a solution matrix of \( M \), where \( A(0) \) and \( Y(0) \) are the constant terms of \( A \) and \( Y \), respectively. The log-growth filtration \( \{ V(M)^\lambda \} \lambda \in \mathbb{R} \) is defined by the orthogonal space of the \( K \)-space \( \text{Sol}(M) = \text{Hom}_{K[x]}(M, A_K(0, 1^-)) \) under the natural bilinear perfect pairing
\[
V(M) \times \text{Sol}(M) \to K.
\]

Then \( V(M)^\lambda = 0 \) for \( \lambda \gg 0 \) by the solvability of \( M \) and the log-growth filtration is a decreasing filtration of \( V(M) \) as \( \varphi \)-modules over \( K \). The following proposition allows one to change the coefficient field \( K \) to a suitable extension \( K' \).

**Proposition 2.1 ([CT09, Proposition 1.10])** Let \( M \) be a \( \varphi \)-module over \( K[x]_0 \). For any extension \( K' \) over \( K \) as a complete discrete valuation field with an extension of Frobenius, there is a canonical isomorphism \( V(M) \otimes_{K[x]} K'[x]_0 \cong V(M) \otimes_K K' \) as log-growth filtered \( \varphi \)-modules.

The induced \( \varphi - \nabla \)-module \( M_\eta = M \otimes_{K[x]} \mathcal{E} \) over \( \mathcal{E} \) is said to be the generic fiber of \( M \), and the \( K \)-module \( V(M) \) is called the special fiber of \( M \).

Let \( M \) be a \( \varphi - \nabla \)-module over \( \mathcal{E} \). We denote by \( M_\tau \) the induced \( \varphi - \nabla \)-module \( M \otimes_{\mathcal{E}} \mathcal{E}_t[X - t]^\lambda_{0} \) over \( \mathcal{E}_t[X - t]^\lambda_{0} \). Applying the theory of Robba [Ro75], we have a decreasing filtration \( \{ M_\lambda \} \lambda \in \mathbb{R} \) of \( M \) as \( \varphi - \nabla \)-modules over \( \mathcal{E} \) which is characterized by the following universal property.

**Theorem 2.2 [Ro75, 2.6, 3.5] (See [CT09, Theorem 3.2].)** For any real number \( \lambda \), \( M/M_\lambda \) is the maximum quotient of \( M \) such that all solutions of log-growth \( \lambda \) of \( M_\tau \) on the generic unit disk come from the solutions of \( (M/M_\lambda)_\tau \).
The filtration \( \{M^\lambda\} \) is called the log-growth filtration of \( M \). Note that \( M^\lambda = M \) for \( \lambda < 0 \) by definition and \( M^\lambda = 0 \) for \( \lambda \gg 0 \) by the solvability. The quotient module \( M/M^0 \) is called the bounded quotient, and, in particular, if \( M^0 = 0 \), then \( M \) is called bounded.

Our main theorem in [CT09] is the following:

**Theorem 2.3 ([CT09, Theorem 6.17])**

1. Let \( M \) be a \( \varphi \-\nabla \)-module over \( E \). If \( \lambda_{\text{max}} \) is the highest Frobenius slope of \( M \), then \( M^\lambda \subset \left(S_{\lambda-\lambda_{\text{max}}(M^\vee)}\right)^\perp\). 

2. Let \( M \) be a \( \varphi \-\nabla \)-module over \( K[[x]]_0 \). If \( \lambda_{\text{max}} \) is the highest Frobenius slope of \( M_\eta \), then \( V(M)^\lambda \subset \left(S_{\lambda-\lambda_{\text{max}}(V(M)^\vee)}\right)^\perp\).

Here \( S^\perp \) denotes the orthogonal space of \( S \) under the natural bilinear perfect pairing \( M \otimes E M^\vee \to E \) or \( V(M) \otimes_K V(M^\vee) \to K \).

We conjectured that equalities hold in Theorem 2.3 if \( M \) is PBQ (Definition 5.1) in [CT09], and proved them if \( M \) is of rank \( \leq 2 \) [CT09, Theorem 7.1, Corollary 7.2].

**Conjecture 2.4 ([CT09, Conjectures 6.8])** Let \( M \) be a \( \varphi \-\nabla \)-module over \( E \).

1. All breaks of log-growth filtration of \( M \) are rational and \( M^\lambda = M^\lambda + \) for any \( \lambda \).

2. Let \( \lambda_{\text{max}} \) be the highest Frobenius slope of \( M \). If \( M/M^0 \) is pure as \( \varphi \)-module (PBQ in Definition 5.1 (1)), then \( M^\lambda = \left(S_{\lambda-\lambda_{\text{max}}(M^\vee)}\right)^\perp\).

We denote Conjecture 2.4 above by \( \text{LGF}_E \).

**Conjecture 2.5 ([CT09, Conjectures 6.9])** Let \( M \) be a \( \varphi \-\nabla \)-module over \( K[[x]]_0 \).

1. All breaks of log-growth filtration of \( V(M) \) are rational and \( V(M)^\lambda = V(M)^\lambda + \) for any \( \lambda \).

2. Let \( \lambda_{\text{max}} \) be the highest Frobenius slope of \( M_\eta \). If \( M_\eta/M_\eta^0 \) is pure as \( \varphi \)-module (PBQ in Definition 5.1 (2)), then \( V(M)^\lambda = \left(S_{\lambda-\lambda_{\text{max}}(V(M)^\vee)}\right)^\perp\).

We denote Conjecture 2.5 above by \( \text{LGF}_{K[[x]]_0} \).

Note that we formulate the theorem and the conjecture in the case where \( \lambda_{\text{max}} = 0 \) in [CT09]. However, the theorem holds for an arbitrary \( \lambda_{\text{max}} \) by Proposition 2.1 (and the conjecture should also hold). Moreover, it suffices to establish the conjecture when the residue field \( k \) of \( V \) is algebraically closed.
In section 7 we will reduce the conjecture \( \text{LGF}_E(1) \) (resp. \( \text{LGF}_{K[[x]]_0}(1) \)) to the conjecture \( \text{LGF}_E(2) \) (resp. \( \text{LGF}_{K[[x]]_0}(2) \)) by applying the proposition below to the PBQ filtration which is introduced in section 5. The following proposition is useful for attacking log-growth questions by induction.

**Proposition 2.6** Let \( 0 \to L \to M \to N \to 0 \) be an exact sequence of \( \varphi\nabla \)-modules over \( E \) (resp. \( K[[x]]_0 \)) and let \( \lambda_{\text{max}} \) be the highest Frobenius slope of \( M \) and \( L \) (resp. \( M_\eta \) and \( L_\eta \)).

1. Suppose that \( L^\lambda = (S_{\lambda_{\text{max}}} - \lambda)^\bot \) for \( \lambda \). Then the induced sequence
   \[
   0 \to L/L^\lambda \to M/M^\lambda \to N/N^\lambda \to 0
   \]
   is exact.

2. Suppose that \( V(L)^\lambda = (S_{\lambda_{\text{max}}} - \lambda)^\bot \) for \( \lambda \). Then the induced sequence
   \[
   0 \to V(L)/V(L)^\lambda \to V(M)/V(M)^\lambda \to V(N)/V(N)^\lambda \to 0
   \]
   is exact.

**Proof.** (1) Since
   \[
   L/L^\lambda \to M/M^\lambda \to N/N^\lambda \to 0
   \]
   is right exact by [CT09, Proposition 3.6], we have only to prove the injectivity of the first morphism. There is an inclusion relation
   \[
   M^\lambda \subset (S_{\lambda_{\text{max}}} - \lambda)^\bot = S_{(\lambda_{\text{max}} - \lambda - \lambda)}(M)
   \]
   by Theorem 2.3 and the equality
   \[
   L^\lambda = (S_{\lambda_{\text{max}}} - \lambda)^\bot = S_{(\lambda_{\text{max}} - \lambda - \lambda)}(L).
   \]
   holds by our hypothesis on \( L \). Since the Frobenius slope filtrations are strict for any morphism, the bottom horizontal morphism in the natural commutative diagram
   \[
   \begin{array}{ccc}
   L/L^\lambda & \to & M/M^\lambda \\
   \downarrow & & \downarrow \\
   L/S_{(\lambda_{\text{max}} - \lambda - \lambda)}(L) & \to & M/S_{(\lambda_{\text{max}} - \lambda - \lambda)}(M)
   \end{array}
   \]
   is injective. Hence we have the desired injectivity.

(2) The proof here is similar to that of (1) on replacing [CT09, Proposition 3.6] by [CT09, Proposition 1.8].
2.3 Dwork's conjecture

We recall Dwork's conjecture. We have proved it in the case where $M$ is of rank $\leq 2$ [CT09, Corollary 7.3].

**Conjecture 2.7** ([Dw73, Conjecture 2], [CT09, Conjecture 4.9]) Let $M$ be a $\varphi$-$\nabla$-module over $K[[x]]_0$. Then the special log-growth is above the generic log-growth polygon (with coincidence at both endpoints).

We denote Conjecture 2.7 above by $\text{LGF}_{Dw}$. We will prove that the conjecture $\text{LGF}_{Dw}$ follows from the conjectures $\text{LGF}_{E}$ and $\text{LGF}_{K[[x]]_0}$ in section 8. There is also a version of Dwork's conjecture without Frobenius structures, we denote it by $\text{LG}_{Dw}$.

Let us recall the definition of the log-growth polygon: the generic log-growth polygon is the piecewise linear curve defined by the vertices

$$(0,0), \left(\dim_{E} \frac{M_{\lambda}^1}{M_{\lambda}^{-}}, \lambda_1 \dim_{E} \frac{M_{\lambda}^{1-}}{M_{\lambda}^{1+}}\right), \ldots, \left(\dim_{E} \frac{M_{\lambda}^{1_i-}}{M_{\lambda}^{1_i+}}, \sum_{j=1}^{i} \lambda_j \dim_{E} \frac{M_{\lambda}^{1_j-}}{M_{\lambda}^{1_j+}}\right),$$

$$\cdots, \left(\dim_{E} M_{\eta}, \sum_{j=1}^{r} \lambda_j \dim_{E} \frac{M_{\lambda_j}^{1-}}{M_{\lambda_j}^{1+}}\right),$$

where $0 = \lambda_1 < \cdots < \lambda_r$ are breaks (i.e., $M_{\lambda}^{-} \neq M_{\lambda}^{+}$) of the log-growth filtration of $M_{\eta}$. The special log-growth polygon is defined in the same way using the log-growth filtration of $V(M)$.

**Remark 2.8** (1) The convention of André's polygon of log-growth [An08] is different from ours. His polygon at the generic fiber is $\sum_{j=1}^{r} \lambda_j \dim_{E} \frac{M_{\lambda_j}^{1-}}{M_{\lambda_j}^{1+}}$ below our polygon in the direction of the vertical axis and the starting point of the polygon is $(\dim_{E} M,0)$, and the same at the special fiber. André proved the conjecture $\text{LG}_{Dw}$ except the coincidence of both endpoints in [An08].

(2) If the special log-growth polygon lies above the generic log-growth polygon in both conventions of André's and ours, then both endpoints coincide with each other. However even if this is the case, we cannot prove $M_{\eta}^{\lambda} = M_{\eta}^{\lambda^+}$ (resp. $V(M)^{\lambda} = V(M)^{\lambda^+}$) for a break $\lambda$.

3 Choices of Frobenius

Let us recall the precise form of equivalence between categories of $\varphi$-$\nabla$-modules with respect to different choices of Frobenius on $E$ (resp. $K[[x]]_0$) (see [Ts98a, Section 3.4] for example). We will use it in the next section.

3.1 Comparison morphism $\vartheta_{\sigma_1, \sigma_2}$

Let $\sigma_1$ and $\sigma_2$ be Frobenius maps on $E$ (resp. $K[[x]]_0$) such that the restriction of each $\sigma_i$ to $K$ is the given Frobenius on $K$. Let $M$ be a $\varphi$-$\nabla$-module. We
define an $E$-linear (resp. $K[[x]]_0$-linear) morphism

$$\vartheta_{\sigma_1, \sigma_2} : \sigma_1^* M \to \sigma_2^* M$$

by

$$\vartheta_{\sigma_1, \sigma_2}(a \otimes m) = a \sum_{n=0}^{\infty} \left( \sigma_2(x) - \sigma_1(x) \right)^n \otimes \frac{1}{n!} \nabla^n \left( \frac{d^n}{dx^n} \right)(m).$$

Since $M$ is solvable and $|\sigma_2(x) - \sigma_1(x)| < 1$, the right hand side converges in $\sigma_2^* M$. As a matrix representation, the transformation matrix is

$$H = \sum_{n=0}^{\infty} \sigma_2(G_n)(\sigma_2(x) - \sigma_1(x))^n \frac{n!}{n}$$

for the induced basis $1 \otimes e_1, \ldots, 1 \otimes e_r$, where $G$ is the matrix of connection, $G_0 = 1$ and $G_{n+1} = GG_n + \frac{d}{dx}G_n$ for $n \geq 0$.

**Proposition 3.1** Let $\sigma_1, \sigma_2, \sigma_3, \sigma$ be Frobenius maps of $E$ (resp. $K[[x]]_0$) as above. Then we have the cocycle conditions:

1. $\vartheta_{\sigma_2, \sigma_3} \circ \vartheta_{\sigma_1, \sigma_2} = \vartheta_{\sigma_1, \sigma_3}$.
2. $\vartheta_{\sigma, \sigma} = \text{id}_{\sigma^* M}$.

**Proposition 3.2** Let $M$ be a $\varphi$-$\nabla$-module pure of slope $\lambda$ over $E$ and let $A$ be the Frobenius matrix of $M$ with respect to a basis. Suppose that $|A - q^n1|_0 \leq q^{-\mu}$ for $\mu \geq \lambda$. Then the representation matrix $H$ of the comparison morphism $\vartheta_{\sigma_1, \sigma_2}$ with respect to the bases which are the pull-backs by $\sigma_1$ and $\sigma_2$ respectively, satisfies $|H - 1|_0 < q^{\lambda - \mu}$.

**Proof.** By replacing the Frobenius $\varphi$ by $q^{-\lambda} \varphi$, we may assume that $\lambda = 0$. The assertion then follows from the fact that under these assumptions the solution matrix $Y$ at the generic point satisfies $Y \equiv 1 \pmod{(X-t)^m \mathcal{O}_E[X-t]}$. Here $n$ is the least integer such that $|m^n| \leq q^{-\mu}$.

### 3.2 Equivalence of categories

Let $R$ be either $E$ or $K[[x]]_0$ and let $\sigma_1$ and $\sigma_2$ be Frobenius maps on $R$ as in the previous subsection. We define a functor

$$\vartheta^*_{\sigma_1, \sigma_2} : (\varphi$-$\nabla$-modules over $(R, \sigma_2)) \to (\varphi$-$\nabla$-modules over $(R, \sigma_1))$$

by $(M, \nabla, \varphi) \mapsto (M, \nabla, \varphi \circ \vartheta_{\sigma_1, \sigma_2})$. Here $\vartheta_{\sigma_1, \sigma_2}$ is defined as in the previous section. The propositions of the previous subsection then give

**Theorem 3.3** $\vartheta^*_{\sigma_1, \sigma_2}$ is an equivalence of categories which preserves tensor products and duals. Moreover, $\vartheta^*_{\sigma_1, \sigma_2}$ preserves the Frobenius slope filtration and the log-growth filtration of $M$ (resp. $V(M)$) for a $\varphi$-$\nabla$-module $M$ over $E$ (resp. $K[[x]]_0$).
4 Boundedness and splitting of the Frobenius slope filtration

4.1 Splitting theorem

**Theorem 4.1** Suppose that the residue field $k$ of $V$ is perfect. A $\varphi\nabla$-module $M$ over $E$ is bounded if and only if $M$ is a direct sum of pure $\varphi\nabla$-modules, that is,

$$M \cong \bigoplus_{i=1}^{r} S_{\lambda_i}(M)/S_{\lambda_i-1}(M)$$

as $\varphi\nabla$-modules, where $\lambda_1 < \lambda_2 < \cdots < \lambda_r$ are Frobenius slopes of $M$.

Since any pure $\varphi\nabla$-module over $E$ is bounded by [CT09, Corollary 6.5]. Hence, Theorem 4.1 above follows from the next proposition.

**Proposition 4.2** Suppose that the residue field $k$ of $V$ is perfect. Let $0 \to L \to M \to N \to 0$ be an exact sequence of $\varphi\nabla$-modules over $E$ such that both $L$ and $N$ are pure of Frobenius slope $\lambda$ and $\nu$, respectively. If one of the conditions

1. $\nu - \lambda < 0$;
2. $\nu - \lambda > 1$;
3. $M$ is bounded and $0 < \nu - \lambda \leq 1$,

holds, then the exact sequence is split, that is, $M \cong L \oplus N$ as $\varphi\nabla$-modules.

In the case (1) the assertion easily follows from the fact that, for $a \in E$ with $|a|_0 < 1$, $a\sigma$ is a contractive operator on the $p$-adic complete field $E$. The rest of this section will be dedicated to proving the assertion in cases (2) and (3).

4.2 Descent of splittings

**Proposition 4.3** Let $0 \to L \to M \to N \to 0$ be an exact sequence of $\varphi$-modules over $E$ such that $L$ and $N$ are pure and the two slopes are different. Let $E'$ be one of the following:

1. $E'$ is a $p$-adic completion of an unramified extension of $E$;
2. $E'$ is the $p$-adic completion of $E \otimes_K K'$ for some extension $K'$ of $K$ as a complete discrete valuation field with an extension $\sigma'$ of $\sigma$ such that, if $G$ is the group of continuous automorphisms of $K'$ over $K$, then the invariant subfield of $K'$ by the action of $G$ is $K$.

If the exact sequence is split over $E'$, then it is split over $E$. The same holds for $\varphi\nabla$-modules over $E$. 

Documenta Mathematica 16 (2011) 33–69
In each case we may assume that \( E \) is the invariant subfield of \( E' \) by the action of continuous automorphism group \( G \). Let \( e_1, \ldots, e_r, e_{r+1}, \ldots, e_{r+s} \) be a basis of \( M \) over \( E \) such that \( e_1, \ldots, e_r \) is a basis of \( L \). Put

\[
\varphi(e_1, \ldots, e_r, e_{r+1}, \ldots, e_{r+s}) = (e_1, \ldots, e_r, e_{r+1}, \ldots, e_{r+s}) \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix},
\]

where \( A_{11} \) is of degree \( r \) and \( A_{22} \) is of degree \( s \), respectively, and all entries of \( A_{11}, A_{12} \) and \( A_{22} \) are contained in \( E \). By the hypothesis of splitting over \( E' \) there exists a matrix \( Y \) with entries in \( E' \) such that

\[
A_{11}\sigma(Y) - YA_{22} + A_{12} = 0.
\]

For any \( \rho \in G \), \( \rho(Y) \) also gives a splitting. Hence \( A_{11}\sigma(Y - \rho(Y)) = (Y - \rho(Y))A_{22} \). By the assumption on slopes, \( \rho(Y) = Y \). Therefore, all entries of \( Y \) are contained in \( E \) and the exact sequence is split over \( E \).

**Definition 4.4** An extension \( E' \) (resp. \( K' \)) of \( E \) (resp. \( K \)) is allowable if \( E' \) is a finitely successive extension of \( E \) (resp. \( K \)) of type in (i) or (ii) (resp. (ii)) of Proposition 4.3.

### 4.3 Preparations

In this subsection we assume that the residue field \( k \) of \( V \) is algebraically closed. Moreover we assume that the Frobenius on \( E \) (resp. \( K[x_0] \)) is defined by \( \sigma(x) = x^q \). For an element \( a = \sum a_n x^n \in E \) (resp. \( K[x] \)) we define the subseries \( a^{(q)} \) by \( \sum a_{qn} x^{qn} \).

**Lemma 4.5** Let \( \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} \) be an invertible matrix of degree \( r + s \) over \( E \) (resp. \( K[x_0] \)) with \( A_{11} \) of degree \( r \) and \( A_{22} \) of degree \( s \) such that the matrix satisfies the conditions:

(i) \( A_{11} = A_{11}^{(q)} \) and \( A_{11} = P^{-1} \) for a matrix \( P \) over \( E \) (resp. \( K[x_0] \)) with \( |P|_0 < 1 \),

(ii) \( A_{22} = A_{22}^{(q)} \) and \( |A_{22} - 1_s|_0 < 1 \).

Suppose that \( A_{12}^{(q)} \neq 0 \). Then there exists an \( r \times s \) matrix \( Y \) over \( E \) (resp. \( K[x_0] \)) with \( |Y|_0 < |A_{12}^{(q)}|_0 \) such that, if one puts \( B = A_{11}\sigma(Y) - YA_{22} + A_{12} \), then \( |B^{(q)}|_0 < |A_{12}^{(q)}|_0 \). Moreover, there exists an \( r \times s \) matrix \( Y \) over \( E \) (resp. \( K[x_0] \)) such that if one defines \( B_{12} \) by

\[
\begin{pmatrix} A_{11} & B_{12} \\ 0 & A_{22} \end{pmatrix} = \begin{pmatrix} 1_r & -Y \\ 0 & 1_s \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} \begin{pmatrix} 1_r & \sigma(Y) \\ 0 & 1_s \end{pmatrix},
\]

then \( B_{12}^{(q)} = 0 \).
Proof. Take a matrix $Y$ such that $\sigma(Y) = -PA^{(q)}_{12}$. Such a $Y$ exists since the residue field $k$ of $V$ is perfect. Then $|Y|_0 < |A^{(q)}_{12}|_0$ and $B = A_{11}\sigma(Y) - YA_{22} + A_{12} = A_{11}PA^{(q)}_{12} - YA_{22} + A_{12} = A_{12} - A^{(q)}_{12} - YA_{22}$. Hence $|B^{(q)}|_0 = |YA^{(q)}_{22}|_0 < |A^{(q)}_{12}|_0$ and we have the first assertion. Applying the first assertion inductively on the value $|A^{(q)}_{12}|_0$, we have a desired matrix $Y$ of the second assertion since $E$ (resp. $K[x]_0$) is complete under the norm $| \cdot |_0$. \hfill $\square$

We give a corollary of the preceding lemma for $\wp$-$\nabla$-modules over $E$.

Proposition 4.6 Let $0 \to L \to M \to N \to 0$ be an exact sequence of $\wp$-$\nabla$-modules over $E$. Suppose that $N$ is pure of Frobenius slope $\nu$ and all Frobenius slopes of $L$ are less than $\nu$. Then there exist an allowable extension $E'$ of $E$ and a basis $e_1, \ldots, e_r, e_{r+1}, \ldots, e_{r+s}$ of $M \otimes_E E'$ with respect to the exact sequence such that, if one fixes an element $x'$ in the ring $O_{E'}$ of integers of $E'$ whose image gives a uniformizer of the residue field of $O_{E'}$ and a Frobenius $\sigma'$ on $E'$ with $\sigma'(x') = x'^q$, then the Frobenius matrix \((A_{11} \quad A_{12} \quad 0 \quad A_{22})\) of $M \otimes_E E'$ with respect to $\sigma'$ (here we use Theorem 3.3) has the following form:

\begin{enumerate}
  \item[(i)] $A_{11} = A^{(q)}_{11}$ and $A_{11} = P^{-1}$ for a matrix $|P|_0 < q^\nu$,
  \item[(ii)] $A_{22} = A^{(q)}_{22}$ and $|A_{22} - q^\nu 1_s|_0 < q^{-\nu}$,
  \item[(iii)] $A^{(q)}_{12} = 0$,
\end{enumerate}

where $A^{(q)}$ is defined by using the parameter $x'$. Moreover, one can replaces the inequality $|A_{22} - q^\nu 1_s|_0 < q^{-\nu}$ in (ii) by the inequality $|A_{22} - q^\nu 1_s|_0 < q^{-\nu} \eta$ for a given $0 < \eta < 1$ (the extension $E'$ depends on $\eta$).

Proof. Since $k$ is algebraically closed, there is a uniformizer $\pi$ of $K$ such that $\sigma(\pi) = \pi$. Let $K_m$ be a Galois extension $K(\pi^{1/m}, \zeta_m)$ of $K$ for a positive integer $m$, where $\zeta_m$ denotes a primitive $m$-th root of unity. Then $\sigma$ on $K_m$ extends on $K_m$. If we choose a positive integer $m$ such that $m/\log_q|\pi|$ is a common multiple of denominators of $\nu$ and the highest Frobenius slope of $L$, then $\nu$ and the highest Frobenius slope of $L$ are contained in $\log_q|K_m^0|$. Hence we may assume that $\nu = 0$ and all Frobenius slopes of the twist $\pi_{\varphi_L}$ of the Frobenius $\varphi_L$ of $L$ are less than or equal to $0$.

Let $A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$ be a Frobenius matrix of $M$ with respect to the given exact sequence. Since any $\wp$-module over $E$ has a cyclic vector [Ts96, Proposition 3.2.1], we may assume that $A_{22} \in \text{GL}_s(O_E)$ by $\nu = 0$. Then there is a matrix $X \in \text{GL}_s(O_E)$ such that $X^{-1}A_{22}\sigma(X) \equiv 1_s \pmod{mO_E}$ for some finite unramified extension $E'$ over $E$ by [Ts98b, Lemma 5.2.2]. By applying the existence of a cyclic vector again, we may assume that the all entries of Frobenius matrix of $L'$ are contained in $mO_E$ by the hypothesis on Frobenius slopes of $L$. 

Documenta Mathematica 16 (2011) 33–69
Now we fix a parameter $x'$ of $E'$ and change a Frobenius $\sigma'$ on $E'$ such that $\sigma'(x') = x'^q$. The hypothesis of the matrices $A_{11}$ and $A_{12}$ are stable by Theorem 3.3. If one replaces the basis $(e_1, \cdots, e_{r+s})$ by $(e_1, \cdots, e_{r+s})A$, then the Frobenius matrix becomes $\sigma'(A)$. Since the hypothesis in Lemma 4.5 hold in our Frobenius matrix $A$, we have the assertion. \hfill \Box

Now a variant of Proposition 4.6 for $\varphi\nabla$-modules over $K[x]_0$, which we use it in section 6, is given.

**Proposition 4.7** Let $0 \rightarrow L \rightarrow M \rightarrow N \rightarrow 0$ be an exact sequence of $\varphi\nabla$-modules over $K[x]_0$. Suppose that $N_q$ is pure of Frobenius slope $\nu$ and all Frobenius slopes of $L_q$ are less than $\nu$. Then there exist an allowable extension $K'$ of $K$ with an extension of Frobenius $\sigma'$ and a basis $e_1, \cdots, e_r, e_{r+1}, \cdots, e_{r+s}$ of $M \otimes_{K[x]_0} K'[x]_0$ with respect to the exact sequence such that the Frobenius matrix \[
\begin{pmatrix}
A_{11} & A_{12} \\
0 & A_{22}
\end{pmatrix}
\] of $M \otimes_{K[x]_0} K'[x]_0$ with respect to $\sigma'$ has the following form:

(i) $A_{11} = A_{11}^{(q)}$ and $A_{11} = P^{-1}$ for a matrix $|P|_0 < q^\nu$,

(ii) $A_{22} = q^\nu 1_\nu$,

(iii) $A_{12}^{(q)} = 0$

**Proof.** We may assume $\mu = 0$ and the highest Frobenius slope of $L_q$ is contained in $\log_q[K'_m^n]$ as in the proof of Proposition 4.6. Then $N$ is a direct sum of copies of the unit object $(K[x]_0, d, \sigma')$’s since $k$ is algebraically closed. In order to find the matrix $P$, we apply the isogeny theorem [Ka79, Theorem 2.6.1] and the existence of a free lattice over $\mathcal{V}[x]$ in [dJ98, Lemma 6.1] for $L'$.

The rest is again same as the proof of Proposition 4.6. \hfill \Box

**Lemma 4.8** Let $\nu$ be a nonnegative rational number. Suppose that $y \in xK[[x]]$ satisfies a Frobenius equation

$$y - q^{-\nu}a \varphi(y) = f.$$ 

for $a \in K$ with $|a| = 1$ and for $f = \sum_n f_n x^n \in xK[[x]]$.

1. Suppose that $f^{(q)} = 0$. If $f \in K[x]_\nu \setminus \{0\}$, then $y \in K[[x]]_\nu \setminus K[[x]]_{\nu-}$, and if $f \in K[x]_\lambda \setminus K[x]_{\lambda-}$ for $\lambda > \nu$, then $y \in K[[x]]_\lambda \setminus K[[x]]_{\lambda-}$.

2. Let $l$ be a nonnegative integer with $q \nmid l$. If $f \in K[x]_0$ and $|f| > |q^\nu f|_0 = q^{-\nu}|f|_0 \neq 0$, then $y \in K[[x]]_\nu \setminus K[[x]]_{\nu-}$.

**Proof.** Since the residue field $k$ of $\mathcal{V}$ is algebraically closed, we may assume that $a = 1$. Formally in $K[\varphi]$, 

$$y = \sum_{n} \sum_{m=0}^{\infty} (q^{-\nu})^n \sigma^m(f_n) x^{q^m n}$$
is a solution of the equation.

(1) If \( q^m n = q^m n' \), then \( m = m' \) and \( n = n' \) because \( q \not\mid n, n' \). Hence, \( y \neq 0 \). By considering a subseries \( \sum_{m=0}^{\infty} (q^{-\nu})^n \sigma^m(f_n)x^m n \) for \( f_n \neq 0 \), \( y \) is of log-growth equal to or greater than \( \nu \). Moreover, we have

\[
| (q^{-\nu})^n \sigma^m(f_n) | / (q^m n + 1) = | f_n | / (n + 1/q^m)^\nu
\]

Hence, if \( f \in K[x]_\nu \), then \( y \) is exactly of log-growth \( \nu \). Suppose \( f \in K[x]_\lambda \setminus K[x]_{\lambda-} \). Since for each \( m, n \)

\[
| (q^{-\nu})^n \sigma^m(f_n) | / (q^m n + 1)^\lambda = | f_n | / (q^m(1-\nu/\lambda) n + 1/q^{nu/\lambda}) ^\lambda,
\]

the log-growth of \( y \) is exactly \( \lambda \).

(2) There exists \( z \in xK[x]_0 \) with \( |z|_0 \leq |q^\nu f|_0 = q^{-\nu} |f|_0 \) such that, if \( g = f - z + q^{-\nu} \sigma(z) = \sum_n g_n x^n \), then \( g^{(q)} = 0 \) and \( g_l \neq 0 \) by the same construction of the proof of Lemma 4.5. The assertion now follows from (1). \( \square \)

### 4.4 Proof of Proposition 4.2

Replacing \( K \) by an extension, we may assume that \( k \) is algebraically closed and that \( \lambda = 0, \nu > 0 \) and \( \nu \in \log_q |K^\times| \) by Proposition 4.3 (see the beginning of proof of Proposition 4.6). We may also assume \( \sigma(x) = x^q \) by Theorem 3.3.

Let \( A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} \) be a Frobenius matrix of \( M \) with respect to the basis which is compatible with the given extension (i.e., the (1, 1)-part (resp. (2, 2)-part) corresponds to \( L \) (resp. \( N \)) and let \( G = \begin{pmatrix} G_{11} & G_{12} \\ 0 & G_{22} \end{pmatrix} \) be the matrix of the connection, respectively. The commutativity of Frobenius and connection (the relation (FC) in section 2.2) gives the relation

\[
1^o \quad \frac{d}{dx} A_{12} + G_{11} A_{12} + G_{12} A_{22} = qx^{q^{-1}}(A_{11} \sigma(G_{12}) + A_{12} \sigma(G_{22}))
\]

of the (1, 2)-part of the matrix. We may assume that

\[
2^o \quad A_{11} = A_{11}^{(q)}, |A_{11} - 1_r|_0 \leq q^{-r} \quad (r \text{ is rank of } L);
\]

\[
3^o \quad A_{22} = A_{22}^{(q)}, |A_{22} - q^{-s} 1_s|_0 \leq q^{-s} \quad (s \text{ is rank of } N);
\]

\[
4^o \quad A_{12}^{(q)} = 0.
\]

by Proposition 4.6 Note that both inequalities \( |G_{11}|_0 < q^{-r} \) and \( |G_{11}|_0 < q^{-r} \) above follow from the relation (FC) in section 2.2 for \( L \) and \( N \), respectively. When \( \nu \neq 1 \), we will first prove \( A_{12} = 0 \) and then prove \( G_{12} = 0 \). When \( \nu = 1 \), we will first prove \( G_{12} = 0 \) and then prove \( A_{12} = 0 \). Hence, we will have a splitting in all cases.
4.4.1 The case where $\nu > 1$

Suppose $\nu > 1$ (and $\lambda = 0$). Assume that $A_{12} \neq 0$. By $4^5$ we have $|\frac{d}{dx}A_{12}|_0 > |qA_{12}|_0 = q^{-1}|A_{12}|_0$. Then $|G_{11}A_{12}|_0 < q^{-1}|A_{12}|_0 < |\frac{d}{dx}A_{12}|_0$ and $|q^{-1}A_{12}(G^2)|_0 < q^{-1}|A_{12}|_0 < |\frac{d}{dx}A_{12}|_0$. On the other hand, $|G_{12}A_{22}|_0 < |q^{-1}A_{12}|_0|G_{12}|_0$ by $\nu > 1$ since $A_{11}$ (resp. $A_{22}$) is a unit matrix (resp. a unit matrix times $q^\nu$) modulo $\mathfrak{m}_G$ (resp. $q^\nu \mathfrak{m}_G$) by $2^5$ (resp. $3^5$). So we have

$$\frac{d}{dx}A_{12} \equiv q^{-1}A_{12}(G_{12}) \pmod{q^{1-\log_2}|\frac{d}{dx}A_{12}|_0\mathfrak{m}_G}$$

But, on comparing the $x$-adic order of both sides, this is seen to be impossible by $2^5$, $3^5$ and $4^5$. Hence $A_{12} = 0$. Now the commutativity of Frobenius and connection (the relation $1^5$) is just

$$G_{12}A_{22} = q^{-1}A_{12}(G_{12}).$$

Since any morphism between pure $\varphi$-modules with different Frobenius slopes are $0$, we have $G_{12} = 0$ by $\nu > 1$.

4.4.2 The case where $0 < \nu < 1$

Suppose $0 < \nu < 1$ (and $\lambda = 0$). Assuming that $A_{12} \neq 0$, we will show the existence of unbounded solutions on the generic disk by applying Lemma 4.8 (2). This is a contradiction to our hypothesis of boundedness of $M$, and thus we must have $A_{12} = 0$. Since $\nu \neq 1$, we again have $G_{12} = 0$ by the slope reason. Therefore, the extension is split. Assume that $A_{12} = \sum_{n=0}^\infty A_{12,n}x^n \neq 0$. Since $|G_{12}A_{22}|_0 = q^{-\nu}|G_{12}|_0$, $|q^{-1}A_{12}|_0 = q^{-1}|G_{12}|_0$, and $|\frac{d}{dx}A_{12}|_0 < q^{-1}|A_{12}|_0$ by $3^5$, $2^5$ and our hypothesis, respectively, the formula $4^5$ gives estimates

$$5^5 \quad q^{-1}|A_{12}|_0 < q^{-\nu}|G_{12}|_0 = |G_{12}A_{22}|_0 = |\frac{d}{dx}A_{12}|_0 \leq |A_{12}|_0.$$

We also claim that

$$6^5 \quad \text{there is a positive integer } m \text{ with } q \nmid m \text{ such that } |\frac{1}{m}\frac{d^m}{dx^m}A_{12}|_0 = |A_{12}|_0$$

by $1^5$. Indeed, let $l$ be an integer such that $|A_{12,l}| = |A_{12}|_0$. When $l > 0$, we put $m = l$. Then the coefficient of $\frac{1}{m}\frac{d^m}{dx^m}A_{12}$ in the $0$-th term $x^0$ is $A_{12,l}$ and we have $|\frac{1}{m}\frac{d^m}{dx^m}A_{12}|_0 = |A_{12,l}| = |A_{12}|_0$. When $l < 0$, we put $m = q^{-l} + l$ (remark that any sufficient large power of $q$ can be replaced by $q^{-l}$). Then the coefficient of $\frac{1}{m}\frac{d^m}{dx^m}A_{12}$ in the $l-m(-q^{-l})$-th term $x^{l-m}$ is $(-1)^m \binom{m-l-1}{m} A_{12,l}$ and we have $|\frac{1}{m}\frac{d^m}{dx^m}A_{12}|_0 = |A_{12,l}| = |A_{12}|_0$ since $(-1)^m \binom{m-l-1}{m}$ is a $p$-adic unit.

In proving the assertion, we will consider the following two cases for $A_{12}$:
(i) \(|\frac{d}{dt}A_{12}|_0 > q^{-\nu}|A_{12}|_0\).

(ii) \(|\frac{d}{dt}A_{12}|_0 \leq q^{-\nu}|A_{12}|_0\). (Hence we have \(|G_{12}|_0 \leq |A_{12}|_0\) by 5°)

In order to prove the existence of unbounded solutions above, let us reorganize the matrix representation by using changes of basis of \(M\), a change of Frobenius and an extension of scalar field. Let us consider the induced \(\varphi\)-\(N\)-module \(M_{\varphi} = M \otimes_{E} E_t[X - t][0]\) over the bounded functions \(E_t[X - t][0]\) at the generic disk. Since \(L_{\tau}\) and \(N_{\tau}\) are pure, we have bounded solution matrices \(Y_{11}\) of \(L\) and \(Y_{22}\) of \(N\), that is,

\[
L : \begin{cases} 
A_{11}(t)\sigma(Y_{11}) = Y_{11}\tau(A_{11}) \\
\frac{d}{dt}Y_{11} = Y_{11}\tau(G_{11}) \\
Y_{11} \in \mathfrak{t}_r + q(X - t)\text{Mat}_r(O_{E_t}[X - t]) \\
A_{22}(t)\sigma(Y_{22}) = Y_{22}\tau(A_{22}) \\
\frac{d}{dt}Y_{22} = Y_{22}\tau(G_{22}) \\
Y_{22} \in \mathfrak{t}_s + q(X - t)\text{Mat}_s(O_{E_t}[X - t])
\end{cases}
\]

by 2° and 3°. Note that \(\tau(f) = \sum_n \frac{n!}{n!}(\frac{1}{n!}f)(t)(X - t)^n\) for \(f \in \mathcal{E}\) and it is an isometry. Consider a change of basis of \(M_{\tau}\) by the matrix \(Y^{-1} = \begin{pmatrix} Y_{11}^{-1} & 0 \\
0 & Y_{22}^{-1} \end{pmatrix}\). Then the new Frobenius matrix and the new connection matrix are as follows:

\[
A^\tau = YA\sigma(Y)^{-1} = \begin{pmatrix} A_{11}(t) & Y_{11}\tau(A_{12})\sigma(Y_{22})^{-1} \\
0 & A_{22}(t) \end{pmatrix} \\
G^\tau = Y\frac{d}{dt}Y^{-1} + YGY^{-1} = \begin{pmatrix} 0 & Y_{11}\tau(G_{12})Y_{22}^{-1} \\
0 & 0 \end{pmatrix}
\]

Let us put \(A_{12} = \sum_n A_{12,n}(X - t)^n\) (resp. \(G_{12}\)) to be the \((1, 2)\)-part of the Frobenius matrix \(A^\tau\) (resp. \(G^\tau\)), and define \(B^\tau_{12} = \sum_{n>0} A_{12,n}(X - t)^n\) by the subseries of positive powers. Then we have

\[
8° \quad |B^\tau_{12}|_0 = |A_{12}|_0 \\
9° \quad |G^\tau_{12}|_0 = |\tau(G_{12})|_0 = |G_{12}|_0.
\]

by 6° and \(Y \equiv 1_{r+s} (\text{mod } q(X - t)O_{E_t}[X - t])\).

Now we consider a change of Frobenius. At first our Frobenius on \(E\) is given by \(\sigma(x) = x^q\). Hence the induced Frobenius on the generic disk is given by \(\sigma(X - t) = ((X - t) + t)^q - t^q\). Let us replace \(\sigma\) by the Frobenius \(\tilde{\sigma}\) defined by \(\tilde{\sigma}(X - t) = (X - t)^q\). Note that

\[
10° \quad \sigma(X - t) - \tilde{\sigma}(X - t) \equiv qt^{q-1}(X - t) \quad (\text{mod } p(X - t)^2O_{E_t}[X - t]).
\]

Since \(\left|\frac{d}{dx}A_{12,x}\right|_0 \leq |n|^{-1}|G_{12}|_0\) and \(|p^n/n| \leq |p|\) for all \(n \geq 1\), the matrix \(H\) of comparison transform \(\partial^\sigma_{\tilde{\sigma}}(M_{\tau})\) in section 3.1 satisfies the congruence relation.
Indeed, in the case (i) for $\tilde{A}$ by the definition of the equivalence (Theorem 3.3), where $\tilde{\sigma}$ is the $12$-part of $\sigma$, can take $m^5$ maximally unramified extension of $E$. By Proposition 2.1 we may replace $\tilde{A}$ by $\tilde{A}$ and let $M_\tau$ with respect to the Frobenius $\tilde{\sigma}$ is

$$\tilde{A} = A^*H = \left( \begin{array}{cc} A_{11}(t) & A_{12} + A_{11}(t)H_{12} \\ 0 & A_{22}(t) \end{array} \right)$$

by the definition of the equivalence (Theorem 3.3), where $H_{12} = \sum_n H_{12,n}(X - t)^n$ is the $(1, 2)$-part of $H$. If we put $\tilde{A}_{12} = \sum_n \tilde{A}_{12,n}(X - t)^n$ to be the $(1, 2)$-part of $\tilde{A}$ and put $B_{12} = \sum_n \tilde{A}_{12,n}(X - t)^n$, then

$$\tilde{A}_{12,m} > q^{-\nu} |B_{12}|_0.$$ 

Indeed, in the case (i) for $A_{12}$, since $\tilde{A}_{12,1} = A_{12,1}^* + A_{11}(t)H_{12,1}$ and $|H_{12,1}|_0 \leq q^{-1}|G|_0$, we have $|\tilde{A}_{12,1}|_0 = |A_{12,1}|_0 = |A_{12,1}^*|_0$ by 5° and 11°. On the other hand, $|B_{12}|_0 \leq \max\{|B_{12}|_0, |H_{12}|_0\} \leq \max\{|A_{12}|_0, |p||G|_0\} < q^{-\nu}|A_{12}|_0$ by 5°, 8° and 11° because of our hypothesis (i), $|A_{12}|_0 > q^{-\nu}|A_{12}|_0$. Hence we can take $m = 1$. In the case (ii), we take a positive integer $m$ such as 6°. Since $|G_{12}|_0 \leq |A_{12}|_0$ by the hypothesis (ii), we have $|\tilde{B}_{12}|_0 \leq \max\{|B_{12}|_0, |H_{12}|_0\} = |A_{12}|_0$ by 8° and 11°.

By Proposition 2.1 we may replace $E_t$ by the $p$-adic completion $\tilde{E}_t^{\text{ur}}$ of the maximally unramified extension of $E_t$. Then we may assume that $\tilde{A}_{11} = 1$, and $\tilde{A}_{22} = q^e 1$, since the solutions of both $(1, 1)$-part and $(2, 2)$-part is 1 modulo $q$ by 2° and 3°. The solution matrix of $M_\tau \otimes_{E_t} \tilde{E}_t^{\text{ur}}$ has a form $Z = \left( \begin{array}{cc} 1_r & Z_{12} \\ 0 & 1_s \end{array} \right)$ satisfying $\tilde{A}|_{X = t}\tilde{\sigma}(Z) = Z\tilde{A}$ and $Z_{12}|_{X = t} = 0$. In particular, $Z_{12}$ satisfies the relation

$$\tilde{\sigma}(Z_{12}) = q^e Z_{12} + \tilde{B}_{12}.$$ 

On applying Lemma 4.8 (2) to $Z_{12}$, one sees that one of entries of $Z_{12}$ must be exactly of log-growth $\nu$ by 12°. Hence the non-vanishing of $A_{12}$ implies that $M$ is unbounded.

This completes the proof for the case $0 < \nu < 1$.

4.4.3 The case where $\nu = 1$

Suppose that $\nu = 1$. Suppose that $G_{12} \neq 0$. Let us develop $G_{12} = \sum_n G_{12,n}x^n$ and let $m$ be the least integer such that $|G_{12,m}| = |G_{12}|_0$. If $A_{12} \neq 0$, we have $|\tilde{A}_{12}|_0 > q^{-1}|A_{12}|_0$ by 4°. So the relation $1^2$ induces a congruence.
\[ 13^\circ \quad \frac{d}{dx} A_{12} + qG_{12} \equiv qx^{q-1}\sigma(G_{12}) \pmod{q^{1+\log_q|G_{12}|\mathcal{O}_E}} \]

by 2\(^c\) and 3\(^c\). This congruence 13\(^c\) also holds when \(A_{12} = 0\).

Suppose that \(m < -1\). The least power of \(x\) which should appear in the right hand side of the congruence 13\(^c\) above is \(qm + q - 1\). Since \(qm + q - 1 < m\), this is precluded by 4\(^c\).

Suppose that \(m = -1\). Then

\[ \tau(G_{12}) = \sum_{n=0}^{\infty} \frac{1}{n!} \left( \frac{d^n}{dx^n} G_{12} \right) (t - t)^n \]

by 7\(^c\) and 14\(^c\). The case where \(A_{12} = 0\) is impossible since \(G_{12} \neq 0\). If \(A_{12} \neq 0\), then we have a solution exactly of log-growth 1 on the generic disk by the similar construction in the case \(m = -1\). This contradicts our hypothesis.

Therefore, we have \(G_{12} = 0\) in any case.

Now we prove \(A_{12} = 0\). Suppose that \(A_{12} \neq 0\). Then the relation 1\(^c\) is

\[ \frac{d}{dx} A_{12} + G_{11} A_{12} = qx^{q-1} A_{12} \sigma(G_{22}). \]

This is impossible by 2\(^c\), 3\(^c\) and 4\(^c\). Hence, \(A_{12} = 0\).

This completes the proof of Proposition 4.2.

Remark 4.9 There is another proof of Proposition 4.2: one can reduce Proposition 4.2 to the case where \(q = p\), that is, the Frobenius \(\sigma\) is a \(p\)-Frobenius. Then, in the proof of the case \(0 < \nu < 1\), it is enough to discuss only in the case \(\frac{d}{dx} A_{12} = \mathcal{O}_E\).
5 PBQ $\varphi$-$\nabla$-modules

5.1 Definition of PBQ $\varphi$-$\nabla$-modules

**Definition 5.1** (Definition of “PBQ $\varphi$-$\nabla$-modules)

1. A $\varphi$-$\nabla$-module $M$ over $E$ is said to be pure of bounded quotient (called PBQ for simplicity) if $M/M^0$ is pure as a $\varphi$-module.

2. A $\varphi$-$\nabla$-module $M$ over $K[[x]]_0$ is said to be pure of bounded quotient (called PBQ for simplicity) if the generic fiber $M_\eta$ of $M$ is PBQ as a $\varphi$-$\nabla$-module over $E$.

The notion “PBQ” depends only on the Frobenius slopes of the bounded quotient of the generic fiber of $\varphi$-$\nabla$-modules. As we saw in Theorem 4.1, the bounded quotient of the generic fiber always admits a splitting by Frobenius slopes when it has different slopes.

**Example 5.2**  

1. A bounded $\varphi$-$\nabla$-module $M$ over $E$ is PBQ if and only if $M$ is pure as a $\varphi$-module. In particular, any $\varphi$-$\nabla$-module $M$ over $E$ of rank 1 is PBQ.

2. Any $\varphi$-$\nabla$-module $M$ over $E$ of rank 2 which is not bounded is PBQ [CT09, Theorem 7.1].

3. Let us fix a Frobenius on $\sigma$ with $\sigma(x) = x^q$. Let $M$ be a $\varphi$-$\nabla$-module over $K[[x]]_0$ with a basis $(e_1, e_2, e_3)$ such that the Frobenius matrix $A$ and the connection matrix $G$ are defined by

$$A = \begin{pmatrix} 1 & -q^{1/2}x & -qx \\ 0 & q^{1/2} & 0 \\ 0 & 0 & q \end{pmatrix}, \quad G = \begin{pmatrix} 0 & \sum_{n=0}^{\infty} q^{n/2} x^{n-1} & \sum_{n=0}^{\infty} x^n q^{n-1} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Then $M_\eta$ is not bounded and $M$ is not PBQ. Indeed, the $K[[x]]_0$-submodule $L$ generated by $e_1$ is a $\varphi$-$\nabla$-submodule of $M$ such that the quotient $(M/L)_\eta$ is bounded and $(M/L)_\eta$ is not pure. On the other hand the dual $M^\vee$ of $M$ is PBQ.

**Proposition 5.3** Any quotient of PBQ $\varphi$-$\nabla$-modules over $E$ (resp. $K[[x]]_0$) is PBQ.

**Proof.** Let $M$ be a PBQ $\varphi$-$\nabla$-module over $E$ and let $M'$ be a quotient of $M$. The assertion follows from that the natural morphism $M/M^0 \to M/(M')^0$ is surjective by [CT09, Corollary 3.5].

5.2 Existence of the maximally PBQ $\varphi$-$\nabla$-submodules over $E$

Proposition 5.4 Suppose that the residue field $k$ of $\mathcal{V}$ is perfect. Let $M$ be a $\varphi$-$\nabla$-module over $E$ with highest Frobenius slope $\lambda_{\text{max}}$ and let $N'$ be a $\varphi$-$\nabla$-submodule of $M/S\lambda_{\text{max}}(M)$. Then there is a unique $\varphi$-$\nabla$-submodule of $N$ of $M$ such that $N$ is PBQ and the natural morphism $N/N^0 \to M/S\lambda_{\text{max}}(M)$ gives an isomorphism between $N/N^0$ and $N'$. When $N = M/S\lambda_{\text{max}}(M)$, we call the corresponding $N$ the maximally PBQ submodule of $M$.

Proof. First we prove the uniqueness of $N$. Let $N_1$ and $N_2$ be a PBQ submodule of $M$ such that both natural morphisms $N_1/N_1^0 \to M/S\lambda_{\text{max}}(M) \leftarrow N_2/N_2^0$ give isomorphisms with $N'$. Let $N$ be the image of $N_1 \oplus N_2 \to M$ $(a, b) \mapsto a + b$. Then $N$ is PBQ by Proposition 5.3. Since $N_1/N_1^0 \oplus N_2/N_2^0 \to N/N^0$ is surjective by [CT09, Proposition 3.6], the natural morphism $N/N^0 \to M/S\lambda_{\text{max}}(M)$ gives an isomorphism with $N'$. If $N_1$ (resp. $N_2$) is not $N$, then the quotient $N/N_1$ (resp. $N/N_2$) has a bounded solution at the generic disk whose Frobenius slope is different from $\lambda_{\text{max}}$. This is impossible because $N$ is PBQ. Hence $N = N_1 = N_2$.

Now we prove the existence of $N$. We use the induction on the dimension of $M$. Let $f : M \to M/M^0$ be a natural surjection. Since $M/M^0$ is bounded, $M/S\lambda_{\text{max}}(M)$ is a direct summand of $M/M^0$ by the maximality of slopes by Theorem 4.1. Put $L = f^{-1}(N')$. If $L$ is PBQ, then one can put $N = L$. If $L$ is not PBQ, then $L$ is a proper submodule of $M$ and there is a PBQ submodule $L'$ of $L$ such that $L'/(L')^0 \cong L/S\lambda_{\text{max}}(L) = N'$ by the induction hypothesis.

\[ \square \]

Corollary 5.5 Suppose that the residue field $k$ of $\mathcal{V}$ is perfect. Let $M$ be a $\varphi$-$\nabla$-module over $E$. Then there is a unique filtration $0 = P_0(M) \subset P_1(M) \subset \cdots \subset P_r(M) = M$ of $\varphi$-$\nabla$-modules over $E$ such that $P_i(M)/P_{i-1}(M)$ is the maximally PBQ submodule of $M/P_{i-1}(M)$ for any $i = 1, \ldots, r$. We call $\{P_i(M)\}$ the PBQ filtration of $M$.

5.3 Existence of the maximally PBQ $\varphi$-$\nabla$-submodules over $K[x]_0$

Theorem 5.6 Suppose that the residue field $k$ of $\mathcal{V}$ is perfect. Let $M$ be a $\varphi$-$\nabla$-module over $K[x]_0$. Then there is a unique $\varphi$-$\nabla$-submodule $N$ of $M$ over $K[x]_0$ such that the generic fiber $N_{\eta}$ of $N$ is the maximally PBQ submodule of the generic fiber $M_{\eta}$ of $M$. We call $N$ the maximally PBQ submodule of $M$.

Proof. The proof of uniqueness of the maximally PBQ submodules is same to the proof of Proposition 5.4.

We prove the existence of the maximally PBQ submodules by induction on the rank of $M$. If $M$ is of rank 1, then the assertion is trivial. For general $M$, if $M$ is PBQ, then there is nothing to prove. Suppose that $M$ is not PBQ. Then there is a direct summand $L_{\eta}$ of $M_{\eta}/M_{\eta}^0$ such that $L_{\eta}$ is pure with the Frobenius slope which is less than the highest slope $\lambda_{\text{max}}$ of $M$ by Theorem
4.1. Consider the composite of natural morphisms \( M \to M'_{\eta}/M'_{\eta} \to L_{\eta} \). It is not injective by Lemma 5.7 below. Put \( M' \) to be the kernel. Then \( M' \) is a \( \varphi \)-\( \nabla \)-submodule of \( M \) such that \( M'_{\eta}/S_{\lambda_{\text{max}}}(M'_{\eta}) \cong M_{\eta}/S_{\lambda_{\text{max}}}(M_{\eta}) \). By the induction hypothesis there is a maximally PBQ submodule \( N \) of \( M' \) which becomes the maximally PBQ submodule \( N \) of \( M \).

**Lemma 5.7** Suppose that the residue field \( k \) of \( V \) is perfect. Let \( M \) be a \( \varphi \)-module over \( K[x]_0 \) such that the highest Frobenius slope of the generic fiber \( M_{p} \) of \( M \) is \( \lambda_{\text{max}} \). Suppose that there exists an injective \( K[x]_0 \)-homomorphism \( f: M \to L_{\eta} \) which is \( \varphi \)-equivariant, i.e., \( \varphi \circ f = f \circ \varphi \) for a pure \( \varphi \)-module \( L_{\eta} \) over \( E \). Then the Frobenius slope of \( L_{\eta} \) is \( \lambda_{\text{max}} \).

**Proof.** In [dJ98, Corollary 8.2] A.J. de Jong proved this assertion when \( L_{\eta} \) is a generic fiber of a rank 1 pure \( \varphi \)-\( \nabla \) module \( L \) over \( K[x]_0 \). (Indeed, he proved a stronger assertion.) We give a sketch of the proof which is due to [dJ98, Propositions 5.5, 6.4 and 8.1]. Our \( E \) (resp. \( \mathcal{E}^\dagger \), resp. \( \mathcal{E}^\dagger \), \( \mathcal{E} \) introduced below) corresponds to \( \Gamma \) (resp. \( \Gamma_2 \), resp. \( \Gamma_2 \), \( \Gamma_2, \ldots \)) in [dJ98]. We also remark that \( \mathcal{E}^\dagger \) is the extended bounded Robba ring \( \mathcal{R}^\text{bd} \) in [Ke08, 2.2]. We may assume that the residue field \( k \) of \( V \) is algebraically closed and all slopes of \( M \) are contained in the value group of \( \log(K^*) \). We may also assume that \( \sigma(x) = x^q \) by Theorem 3.3. Let us define \( K \)-algebras

\[
\mathcal{E} = \left\{ \sum_{n \in \mathbb{Q}} a_n x^n \middle| a_n \in K, \sup_n |a_n| < \infty, |a_n| \to -\infty (n \to -\infty), \{n \mid |a_n| \geq \alpha \} \text{ is a well-ordered set with respect to the order} \leq \text{ for any} \alpha \in \mathbb{R} \right\}
\]

\[
\mathcal{E}^\dagger = \left\{ \sum_{n \in \mathbb{Q}} a_n x^n \in \mathcal{E} \middle| |a_n| n^q \to 0 (n \to -\infty) \text{ for some} 0 < q < 1 \right\}.
\]

Both \( \mathcal{E} \) and \( \mathcal{E}^\dagger \) are discrete valuation fields such that both ring of integers have a same residue field

\[
k((\mathcal{E} \cap \mathcal{R}^\text{bd}))(x) = \left\{ \sum_{n \in \mathbb{Q}} a_n x^n \middle| a_n \in k, \{n \mid a_n \neq 0 \} \text{ is a well-ordered set with respect to the order} \leq \right\},
\]

which includes an algebraic closure of \( k((x)) \) [Ke01], and that the \( p \)-adic completion of \( \mathcal{E}^\dagger \) is \( \mathcal{E} \). \( \mathcal{E} \) naturally extends to \( \mathcal{E}^\dagger \) by \( \sigma(\sum_n a_n x^n) = \sum_n \sigma(a_n) x^{qn} \). Put

\[
\mathcal{E}^\dagger = \mathcal{E}^\dagger \cap \mathcal{E}.
\]

Then \( \mathcal{E}^\dagger \) is stable under \( \sigma \) and the \( K \)-derivation \( d/dx \). We also denote by \( \mathcal{O}_{\mathcal{E}^\dagger} \) the ring of integer of \( \mathcal{E}^\dagger \).

By explicit calculations we have the following sublemmas.
Sublemma 5.8 For $0 < \eta < 1$ and for $\sum_{n \in \mathbb{Q}} a_n x^n \in \tilde{E}^\dagger$, let us consider a condition:

$$(*)_{\eta} : \sup_n |a_n| \max\{\eta^n, 1\} \leq 1.$$ 

If $f$ and $g$ in $\tilde{E}^\dagger$ satisfy the condition $(*_{\eta})$, then so are $f + g$ and $fg$. Moreover, if $f = \sum_n a_n x^n$ satisfies the condition $(*_{\eta})$ and $|a_0| = 1$, then so is $f^{-1}$.

Note that, if $\eta < \mu$, then the condition $(*)_{\eta}$ implies the condition $(*)_{\mu}$.

Sublemma 5.9 (1) Let $A = 1 + B$ be a square matrix such that $1$ is the unit matrix and all entries of $B$ contained in $m^n \mathcal{O}_{\tilde{E}}$, for a positive integer $n$. Suppose that all entries of $A$ satisfy the condition $(*)_{\eta}$ in Sublemma 5.8. Then there is a matrix $Y = 1 + Z$ with $A \sigma(Y) = Y$ such that all entries of $Z$ are contained in $m^n \mathcal{O}_{\tilde{E}}$ and satisfy the condition $(*)_{\eta}$.

(2) Let $C$ be a matrix such that all entries are contained in $m^n \mathcal{O}_{\tilde{E}}$, for a nonnegative integer $n$ and satisfy the condition $(*)_{\eta}$. Then there is a matrix $Z$ satisfying $\sigma(Z) - Z = C$ such that all entries of $Z$ are contained in $m^n \mathcal{O}_{\tilde{E}}$ and satisfy the condition $(*)_{\eta}$.

Proof. (1) follows from (2) by considering a congruence equation $A \sigma(Y) \equiv Y$ (mod $m^n \mathcal{O}_{\tilde{E}}$) inductively on $l$.

(2) Since the residue field $k$ of $\mathcal{V}$ is perfect, $\sigma$ is bijective. Put $C = \sum_n C_n x^n = C_- + C_0 + C_+$, where they are subseries of negative powers, a constant term, and subseries of positive powers, respectively. The series $Z_- = \sum_{n < 0} \sum_{i=1}^{\infty} \sigma^{-i}(C_n) x^{n+q^i}$ converges and all entries of $Z_-$ satisfies the condition $(*)_{\eta}$, and the equation $\sigma(Z_-) = Z_- = C$. Since $k$ is algebraically closed, there is a matrix $Z_0$ over $\mathcal{V}$ with $|Z_0| \leq |C_0|$ such that $\sigma(Z_0) - Z_0 = C_0$. The series $Z_+ = -\sum_{n=0}^{\infty} \sigma(C_+)$ converges and satisfies $\sigma(Z_+) = Z_+ = C_+$. Hence, $Z = Z_- + Z_0 + Z_+$ is the desired solution.

If $N^\dagger$ is a $\varphi$-$\nabla$-submodule of $M \otimes_{K[[x]]} \mathcal{E}^\dagger$ over $\mathcal{E}^\dagger$, then there is a $\varphi$-$\nabla$-submodule $N$ of $M$ over $K[[x]]$ with $N \otimes_{K[[x]]} \mathcal{E}^\dagger \cong N$ by [dJ98, Proposition 6.4]. Hence, the induced morphism $M \otimes_{K[[x]]} \mathcal{E}^\dagger \rightarrow L_\eta$ is also injective. Moreover, since $\mathcal{E}^\dagger \otimes_{\mathcal{E}} \mathcal{E} \rightarrow \mathcal{E}$ is injective (the similar proof of [dJ98, Proposition 8.1] works), the induced morphism $M \otimes_{K[[x]]} \mathcal{E}^\dagger \rightarrow L_\eta \otimes_{\mathcal{E}} \mathcal{E}$ is again injective. Let $\lambda_1 < \cdots < \lambda_r (= \lambda_{\max})$ be Frobenius slopes of $M_\eta$. One can prove that there exists an increasing filtration $0 = M_0 \subsetneq M_1 \subsetneq \cdots \subsetneq M_r = M \otimes_{K[[x]]} \mathcal{E}^\dagger$ of $\varphi$-modules over $\tilde{E}^\dagger$ such that $(M_i/M_{i-1}) \otimes_{\mathcal{E}^\dagger} \mathcal{E}$ is pure of slope $\lambda_{r-i+1}$. This existence of filtration of opposite direction corresponds to Proposition 5.5 in [dJ98]. Indeed, since the residue field $k((x^q))$ includes an algebraic closure of $k((x))$, there is a basis of $M \otimes_{K[[x]]} \mathcal{E}^\dagger$ such that the Frobenius matrix of
$M \otimes_{K[x]} \tilde{E}$ has a form

$$
\begin{pmatrix}
q^{\lambda_1}1 & \cdots & 0 \\
0 & \ddots & \ddots \\
q^{\lambda_n}1 & \cdots & q^{\lambda_1}
\end{pmatrix} + (a \text{ square matrix with entries in } m^\infty \mathcal{O}_E)
$$

by Dieudonné-Manin classification of $\varphi$-modules and the density of $\tilde{E}$ in $E$. Here $q^\lambda$ is an element of $K$ with $\log_q|q^\lambda| = -\lambda$, 1 is the unit matrix with a certain size (the first matrix is a diagonal matrix), and $n$ is sufficiently large. One can find a basis of $M \otimes_{K[x]} \tilde{E}$ such that the Frobenius matrix of $M \otimes_{K[x]} \tilde{E}$ is a lower triangle matrix

$$
\begin{pmatrix}
q^{\lambda_1}1 & 0 \\
* & \ddots & q^{\lambda_n}1
\end{pmatrix}
$$

by Sublemmas 5.8 and 5.9. Hence, one has a filtration of opposite direction. Since $M_1$ is pure of slope $\lambda_r = \lambda_{\text{max}}$ and the inclusion $M_1 \subset \eta \otimes_E \tilde{E}$ is $\varphi$-equivariant, the slope of $L_\eta$ must be $\lambda_{\text{max}}$. 

\begin{corollary}
Suppose that the residue field $k$ of $V$ is perfect. Let $M$ be a $\varphi$-$\nabla$-module over $K[x]_0$. Then there is a unique filtration $0 = P_0(M) \subset P_1(M) \subset \cdots \subset P_r(M) = M$ as $\varphi$-$\nabla$-modules over $K[x]_0$ such that $P_i(M)/P_{i-1}(M)$ is the maximally PBQ submodule of $M/P_{i-1}(M)$ for any $i = 1, \ldots, r$. We call \{ $P_i(M)$ \} the PBQ filtration of $M$.
\end{corollary}

\begin{example}
Let $M$ be a $\varphi$-$\nabla$-module over $K[x]_0$ which is introduced in Example 5.2 (3). If $P_i(M)$ is a $\varphi$-$\nabla$-submodule of $M$ over $K[x]_0$ generated by $e_1$ and $e_3$, the sequence $0 = P_0(M) \subset P_1(M) \subset P_2(M) = M$ is the PBQ filtration of $M$.
\end{example}

\section{Log-growth and Frobenius Slope Filtration for HPBQ $\varphi$-$\nabla$-Modules over $K[x]_0$}

\subsection{Log-growth for HPBQ $\varphi$-$\nabla$-modules}

\begin{definition}
(1) A $\varphi$-$\nabla$-module $M$ over $K[x]_0$ is horizontal of bounded quotient (HBQ for simplicity) if there is a quotient $N$ of $M$ as a $\varphi$-$\nabla$-module over $K[x]_0$ such that the canonical surjection induces an isomorphism $M_{\eta}/M_{\eta}^0 \cong N_{\eta}$ at the generic fiber.

(2) A $\varphi$-$\nabla$-module $M$ over $K[x]_0$ is horizontally pure of bounded quotient (HPBQ for simplicity) if $M$ is PBQ and HBQ.
\end{definition}

\begin{example}
(1) A bounded $\varphi$-$\nabla$-module $M$ over $K[x]_0$ is HBQ. A bounded $\varphi$-$\nabla$-module $M$ over $K[x]_0$ is HPBQ if and only if $M_{\eta}$ is pure as a $\varphi$-module.
\end{example}
(2) Let $M$ be a $\varphi$-$\nabla$-module over $K[x]_0$ of rank 2 which arises from the first crystalline cohomology of a projective smooth family $E$ of elliptic curves over $\text{Spec} \ k[x]$. Then $M$ is HBQ if and only if either (i) $E$ is a non-isotrivial family over $\text{Spec} \ k[x]$ and the special fiber $E_s$ of $E$ is ordinary or (ii) $E$ is an isotrivial family over $\text{Spec} \ k[x]$. In the case (i) $M$ is HPBQ, but in the case (ii) $M$ is HPBQ if and only if $E$ is an isotrivial family of supersingular elliptic curves.

(3) Let $M$ be a $\varphi$-$\nabla$-module over $K[x]_0$ which is introduced in Example 5.2 (3). Then $M$ is HBQ but is not HPBQ. The dual $M^\prime$ of $M$ is HPBQ.

**Proposition 6.3** Let $M$ be a $\varphi$-$\nabla$-module over $K[x]_0$. Then $M$ is HBQ if and only if 
\[
\dim_K V(M)/V(M)^0 = \dim_K M_{\eta}/M_{\eta}^0.
\]
Moreover, when $M$ is HBQ, the natural pairing $M \otimes_K \text{Sol}_0(M) \to K[x]_0$ induces an isomorphism 
\[
M_{\eta}/M_{\eta}^0 \cong V(M)/V(M)^0 \otimes_K \mathcal{E}
\]
as $\varphi$-$\nabla$-modules.

**Proof.** Suppose that $M$ is HBQ. Let $N$ be the quotient as in Definition 6.1 (1). Since $N_\eta$ is bounded, we have $V(N)^0 = 0$ by Christol’s transfer theorem (see [CT09, Proposition 4.3]) and $\dim_K V(M)/V(M)^0 \geq \dim_K V(N)/V(N)^0 = \text{rank}_K K[x]_0 \cap N = \dim_K M_{\eta}/M_{\eta}^0$. On the other hand, one knows an inequality $\dim_K V(M)/V(M)^0 \leq \dim_K M_{\eta}/M_{\eta}^0$ by [CT09, Proposition 4.10]. Hence, we have an equality $\dim_K V(M)/V(M)^0 = \dim_K M_{\eta}/M_{\eta}^0$. Now we prove the inverse. The natural pairing $M \otimes_K \text{Sol}_0(M) \to K[x]_0$ induces the surjection $M \to V(M)/V(M)^0 \otimes_K K[x]_0$. If $\dim_K V(M)/V(M)^0 = \dim_K M_{\eta}/M_{\eta}^0$, we have an isomorphism $M_{\eta}/M_{\eta}^0 \cong V(M)/V(M)^0 \otimes_K \mathcal{E}$ since $V(M)/V(M)^0 \otimes_K \mathcal{E}$ is bounded. 

Since any quotient of bounded $\varphi$-$\nabla$-modules over $\mathcal{E}$ is again bounded, the proposition below follows from the chase of commutative diagrams.

**Proposition 6.4** Any quotient of HBQ $\varphi$-$\nabla$-modules over $K[x]_0$ is HBQ. In particular, any quotient of HPBQ modules is HPBQ.

**Proof.** We may assume that the residue field of $\mathcal{V}$ is algebraically closed and $q^{\lambda_{\max}} \in K$. Since $M$ is HBQ, there is a surjection $M \to V(M)/V(M)^0 \otimes_K K[x]_0$ by Proposition 6.3 whose kernel is denoted by $L$. Then $M_{\eta}/M_{\eta}^0 = L_\eta$. If $f : \hat{M} \to \hat{N}$ is the given surjection, $N/f(L)$ is a quotient of $V(M)/V(M)^0 \otimes_K K[x]_0$ and hence a direct sum of copies of $(K[x]_0, q^\lambda \sigma, d)$ for some $\lambda$. Since $f$ gives a surjection from $M_{\eta}^0$ to $N_{\eta}^0$ by [CT09, Proposition 3.6], we have 
\[
\dim_K V(N)/V(N)^0 \geq \text{rank}_K K[x]_0 N/f(L) = N_{\eta}/N_{\eta}^0.
\]
On the other hand, $\dim_K V(N)/V(N)^0 \leq \dim_{\mathcal{E}} N_{\eta}/N_{\eta}^0$ by [CT09, Proposition 4.10]. Hence $\dim_K V(N)/V(N)^0 = \dim_{\mathcal{E}} N_{\eta}/N_{\eta}^0$. The rest follows from Proposition 5.3.

Note that the notion PBQ is determined only by the generic fiber. On the other hand, for “HPBQ”, the bounded quotient is horizontal.

**Theorem 6.5** Let $M$ be a $\varphi$-$\nabla$-module $M$ over $K[[x]]_0$ which is HPBQ. Then the conjecture $\text{LGF}_{K[[x]]_0}$ (see 2.5) holds for $M$.

**Proof.** We have only to prove the conjecture $\text{LGF}_{K[[x]]_0}(2)$ for $M$. Then the property of Frobenius slopes implies the conjecture the conjecture $\text{LGF}_{K[[x]]_0}(1)$ for $M$. We may assume that the residue field of $V$ is algebraically closed and all Frobenius slopes of $V(M)$ are contained in the valued group $\log_q |K^\times|$ by Proposition 2.1. We may also assume that our Frobenius $\sigma$ is defined by $\sigma(x) = x^q$ by Theorem 3.3. Let us denote by $\lambda_{\text{max}}$ the highest Frobenius slope of $M_{\eta}$ (i.e., the highest Frobenius slope of $V(M)$). Let $0 = M_0 \subseteq M_1 \subseteq \ldots \subseteq M_r = M$ be a filtration of $M$ as $\varphi$-$\nabla$-modules over $K[[x]]_0$ such that $M_i/M_{i-1}$ is irreducible (i.e., it has no nontrivial $\varphi$-$\nabla$-submodule over $K[[x]]_0$). We will prove the induction on $r$. If $r = 1$, then $M \cong (K[[x]]_0, q^{\lambda_{\text{max}}}, d)$ and the assertion is trivial.

Now suppose $r > 1$. We may also assume $\dim_K V(M)/V(M)^0 = 1$, hence $M_r/M_{r-1} \cong (K[[x]]_0, q^{\lambda_{\text{max}}}, d)$. Indeed, suppose that $s = \dim_K V(M)/V(M)^0 > 1$. By our assumption, there is a $\varphi$-$\nabla$-submodule $L'$ over $K[[x]]_0$ such that the highest Frobenius slope of $L'$ is $\lambda_{\text{max}}$ with multiplicity 1 (note that $L'$ is $M_{r-1}$). Take the maximally PBQ submodule $L$ of $L'$. Then $L$ is HPBQ such that the highest Frobenius slope is $\lambda_{\text{max}}$ with multiplicity 1. Since both highest Frobenius slopes of $L$ and $M/L$ are $\lambda_{\text{max}}$, the assertion follows from the induction hypothesis by Propositions 2.6 and 6.4.

Since all Frobenius slopes of $(M_{r-1})_0$ are less than $\lambda_{\text{max}}$, one can take a basis $e_1, \ldots, e_s$ of $M$ such that the Frobenius matrix $A = \begin{pmatrix} A_1 & B \\ 0 & q^{\lambda_{\text{max}}} \end{pmatrix}$ ($A_1$ is the Frobenius matrix of $M_{r-1}$) satisfies (i) all entries of $A_1$ are contained in $K[[x]]_0 \cap x^q K[[x]]$ and (ii) all entries of $B$ are contained in $xK[[x]]_0 \cup x^q K[[x]] \cup \{0\}$ by Proposition 4.7. Moreover $B \neq 0$ by Lemma 6.6 below since $M$ is PBQ. Let $G$ be the matrix of connection of $M$. Then the identification

$$\text{Sol}(M) = \left\{ y \in \mathcal{A}_{K}(0, 1^-) \mid \frac{d}{dx} y = yG \right\}$$

is given by $f \mapsto (f(e_1), \ldots, f(e_s))$. The inclusion relation in Theorem 2.3 for the solution space is

$$\text{Sol}_{\lambda}(M) \supset S_{\lambda - \lambda_{\text{max}}} (\text{Sol}(M)).$$

Then it is sufficient to prove the inclusion is equal for all $\lambda$. The $\varphi$-module is a direct sum of 1-dimensional $\varphi$-spaces, on which $\varphi$ acts by $q^{d}\sigma$ for some rational
number $\delta$ such that $\lambda_{\max} - \delta$ is a Frobenius slope of $M$, by our assumption of $K$. Let $f \in \text{Sol}_1(M)$ with $\varphi(f) = q^\delta f$. Then the restriction of $f$ on $M_{r-1}$ gives a $(\varphi, \frac{d}{dx})$-equivariant morphism

$$M_{r-1} \to (\mathcal{A}_K(0,1^-), q^{-\delta} \sigma, d).$$

The kernel $L$ of $f$ is a $\varphi$-$\nabla$-module over $K[x]_0$ and $f$ is a solution of $M/L$ of log-growth $\lambda$.

Suppose that $L \neq 0$. Then the length of $M/L$ is smaller than $M$ and $M/L$ is HPBQ by Proposition 6.4. Considering $f$ as a solution of $M/L$, we have $\delta \leq \lambda - \lambda_{\max}$ by the hypothesis of induction. 

Suppose that $L = 0$. The Frobenius relation $\varphi(f) = q^\delta f$ is equivalent to

$$q^{-\delta} \sigma(f(e_1), \cdots, f(e_s)) = (f(e_1), \cdots, f(e_s))A.$$

By the assumption of $A_1$ we have $f(e_1) \in \mathcal{A}_K(0,1^-) \cap x^q K[x^q]$. Let us focus on the $s$-th entry, then it is

$$q^{-\delta} \sigma(f(e_s)) = q^{\lambda_{\max}} f(e_s) + (f(e_1), \cdots, f(e_{s-1}))(\sigma, d).$$

Since the highest Frobenius slope of $M_{r-1}$ is less than $\lambda_{\max}$, the log-growth of the restriction of $f$ on $M_{r-1}$ is of log-growth less than $\lambda_{\max} + \delta$, and so is $(f(e_1), \cdots, f(e_{s-1}))(\sigma, d)$. Since $f$ is injective, $(f(e_1), \cdots, f(e_{s-1}))B \in \mathcal{A}_K(0,1^-) \setminus x^q K[x^q]$ is not 0. Hence, $f(e_s)$ is exactly of log-growth $\lambda_{\max} + \delta$ by Lemma 4.8 (1). This provides an inequality $\lambda_{\max} + \delta \leq \lambda$, and we have $\delta \leq \lambda - \lambda_{\max}$.

Therefore, $f \in S_{\lambda - \lambda_{\max}}(\text{Sol}(M))$. This completes the proof of Theorem 6.5.

\begin{lemma}
Let $0 \to L \to M \to N \to 0$ be an exact sequence of $\varphi$-$\nabla$-modules over $K[x]_0$. If the exact sequence is split as $\varphi$-modules, then it is split as $\varphi$-$\nabla$-modules.
\end{lemma}

\begin{proof}
Let $A = \begin{pmatrix} A_1 & B \\ 0 & A_2 \end{pmatrix}$ and $G = \begin{pmatrix} G_1 & H \\ 0 & G_2 \end{pmatrix}$ be the matrices of Frobenius and connection, respectively. We should prove that $B = 0$ implies $H = 0$. It is sufficient to prove the assertion above as $\mathcal{A}_K(0,1^-)$-modules with Frobenius and connection. Solving the differential modules $L$ and $N$, we may assume that $A_1$ and $A_2$ are constant regular matrices and $G_1 = G_2 = 0$. Then the horizontality of Frobenius structure means he relation

$$HA_2 = qx^{q-1}A_1 \sigma(H).$$

Then we have $H = 0$ by comparing the $x$-adic order of both sides.
\end{proof}
6.2 Equislope $\varphi$-$\nabla$-modules over $K\llbracket x \rrbracket_0$

**Definition 6.7** A $\varphi$-$\nabla$-module $M$ over $K\llbracket x \rrbracket_0$ is equislope if there is an increasing filtration $\{S_\lambda(M)\}_{\lambda \in \mathbb{R}}$ of $\varphi$-$\nabla$-module over $K\llbracket x \rrbracket_0$ such that $S_\lambda(M) \otimes \mathcal{E}$ gives the Frobenius slope filtration of the generic fiber $M_\eta$ of $M$. We also call $\{S_\lambda(M)\}_{\lambda \in \mathbb{R}}$ the Frobenius slope filtration of $M$.

By [Ka79, 2.6.2] (see [CT09, Theorem 6.21]) we have

**Proposition 6.8** A $\varphi$-$\nabla$-module $M$ over $K\llbracket x \rrbracket_0$ is equislope if and only if both the special polygon and generic polygon of Frobenius slopes of $M$ coincides with each other.

**Corollary 6.9** Any subquotients, direct sums, extensions, tensor products, duals of equislope $\varphi$-$\nabla$-modules over $K\llbracket x \rrbracket_0$ are equislope.

**Proposition 6.10** Let $M$ be an equislope $\varphi$-$\nabla$-module over $K\llbracket x \rrbracket_0$.

1. $M$ is HBQ. In particular, if $M$ is PBQ, then $M$ is HPBQ.

2. If $V(M)/V(M)^0$ is pure as a $\varphi$-module, then $M$ is HPBQ.

**Proof.** (1) We may assume that the residue field of $V$ is algebraically closed and all slopes of $M_\eta$ are contained in the value group $\log_\eta(K^\times)$ of $K^\times$ by Proposition 2.1. Let us take a $\varphi$-$\nabla$-submodule $L$ such that its generic fiber $L_\eta$ is $M_\eta^0$. Such an $L$ exists by Lemma 6.11 below. Since $(M/L)_\eta \cong M_\eta/L_\eta$ is bounded, $M$ is HBQ by definition.

2. The assertion follows from (1) and Proposition 6.3.

**Lemma 6.11** Let $M$ be an equislope $\varphi$-$\nabla$-module over $K\llbracket x \rrbracket_0$. Suppose that the residue field of $V$ is algebraically closed and all slopes of $M_\eta$ are contained in the valued group $\log_\eta(K^\times)$. The map taking generic fibers gives a bijection from the set of $\varphi$-$\nabla$-submodules of $M$ to the set of $\varphi$-$\nabla$-submodules of $M_\eta$.

**Proof.** Since the functor from the category $\varphi$-$\nabla$-module over $K\llbracket x \rrbracket_0$ to the category $\varphi$-$\nabla$-module over $\mathcal{E}$ is fully faithful, it is sufficient to prove the surjectivity [dJ98, Theorem 1.1].

We may assume that $\sigma(x) = x^q$ by Theorem 3.3. We use the induction on the number of Frobenius slopes of $M$ in order to prove the existence of a submodule $N$ over $K\llbracket x \rrbracket_0$ for a given submodule $N_\eta$ over $\mathcal{E}$. Suppose that $M$ is pure of slope $\lambda$. There are a basis $e_1, \ldots, e_r$ of $M$ such that the Frobenius matrix is $q^\lambda 1_r$ since $M$ is bounded. Let $N_\eta$ be a $\varphi$-$\nabla$-submodule of $M_\eta$ over $\mathcal{E}$ which is generated by $(e_1, \ldots, e_r)P$ for $P \in \text{Mat}_{rs}(\mathcal{E})$ with $s = \dim_{\mathbb{C}} N_\eta$. Since $N_\eta$ is a $\varphi$-submodule, there is a $B \in \text{GL}_s(\mathcal{E})$ such that $q^\lambda \sigma(P) = PB$. Since $\text{rank}(P) = s$, there is a regular minor $Q$ of $P$ of degree $s$ such that $q^\lambda \sigma(Q) = QB$. If one puts $R = PQ^{-1} \in \text{Mat}_{rs}(\mathcal{E})$, then $\sigma(R) = R$. Hence, $R \in \text{Mat}_{rs}(K)$. Since $(e_1, \ldots, e_r)R$ is a basis of $N_\eta$ such that $(e_1, \ldots, e_r)R$
are included in $M$, the submodule $N$ is given by the $K[[x]]_0$-submodule of $M$
generated by $(e_1, \ldots, e_r)R$.

Let $\lambda_1$ be the first slopes of $M_n$. By the induction hypothesis there are a $\varphi$-$\nabla$-submodule $N_1$ of $S_{\lambda_1}(M)$ such that the generic fiber $(N_1)_n$ of $N_1$ is $N_n \cap S_{\lambda_1}(M_n)$ and a $\varphi$-$\nabla$-submodule $N_2$ of $M/S_{\lambda_1}(M)$ such that the generic fiber of $N_2$ is $N_n/(S_{\lambda_1}(M_n) \cap N_n) = N_n/(N_1)_n$. Let $N_3$ be the inverse image of $N_2$ by the surjection $M/N_1 \to M/S_{\lambda_1}(M)$. Since the intersection of $N_n/(N_1)_n$ and $S_{\lambda_1}(M_n)/(N_1)_n$ is 0 in $M_n/(N_1)_n$, $(N_3)_n$ is a direct sum of $N_n/(N_1)_n$ and $S_{\lambda_1}(M_n)/(N_1)_n$. By applying the fully faithfulness of the functor from the category of $\varphi$-$\nabla$-modules over $K[[x]]_0$ to the category of $\varphi$-$\nabla$-modules over $\mathcal{E}$ [dJ98, Theorem 1.1], there is a direct summand $N_4$ of $N_3$ as $\varphi$-$\nabla$-module over $K[[x]]_0$ such that the generic fiber of $N_4$ is $N_n/(N_1)_n$. Then the inverse image $N$ of $N_4$ by the surjection $M \to M/N_1$ is our desired one. 

\begin{theorem}
The conjecture $\text{LGF}_{K[[x]]_0}$ (see 2.5) holds for any equislope and PBQ $\varphi$-$\nabla$-module over $K[[x]]_0$.
\end{theorem}

\begin{proof}
The assertion follows from Theorem 6.5 and Proposition 6.10 (1).
\end{proof}

7 Log-growth filtration and Frobenius filtration at the generic point

7.1 The log-growth of PBQ $\varphi$-$\nabla$-modules over $\mathcal{E}$

\begin{theorem}
The conjecture $\text{LGF}_\mathcal{E}$ (see 2.4) holds for any PBQ $\varphi$-$\nabla$-module over $\mathcal{E}$.
\end{theorem}

\begin{proof}
Let $M$ be a PBQ $\varphi$-$\nabla$-module over $\mathcal{E}$ such that $\lambda_{\text{max}}$ is the highest Frobenius slope of $M$, and let us consider a $\varphi$-$\nabla$-module $M_\tau = M \otimes \mathcal{E}_t[X-t]_0$ over the $\mathcal{E}_t$-algebra $\mathcal{E}_t[X-t]_0$ of bounded functions on the generic disk. Then $M_\tau$ is equislope since $\{(S_{\lambda}(M)_\tau)\}$ gives a Frobenius slope filtration of $M_\tau$. Moreover, since $M$ is PBQ, $\text{Sol}_0(M, \mathcal{A}_t, (t,1^-))$ is a pure $\varphi$-module. Hence $V(M_\tau)/V(M_\tau)^{\perp}$ is pure, and $M_\tau$ is HPBQ by Proposition 6.10 (2). Applying Theorem 6.5 to $M_\tau$, we have

$$\dim_{\mathcal{E}} M/M^\lambda = \dim_{\mathcal{E}_t} \text{Sol}_{\lambda}(M, \mathcal{A}_t, (t,1^-)) = \dim_{\mathcal{E}_t} V(M_\tau)/V(M_\tau)^{\perp}$$

$$= \dim_{\mathcal{E}_t} V(M_\tau)^{\perp} - \dim_{\mathcal{E}_t} (S_{\lambda-\lambda_{\text{max}}}(V(M_\tau)^{\perp}))^{\perp}$$

$$= \dim_{\mathcal{E}} M^{\vee} - \dim_{\mathcal{E}} (S_{\lambda-\lambda_{\text{max}}}(M^{\vee}))^{\perp}$$

$$= \dim_{\mathcal{E}} M^{\vee}/(S_{\lambda-\lambda_{\text{max}}}(M^{\vee}))^{\perp}$$

for any $\lambda$. Hence, $M^\lambda = (S_{\lambda-\lambda_{\text{max}}}(M^{\vee}))^{\perp}$ by Theorem 2.3. Therefore, the conjecture $\text{LGF}_\mathcal{E}$ holds for $M$. 

\end{proof}
7.2 Rationality of breaks of log-growth filtrations

Theorem 7.2 Let $M$ be a $\varphi\nabla$-module over $\mathcal{E}$ and let $\lambda$ be a break of log-growth filtration of $M$, i.e., $M^\lambda^+ \supseteq M^{\lambda^+}$. Then $\lambda$ is rational and $M^\lambda = M^{\lambda^+}$. In other words, the conjecture $\text{LGF}_E(1)$ (see 2.4) holds for any $\varphi\nabla$-modules over $\mathcal{E}$.

Proof. We may assume that the residue field $k$ of $V$ is perfect by Proposition 2.1. Suppose that $\lambda_{\text{max}}$ be the maximal Frobenius slope of $M$. If $M$ is PBQ, then $M^{\lambda} = (S_{\lambda - \lambda_{\text{max}}} (M^\lambda))^\perp = S_{\lambda_{\text{max}} - \lambda}(M)$ for any $\lambda$ by Theorem 7.1. Then we have

$$M^{\lambda^+} = \cup_{\mu > \lambda} S_{\lambda_{\text{max}} - \mu}(M) = \cup_{\mu > \lambda} S_{\lambda_{\text{max}} - \mu}(M) = S_{\lambda_{\text{max}} - \lambda}(M) = M^\lambda.$$

If $\lambda$ is a break of log-growth filtration, then

$$S_{\lambda_{\text{max}} - \lambda}(M) = S_{\lambda_{\text{max}} - \lambda}(M) = M^{\lambda^+} \supseteq M^\lambda = S_{\lambda_{\text{max}} - \lambda}(M)$$

and $\lambda$ is also a Frobenius slope filtration. Hence $\lambda$ is rational. For a general $M$, we use the induction on the length of the PBQ filtration of $M$. Let $L$ be the maximally PBQ submodule of $M$ and suppose $N = M/L$. Then we have the assertion by Proposition 2.6 (1), the PBQ case and the induction hypothesis on $L$ and $N$.  

Proposition 7.3 Suppose that the residue field of $V$ is perfect. Let $M$ be a $\varphi\nabla$-module over $K[[x]]_0$ and let $\lambda$ be a break of log-growth filtration of $V(M)$, i.e., $V(M)^\lambda^+ \supseteq V(M)^{\lambda^+}$, and let $\{P_i(M)\}$ be the PBQ filtration of $M$. Suppose that the conjecture $\text{LGF}_{K[[x]]_0}(2)$ (see 2.5) holds for all $P_i(M)/P_{i-1}(M)$. Then $\lambda$ is rational and $V(M)^\lambda = V(M)^{\lambda^+}$. In particular, the conjecture $\text{LGF}_{K[[x]]_0}(2)$ implies the conjecture $\text{LGF}_{K[[x]]_0}(1)$ for any $\varphi\nabla$-modules over $K[[x]]_0$.

Proof. The proof is similar to that of Theorem 7.2 by replacing Proposition 2.6 (1) by Proposition 2.6 (2).

8 Toward Dwork’s conjecture $\text{LGF}_{Dw}$

8.1 The comparison at the special point and Dwork’s conjecture $\text{LGF}_{Dw}$

Theorem 8.1 The conjecture $\text{LGF}_{K[[x]]_0}(2)$ (see 2.5) implies the conjecture $\text{LGF}_{Dw}$ (see 2.7), that is, the special log-growth polygon lies above the generic log-growth polygon (and they have the same endpoints).

The theorem above follows from the proposition below by Proposition 2.1.
Proposition 8.2 Suppose that the residue field $k$ of $V$ is perfect. Let $M$ be a $\varphi\nabla$-module over $K[x]_0$ and let $\{P_i(M)\}$ be the PBQ filtration of $M$. Suppose that the conjecture $\text{LGF}_{K[x]}(2)$ (see 2.5) holds for all $P_i(M)/P_{i-1}(M)$. Then the special log-growth polygon of $M$ lies above the generic log-growth polygon of $M$ (and they have the same endpoints).

Proof. For the PBQ $\varphi\nabla$-modules arising from the PBQ filtration of $M$, the log-growth polygons at the generic (resp. special) fiber coincides with the Newton polygon of Frobenius slopes of the dual at the generic (resp. special) fiber under the suitable shifts of Frobenius actions by Theorem 7.1 (resp. our hypothesis). The assertion follows from Proposition 2.6, Lemma 8.3 below and the fact that the special Newton polygon of Frobenius slopes is above the generic Newton polygon of Frobenius slopes and they have the same endpoints.

Lemma 8.3 Let $0 \to L \to M \to N \to 0$ be an exact sequence of $\varphi\nabla$-modules over $K[x]_0$ such that the induced sequences

\[
0 \to L_\eta/L_\eta^\lambda \to M_\eta/M_\eta^\lambda \to N_\eta/N_\eta^\lambda \to 0 \\
0 \to V(L)/V(L)^\lambda \to V(M)/V(M)^\lambda \to V(N)/V(N)^\lambda \to 0
\]

on both the generic fiber and the special fiber are exact for any $\lambda$.

1. If the special log-growth polygon lies above the generic log-growth polygon (the endpoints might be different) for both $L$ and $N$, then the same holds for $M$.

2. If the special log-growth polygon and the generic log-growth polygon have the same endpoints for both $L$ and $N$, then the same holds for $M$.

3. Suppose that the special log-growth polygon lies above the generic log-growth polygon for both $L$ and $N$. Then both the special and the generic log-growth polygons coincide with each other for $M$ if and only if the same hold for $L$ and $N$.

Proof. Let $r$ be the rank of $M$. Let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_r$ be breaks of log-growth filtration of $M_\eta$ with multiplicities, and put $b_0(M_\eta) = 0$ and

\[b_j(M_\eta) = \lambda_1 + \cdots + \lambda_j\]

for $1 \leq j \leq r$. Then the generic log-growth polygon of $M$ is a polygon which connects points $(0, b_0(M_\eta))$, $(1, b_1(M_\eta))$, $\cdots$, $(r, b_r(M_\eta))$ by lines. We also define $b_j(V(M))$ for the special log-growth of $M$. Then the exactness for any $\lambda$ implies the equality

\[b_j(M_\eta) = \min \left\{ b_i(L_\eta) + b_k(N_\eta) \mid 0 \leq i \leq \text{rank } L, \ 0 \leq k \leq \text{rank } N, \ i + k = j \right\}\]
for all $0 \leq j \leq r$, and the same holds for the special log-growth. The special log-growth polygon lies above the generic log-growth polygon for $M$ if and only if $b_j(M) \leq b_j(V(M))$ for all $j$, the special log-growth polygon and the generic log-growth polygon have the same endpoints for $M$ if and only if $b_r(M) = b_r(V(M))$, and both the special and the generic log-growth polygons coincide with each other for $M$ if and only if $b_j(M) = b_j(V(M))$ for all $j$. Hence we have the assertions.

**Remark 8.4** If $L$ is supposed to be HPBQ in the short exact sequence of the previous lemma, then the induced sequences are automatically exact for all $\lambda$: in fact one has Theorems 7.1 and 6.5 and can apply Proposition 2.6.

**Remark 8.5** If one assumes that the conjecture $\text{LGF}_{K^s_0} (2)$ (see 2.5) for any PBQ $\varphi$-$\nabla$-module over $K[x]_0$ of rank $\leq r$, then the proofs of Proposition 7.3 and Theorem 8.1 works for any $\varphi$-$\nabla$-module over $K[x]_0$ of rank $\leq r$.

### 8.2 Dwork’s conjecture in the HBQ cases

**Lemma 8.6** Let $M$ be a HBQ $\varphi$-$\nabla$-module over $K[x]_0$ and let $N$ be a $\varphi$-$\nabla$-submodule of $M$ over $K[x]_0$ which is PBQ. Then $N$ is HPBQ. In particular, suppose that the residue field of $V$ is perfect and let $\{P_i(M)\}$ be the PBQ filtration of $M$, then $P_i(M)/P_{i-1}(M)$ is HPBQ for all $i$.

**Proof.** We have $\dim_k V(M)/V(M)^0 = \dim_k M_n/M_n^0$ and $\dim_k V(M/N)/V(M/N)^0 = \dim_k (M/N)_n/(M/N)^0_n$ by Proposition 6.3 since the quotient $M/N$ is HBQ by Proposition 6.4. Comparing the induced exact sequence $0 \to N_n/N_0^0 \to M_n/M_0^0 \to (M/N)_n/(M/N)^0_n \to 0$ at the generic point by Theorem 7.1 and Proposition 2.6 (1) to the corresponding right exact sequence at the special point, we have an inequality $\dim_k V(N)/V(N)^0 \geq \dim_k V(N)/V(N)^0$. On the contrary, we know the inequality $\dim_k V(N)/V(N)^0 \leq \dim_k N_n/N_0^0$ by [CT09, Proposition 4.10]. Hence, $\dim_k V(N)/V(N)^0 = \dim_k N_n/N_0^0$ and $N$ is HPBQ. The rest follows from the first part and Proposition 6.4. \qed

**Theorem 8.7** Let $M$ be a HBQ $\varphi$-$\nabla$-module over $K[x]_0$. Then the conjecture $\text{LGF}_{K[x]_0} (1)$ (see 2.5) and the conjecture $\text{LGF}_{Dw}$ (see 2.7) hold for $M$.

**Proof.** The assertions follows from the similar arguments of Theorems 7.2 and 8.1, respectively, by using Theorem 6.5 and Lemma 8.6. \qed

### 8.3 When do the generic and special log-growth polygons coincide?

**Theorem 8.8** Let $M$ be a $\varphi$-$\nabla$-module over $K[x]_0$. The special log-growth polygon and the generic log-growth polygon coincide with each other if and only if $M$ is equislope.
Proof. We may assume that the residue field of $\mathcal{V}$ is algebraically closed by Proposition 2.1. Let $(P_i(M))$ be the PBQ filtration of $M$ (Theorem 5.6). Each condition (i) the coincidence of special and generic log-growth polygons or (ii) equislope implies that $P_i(M)/P_{i-1}(M)$ is HPBQ and $M/P_i(M)$ is HBQ for all $i$ by Propositions 6.3, 6.4, and Lemma 8.6 for (i) and by Corollary 6.9 and Proposition 6.10 (1) for (ii). Then we can apply Lemma 8.3 (3) inductively on $i$ by Remark 8.4 and Theorem 8.7. Hence it is sufficient to prove the assertion when $M$ is HPBQ by Corollary 6.9. Then the coincidence of the log-growth filtration and the Frobenius slope filtration both at the special point (Theorem 8.7) and at the generic point (Theorem 7.1) implies our desired equivalence.

Example 8.9

(1) Let $M$ be a $\varphi$-$\nabla$-module over $K[[x]]_0$ such that $M_\eta$ is bounded. Then there is a $\varphi$-module $L$ over $K$ such that $M \cong L \otimes_K K[[x]]_0$ by Christol’s transfer theorem (see [CT09, Proposition 4.3]). Hence, $M$ is equislope.

(2) Let $M$ be a $\varphi$-$\nabla$-module over $K[[x]]_0$ of rank 2 such that $M_\eta$ is not bounded. Then we have identities $M^\lambda = (S_{\lambda - \lambda_{\text{max}}}(M^\vee))^\perp$ and $V(M)^\lambda = (S_{\lambda - \lambda_{\text{max}}}(V(M^\vee)))^\perp$ for any $\lambda$ [CT09, Theorem 7.1], where $\lambda_{\text{max}}$ is the highest Frobenius slope of $M_\eta$. Hence the special log-growth polygon and the generic log-growth polygon coincident with each other if and only if $M$ is equislope.

References

[An08] André, Y. Dwork’s conjecture on the logarithmic growth of solutions of $p$-adic differential equations, Compos. Math. 144 (2008), no. 2, 484–494.

[CT09] Chiarellotto, B., Tsuzuki, N., Logarithmic growth and Frobenius filtrations for solutions of $p$-adic differential equations, J. Math. Inst. Jussieu 8 (2009), 465–505.

[Ch83] Christol, G. Modules différentiels et équations différentielles $p$-adiques, Queen’s papers in Pure and Applied Math., 66, Queen’s Univ., Kingston ON, 1983.

[Dw73] Dwork, B.M. On $p$-adic differential equations III, Invent. Math. 20 (1973), 35–45.

[Dw82] Dwork, B.M. Lectures on $p$-adic differential equations, Grundlehren der mathematischen Wissenschaften, Volume 253 (Springer, 1982).

[dJ98] de Jong, A.J., Homomorphisms of Barsotti-Tate groups and crystals in positive characteristic, Invent. Math. 134 (1998), no. 2, 301–333.

[Ka73] Katz, N.M., Travaux de Dwork, Séminaire Bourbaki (1971/1972), Exp. No. 409, 167–200, Lecture Notes in Math., 317 (1973), Springer, Berlin.
[Ka79] Katz, N.M., Slope filtrations of $F$-isocrystals, Astérisque 63 (1979), 113–164.

[Ke01] Kedlaya, K.S., The algebraic closure of the power series field in positive characteristic, Proc. Amer. Math. Soc. 129 No. 12 (2002), 3461–3470.

[Ke08] Kedlaya, K.S., Slope filtration for relative Frobenius, Astérisque 319 (2008), 259–301.

[Ke09] Kedlaya, K.S., $p$-adic Differential Equations, book draft (version of 19 Jan 09).

[Ro75] Robba, Ph. On the Index of $p$-adic differential operators I, Ann. of Math. 101 (1975), 280–316.

[Ts96] Tsuzuki, N., The overconvergence of morphisms of étale $\phi$-$\nabla$-spaces on a local field, Compositio Math. 103 no. 2 (1996), 227–239.

[Ts98a] Tsuzuki, N. Slope filtration of quasi-unipotent overconvergent $F$-isocrystals, Annales de l’institut Fourier, Grenoble 48 (1998), 379–412.

[Ts98b] Tsuzuki, N. Finite local monodromy of overconvergent unit-root $F$-isocrystals on a curve, Amer. J. Math. 120 no. 6 (1998), 1165–1190.

Bruno Chiarellotto
Dipartimento Matematica
Università di Padova
Via Trieste 63
35121 Padova
Italy
chiarbru@math.unipd.it

Nobuo Tsuzuki
Mathematical Institute
Tohoku University
Aramaki Aza-Aoba 6-3
Aobaku Sendai 980-8578
Japan
tsuuki@math.tohoku.ac.jp
Multigraded Factorial Rings
and Fano Varieties with Torus Action

Jürgen Hausen, Elaine Herppich, Hendrik Süß

Received: November 15, 2009
Revised: December 12, 2010

Communicated by Thomas Peternell

Abstract. In a first result, we describe all finitely generated factorial algebras over an algebraically closed field of characteristic zero that come with an effective multigrading of complexity one by means of generators and relations. This enables us to construct systematically varieties with free divisor class group and a complexity one torus action via their Cox rings. For the Fano varieties of this type that have a free divisor class group of rank one, we provide explicit bounds for the number of possible deformation types depending on the dimension and the index of the Picard group in the divisor class group. As a consequence, one can produce classification lists for fixed dimension and Picard index. We carry this out exemplarily in the following cases. There are 15 non-toric surfaces with Picard index at most six. Moreover, there are 116 non-toric threefolds with Picard index at most two; nine of them are locally factorial, i.e. of Picard index one, and among these one is smooth, six have canonical singularities and two have non-canonical singularities. Finally, there are 67 non-toric locally factorial fourfolds and two one-dimensional families of non-toric locally factorial fourfolds. In all cases, we list the Cox rings explicitly.

2010 Mathematics Subject Classification: 13A02, 13F15, 14J45
Keywords and Phrases: Multigraded factorial ring, Fano variety, torus action

Documenta Mathematica 16 (2011) 71–109
Let $K$ be an algebraically closed field of characteristic zero. A first aim of this paper is to determine all finitely generated factorial $K$-algebras $R$ with an effective complexity one multigrading $R = \oplus_{u \in M} R_u$ satisfying $R_0 = K$; here effective complexity one multigrading means that with $d := \dim R$ we have $M \cong \mathbb{Z}_{d-1}$ and the $u \in M$ with $R_u \neq 0$ generate $M$ as a $\mathbb{Z}$-module. Our result extends work by Mori [23] and Ishida [17], who settled the cases $d = 2$ and $d = 3$.

An obvious class of multigraded factorial algebras as above is given by polynomial rings. A much larger class is obtained as follows. Take a sequence $A = (a_0, \ldots, a_r)$ of vectors $a_i \in K^2$ such that $(a_i, a_k)$ is linearly independent whenever $k \neq i$, a sequence $n = (n_0, \ldots, n_r)$ of positive integers and a family $L = (l_{ij})$ of positive integers, where $0 \leq i \leq r$ and $1 \leq j \leq n_i$. For every $0 \leq i \leq r$, we define a monomial

$$f_i := T_{i1}^{l_{i1}} \cdots T_{in_i}^{l_{in_i}} \in K[T_{ij}; \ 0 \leq i \leq r, \ 1 \leq j \leq n_i],$$

for any two indices $0 \leq i, j \leq r$, we set $\alpha_{ij} := \det(a_i, a_j)$, and for any three indices $0 \leq i < j < k \leq r$, we define a trinomial

$$g_{i,j,k} := \alpha_{jk} f_i + \alpha_{ki} f_j + \alpha_{ij} f_k \in K[T_{ij}; \ 0 \leq i \leq r, \ 1 \leq j \leq n_i].$$

Note that the coefficients of $g_{i,j,k}$ are all nonzero. The triple $(A, n, L)$ then defines a $K$-algebra

$$R(A, n, L) := K[T_{ij}; \ 0 \leq i \leq r, \ 1 \leq j \leq n_i] / (g_{i,i+1,i+2}; 0 \leq i \leq r - 2).$$

It turns out that $R(A, n, L)$ is a normal complete intersection, see Proposition 1.2. In particular, it is of dimension

$$\dim R(A, n, L) = n_0 + \ldots + n_r - r + 1.$$ 

If the triple $(A, n, L)$ is admissible, i.e., the numbers $\gcd(l_{i1}, \ldots, l_{in_i})$, where $0 \leq i \leq r$, are pairwise coprime, then $R(A, n, L)$ admits a canonical effective complexity one grading by a lattice $K$, see Construction 1.7. Our first result is the following.

**Theorem 1.9.** Up to isomorphy, the finitely generated factorial $K$-algebras with an effective complexity one grading $R = \oplus_M R_u$ and $R_0 = K$ are

(i) the polynomial algebras $K[T_1, \ldots, T_d]$ with a grading $\deg(T_i) = u_i \in \mathbb{Z}_{d-1}$ such that $u_1, \ldots, u_d$ generate $\mathbb{Z}_d$ as a lattice and the convex cone on $\mathbb{Q}^d$ generated by $u_1, \ldots, u_d$ is pointed,

(ii) the $(K \times \mathbb{Z}^m)$-graded algebras $R(A, n, L)[S_1, \ldots, S_m]$, where $R(A, n, L)$ is the $K$-graded algebra defined by an admissible triple $(A, n, L)$ and $\deg S_j \in \mathbb{Z}^m$ is the $j$-th canonical base vector.
The further paper is devoted to normal (possibly singular) \(d\)-dimensional Fano varieties \(X\) with an effective action of an algebraic torus \(T\). In the case \(\dim T = d\), we have the meanwhile extensively studied class of toric Fano varieties, see [3], [27] and [4] for the initiating work. Our aim is to show that the above Theorem provides an approach to classification results for the case \(\dim T = d - 1\), that means Fano varieties with a complexity one torus action. Here, we treat the case of divisor class group \(\Cl(X) \cong \mathbb{Z}\); note that in the toric setting this gives precisely the weighted projective spaces. The idea is to consider the Cox ring

\[
\mathcal{R}(X) = \bigoplus_{D \in \Cl(X)} \Gamma(X, \mathcal{O}_X(D)).
\]

The ring \(\mathcal{R}(X)\) is factorial, finitely generated as a \(\mathbb{K}\)-algebra and the \(T\)-action on \(X\) gives rise to an effective complexity one multigrading of \(\mathcal{R}(X)\) refining the \(\Cl(X)\)-grading, see [5] and [15]. Consequently, \(\mathcal{R}(X)\) is one of the rings listed in the first Theorem. Moreover, \(X\) can be easily reconstructed from \(\mathcal{R}(X)\); it is the homogeneous spectrum with respect to the \(\Cl(X)\)-grading of \(\mathcal{R}(X)\). Thus, in order to construct Fano varieties, we firstly have to figure out the Cox rings among the rings occurring in the first Theorem and then find those, which belong to a Fano variety; this is done in Propositions 1.11 and 2.5.

In order to produce classification results via this approach, we need explicit bounds on the number of deformation types of Fano varieties with prescribed discrete invariants. Besides the dimension, in our setting, a suitable invariant is the Picard index \([\Cl(X) : \Pic(X)]\). Denoting by \(\xi(\mu)\) the number of primes less or equal to \(\mu\), we obtain the following bound, see Corollary 2.2: for any pair \((d, \mu) \in \mathbb{Z}_{>0}^2\), the number \(\delta(d, \mu)\) of different deformation types of \(d\)-dimensional Fano varieties with a complexity one torus action such that \(\Cl(X) \cong \mathbb{Z}\) and \(\mu = [\Cl(X) : \Pic(X)]\) hold is bounded by

\[
\delta(d, \mu) \leq (6d\mu)^{2(3d\mu) + d - 2} \xi(\mu)^2 + 2\xi((d+2)\mu)^2 + 2d + 2.
\]

In particular, we conclude that for fixed \(\mu \in \mathbb{Z}_{>0}\), the number \(\delta(d)\) of different deformation types of \(d\)-dimensional Fano varieties with a complexity one torus action \(\Cl(X) \cong \mathbb{Z}\) and Picard index \(\mu\) is asymptotically bounded by \(d^{4d}\) with a constant \(A\) depending only on \(\mu\), see Corollary 2.4.

In fact, in Theorem 2.1 we even obtain explicit bounds for the discrete input data of the rings \(R(A, n, L)[S_1, \ldots, S_m]\). This allows us to construct all Fano varieties \(X\) with prescribed dimension and Picard index that come with an effective complexity one torus action and have divisor class group \(\mathbb{Z}\). Note that, by the approach, we get the Cox rings of the resulting Fano varieties \(X\) for free. In Section 3, we give some explicit classifications. We list all non-toric surfaces \(X\) with Picard index at most six and the non-toric threefolds \(X\) with Picard index up at most two. They all have a Cox ring defined by a single relation; in fact, for surfaces the first Cox ring with more than one relation
occurs for Picard index 29, and for the threefolds this happens with Picard index 3, see Proposition 3.5 as well as Examples 3.4 and 3.7. Moreover, we determine all locally factorial fourfolds \( X \), i.e. those of Picard index one: 67 of them occur sporadic and there are two one-dimensional families. Here comes the result on the locally factorial threefolds; in the table, we denote by \( w_i \) the \( \text{Cl}(X) \)-degree of the variable \( T_i \).

**Theorem 3.2.** The following table lists the Cox rings \( \mathcal{R}(X) \) of the three-dimensional locally factorial non-toric Fano varieties \( X \) with an effective two torus action and \( \text{Clim}(X) = \mathbb{Z} \).

| No. | \( \mathcal{R}(X) \) | \( (w_1, \ldots, w_5) \) | \( (-K_X)^3 \) |
|-----|----------------------|----------------------|------------------|
| 1   | \( \mathbb{K}[T_1, \ldots, T_5] / \left< T_1 T_2^2 + T_3^3 + T_4^2 \right> \) | (1, 1, 2, 3, 1) | 8                |
| 2   | \( \mathbb{K}[T_1, \ldots, T_5] / \left< T_1 T_2 T_3^3 + T_3^3 + T_4^2 \right> \) | (1, 1, 1, 2, 3) | 8                |
| 3   | \( \mathbb{K}[T_1, \ldots, T_5] / \left< T_1 T_2 T_3^3 + T_3^3 + T_4^2 \right> \) | (1, 1, 1, 2, 3) | 8                |
| 4   | \( \mathbb{K}[T_1, \ldots, T_5] / \left< T_1 T_2 + T_3 T_4 + T_5^2 \right> \) | (1, 1, 1, 1) | 54               |
| 5   | \( \mathbb{K}[T_1, \ldots, T_5] / \left< T_1 T_2^2 + T_3 T_4 + T_5^2 \right> \) | (1, 1, 1, 1, 1) | 24               |
| 6   | \( \mathbb{K}[T_1, \ldots, T_5] / \left< T_1 T_2^2 + T_3 T_4^2 + T_5^2 \right> \) | (1, 1, 1, 1) | 4                |
| 7   | \( \mathbb{K}[T_1, \ldots, T_5] / \left< T_1 T_2^2 + T_3 T_4^2 + T_5^2 \right> \) | (1, 1, 1, 1, 2) | 16               |
| 8   | \( \mathbb{K}[T_1, \ldots, T_5] / \left< T_1 T_2^2 + T_3 T_4^2 + T_5^2 \right> \) | (1, 1, 1, 1, 3) | 2                |
| 9   | \( \mathbb{K}[T_1, \ldots, T_5] / \left< T_1 T_2^2 + T_3 T_4^2 + T_5^2 \right> \) | (1, 1, 1, 1, 3) | 2                |

Note that each of these varieties \( X \) is a hypersurface in the respective weighted projective space \( \mathbb{P}(w_1, \ldots, w_5) \). Except number 4, none of them is quasismooth in the sense that Spec \( \mathcal{R}(X) \) is singular at most in the origin; quasismooth hypersurfaces of weighted projective spaces were studied in [21] and [7]. In Section 4, we take a closer look at the singularities of the threefolds listed above. It turns out that number 1, 3, 5, 7 and 9 are singular with only canonical singularities and all of them admit a crepant resolution. Number 6 and 8 are singular with non-canonical singularities but admit a smooth relative minimal model. Number two is singular with only canonical singularities, one of them of type \( \text{cA}_1 \), and it admits only a singular relative minimal model. Moreover, in all cases, we determine the Cox rings of the resolutions.

The authors would like to thank Ivan Arzhantsev for helpful comments and discussions and also the referee for valuable remarks and many references.
1 UFDs with complexity one multigrading

As mentioned before, we work over an algebraically closed field $\mathbb{K}$ of characteristic zero. In Theorem 1.9, we describe all factorial finitely generated $\mathbb{K}$-algebras $R$ with an effective complexity one grading and $R_0 = \mathbb{K}$. Moreover, we characterize the possible Cox rings among these algebras, see Proposition 1.11.

First we recall the construction sketched in the introduction.

**Construction 1.1.** Consider a sequence $A = (a_0, \ldots, a_r)$ of vectors $a_i = (b_i, c_i)$ in $\mathbb{K}^2$ such that any pair $(a_i, a_k)$ with $k \neq i$ is linearly independent, a sequence $n = (n_0, \ldots, n_r)$ of positive integers and a family $L = (l_{ij})$ of positive integers, where $0 \leq i \leq r$ and $1 \leq j \leq n_i$. For every $0 \leq i \leq r$, define a monomial

$$f_i := T_{i1}^{l_{i1}} \cdots T_{in_i}^{l_{in_i}} \in \mathbb{K}[T_{ij}; 0 \leq i \leq r, 1 \leq j \leq n_i],$$

for any two indices $0 \leq i, j \leq r$, set $\alpha_{ij} := \det(a_i, a_j) = b_i c_j - b_j c_i$ and for any three indices $0 \leq i < j < k \leq r$ define a trinomial

$$g_{i,j,k} := \alpha_{jk} f_i + \alpha_{ki} f_j + \alpha_{ij} f_k \in \mathbb{K}[T_{ij}; 0 \leq i \leq r, 1 \leq j \leq n_i].$$

Note that the coefficients of this trinomial are all nonzero. The triple $(A, n, L)$ then defines a ring

$$R(A, n, L) := \mathbb{K}[T_{ij}; 0 \leq i \leq r, 1 \leq j \leq n_i] / (g_{i+1,i+2}; 0 \leq i \leq r - 2).$$

**Proposition 1.2.** For every triple $(A, n, L)$ as in 1.1, the ring $R(A, n, L)$ is a normal complete intersection of dimension

$$\dim R(A, n, L) = n - r + 1, \quad n := n_0 + \ldots + n_r.$$

**Lemma 1.3.** In the setting of 1.1, one has for any $0 \leq i < j < k < l \leq r$ the identities

$$g_{i,k,l} = \alpha_{kl} \cdot g_{i,j,k} + \alpha_{ik} \cdot g_{j,k,l}, \quad g_{i,j,l} = \alpha_{jl} \cdot g_{i,j,k} + \alpha_{ij} \cdot g_{j,k,l}.$$ 

In particular, every trinomial $g_{i,j,k}$, where $0 \leq i < j < k \leq r$ is contained in the ideal $(g_{i+1,i+2}; 0 \leq i \leq r - 2)$.

**Proof.** The identities are easily obtained by direct computation; note that for this one may assume $a_j = (1, 0)$ and $a_k = (0, 1)$. The supplement then follows by repeated application of the identities.

**Lemma 1.4.** In the notation of 1.1 and 1.2, set $X := V(\mathbb{K}^n, g_{0}, \ldots, g_{r-2})$, and let $z \in X$. If we have $f_i(z) = f_j(z) = 0$ for two $0 \leq i < j \leq r$, then $f_k(z) = 0$ holds for all $0 \leq k \leq r$.

**Proof.** If $i < k < j$ holds, then, according to Lemma 1.3, we have $g_{i,k,j}(z) = 0$, which implies $f_k(z) = 0$. The cases $k < i$ and $j < k$ are obtained similarly.
Proof of Proposition 1.2. Set \( X := V(\mathbb{K}^n; g_0, \ldots, g_{r-2}) \), where \( g_i := g_{i,i+1,i+2} \). Then we have to show that \( X \) is a connected complete intersection with at most normal singularities. In order to see that \( X \) is connected, set \( \ell := \prod n_i \prod l_{ij} \) and \( \zeta_{ij} := \ell n_i^{-1} l_{ij}^{-1} \). Then \( X \subseteq \mathbb{K}^n \) is invariant under the \( \mathbb{K}^* \)-action given by

\[
t \cdot z := (t^\zeta_{ij} z_{ij})
\]

and the point \( 0 \in \mathbb{K}^n \) lies in the closure of any orbit \( \mathbb{K}^* x \subseteq X \), which implies connectedness. To proceed, consider the Jacobian \( J_g \) of \( g := (g_0, \ldots, g_{r-2}) \). According to Serre’s criterion, we have to show that the set of points of \( z \in X \) with \( J_g(z) \) not of full rank is of codimension at least two in \( X \). Note that the Jacobian \( J_g \) is of the shape

\[
J_g = \begin{pmatrix}
\delta_{00} & \delta_{01} & \delta_{02} & 0 & 0 \\
0 & \delta_{11} & \delta_{12} & \delta_{13} & 0 \\
0 & 0 & \delta_{r-3r-3} & \delta_{r-3r-2} & \delta_{r-3r-1} & 0 \\
0 & 0 & \delta_{r-2r-2} & \delta_{r-2r-1} & \delta_{r-2r} & 0
\end{pmatrix}
\]

where \( \delta_{il} \) is a nonzero multiple of the gradient \( \delta_i := \text{grad } f_i \). Consider \( z \in X \) with \( J_g(z) \) not of full rank. Then \( \delta_i(z) = 0 = \delta_k(z) \) holds with some \( 0 \leq i < k \leq r \). This implies \( z_{ij} = 0 = z_{kl} \) for some \( 1 \leq j \leq n_i \) and \( 1 \leq l \leq n_k \). Thus, we have \( f_i(z) = 0 = f_k(z) \). Lemma 1.4 gives \( f_s(z) = 0 \), for all \( 0 \leq s \leq r \). Thus, some coordinate \( z_{st} \) must vanish for every \( 0 \leq s \leq r \). This shows that \( z \) belongs to a closed subset of \( X \) having codimension at least two in \( X \). \( \Box \)

Lemma 1.5. Notation as in 1.1. Then the variable \( T_{ij} \) defines a prime ideal in \( R(A, n, L) \) if and only if the numbers \( \gcd(l_{k1}, \ldots, l_{kn}) \), where \( k \neq i \), are pairwise coprime.

Proof. We treat exemplarily \( T_{01} \). Using Lemma 1.3, we see that the ideal of relations of \( R(A, n, L) \) can be presented as follows

\[
\langle g_{s,s+1,s+2}; \ 0 \leq s \leq r - 2 \rangle = \langle g_{0,s,s+1}; \ 1 \leq s \leq r - 1 \rangle.
\]

Thus, the ideal \( \langle T_{01} \rangle \subseteq R(A, n, L) \) is prime if and only if the following binomial ideal is prime

\[
a := \langle \alpha_{s+1} f_s + \alpha_0 f_{s+1}; \ 1 \leq s \leq r - 1 \rangle \subseteq \mathbb{K}[T_{ij}; \ (i,j) \neq (0,1)].
\]

Set \( l_i := (l_{i1}, \ldots, l_{im}) \). Then the ideal \( a \) is prime if and only if the following family can be complemented to a lattice basis

\[
(l_1, -l_2, 0, \ldots, 0), \ldots, (0, \ldots, 0, l_{r-1}, -l_r).
\]

This in turn is equivalent to the statement that the numbers \( \gcd(l_{k1}, \ldots, l_{kn}) \), where \( 1 \leq k \leq r \), are pairwise coprime. \( \Box \)
**Definition 1.6.** We say that a triple \((A, n, L)\) as in 1.1 is *admissible* if the numbers \(\gcd(l_{i1}, \ldots, l_{in})\), where \(0 \leq i \leq r\), are pairwise coprime.

**Construction 1.7.** Let \((A,n, L)\) be an admissible triple and consider the following free abelian groups

\[
E := \bigoplus_{i=0}^{r} \bigoplus_{j=1}^{n_i} \mathbb{Z} \cdot e_{ij}, \\
K := \bigoplus_{j=1}^{n_0} \mathbb{Z} \cdot u_{0j} \bigoplus \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{n_i-1} \mathbb{Z} \cdot u_{ij}
\]

and define vectors \(u_{in_i} := u_{01} + \ldots + u_{0r} - u_{i1} - \ldots - u_{in_i-1} \in K\). Then there is an epimorphism \(\lambda: E \to K\) fitting into a commutative diagram with exact rows

\[
\begin{array}{ccc}
0 & \xrightarrow{\alpha} & E \\
\cong & & \xrightarrow{\lambda} K \\
0 & \xrightarrow{\beta} & K
\end{array}
\]

Define a \(K\)-grading of \(\mathbb{K}[T_{ij}; 0 \leq i \leq r, 1 \leq j \leq n_i]\) by setting \(\deg(T_{ij}) := \lambda(e_{ij})\). Then every \(f_i = T_{i1}^{l_{i1}} \cdots T_{in_i}^{l_{in_i}}\) is \(K\)-homogeneous of degree

\[
\deg f_i = l_{i1} \lambda(e_{i1}) + \ldots + l_{in_i} \lambda(e_{in_i}) = l_{01} \lambda(e_{01}) + \ldots + l_{0n_0} \lambda(e_{0n_0}) \in K.
\]

Thus, the polynomials \(g_{i,j,k}\) of 1.1 are all \(K\)-homogeneous of the same degree and we obtain an effective \(K\)-grading of complexity one of \(R(A, n, L)\).

**Proof.** Only for the existence of the commutative diagram there is something to show. Write for short \(l_i := (l_{i1}, \ldots, l_{in_i})\). By the admissibility condition, the vectors \(e_i := (0, \ldots, 0, l_{i1}, -l_{i1+1}, 0, \ldots, 0)\), where \(0 \leq i \leq r-1\), can be completed to a lattice basis for \(E\). Consequently, we find an epimorphism \(\lambda: E \to K\) having precisely \(\text{lin}(v_0, \ldots, v_{r-1})\) as its kernel. By construction, \(\ker(\lambda) = \alpha(\ker(\eta))\). Using this, we obtain the induced morphism \(\beta: K \to K\) and the desired properties. \(\square\)

**Lemma 1.8.** Notation as in 1.7. Then \(R(A, n, L)_0 = \mathbb{K}\) and \(R(A, n, L)^* = \mathbb{K}^*\) hold. Moreover, the \(T_{ij}\) define pairwise nonassociated prime elements in \(R(A, n, L)\).

**Proof.** The fact that all elements of degree zero are constant is due to the fact that all degrees \(\deg(T_{ij}) = u_{ij} \in K\) are non-zero and generate a pointed convex cone in \(K_Q\). As a consequence, we obtain that all units in \(R(A, n, L)\) are constant. The \(T_{ij}\) are prime by the admissibility condition and Lemma 1.5, and they are pairwise nonassociated because they have pairwise different degrees and all units are constant. \(\square\)
Theorem 1.9. Up to isomorphy, the finitely generated factorial \( \mathbb{K} \)-algebras with an effective complexity one grading \( R = \oplus M R_u \) and \( R_0 = \mathbb{K} \) are

(i) the polynomial algebras \( \mathbb{K}[T_1, \ldots, T_d] \) with a grading \( \text{deg}(T_i) = u_i \in \mathbb{Z}^{d-1} \) such that \( u_1, \ldots, u_d \) generate \( \mathbb{Z}^{d-1} \) as a lattice and the convex cone on \( \mathbb{Q}^{d-1} \) generated by \( u_1, \ldots, u_d \) is pointed,

(ii) the \( (K \times \mathbb{Z}^m) \)-graded algebras \( R(A, n, L)[S_1, \ldots, S_m] \), where \( R(A, n, L) \) is the \( K \)-graded algebra defined by an admissible triple \( (A, n, L) \) as in 1.1 and 1.7 and \( \text{deg} S_j \in \mathbb{Z}^m \) is the \( j \)-th canonical base vector.

Proof. We first show that for any admissible triple \( (A, n, L) \) the ring \( R(A, n, L) \) is a unique factorization domain. If \( l_{ij} = 1 \) holds for any two \( i, j \), then, by [15, Prop. 2.4], the ring \( R(A, n, L) \) is the Cox ring of a space \( \mathbb{P}_1(A, n) \) and hence is a unique factorization domain.

Now, let \( (A, n, L) \) be arbitrary admissible data and let \( \lambda : E \to K \) be an epimorphism as in 1.7. Set \( n := n_0 + \ldots + n_r \) and consider the diagonalizable groups

\[ T^n := \text{Spec } \mathbb{K}[E], \quad H := \text{Spec } \mathbb{K}[K], \quad H_0 := \text{Spec } \mathbb{K}[\oplus_{i,j} \mathbb{Z}/l_{ij} \mathbb{Z}]. \]

Then \( T^n = (K^*)^n \) is the standard \( n \)-torus and \( H_0 \) is the direct product of the cyclic subgroups \( H_{ij} := \text{Spec } \mathbb{K}[\mathbb{Z}/l_{ij} \mathbb{Z}] \). Moreover, the diagram in 1.7 gives rise to a commutative diagram with exact rows

\[
\begin{array}{ccccccccc}
0 & \longrightarrow & T^n & \stackrel{(t_{ij}) \longleftarrow (t_{ij})}{\longrightarrow} & T^n & \longrightarrow & H_{0} & \longrightarrow & 0 \\
& & \uparrow & & \uparrow & & \uparrow & & \\
0 & \longrightarrow & H & \longrightarrow & H & \longrightarrow & H_0 & \longrightarrow & 0
\end{array}
\]

where \( t_{ij} = \chi^{e_{ij}} \) are the coordinates of \( T^n \) corresponding to the characters \( e_{ij} \in E \) and the maps \( i, j \) are the closed embeddings corresponding to the epimorphisms \( \eta, \lambda \) respectively.

Setting \( \text{deg} T_{ij} := e_{ij} \) defines an action of \( T^n \) on \( \mathbb{K}^n = \text{Spec } \mathbb{K}[T_{ij}] \); in terms of the coordinates \( z_{ij} \) corresponding to \( T_{ij} \) this action is given by \( t \cdot z = (t_{ij} z_{ij}) \).

The torus \( H \) acts effectively on \( \mathbb{K}^n \) via the embedding \( j : H \to T^n \). The generic isotropy group of \( H \) along \( V(\mathbb{K}^n, T_{ij}) \) is the subgroup \( H_{ij} \subseteq H \) corresponding to \( K \to K/\lambda(E_{ij}) \), where \( E_{ij} \subseteq E \) denotes the sublattice generated by all \( e_{kl} \) with \( (k, l) \neq (i, j) \); recall that we have \( K/\lambda(E_{ij}) \cong \mathbb{Z}/l_{ij} \mathbb{Z} \).

Now, set \( t'_{ij} := 1 \) for any two \( i, j \) and consider the spectra \( X := \text{Spec } R(A, n, L) \) and \( X' := \text{Spec } R(A, n, L') \). Then the canonical surjections \( \mathbb{K}[T_{ij}] \to R(A, n, L) \) and \( \mathbb{K}[T_{ij}] \to R(A, n, L') \) define embeddings \( X \to \mathbb{K}^n \) and \( X' \to \mathbb{K}^n \). These
embeddings fit into the following commutative diagram

\[
\begin{array}{ccc}
\mathbb{K}^n & \xleftarrow{(z_{ij})^*} & \mathbb{K}^n \\
\uparrow & & \uparrow \\
X' & \xleftarrow{\pi} & X
\end{array}
\]

The action of \( H \) leaves \( X \) invariant and the induced \( H \)-action on \( X \) is the one given by the \( K \)-grading of \( R(A, n, L) \). Moreover, \( \pi : \mathbb{K}^n \to \mathbb{K}^n \) is the quotient map for the induced action of \( H_0 \subseteq H \) on \( \mathbb{K}^n \), we have \( X = \pi^{-1}(X') \), and hence the restriction \( \pi : X \to X' \) is a quotient map for the induced action of \( H_0 \) on \( X \).

Removing all subsets \( V(X; T_{ij}, T_{kl}) \), where \((i, j) \neq (k, l)\) from \( X \), we obtain an open subset \( U \subseteq X \). By Lemma 1.8, the complement \( X \setminus U \) is of codimension at least two and each \( V(U, T_{ij}) \) is irreducible. By construction, the only isotropy groups of the \( H \)-action on \( U \) are the groups \( H_{ij} \) of the points of \( V(U, T_{ij}) \). The image \( U' := \pi(U) \) is open in \( X' \), the complement \( X' \setminus U' \) is as well of codimension at least two and \( H/H_0 \) acts freely on \( U' \). According to [22, Cor. 5.3], we have two exact sequences fitting into the following diagram

\[
\begin{array}{cccc}
1 & \to & \text{Pic}(U') & \\
& & \downarrow \pi^* & \\
& & \text{Pic}(H_0) & \xrightarrow{\beta} \text{Pic}(U) \\
1 & \xrightarrow{\alpha} & \text{Pic}_{H_0}(U) & \\
& & \downarrow \beta & \\
& & \Pi_{i,j} \mathbb{K}(H_{ij}) & \\
\end{array}
\]

Since \( X' \) is factorial, the Picard group \( \text{Pic}(U') \) is trivial and we obtain that \( \delta \) is injective. Since \( H_0 \) is the direct product of the isotropy groups \( H_{ij} \) of the Luna strata \( V(U, T_{ij}) \), we see that \( \delta \circ \alpha \) is an isomorphism. It follows that \( \delta \) is surjective and hence an isomorphism. This in turn shows that \( \alpha \) is an isomorphism. Now, every bundle on \( U \) is \( H \)-linearizable. Since \( H_0 \) acts as a subgroup of \( H \), we obtain that every bundle is \( H_0 \)-linearizable. It follows that \( \beta \) is surjective and hence \( \text{Pic}(U) \) is trivial. We conclude \( \text{Cl}(X) = \text{Pic}(U) = 0 \), which means that \( R(A, n, L) \) admits unique factorization.

The second thing we have to show is that any finitely generated factorial \( \mathbb{K} \)-algebra \( R \) with an effective complexity one multigrading satisfying \( R_0 = \mathbb{K} \) is as claimed. Consider the action of the torus \( G \) on \( X = \text{Spec} R \) defined by the multigrading, and let \( X_0 \subseteq X \) be the set of points having finite isotropy \( G_x \).
Then [15, Prop 3.3] provides a graded splitting
\[ R \cong R'[S_1, \ldots, S_m], \]
where the variables \( S_j \) are identified with the homogeneous functions defining the prime divisors \( E_j \) inside the boundary \( X \setminus X_0 \) and \( R' \) is the ring of functions of \( X_0 \), which are invariant under the subtorus \( G_0 \subseteq G \) generated by the generic isotropy groups \( G_j \) of \( E_j \).

Since \( R'_0 = R_0 = K \) holds, the orbit space \( X_0/G \) has only constant functions and thus is a space \( \mathbb{P}_1(A, n) \) as constructed in [15, Section 2]. This allows us to proceed exactly as in the proof of Theorem [15, Thm 1.3] and gives \( P \) and thus is a space.

**Remark 1.10.** Let \((A, n, L)\) be an admissible triple with \( n = (1, \ldots, 1) \). Then \( K = \mathbb{Z} \) holds, the admissibility condition just means that the numbers \( l_{ij} \) are pairwise coprime and we have
\[ \dim R(A, n, L) = n_0 + \ldots + n_r - r + 1 = 2. \]
Consequently, for two-dimensional rings, Theorem 1.9 specializes to Mori’s description of almost geometrically graded two-dimensional unique factorization domains provided in [23].

**Proposition 1.11.** Let \((A, n, L)\) be an admissible triple, consider the associated \((K \times \mathbb{Z}^m)\)-graded ring \( R(A, n, L)[S_1, \ldots, S_m] \) as in Theorem 1.9 and let \( \mu: K \times \mathbb{Z}^m \to K' \) be a surjection onto an abelian group \( K' \). Then the following statements are equivalent.

(i) The \( K' \)-graded ring \( R(A, n, L)[S_1, \ldots, S_m] \) is the Cox ring of a projective variety \( X' \) with \( \text{Cl}(X') \cong K' \).

(ii) For every pair \( i, j \) with \( 0 \leq i \leq r \) and \( 1 \leq j \leq n_i \), the group \( K' \) is generated by the elements \( \mu(\lambda(e_{ij})) \) and \( \mu(e_s) \), where \( (i, j) \neq (k, l) \) and \( 1 \leq s \leq m, \) for every \( 1 \leq t \leq m, \) the group \( K' \) is generated by the elements \( \mu(\lambda(e_{ij})) \) and \( \mu(e_s) \), where \( 0 \leq i \leq r, 1 \leq j \leq n_i \) and \( s \neq t, \) and, finally the following cone is of full dimension in \( K'_Q^1 \):
\[ \bigcap_{(k, l)} \text{cone}(\mu(\lambda(e_{ij})), \mu(e_s); (i, j) \neq (k, l)) \cap \bigcap_t \text{cone}(\mu(\lambda(e_{ij})), \mu(e_s); s \neq t). \]

**Proof.** Suppose that (i) holds, let \( p: \tilde{X}' \to X' \) denote the universal torsor and let \( X'' \subseteq X' \) be the set of smooth points. According to [14, Prop. 2.2], the group \( H' = \text{Spec } K[K'] \) acts freely on \( p^{-1}(X'') \), which is a big open subset of the total coordinate space \( \text{Spec } R(A, n, L)[S_1, \ldots, S_m] \). This implies the first condition of (ii). Moreover, by [14, Prop. 4.1], the displayed cone is the moving cone of \( X' \) and hence of full dimension. Conversely, if (ii) holds, then the \( K' \)-graded ring \( R(A, n, L)[S_1, \ldots, S_m] \) can be made into a bunched ring and hence is the Cox ring of a projective variety, use [14, Thm. 3.6].
Theorem 2.1

The following result provides bounds for the discrete data of the Cox ring.

Moreover, for \( n \) we have

Here, we may (and will) assume \( n_0 \geq \ldots \geq n_r \geq 1 \). With \( n := n_0 + \ldots + n_r \), we have \( n + m = d + r \). For the degrees of the variables in \( \text{Cl}(X) \cong \mathbb{Z} \), we write \( w_{ij} := \deg T_{ij} \) for \( 0 \leq i \leq r \), \( 1 \leq j \leq n_i \), and \( u_k := \deg S_k \) for \( 1 \leq k \leq m \). Moreover, for \( \mu \in \mathbb{Z}_{\geq 0} \), we denote by \( \xi(\mu) \) the number of primes in \( \{2, \ldots, \mu\} \). The following result provides bounds for the discrete data of the Cox ring.

Theorem 2.1. In the above situation, fix the dimension \( d = \dim(X) \) and the Picard index \( \mu = |\text{Cl}(X) : \text{Pic}(X)| \). Then we have

for \( 1 \leq k \leq m \).

Moreover, for the degree \( \gamma \) of the relations, the weights \( w_{ij} \) and the exponents \( l_{ij} \), where \( 0 \leq i \leq r \) and \( 1 \leq j \leq n_i \), one obtains the following.

(i) Suppose that \( r = 0, 1 \) holds. Then \( n + m = d + 1 \) holds and one has the bounds

and the Picard index is given by

(ii) Suppose that \( r \geq 2 \) and \( n_0 = 1 \) hold. Then \( r \leq \xi(\mu) - 1 \) and \( n = r + 1 \) and \( m = d - 1 \) hold and one has

and the Picard index is given by

(iii) Suppose that \( r \geq 2 \) and \( n_0 > n_1 = 1 \) hold. Then we may assume \( l_{11} > \ldots > l_{r1} \geq 2 \), we have \( r \leq \xi(3d\mu) - 1 \) and \( n_0 + m = d \) and the bounds

\[
\begin{align*}
  w_{01}, \ldots, w_{0n_0} &\leq \mu, \\
  l_{01}, \ldots, l_{0n_0} &< 6d\mu, \\
  w_{11}, l_{21} &< 2d\mu, \\
  w_{21}, l_{11} &< 3d\mu,
\end{align*}
\]
The Picard index is given by
\[ \mu = \text{lcm}(w_{ij}, \text{gcd}(w_{11}, \ldots, w_{1r}), u_k; 1 \leq j \leq n_0, 1 \leq k \leq m). \]

(iv) Suppose that \( n_1 > n_2 = 1 \) holds. Then we may assume \( l_{21} > \ldots > l_{1r} \geq 2 \), we have \( r \leq \xi(2(d+1)\mu) - 1 \) and \( n_0 + n_1 + m = d + 1 \) and the bounds
\[ w_{ij} \leq \mu \quad \text{for } i = 0, 1 \text{ and } 1 \leq j \leq n_i, \quad w_{21} < (d+1)\mu, \]
\[ l_{ij} < 2(d+1)\mu \quad \text{for } 0 \leq i \leq r \text{ and } 1 \leq j \leq n_i, \]
\[ l_{11} \cdots l_{1r} \mid \gamma < 2(d+1)\mu, \]
and the Picard index is given by
\[ \mu = \text{lcm}(w_{ij}, u_k; 0 \leq i \leq s, 1 \leq j \leq n_i, 1 \leq k \leq m). \]

(v) Suppose that \( n_2 > 1 \) holds and let \( s \) be the maximal number with \( n_s > 1 \). Then one may assume \( l_{s+1,1} > \ldots > l_{1r} \geq 2 \), we have \( r \leq \xi((d+2)\mu) - 1 \) and \( n_0 + \ldots + n_s + m = d + s \) and the bounds
\[ w_{ij} \leq \mu \quad \text{for } 0 \leq i \leq s, \]
\[ w_{ij}, l_{ij} < (d+2)\mu \quad \text{for } 0 \leq i \leq r \text{ and } 1 \leq j \leq n_i, \]
\[ l_{s+1,1} \cdots l_{1r} \mid \gamma < (d+2)\mu, \]
and the Picard index is given by
\[ \mu = \text{lcm}(w_{ij}, u_k; 0 \leq i \leq s, 1 \leq j \leq n_i, 1 \leq k \leq m). \]

Putting all the bounds of the theorem together, we obtain the following (raw) bound for the number of deformation types.

**Corollary 2.2.** For any pair \((d, \mu) \in \mathbb{Z}^2_+\), the number \( \delta(d, \mu) \) of different deformation types of \( d \)-dimensional Fano varieties with a complexity one torus action such that \( \text{Cl}(X) \cong \mathbb{Z} \) and \( [\text{Cl}(X) : \text{Pic}(X)] = \mu \) hold is bounded by
\[ \delta(d, \mu) \leq (6d\mu)^{2((3d\mu) + d - 2)\mu(\mu)^2 + 2((d+1)\mu + 2d + 2).} \]

**Proof.** By Theorem 2.1 the discrete data \( r, n, L \) and \( m \) occurring in \( \mathcal{R}(X) \) are bounded as in the assertion. The continuous data in \( \mathcal{R}(X) \) are the coefficients \( \alpha_{ij} \); they stem from the family \( A = (a_0, \ldots, a_r) \) of points \( a_i \in \mathbb{K}^2 \). Varying the \( a_i \) provides flat families of Cox rings and hence, by passing to the homogeneous spectra, flat families of the resulting Fano varieties \( X \).
Corollary 2.3. Fix $d \in \mathbb{Z}_{>0}$. Then the number $\delta(\mu)$ of different deformation types of $d$-dimensional Fano varieties with a complexity one torus action, $\text{Cl}(X) \cong \mathbb{Z}$ and Picard index $\mu := [\text{Cl}(X) : \text{Pic}(X)]$ is asymptotically bounded by $\mu^A d^2/\log^2 \mu$ with a constant $A$ depending only on $d$.

Corollary 2.4. Fix $\mu \in \mathbb{Z}_{>0}$. Then the number $\delta(d)$ of different deformation types of $d$-dimensional Fano varieties with a complexity one torus action, $\text{Cl}(X) \cong \mathbb{Z}$ and Picard index $\mu := [\text{Cl}(X) : \text{Pic}(X)]$ is asymptotically bounded by $d^{A_d}$ with a constant $A$ depending only on $\mu$.

We first recall the necessary facts on Cox rings, for details, we refer to [14]. Let $X$ be a complete $d$-dimensional variety with divisor class group $\text{Cl}(X) \cong \mathbb{Z}$. Then the Cox ring $\mathcal{R}(X)$ is finitely generated and the total coordinate space $\mathcal{X} := \text{Spec} \mathcal{R}(X)$ is a factorial affine variety coming with an action of $\mathbb{K}^*$ defined by the $\text{Cl}(X)$-grading of $\mathcal{R}(X)$. Choose a system $f_1, \ldots, f_\nu$ of homogeneous pairwise nonassociated prime generators for $\mathcal{R}(X)$. This provides an $\mathbb{K}^*$-equivariant embedding

$$\mathcal{X} \rightarrow \mathbb{K}^\nu, \quad \varpi \mapsto (f_1(\varpi), \ldots, f_\nu(\varpi)).$$

where $\mathbb{K}^*$ acts diagonally with the weights $w_j = \deg(f_j) \in \text{Cl}(X) \cong \mathbb{Z}$ on $\mathbb{K}^\nu$. Moreover, $X$ is the geometric $\mathbb{K}^*$-quotient of $\check{X} := \mathcal{X} \setminus \{0\}$, and the quotient map $p: \check{X} \rightarrow X$ is a universal torsor. By the local divisor class group $\text{Cl}(X, x)$ of a point $x \in X$, we mean the group of Weil divisors $\text{WDiv}(X)$ modulo those that are principal near $x$.

Proposition 2.5. For any $\varpi = (\varpi_1, \ldots, \varpi_\nu) \in \check{X}$ the local divisor class group $\text{Cl}(X, x)$ of $x := p(\varpi)$ is finite of order $\gcd(w_i; \varpi_i \neq 0)$. The index of the Picard group $\text{Pic}(X)$ in $\text{Cl}(X)$ is given by

$$[\text{Cl}(X) : \text{Pic}(X)] = \text{lcm}_{x \in X}([\text{Cl}(X, x)]).$$

Suppose that the ideal of $\mathcal{X} \subseteq \mathbb{K}^\nu$ is generated by $\text{Cl}(X)$-homogeneous polynomials $g_1, \ldots, g_\nu - d - 1$ of degree $\gamma_j := \deg(g_j)$. Then one obtains

$$-\mathcal{K}_X = \sum_{i=1}^{\nu} w_i - \sum_{j=1}^{\nu-d-1} \gamma_j, \quad (-\mathcal{K}_X)^d = \left( \sum_{i=1}^{\nu} w_i - \sum_{j=1}^{\nu-d-1} \gamma_j \right)^d \prod_{i=1}^\nu \gamma_{i-d-1}$$

for the anticanonical class $-\mathcal{K}_X \in \text{Cl}(X) \cong \mathbb{Z}$. In particular, $X$ is a Fano variety if and only if the following inequality holds

$$\sum_{j=1}^{\nu-d-1} \gamma_j < \sum_{i=1}^{\nu} w_i.$$

Proof. Using [14, Prop. 2.2, Thm. 4.19], we observe that $X$ arises from the bunched ring $(R, \mathfrak{g}, \Phi)$, where $R = \mathcal{R}(X)$, $\mathfrak{g} = (f_1, \ldots, f_\nu)$ and $\Phi = \{Q_{\geq 0}\}$.
The descriptions of local class groups, the Picard index and the anticanonical class are then special cases of [14, Prop. 4.7, Cor. 4.9 and Cor. 4.16]. The anticanonical self-intersection number is easily computed in the ambient weighted projective space $\mathbb{P}(w_1, \ldots, w_r)$, use [14, Constr. 3.13, Cor. 4.13].

**Remark 2.6.** If the ideal of $\mathbf{X} \subseteq \mathbb{K}^r$ is generated by $\text{Cl}(X)$-homogeneous polynomials $g_1, \ldots, g_{r-d-1}$, then [14, Constr. 3.13, Cor. 4.13] show that $X$ is a well formed complete intersection in the weighted projective space $\mathbb{P}(w_1, \ldots, w_r)$ in the sense of [16, Def. 6.9].

We turn back to the case that $X$ comes with a complexity one torus action as at the beginning of this section. We consider the case $n_0 = \ldots = n_r = 1$, that means that each relation $g_{i,j,k}$ of the Cox ring $\mathcal{R}(X)$ depends only on three variables. Then we may write $T_i$ instead of $T_{i1}$ and $w_i$ instead of $w_{ii}$, etc.. In this setting, we obtain the following bounds for the numbers of possible varieties $X$ (Fano or not).

**Proposition 2.7.** For any pair $(d, \mu) \in \mathbb{Z}^2_{>0}$ there is, up to deformation, only a finite number of complete $d$-dimensional varieties with divisor class group $\mathbb{Z}$, Picard index $[\text{Cl}(X) : \text{Pic}(X)] = \mu$ and Cox ring $\mathbb{K}[T_0, \ldots, T_r, S_1, \ldots, S_m] / \langle \alpha_1, \ldots, \alpha_{r+2}, T_i, T_i, T_{i,i+1} \rangle : 0 \leq i \leq r - 2).

In this situation we have $r \leq \xi(\mu) - 1$. Moreover, for the weights $w_i := \text{deg } T_i$, where $0 \leq i \leq r$ and $u_k := \text{deg } S_k$, where $1 \leq k \leq m$, the exponents $l_i$ and the degree $\gamma := l_0 w_0$ of the relation one has

$$l_0 \cdots l_r \mid \gamma, \quad l_0 \cdots l_r \mid \mu, \quad w_i \leq \mu^{\xi(\mu)-1}, \quad u_k \leq \mu.$$

**Proof.** Consider the total coordinate space $\mathbf{X} \subseteq \mathbb{K}^{r+1+n}$ and the universal torsor $p : \overline{X} \to X$ as discussed before. For each $0 \leq i \leq r$ fix a point $\overline{x}(i) = (\overline{X}_0, \ldots, \overline{X}_r, 0, \ldots, 0)$ in $\overline{X}$ such that $\overline{x}_i = 0$ and $\overline{x}_j \neq 0$ for $j \neq i$ hold. Then, denoting $x(i) := p(\overline{x}(i))$, we obtain

$$\gcd(w_j; j \neq i) = |\text{Cl}(X, x(i))| \mid \mu.$$

Consider $i, j$ with $j \neq i$. Since all relations are homogeneous of the same degree, we have $l_i w_i = l_j w_j$. Moreover, by the admissibility condition, $l_i$ and $l_j$ are coprime. We conclude $l_i | w_j$ for all $j \neq i$ and hence $l_i | \gcd(w_j; j \neq i)$. This implies

$$l_0 \cdots l_r \mid l_0 w_0 = \gamma, \quad l_0 \cdots l_r \mid \mu.$$

We turn to the bounds for the $w_i$, and first verify $w_0 \leq \mu^r$. Using the relation $l_i w_i = l_0 w_0$, we obtain for every $l_i$ a presentation

$$l_i = l_0 \frac{w_0 \cdots w_{i-1}}{w_1 \cdots w_i} = \eta_i \frac{\gcd(w_0, \ldots, w_{i-1})}{\gcd(w_0, \ldots, w_i)}$$
with suitable integers $1 \leq \eta_i \leq \mu$. In particular, the very last fraction is bounded by $\mu$. This gives the desired estimate:
\[
w_0 = \frac{w_0}{\gcd(w_0, w_1)} \cdots \frac{\gcd(w_0, \ldots, w_{r-2})}{\gcd(w_0, \ldots, w_{r-1})} \leq \mu^r.
\]

Similarly, we obtain $w_i \leq \mu^r$ for $1 \leq i \leq r$. Then we only have to show that $r+1$ is bounded by $\xi(\mu)$, but this follows immediately from the fact that $l_0, \ldots, l_r$ are pairwise coprime.

Finally, to estimate the $u_k$, consider the points $\pi(k) \in \tilde{X}$ having the $(r+k)$-th coordinate one and all others zero. Set $x(k) := p(\pi(k))$. Then $\text{Cl}(X, x(k))$ is of order $u_k$, which implies $u_k \leq \mu$.

**Lemma 2.8.** Consider the ring $\mathbb{K}[T_{ij}; 0 \leq i \leq 2, 1 \leq j \leq n_1|S_1, \ldots, S_k]/(g)$ where $n_0 \geq n_1 \geq n_2 \geq 1$ holds. Suppose that $g$ is homogeneous with respect to a $\mathbb{Z}$-grading of $\mathbb{K}[T_{ij}, S_k]$ given by $\deg T_{ij} = w_{ij} \in \mathbb{Z}_{>0}$ and $\deg S_k = u_k \in \mathbb{Z}_{>0}$, and assume
\[
\deg g < 2 \sum_{i=0}^{n_1} \sum_{j=1}^{w_{ij}} + \sum_{i=1}^{m} u_i.
\]

Let $\mu \in \mathbb{Z}_{>1}$, assume $w_{ij} \leq \mu$ whenever $n_i > 1$, $1 \leq j \leq n_i$ and $u_k \leq \mu$ for $1 \leq k \leq m$ and set $d := n_0 + n_1 + n_2 + m - 2$. Depending on the shape of $g$, one obtains the following bounds.

(i) Suppose that $g = \eta_0 T_{01}^{l_{00}} \cdots T_{0n_0}^{l_{0n_0}} + \eta_1 T_{11}^{l_{11}} + \eta_2 T_{21}^{l_{21}}$ with $n_0 > 1$ and coefficients $\eta_i \in \mathbb{K}^*$ holds, we have $l_{11} \geq l_{21} \geq 2$ and $l_{11}, l_{21}$ are coprime. Then, one has
\[
w_{11}, l_{21} < 2d\mu, \quad w_{21}, l_{11} < 3d\mu, \quad \deg g < 6d\mu.
\]

(ii) Suppose that $g = \eta_0 T_{01}^{l_{01}} \cdots T_{0n_0}^{l_{0n_0}} + \eta_1 T_{11}^{l_{11}} \cdots T_{1n_1}^{l_{1n_1}} + \eta_2 T_{21}^{l_{21}}$ with $n_1 > 1$ and coefficients $\eta_i \in \mathbb{K}^*$ holds and we have $l_{21} \geq 2$. Then one has
\[
w_{21} < (d+1)\mu, \quad \deg g < 2(d+1)\mu.
\]

**Proof.** We prove (i). Set for short $c := (n_0 + m)\mu = d\mu$. Then, using homogeneity of $g$ and the assumed inequality, we obtain
\[
l_{11} w_{11} = l_{21} w_{21} = \deg g < 2 \sum_{i=0}^{n_1} \sum_{j=1}^{w_{ij}} + \sum_{i=1}^{m} u_i \leq c + w_{11} + w_{21}.
\]

Since $l_{11}$ and $l_{21}$ are coprime, we have $l_{11} > l_{21} > 2$. Plugging this into the above inequalities, we arrive at $2w_{11} < c + w_{21}$ and $w_{21} < c + w_{11}$. We conclude
w_{11} < 2c \text{ and } w_{21} < 3c. \text{ Moreover, } l_{11}w_{11} = l_{21}w_{21} \text{ and } \gcd(l_{11}, l_{21}) = 1 \text{ imply } l_{11}|w_{21} \text{ and } l_{21}|w_{11}. \text{ This shows } l_{11} < 3c \text{ and } l_{21} < 2c. \text{ Finally, we obtain }

\text{deg } g < c + w_{11} + w_{21} < 6c.

We prove (ii). Here we set \( c := (n_0 + n_1 + m)\mu = (d+1)\mu \). Then the assumed inequality gives

\[ l_{21}w_{21} = \text{deg } g < \sum_{i=0}^{1} n_i \sum_{j=1}^{w_{ij}} + \sum_{i=1}^{m} u_i + w_{21} \leq c + w_{21}. \]

Since we assumed \( l_{21} \geq 2 \), we can conclude \( w_{21} < c \). This in turn gives us \( \text{deg } g < 2c \) for the degree of the relation. \( \square \)

**Proof of Theorem 2.1.** As before, we denote by \( \overline{X} \subseteq \mathbb{K}^{n+m} \) the total coordinate space and by \( p: \overline{X} \to X \) the universal torsor.

We first consider the case that \( X \) is a toric variety. Then the Cox ring is a polynomial ring, \( \mathcal{R}(X) = \mathbb{K}[S_1, \ldots, S_m] \). For each \( 1 \leq k \leq m \), consider the point \( \overline{x}(k) \in \overline{X} \) having the \( k \)-th coordinate one and all others zero and set \( x(k) := p(\overline{x}(k)) \). Then, by Proposition 2.5, the local class group \( \text{Cl}(X, x(k)) \) is of order \( u_k \) where \( u_k := \text{deg } S_k \). This implies \( u_k \leq \mu \) for \( 1 \leq k \leq m \) and settles Assertion (i).

Now we treat the non-toric case, which means \( r \geq 2 \). Note that we have \( n \geq 3 \).

The case \( n_0 = 1 \) is done in Proposition 2.7. So, we are left with \( n_0 > 1 \). For every \( i \) with \( n_i > 1 \) and every \( 1 \leq j \leq n_i \), there is the point \( \overline{x}(i,j) \in \overline{X} \) with \( ij \)-coordinate \( T_{ij} \) equal to one and all others equal to zero, and thus we have the point \( x(i,j) := p(\overline{x}(i,j)) \in X \). Moreover, for every \( 1 \leq k \leq m \), we have the point \( \overline{x}(k) \in \overline{X} \) having the \( k \)-coordinate \( S_k \) equal to one and all others zero; we set \( x(k) := p(\overline{x}(k)) \). Proposition 2.5 provides the bounds

\[ w_{ij} = \text{deg } T_{ij} = |\text{Cl}(X, x(i,j))| \leq \mu \quad \text{for } n_i > 1, 1 \leq j \leq n_i, \]

\[ u_k = \text{deg } S_k = |\text{Cl}(X, x(k))| \leq \mu \quad \text{for } 1 \leq k \leq m. \]

Let \( 0 \leq s \leq r \) be the maximal number with \( n_s > 1 \). Then \( g_{s-2,s-1,s} \) is the last polynomial such that each of its three monomials depends on more than one variable. For any \( t \geq s \), we have the “cut ring”

\[ R_t := \mathbb{K}[T_{ij}; 0 \leq i \leq t, 1 \leq j \leq n_i][S_1, \ldots, S_m] / \langle g_{i,i+1,i+2}; 0 \leq i \leq t-2 \rangle \]

where the relations \( g_{i,i+1,i+2} \) depend on only three variables as soon as \( i > s \).
holds. For the degree $\gamma$ of the relations we have
\[
(r-1)\gamma = (t-1)\gamma + (r-t)\gamma
\]
\[
= (t-1)\gamma + l_{i+1,1}w_{i+1,1} + \ldots + l_{r1}w_{r1}
\]
\[
< \sum_{i=0}^{r} \sum_{j=1}^{n_i} w_{ij} + \sum_{i=1}^{m} u_i
\]
\[
= \sum_{i=0}^{r} \sum_{j=1}^{n_i} w_{ij} + w_{t+1,1} + \ldots + w_{r1} + \sum_{i=1}^{m} u_i.
\]
Since $l_{i1}w_{i1} > w_{i1}$ holds in particular for $t + 1 \leq i \leq r$, we derive from this the inequality
\[
\gamma < \frac{1}{t-1} \left( \sum_{i=0}^{r} \sum_{j=1}^{n_i} w_{ij} + \sum_{i=1}^{m} u_i \right).
\]

To obtain the bounds in Assertions (iii) and (iv), we consider the cut ring $R_t$ with $t = 2$ and apply Lemma 2.8; note that we have $d = n_0 + n_1 + n_2 + m - 2$ for the dimension $d = \dim(X)$ and that $l_{22} \geq 0$ is due to the fact that $X$ is non-toric. The bounds $w_{ij}, l_{ij} < 6d\mu$ in Assertion (iii) follow from $l_{i1}w_{i1} = \gamma < 6d\mu$ and $l_{11} < 2d\mu$ follows from $l_{11} | w_{21}$ for $3 \leq i \leq r$. Moreover, $l_{i1} | w_{i1}$ for $2 \leq i \leq r$ implies $l_{11} \cdots l_{r1} | \gamma = l_{11}w_{11}$. Similarly $w_{ij}, l_{ij} < 2(d+1)\mu$ in Assertion (iv) follow from $l_{ij}w_{ij} = \gamma < 2(d+1)\mu$ and $l_{21} \cdots l_{r1} | \gamma = l_{21}w_{21}$ follows from $l_{i1} | w_{21}$ for $3 \leq i \leq r$. The bounds on $r$ in (iii) in (iv) are as well consequences of the admissibility condition.

To obtain the bounds in Assertion (v), we consider the cut ring $R_t$ with $t = s$. Using $n_i = 1$ for $i \geq t + 1$, we can estimate the degree of the relation as follows:
\[
\gamma \leq \frac{(n_0 + \ldots + n_t + m)\mu}{t-1} = \frac{(d+t)\mu}{t-1} \leq (d+2)\mu.
\]
Since we have $w_{ij}, l_{ij} \leq \deg g_0$ for any $0 \leq i \leq r$ and any $1 \leq j \leq n_i$, we see that all $w_{ij}$ and $l_{ij}$ are bounded by $(d+2)\mu$. As before, $l_{s+1,1} \cdots l_{r1} | \gamma$ is a consequence of $l_{i1} | \gamma$ for $i = s + 2, \ldots, r$ and also the bound on $r$ follows from the admissibility condition.

Finally, we have to express the Picard index $\mu$ in terms of the weights $w_{ij}$ and $u_k$ as claimed in the Assertions. This is a direct application of the formula of Proposition 2.5. Observe that it suffices to work with the $p$-images of the following points: For every $0 \leq i \leq r$ with $n_i > 1$ take a point $\overline{\tau}(i,j) \in \overline{X}$ with $ij$-coordinate $T_{ij}$ equal to one and all others equal to zero, for every $0 \leq i \leq r$ with $n_i = 1$ whenever $n_i = 1$ take $\overline{\tau}(i,j) \in \overline{X}$ with $ij$-coordinate $T_{ij}$ equal to zero, all other $T_{st}$ equal to one and coordinates $S_k$ equal to zero, and, for every $1 \leq k \leq m$, take a point $\overline{\tau}(k) \in \overline{X}$ having the $k$-coordinate $S_k$ equal to one and all others zero.

\[\square\]
We conclude the section with discussing some aspects of the not necessarily Fano varieties of Proposition 2.7. Recall that we considered admissible triples $(A, n, L)$ with $n_0 = \ldots = n_r = 1$ and thus rings $R$ of the form $\mathbb{K}[T_0, \ldots , T_r, S_1, \ldots , S_m] / (\alpha_{i+1} t_{i+1} + \alpha_{i+2} t_{i+2} + \alpha_{i+3} t_{i+3}, 0 \leq i \leq r - 2)$.

**Proposition 2.9.** Suppose that the ring $R$ as above is the Cox ring of a non-toric variety $X$ with $\text{Cl}(X) = \mathbb{Z}$. Then we have $m \geq 1$ and $\mu := [\text{Cl}(X) : \text{Pic}(X)] \geq 30$. Moreover, if $X$ is a surface, then we have $m = 1$ and $w_i = l_i^{-1} l_0 \cdots l_r$.

**Proof.** The homogeneity condition $l_i w_i = l_j w_j$ together with the admissibility condition $\gcd(l_i, l_j) = 1$ for $0 \leq i \neq j \leq r$ gives us $l_i \mid \gcd(w_i; j \neq i)$. Moreover, by Proposition 1.11, every set of $m + r$ weights $w_i$ has to generate the class group $\mathbb{Z}$, so they must have greatest common divisor one. Since $X$ is non-toric, $l_i \geq 2$ holds and we obtain $m \geq 1$. To proceed, we infer $l_0 \cdots l_r \mid \mu$ and $l_0 \cdots l_r \mid \deg g_{ijk}$ from Proposition 2.5. As a consequence, the minimal value for $\mu$ and $\deg g_{ijk}$ is obviously $2 \cdot 3 \cdot 5 = 30$, what really can be received as the following example shows. Note that if $X$ is a surface we have $m = 1$ and $\gcd(w_i; 0 \leq i \leq r) = 1$. Thus, $l_i w_i = l_j w_j$ gives us $\deg g_{ijk} = l_0 \cdots l_r$ and $w_i = l_i^{-1} l_0 \cdots l_r$. \qed

The bound $[\text{Cl}(X) : \text{Pic}(X)] \geq 30$ given in the above proposition is even sharp; the surface discussed below realizes it.

**Example 2.10.** Consider $X$ with $R(X) = \mathbb{K}[T_0, T_1, T_2, T_3]/(g)$ with $g = T_0^2 + T_1^3 + T_2^5$ and the grading

$$\deg T_0 = 15, \quad \deg T_1 = 10, \quad \deg T_2 = 6, \quad \deg T_3 = 1.$$  

Then we have $\gcd(15, 10) = 5$, $\gcd(15, 6) = 3$ and $\gcd(10, 6) = 2$ and therefore $[\text{Cl}(X) : \text{Pic}(X)] = 30$. Further $X$ is Fano because of

$$\deg g = 30 < 32 = \deg T_0 + \ldots + \deg T_3.$$  

Let us have a look at the geometric meaning of the condition $n_0 = \ldots = n_r = 1$. For a variety $X$ with an action of a torus $T$, we denote by $X_0 \subseteq X$ the union of all orbits with at most finite isotropy. Then there is a possibly non-separated orbit space $X_0/T$; we call it the maximal orbit space. From [15], we infer that $n_0 = \ldots = n_r = 1$ holds if and only if $X_0/T$ is separated. Combining this with Propositions 2.7 and 2.9 gives the following.

**Corollary 2.11.** For any pair $(d, \mu) \in \mathbb{Z}_{>0}^2$ there is, up to deformation, only a finite number of $d$-dimensional complete varieties $X$ with a complexity one torus action having divisor class group $\mathbb{Z}$, Picard index $[\text{Cl}(X) : \text{Pic}(X)] = \mu$ and maximal orbit space $\mathbb{P}_1$ and for each of these varieties the complement $X \setminus X_0$ contains divisors.
Finally, we present a couple of examples showing that there are also non-Fano varieties with a complexity one torus action having divisor class group \( \mathbb{Z} \) and maximal orbit space \( \mathbb{P}_1 \).

**Example 2.12.** Consider \( X \) with \( \mathcal{R}(X) = \mathbb{K}[T_0, T_1, T_2, T_3]/(g) \) with \( g = T_0^2 + T_1^3 + T_2^7 \) and the grading

\[
\deg T_0 = 21, \quad \deg T_1 = 14, \quad \deg T_2 = 6, \quad \deg T_3 = 1.
\]

Then we have \( \gcd(21, 14) = 7, \gcd(21, 6) = 3 \) and \( \gcd(14, 6) = 2 \) and therefore \( [\text{Cl}(X) : \text{Pic}(X)] = 42 \). Moreover, \( X \) is not Fano, because its canonical class \( K_X \) is trivial

\[
K_X = \deg g - \deg T_0 - \ldots - \deg T_3 = 0.
\]

**Example 2.13.** Consider \( X \) with \( \mathcal{R}(X) = \mathbb{K}[T_0, T_1, T_2, T_3]/(g) \) with \( g = T_0^2 + T_1^3 + T_2^{11} \) and the grading

\[
\deg T_0 = 33, \quad \deg T_1 = 22, \quad \deg T_2 = 6, \quad \deg T_3 = 1.
\]

Then we have \( \gcd(22, 33) = 11, \gcd(33, 6) = 3 \) and \( \gcd(22, 6) = 2 \) and therefore \( [\text{Cl}(X) : \text{Pic}(X)] = 66 \). The canonical class \( K_X \) of \( X \) is even ample:

\[
K_X = \deg g - \deg T_0 - \ldots - \deg T_3 = 4.
\]

The following example shows that the Fano assumption is essential for the finiteness results in Theorem 2.1.

**Remark 2.14.** For any pair \( p, q \) of coprime positive integers, we obtain a locally factorial \( \mathbb{K}^* \)-surface \( X(p, q) \) with \( \text{Cl}(X) = \mathbb{Z} \) and Cox ring

\[
\mathcal{R}(X(p, q)) = \mathbb{K}[T_{01}, T_{02}, T_{11}, T_{21}]/(g), \quad g = T_0T_0^{pq-1}T_1^p + T_2^q;
\]

the \( \text{Cl}(X) \)-grading is given by \( \deg T_{01} = \deg T_{02} = 1, \deg T_{11} = p \) and \( \deg T_{21} = q \). Note that \( \deg g = pq \) holds and for \( p, q \geq 3 \), the canonical class \( K_X \) satisfies

\[
K_X = \deg g - \deg T_{01} - \deg T_{02} - \deg T_{11} - \deg T_{21} = pq - 2 - p - q \geq 0.
\]

### 3 Classification results

In this section, we give classification results for Fano varieties \( X \) with \( \text{Cl}(X) \cong \mathbb{Z} \) that come with a complexity one torus action; note that they are necessarily rational. The procedure to obtain classification lists for prescribed dimension \( d = \dim X \) and Picard index \( \mu = [\text{Cl}(X) : \text{Pic}(X)] \) is always the following. By Theorem 1.9, we know that their Cox rings are of the form \( \mathcal{R}(X) \cong \mathcal{R}(A, n, L)[S_1, \ldots, S_m] \) with admissible triples \( (A, n, L) \). Note that for the family \( A = (a_0, \ldots, a_r) \) of points \( a_i \in \mathbb{K}^2 \), we may assume

\[
a_0 = (1, 0), \quad a_1 = (1, 1), \quad a_2 = (0, 1).
\]

---

*Documenta Mathematica 16 (2011) 71–109*
The bounds on the input data of \((A,n,L)\) provided by Theorem 2.1 as well as the criteria of Propositions 1.11 and 2.5 allow us to generate all the possible Cox rings \(\mathcal{R}(X)\) of the Fano varieties \(X\) in question for fixed dimension \(d\) and Picard index \(\mu\). Note that \(X\) can be reconstructed from \(\mathcal{R}(X) = R(A,n,L)[S_1,\ldots,S_n]\) as the homogeneous spectrum with respect to the \(\text{Cl}(X)\)-grading. Thus \(X\) is classified by its Cox ring \(\mathcal{R}(X)\).

In the following tables, we present the Cox rings as \(\mathbb{K}[T_1,\ldots,T_s]\) modulo relations and fix the \(\mathbb{Z}\)-gradings by giving the weight vector \((w_1,\ldots,w_s)\), where \(w_i := \deg T_i\). The first classification result concerns surfaces.

**Theorem 3.1.** Let \(X\) be a non-toric Fano surface with an effective \(\mathbb{K}^*\)-action such that \(\text{Cl}(X) = \mathbb{Z}\) and \(\text{Cl}(X) : \text{Pic}(X) \leq 6\) hold. Then its Cox ring is precisely one of the following.

\[
\begin{array}{|c|c|c|c|}
\hline
\text{No.} & \mathcal{R}(X) & (w_1,\ldots,w_d) & (-K_X)^2 \\
\hline
1 & \mathbb{K}[T_1,\ldots,T_4]/(T_1^2 T_2^3 + T_3^3 + T_4^2) & (1,1,2,3) & 1 \\
\hline
2 & \mathbb{K}[T_1,\ldots,T_4]/(T_1^2 T_2 + T_3^3 + T_4^2) & (1,2,2,3) & 2 \\
\hline
3 & \mathbb{K}[T_1,\ldots,T_4]/(T_1^3 T_2 + T_3^3 + T_4^2) & (1,3,2,3) & 3 \\
\hline
4 & \mathbb{K}[T_1,\ldots,T_4]/(T_1 T_2^2 + T_3^3 + T_4^2) & (1,3,2,5) & 1/3 \\
\hline
5 & \mathbb{K}[T_1,\ldots,T_4]/(T_1 T_2^2 + T_3^3 + T_4^2) & (1,3,2,5) & 1/3 \\
\hline
6 & \mathbb{K}[T_1,\ldots,T_4]/(T_1^2 T_2 + T_3^3 + T_4^2) & (1,4,2,3) & 4 \\
\hline
7 & \mathbb{K}[T_1,\ldots,T_4]/(T_1^2 T_2 + T_3^3 + T_4^2) & (1,4,2,5) & 1 \\
\hline
\end{array}
\]
| No. | \( \mathbb{K}[T_1, \ldots , T_4]/(T_1T_2 + T_3^5 + T_4^5) \) | (1, 5, 2, 3) | 5 |
|-----|---------------------------------|--------------|-----|
| 9   | \( \mathbb{K}[T_1, \ldots , T_4]/(T_1^5T_2 + T_3^5 + T_4^5) \) | (1, 5, 2, 5) | 9/5 |
| 10  | \( \mathbb{K}[T_1, \ldots , T_4]/(T_1^5T_2 + T_3^5 + T_4^5) \) | (1, 5, 2, 7) | 1/5 |
| 11  | \( \mathbb{K}[T_1, \ldots , T_4]/(T_1^5T_2 + T_3^4 + T_4^5) \) | (1, 5, 3, 4) | 1/5 |

\[ \text{[Cl}(X) : \text{Pic}(X)] = 6 \]

| No. | \( \mathcal{R}(X) \) | \((w_1, \ldots , w_4)\) | \((-K_X)^2\) |
|-----|---------------------------|-------------------|--------------|
| 12  | \( \mathbb{K}[T_1, \ldots , T_4]/(T_1T_2 + T_3^5 + T_4^5) \) | (1, 6, 2, 5) | 8/3 |
| 13  | \( \mathbb{K}[T_1, \ldots , T_4]/(T_1^5T_2 + T_3^5 + T_4^5) \) | (1, 6, 2, 7) | 2/3 |
| 14  | \( \mathbb{K}[T_1, \ldots , T_4]/(T_1^5T_2 + T_3^3 + T_4^3) \) | (1, 6, 3, 4) | 2/3 |
| 15  | \( \mathbb{K}[T_1, \ldots , T_4]/(T_1^5T_2 + T_3^3 + T_4^5) \) | (1, 3, 4, 6) | 2/3 |

**Proof.** As mentioned, Theorems 1.9, 2.1 and Propositions 1.11, 2.5 produce a list of all Cox rings of surfaces with the prescribed data. Doing this computation, we obtain the list of the assertion. Note that none of the Cox rings listed is a polynomial ring and hence none of the resulting surfaces \( X \) is a toric variety.

To show that different members of the list are not isomorphic to each other, we use the following two facts. Firstly, observe that any two minimal systems of homogeneous generators of the Cox ring have (up to reordering) the same list of degrees, and thus the list of generator degrees is invariant under isomorphism (up to reordering). Secondly, by Construction 1.7, the exponents \( l_{ij} > 1 \) are precisely the orders of the non-trivial isotropy groups of one-codimensional orbits of the action of the torus \( T \) on \( X \). Using both principles and going through the list, we see that different members \( X \) cannot be \( T \)-equivariantly isomorphic to each other. Since all listed \( X \) are non-toric, the effective complexity one torus action on each \( X \) corresponds to a maximal torus in the linear algebraic group \( \text{Aut}(X) \). Any two maximal tori in the automorphism group are conjugate, and thus we can conclude that two members are isomorphic if and only if they are \( T \)-equivariantly isomorphic.

We remark that in [28, Section 4], log del Pezzo surfaces with an effective \( \mathbb{K} \)-action and Picard number 1 and Gorenstein index less than 4 were classified. The above list contains six such surfaces, namely no. 1-4, 6 and 8; these are exactly the ones where the maximal exponents of the monomials form a platonic triple, i.e., are of the form \((1, k, l), (2, 2, k), (2, 3, 3), (2, 3, 4)\) or \((2, 3, 5)\). The remaining ones, i.e., no. 5, 7, and 9-15 have non-log-terminal and thus
non-rational singularities; to check this one may compute the resolutions via resolution of the ambient weighted projective space as in [14, Ex. 7.5].

With the same scheme of proof as in the surface case, one establishes the following classification results on Fano threefolds.

Theorem 3.2. Let $X$ be a three-dimensional locally factorial non-toric Fano variety with an effective two torus action such that $\text{Cl}(X) = \mathbb{Z}$ holds. Then its Cox ring is precisely one of the following.

| No. | $\mathcal{R}(X)$ | $(w_1, \ldots, w_5)$ | $(-K_X)^3$ |
|-----|------------------|----------------------|-------------|
| 1   | $\mathbb{K}[T_1, \ldots, T_5]/\langle T_1 T_2^3 + T_3^3 + T_4^2 \rangle$ | $(1, 1, 2, 3, 1)$ | 8 |
| 2   | $\mathbb{K}[T_1, \ldots, T_5]/\langle T_1 T_2 T_3^4 + T_4^3 + T_5^2 \rangle$ | $(1, 1, 1, 2, 3)$ | 8 |
| 3   | $\mathbb{K}[T_1, \ldots, T_5]/\langle T_1 T_2^2 T_3^3 + T_4^3 + T_5^2 \rangle$ | $(1, 1, 1, 2, 3)$ | 8 |
| 4   | $\mathbb{K}[T_1, \ldots, T_5]/\langle T_1 T_2 + T_3 T_4 + T_5^2 \rangle$ | $(1, 1, 1, 1, 1)$ | 54 |
| 5   | $\mathbb{K}[T_1, \ldots, T_5]/\langle T_1 T_2^2 + T_3 T_4^2 + T_5^2 \rangle$ | $(1, 1, 1, 1, 1)$ | 24 |
| 6   | $\mathbb{K}[T_1, \ldots, T_5]/\langle T_1 T_2^3 + T_3 T_4^3 + T_5^2 \rangle$ | $(1, 1, 1, 1, 1)$ | 4 |
| 7   | $\mathbb{K}[T_1, \ldots, T_5]/\langle T_1 T_3^2 + T_3 T_4^3 + T_5^2 \rangle$ | $(1, 1, 1, 1, 2)$ | 16 |
| 8   | $\mathbb{K}[T_1, \ldots, T_5]/\langle T_1 T_2^3 + T_3 T_4^3 + T_5^2 \rangle$ | $(1, 1, 1, 1, 3)$ | 2 |
| 9   | $\mathbb{K}[T_1, \ldots, T_5]/\langle T_1 T_2^5 + T_3 T_4^4 + T_5^3 \rangle$ | $(1, 1, 1, 1, 3)$ | 2 |

The singular threefolds listed in this theorem are rational degenerations of smooth Fano threefolds from [18]. The (smooth) general Fano threefolds of the corresponding families are non-rational see [12] for no. 1-3, [8] for no. 5, [20] for no. 6, [30, 29] for no. 7 and [19] for no. 8-9. Even if one allows certain mild singularities, one still has non-rationality in some cases, see [13], [9, 25], [10], [6].

Theorem 3.3. Let $X$ be a three-dimensional non-toric Fano variety with an effective two torus action such that $\text{Cl}(X) = \mathbb{Z}$ and $[\text{Cl}(X) : \text{Pic}(X)] = 2$ hold. Then its Cox ring is precisely one of the following.

| No. | $\mathcal{R}(X)$ | $(w_1, \ldots, w_5)$ | $(-K_X)^3$ |
|-----|------------------|----------------------|-------------|
| 1   | $\mathbb{K}[T_1, \ldots, T_5]/\langle T_1 T_2^3 + T_3^3 + T_4^2 \rangle$ | $(1, 2, 2, 3, 1)$ | 27/2 |
| 2   | $\mathbb{K}[T_1, \ldots, T_5]/\langle T_1 T_2^3 + T_3^2 + T_4^2 \rangle$ | $(1, 2, 2, 5, 1)$ | 1/2 |
| 3   | $\mathbb{K}[T_1, \ldots, T_5]/\langle T_1 T_2^3 + T_3^3 + T_4^2 \rangle$ | $(1, 2, 2, 5, 1)$ | 1/2 |
| 4   | $\mathbb{K}[T_1, \ldots, T_5]/\langle T_1 T_2^3 + T_3^3 + T_4^2 \rangle$ | $(1, 2, 2, 3, 2)$ | 16 |
| $m$ | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1^m T_2^3 + T_3^3 + T_4^2 \rangle$ | $\text{dim}$ |
|-----|---------------------------------------------------|--------|
| 5   | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1^5 T_2^3 + T_3^3 + T_4^2 \rangle$ | 2 |
| 6   | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1^6 T_2^3 + T_3^3 + T_4^3 \rangle$ | 2 |
| 7   | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1^7 T_2^3 + T_3^3 + T_4^3 \rangle$ | 27/2 |
| 8   | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2^2 + T_3^3 + T_4^3 \rangle$ | 1/2 |
| 9   | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2^2 + T_3^3 + T_4^3 \rangle$ | 1/2 |
| 10  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 1/2 |
| 11  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 2 |
| 12  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 2 |
| 13  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 2 |
| 14  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 2 |
| 15  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 2 |
| 16  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 16 |
| 17  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 2 |
| 18  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 2 |
| 19  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 81/2 |
| 20  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 1/2 |
| 21  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 5/2 |
| 22  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 5/2 |
| 23  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 5/2 |
| 24  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 5/2 |
| 25  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 27 |
| 26  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 3/2 |
| 27  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 3/2 |
| 28  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 3/2 |
| 29  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 8 |
| 30  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 8 |
| 31  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 1 |
| 32  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 1 |
| 33  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 1 |
| 34  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 1 |
| 35  | $\mathbb{K}[T_1,\ldots,T_5]/\langle T_1 T_2 T_3^3 + T_4^3 + T_5^3 \rangle$ | 27 |
| $\aleph$ | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2^3 + T_2T_4^2 + T_5^2)$ | $\{1, 1, 2, 2, 1\}$ | $3/2$ |
|---|---|---|---|
| 37 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2^2 + T_3T_4 + T_5^2)$ | $\{1, 1, 2, 2, 2\}$ | $16$ |
| 38 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2^3 + T_3T_4^2 + T_5^2)$ | $\{1, 1, 2, 2, 2\}$ | $6$ |
| 39 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2^2 + T_3T_4^2 + T_5^2)$ | $\{1, 1, 2, 2, 2\}$ | $6$ |
| 40 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2^2 + T_3T_4^2 + T_5^2)$ | $\{1, 1, 2, 2, 2\}$ | $27/2$ |
| 41 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2^2 + T_3T_4^2 + T_5^2)$ | $\{1, 1, 2, 2, 2\}$ | $32$ |
| 42 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2^2 + T_3T_4^2 + T_5^2)$ | $\{1, 1, 2, 2, 2\}$ | $4$ |
| 43 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2^2 + T_3T_4^2 + T_5^2)$ | $\{1, 1, 2, 2, 4\}$ | $32$ |
| 44 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2^2 + T_3T_4 + T_5^2)$ | $\{1, 1, 2, 2, 5\}$ | $1/2$ |
| 45 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2^2 + T_3T_4 + T_5^2)$ | $\{1, 1, 2, 2, 5\}$ | $1/2$ |
| 46 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2^2 + T_3T_4 + T_5^2)$ | $\{1, 1, 2, 2, 5\}$ | $1/2$ |
| 47 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2^2 + T_3T_4 + T_5^2)$ | $\{1, 1, 2, 2, 5\}$ | $1/2$ |
| 48 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2 + T_3T_4 + T_5^2)$ | $\{1, 2, 1, 1\}$ | $48$ |
| 49 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2 + T_3T_4 + T_5^2)$ | $\{1, 2, 1, 1\}$ | $27$ |
| 50 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2 + T_3T_4 + T_5^2)$ | $\{1, 2, 1, 1\}$ | $10$ |
| 51 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2 + T_3T_4 + T_5^2)$ | $\{1, 2, 1, 1\}$ | $10$ |
| 52 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2 + T_3T_4 + T_5^2)$ | $\{1, 2, 1, 1\}$ | $10$ |
| 53 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2 + T_3T_4 + T_5^2)$ | $\{1, 2, 1, 1\}$ | $3/2$ |
| 54 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2 + T_3T_4 + T_5^2)$ | $\{1, 2, 1, 1\}$ | $32$ |
| 55 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2 + T_3T_4 + T_5^2)$ | $\{1, 2, 1, 1\}$ | $6$ |
| 56 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2 + T_3T_4 + T_5^2)$ | $\{1, 2, 1, 1\}$ | $6$ |
| 57 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2 + T_3T_4 + T_5^2)$ | $\{1, 2, 1, 2, 3\}$ | $27/2$ |
| 58 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2 + T_3T_4 + T_5^2)$ | $\{1, 2, 1, 2, 4\}$ | $4$ |
| 59 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2 + T_3T_4 + T_5^2)$ | $\{1, 2, 1, 2, 4\}$ | $4$ |
| 60 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2 + T_3T_4 + T_5^2)$ | $\{1, 2, 1, 2, 4\}$ | $4$ |
| 61 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2^3 + T_3T_4 + T_5^2)$ | $\{1, 2, 1, 2, 5\}$ | $1/2$ |
| 62 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2^3 + T_3T_4 + T_5^2)$ | $\{1, 2, 1, 2, 5\}$ | $1/2$ |
| 63 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2^3 + T_3T_4 + T_5^2)$ | $\{1, 2, 1, 2, 5\}$ | $1/2$ |
| 64 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2 + T_3T_4 + T_5^2)$ | $\{1, 2, 1, 2, 5\}$ | $32$ |
| 65 | $\mathbb{k}[T_1, \ldots, T_5]/(T_1T_2 + T_3T_4 + T_5^2)$ | $\{1, 2, 2, 2, 1\}$ | $6$ |
| No. | Expression                                                                 | 1st Degree | 2nd Degree | 3rd Degree | 4th Degree | 5th Degree | 6th Degree | 7th Degree |
|-----|---------------------------------------------------------------------------|------------|------------|------------|------------|------------|------------|------------|
| 68  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1^2T_2 + T_3^2T_4 + T_5^2)$             | 1, 2, 2, 2, 3 | 2          |            |            |            |            |            |
| 69  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1^2T_2^3 + T_3^4T_4 + T_5^2)$           | 1, 2, 2, 2, 3 | 2          |            |            |            |            |            |
| 70  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1^2T_2^3 + T_3^4T_4 + T_5^2)$           | 1, 2, 2, 2, 3 | 2          |            |            |            |            |            |
| 71  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1^2T_2^3 + T_3^4T_4 + T_5^2)$           | 1, 2, 2, 2, 3 | 2          |            |            |            |            |            |
| 72  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1^2T_2^3 + T_3^4T_4 + T_5^2)$           | 1, 2, 2, 2, 3 | 2          |            |            |            |            |            |
| 73  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 1, 4, 6 | 1, 2       |            |            |            |            |            |
| 74  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 1, 4, 6 | 1, 2       |            |            |            |            |            |
| 75  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 1, 4, 6 | 1, 2       |            |            |            |            |            |
| 76  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 1, 4, 6 | 1, 2       |            |            |            |            |            |
| 77  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 1, 4, 6 | 1, 2       |            |            |            |            |            |
| 78  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 1, 4, 6 | 1, 2       |            |            |            |            |            |
| 79  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 1, 4, 6 | 1, 2       |            |            |            |            |            |
| 80  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 1, 4, 6 | 1, 2       |            |            |            |            |            |
| 81  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 2, 2, 3 | 27/2       |            |            |            |            |            |
| 82  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 2, 2, 3 | 27/2       |            |            |            |            |            |
| 83  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 2, 2, 3 | 27/2       |            |            |            |            |            |
| 84  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 2, 2, 3 | 27/2       |            |            |            |            |            |
| 85  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 2, 2, 3 | 27/2       |            |            |            |            |            |
| 86  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 2, 2, 3 | 27/2       |            |            |            |            |            |
| 87  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 2, 2, 3 | 27/2       |            |            |            |            |            |
| 88  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 2, 2, 3 | 27/2       |            |            |            |            |            |
| 89  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 2, 2, 3 | 27/2       |            |            |            |            |            |
| 90  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 2, 2, 3 | 27/2       |            |            |            |            |            |
| 91  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 2, 2, 3 | 27/2       |            |            |            |            |            |
| 92  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 2, 2, 3 | 27/2       |            |            |            |            |            |
| 93  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 2, 2, 3 | 27/2       |            |            |            |            |            |
| 94  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 2, 2, 3 | 27/2       |            |            |            |            |            |
| 95  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 2, 2, 3 | 27/2       |            |            |            |            |            |
| 96  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 2, 2, 3 | 27/2       |            |            |            |            |            |
| 97  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 2, 2, 3 | 27/2       |            |            |            |            |            |
| 98  | $\mathbb{K}[T_1, \ldots, T_5]/(T_1T_2T_3^4 + T_4^3 + T_5^2)$             | 1, 1, 2, 2, 3 | 27/2       |            |            |            |            |            |
The varieties no. 2, 3 and 25, 26 are rational degenerations of quasi-smooth varieties from the list in [16]. In [11] the non-rationality of a general (quasi-smooth) element of the corresponding family was proved.

The varieties listed so far might suggest that we always obtain only one relation in the Cox ring. We discuss now some examples, showing that for a Picard index big enough, we need in general more than one relation, where this refers always to a presentation as in Theorem 1.9 (ii).

Example 3.4. A Fano $\mathbb{K}^*$-surface $X$ with $\text{Cl}(X) = \mathbb{Z}$ such that the Cox ring $\mathcal{R}(X)$ needs two relations. Consider the $\mathbb{Z}$-graded ring

$$R = \mathbb{K}[T_{01}, T_{02}, T_{11}, T_{21}, T_{31}] / \langle g_0, g_1 \rangle,$$

where the degrees of $T_{01}, T_{02}, T_{11}, T_{21}, T_{31}$ are 29, 1, 6, 10, 15, respectively, and the relations $g_0, g_1$ are given by

$$g_0 := T_{01}T_{02} + T_{11}^3 + T_{21}^3, \quad g_1 := \alpha_{23}T_{11}^5 + \alpha_{31}T_{21}^3 + \alpha_{12}T_{31}^2.$$

Then $R$ is the Cox ring of a Fano $\mathbb{K}^*$-surface. Note that the Picard index is given by $[\text{Cl}(X) : \text{Pic}(X)] = \text{lcm}(29, 1) = 29.$

Proposition 3.5. Let $X$ be a non-toric Fano surface with an effective $\mathbb{K}^*$-action such that $\text{Cl}(X) \cong \mathbb{Z}$ and $[\text{Cl}(X) : \text{Pic}(X)] < 29$ hold. Then the Cox ring of $X$ is of the form

$$\mathcal{R}(X) \cong \mathbb{K}[T_1, \ldots, T_5] / \langle T_1^{l_1} T_2^{l_2} + T_3^{l_3} + T_4^{l_4} \rangle.$$

Proof. The Cox ring $\mathcal{R}(X)$ is as in Theorem 1.9, and, in the notation used there, we have $n_0 + \ldots + n_r + m = 2 + r$. This leaves us with the possibilities $n_0 = m = 1$ and $n_0 = 2, m = 0$. In the first case, Proposition 2.9 tells us that the Picard index of $X$ is at least 30.
So, consider the case $n_0 = 2$ and $m = 0$. Then, according to Theorem 1.9, the Cox ring $R(X)$ is $\mathbb{K}[T_{01}, T_{02}, T_1, \ldots, T_r]$ divided by relations

$$g_{0,1,2} = T_{01}^{i_0} T_{02}^{i_1} + T_1^{i_2}, \quad g_{i_i+1} = \alpha_{i+1} T_i^{l_i} + \alpha_{i+2} T_{i+1}^{l_{i+1}} + \alpha_{i+3} T_{i+2}^{l_{i+2}},$$

where $1 \leq i \leq r - 2$. We have to show that $r = 2$ holds. Set $\mu := [\text{Cl}(X) : \text{Pic}(X)]$ and let $\gamma \in \mathbb{Z}$ denote the degree of the relations. Then we have $\gamma = w_i l_i$ for $1 \leq i \leq r$, where $w_i := \deg T_i$. With $w_{0i} := \deg T_{0i}$, Proposition 2.5 gives us

$$(r - 1) \gamma < w_{01} + w_{02} + w_1 + \ldots + w_r.$$ 

We claim that $w_{01}$ and $w_{02}$ are coprime. Otherwise they had a common prime divisor $p$. This $p$ divides $\gamma = l_i w_i$. Since $l_1, \ldots, l_r$ are pairwise coprime, $p$ divides at least $r - 1$ of the weights $w_1, \ldots, w_r$. This contradicts the Cox ring condition that any $r + 1$ of the $r + 2$ weights generate the class group $\mathbb{Z}$. Thus, $w_{01}$ and $w_{02}$ are coprime and we obtain

$$\mu \geq \text{lcm}(w_{01}, w_{02}) = w_{01} \cdot w_{02} \geq w_{01} + w_{02} - 1.$$ 

Now assume that $r \geq 3$ holds. Then we can conclude

$$2 \gamma < w_{01} + w_{02} + w_1 + w_2 + w_3 \leq \mu + 1 + \gamma \left( \frac{1}{l_1} + \frac{1}{l_2} + \frac{1}{l_3} \right)$$

Since the numbers $l_i$ are pairwise coprime, we obtain $l_1 \geq 5$, $l_2 \geq 3$ and $l_3 \geq 2$. Moreover, $l_i w_i = l_i w_j$ implies $l_i \mid w_j$ and hence $l_1 l_2 l_3 \mid \gamma$. Thus, we have $\gamma \geq 30$. Plugging this in the above inequality gives

$$\mu \geq \gamma \left( 2 - \frac{1}{l_1} - \frac{1}{l_2} - \frac{1}{l_3} \right) - 1 = 29.$$ 

The Fano assumption is essential in this result; if we omit it, then we may even construct locally factorial surfaces with a Cox ring that needs more then one relation.

**Example 3.6.** A locally factorial $\mathbb{K}^*$-surface $X$ with $\text{Cl}(X) = \mathbb{Z}$ such that the Cox ring $R(X)$ needs two relations. Consider the $\mathbb{Z}$-graded ring

$$R = \mathbb{K}[T_{01}, T_{02}, T_{11}, T_{21}, T_{31}] / (g_0, g_1),$$

where the degrees of $T_{01}, T_{02}, T_{11}, T_{21}, T_{31}$ are 1, 1, 6, 10, 15, respectively, and the relations $g_0, g_1$ are given by

$$g_0 := T_{01}^{3} T_{02}^{2} + T_{11}^{5} + T_{21}^{3}, \quad g_1 := \alpha_2 T_{11}^{5} + \alpha_3 T_{21}^{3} + \alpha_4 T_{31}^{2}.$$ 

Then $R$ is the Cox ring of a non Fano $\mathbb{K}^*$-surface $X$ of Picard index one, i.e, $X$ is locally factorial.
For non-toric Fano threefolds $X$ with an effective 2-torus action $\text{Cl}(X) \cong \mathbb{Z}$, the classifications 3.2 and 3.3 show that for Picard indices one and two we only obtain hypersurfaces as Cox rings. The following example shows that this stops at Picard index three.

**Example 3.7.** A Fano threefold $X$ with $\text{Cl}(X) = \mathbb{Z}$ and a 2-torus action such that the Cox ring $R = \mathbb{K}[T_{01}, T_{02}, T_{11}, T_{12}, T_{21}, T_{31}] / \langle g_0, g_1 \rangle$ needs two relations. Consider

$$R = \mathbb{K}[T_{01}, T_{02}, T_{11}, T_{12}, T_{21}, T_{31}] / \langle g_0, g_1 \rangle$$

where the degrees of $T_{01}, T_{02}, T_{11}, T_{12}, T_{21}, T_{31}$ are 1, 1, 3, 3, 2, 3, respectively, and the relations are given by

$$g_0 = T_{01}^2 T_{02}^2 + T_{11} T_{12}^2 + T_{31}^3, \quad g_1 = \alpha_{23} T_{11} T_{12} + \alpha_{31} T_{21}^2 + \alpha_{12} T_{31}^2.$$ 

Then $R$ is the Cox ring of a Fano threefold with a 2-torus action. Note that the Picard index is given by

$$[\text{Cl}(X) : \text{Pic}(X)] = \text{lcm}(1, 1, 3, 3) = 3.$$ 

Finally, we turn to locally factorial Fano fourfolds. Here we observe more than one relation in the Cox ring even in the locally factorial case.

**Theorem 3.8.** Let $X$ be a four-dimensional locally factorial non-toric Fano variety with an effective three torus action such that $\text{Cl}(X) = \mathbb{Z}$ holds. Then its Cox ring is precisely one of the following.

| No. | $\mathcal{R}(X)$ | $(w_1, \ldots, w_6)$ | $(-K_X)^4$ |
|-----|-----------------|---------------------|------------|
| 1   | $\mathbb{K}[T_1, \ldots, T_6]/(T_1 T_2^2 + T_3^2 + T_4^2)$ | $(1, 1, 2, 3, 1, 1)$ | 81         |
| 2   | $\mathbb{K}[T_1, \ldots, T_6]/(T_1 T_2^3 + T_3^2 + T_4^2)$ | $(1, 1, 2, 5, 1, 1)$ | 1          |
| 3   | $\mathbb{K}[T_1, \ldots, T_6]/(T_1 T_2 T_3^2 + T_3^2 + T_4^2)$ | $(1, 1, 2, 5, 1, 1)$ | 1          |
| 4   | $\mathbb{K}[T_1, \ldots, T_6]/(T_1 T_2 T_3 + T_4 + T_5^2)$ | $(1, 1, 1, 2, 3, 1)$ | 81         |
| 5   | $\mathbb{K}[T_1, \ldots, T_6]/(T_1 T_2 T_3^2 + T_3^2 + T_5^2)$ | $(1, 1, 1, 2, 5, 1)$ | 1          |
| 6   | $\mathbb{K}[T_1, \ldots, T_6]/(T_1 T_2 T_3^2 + T_4 + T_5^2)$ | $(1, 1, 1, 2, 5, 1)$ | 1          |
| 7   | $\mathbb{K}[T_1, \ldots, T_6]/(T_1 T_2 T_3^2 + T_4 + T_5^2)$ | $(1, 1, 1, 2, 5, 1)$ | 1          |
| 8   | $\mathbb{K}[T_1, \ldots, T_6]/(T_1 T_2 T_3^2 + T_4 + T_5^2)$ | $(1, 1, 1, 2, 5, 1)$ | 1          |
| 9   | $\mathbb{K}[T_1, \ldots, T_6]/(T_1 T_2 T_3^2 + T_4 + T_5^2)$ | $(1, 1, 1, 2, 5, 1)$ | 1          |
| 10  | $\mathbb{K}[T_1, \ldots, T_6]/(T_1 T_2 T_3^2 + T_4 + T_5^2)$ | $(1, 1, 1, 2, 5, 1)$ | 1          |
| 11  | $\mathbb{K}[T_1, \ldots, T_6]/(T_1 T_2 T_3^2 + T_4 + T_5^2)$ | $(1, 1, 1, 2, 5, 1)$ | 1          |
| 12  | $\mathbb{K}[T_1, \ldots, T_6]/(T_1 T_2 + T_3 T_4 + T_5^2)$ | $(1, 1, 1, 1, 1, 1)$ | 512        |
| 13  | $\mathbb{K}[T_1, \ldots, T_6]/(T_1 T_2^2 + T_3 T_4^2 + T_5^3)$ | $(1, 1, 1, 1, 1, 1)$ | 243        |

*Documenta Mathematica 16 (2011) 71–109*
| 14 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,1,1) | 64 |
| 15 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,1,1) | 5 |
| 16 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,1,1) | 5 |
| 17 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1^3T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,1,1) | 5 |
| 18 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,1) | 162 |
| 19 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,1) | 3 |
| 20 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,1) | 3 |
| 21 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,1,3) | 32 |
| 22 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,1,3) | 32 |
| 23 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,4,1) | 2 |
| 24 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,4,1) | 2 |
| 25 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1^3T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,4,1) | 2 |
| 26 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,3) | 81 |
| 27 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,3) | 81 |
| 28 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,5) | 1 |
| 29 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,5) | 1 |
| 30 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,5) | 1 |
| 31 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,5) | 1 |
| 32 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,5) | 1 |
| 33 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,5) | 1 |
| 34 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,5) | 1 |
| 35 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,5) | 1 |
| 36 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,5) | 1 |
| 37 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,5) | 1 |
| 38 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,5) | 1 |
| 39 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,5) | 1 |
| 40 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,5) | 1 |
| 41 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,5) | 1 |
| 42 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,5) | 1 |
| 43 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,5) | 1 |
| 44 | $\mathbb{K}[T_1, \ldots, T_6]/(T_1T_2^3 + T_3^3T_4^3 + T_5^3)\rangle$ | (1,1,1,1,2,5) | 1 |
By the result of [26], the singular quintics of this list are rational degenerations of smooth non-rational Fano fourfolds.
4 Geometry of the locally factorial threefolds

In this section, we take a closer look at the (factorial) singularities of the Fano varieties $X$ listed in Theorem 3.2. Recall that the discrepancies of a resolution $\varphi: \tilde{X} \to X$ of a singularity are the coefficients of $K_{\tilde{X}} - \varphi^*K_X$, where $K_X$ and $K_{\tilde{X}}$ are canonical divisors such that $K_{\tilde{X}} - \varphi^*K_X$ is supported on the exceptional locus of $\varphi$. A resolution is called crepant, if its discrepancies vanish and a singularity is called canonical (terminal), if it admits a resolution with nonnegative (positive) discrepancies. By a relative minimal model we mean a projective morphism $\tilde{X} \to X$ such that $\tilde{X}$ has at most terminal singularities and its relative canonical divisor is relatively nef.

Theorem 4.1. For the nine 3-dimensional Fano varieties listed in Theorem 3.2, we have the following statements.

(i) No. 4 is a smooth quadric in $\mathbb{P}^4$.

(ii) Nos. 1, 3, 5, 7 and 9 are singular with only canonical singularities and all admit a crepant resolution.

(iii) Nos. 6 and 8 are singular with non-canonical singularities but admit a smooth relative minimal model.

(iv) No. 2 is singular with only canonical singularities, one of them of type $cA_1$, and admits only a singular relative minimal model.

The Cox ring of the relative minimal model $\tilde{X}$ as well as the the Fano degree of $X$ itself are given in the following table.

| No. | $\mathcal{R}(\tilde{X})$ | $(-K_X)^3$ |
|-----|--------------------------|-------------|
| 1   | $\mathbb{K}[T_1, \ldots, T_{14}] / (T_1 T_2 T_3^2 T_4^3 T_5 T_6 + T_7 T_8 T_9 + T_{10} T_{11})$ | 8 |
| 2   | $\mathbb{K}[T_1, \ldots, T_6] / (T_1 T_2 T_3^2 T_4 + T_5 T_6 T_7 + T_8)$ | 8 |
| 3   | $\mathbb{K}[T_1, \ldots, T_8] / (T_1 T_2^2 T_3^3 + T_4 T_5^3 + T_6 T_7^2)$ | 8 |
| 4   | $\mathbb{K}[T_1, \ldots, T_9] / (T_1 T_2 + T_3 T_4 + T_5^2)$ | 54 |
| 5   | $\mathbb{K}[T_1, \ldots, T_9] / (T_1 T_2^2 + T_3 T_4^2 + T_5^3)$ | 24 |
| 6   | $\mathbb{K}[T_1, \ldots, T_9] / (T_1 T_2^3 + T_3 T_4^3 + T_5^4)$ | 4 |
| 7   | $\mathbb{K}[T_1, \ldots, T_9] / (T_1 T_2^3 + T_3 T_4^3 + T_5^2 T_6)$ | 16 |
| 8   | $\mathbb{K}[T_1, \ldots, T_9] / (T_1 T_2^3 + T_3 T_4^3 + T_5^2 T_6)$ | 2 |
| 9   | $\mathbb{K}[T_1, \ldots, T_{46}] / \left( T_1 T_2 T_3 T_4 T_5^2 T_6^2 T_7^3 T_8^4 T_9^5 + T_{10} + \ldots \right)$ | 2 |
For the proof, it is convenient to work in the language of polyhedral divisors introduced in [1] and [2]. As we are interested in rational varieties with a complexity one torus action, we only have to consider polyhedral divisors on the projective line \( Y = \mathbb{P}^1 \). This considerably simplifies the general definitions and allows us to give a short summary. In the sequel, \( N \cong \mathbb{Z}^n \) denotes a lattice and \( M = \text{Hom}(N, \mathbb{Z}) \) its dual. For the associated rational vector spaces we write \( N_\mathbb{Q} \) and \( M_\mathbb{Q} \). A polyhedral divisor on the projective line \( Y := \mathbb{P}^1 \) is a formal sum

\[
D = \sum_{y \in Y} D_y \cdot y,
\]

where the coefficients \( D_y \subseteq N_\mathbb{Q} \) are (possibly empty) convex polyhedra all sharing the same tail (i.e., recession) cone \( D_Y = \sigma \subseteq N_\mathbb{Q} \), and only finitely many \( D_y \) differ from \( \sigma \). The locus of \( D \) is the open subset \( Y(D) \subseteq Y \) obtained by removing all points \( y \subseteq Y \) with \( D_y = \emptyset \). For every \( u \in \sigma^\vee \cap M \) we have the evaluation

\[
D(u) := \sum_{y \in Y} \min_{v \in D_y} \langle u, v \rangle \cdot y,
\]

which is a usual rational divisor on \( Y(D) \). We call the polyhedral divisor \( D \) on \( Y \) proper if \( \text{deg} \, D \subseteq \sigma \) holds, where the polyhedral degree is defined by

\[
\text{deg} \, D := \sum_{y \in Y} D_y.
\]

Every proper polyhedral divisor \( D \) on \( Y \) defines a normal affine variety \( X(D) \) of dimension \( \text{rk}(N) + 1 \) coming with an effective action of the torus \( T = \text{Spec} \mathbb{K}[M] \): set \( X(D) := \text{Spec} A(D) \), where

\[
A(D) := \bigoplus_{u \in \sigma^\vee \cap M} \Gamma(Y(D), \mathcal{O}(D(u))) \subseteq \bigoplus_{u \in M} \mathbb{K}(Y) \cdot \chi^u.
\]

A divisorial fan, is a finite set \( \Xi \) of polyhedral divisors \( D \) on \( Y \), all having their polyhedral coefficients \( D_y \) in the same \( N_\mathbb{Q} \) and fulfilling certain compatibility conditions, see [2]. In particular, for every point \( y \in Y \), the slice

\[
\Xi_y := \{ D_y; D \in \Xi \}
\]

must be a polyhedral subdivision. The tail fan is the set \( \Xi_Y \) of the tail cones \( D_Y \) of the \( D \in \Xi \); it is a fan in the usual sense. Given a divisorial fan \( \Xi \), the affine varieties \( X(D) \), where \( D \in \Xi \), glue equivariantly together to a normal variety \( X(\Xi) \), and we obtain every rational normal variety with a complexity one torus action this way.

Smoothness of \( X = X(\Xi) \) is checked locally. For a proper polyhedral divisor \( D \) on \( Y \), we infer the following from [28, Theorem 3.3]. If \( Y(D) \) is affine,
then $X(D)$ is smooth if and only if $\text{cone}(\{1\} \times D_y) \subseteq \mathbb{Q} \times N_{\mathbb{Q}}$, the convex, polyhedral cone generated by $\{1\} \times D_y$, is regular for every $y \in Y(D)$. If $Y(D) = Y$ holds, then $X(D)$ is smooth if and only if there are $y, z \in Y$ such that $D = D_y + D_z$ holds and $\text{cone}(\{1\} \times D_y) + \text{cone}(\{-1\} \times D_z)$ is a regular cone in $\mathbb{Q} \times N_{\mathbb{Q}}$. Similarly to toric geometry, singularities of $X(D)$ are resolved by means of subdividing $D$. This means to consider divisorial fans $\Xi$ such that for any $y \in Y$, the slice $\Xi_y$ is a subdivision of $D_y$. Such a $\Xi$ defines a dominant morphism $X(\Xi) \rightarrow X(D)$ and a slight generalization of [2, Thm. 7.5.] yields that this morphism is proper.

**Proposition 4.2.** The 3-dimensional Fano varieties No. 1-8 listed in Theorem 3.2 and their relative minimal models arise from divisorial fans having the following slices and tail cones.
The above table should be interpreted as follows. The first three pictures in each row are the slices at 0, 1 and \( \infty \) and the last one is the tail fan. The divisorial fan of the fano variety itself is given by the solid polyhedra in the pictures. Here, all polyhedra of the same gray scale belong to the same polyhedral divisor. The subdivisions for the relative minimal models are sketched with dashed lines. In general, polyhedra with the same tail cone belong all to a unique polyhedral divisor with complete locus. For the white cones inside the tail fan we have another rule: for every polyhedron \( \Delta \in \Xi_y \) with the given white cone as its tail there is a polyhedral divisor \( \Delta \cdot y + \emptyset \cdot z \in \Xi \), with \( z \in \{0, 1, \infty\} \setminus \{y\} \). Here, different choices of \( z \) lead to isomorphic varieties, only the affine covering given by the \( X(D) \) changes.

In order to prove Theorem 4.1, we also have to understand invariant divisors on \( X = X(\Xi) \) in terms of \( \Xi \), see [15, Prop. 4.11 and 4.12] for details. A first type of invariant prime divisors, is in bijection \( D_{y,v} \leftrightarrow (y,v) \) with the vertices \( (y,v) \), where \( y \in Y \) and \( v \in \Xi_y \) is of dimension zero. The order of the generic isotropy
group along $D_{y,v}$ equals the minimal positive integer $\mu(v)$ with $\mu(v)v \in N$. A second type of invariant prime divisors, is in $D_\varrho \leftrightarrow \varrho$ with the extremal rays $\varrho \in \Xi_Y$, where a ray $\varrho \in \Xi_Y$ is called extremal if there is a $D \in \Xi$ such that $\varrho \subseteq D \land \deg D \cap \varrho = 0$ holds. The set of extremal rays is denoted by $\Xi_Y^\times$.

The divisor of a semi-invariant function $f \cdot \chi^u \in \mathbb{K}(X)$ is then given by

$$\text{div}(f \cdot \chi^u) = -\sum_{y \in Y} \sum_{v \in \Xi_Y^{(0)}} \mu(v) \cdot ((v,u) + \text{ord}_y f) \cdot D_{y,v} - \sum_{\varrho \in \Xi_Y^\times} \langle n_\varrho, u \rangle \cdot D_\varrho.$$

Next we describe the canonical divisor. Choose a point $y_0 \in Y$ such that $\Xi_{y_0} = \Xi_Y$ holds. Then a canonical divisor on $X = X(\Xi)$ is given by

$$K_X = (s-2) \cdot y_0 - \sum_{\Xi_y \neq \Xi_Y} \sum_{v \in \Xi_Y^{(0)}} D_{y,v} - \sum_{\varrho \in \Xi_Y^\times} E_\varrho.$$

**Proposition 4.3.** Let $D$ be a proper polyhedral divisor with $Y(D) = \mathbb{P}_1$, let $\Xi$ be a refinement of $D$ and denote by $y_1, \ldots, y_s \in Y$ the points with $\Xi_{y_i} \neq \Xi_Y$. Then the associated morphism $\varphi: X(\Xi) \to X(D)$ satisfies the following.

(i) The prime divisors in the exceptional locus of $\varphi$ are the divisors $D_{y_i,v}$ and $D_\varrho$ corresponding to $v \in \Xi_{y_i}^{(0)} \setminus D_{y_i}^{(0)} \setminus \Xi_{y_i}^{\varrho} \setminus D_{y_i}^{\varrho}$ respectively.

(ii) The discrepancies along the prime divisors $D_{y_i,v}$ and $D_\varrho$ of (i) are computed as

$$d_{y_i,v} = -\mu(v) \cdot (\langle v, u \rangle + \alpha_y) - 1, \quad d_\varrho = -\langle v_\varrho, u \rangle - 1,$$

where the numbers $\alpha_i$ are determined by

$$\begin{pmatrix}
-1 & -1 & \cdots & -1 & 0 \\
\mu(v_1^1) & 0 & \cdots & 0 & \mu(v_1^1)v_1^1 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\mu(v_r^1) & 0 & \cdots & 0 & \mu(v_r^1)v_r^1 \\
0 & 0 & \cdots & \mu(v_1^s) & \mu(v_1^s)v_1^s \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & \mu(v_r^s) & \mu(v_r^s)v_r^s \\
0 & 0 & \cdots & 0 & n_{y_1} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & n_{y_s}
\end{pmatrix} \begin{pmatrix}
\alpha_{y_1} \\
\vdots \\
\alpha_{y_s} \\
u
\end{pmatrix} = \begin{pmatrix}
2-s \\
1 \\
1 \\
1
\end{pmatrix}.$$
Then our formulæ for \( \operatorname{div}(f \cdot \chi_u) \) and \( K_X \) provide a row for every vertex \( v_i \in \Xi_y, \) \( i = 0, \ldots, s, \) and for every extremal ray \( q_i \in \Xi^\times, \) and \( \ell^{-1}(\alpha, u) \) is the (unique) solution of the above system.

Note, that in the above Proposition, the variety \( X(D) \) is \( \mathbb{Q} \)-Gorenstein if and only if the linear system of equations has a solution.

**Proof of Theorem 4.1 and Proposition 4.2.** We exemplarily discuss variety number eight. Recall that its Cox ring is given as

\[
\mathcal{R}(X) = K[T_1, \ldots, T_5]/(T_1 T_2^5 + T_3 T_4^5 + T_5^2)
\]

with the degrees \( 1, 1, 1, 1, 3. \) In particular, \( X \) is a hypersurface of degree 6 in \( \mathbb{P}(1, 1, 1, 1, 3) \), and the self-intersection of the anti-canonical divisor can be calculated as

\[
(-K_X^3) = 6 \cdot \frac{(1 + 1 + 1 + 1 + 3 - 6)^3}{1 \cdot 1 \cdot 1 \cdot 1 \cdot 3} = 2.
\]

The embedding \( X \subseteq \mathbb{P}(1, 1, 1, 1, 3) \) is equivariant, and thus we can use the technique described in [1, Sec. 11] to calculate a divisorial fan \( \Xi \) for \( X \). The result is the following divisorial fan; we draw its slices and indicate the polyhedral divisors with affine locus by colouring their tail cones \( D_{\mathbb{Y}} \in \Xi_{\mathbb{Y}} \) white:

One may also use [15, Cor. 4.9.] to verify that \( \Xi \) is the right divisorial fan: it computes the Cox ring in terms of \( \Xi \), and, indeed, we obtain again \( \mathcal{R}(X) \). Now we subdivide and obtain a divisorial fan having the refined slices as indicated in the following picture.

Here, the white ray \( Q_{\geq 0} \cdot (1, 0) \) indicates that the polyhedral divisors with that tail have affine loci. According to [15, Cor. 4.9.], the corresponding Cox ring is given by

\[
\mathcal{R}(\tilde{X}) = K[T_1, \ldots, T_7]/(T_1 T_2^5 + T_3 T_4^5 + T_5^2 T_6).
\]
We have to check that $\tilde{X}$ is smooth. Let us do this explicitly for the affine chart defined by the polyhedral divisor $\mathcal{D}$ with tail cone $\mathcal{D}_Y = \text{cone}((1, 2), (3, 1))$. Then $\mathcal{D}$ is given by
\[
\mathcal{D} = \left(\left(\frac{3}{5}, \frac{1}{5}\right) + \sigma\right) \cdot \{0\} + \left(\left[-\frac{1}{2}, 0\right] \times 0 + \sigma\right) \cdot \{\infty\}.
\]
Thus, $\text{cone}(\{1\} \times \mathcal{D}_0) + \text{cone}(\{-1\} \times \mathcal{D}_\infty)$ is generated by $(5, 3, 1)$, $(-2, -1, 0)$ and $(-1, 0, 0)$; in particular, it is a regular cone. This implies smoothness of the affine chart $X(\mathcal{D})$. Furthermore, we look at the affine charts defined by the polyhedral divisors $\mathcal{D}$ with tail cone $\mathcal{D}_Y = \text{cone}(1, 0)$. Since they have affine locus, we have to check $\text{cone}(\{1\} \times \mathcal{D}_y)$, where $y \in Y$. For $y \neq 0, 1$, we have $\mathcal{D}_y = \mathcal{D}_Y$. In this case, $\text{cone}(\{1\} \times \mathcal{D}_y)$ is generated by $(1, 1, 0)$, $(0, 1, 0)$ and is regular. For $y = 0$, we obtain that $\text{cone}(\{1\} \times \mathcal{D}_y)$ is generated by $(5, 3, 1)$, $(1, 0, 0)$, $(0, 1, 0)$ and this is regular. For $y = 1$ we get the same result. Hence, the polyhedral divisors with tail cone $\mathcal{D}_y = \text{cone}(1, 0)$ give rise to smooth affine charts.

Now we compute the discrepancies according to Proposition 4.3. The resolution has two exceptional divisors $\mathcal{D}_{\infty, 0}$ and $E_{(1, 0)}$. We work in the chart defined by the divisor $\mathcal{D} \in \Xi$ with tail cone $\mathcal{D}_Y = \text{cone}((1, 2), (1, 0))$. The resulting system of linear equations and its unique solution are given by
\[
\begin{pmatrix}
-1 & -1 & -1 & 0 & 0 & 1
5 & 0 & 0 & 3 & 1 & 1
0 & 1 & 0 & 0 & 0 & 1
0 & 5 & 0 & 0 & -1 & 1
0 & 0 & 2 & -1 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
\alpha_0 \\
\alpha_1 \\
\alpha_\infty \\
u
\end{pmatrix}
= \begin{pmatrix}
0 \\
1 \\
0 \\
-1 \\
4
\end{pmatrix}.
\]

The formula for the discrepancies yields $d_{\infty, 0} = -1$ and $d_{(1, 0)} = -2$. In particular, $X$ has non-canonical singularities. By a criterion from [24, Sec. 3.4.], we know that $D_{\infty, 0} + 2 \cdot E_{(1, 0)}$ is a nef divisor. It follows that $\tilde{X}$ is a minimal model over $X$. 

References

[1] K. Altmann, J. Hausen: Polyhedral divisors and algebraic torus actions. Math. Ann. 334 (2006), no. 3, 557–607.
[2] K. Altmann, J. Hausen, H. Süß: Gluing Affine Torus Actions Via Divisorsial Fans. Transformation Groups 13 (2008), no. 2, 215–242.
[3] V.V. Batyrev: Toric Fano threefolds. Izv. Akad. Nauk SSSR Ser. Mat. 45 (1981), no. 4, 704–717, 927.
[4] V.V. Batyrev: On the classification of toric Fano 4-folds. J. Math. Sci. New York 94, 1021–1050 (1999).
[5] F. Berchtold, J. Hausen: Homogeneous coordinates for algebraic varieties. J. Algebra 266 (2003), no. 2, 636–670.

[6] Cheltsov, Ivan; Park, Jihun Sextic double solids. In: Bogomolov, Fedor (ed.) et al.: Cohomological and geometric approaches to rationality problems. New Perspectives. Boston, MA: Birkhäuser. Progress in Mathematics 282, 75–132 (2010).

[7] J.J. Chen, J.A. Chen, M. Chen: On quasismooth weighted complete intersections. Preprint, arXiv:0908.1439.

[8] C.H. Clemens, P.A. Griffiths: The intermediate Jacobian of the cubic threefold. Ann. Math. (2) 95 (1972), 281–356.

[9] A. Corti: Singularities of linear systems and 3-fold birational geometry. Explicit birational geometry of 3-folds, 259–312, London Math. Soc. Lecture Note Ser., 281, Cambridge Univ. Press, Cambridge, 2000.

[10] A. Corti, M. Mella: Birational geometry of terminal quartic 3-folds I. Am. J. Math. 126 (2004), No. 4, 739–761.

[11] A. Corti, A. Pukhlikov, M. Reid: Fano 3-fold hypersurfaces. Explicit birational geometry of 3-folds, 175–258, London Math. Soc. Lecture Note Ser., 281, Cambridge Univ. Press, Cambridge, 2000.

[12] M.M. Grinenko: Mori structures on a Fano threefold of index 2 and degree 1. Tr. Mat. Inst. Steklova 246 (2004), Algebr. Geom. Metody, Svyazi i Prilozhi, 116–141; translation in Proc. Steklov Inst. Math. 2004, no. 3 (246), 103–128.

[13] M.M. Grinenko: Birational automorphisms of a three-dimensional double cone. Mat. Sb. 189 (1998), no. 7, 37–52; translation in Sb. Math. 189 (1998), no. 7-8, 991–1007.

[14] J. Hausen: Cox rings and combinatorics II. Mosc. Math. J., 8 (2008), 711–757.

[15] J. Hausen, H. Süss: The Cox ring of an algebraic variety with torus action. Adv. Math. 225 (2010), no. 2, 977–1012.

[16] A.R. Iano-Fletcher: Working with weighted complete intersections. Explicit birational geometry of 3-folds, 101–173, London Math. Soc. Lecture Note Ser., 281, Cambridge Univ. Press, Cambridge, 2000.

[17] M.-N. Ishida: Graded factorial rings of dimension 3 of a restricted type. J. Math. Kyoto Univ. 17 (1977), no. 3, 441–456.

[18] V.A. Iskovskih: Fano threefolds. II. Izv. Akad. Nauk SSSR Ser. Mat. 42 (1978), no. 3, 506–549.

[19] V.A. Iskovskih: Birational automorphisms of three-dimensional algebraic varieties. Current problems in mathematics, Vol. 12 (Russian), pp. 159–236, 239 (loose errata), VINITI, Moscow, 1979.
[20] V.A. Iskovskih, Yu.I. Manin: Three-dimensional quartics and counterexamples to the Lüroth problem. Mat. Sb. (N.S.) 86 (128) (1971), 140–166.

[21] J.M. Johnson, J. Kollár: Fano hypersurfaces in weighted projective 4-spaces. Experiment. Math. 10 (2001), no. 1, 151–158.

[22] F. Knop, H Kraft, T. Vust: The Picard group of a $G$-variety. In: Algebraic Transformation Groups and Invariant Theory, DMV Seminar, Band 3, Birkhäuser.

[23] S. Mori: Graded factorial domains. Japan J. Math. 3 (1977), no. 2, 223–238.

[24] L. Petersen, H. Süß: Torus invariant divisors. Preprint, arXiv:0811.0517 (2008), to appear in Israel J. Math.

[25] A.V. Pukhlikov: Birational automorphisms of a three-dimensional quartic with a simple singularity. Mat. Sb. (N.S.) 135(177) (1988), no. 4, 472–496, 559; translation in Math. USSR-Sb. 63 (1989), no. 2, 457–482.

[26] A.V. Pukhlikov: Birational isomorphisms of four-dimensional quintics. Invent. Math. 87, no. 2 (1987), 303–329.

[27] K. Watanabe, M. Watanabe: The classification of Fano 3-folds with torus embeddings. Tokyo J. Math. 5, 37–48 (1982).

[28] H. Süß: Canonical divisors on $T$-varieties. Preprint, arXiv:0811.0626v1.

[29] A.S. Tikhomirov: The intermediate Jacobian of double $\mathbb{P}^3$ that is branched in a quartic. Izv. Akad. Nauk SSSR Ser. Mat. 44 (1980), no. 6, 1329–1377, 1439.

[30] Voisin, Claire: Sur la jacobienne intermédiaire du double solide d’indice deux. Duke Math. J. 57 (1988), no. 2, 629–646.

Jürgen Hausen  
Mathematisches Institut  
Universität Tübingen  
Auf der Morgenstelle 10  
72076 Tübingen  
Germany

Elaine Herppich  
Mathematisches Institut  
Universität Tübingen  
Auf der Morgenstelle 10  
72076 Tübingen  
Germany

Hendrik Süß  
Institut für Mathematik  
LS Algebra und Geometrie  
BTU Cottbus  
PF 10 13 44  
03013 Cottbus  
Germany
ON THE STRUCTURE OF COVERS OF SOFIChifts

RUNE JOHANSEN

Received: June 6, 2010
Revised: March 1, 2011

Communicated by Joachim Cuntz

ABSTRACT. A canonical cover generalizing the left Fischer cover to arbitrary sofic shifts is introduced and used to prove that the left Krieger cover and the past set cover of a sofic shift can be divided into natural layers. These results are used to find the range of a flow-invariant and to investigate the ideal structure of the universal C∗-algebra associated to a sofic shift space.

2010 Mathematics Subject Classification: Primary: 37B10; Secondary: 46L05
Keywords and Phrases: shift space, sofic shift, Krieger cover, Fischer cover

1 INTRODUCTION

Shifts of finite type have been completely classified up to flow equivalence by Boyle and Huang [4, 6, 15], but very little is known about the classification of the class of sofic shift spaces introduced by Weiss [38], even though they are a natural first generalization of shifts of finite type. The purpose of this paper is to investigate the structure of - and relationships between - various standard presentations (the Fischer cover, the Krieger cover, and the past set cover) of sofic shift spaces. These results are used to find the range of the flow-invariant introduced in [1], and to investigate the ideal structure of the C∗-algebras associated to sofic shifts. In this way, the present paper can be seen as a continuation of the strategy applied in [10, 11, 30], where invariants for shift spaces are extracted from the associated C∗-algebras.

Section 2 recalls the definitions of shift spaces, labelled graphs, and covers to make the paper self contained. Section 3 introduces a canonical and flow-invariant cover generalizing the left Fischer cover to arbitrary sofic shifts.
Section 4 introduces the concept of a foundation of a cover, which is used to prove that the left Krieger cover and the past set cover can be divided into natural layers and to show that the left Krieger cover of an arbitrary sofic shift can be identified with a subgraph of the past set cover.

In Section 5, the structure of the layers of the left Krieger cover of an irreducible sofic shift is used to find the range of the flow-invariant introduced in [1]. Section 6 uses the results about the structure of covers of sofic shifts to investigate ideal lattices of the associated $C^*$-algebras. Additionally, it is proved that Condition (\( \ast \)) introduced by Carlsen and Matsumoto [12] holds if and only if the left Krieger cover is the maximal essential subgraph of the past set cover.

Acknowledgements. This work was supported by the Danish National Research Foundation (DNRF) through the Centre for Symmetry and Deformation. The author would like to thank David Pask, Toke Meier Carlsen, and Søren Eilers for interesting discussions and helpful comments. The author would also like to thank the anonymous referee for useful comments improving the exposition and to thank the University of Wollongong, Australia and the University of Tokyo, Japan where parts of the research for this paper were carried out during visits funded by Rejselegat for Matematikere.

2 Background

Shift spaces. Here, a short introduction to the definition and properties of shift spaces is given to make the present paper self-contained; for a thorough treatment of shift spaces see [21]. Let $\mathcal{A}$ be a finite set with the discrete topology. The full shift over $\mathcal{A}$ consists of the space $\mathcal{A}^\mathbb{Z}$ endowed with the product topology and the shift map $\sigma: \mathcal{A}^\mathbb{Z} \to \mathcal{A}^\mathbb{Z}$ defined by $\sigma(x)_i = x_{i+1}$ for all $i \in \mathbb{Z}$. Let $\mathcal{A}^*$ be the collection of finite words (also known as blocks) over $\mathcal{A}$. A subset $X \subseteq \mathcal{A}^\mathbb{Z}$ is called a shift space if it is invariant under the shift map and closed. For each $F \subseteq \mathcal{A}^*$, define $X_F$ to be the set of bi-infinite sequences in $\mathcal{A}^\mathbb{Z}$ which do not contain any of the forbidden words from $F$. A subset $X \subseteq \mathcal{A}^\mathbb{Z}$ is a shift space if and only if there exists $F \subseteq \mathcal{A}^*$ such that $X = X_F$ (cf. [21, Proposition 1.3.4]). $X$ is said to be a shift of finite type (SFT) if this is possible for a finite set $F$.

The language of a shift space $X$ is defined to be the set of all words which occur in at least one $x \in X$, and it is denoted $\mathcal{B}(X)$. $X$ is said to be irreducible if there for every $u, w \in \mathcal{B}(X)$ exists $v \in \mathcal{B}(X)$ such that $uvw \in \mathcal{B}(X)$. For each $x \in X$, define the left-ray of $x$ to be $x^- = \cdots x_{-2}x_{-1}$ and define the right-ray of $x$ to be $x^+ = x_0x_1x_2 \cdots$. The sets of all left-rays and all right-rays are, respectively, denoted $X^-$ and $X^+$.

A bijective, continuous, and shift commuting map between two shift spaces is called a conjugacy, and when such a map exists, the two shift spaces are said to be conjugate. Shift spaces $(X, \sigma_X)$ and $(Y, \sigma_Y)$ are said to be flow equivalent if the corresponding suspension flows $SX$ and $SY$ are topologically equivalent. Flow equivalence is generated by conjugacy and symbol expansion [35].
Graphs. For countable sets $E^0$ and $E^1$, and maps $r, s : E^1 \to E^0$ the quadruple $E = (E^0, E^1, r, s)$ is called a directed graph. The elements of $E^0$ and $E^1$ are, respectively, the vertices and the edges of the graph. For each edge $e \in E^1$, $s(e)$ is the vertex where $e$ starts, and $r(e)$ is the vertex where $e$ ends. A path $\lambda = e_1 \cdots e_n$ is a sequence of edges such that $r(e_i) = s(e_{i+1})$ for all $i \in \{1, \ldots, n - 1\}$. For each $n \in \mathbb{N}_0$, the set of paths of length $n$ is denoted $E^n$, and the set of all finite paths is denoted $E^*$. Extend the maps $r$ and $s$ to $E^*$ by defining $s(e_1 \cdots e_n) = s(e_1)$ and $r(e_1 \cdots e_n) = r(e_n)$. A circuit is a path $\lambda$ with $r(\lambda) = s(\lambda)$ and $|\lambda| > 0$. For $u, v \in E^0$, $u$ is said to be connected to $v$ if there is a path $\lambda \in E^*$ such that $s(\lambda) = u$ and $r(\lambda) = v$, and this is denoted by $u \geq v$ [21, Section 4.4]. A vertex is said to be maximal, if it is connected to all other vertices. $E$ is said to be irreducible if all vertices are maximal. If $E$ has a unique maximal vertex, this vertex is said to be the root of $E$. $E$ is said to be essential if every vertex emits and receives an edge. For a finite essential directed graph $E$, the edge shift $(X_E, \sigma_E)$ is defined by

$$X_E = \{ x \in (E^0)^{\mathbb{Z}} \mid r(x_i) = s(x_{i+1}) \text{ for all } i \in \mathbb{Z} \}.$$ 

A labelled graph $(E, \mathcal{L})$ over an alphabet $A$ consists of a directed graph $E$ and a surjective labelling map $\mathcal{L} : E^1 \to A$. Extend the labelling map to $E^* \to A^*$ by defining $\mathcal{L}(e_1 \cdots e_n) = \mathcal{L}(e_1) \cdots \mathcal{L}(e_n) \in A^*$. For a finite essential labelled graph $(E, \mathcal{L})$, define the shift space $(X_{(E, \mathcal{L})}, \sigma)$ by

$$X_{(E, \mathcal{L})} = \{ (\mathcal{L}(x_i))_i \in A^{\mathbb{Z}} \mid x \in X_E \}.$$ 

The labelled graph $(E, \mathcal{L})$ is said to be a presentation of the shift space $X_{(E, \mathcal{L})}$, and a representative of a word $w \in B(X_{(E, \mathcal{L})})$ is a path $\lambda \in E^*$ such that $\mathcal{L}(\lambda) = w$. Representatives of rays are defined analogously. If $H \subseteq E^0$ then the subgraph of $(E, \mathcal{L})$ induced by $H$ is the labelled subgraph of $(E, \mathcal{L})$ with vertices $H$ and edges $\{ e \in E^1 \mid s(e), r(e) \in H \}$.

Sofic shifts. A function $\pi : X_1 \to X_2$ between shift spaces $X_1$ and $X_2$ is said to be a factor map if it is continuous, surjective, and shift commuting. A shift space is called sofic [38] if it is the image of an SFT under a factor map. A shift space is sofic if and only if it can be presented by a finite labelled graph [14]. A sofic shift space is irreducible if and only if it can be presented by an irreducible labelled graph (see [21, Section 3.1]). Let $(E, \mathcal{L})$ be a finite labelled graph and let $\pi_{\mathcal{L}} : X_E \to X_{(E, \mathcal{L})}$ be the factor map induced by the labelling map $\mathcal{L} : E^1 \to A$ then the SFT $X_E$ is called a cover of the sofic shift $X_{(E, \mathcal{L})}$, and $\pi_{\mathcal{L}}$ is called the covering map.

A presentation $(E, \mathcal{L})$ of a sofic shift space $X$ is said to be left-resolving if no vertex in $E^0$ receives two edges with the same label. Fischer proved [14] that, up to labelled graph isomorphism, every irreducible sofic shift has a unique left-resolving presentation with fewer vertices than any other left-resolving presentation. This is called the left Fischer cover of $X$, and it is denoted $(F, \mathcal{L}_F)$. An irreducible sofic shift is said to have almost finite type (AFT) [22, 33] if the left Fischer cover is right-closing (see e.g. [21, Definition 5.1.4]).
For \( x \in \mathcal{B}(X) \cup X^+ \), define the \textit{predecessor set} of \( x \) to be the set of left-rays which may precede \( x \) in \( X \) (see [17, Sections I and III] and [21, Exercise 3.2.8]). The \textit{follower set} of a left-ray or word is defined analogously. Let \((E, \mathcal{L})\) be a labelled graph presenting \( X \) and let \( v \in E^0 \). Define the \textit{predecessor set} of \( v \) to be the set of left-rays in \( X \) which have a presentation terminating at \( v \). This is denoted \( P_\infty^E(v) \), or just \( P_\infty(v) \) when \((E, \mathcal{L})\) is understood from the context. The presentation \((E, \mathcal{L})\) is said to be \textit{predecessor-separated} if \( P_\infty^E(u) \neq P_\infty^E(v) \) when \( u, v \in E^0 \) and \( u \neq v \).

The \textit{left Krieger cover} of the shift space \( X \) is the labelled graph \((K, \mathcal{L}_K)\) where \( K^0 = \{ P_\infty(x^+) \mid x^+ \in X^+ \} \), and where there is an edge labelled \( a \in A \) from \( P \in K^0 \) to \( P' \in K^0 \) if and only if there exists \( x^+ \in X^+ \) such that \( P = P_\infty(ax^+) \) and \( P' = P_\infty(x^+) \). The \textit{past set cover} of the shift space \( X \) is the labelled graph \((W, \mathcal{L}_W)\) where \( W^0 = \{ P_\infty(w) \mid w \in \mathcal{B}(X) \} \) and where the edges and labels are constructed as in the Krieger cover. A shift space is sofic if and only if the number of predecessor sets is finite [19, §2], so the left Krieger cover is finite exactly when the shift space is sofic. The left Fischer cover, the left Krieger cover, and the past set cover are left-resolving and predecessor-separated presentations of \( X \).

The right Krieger cover and the future set cover are right-resolving and follower-separated covers defined analogously to the left Krieger cover and the past set cover, respectively. Every result developed for left-resolving covers in the following has an analogue for the corresponding right-resolving cover. These results can easily be obtained by considering the transposed shift space \( X^T \) (see e.g. [21, p. 39]).

3 Generalizing the Fischer cover

Jonoska [16] proved that a reducible sofic shift does not necessarily have a unique minimal left-resolving presentation. The aim of this section is to define a generalization of the left Fischer cover as the subgraph of the left Krieger cover induced by a certain subset of vertices. Let \( X \) be a sofic shift space, and let \((K, \mathcal{L}_K)\) be the left Krieger cover of \( X \). A predecessor set \( P \in K^0 \) is said to be \textit{non-decomposable} if \( V \subseteq K^0 \) and \( P = \bigcup_{Q \in V} Q \) implies that \( P \in V \).

**Lemma 3.1.** If \( P \in K^0 \) is non-decomposable then the subgraph of \((K, \mathcal{L}_K)\) induced by \( K^0 \setminus \{ P \} \) is not a presentation of \( X \).

**Proof.** Let \( E \) be the subgraph of \( K \) induced by \( K^0 \setminus \{ P \} \). Choose \( x^+ \in X^+ \) such that \( P = P_\infty(x^+) \). Let \( V \subseteq K^0 \setminus \{ P \} \) be the set of vertices where a presentation of \( x^+ \) can start. Then \( Q \subseteq P_\infty(x^+) = P \) for each \( Q \in V \), and by assumption, there exists \( y^- \in P \setminus \bigcup_{Q \in V} Q \). Hence, there is no presentation of \( y^- x^+ \) in \((E, \mathcal{L}_K|_E)\).

Lemma 3.1 shows that a subgraph of the left Krieger cover which presents the same shift must contain all the non-decomposable vertices. The next example shows that this subgraph is not always large enough.
Example 3.2. It is easy to check that the labelled graph in Figure 1 is the left Krieger cover of a reducible sofic shift $X$. Note that the predecessor set $P$ is decomposable since $P = P_1 \cup P_2$, and that the graph obtained by removing the vertex $P$ and all edges starting at or terminating at $P$ is not a presentation of the same sofic shift since there is no presentation of $f^\infty dbjk^\infty$ in this graph. Note that there is a path from $P$ to the vertex $P'$ which is non-decomposable.

Together with Lemma 3.1, this example motivates the following definition.

Definition 3.3. The generalized left Fischer cover $(G, L_G)$ of a sofic shift $X$ is defined to be the subgraph of the left Krieger cover induced by $G^0 = \{ P \in K^0 \mid P \geq P', P' \text{ non-decomposable} \}$.

The following proposition justifies the term generalized left Fischer cover.

Proposition 3.4.

(i) The generalized left Fischer cover of a sofic shift $X$ is a left-resolving and predecessor-separated presentation of $X$.

(ii) If $X$ is an irreducible sofic shift then the generalized left Fischer cover is isomorphic to the left Fischer cover.

(iii) If $X_1, X_2$ are sofic shifts with disjoint alphabets then the generalized left Fischer cover of $X_1 \cup X_2$ is the disjoint union of the generalized left Fischer covers of $X_1$ and $X_2$.

Proof. Given $y^- \in X^-$, choose $x^+ \in X^+$ such that $y^- \in P_\infty(x^+) = P$. By definition of the generalized left Fischer cover, there exist vertices $P_1, \ldots, P_n \in G^0$ such that $P = \bigcup_{i=1}^n P_i$. Choose $i$ such that $y^- \in P_i$. By construction, the left Krieger cover contains a path labelled $y^-$ terminating at $P_i$. Since
Thus, if \( y \in P \), this is also a path in the generalized left Fischer cover. This proves that the generalized left Fischer cover is a presentation of \( X^- \), and hence also a presentation of \( X \). Since the left Krieger cover is left-resolving and predecessor-separated, so is the generalized left Fischer cover.

Let \( X \) be an irreducible sofic shift, and identify the left Fischer cover \((F, L_F)\) with the top irreducible component of the left Krieger cover \((K, L_K)\) [19, Lemma 2.7]. By the construction of the generalized left Fischer cover, it follows that the left Fischer cover is a subgraph of the generalized left Fischer cover. Let \( x^+ \in X^+ \) such that \( P = P_\infty(x^+) \) is non-decomposable. Let \( S \subseteq F^0 \) be the set of vertices where a presentation of \( x^+ \) in \((F, L_F)\) can start. Then \( P = \bigcup_{v \in S} P_\infty(v) \), so \( P \subseteq F^0 \) by assumption.

Since \( X_1 \) and \( X_2 \) have no letters in common, the left Krieger cover of \( X_1 \cup X_2 \) is just the disjoint union of the left Krieger covers of \( X_1 \) and \( X_2 \). The generalized left Fischer cover inherits this property from the left Krieger cover.

The shift consisting of two non-interacting copies of the even shift is a simple example where the generalized left Fischer cover is a proper subgraph of the left Krieger cover.

**Lemma 3.5.** Let \( X \) be a sofic shift with left Krieger cover \((K, L_K)\). If there is an edge labelled \( a \) from a non-decomposable \( P \in K^0 \) to a decomposable \( Q \in K^0 \), then there exists a non-decomposable \( Q' \in K^0 \) and an edge labelled \( a \) from \( P \) to \( Q' \).

**Proof.** Choose \( x^+ \in X^+ \) such that \( P = P_\infty(ax^+) \) and \( Q = P_\infty(x^+) \). Since \( Q \) is decomposable, there exist \( n > 1 \) and non-decomposable \( Q_1, \ldots, Q_n \in K^0 \setminus \{Q\} \) such that \( Q = Q_1 \cup \cdots \cup Q_n \). Let \( S \) be the set of predecessor sets \( P' \in K^0 \) for which there is an edge labelled \( a \) from \( P' \) to \( Q_j \) for some \( 1 \leq j \leq n \). Given \( y^- \in P \), \( y^-ax^+ \in X \), so \( y^-a \in Q \). Choose \( 1 \leq i \leq n \) such that \( y^-a \in Q_i \). By construction, there exists \( P' \in S \) such that \( y^- \in P' \). Reversely, if \( y^- \in P'' \in S \) then there is an edge labelled \( a \) from \( P'' \) to \( Q_i \) for some \( 1 \leq i \leq n \), so \( y^-a \in Q_i \subseteq Q \). This implies that \( y^-ax^+ \in X \), so \( y^- \in P \). Thus \( P = \bigcup_{P' \in S} P' \), but \( P \) is non-decomposable, so this means that \( P \subseteq S \). Hence, there is an edge labelled \( a \) from \( P \) to \( Q_i \) for some \( i \), and \( Q_i \) is non-decomposable.

The following proposition is an immediate consequence of this result and the definition of the generalized left Fischer cover.

**Proposition 3.6.** The generalized left Fischer cover is essential.

The left Fischer cover of an irreducible sofic shift \( X \) is minimal in the sense that no other left-resolving presentation of \( X \) has fewer vertices. This is not always the case for the generalized left Fischer cover.

**Canonical.** Krieger proved that a conjugacy \( \Phi: X_1 \to X_2 \) between sofic shifts with left Krieger covers \((K_1, L_1)\) and \((K_2, L_2)\), respectively, induces a conjugacy \( \varphi: X_{K_1} \to X_{K_2} \) such that \( \Phi \circ \pi_1 = \pi_2 \circ \varphi \) when \( \pi_i: X_{K_i} \to X_i \) is the
covering map of the left Krieger cover of $X_1$ [19]. A cover with this property is said to be canonical. The next goal is to prove that the generalized left Fischer cover is canonical. This will be done by using results and methods used by Nasu [34] to prove that the left Krieger cover is canonical.

**Definition 3.7 (Bipartite code).** When $A, C, D$ are alphabets, an injective map $f: A \to CD$ is called a bipartite expression. If $X_1, X_2$ are shift spaces with alphabets $A_1$ and $A_2$, respectively, and if $f_1: A_1 \to CD$ is a bipartite expression then a map $\Phi: X_1 \to X_2$ is said to be a bipartite code induced by $f_1$ if there exists a bipartite expression $f_2: A_2 \to DC$ such that one of the following two conditions is satisfied:

(i) If $x \in X_1$, $y = \Phi(x)$, and $f_1(x_i) = c_i d_i$ with $c_i \in C$ and $d_i \in D$ for all $i \in \mathbb{Z}$ then $f_2(y_i) = d_i c_{i+1}$ for all $i \in \mathbb{Z}$.

(ii) If $x \in X_1$, $y = \Phi(x)$, and $f_1(x_i) = c_i d_i$ with $c_i \in C$ and $d_i \in D$ for all $i \in \mathbb{Z}$ then $f_2(y_i) = d_i-1 c_i$ for all $i \in \mathbb{Z}$.

A mapping $\Phi: X_1 \to X_2$ is called a bipartite code, if it is the bipartite code induced by some bipartite expression.

It is clear that a bipartite code is a conjugacy and that the inverse of a bipartite code is a bipartite code.

**Theorem 3.8 (Nasu [34, Thm. 2.4]).** Any conjugacy between shift spaces can be decomposed into a product of bipartite codes.

Let $\Phi: X_1 \to X_2$ be a bipartite code corresponding to bipartite expressions $f_1: A_1 \to CD$ and $f_2: A_2 \to DC$, and use the bipartite expressions to recode $X_1$ and $X_2$ to

\[
\hat{X}_1 = \{(f_1(x_i)), x \in X_1\} \subseteq (CD)^\mathbb{Z}
\]

\[
\hat{X}_2 = \{(f_2(x_i)), x \in X_2\} \subseteq (DC)^\mathbb{Z}.
\]

For $i \in \{1, 2\}$, $f_i$ induces a one-block conjugacy from $X_i$ to $\hat{X}_i$, and $\Phi$ induces a bipartite code $\hat{\Phi}: \hat{X}_1 \to \hat{X}_2$ which commutes with these conjugacies. If $\Phi$ satisfies condition (i) in the definition of a bipartite code then $(\hat{\Phi}(\hat{x}))_i = d_i c_{i+1}$ when $\hat{x} = (c_i d_i)_{i \in \mathbb{Z}} \in \hat{X}_1$. If it satisfies condition (ii) then $(\hat{\Phi}(\hat{x}))_i = d_{i-1} c_i$ when $\hat{x} = (c_i d_i)_{i \in \mathbb{Z}} \in \hat{X}_1$. The shifts $\hat{X}_1$ and $\hat{X}_2$ will be called the recoded shifts of the bipartite code, and $\hat{\Phi}$ will be called the recoded bipartite code.

A labelled graph $(G, \mathcal{L})$ is said to be bipartite if $G$ is a bipartite graph (i.e. the vertex set can be partitioned into two sets $(G^0)_1$ and $(G^0)_2$ such that no edge has its range and source in the same set). When $(G, \mathcal{L})$ is a bipartite labelled graph over an alphabet $A$, define two graphs $G_1$ and $G_2$ as follows: For $i \in \{1, 2\}$, the vertex set of $G_i$ is $(G^0)_i$, the edge set is the set of paths of length 2 in $(G, \mathcal{L})$ for which both range and source are in $(G^0)_i$, and the range and source maps are inherited from $G$. For $i \in \{1, 2\}$, define $\mathcal{L}_i: G_i \to A^2$.
by \( L(e,f) = L(e)L(f) \). The pair \((G_1,L_1), (G_2,L_2)\) is called the induced pair of labelled graphs of \((G,L)\). This decomposition is not necessarily unique, but whenever a bipartite labelled graph is mentioned, it will be assumed that the induced graphs are specified.

**Remark 3.9 (Nasu [34, Remark 4.2])**. Let \((G,L)\) be a bipartite labelled graph for which the induced pair of labelled graphs is \((G_1,L_1), (G_2,L_2)\). Let \(X_1\) and \(X_2\) be the sofic shifts presented by these graphs, and let \(X_{G_1}, X_{G_2}\) be the edge shifts generated by \(G_1, G_2\). The natural embedding \(f: G_1 \to (G_1)^2\) is a bipartite expression which induces two bipartite codes \(\varphi_{\pm}: X_{G_1} \to X_{G_2}\) such that \((\varphi_+(x))_i = f_i e_i + 1\) and \((\varphi_-(x))_i = f_{i-1} e_i\) when \(x = (e_i f_i)_{i \in \mathbb{Z}} \in X_{G_2}\). Similarly, the embedding \(F: L_1(G_1) \to (L(G_1))^2\) is a bipartite expression which induces bipartite codes \(\Phi_{\pm}: X_1 \to X_2\) such that \((\Phi_+(x))_i = b_i a_i + 1\) and \((\Phi_-(x))_i = b_{i-1} a_i\) when \(x = (a_i b_i)_{i \in \mathbb{Z}} \in X_1\). By definition, \(\Phi_{\pm} \circ \pi_1 = \pi_2 \circ \varphi_{\pm}\) when \(\pi_1: X_{G_1} \to X_1, \pi_2: X_{G_2} \to X_2\) are the covering maps. The bipartite codes \(\varphi_{\pm}\) and \(\Phi_{\pm}\) are called the standard bipartite codes induced by \((G,L)\).

**Lemma 3.10 (Nasu [34, Cor. 4.6 (1)]).** Let \(\Phi: X_1 \to X_2\) be a bipartite code between sofic shifts \(X_1, X_2\). Let \(X_1, X_2\) be the recoded shifts of \(X_1\) and \(X_2\) respectively, and let \(K_1, L_1\) and \(K_2, L_2\) be the left Krieger covers of \(X_1\) and \(X_2\) respectively. Then there exists a sofic shift \(\hat{X}\) for which the left Krieger cover is a bipartite labelled graph such that the induced pair of labelled graphs is \((K_1,L_1), (K_2,L_2)\) and such that the recoded bipartite code \(\hat{\Phi}: \hat{X}_1 \to \hat{X}_2\) of \(\Phi\) is one of the standard bipartite codes \(\Phi_{\pm}\) induced by the left Krieger cover of \(\hat{X}\) as defined in Remark 3.9.

The proof of the following theorem is very similar to the proof of the corresponding result by Nasu [34, Thm. 3.3] for the left Krieger cover.

**Theorem 3.11.** The generalized left Fischer cover is canonical.

**Proof.** Let \(\Phi: X_1 \to X_2\) be a bipartite code. Let \(\hat{X}_1, \hat{X}_2\) be the recoded shifts, let \(K_1, L_1\), \(K_2, L_2\) be the corresponding left Krieger covers, and let \(\hat{\Phi}: \hat{X}_1 \to \hat{X}_2\) be the recoded bipartite code. Use Lemma 3.10 to find a sofic shift \(\hat{X}\) such that the left Krieger cover \((K,L)\) of \(\hat{X}\) is a bipartite labelled graph for which the induced pair of labelled graphs is \((K_1,L_1), (K_2,L_2)\). Let \((G_1,L_1), (G_2,L_2)\) and \((G,L)\) be the generalized left Fischer covers of respectively \(X_1, X_2\), and \(\hat{X}\).

The labelled graph \((G,L)\) is bipartite since \(G\) is a subgraph of \(K\). Note that a predecessor set \(P\) in \(K_0\) or \(K^0\) is decomposable if and only if the corresponding predecessor set in \(K^0\) is decomposable. If \(i \in \{1,2\}\) and \(Q \subseteq G_i \subseteq K_i^0\) then there is a path in \(K_i\) from \(Q\) to a non-decomposable \(P \subseteq K_i^0\). By considering the corresponding path in \(K\), it is clear that the vertex in \(K_i^0\) corresponding to \(Q\) is in \(G^0\). Conversely, if \(Q \subseteq G^0\) then there is a path in \(K\) from \(Q\) to a non-decomposable \(P \subseteq K^0\). If \(P\) and \(Q\) belong to the same partition \(K_i^0\) then the vertex in \(K_i\) corresponding to \(Q\) is in \(G_i^0\) by definition. On the other hand, if \(Q\) corresponds to a vertex in \(K_i\) and if \(P\) belongs to the other partition
then Lemma 3.5 shows that there exists a non-decomposable \( P' \) in the same partition as \( Q \) and an edge from \( P \) to \( P' \) in \( K \). Hence, there is also a path in \( K \) from the vertex corresponding to \( Q \) to the vertex corresponding to \( P' \), so \( Q \in G^0_{11} \). This proves that the pair of induced labelled graphs of \((G, \mathcal{L})\) is \((G_1, \mathcal{L}_1), (G_2, \mathcal{L}_2)\).

Let \( \Psi_{\pm} : X_1 \to X_2 \) be the standard bipartite codes induced by \((G, \mathcal{L})\). Remark 3.9 shows that there exist bipartite codes \( \hat{\psi}_{\pm} : X_{G_1} \to X_{G_2} \) such that \( \Psi_{\pm} \circ \hat{\pi}_1|_{X_{G_1}} = \hat{\pi}_2|_{X_{G_2}} \circ \hat{\psi}_{\pm} \). The labelled graph \((G, \mathcal{L})\) presents the same sofic shift as \((K, \mathcal{L})\), so they both induce the same standard bipartite codes from \( X_1 \) to \( X_2 \), and by Lemma 3.10, \( \hat{\Phi} \) is one of these standard bipartite codes, so \( \hat{\Phi} = \hat{\Psi}_{\pm} \). In particular, there exists a bipartite code \( \hat{\psi} : X_{G_1} \to X_{G_2} \) such that \( \hat{\Phi} \circ \hat{\pi}_1|_{X_{G_2}} = \hat{\pi}_2|_{X_{G_2}} \circ \hat{\psi} \).

By recoding \( \hat{X}_1 \) to \( X_1 \) and \( \hat{X}_2 \) to \( X_2 \) via the bipartite expressions inducing \( \Phi \), this gives a bipartite code \( \hat{\psi} \) such that \( \hat{\Phi} \circ \hat{\pi}_1 = \hat{\pi}_2 \circ \hat{\psi} \) when \( \pi_1, \pi_2 \) are the covering maps of the generalized left Fischer covers of \( X_1 \) and \( X_2 \) respectively. By Theorem 3.8, any conjugacy can be decomposed as a product of bipartite codes, so this proves that the generalized left Fischer cover is canonical.

**Theorem 3.12.** For \( i \in \{1, 2\} \), let \( X_i \) be a sofic shift with generalized left Fischer cover \((G_i, \mathcal{L}_i)\) and covering map \( \pi_i : X_{G_i} \to X_i \). If \( \Phi : SX_1 \to SX_2 \) is a flow equivalence then there exists a unique flow equivalence \( \varphi : SX_{G_1} \to SX_{G_2} \) such that \( \Phi \circ S\pi_1 = S\pi_2 \circ \varphi \).

**Proof.** In [5] it is proved that the left Krieger cover respects symbol expansion: if \( X \) is a sofic shift with alphabet \( A \), \( a \in A \), \( \bullet \) is some symbol not in \( A \), and if \( \hat{X} \) is obtained from \( X \) via a symbol expansion which inserts \( a \bullet \) after each \( a \) then the left Krieger cover of \( \hat{X} \) is obtained by replacing each edge labelled \( a \) in the left Krieger cover of \( X \) by two edges in sequence labelled \( a \) and \( \bullet \) respectively. Clearly, the generalized left Fischer cover inherits this property. By [5], any canonical cover which respects flow equivalence has the desired property, so the result follows from Theorem 3.11.

4 Foundations and layers of covers

Let \( \mathcal{E} = (E, \mathcal{L}) \) be a finite left-resolving and predecessor-separated labelled graph. For each \( V \subseteq E^0 \) and each word \( w \) over the alphabet \( A \) of \( \mathcal{L} \) define

\[ wV = \{ u \in E^0 \mid u \text{ is the source of a path labelled } w \text{ terminating in } V \}. \]

**Definition 4.1.** Let \( S \) be a subset of the power set \( \mathcal{P}(E^0) \), and let \( \sim \) be an equivalence relation on \( S \). The pair \((S, \sim)\) is said to be past closed if

- \( \{ v \} \in S \),
- \( \{ u \} \sim \{ v \} \) implies \( u = v \),
- \( aV \neq \emptyset \) implies \( aV \in S \), and
they present the same sofic shift.

Proof. By assumption, there is a bijection between \( [V] \) and \( [aV] \). For each \( V \in S \), let \( [V] \) denote the equivalence class of \( V \) with respect to \( \sim \). When \( a \in A \) and \( V \in S \), \([V] \) is said to receive \( a \) if \( aV \neq \emptyset \). For each \( [V] \in S/\sim \), define \( ||V|| = \min_{V \in [V]} ||V|| \).

Definition 4.2. Define \( G(\mathcal{E}, S, \sim) \) to be the labelled graph with vertex set \( S/\sim \) for which there is an edge labelled \( a \) from \([aV]\) to \([V]\) whenever \([V]\) receives \( a \). For each \( n \in \mathbb{N} \), the \( n \)th layer of \( G(\mathcal{E}, S, \sim) \) is the labelled subgraph induced by \( S_n = \{ [V] \in S/\sim \mid n = ||V|| \} \). \( \mathcal{E} \) is said to be a foundation of any labelled graph isomorphic to \( G(\mathcal{E}, S, \sim) \).

If a labelled graph \( \mathcal{H} \) is isomorphic to \( G(\mathcal{E}, S, \sim) \) then the subgraph of \( \mathcal{H} \) corresponding to the \( n \)th layer of \( G(\mathcal{E}, S, \sim) \) is be said to be the \( n \)th layer of \( \mathcal{H} \) with respect to \( \mathcal{E} \), or simply the \( n \)th layer if \( \mathcal{E} \) is understood from the context.

Proposition 4.3. \( \mathcal{E} \) and \( G(\mathcal{E}, S, \sim) \) present the same sofic shift, and \( \mathcal{E} \) is labelled graph isomorphic to the first layer of \( G(\mathcal{E}, S, \sim) \).

Proof. By assumption, there is a bijection between \( E^0 \) and the set of vertices in the first layer of \( G(\mathcal{E}, S, \sim) \). By construction, there is an edge labelled \( a \) from \( u \) to \( v \) in \( \mathcal{E} \) if and only if there is an edge labelled \( a \) from \([u]\) to \([v]\) in \( G(\mathcal{E}, S, \sim) \). Every finite word presented by \( G(\mathcal{E}, S, \sim) \) is also presented by \( \mathcal{E} \), so they present the same sofic shift.

The following proposition motivates the use of the term layer by showing that edges can never go from higher to lower layers.

Proposition 4.4. If \( [V] \in S/\sim \) receives \( a \in A \) then \( ||aV|| \leq ||V|| \). If \( G(\mathcal{E}, S, \sim) \) has an edge from a vertex in the \( n \)th layer to a vertex in the \( m \)th layer then \( m \leq n \).

Proof. Choose \( V \in [V] \) such that \( ||V|| = ||V|| \). Each \( u \in aV \) emits at least one edge labelled \( a \) terminating in \( V \), and \( \mathcal{E} \) is left-resolving, so \( ||aV|| \leq ||aV|| \leq ||V|| = ||V|| \). The second statement follows from the definition of \( G(\mathcal{E}, S, \sim) \).

Example 4.5. Let \( (F, \mathcal{L}_F) \) be the left Fischer cover of an irreducible sofic shift \( X \). For each \( x^+ \in X^+ \), define \( s(x^+) \subseteq F^0 \) to be the set of vertices where a presentation of \( x^+ \) can start. \( S = \{ s(x^+) \mid x^+ \in X^+ \} \subseteq \mathcal{P}(F^0) \) is past closed since each vertex in the left Fischer cover is the predecessor set of an intrinsically synchronizing right-ray, so the multiplicity set cover of \( X \) can be defined to be \( G((F, \mathcal{L}_F), S, \sim) \). An analogous cover can be defined by considering the vertices where presentations of finite words can start. Thomsen [37] constructs the derived shift space \( \partial X \) of \( X \) using right-resolving graphs, but an analogous construction works for left-resolving graphs. The procedure from [37, Example 6.10] shows that this \( \partial X \) is presented by the labelled graph obtained by removing the left Fischer cover from the multiplicity set cover.
Let $X$ be a sofic shift, and let $(K, \mathcal{L}_K)$ be the left Krieger cover of $X$. In order to use the preceding results to investigate the structure of the left Krieger cover and the past set cover, define an equivalence relation on $P(K^0)$ by $U \sim_U V$ if and only if $\bigcup_{P \in U} P = \bigcup_{Q \in V} Q$. Clearly, $\{P\} \sim_U \{Q\}$ if and only if $P = Q$. If $U, V \subseteq K^0$, $a \in A$, $aV \neq \emptyset$, and $U \sim_U V$ then $aU \sim_U aV$ by the definition of the left Krieger cover.

**Theorem 4.6.** For a sofic shift $X$, the generalized left Fischer cover $(G, \mathcal{L}_G)$ is a foundation of the left Krieger cover $(K, \mathcal{L}_K)$, and no smaller subgraph is a foundation.

**Proof.** Define $S = \{V \subseteq G^0 \mid \exists x^+ \in X^+ \text{ such that } P_\infty(x^+) = \bigcup_{P \in V} P\}$. Note that $\{P\} \in S$ for every $P \in G^0$. If $x^+ \in X^+$ with $P_\infty(x^+) = \bigcup_{P \in V} P$ and if $aV \neq \emptyset$ for some $a \in A$ then $ax^+ \in X^+$ and $P_\infty(ax^+) = \bigcup_{P \in aV} P$. This proves that the pair $(S, \sim_U)$ is past closed, so $\mathcal{G}((G, \mathcal{L}_G), S, \sim_U)$ is well defined. Since $(G, \mathcal{L}_G)$ is a presentation of $X$, there is a bijection $\varphi: S/\sim_U \rightarrow K^0$ defined by $\varphi([V]) = \bigcup_{P \in V} P$. By construction, there is an edge labelled $a$ from $[U]$ to $[V]$ in $\mathcal{G}((G, \mathcal{L}_G), S, \sim_U)$ if and only if there exists $x^+ \in X^+$ such that $P_\infty(ax^+) = \bigcup_{P \in U} P$ and $P_\infty(x^+) = \bigcup_{Q \in V} Q$, so $\mathcal{G}((G, \mathcal{L}_G), S, \sim_U)$ is isomorphic to $(K, \mathcal{L}_K)$. It follows from Lemma 3.1 that no proper subgraph of $(G, \mathcal{L}_G)$ can be a foundation of the left Krieger cover.

The example from [12, Section 4] shows that the left Krieger cover can be a proper subgraph of the past set cover. The following lemma will be used to further investigate this relationship.

**Lemma 4.7.** Let $X$ be a sofic shift. For every right-ray $x^+ = x_1x_2x_3\ldots \in X^+$ there exists $n \in \mathbb{N}$ such that $P_\infty(x^+) = P_\infty(x_1x_2\ldots x_k)$ for all $k \geq n$.

**Proof.** It is clear that $P_\infty(x_1) \supseteq P_\infty(x_1x_2) \supseteq \cdots \supseteq P_\infty(x^+)$. Since $X$ is sofic, there are only finitely many different predecessor sets of words, so there must exist $n \in \mathbb{N}$ such that $P_\infty(x_1x_2\ldots x_k) = P_\infty(x_1x_2\ldots x_n)$ for all $k \geq n$. If $y^- \in P_\infty(x_1x_2\ldots x_n)$ is given, then $y^-x_1x_2\ldots x_k \in X$ for all $k \geq n$, so $y^-x^+$ contains no forbidden words, and therefore $y^- \in P_\infty(x^+)$. Since $y^-$ was arbitrary, $P_\infty(x^+) = P_\infty(x_1x_2\ldots x_n)$.

**Theorem 4.8.** For a sofic shift $X$, the generalized left Fischer cover $(G, \mathcal{L}_G)$ and the left Krieger cover $(K, \mathcal{L}_K)$ are both foundations of the past set cover $(W, \mathcal{L}_W)$.

**Proof.** Define $S = \{V \subseteq G^0 \mid \exists w \in B(X) \text{ such that } P_\infty(w) = \bigcup_{P \in V} P\}$, and use Lemma 4.7 to conclude that $S$ contains $\{P\}$ for every $P \in G^0$. By arguments analogous to the ones used in the proof of Theorem 4.6, it follows that $\mathcal{G}((G, \mathcal{L}_G), S, \sim_U)$ is isomorphic to $(W, \mathcal{L}_W)$. To see that $(K, \mathcal{L}_K)$ is also a foundation, define $T = \{V \subseteq K^0 \mid \exists w \in B(X) \text{ such that } P_\infty(w) = \bigcup_{P \in V} P\}$, and apply arguments analogous to the ones used above to prove that $(W, \mathcal{L}_W)$ is isomorphic to $\mathcal{G}((K, \mathcal{L}_K), T, \sim_U)$.

---

*Documenta Mathematica 16 (2011) 111–131*
In the following, the $n$th layer of the left Krieger cover (past set cover) will always refer to the $n$th layer with respect to the generalized left Fischer cover $(G, L_G)$. For a right-ray (word) $x$, $P_\infty(x)$ is a vertex in the $n$th layer of the left Krieger cover (past set cover) for some $n \in \mathbb{N}$, and such an $x$ is said to be $1/n$-synchronizing. Note that $x$ is $1/n$-synchronizing if and only if $n$ is the smallest number such that there exist $P_1, \ldots, P_n \in G^0$ with $\bigcup_{i=1}^n P_i = P_\infty(x)$.

In an irreducible sofic shift with left Fischer cover $(F, L_F)$, this happens if and only if $n$ is the smallest number such that there exist $u_1, \ldots, u_n \in F^0$ with $\bigcup_{i=1}^n P_\infty(u_i) = P_\infty(x)$.

Example 4.9. Figures 2 and 3 show, respectively, the left Fischer and the left Krieger cover of the 3-charge constrained shift (see e.g. [21, Example 1.2.7] for the definition of charge constrained shifts). There are 3 vertices in the second layer of the left Krieger cover and two in the third. Note how the left Fischer cover can be identified with the first layer of the left Krieger cover.

Corollary 4.10. If the left Krieger cover of a sofic shift is reducible then so is the past set cover.

Proof. This follows from Proposition 4.4 and Theorem 4.8.
5 The range of a flow invariant

Let $E$ be a directed graph. Vertices $u, v \in E^0$ properly communicate [1] if there are paths $\mu, \lambda \in E^*$ of length greater than or equal to 1 such that $s(\mu) = u$, $r(\mu) = v$, $s(\lambda) = v$, and $r(\lambda) = u$. This relation is used to construct maximal disjoint subsets of $E^0$, called proper communication sets of vertices, such that $u, v \in E^0$ properly communicate if and only if they belong to the same subset. The proper communication graph $PC(E)$ is defined to be the directed graph for which the vertices are the proper communication sets of vertices of $E$ and for which there is an edge from one proper communication set to another if and only if there is a path from a vertex in the first set to a vertex in the second. The proper communication graph of the left Krieger cover of a sofic shift space is a flow-invariant [1].

Let $X$ be an irreducible sofic shift with left Fischer cover $(F, L_F)$ and left Krieger cover $(K, L_K)$, and let $E$ be the proper communication graph of $K$.

By construction, $E$ is finite and contains no circuit. The left Fischer cover is isomorphic to an irreducible subgraph of $(K, L_K)$ corresponding to a root $r \in E^0$ [19, Lemma 2.7], and by definition, there is an edge from $u \in E^0$ to $v \in E^0$ whenever $u > v$. The following proposition gives the range of the flow-invariant by proving that all such graphs can occur.

**Proposition 5.1.** Let $E$ be a finite directed graph with a root and without circuits. $E$ is the proper communication graph of the left Krieger cover of an AFT shift if there is an edge from $u \in E^0$ to $v \in E^0$ whenever $u > v$.

**Proof.** Let $E$ be an arbitrary finite directed graph which contains no circuit and which has a root $r$, and let $\bar{E}$ be the directed graph obtained from $E$ by adding an edge from $u \in E^0$ to $v \in E^0$ whenever $u > v$. The goal is to construct a labelled graph $(F, L_F)$ which is the left Fischer cover of an irreducible sofic shift with the desired properties. For each $v \in E^0$, let $l(v)$ be the length of the longest path from $r$ to $v$. This is well-defined since $E$ does not contain any circuits. For each $v \in E^0$, define $n(v) = 2^{l(v) + 1}$ vertices $v_1, \ldots, v_{n(v)} \in F^0$. The single vertex corresponding to the root $r \in E^0$ is denoted $r_1$. For each $v \in E^0$, draw a loop of length 1 labelled $a_v$ at each of the vertices $v_1, \ldots, v_{n(v)} \in F^0$. If there is an edge from $u \in E^0$ to $v \in E^0$ then $l(v) > l(u)$. From each vertex $u_1, \ldots, u_{n(v)}$ draw $n(u, v) = n(v)/n(u) = 2^{l(v) - l(u)} \geq 2$ edges labelled $a_{u_1,v}, \ldots, a_{u_{n(v)},v}$ such that every vertex $v_1, \ldots, v_{n(v)}$ receives exactly one of these edges. For each sink $v \in E^0$ draw a uniquely labelled edge from each vertex $v_1, \ldots, v_{n(v)}$ to $r_1$. This finishes the construction of $(F, L_F)$.

By construction, $F$ is irreducible, right-resolving, and left-resolving. Additionally, it is predecessor-separated because there is a uniquely labelled path to every vertex in $E^0$ from $r_1$. Thus, $(F, L_F)$ is the left Fischer cover of an AFT shift $X$. Let $(K, L_K)$ be the left Krieger cover of $X$.

For every $v \in E^0$, $P_\infty(a_v^\infty) = \bigcup_{i=1}^{n(v)} P_\infty(v_i)$ and no smaller set of vertices has this property, so $P_\infty(a_v^\infty)$ is a vertex in the $n(v)$th layer of the left Krieger cover. There is clearly a loop labelled $a_v$ at the vertex $P_\infty(a_v^\infty)$, so it belongs
to a proper communication set of vertices. Furthermore, $ba^\infty_v \in X^+$ if and only if $b = a_v$ or $b = a^i_u, v$ for some $u \in E^0$ and $1 \leq i \leq n(u, v)$. By construction, $P^\infty_u (a^i_u, a^\infty_v) = \bigcup_{i=1}^n (u, v) P^\infty (a_u)$, so there is an edge from $P^\infty (a^\infty_v)$ if and only if there is an edge from $u$ to $v$ in $E$. This proves that $E$, and hence also $\tilde{E}$, are subgraphs of the proper communication graph of $K$.

Since the edges which terminate at $r_1$ are uniquely labelled, any $x^+ \in X^+$ which contains one of these letters must be intrinsically synchronizing. If $x^+ \in X^+$ does not contain any of these letters then $x^+$ must be eventually periodic with $x^+ = wa^\infty_v$ for some $v \in E^0$ and $w \in \mathcal{B}(X)$. Thus, $K$ only has the vertices described above, and therefore the proper communication graph of $K$ is $\tilde{E}$.

**Example 5.2.** To illustrate the construction used in the proof of Proposition 5.1, let $E$ be the directed graph drawn in Figure 4. $E$ has a unique maximal vertex $r$ and contains no circuit, so it is the proper communication graph of

![Figure 4: A directed graph with root $r$ and without circuits.](image)

![Figure 5: Left Fischer cover of the sofic shift $X$ considered in Example 5.2.](image)
Figure 6: Left Krieger cover of the shift space $X$ considered in Example 5.2. The structure of the irreducible component corresponding to the left Fischer cover has been suppressed.

the left Krieger cover of an irreducible sofic shift. Note that $l(x) = l(y) = 1$ and that $l(z) = 2$. Figure 5 shows the left Fischer cover of a sofic shift $X$ constructed using the method from the proof of Proposition 5.1. Note that the top and bottom vertices should be identified, and that the labelling of the edges terminating at $r_1$ has been suppressed. Figure 6 shows the left Krieger cover of $X$, but the structure of the irreducible component corresponding to the left Fischer cover has been suppressed to emphasize the structure of the higher layers.

In [1], it was also remarked that an invariant analogous to the one discussed in Proposition 5.1 is obtained by considering the proper communication graph of the right Krieger cover. The following example shows that the two invariants may carry different information.

**Example 5.3.** The labelled graph in Figure 7 is left-resolving, irreducible, and predecessor-separated, so it is the left Fischer cover of an irreducible sofic shift. Similarly, the labelled graph in Figure 8 is irreducible, right-resolving and follower-separated, so it is the right Fischer cover of an irreducible sofic shift. By considering the edges labelled $d$, it is easy to see that the two graphs present the same sofic shift space $X$.

Every right-ray which contains a letter different from $a$ or $a'$ is intrinsically synchronizing, so consider a right-ray $x^+ \in X^+$ such that $(x^+)_i \in \{a, a'\}$ for all $i \in \mathbb{N}$. By considering Figure 7, it is clear that $P_\infty(x^+) = P_\infty(a) \cup P_\infty(v) \cup P_\infty(y) = P_\infty(y)$, so $P(x^+)$ is also in the first layer of the left Krieger cover. Hence, the proper communication graph has only one vertex and no edges.

Every left-ray containing a letter different from $a$ or $a'$ is intrinsically synchronizing, so consider the left-ray $a^\infty \in X^-$. Figure 8 shows that $F_\infty(a^\infty) =$
$F_\infty(u') \cup F_\infty(v')$ and that no single vertex $y'$ in the right Fischer cover has $F_\infty(y') = F_\infty(a^\infty)$, so there is a vertex in the second layer of the right Krieger cover. In particular, the corresponding proper communication graph is non-trivial.

6 \hspace{1cm} \textit{C*-Algebras associated to sofic shift spaces}

Cuntz and Krieger [13] introduced a class of $C^*$-algebras which can naturally be viewed as the universal $C^*$-algebras associated to shifts of finite type. This was generalized by Matsumoto [23] who associated two $C^*$-algebras $\mathcal{O}_X$ and $\mathcal{O}_X^*$ to every shift space $X$, and these Matsumoto algebras have been studied intensely [8, 18, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32]. The two Matsumoto algebras $\mathcal{O}_X$ and $\mathcal{O}_X^*$ are generated by elements satisfying the same relations,
but they are not isomorphic in general [12]. This paper will follow the approach
of Carlsen in [9] where a universal C∗-algebra O_\tilde{X} is associated to every one-
sided shift space \tilde{X}. This also gives a way to associate C∗-algebras to every
two-sided shift since a two-sided shift X corresponds to two one-sided shifts
X^+ and X^-.

Ideal lattices. Let X be a sofic shift space and let O_X be the universal
C∗-algebra associated to the one-sided shift X^+ as defined in [9]. Carlsen
proved that O_X is isomorphic to the Cuntz-Krieger algebra of the left Krieger
cover of X [8], so the lattice of gauge invariant ideals in O_X is given by the
proper communication graph of the left Krieger cover of X [3, 20], and all
ideals are given in this way if the left Krieger cover satisfies Condition (K)
[36, Theorem 4.9]. Hence, Proposition 4.4 and Theorem 4.6 can be used to
investigate the ideal lattice of O_X. For a reducible sofic shift, a part of the
ideal lattice is given by the structure of the generalized left Fischer cover, which
is reducible, but if X is an irreducible sofic shift, and the left Krieger cover of
X satisfies Condition (K) then the fact that the left Krieger cover has a unique
top component implies that O_X will always have a unique maximal ideal. The
following proposition shows that all these lattices can be realized.

**Proposition 6.1.** Any finite lattice of ideals with a unique maximal ideal is
the ideal lattice of the universal C∗-algebra O_X associated to an AFT shift X.

**Proof.** Let E be a finite directed graph without circuits and with a unique
maximal vertex. Consider the following slight modification of the algorithm
from the proof of Proposition 5.1. For each v \in E, draw two loops of length 1
at each vertex v_1, \ldots, v_n(v) associated to v: One labelled a_v and one labelled
a_v'. The rest of the construction is as before. Let (K, L_K) be the left Krieger
cover of the corresponding sofic shift. As before, the proper communication
graph of K is given by E, and now (K, L_K) satisfies Condition (K), so there is
a bijective correspondence between the hereditary subsets of E^0 and the ideals
of C*(K) \cong O_X. Since E was arbitrary, any finite ideal lattice with a unique
maximal ideal can be obtained in this way.

The C∗-algebras O_X^+ and O_X^−. Every two-sided shift space X corresponds
to two one-sided shift spaces X^+ and X^−, and this gives two natural ways to
associate a universal C∗-algebra to X. The next goal is to show that these two
C∗-algebras may carry different information about the shift space. Let O_X^− be the universal C∗-algebra associated to the one-sided shift space (X^T)^+ as
defined in [9]. The left Krieger cover of X^T is the transpose of the right Krieger
cover of X, so by [8], O_X^− is isomorphic to the Cuntz-Krieger algebra of the
transpose of the right Krieger cover of X.

**Example 6.2.** Let X be the sofic shift from Example 5.3. Note that the left and right Krieger covers of X both satisfy Condition (K) from [36], so the
corresponding proper communication graphs completely determine the ideal
lattices of O_X^+ and O_X^−. The proper communication graph of the left Krieger
cover (K, L_K) of X is trivial, so O_X^+ is simple, while there are precisely two
vertices in the proper communication graph of the right Krieger cover of $X$, so there is exactly one non-trivial ideal in $O_{X^-}$. In particular, $O_{X^+}$ and $O_{X^-}$ are not isomorphic.

Consider the edge shift $Y = X_K$. This is an SFT, and the left and right Krieger covers of $Y$ are both $(K, \mathcal{L}_{1,0})$, where $\mathcal{L}_{1,0}$ is the identity map on the edge set $K^I$.

By [8], $O_{X^+}$ and $O_{Y^+}$ are isomorphic to $C^*(K)$. Similarly, $O_{Y^-}$ is isomorphic to $C^*(K^T)$ and $K^T$ is an irreducible graph satisfying Condition (K), so $O_{Y^-}$ is simple. In particular, $O_{Y^-}$ is not isomorphic to $O_{X^-}$. This shows that the $C^*$-algebras associated to $X^+$ and $X^-$ are not always isomorphic, and that there can exist a shift space $Y$ such that $O_{Y^+}$ is isomorphic to $O_{X^+}$ while $O_{Y^-}$ is not isomorphic to $O_{X^-}$.

**An investigation of Condition (\#).** In [12], two $C^*$-algebras $O_X$ and $O_{X^+}$ are associated to every two-sided shift space $X$. The $C^*$-algebras $O_X$, $O_{X^+}$, and $O_{X^-}$ are generated by partial isometries satisfying the same relations, but $O_{X^+}$ is always universal unlike $O_X$ [9]. In [12], it is proved that $O_X$ and $O_{X^+}$ are isomorphic when $X$ satisfies a condition called Condition (\#). The example from [12, Section 4] shows that not all sofic shift spaces satisfy this condition by constructing a sofic shift where the left Krieger cover and the past set cover are not isomorphic. The relationship between Condition (\#) and the structure of the left Krieger cover and the past set cover is further clarified by the following main result. For each $l \in \mathbb{N}$ and $w \in \mathcal{B}(X)$ define $P_l(w) = \{v \in \mathcal{B}(X) \mid vw \in \mathcal{B}(X), |v| \leq l\}$. Two words $u, v \in \mathcal{B}(X)$ are said to be $l$-past equivalent if $P_l(v) = P_l(w)$. For $x^+ \in X^+$, $P_l(x^+)$ and $l$-past equivalence are defined analogously.

**Condition (\#).** For every $l \in \mathbb{N}$ and every infinite $F \subseteq \mathcal{B}(X)$ such that $P_l(u) = P_l(v)$ for all $u, v \in F$ there exists $x^+ \in X^+$ such that $P_l(w) = P_l(x^+)$ for all $w \in F$.

**Lemma 6.3.** A vertex $P$ in the past set cover of a sofic shift $X$ is in an essential subgraph if and only if there exist infinitely many $w \in \mathcal{B}(X)$ such that $P_\infty(w) = P$.

**Proof.** Let $P$ be a vertex in an essential subgraph of the past set cover of $X$, and let $x^+ \in X^+$ be a right ray with a presentation starting at $P$. Given $n \in \mathbb{N}$, there exists $w_n \in \mathcal{B}(X)$ such that $P = P_\infty(x_1x_2\ldots x_nw_n)$. To prove the converse, let $P$ be a vertex in the past set cover for which there exist infinitely many $w \in \mathcal{B}(X)$ such that $P = P_\infty(w)$. For each $w$, there is a path labelled $w_{[1,|w|-1]}$ starting at $P$. There are no sources in the past set cover, so this implies that $P$ is not stranded. \(\square\)

**Proposition 6.4.** A sofic shift $X$ satisfies Condition (\#) if and only if the left Krieger cover is the maximal essential subgraph of the past set cover.

**Proof.** Assume that $X$ satisfies Condition (\#). Let $P$ be a vertex in an essential subgraph of the past set cover and define $F = \{w \in \mathcal{B}(X) \mid P_\infty(w) = P\}$. Choose $m \in \mathbb{N}$ such that for all $x, y \in \mathcal{B}(X) \cup X^+$, $P_\infty(x) = P_\infty(y)$ if and only
if $P_m(x) = P_m(y)$. By Lemma 6.3, $F$ is an infinite set, so Condition (†) can be used to choose $x^+ \in X^+$ such that $P_m(x^+) = P_m(w)$ for all $w \in F$. By the choice of $m$, this means that $P_\infty(x^+) = P_\infty(w) = P$ for all $w \in F$, so $P$ is a vertex in the left Krieger cover.

To prove the other implication, assume that the left Krieger cover is the maximal essential subgraph of the past set cover. Let $l \in \mathbb{N}$ be given, and consider an infinite set $F \subseteq B(X)$ for which $P_l(u) = P_l(v)$ for all $u, v \in F$. Since $X$ is sofic, there are only finitely many different predecessor sets, so there must exist $w \in F$ such that $P_\infty(w) = P_\infty(v)$ for infinitely many $v \in F$. By Lemma 6.3, this proves that $P = P_\infty(w)$ is a vertex in the maximal essential subgraph of the past set cover. By assumption, this means that it is a vertex in the left Krieger cover, so there exists $x^+ \in X^+$ such that $P_\infty(w) = P_\infty(x^+)$. In particular, $P_l(x^+) = P_l(w) = P_l(v)$ for all $v \in F$, so Condition (†) is satisfied.

In [2], it was proved that $O_X^+$ is isomorphic to the Cuntz-Krieger algebra of the past set cover of $X$ when $X$ satisfies a condition called Condition (I). According to Carlsen [7], a proof similar to the proof which shows that $O_X^+$ is isomorphic to the Cuntz-Krieger algebra of the left Krieger cover of $X$ should prove that $O_{X^+}^+$ is isomorphic to the Cuntz-Krieger algebra of the subgraph of the past set cover of $X$ induced by the vertices $P$ for which there exist infinitely many words $w$ such that $P_\infty(w) = P$. Using Lemma 6.3, this shows that $O_X^+$ is always isomorphic to the Cuntz-Krieger algebra of the maximal essential subgraph of the past set cover of $X$.

References

[1] T. Bates, S. Eilers, and D. Pask. Reducibility of covers of AFT shifts. *Israel J. Math.*, 2011 (in press).

[2] T. Bates and D. Pask. $C^*$-algebras of labelled graphs. *J. Operator Theory*, 57(1):207–226, 2007.

[3] T. Bates, D. Pask, I. Raeburn, and W. Szymański. The $C^*$-algebras of row-finite graphs. *New York J. Math.*, 6:307–324 (electronic), 2000.

[4] M. Boyle. Flow equivalence of shifts of finite type via positive factorizations. *Pacific J. Math.*, 204(2):273–317, 2002.

[5] M. Boyle, T. M. Carlsen, and S. Eilers. Flow equivalence of sofic shifts. *In preparation*.

[6] M. Boyle and D. Huang. Poset block equivalence of integral matrices. *Trans. Amer. Math. Soc.*, 355(10):3861–3886 (electronic), 2003.

[7] T. M. Carlsen. Personal communication.

[8] T. M. Carlsen. On $C^*$-algebras associated with sofic shifts. *J. Operator Theory*, 49(1):203–212, 2003.
[9] T. M. Carlsen. Cuntz-Pimsner $C^*$-algebras associated with subshifts. *Internat. J. Math.*, 19(1):47–70, 2008.

[10] T. M. Carlsen and S. Eilers. Augmenting dimension group invariants for substitution dynamics. *Ergodic Theory Dynam. Systems*, 24(4):1015–1039, 2004.

[11] T. M. Carlsen and S. Eilers. Matsumoto $K$-groups associated to certain shift spaces. *Doc. Math.*, 9:639–671 (electronic), 2004.

[12] T. M. Carlsen and K. Matsumoto. Some remarks on the $C^*$-algebras associated with subshifts. *Math. Scand.*, 95(1):145–160, 2004.

[13] J. Cuntz and W. Krieger. A class of $C^*$-algebras and topological Markov chains. *Invent. Math.*, 56(3):251–268, 1980.

[14] R. Fischer. Sofic systems and graphs. *Monatsh. Math.*, 80(3):179–186, 1975.

[15] D. Huang. Flow equivalence of reducible shifts of finite type. *Ergodic Theory Dynam. Systems*, 14(4):695–720, 1994.

[16] N. Jonoska. Sofic shifts with synchronizing presentations. *Theoret. Comput. Sci.*, 158(1-2):81–115, 1996.

[17] N. Jonoska and B. Marcus. Minimal presentations for irreducible sofic shifts. *IEEE Trans. Inform. Theory*, 40(6):1818–1825, 1994.

[18] Y. Katayama, K. Matsumoto, and Y. Watatani. Simple $C^*$-algebras arising from $\beta$-expansion of real numbers. *Ergodic Theory Dynam. Systems*, 18(4):937–962, 1998.

[19] W. Krieger. On sofic systems. I. *Israel J. Math.*, 48(4):305–330, 1984.

[20] A. Kumjian, D. Pask, I. Raeburn, and J. Renault. Graphs, groupoids, and Cuntz-Krieger algebras. *J. Funct. Anal.*, 144(2):505–541, 1997.

[21] D. Lind and B. Marcus. *An introduction to symbolic dynamics and coding*. Cambridge University Press, Cambridge, 1995.

[22] B. Marcus. Sofic systems and encoding data. *IEEE Trans. Inform. Theory*, 31(3):366–377, 1985.

[23] K. Matsumoto. On $C^*$-algebras associated with subshifts. *Internat. J. Math.*, 8(3):357–374, 1997.

[24] K. Matsumoto. $K$-theory for $C^*$-algebras associated with subshifts. *Math. Scand.*, 82(2):237–255, 1998.

[25] K. Matsumoto. Dimension groups for subshifts and simplicity of the associated $C^*$-algebras. *J. Math. Soc. Japan*, 51(3):679–698, 1999.
[26] K. Matsumoto. Relations among generators of $C^*$-algebras associated with subshifts. *Internat. J. Math.*, 10(3):385–405, 1999.

[27] K. Matsumoto. A simple $C^*$-algebra arising from a certain subshift. *J. Operator Theory*, 42(2):351–370, 1999.

[28] K. Matsumoto. On automorphisms of $C^*$-algebras associated with subshifts. *J. Operator Theory*, 44(1):91–112, 2000.

[29] K. Matsumoto. Stabilized $C^*$-algebras constructed from symbolic dynamical systems. *Ergodic Theory Dynam. Systems*, 20(3):821–841, 2000.

[30] K. Matsumoto. Bowen-Franks groups for subshifts and Ext-groups for $C^*$-algebras. *K-Theory*, 23(1):67–104, 2001.

[31] K. Matsumoto. $C^*$-algebras associated with presentations of subshifts. *Doc. Math.*, 7:1–30 (electronic), 2002.

[32] K. Matsumoto, Y. Watatani, and M. Yoshida. KMS states for gauge actions on $C^*$-algebras associated with subshifts. *Math. Z.*, 228(3):489–509, 1998.

[33] M. Nasu. An invariant for bounded-to-one factor maps between transitive sofic subshifts. *Ergodic Theory Dynam. Systems*, 5(1):89–105, 1985.

[34] M. Nasu. Topological conjugacy for sofic systems. *Ergodic Theory Dynam. Systems*, 6(2):265–280, 1986.

[35] B. Parry and D. Sullivan. A topological invariant of flows on 1-dimensional spaces. *Topology*, 14(4):297–299, 1975.

[36] I. Raeburn. *Graph algebras*, volume 103 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2005.

[37] K. Thomsen. On the structure of a sofic shift space. *Trans. Amer. Math. Soc.*, 356(9):3557–3619 (electronic), 2004.

[38] B. Weiss. Subshifts of finite type and sofic systems. *Monatsh. Math.*, 77:462–474, 1973.

Rune Johansen
Department of
Mathematical Sciences
University of Copenhagen
Universitetsparken 5
2100 København Ø
Denmark
Ergodic Properties and KMS Conditions
on $C^*$-Symbolic Dynamical Systems

Kengo Matsumoto

Received: March 3, 2010
Revised: March 12, 2011
Communicated by Joachim Cuntz

Abstract. A $C^*$-symbolic dynamical system $(\mathcal{A}, \rho, \Sigma)$ consists of a unital $C^*$-algebra $\mathcal{A}$ and a finite family $\{\rho_\alpha\}_{\alpha \in \Sigma}$ of endomorphisms $\rho_\alpha$ of $\mathcal{A}$ indexed by symbols $\alpha$ of $\Sigma$ satisfying some conditions. The endomorphisms $\rho_\alpha, \alpha \in \Sigma$ yield both a subshift $\Lambda_\rho$ and a $C^*$-algebra $\mathcal{O}_\rho$. We will study ergodic properties of the positive operator $\lambda_\rho = \sum_{\alpha \in \Sigma} \rho_\alpha$ on $\mathcal{A}$. We will next introduce KMS conditions for continuous linear functionals on $\mathcal{O}_\rho$ under gauge action at inverse temperature taking its value in complex numbers. We will study relationships among the eigenvectors of $\lambda_\rho$ in $\mathcal{A}^*$, the continuous linear functionals on $\mathcal{O}_\rho$ satisfying KMS conditions and the invariant measures on the associated one-sided shifts. We will finally present several examples of continuous linear functionals satisfying KMS conditions.

2000 Mathematics Subject Classification: Primary 46L55, 37B10; Secondary 37D35.
Keywords and Phrases: $C^*$-algebra, symbolic dynamics, subshift, ergodic, KMS condition, invariant measure.

1. Introduction

D. Olesen and G. K. Pedersen [37] have shown that the $C^*$-dynamical system $(\mathcal{O}_N, \alpha, R)$ for the Cuntz algebra $\mathcal{O}_N$ with gauge action $\alpha$ admits a KMS state at the inverse temperature $\gamma$ if and only if $\gamma = \log N$, and the admitted KMS state is unique. By Enomoto-Fujii-Watatani [9], the result has been generalized to the Cuntz-Krieger algebras $\mathcal{O}_A$ as $\gamma = \log r_A$, where $r_A$ is the Perron-Frobenius eigenvalue for the irreducible matrix $A$ with entries in $\{0,1\}$. These results are generalized to several classes of $C^*$-algebras having gauge actions (cf. [7], [10], [11], [15], [17], [18], [27], [35], [36], [41], etc.).
Cuntz-Krieger algebras are considered to be constructed by finite directed graphs which yield an important class of symbolic dynamics called shifts of finite type. In [29], the author has generalized the notion of finite directed graphs to a notion of labeled Bratteli diagrams having shift like maps, which we call \( \lambda \)-graph systems. A \( \lambda \)-graph system \( \mathcal{L} \) gives rise to both a subshift \( \Lambda_{\mathcal{L}} \) and a \( C^* \)-algebra \( \mathcal{O}_{\mathcal{L}} \) with gauge action. Some topological conjugacy invariants of subshifts have been studied through the \( C^* \)-algebras constructed from \( \lambda \)-graph systems \([30]\).

A \( C^* \)-symbolic dynamical system is a generalization of both a \( \lambda \)-graph system and an automorphism of a unital \( C^* \)-algebra \([31]\). It is a finite family \( \{\rho_{\alpha}\}_{\alpha \in \Sigma} \) of endomorphisms indexed by a finite set \( \Sigma \) of a unital \( C^* \)-algebra \( \mathcal{A} \) such that \( \rho_{\alpha}(Z_{\mathcal{A}}) \subset Z_{\mathcal{A}} \) for \( \alpha \in \Sigma \) and \( \sum_{\alpha \in \Sigma} \rho_{\alpha}(1) \geq 1 \) where \( Z_{\mathcal{A}} \) denotes the center of \( \mathcal{A} \).

A finite directed labeled graph \( G \) gives rise to a \( C^* \)-symbolic dynamical system \( (\mathcal{A}_G, \rho^G, \Sigma) \) such that \( \mathcal{A}_G = C^N \) for some \( N \in \mathbb{N} \). A \( \lambda \)-graph system \( \Sigma \) also gives rise to a \( C^* \)-symbolic dynamical system \( (\mathcal{A}_{\Sigma}, \rho^\Sigma, \Sigma) \) such that \( \mathcal{A}_{\Sigma} = C(\Omega_{\Sigma}) \) for some compact Hausdorff space \( \Omega_{\Sigma} \) with \( \dim \Omega_{\Sigma} = 0 \). A \( C^* \)-symbolic dynamical system \( (\mathcal{A}, \rho, \Sigma) \) yields a subshift denoted by \( \Lambda_{\rho} \) over \( \Sigma \) and a Hilbert \( C^* \)-bimodule \( (\phi_{\rho}, \mathbb{H}^\rho_{\Sigma}) \) over \( \mathcal{A} \). By using general construction of \( C^* \)-algebras from Hilbert \( C^* \)-bimodules established by M. Pimsner \([40]\), a \( C^* \)-algebra denoted by \( \mathcal{O}_{\rho} \) from \( (\phi_{\rho}, \mathbb{H}^\rho_{\Sigma}) \) has been introduced in \([31]\). The \( C^* \)-algebra \( \mathcal{O}_{\rho} \) is realized as the universal \( C^* \)-algebra generated by partial isometries \( S_{\alpha}, \alpha \in \Sigma \) and \( x \in \mathcal{A} \) subject to the relations:

\[
\sum_{\gamma \in \Sigma} S_{\gamma} S_{\gamma}^* = 1, \quad S_{\alpha} S_{\alpha}^* x = x S_{\alpha} S_{\alpha}^*, \quad S_{\alpha}^* x S_{\alpha} = \rho_\alpha(x)
\]

for all \( x \in \mathcal{A} \) and \( \alpha \in \Sigma \). We call the algebra \( \mathcal{O}_{\rho} \) the \( C^* \)-symbolic crossed product of \( \mathcal{A} \) by the subshift \( \Lambda_{\rho} \). The gauge action on \( \mathcal{O}_{\rho} \) denoted by \( \hat{\rho} \) is defined by

\[
\hat{\rho}_{z}(x) = x, \quad x \in \mathcal{A} \quad \text{and} \quad \hat{\rho}_{z}(S_{\alpha}) = z S_{\alpha}, \quad \alpha \in \Sigma
\]

for \( z \in \mathbb{C}, |z| = 1 \). If \( \mathcal{A} = C(X) \) with \( \dim X = 0 \), there exists a \( \lambda \)-graph system \( \mathcal{L} \) such that \( \Lambda_{\rho} \) is the subshift presented by \( \mathcal{L} \) and \( \mathcal{O}_{\rho} \) is the \( C^* \)-algebra \( \mathcal{O}_{\mathcal{L}} \) associated with \( \mathcal{L} \). If in particular, \( \mathcal{A} = C^N \), the subshift \( \Lambda_{\rho} \) is a sofic shift and \( \mathcal{O}_{\rho} \) is a Cuntz-Krieger algebra. If \( \Sigma = \{\alpha\} \) an automorphism \( \alpha \) of a unital \( C^* \)-algebra \( \mathcal{A} \), the \( C^* \)-algebra \( \mathcal{O}_{\rho} \) is the ordinary \( C^* \)-crossed product \( \mathcal{A} \rtimes_{\alpha} \mathbb{Z} \).

Throughout the paper, we will assume that the \( C^* \)-algebra \( \mathcal{A} \) is commutative. For a \( C^* \)-symbolic dynamical system \( (\mathcal{A}, \rho, \Sigma) \), define the positive operator \( \lambda_\rho \) on \( \mathcal{A} \) by

\[
\lambda_\rho(x) = \sum_{\alpha \in \Sigma} \rho_\alpha(x), \quad x \in \mathcal{A}.
\]

We set for a complex number \( \beta \in \mathbb{C} \) the eigenvector space of \( \lambda_\rho \)

\[
\mathcal{E}_\beta(\rho) = \{ \varphi \in \mathcal{A}^* \mid \varphi \circ \lambda_\rho = \beta \varphi \}.
\]

(1.1)

Let \( Sp(\rho) \) be the set of eigenvalues of \( \lambda_\rho \) defined by

\[
Sp(\rho) = \{ \beta \in \mathbb{C} \mid \mathcal{E}_\beta(\rho) \neq \{0\} \}.
\]

(1.2)
Let $r_\rho$ denote the spectral radius of $\lambda_\rho$ on $\mathcal{A}$. We set $T_\rho = \frac{1}{r_\rho} \lambda_\rho$. ($\mathcal{A}, \rho, \Sigma$) is said to be power-bounded if the sequence $\|T_\rho^k\|$, $k \in \mathbb{N}$ is bounded. A state $\varphi$ on $\mathcal{A}$ is said to be invariant if $\varphi \circ T_\rho = \varphi$. If an invariant state is unique, ($\mathcal{A}, \rho, \Sigma$) is said to be uniquely ergodic. If $\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} T_\rho^k(a)$ exists in $\mathcal{A}$ for $a \in \mathcal{A}$, ($\mathcal{A}, \rho, \Sigma$) is said to be mean ergodic. If there exists no nontrivial ideal of $\mathcal{A}$ invariant under $\lambda_\rho$, ($\mathcal{A}, \rho, \Sigma$) is said to be irreducible. It will be proved that a mean ergodic and irreducible ($\mathcal{A}, \rho, \Sigma$) is uniquely ergodic and power-bounded (Theorem 3.12).

Let $A = [A(i, j)]_{i,j=1}^N$ be an irreducible matrix with entries in $\{0, 1\}$, and $S_i, i = 1, \ldots, N$ be the canonical generating family of partial isometries of the Cuntz-Krieger algebra $\mathcal{O}_A$. Let $\mathcal{A}_A$ be the $C^*$-subalgebra of $\mathcal{O}_A$ generated by the projections $S_j S_i^*, i, j = 1, \ldots, N$. Put $\Sigma = \{1, \ldots, N\}$ and $\rho_\Sigma^A(x) = S_i^* x S_i, x \in \mathcal{A}_A, i \in \Sigma$. Then the triplet ($\mathcal{A}_A, \rho_\Sigma^A, \Sigma$) yields an example of $C^*$-symbolic dynamical system such that its $C^*$-symbolic crossed product $\mathcal{O}_A, \rho_\Sigma$ is the Cuntz-Krieger algebra $\mathcal{O}_A$. The above space $\mathcal{E}_\beta(\rho)$ is identified with the eigenvector space of the matrix $A$ for an eigenvalue $\beta$. By Enomoto-Fujii-Watatani [9], a tracial state $\varphi \in \mathcal{E}_\beta(\rho^A)$ on $\mathcal{A}_A$ extends to a KMS state for gauge action on $\mathcal{O}_A$ if and only if $\beta = r_\rho$ the Perron-Frobenius eigenvalue, and its inverse temperature is $\log r_\rho$. The admitted KMS state is unique.

In this paper, we will study the space $\mathcal{E}_\beta(\rho)$ of a general $C^*$-symbolic dynamical system ($\mathcal{A}, \rho, \Sigma$) for a general eigenvalue $\beta$ in $\mathbb{C}$ not necessarily maximum eigenvalue and then introduce KMS condition for inverse temperature taking its value in complex numbers. In this generalization, we will study possibility of extension of a continuous linear functional on $\mathcal{A}$ belonging to the eigenvector space $\mathcal{E}_\beta(\rho)$ to the whole algebra $\mathcal{O}_\rho$ as a continuous linear functional satisfying KMS condition. For a $C^*$-algebra with a continuous action of the one-dimensional torus group $\mathbf{T} = \mathbb{R}/2\pi \mathbb{Z}$ and a complex number $\beta$ in $\mathbb{C}$, we will introduce KMS condition for a continuous linear functional without assuming its positivity at inverse temperature $\log \beta$. Let $\mathcal{B}$ be a $C^*$-algebra and $\alpha : \mathbf{T} \rightarrow \text{Aut}(\mathcal{B})$ be a continuous action of $\mathbf{T}$ to the automorphism group Aut($\mathcal{B}$). We write a complex number $\beta$ with $|\beta| > 1$ as $\beta = r e^{i\theta}$ where $r > 1, \theta \in \mathbb{R}$. Denote by $\mathcal{B}^*$ the Banach space of all complex valued continuous linear functionals on $\mathcal{B}$.

**Definition.** A continuous linear functional $\varphi \in \mathcal{B}^*$ is said to satisfy KMS condition at $\log \beta$ if $\varphi$ satisfies the condition

$$\varphi(\alpha_t(x)y) = \varphi(\alpha_0(x)y), \quad x \in \mathcal{B}^*, y \in \mathcal{B},$$

where $\mathcal{B}^*$ is the set of analytic elements of the action $\alpha : \mathbf{T} \rightarrow \text{Aut}(\mathcal{B})$ (cf.[3]).

We will prove

**Theorem 1.1.** Let ($\mathcal{A}, \rho, \Sigma$) be an irreducible and power-bounded $C^*$-symbolic dynamical system. Let $\beta \in \mathbb{C}$ be a complex number with $|\beta| > 1$. Then

1. If $\beta \in \text{Sp}(\rho)$ and $|\beta| = r_\rho$ the spectral radius of the positive operator $\lambda_\rho : \mathcal{A} \rightarrow \mathcal{A}$, then there exists a nonzero continuous linear functional
on $O_{\rho}$ satisfying KMS condition at $\log \beta$ under gauge action. The converse implication holds if $(A, \rho, \Sigma)$ is mean ergodic.

(ii) Under the condition $|\beta| = r_{\rho}$, there exists a linear isomorphism between the space $E_{\beta}(\rho)$ of eigenvectors of continuous linear functionals on $A$ and the space $KMS_{\beta}(O_{\rho})$ of continuous linear functionals on $O_{\rho}$ satisfying KMS condition at $\log \beta$.

(iii) If $(A, \rho, \Sigma)$ is uniquely ergodic, there uniquely exists a state on $O_{\rho}$ satisfying KMS condition at $\log r_{\rho}$.

(iv) If in particular $(A, \rho, \Sigma)$ is mean ergodic, then $\dim E_{\beta}(\rho) \leq 1$ for all $\beta \in \mathbb{C}$.

In the proof of the above theorem, a Perron-Frobenius type theorem is proved (Theorem 3.13).

Let $D_{\rho}$ be the $C^*$-subalgebra of $O_{\rho}$ generated by all elements of the form: $S_{\alpha_{1}} \cdots S_{\alpha_{k}} x S_{\alpha_{k}}^{*} \cdots S_{\alpha_{1}}^{*}$ for $x \in A, \alpha_{1}, \ldots, \alpha_{k} \in \Sigma$. Let $\phi_{\rho}$ be the endomorphism on $D_{\rho}$ defined by $\phi_{\rho}(y) = \sum_{\alpha \in \Sigma} S_{\alpha} y S_{\alpha}^{*} y \in D_{\rho}$, which comes from the left-shift on the underlying shift space $\Lambda_{\rho}$. Suppose that $(A, \rho, \Sigma)$ is uniquely ergodic.

The restriction of the unique KMS state on $O_{\rho}$ is not necessarily a $\phi_{\rho}$-invariant state. We will clarify a relationship between KMS states on $O_{\rho}$ and $\phi_{\rho}$-invariant states on $D_{\rho}$ as in the following way:

**Theorem 1.2.** Assume that $(A, \rho, \Sigma)$ is irreducible and mean ergodic. Let $\tau$ be the restriction to $D_{\rho}$ of the unique KMS state on $O_{\rho}$ at $\log r_{\rho}$ and $x_{\rho}$ be a positive element of $A$ defined by the limit of the mean

$$\lim_{n \to \infty} \frac{1}{n} (1 + T_{\rho}(1) + \cdots + T_{\rho}^{n-1}(1)).$$

Let $\mu_{\rho}$ be a linear functional on $D_{\rho}$ defined by

$$\mu_{\rho}(y) = \tau(y x_{\rho}), \quad y \in D_{\rho}.$$

(i) $\mu_{\rho}$ is a faithful, $\phi_{\rho}$-invariant and ergodic state on $D_{\rho}$ in the sense that the formula

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu_{\rho}(\phi_{\rho}^{k}(y)x) = \mu_{\rho}(y) \mu_{\rho}(x), \quad x, y \in D_{\rho}$$

holds.

(ii) $\mu_{\rho}$ gives rise to a unique $\phi_{\rho}$-invariant probability measure absolutely continuous with respect to the probability measure for the state $\tau$.

(iii) $\mu_{\rho}$ is equivalent to the state $\tau$ as a measure on $D_{\rho}$.

For a $C^*$-symbolic dynamical system $(A_{A}, \rho^{A}, \Sigma)$ coming from an irreducible matrix $A = [A(i, j)]_{i,j=1}^{N}$ with entries in $\{0, 1\}$, the subalgebra $D_{\rho^{A}}$ is nothing but the commutative $C^*$-algebra $C(X_{A})$ of all continuous functions on the right one-sided topological Markov shift $X_{A}$. As $\phi_{\rho^{A}}$ corresponds to the left-shift $\sigma_A$ on $X_A$, the above unique $\phi_{\rho^{A}}$-invariant state $\tau$ is the Parry measure on $X_{A}$. The positive element $x_{\rho^{A}}$ is given by the positive Perron eigenvector.
ergodic properties of the operator $T$ by Section 4, we will study extendability of a linear functional belonging to $C$ consisting of a unital $A$ for all $x$ such that $\Phi \circ O = \rho$ called $(\sigma, \alpha)$-invariant states on $O$ to the subalgebra $D$ of $O$, which will extend to $O$. In Section 5, we will prove Theorem 1.1. In Section 6, we will study a relationship between KMS states and $\phi_\beta$-invariant states on $O$, to prove Theorem 1.2. In Section 7, we will present several examples of continuous linear functionals on $O_\beta$ satisfying KMS conditions.

2. C*-symbolic dynamical systems and their crossed products

Let $A$ be a unital C*-algebra. In what follows, an endomorphism of $A$ means a $*$-endomorphism of $A$ that does not necessarily preserve the unit 1 of $A$. Denote by $Z_A$ the center $\{x \in A \mid ax = xa \text{ for all } a \in A\}$ of $A$. Let $\Sigma$ be a finite set. A finite family of nonzero endomorphisms $\rho_\alpha, \alpha \in \Sigma$ of $A$ indexed by elements of $\Sigma$ is said to be essential if $\rho_\alpha(Z_A) \subset Z_A$ for $\alpha \in \Sigma$ and $\sum_{\alpha \in \Sigma} \rho_\alpha(1) \geq 1$. If in particular, $A$ is commutative, the family $\rho_\alpha, \alpha \in \Sigma$ is essential if and only if $\sum_{\alpha \in \Sigma} \rho_\alpha(1) \geq 1$. We remark that the definition in [31] of “essential” for $\rho_\alpha, \alpha \in \Sigma$ is weaker than the above definition. It is said to be faithful if for any nonzero $x \in A$ there exists a symbol $\alpha \in \Sigma$ such that $\rho_\alpha(x) \neq 0$.

Definition ([31]). A C*-symbolic dynamical system is a triplet $(A, \rho, \Sigma)$ consisting of a unital C*-algebra $A$ and an essential, faithful finite family $\{\rho_\alpha\}_{\alpha \in \Sigma}$ of endomorphisms of $A$.

Two C*-symbolic dynamical systems $(A, \rho, \Sigma)$ and $(A', \rho', \Sigma')$ are said to be isomorphic if there exist an isomorphism $\Phi : A \rightarrow A'$ and a bijection $\pi : \Sigma \rightarrow \Sigma'$ such that $\Phi \circ \rho_\alpha = \rho'_{\pi(\alpha)} \circ \Phi$ for all $\alpha \in \Sigma$. For an automorphism $\alpha$ of a unital C*-algebra $A$, by setting $\Sigma = \{\alpha\}, \rho_\alpha = \alpha$ the triplet $(A, \rho, \Sigma)$ becomes a C*-symbolic dynamical system. A C*-symbolical dynamical system $(A, \rho, \Sigma)$ yields a subshift $\Lambda_\rho$ over $\Sigma$ such that a word $\alpha_1 \cdots \alpha_k$ of $\Sigma$ is admissible for $\Lambda_\rho$ if and only if $\rho_{\alpha_1} \circ \cdots \circ \rho_{\alpha_k} \neq 0$ ([31, Proposition 2.1]). We say that a subshift $\Lambda$ acts on a C*-algebra $A$ if there exists a C*-symbolical dynamical system $(A, \rho, \Sigma)$ such that the associated subshift $\Lambda_\rho$ is $\Lambda$.

For a C*-symbolical dynamical system $(A, \rho, \Sigma)$ the C*-algebra $O_\rho$ has been originally constructed in [31] as a C*-algebra from a Hilbert C*-bimodule by using a Pimsner’s general construction of Hilbert C*-bimodule algebras [40] (cf. [16] etc.). It is called the C*-symbolic crossed product of $A$ by the subshift $\Lambda_\rho$, and realized as the universal C*-algebra $C^*(x, S_\alpha ; x \in A, \alpha \in \Sigma)$ generated by $x \in A$ and partial isometries $S_\alpha, \alpha \in \Sigma$ subject to the following relations called $(\rho)$:

$$\sum_{\gamma \in \Sigma} S_\gamma S_\gamma^* = 1, \quad S_\alpha S_\alpha^* x = x S_\alpha S_\alpha^*, \quad S_\alpha^* x S_\alpha = \rho_\alpha(x)$$

for all $x \in A$ and $\alpha \in \Sigma$. 
Let $G = (G, \lambda)$ be a left-resolving finite labeled graph with underlying finite directed graph $G = (V, E)$ and labeling map $\lambda : E \to \Sigma$ (see [28, p.76]). Denote by $v_1, \ldots, v_N$ the vertex set $V$. Assume that every vertex has both an incoming edge and an outgoing edge. Consider the $N$-dimensional commutative $C^*$-algebra $A_G = CE_1 \oplus \cdots \oplus CE_N$ where each minimal projection $E_i$ corresponds to the vertex $v_i$ for $i = 1, \ldots, N$. Define an $N \times N$-matrix for $\alpha \in \Sigma$ by

$$A^G(i, \alpha, j) = \begin{cases} 1 & \text{if there exists an edge } e \text{ from } v_i \text{ to } v_j \text{ with } \lambda(e) = \alpha, \\ 0 & \text{otherwise} \end{cases}$$

(2.1)

for $i, j = 1, \ldots, N$. We set $\rho^G_\alpha(E_i) = \sum_{j=1}^N A^G(i, \alpha, j)E_j$ for $i = 1, \ldots, N$. Then $\rho^G_\alpha, \alpha \in \Sigma$ define endomorphisms of $A_G$ such that $(A_G, \rho^G_\alpha, \Sigma)$ is a $C^*$-symbolic dynamical system for which the subshift $\Lambda_0^G$ is the sofic shift presented by $G$. Conversely, for a $C^*$-symbolic dynamical system $(A, \rho, \Sigma)$, if $A$ is $C^\infty$, there exists a left-resolving labeled graph $G$ such that $A = A_G$ and $\Lambda_0 = \Lambda_G$ the sofic shift presented by $G$ ([31, Proposition 2.2]). Put $A_G(i, j) = \sum_{\alpha \in \Sigma} A^G(i, \alpha, j), i, j = 1, \ldots, N$. The $N \times N$ matrix $A_G = [A^G(i, j)]_{i,j=1,\ldots,N}$ is called the underlying nonnegative matrix for $G$. Consider the matrix $A^{[2]}_G = [A^{[2]}_G(e, f)]_{e,f \in E}$ indexed by edges $E$ whose entries are in $\{0, 1\}$ by setting

$$A^{[2]}_G(e, f) = \begin{cases} 1 & \text{if } f \text{ follows } e, \\ 0 & \text{otherwise}. \end{cases}$$

(2.2)

The $C^*$-algebra $O_{\rho^G}$ for the $C^*$-symbolic dynamical system $(A_G, \rho^G, \Sigma)$ is the Cuntz-Krieger algebra $O_{A_G^\Sigma}$ (cf. [30, Proposition 7.1], [1]).

More generally let $\Sigma$ be a $\lambda$-graph system $(V, E, \lambda, \iota)$ over $\Sigma$. We equip each vertex set $V_l$ with discrete topology. We denote by $\Omega_\Sigma$ the compact Hausdorff space with $\dim \Omega_\Sigma = 0$ of the projective limit $V_0 \leftarrow V_1 \leftarrow V_2 \leftarrow \cdots$ as in [30, Section 2]. Since the algebra $C(V_l)$ denoted by $A_{E_l,1}$ of all continuous functions on $V_l$ is the commutative finite dimensional algebra, the commutative $C^*$-algebra $C(\Omega_\Sigma)$ is an AF-algebra, that is denoted by $A_{E_\Sigma}$. We then have a $C^*$-symbolic dynamical system $(A_{E_\Sigma}, \rho^E_\Sigma)$ such that the subshift $\Lambda_{\rho^E_\Sigma}$ coincides with the subshift $A_{E_\Sigma}$ presented by $\Sigma$. Conversely, for a $C^*$-symbolic dynamical system $(A, \rho, \Sigma)$, if the algebra $A$ is $C(X)$ with $\dim X = 0$, there exists a $\lambda$-graph system $\Sigma$ over $\Sigma$ such that the associated $C^*$-symbolic dynamical system $(A_{E_\Sigma}, \rho^E_\Sigma)$ is isomorphic to $(A, \rho, \Sigma)$ ([31, Theorem 2.4]). The $C^*$-algebra $O_{\rho^E_\Sigma}$ is the $C^*$-algebra $O_{\Sigma}$ associated with the $\lambda$-graph system $\Sigma$.

Let $\alpha$ be an automorphism of a unital $C^*$-algebra $A$. Put $\Sigma = \{\alpha\}$ and $\rho_\alpha = \alpha$. The $C^*$-algebra $O_{\rho}$ for the $C^*$-symbolic dynamical system $(A, \rho, \Sigma)$ is the ordinary $C^*$-crossed product $A \times_\alpha \mathbb{Z}$.

In what follows, for a subset $F$ of a $C^*$-algebra $B$, we will denote by $C^*(F)$ the $C^*$-subalgebra of $B$ generated by $F$.

Let $(A, \rho, \Sigma)$ be a $C^*$-symbolic dynamical system over $\Sigma$ and $\Lambda$ the associated subshift $\Lambda_\rho$. We denote by $B_\mu(\Lambda)$ the set of admissible words $\mu$ of $\Lambda$ with length
|µ| = k. Put B∗(A) = ∪∞n=k Bn(A), where B0(A) consists of the empty word.
Let S, α ∈ Σ be the partial isometries in Oρ satisfying the relation (ρ). For
µ = (µ1, . . . , µk) ∈ Bk(A), we put Sµ = Sµ1 · · · Sµk and ρµ = ρµk · · · ρµ1. In
the algebra Oρ, we set for k ∈ ℤ+,

\[ D^k_ρ = C^∗(S_µxS_ν^* : µ ∈ B_k(A), x ∈ A), \]

\[ D_ρ = C^∗(S_µxS_ν^* : µ ∈ B_1(A), x ∈ A), \]

\[ F^k_ρ = C^∗(S_µxS_ν^* : µ, ν ∈ B_k(A), x ∈ A) \quad \text{and} \]

\[ F_ρ = C^∗(S_µxS_ν^* : µ, ν ∈ B_1(A), |µ| = |ν|, x ∈ A). \]

The identity S_µxS_ν^* = ∑α∈Σ S_µαρ_α(x)S_ν^* for x ∈ A, µ, ν ∈ Bk(A) holds so
that the algebra F^k_ρ is embedded into the algebra F^{k+1}_ρ such that ∪k∈Z F^k_ρ is
dense in F_ρ. Similarly D^k_ρ is embedded into the algebra D^{k+1}_ρ such that
∪k∈Z D^k_ρ is dense in D_ρ. The gauge action ˆρ of the one-dimensional torus
group T = \{z ∈ C | |z| = 1\} on Oρ is defined by ˆρz(x) = x for x ∈ A and
ˆρz(S_α) = zS_α for α ∈ Σ. The fixed point algebra of Oρ under ˆρ is denoted by
(Oρ)^{ˆρ}. Let E_ρ : Oρ → (Oρ)^{ˆρ} be the conditional expectation defined by

\[ E_ρ(X) = \int_{z ∈ T} ˆρz(X)dz, \quad X ∈ Oρ. \quad (2.3) \]

It is routine to check that (Oρ)^{ˆρ} = F_ρ.

DEFINITION ([33]). A C^∗-symbolic dynamical system (A, ρ, Σ) satisfies condition (I) if there exists a unital increasing sequence

A_0 ⊂ A_1 ⊂ · · · ⊂ A

of C∗-subalgebras of A such that ρ_α(A_i) ⊂ A_{i+1} for all l ∈ Z, α ∈ Σ, the union
∪l∈Z A_l is dense in A and for e > 0, k, l ∈ N with k ≤ l and X_0 ∈ F^l_ρ =
C^∗(S_µxS_ν^* : µ, ν ∈ B_l(A), x ∈ A_i), there exists an element g ∈ D_ρ ∩ A_i′ (= \{y ∈
D_ρ | yα = ay for α ∈ A_i\}) with 0 ≤ e ≤ 1 such that

(i) \|X_0φ^k_ρ(g)\| ≥ \|X_0\| − e,

(ii) φ^m_ρ(g) = 0 for all m = 1, 2, . . . , k, where φ^m_ρ(X) = ∑µ∈B_m(A) S_µXS_µ^*.

As the element g belongs to the diagonal subalgebra F_ρ of F_ρ, the condition
(I) is intrinsically determined by (A, ρ, Σ) by virtue of [31, Lemma 4.1]. The
condition (I) for (A, ρ, Σ) yields the uniqueness of the C^∗-algebra O_ρ under the
relations (ρ) ([33]).

If a λ-graph system L over Σ satisfies condition (I), then (A_L, ρ^L, Σ) satisfies
condition (I) (cf. [30, Lemma 4.1]).

Recall that the positive operator λ_ρ : A → A is defined by λ_ρ(x) =
∑α∈Σ ρ_α(x), x ∈ Σ. Then a C∗-symbolic dynamical system (A, ρ, Σ) is said to
be irreducible, if there exists no nontrivial ideal of A invariant under λ_ρ. It
has been shown in [31] that if (A, ρ, Σ) satisfies condition (I) and is irreducible,
then the C^∗-algebra O_ρ is simple.

Interesting examples of (A, ρ, Σ) in [31], [34] which we have seen from the
view point of symbolic dynamics come from ones for which A is commutative.
Hence we assume that the algebra A is commutative so that A is written as
3. Ergodicity and Perron-Frobenius type theorem

In this section, we will study ergodic properties of a C*-symbolic dynamical system \((\mathcal{A}, \rho, \Sigma)\) and prove a Perron-Frobenius type theorem.

Let \(\mathcal{A}^*\) denote the Banach space of all complex valued continuous linear functionals on \(\mathcal{A}\). For \(\beta \in \mathbb{C}\) with \(\beta \neq 0\), set

\[
\mathcal{E}_\beta(\rho) = \{ \varphi \in \mathcal{A}^* \mid \varphi \circ \lambda_\rho(a) = \beta \varphi(a) \text{ for all } a \in \mathcal{A} \}.
\]

It is possible that \(\mathcal{E}_\beta(\rho)\) is \(\{0\}\). A nonzero continuous linear functional \(\varphi\) in \(\mathcal{E}_\beta(\rho)\) is called an eigenvector of the operator \(\lambda_\rho^\ast\) with respect to the eigenvalue \(\beta\). Let \(r_\rho\) be the spectral radius of the positive operator \(\lambda_\rho : \mathcal{A} \to \mathcal{A}\). Since \(\lambda_\rho^k(1) \geq 1, k \in \mathbb{N}\), one sees that \(r_\rho \geq 1\). As \(\text{Sp}(\lambda_\rho) = \text{Sp}(\lambda_\rho^*)\) (cf. \([8, VI. 2.7]\)), we note \(r_\rho = r(\lambda_\rho^*)\). Let \(S(\mathcal{A})\) denote the state space of \(\mathcal{A}\).

**Lemma 3.1.** \((\mathcal{A}, \rho, \Sigma)\) is irreducible if and only if for a state \(\varphi\) on \(\mathcal{A}\) and a nonzero element \(x \in \mathcal{A}\), there exists a natural number \(n\) such that \(\varphi(\lambda_\rho^n(x^*x)) > 0\).

**Proof.** Suppose that \((\mathcal{A}, \rho, \Sigma)\) is irreducible. For a state \(\varphi\) on \(\mathcal{A}\), put

\[
I_\varphi = \{ x \in \mathcal{A} \mid \varphi(\lambda_\rho^n(x^*x)) = 0 \text{ for all } n \in \mathbb{N} \}
\]

which is an ideal of \(\mathcal{A}\) because \(\mathcal{A}\) is commutative. The Schwarz type inequality

\[
\lambda_\rho^n(\lambda_\rho(x)^*\lambda_\rho(x)) \leq \|\lambda_\rho\|\lambda_\rho^{n+1}(x^*x) \quad \text{for} \quad x \in \mathcal{A}
\]

implies that \(I_\varphi\) is \(\lambda_\rho\)-invariant. Hence \(I_\varphi\) is trivial.

Conversely, let \(I\) be an ideal of \(\mathcal{A}\) invariant under \(\lambda_\rho\). Put \(\mathcal{B} = \mathcal{A}/I\). Denote by \(q : \mathcal{A} \to \mathcal{B}\) the quotient map. Take \(\psi \in S(\mathcal{B})\) a state. Put \(\varphi = \psi \circ q\). For \(y \in I\), as \(\varphi(\lambda_\rho^n(y^*y)) = 0, n \in \mathbb{N}\), one sees that \(y = 0\) and hence \(I = \{0\}\) by the hypothesis. Hence \((\mathcal{A}, \rho, \Sigma)\) is irreducible.

We denote by \(T_\rho : \mathcal{A} \to \mathcal{A}\) the positive operator \(\frac{1}{r_\rho} \lambda_\rho\). The spectral radius of \(T_\rho\) is \(1\). A state \(\tau\) on \(\mathcal{A}\) is called an invariant state if \(\tau \circ T_\rho = \tau\) on \(\mathcal{A}\), equivalently \(\tau \in \mathcal{E}_{T_\rho}(\rho)\).

**Corollary 3.2.** Suppose that \((\mathcal{A}, \rho, \Sigma)\) is irreducible. Then any positive eigenvector of \(\lambda_\rho^\ast\) for a nonzero eigenvalue is faithful.

**Proof.** Let \(\varphi \in \mathcal{E}_\beta(\rho)\) be a positive linear functional for some nonzero \(\beta \in \mathbb{C}\). Since \(\varphi(\lambda_\rho(1)) = \beta \varphi(1)\), one has \(\beta > 0\). By the preceding lemma, one has \(\varphi(x^*x) > 0\) for nonzero \(x \in \mathcal{A}\).}

Yasuo Watatani has kindly informed to the author that the lemma below, which is seen from \([41, Theorem 2.5]\), is needed in the proof of Lemma 3.4. In our restrictive situation, we may directly prove it as in the following way.
Lemma 3.3. The spectral radius $r_p$ of the operator $\lambda_p$ is contained in the spectrum $\text{Sp}(\lambda_p)$ of $\lambda_p$.

Proof. The resolvent $R(z) = (z - \lambda_p)^{-1}$ for $\lambda_p$ has the expansion $R(z) = \sum_{n=0}^{\infty} \frac{\lambda_p^n}{z^{n+1}}$ for $z \in \mathbb{C}, |z| > r_p$ which converges in norm. We note that the family $\{ R(z) \}_{z > r_p}$ is not uniformly bounded. Otherwise, there exists a constant $M > 0$ such that $\|R(z)\| < M$ for $z \in \mathbb{C}, |z| > r_p$. By the compactness of $\text{Sp}(\lambda_p)$, we may find $z_0 \in \text{Sp}(\lambda_p)$ with $|z_0| = r_p$. Take $z_n \notin \text{Sp}(\lambda_p)$ satisfying $\lim_{n \to \infty} z_n = z_0$ and $|z_n| > r_p$. The resolvent equation $R(z_n) - R(z_n)R(z_m) = (z_n - z_m)R(z_m)$ implies the inequality $\|R(z_n) - R(z_m)\| \leq |z_n - z_m|M^2$ so that there exists a bounded linear operator $R_0 = \lim_{n \to \infty} R(z_n)$ on $\mathcal{A}$. The equality $(z_n - \lambda_p)R(z_n)x = x, x \in \mathcal{A}$ implies $(z_0 - \lambda_p)R_0x = x, x \in \mathcal{A}$ and hence $z_0 \notin \text{Sp}(\lambda_p)$ a contradiction. Thus there exists $r_n \in \mathbb{C}$ such that $|r_n| \notin \text{Sp}(\lambda_p)$ and $|r_n| \downarrow r_p$ and $\lim_{n \to \infty} \|R(r_n)f\| = \infty$ for some $f \in \mathcal{A}$. We may assume that $\rho > 0$. For a state $\varphi$ on $\mathcal{A}$, one has

$$\|\varphi(R(r_n)f)\| \leq \sum_{k=0}^{\infty} \frac{\varphi(z_n^k)}{|r_n|^{k+1}} = \varphi(R(|r_n|)f).$$

Denote by $w(y)$ the numerical radius of an element $y \in \mathcal{A}$, which is defined by

$$w(y) = \sup\{\varphi(y) \mid \varphi \in \text{S}(\mathcal{A})\}.$$ 

As the inequalities $\frac{1}{2}\|y\| \leq w(y) \leq \|y\|$ always hold (cf. [13, p.95]), one sees

$$\frac{1}{2}\|R(r_n)f\| \leq w(R(r_n)f) \leq w(R(|r_n|)f) \leq \|R(|r_n|)f\|$$

so that

$$\lim_{n \to \infty} \|R(|r_n|)f\| = \infty.$$ 

If $r_p \notin \text{Sp}(\lambda_p)$, the condition $|r_n| \notin \text{Sp}(\lambda_p)$ means that $R(|r_n|) \uparrow R(r_p)$ because $R(z)$ increases for $z \downarrow r_p$. Hence $\lim_{n \to \infty} \|R(|r_n|)f\| = \|R(r_p)f\| < \infty$, a contradiction. Therefore we conclude $r_p \in \text{Sp}(\lambda_p)$. \hfill \Box

The following lemma is crucial.

Lemma 3.4. Suppose that $(\mathcal{A}, \rho, \Sigma)$ is irreducible. Then there exists a faithful invariant state on $\mathcal{A}$.

Proof. We denote by $R^r(t)$ the resolvent of $\lambda^r_p : \mathcal{A}^* \to \mathcal{A}$ defined by $R^r(t)\varphi = (t - \lambda^r_p)^{-1}\varphi$ for $\varphi \in \mathcal{A}^*, t > r(\lambda^r_p)$. As $r_p = r(\lambda^r_p)$, there exists $\varphi_0 \in \mathcal{A}^*$ such that $\|R^r(t)\varphi_0\|$ is unbounded for $t \downarrow r_p$ by Lemma 3.3. We may assume that $\varphi_0$ is a state on $\mathcal{A}$. Put

$$\varphi_n = \frac{R^r(t_p + \frac{1}{n})\varphi_0}{\|R^r(t_p + \frac{1}{n})\varphi_0\|} \quad \text{for} \quad n = 1, 2, \ldots.$$ 

Since $R^r(t)$ is positive for $t > r_p$, each $\varphi_n$ is a state on $\mathcal{A}$ so that there exists a weak* cluster point $\varphi_\infty \in \mathcal{S}(\mathcal{A})$ of the sequence $\{\varphi_n\}$ in $\mathcal{S}(\mathcal{A})$. As we see

$$(r_p - \lambda^r_p)\varphi_n = -\frac{1}{n} \varphi_n + \frac{\varphi_0}{\|R^r(t_p + \frac{1}{n})\varphi_0\|}.$$
we get \( r_\rho \varphi_\infty = \lambda_\rho \varphi_\infty \) so that \( \varphi_\infty \in E_{r_\rho}(\rho) \). By Corollary 3.2, one knows that \( \varphi_\infty \) is faithful on \( \mathcal{A} \).

**Definition.** A \( C^* \)-symbolic dynamical system \((\mathcal{A}, \rho, \Sigma)\) is said to be **uniquely ergodic** if there exists a unique invariant state on \( \mathcal{A} \). Denote by \( \tau \) the unique invariant state.

If \((\mathcal{A}, \rho, \Sigma)\) is irreducible and uniquely ergodic, the unique invariant state \( \tau \) is automatically faithful because any invariant state is faithful.

There is an example of a \( C^* \)-symbolic dynamical system \((\mathcal{A}, \rho, \Sigma)\) for which a unique invariant state is not faithful, unless \((\mathcal{A}, \rho, \Sigma)\) is irreducible. Let \( \mathcal{A} = \mathbb{C} \oplus \mathbb{C}, \Sigma = \{1, 2\} \) and \( \rho_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \rho_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \). Then \( \lambda_\rho = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \)

\( r_\rho = 2 \) and \( T_\rho = \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix} \). The vector \( \begin{bmatrix} 1 \\ 0 \end{bmatrix} \) is a unique invariant state on \( \mathcal{A} \), that is not faithful.

We will see, in Section 7, that the \( C^* \)-symbolic dynamical system \((\mathcal{A}_G, \rho^G, \Sigma)\) for a finite labeled graph \( G \) is uniquely ergodic if and only if the underlying nonnegative matrix \( \mathcal{A}_G \) is irreducible.

We will next consider the eigenvector space of the operator \( \lambda_\rho \) on \( \mathcal{A} \). We are assuming that the algebra \( \mathcal{A} \) is commutative so that \( \mathcal{A} \) is written as \( C(\Omega) \) for some compact Hausdorff space \( \Omega \).

**Lemma 3.5.** Assume that \((\mathcal{A}, \rho, \Sigma)\) is irreducible.

(i) If \( T_\rho \) has a nonzero fixed element in \( \mathcal{A} \), then \( T_\rho \) has a nonzero positive fixed element in \( \mathcal{A} \).

(ii) A nonzero positive fixed element by \( T_\rho \) in \( \mathcal{A} \) must be strictly positive.

(iii) If there exist two nonzero positive fixed elements by \( T_\rho \) in \( \mathcal{A} \), then one is a scalar multiple of the other.

(iv) The dimension of the space consisting of the fixed elements by \( T_\rho \) is at most one.

**Proof.**

(i) Let \( y \in \mathcal{A} \) be a nonzero fixed element by \( T_\rho \). Since \( y^+ \) is also fixed by \( T_\rho \), we may assume that \( y = y^+ \). Denote by \( y = y^+ - y^- \) with \( y^+, y^- \geq 0 \) the Jordan decomposition of \( y \). We have \( y^+ \geq y \) and hence \( T_\rho(y^+) \geq T_\rho(y) = y \).

As \( T_\rho(y^+) \geq 0 \), one sees that \( T_\rho(y^+) \geq y^+ \). Now \((\mathcal{A}, \rho, \Sigma)\) is irreducible so that there exists a faithful invariant state \( \tau \) on \( \mathcal{A} \). Since \( \tau(T_\rho(y^+) - y^+) = 0 \), one has \( T_\rho(y^+) = y^+ \). Similarly we have \( T_\rho(y^-) = y^- \). As \( y \neq 0 \), either \( y^+ \) or \( y^- \) is not zero.

(ii) Let \( y \in \mathcal{A} \) be a nonzero fixed positive element by \( T_\rho \). Suppose that there exists \( \omega_0 \in \Omega \) such that \( y(\omega_0) = 0 \). Let \( I_y \) be the closed ideal of \( \mathcal{A} \) generated by \( y \). For a nonzero positive element \( f \in \mathcal{A} \) we have

\[ T_\rho(fy) \leq ||f||T_\rho(y) = ||f||y \]

so that \( T_\rho(fy) \) belongs to \( I_y \). As the ideal \( I_y \) is approximated by linear combinations of the elements of the form \( fy, f \in \mathcal{A}, f \geq 0 \), the ideal \( I_y \) is invariant under \( T_\rho \). Now \((\mathcal{A}, \rho, \Sigma)\) is irreducible so that \( I_y = \mathcal{A} \). As any element of \( \mathcal{A} \) vanishes at \( \omega_0 \), a contradiction.

<Documenta Mathematica 16 (2011) 133–175>
(iii) Let $x, y \in \mathcal{A}$ be nonzero positive fixed elements by $T_\rho$. By the above discussions, they are strictly positive. Set $c_0 = \min \{ \frac{\omega(x)}{\omega(y)} \mid \omega \in \Omega \}$. The function $x - c_0y$ is positive element but not strictly positive, so that it must be zero.

(iv) Let $y \in \mathcal{A}$ be a fixed element under $T_\rho$, which is written as the Jordan decomposition $y = y_1 - y_2 + i(y_3 - y_4)$ for some positive elements $y_i, i = 1, 2, 3, 4$ in $\mathcal{A}$. By the above discussions, all the elements $y_i, i = 1, 2, 3, 4$ are fixed under $T_\rho$ and they are strictly positive if it is nonzero. Hence (iii) implies the desired assertion.

**Definition.** A $C^*$-symbolic dynamical system $(\mathcal{A}, \rho, \Sigma)$ is said to satisfy (FP) if there exists a nonzero fixed element in $\mathcal{A}$ under $T_\rho$. If in particular, $(\mathcal{A}, \rho, \Sigma)$ is irreducible, a nonzero fixed element can be taken as a strictly positive element in $\mathcal{A}$ by the previous lemma.

**Lemma 3.6.** Assume that $(\mathcal{A}, \rho, \Sigma)$ is irreducible and satisfies (FP).

(i) If there exists a state $\psi$ in $E_\rho(\rho)$ for some $\beta \in \mathbb{C}$ with $\beta \neq 0$, then we have $\beta = r_\rho$.

(ii) If in particular, $(\mathcal{A}, \rho, \Sigma)$ is uniquely ergodic, the eigenspace $E_{r_\rho}(\rho)$ is of one-dimensional.

**Proof.** (i) Suppose that there exists a state $\psi$ in $E_\rho(\rho)$ for some $\beta \in \mathbb{C}$ with $\beta \neq 0$. Let $x_0 \in \mathcal{A}$ be a nonzero fixed element by $T_\rho$. One may take it to be strictly positive by the preceding lemma. Since $\lambda_\rho(x_0) = r_\rho x_0$, one has

$$\beta \psi(x_0) = \psi(\lambda_\rho(x_0)) = r_\rho \psi(x_0).$$

By Corollary 3.2, one has $\psi(x_0) > 0$ so that $\beta = r_\rho$.

(ii) Take an arbitrary $\varphi \in E_{r_\rho}(\rho)$. Put $\varphi^*(x) = \varphi(x^*), x \in \mathcal{A}$ and hence $\varphi^* \in E_{r_\rho}(\rho)$. Both of the continuous linear functionals $\varphi_{\text{Re}} = \frac{1}{2}(\varphi + \varphi^*)$ and $\varphi_{\text{Im}} = \frac{1}{2i}(\varphi - \varphi^*)$ belong to $E_{r_\rho}(\rho)$ which come from real valued measures on $\Omega$. Put $\psi = \varphi_{\text{Re}}$. Let $\psi = \psi_+ - \psi_-$ be the Jordan decomposition of $\psi$, where $\psi_+, \psi_- \in \mathcal{A}$ are positive linear functionals on $\mathcal{A}$, $\psi_+ \geq \psi_-$, $\psi_+ \geq \psi_-$ are positive linear functionals on $\mathcal{A}$. Since $\psi_+ \geq \psi_-$, one has $T_\rho \psi_+ = T_\rho \psi_+ = T_\rho \psi_+ = \psi$. As $T_\rho \psi_+$ is positive, one has $T_\rho \psi_+ > \psi_+$. Now $(\mathcal{A}, \rho, \Sigma)$ is irreducible and satisfies (FP) so that one finds a strictly positive element $x_0 \in \mathcal{A}$ fixed by $T_\rho$. Then $\hat{\psi} = T_\rho \psi_+ - \psi_+$ is a positive linear functional satisfying $\hat{\psi}(x_0) = 0$. It follows that $\hat{\psi} = 0$ so that $T_\rho \psi_+ = \psi_+$. Similarly we have $T_\rho \psi_- = \psi_-$. As both $\psi_+, \psi_- \in \mathcal{A}$ are positive linear functionals on $\mathcal{A}$, the unique ergodicity asserts that there exist $0 \leq c_+, c_- \in \mathbb{R}$ such that $\psi_+ = c_+ \tau, \psi_- = c_- \tau$. By putting $c_{\text{Re}} = c_+ - c_-$, one has $\varphi_{\text{Re}} = c_{\text{Re}} \tau$ and similarly $\varphi_{\text{Im}} = c_{\text{Im}} \tau$. Therefore we have

$$\varphi = (c_{\text{Re}} + ic_{\text{Im}}) \tau.$$ 

Hence any continuous linear functional fixed by $T_\rho$ is a scalar multiple of $\tau$, so that

$$\dim E_{r_\rho}(\rho) = 1.$$
A $C^*$-symbolic dynamical system $(\mathcal{A}, \rho, \Sigma)$ is said to be power-bounded if the sequence $\{\|T^n\| \mid n \in \mathbb{N}\}$ is bounded. As $T_p^k : \mathcal{A} \rightarrow \mathcal{A}$ is completely positive, the equalities $\|T_p^k\| = |T_p^k(1)| = \frac{1}{\varphi_p} \sum_{\mu \in B_k(\Lambda)} \rho_p(1)$ hold. We remark that for an irreducible matrix $A = [\sigma(i,j)]_{i,j=1}^N$ with entries in $\{0,1\}$, the associated $C^*$-symbolic dynamical system $(\mathcal{A}, \rho^A, \Sigma)$ defined in the Cuntz-Krieger algebra $\mathcal{O}_\Lambda$ is power-bounded. One indeed sees that there is a constant $d > 0$ such that
\[
\sum_{i,j=1}^N A^k(i,j) \leq d \cdot r^k_A \quad \text{(cf. [28, Proposition 4.2.1])}.
\]
Hence
\[
\|\lambda^k_A(1)\| = \max_i \sum_{j=1}^N A^k(i,j) \leq d \cdot r^k_A.
\]

**Lemma 3.7.** Assume that $(\mathcal{A}, \rho, \Sigma)$ is irreducible. If $(\mathcal{A}, \rho, \Sigma)$ satisfies (FP), then $(\mathcal{A}, \rho, \Sigma)$ is power-bounded.

**Proof.** As $(\mathcal{A}, \rho, \Sigma)$ is irreducible and satisfies (FP), there exists a strictly positive fixed element $x_0$ of $\mathcal{A}$ under $T_p$. Since $\Omega$ is compact, one finds positive constants $c_1, c_2$ such that $0 < c_1 < x_0(\omega) < c_2$ for all $\omega \in \Omega$. It follows that
\[
c_1 T^n_p(1) = T^n_p(c_1) \leq T^n_p(x_0) = x_0 \leq c_2, \quad n \in \mathbb{N}.
\]
Thus we have $\|T_p^n\| = \|T_p^n(1)\| \leq \frac{c_2}{c_1}$ for $n \in \mathbb{N}$. \hfill $\square$

We define the mean operator $M_n : \mathcal{A} \rightarrow \mathcal{A}$ for $n \in \mathbb{N}$ by setting
\[
M_n(a) = \frac{a + T_p(a) + T_p^2(a) + \cdots + T_p^{n-1}(a)}{n}, \quad a \in \mathcal{A}. \quad (3.1)
\]

**Definition.** A $C^*$-symbolic dynamical system $(\mathcal{A}, \rho, \Sigma)$ is said to be mean ergodic if for $a \in \mathcal{A}$ the limit $\lim_{n \rightarrow \infty} M_n(a)$ exists in $\mathcal{A}$ under norm-topology. For a mean ergodic $(\mathcal{A}, \rho, \Sigma)$, the limit $\lim_{n \rightarrow \infty} M_n(1)$ exists in $\mathcal{A}$ under norm-topology, which we denote by $x_\rho \in \mathcal{A}$.

**Lemma 3.8.** Assume that $(\mathcal{A}, \rho, \Sigma)$ is irreducible. For a mean ergodic $(\mathcal{A}, \rho, \Sigma)$, we have for $a \in \mathcal{A}$,
\[
\lim_{n \rightarrow \infty} M_n(a) = \lim_{n \rightarrow \infty} M_n(T_p(a)) = \lim_{n \rightarrow \infty} T_p(M_n(a)). \quad (3.2)
\]
In particular $x_\rho$ is a nonzero positive element which satisfies $x_\rho = T_p(x_\rho)$ and $\tau(x_\rho) = 1$ for an invariant state $\tau \in \mathcal{E}_{r_p}(\rho)$.

**Proof.** For $a \in \mathcal{A}$, the equality $T_p(M_n(a)) = M_n(T_p(a))$ is clear. As
\[
(n + 1)M_{n+1}(a) - nM_n(a) = T_p^n(a),
\]
on one has
\[
\frac{1}{n} T_p^n(a) = M_{n+1}(a) - M_n(a) + \frac{1}{n} M_{n+1}(a)
\]
so that \( \lim_{n \to \infty} \frac{1}{n} T^n_\rho(a) = 0 \). By the equality
\[
T_\rho(M_n(a)) - M_n(a) = \frac{1}{n} (T^n_\rho(a) - a)
\]
we have
\[
\lim_{n \to \infty} (T_\rho(M_n(a)) - M_n(a)) = \lim_{n \to \infty} \frac{1}{n} (T^n_\rho(a) - a) = 0.
\]
Take a faithful invariant state \( \tau \) on \( \mathcal{A} \), we have
\[
\tau(x_p) = \lim_{n \to \infty} \tau(M_n(1)) = \tau(1) = 1.
\]

**Proposition 3.9.** Assume that \((\mathcal{A}, \rho, \Sigma)\) is irreducible. If \((\mathcal{A}, \rho, \Sigma)\) is mean ergodic, there exists a faithful invariant state \( \tau \) on \( \mathcal{A} \) such that
\[
\lim_{n \to \infty} M_n(a) = \tau(a)x_\rho, \quad a \in \mathcal{A}.
\]

**Proof.** For \( a \in \mathcal{A} \), the limit \( \Phi(a) = \lim_{n \to \infty} M_n(a) \) is fixed by \( T_\rho \) so that it is a scalar multiple of \( x_\rho \) by Lemma 3.5 (iv). One may put
\[
\Phi(a) = \tau(a)x_\rho \quad \text{for some } \tau(a) \in \mathbb{C}.
\]
It is easy to see that \( \tau : \mathcal{A} \to \mathbb{C} \) is a state. As \( \Phi(T_\rho(a)) = \Phi(a) \), one sees \( \tau(T_\rho(a)) = \tau(a) \) for \( a \in \mathcal{A} \). Hence \( \tau \) is an invariant state on \( \mathcal{A} \). Now \((\mathcal{A}, \rho, \Sigma)\) is irreducible, the invariant state is faithful. \( \square \)

Hence the following corollary is clear.

**Corollary 3.10.** Assume that \((\mathcal{A}, \rho, \Sigma)\) is irreducible. Then the following two assertions are equivalent:

(i) \((\mathcal{A}, \rho, \Sigma)\) is mean ergodic.

(ii) There exist an invariant state \( \tau \) on \( \mathcal{A} \) and a positive element \( x_0 \in \mathcal{A} \) with \( \tau(x_0) = 1 \) such that \( \lim_{n \to \infty} M_n(a) = \tau(a)x_0 \) for \( a \in \mathcal{A} \).

In this case \( x_0 \) is given by \( \lim_{n \to \infty} M_n(1)(= x_\rho) \), and \( \tau \) is faithful.

**Theorem 3.11.** Assume that \((\mathcal{A}, \rho, \Sigma)\) is irreducible. Then the following two assertions are equivalent:

(i) \((\mathcal{A}, \rho, \Sigma)\) is mean ergodic.

(ii) \((\mathcal{A}, \rho, \Sigma)\) is uniquely ergodic and satisfies (FP).

**Proof.** (i) \(\Rightarrow\) (ii): Suppose that \((\mathcal{A}, \rho, \Sigma)\) is mean ergodic. Put \( \Phi(a) = \lim_{n \to \infty} M_n(a) \) for \( a \in \mathcal{A} \). The element \( x_\rho = \Phi(1) \) is a nonzero fixed element of \( \mathcal{A} \) under \( T_\rho \). By the previous corollary, there exists an invariant state \( \tau \) on \( \mathcal{A} \) satisfying \( \Phi(a) = \tau(a)x_\rho \) for \( a \in \mathcal{A} \). For any invariant state \( \psi \) on \( \mathcal{A} \), we have \( \psi \circ M_n(a) = \psi(a) \) for \( a \in \mathcal{A} \). Hence \( \psi(\Phi(a)) = \psi(a) \) so that \( \psi(a) = \psi(\tau(a)x_\rho) = \tau(a)\psi(x_\rho) \). Since \( \psi(x_\rho) = 1 \), we obtain \( \psi(a) = \tau(a) \). Therefore \( \psi = \tau \) so that \((\mathcal{A}, \rho, \Sigma)\) is uniquely ergodic.

(ii) \(\Rightarrow\) (i): Suppose that \((\mathcal{A}, \rho, \Sigma)\) is uniquely ergodic and satisfies (FP). By Lemma 3.7, \((\mathcal{A}, \rho, \Sigma)\) is power-bounded. Hence the sequence \( \left( \frac{1}{n} \sum_{k=0}^{n-1} T_\rho^{k} \right)_{n \in \mathbb{N}} \)
is uniformly bounded. This means that $T_\rho : A \to A$ is Cesàro bounded (cf. [22, p.72]). As $\lim_{n \to \infty} T_n^{-1}(a) = 0$ for $a \in A$, the operator $T_\rho : A \to A$ satisfies the assumption of [22, p.74 Theorem 1.4]. To prove mean ergodicity, it suffices to show that $F = \{ x \in A \mid T_\rho x = x \}$ separates $F^* = \{ \varphi \in A^* \mid \varphi \circ T_\rho = \varphi \}$. By Lemma 3.6, one knows that $F^* = C\tau$, where $\tau$ is a unique faithful invariant state on $A$. Hence if $\varphi = c\tau \in F^*$ is nonzero, then $c \neq 0$ and $\varphi(x_\rho) = c\tau(x_\rho) = c \neq 0$. This implies that $F$ separates $F^*$. Thus by [22, p.74 Theorem 1.4], $(A, \rho, \Sigma)$ is mean ergodic.

\[ \square \]

**Remark.** In [22, p.179], it is shown that a mean ergodic irreducible “Markov operator” is uniquely ergodic. In our situation, the operator $T_\rho$ does not necessarily satisfy $T_\rho(1) = 1$. Hence the operator $T_\rho$ is not necessarily a Markov operator.

We summarize results obtained in this section as in the following way:

**Theorem 3.12.** Assume that $(A, \rho, \Sigma)$ is irreducible. Then the following implications hold:

\[
(ME) \iff (UE) + (FP) \Rightarrow (FP) \Rightarrow (PB)
\]

\[
\Downarrow
\]

\[
\dim \mathcal{E}_{r_\rho}(\rho) = 1 \Rightarrow (UE),
\]

where (ME) means mean ergodic, (UE) means uniquely ergodic, and (PB) means power-bounded.

If in particular $(A, \rho, \Sigma)$ is irreducible and mean ergodic, the following Perron-Frobenius type theorem holds.

**Theorem 3.13.** Assume that $(A, \rho, \Sigma)$ is irreducible and mean ergodic.

(i) There exists a unique pair of a faithful state $\tau$ on $A$ and a strictly positive element $x_\rho$ in $A$ satisfying the conditions:

\[
\tau \circ \lambda_\rho = r_\rho \tau, \quad \lambda_\rho(x_\rho) = r_\rho x_\rho \quad \text{and} \quad \tau(x_\rho) = 1,
\]

where $r_\rho$ is the spectral radius of the positive operator $\lambda_\rho$ on $A$.

(ii) If there exists a continuous linear functional $\psi$ on $A$ satisfying

\[
\psi \circ \lambda_\rho = r_\rho \psi,
\]

then $\psi = c\tau$ for some complex number $c \in \mathbb{C}$.

(iii) If there exists a state $\varphi$ on $A$ and a complex number $\beta \in \mathbb{C}$ with $\beta \neq 0$

satisfying

\[
\varphi \circ \lambda_\rho = \beta \varphi,
\]

then $\varphi = \tau$ and $\beta = r_\rho$. 
Ergodic Properties and KMS Conditions

(iv) For any $a \in A$, the limit $\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \frac{\lambda_k(a)}{r_k} \in A$ exists in the norm topology such that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \frac{\lambda_k(a)}{r_k} = \tau(a)x.$$  

Proof. Under the assumption that $(A, \rho, \Sigma)$ is irreducible, mean ergodicity is equivalent to unique ergodicity with (FP). (i) and (iv) follows from Corollary 3.10 and unique ergodicity. (ii) follows from Lemma 3.6 (ii). (iii) follows from Lemma 3.6 (i) and unique ergodicity. □

4. Extension of eigenvectors to $F_\rho$

In this section, we will study extendability of an eigenvector in $E_\beta(\rho)$ to the subalgebra $F_\rho$. We fix a $C^*$-symbolic dynamical system $(A, \rho, \Sigma)$ satisfying condition (I) henceforth.

**Lemma 4.1.** Fix a nonnegative integer $k \in \mathbb{Z}_+$. For any element $x \in F^k_\rho$ there uniquely exists $x_{\mu, \nu}$ in $A$ for each $\mu, \nu \in B_k(\Lambda)$ such that

$$x = \sum_{\mu, \nu \in B_k(\Lambda)} S_{\mu} x_{\mu, \nu} S^*_\nu \quad \text{and} \quad x_{\mu, \nu} = \rho(1)x_{\mu, \nu}(1). \quad (4.1)$$

If in particular $x$ belongs to $D_k^\rho$, there uniquely exists $x_\mu$ in $A$ for each $\mu \in B_k(\Lambda)$ such that

$$x = \sum_{\mu \in B_k(\Lambda)} S_{\mu} x_{\mu} S^*_\mu \quad \text{and} \quad x_\mu = \rho(1)x_{\mu}(1). \quad (4.2)$$

Proof. For an element $x$ in $F^k_\rho$ and $\mu, \nu \in B_k(\Lambda)$, put $x_{\mu, \nu} = S^*_\mu x S_\nu$ that belongs to $A$ and satisfies the equalities (4.1). □

We set

$$D^{alg}_\rho = \text{the algebraic linear span of } S_{\mu} a S^*_\mu \text{ for } \mu \in B_*(\Lambda), a \in A,$$

and

$$F^{alg}_\rho = \text{the algebraic linear span of } S_{\mu} a S^*_\nu \text{ for } \mu, \nu \in B_*(\Lambda), |\mu| = |\nu|, a \in A.$$  

Hence $D^{alg}_\rho = \bigcup_{k=0}^{\infty} D^k_\rho$ and $F^{alg}_\rho = \bigcup_{k=0}^{\infty} F^k_\rho$. They are dense $*$-subalgebras of $D_\rho$ and $F_\rho$, respectively.

**Lemma 4.2.** For $\beta \in \mathbb{C}$ with $|\beta| > 1$ and $\varphi \in E_\beta(\rho)$ on $A$, put

$$\hat{\varphi}(S_{\mu} a S^*_\mu) = \frac{1}{|\beta|^{|\mu|}} \varphi(a_{\rho_{\mu}}(1)), \quad a \in A, \mu \in B_*(\Lambda). \quad (4.3)$$

Then $\hat{\varphi}$ is a well-defined (not necessarily continuous) linear functional on $D^{alg}_\rho$, that is an extension of $\varphi$.  

**Documenta Mathematica** 16 (2011) 133–175
Proof. By the expansion (4.2) for an element $x \in \mathcal{D}^k_\rho$, the following definition of $\varphi_k(x)$ yields a linear functional $\varphi_k$ on $\mathcal{D}^k_\rho$

$$\varphi_k(x) = \sum_{\mu \in B_k(\Lambda)} \frac{1}{\beta_k^{|\mu|}} \varphi(x_\mu). \quad (4.4)$$

We will show that $\varphi_k = \varphi_{k+1}$ on $\mathcal{D}^k_\rho$. As $S_\mu x_\mu S_\mu^* = \sum_{a \in \Sigma} S_{\mu a} \rho(a)(x_\mu) S_{\mu a}^*$ and $\rho_{\mu a}(1) \rho_0(x_\mu) \rho_{\mu a}(1) = 1 S_{\mu a} \rho_0(1) x_\mu \rho_{\mu a}(1) S_{\mu a} = \rho_0(x_\mu)$, the following expression of $x$ in $\mathcal{D}^{k+1}_\rho$

$$x = \sum_{\mu \in B_k(\Lambda), a \in \Sigma} S_{\mu a} \rho_0(x_\mu) S_{\mu a}$$

is the unique expression of (4.2). Hence we obtain

$$\varphi_{k+1}(x) = \sum_{\mu \in B_k(\Lambda), a \in \Sigma} \frac{1}{\beta_{k+1}^{|\mu|}} \varphi(\rho_0(x_\mu)) = \frac{1}{\beta_k} \sum_{\mu \in B_k(\Lambda)} \varphi(x_\mu) = \varphi_k(x).$$

The family $\{\varphi_k\}_{k \in \mathbb{Z}_+}$ of linear functionals on the subalgebras $\{\mathcal{D}^k_\rho\}_{k \in \mathbb{Z}_+}$ yields a linear functional on the algebra $\mathcal{D}^\text{alg}_\rho$. We denote it by $\varphi$. As the expansion $a = \sum_{a \in \Sigma} S_{\alpha a} \rho_0(a) S_{\alpha a}^*$ for $a \in \mathcal{A}$ is the unique expansion of $a$ in (4.2) as an element of $\mathcal{D}^1_\rho$, we have $\varphi(a) = \sum_{a \in \Sigma} \varphi(\rho_0(a)) = \varphi(a)$ so that $\varphi = \varphi$ on $\mathcal{A}$.

We will extend $\lambda_\rho$ on $\mathcal{A}$ to $\mathcal{F}_\rho$, such as

$$\lambda_\rho(x) = \sum_{a \in \Sigma} S_{\alpha a} x S_{\alpha a} \quad \text{for } x \in \mathcal{F}_\rho.$$

**Lemma 4.3.** Let $\psi$ be a linear functional on $\mathcal{F}^\text{alg}_\rho$ such that its restriction to $\mathcal{A}$ is continuous. Then the following three conditions are equivalent:

(i) $\psi$ is tracial and $\psi \circ \lambda_\rho(x) = \beta \psi(x)$ for $x \in \mathcal{F}^\text{alg}_\rho$.

(ii) $\psi(S_\mu x_\mu S_\mu^*) = \delta_{\mu, \nu, 1} \frac{1}{|\mu|} \psi(x S_\mu^* S_\mu)$ for $x \in \mathcal{F}^\text{alg}_\rho, \mu, \nu \in B_s(\Lambda)$ with $|\mu| = |\nu|$.

(iii) There exists $\varphi \in \mathcal{E}_\beta(\rho)$ such that

$$\psi(S_{\mu a} S_{\nu a}^*) = \delta_{\mu, \nu} \frac{1}{|\mu a|} \varphi(\rho_0(1)) \text{ for } a \in \mathcal{A}, \mu, \nu \in B_s(\Lambda) \text{ with } |\mu| = |\nu|.$$  

**Proof.** (i) $\Rightarrow$ (ii): The equation (i) implies that for $k \in \mathbb{N}$,

$$\psi(x) = \frac{1}{\beta_k} \sum_{\gamma \in B_k(\Lambda)} \psi(S_{\gamma a} S_{\gamma a}^*) \quad x \in \mathcal{F}^\text{alg}_\rho.$$

It then follows that for $\mu, \nu \in B_k(\Lambda)$

$$\psi(S_{\mu a} S_{\nu a}^*) = \frac{1}{\beta_k} \sum_{\gamma \in B_k(\Lambda)} \psi(S_{\gamma a} S_{\mu a} S_{\nu a}^*) = \delta_{\mu, \nu} \frac{1}{|\mu a|} \psi(x S_{\mu a}^* S_{\mu a}).$$
(ii) ⇒ (iii): Define a linear functional \( \varphi \) on \( \mathcal{A} \) by the restriction of \( \psi \) to the subalgebra \( \mathcal{A} \). By the equation (ii) for \( a \in \mathcal{A} \) and hence \( S^*_\alpha aS_\alpha \in \mathcal{A} \), we see

\[
\psi(S_\alpha S^*_\alpha a) = \psi(S_\alpha S^*_\alpha aS^*_\alpha S_\alpha) = \frac{1}{\beta} \psi(S^*_\alpha aS_\alpha S^*_\alpha S_\alpha) = \frac{1}{\beta} \psi(S^*_\alpha aS_\alpha)
\]

so that \( \varphi \in \mathcal{E}_\beta(\rho) \). The equation (iii) is clear.

(iii) ⇒ (i): We will see that \( \psi \) is tracial. Let \( x, y \in \mathcal{F}_\rho^k \) be expanded as in (4.1) so that

\[
x = \sum_{\mu, \nu \in B_\Lambda(\Lambda)} S_\mu x_{\mu, \nu} S^*_\nu, \quad y = \sum_{\mu, \nu \in B_\Lambda(\Lambda)} S_\mu y_{\mu, \nu} S^*_\nu.
\]

We have

\[
xy = \sum_{\mu, \nu, \gamma \in B_\Lambda(\Lambda)} S_\mu x_{\mu, \nu} \rho_\mu(1) y_{\nu, \gamma} S^*_\gamma = \sum_{\mu \in B_\Lambda(\Lambda)} S_\mu \left( \sum_{\nu \in B_\Lambda(\Lambda)} x_{\mu, \nu} y_{\nu, \gamma} \right) S^*_\gamma
\]

and

\[
\sum_{\nu \in B_\Lambda(\Lambda)} x_{\mu, \nu} y_{\nu, \gamma} = \rho_\mu(1) \left( \sum_{\nu \in B_\Lambda(\Lambda)} x_{\mu, \nu} y_{\nu, \gamma} \right) \rho_\nu(1), \text{ similarly}
\]

\[
yx = \sum_{\eta, \nu \in B_\Lambda(\Lambda)} S_\eta \left( \sum_{\gamma \in B_\Lambda(\Lambda)} y_{\eta, \gamma} x_{\gamma, \nu} \right) S^*_\nu
\]

and

\[
\sum_{\gamma \in B_\Lambda(\Lambda)} y_{\eta, \gamma} x_{\gamma, \nu} = \rho_\eta(1) \left( \sum_{\gamma \in B_\Lambda(\Lambda)} y_{\eta, \gamma} x_{\gamma, \nu} \right) \rho_\nu(1).
\]

It follows that

\[
\psi(xy) = \sum_{\mu, \nu \in B_\Lambda(\Lambda)} \frac{1}{\beta \kappa} \varphi(x_{\mu, \nu} y_{\mu, \nu}) = \sum_{\gamma, \eta \in B_\Lambda(\Lambda)} \frac{1}{\beta \kappa} \varphi(y_{\eta, \gamma} x_{\gamma, \eta}) = \psi(yx).
\]

Hence \( \psi \) is tracial on \( \mathcal{F}_\rho^k \).

We will finally show that the equality in (i) holds. For \( S_\mu aS^*_\nu \in \mathcal{F}_\rho^k \) with \( a \in \mathcal{A}, \mu = (\mu_1, \ldots, \mu_k), \nu = (\nu_1, \ldots, \nu_k) \in B_\Lambda(\Lambda) \), put \( \mu_{[2,k]} = (\mu_2, \ldots, \mu_k), \nu_{[2,k]} = (\nu_2, \ldots, \nu_k) \in \hat{B}_{k-1}(\Lambda) \). One has

\[
\sum_{a \in \Sigma} \psi(S^*_\alpha(S_\mu aS^*_\nu)S_\beta) = \delta_{\mu_{[2,k]} \nu_{[2,k]}} \varphi(\rho_\mu(1) a \rho_{\nu_{[2,k]}(1)} S^*_\nu)
\]

\[
= \delta_{\mu_{[2,k]} \nu_{[2,k]}} \varphi(\rho_\mu(1) a \rho_{\nu_{[2,k]}(1)} S^*_\nu)
\]

\[
= \delta_{\mu_{[2,k]} \nu_{[2,k]}} \varphi(\rho_\mu(1) a \rho_{\nu_{[2,k]}(1)} S^*_\nu)
\]

\[
= \delta_{\mu_{[2,k]} \nu_{[2,k]}} \varphi(\rho_\mu(1) a \rho_{\nu_{[2,k]}(1)} S^*_\nu)
\]

\[
= \delta_{\mu_{[2,k]} \nu_{[2,k]}} \varphi(\rho_\mu(1) a \rho_{\nu_{[2,k]}(1)} S^*_\nu)
\]

\[
= \beta \psi(S_\mu aS^*_\nu).
\]

\[\square\]

Let \( E_D : \mathcal{F}_\rho \to \mathcal{D}_\rho \) denote the expectation satisfying

\[
E_D(S_\mu aS^*_\nu) = \delta_{\mu, \nu} \rho_\mu aS^*_\nu, \quad a \in \mathcal{A}, \quad \mu, \nu \in B_\Lambda(\Lambda), |\mu| = |\nu|.
\]

Once we have an extension \( \tilde{\varphi} \) to \( \mathcal{D}_\rho \) of \( \varphi \in \mathcal{E}_\beta(\rho) \), \( \tilde{\varphi} \) has a further extension to \( \mathcal{F}_\rho \) by \( \varphi \circ E_D \). The extension \( \varphi \circ E_D \) on \( \mathcal{F}_\rho \) is continuous if \( \tilde{\varphi} \) is so on \( \mathcal{D}_\rho \). It satisfies

\[
\hat{\varphi} \circ E_D(S_\mu aS^*_\nu) = \delta_{\mu, \nu} \frac{1}{\beta |\mu|} \varphi(\rho_\mu(1)) \tag{4.5}
\]
for \( a \in \mathcal{A}, \mu, \nu \in B_\alpha(\Lambda) \) with \( |\mu| = |\nu| \). Hence the extension of a continuous linear functional on \( D_\rho \) to \( F_\rho \) is automatic. We have only to study extension of a linear functional \( \varphi \in \mathcal{E}_\beta(\rho) \) on \( \mathcal{A} \) to \( D_\rho \).

The condition (iii) of Lemma 4.3 is equivalent to \( \psi = \tilde{\varphi} \circ E_D \) where \( \tilde{\varphi} \) is a linear functional on \( D_\rho^\text{alg} \) obtained from \( \varphi \in \mathcal{E}_\beta(\rho) \) as in Lemma 4.2, and so that \( \psi \) is continuous if and only if \( \tilde{\varphi} \) is continuous. We call the extension \( \tilde{\varphi} \) on \( D_\rho^\text{alg} \) of \( \varphi \in \mathcal{E}_\beta(\rho) \) the canonical extension of \( \varphi \).

Lemma 4.4. Suppose that \((\mathcal{A}, \rho, \Sigma)\) is irreducible and power-bounded. For \( \beta \in C \) with \( |\beta| = r_\beta > 1 \), a (not necessarily positive) continuous linear functional \( \varphi \in \mathcal{E}_\beta(\rho) \) on \( \mathcal{A} \) extends to a continuous linear functional \( \tilde{\varphi} \) on \( D_\rho \), satisfying (4.3).

Proof. As \((\mathcal{A}, \rho, \Sigma)\) is irreducible, we may take a faithful invariant state \( \tau \) on \( \mathcal{A} \), which we will fix. By the hypothesis that \((\mathcal{A}, \rho, \Sigma)\) is power-bounded, there exists a positive number \( M \) such that \( \frac{\|\lambda^K(1)\|}{r^K} < M \) for all \( k \in \mathbb{N} \). By [43, Theorem 4.2], there exists a partial isometry \( v \in \mathcal{A}^* \) and a positive linear functional \( \psi \in \mathcal{A}^* \) such that

\[ \varphi(a) = \psi(av), \quad a \in \mathcal{A}. \]

If in particular a linear functional \( \varphi \in \mathcal{E}_\beta(\rho) \) is positive on \( \mathcal{A} \), it always extends to a continuous linear functional on \( D_\rho \) as in the following way:
Lemma 4.5. Let $\beta \in \mathbb{C}$ be $|\beta| > 1$. If $\varphi \in \mathcal{E}_\beta(\rho)$ is a positive linear functional on $\mathcal{A}$, then $\beta$ becomes a positive real number and the canonical extension $\tilde{\varphi}$ to $\mathcal{D}_\rho$ is continuous on $\mathcal{D}_\rho$.

Proof. One may assume that $\varphi \neq 0$ and $\varphi(1) = 1$. We have $\beta = \beta \varphi(1) = \varphi(\lambda_\rho(1)) \geq 1$. For $k \in \mathbb{N}$, define a linear functional $\varphi_k$ on $\mathcal{D}_\rho^k$ by (4.4). Since $x = \sum_{\mu \in B_k(\Lambda)} S(0) S(1)$, we have by (4.6),

$$|\varphi(\rho_\mu(1)x_\mu\rho_\mu(1))| \leq \varphi_\rho(1)^{1/2} \varphi(\rho_\mu(1)x_\mu\rho_\mu(1))^{1/2} \leq \|\varphi\rho(1)|$$

it follows that

$$|\varphi_k(x)| \leq \frac{1}{|\beta|^k} \sum_{\mu \in B_k(\Lambda)} |\varphi(\rho_\mu(1)x_\mu\rho_\mu(1))| \leq \frac{1}{|\beta|^k} \|x\| \|\varphi(\lambda_\rho^k(1))\| = \|x\|^2.$$

Therefore $\{\varphi_k\}_{k \in \mathbb{N}}$ extends to a state on $\mathcal{D}_\rho$.

We are now assuming that $(\mathcal{A}, \rho, \Sigma)$ is irreducible. By Lemma 3.4, there exists a faithful invariant state $\tau \in \mathcal{E}_{\tau_{\rho}}(\rho)$ on $\mathcal{A}$. By the previous lemma, the canonical extension $\tilde{\tau}$ is continuous on $\mathcal{D}_\rho$ which satisfies

$$\tilde{\tau}(S(0) S(1)) = \frac{1}{r(\rho)^k} \tau(\rho_\mu(1)), \quad a \in \mathcal{A}, \mu \in B_k(\Lambda). \quad (4.7)$$

Lemma 4.6. For a faithful invariant state $\tau \in \mathcal{E}_{\tau_{\rho}}(\rho)$ on $\mathcal{A}$, the canonical extension $\tilde{\tau}$ is faithful on $\mathcal{D}_\rho$.

Proof. Suppose that $\tilde{\tau}$ is not faithful on $\mathcal{D}_\rho$. Put

$$I_\rho = \{x \in \mathcal{D}_\rho \mid \tilde{\tau}(x^* x) = 0\}.$$

Since $\tilde{\tau}$ is tracial on $\mathcal{D}_\rho$, $I_\rho$ is a nonzero ideal of $\mathcal{D}_\rho$. By Lemma 4.3, the equality $\tilde{\tau} \circ \lambda_\rho = r(\rho) \tilde{\tau}$ holds on $\mathcal{D}_\rho$ so that $I_\rho$ is $\lambda_\rho$-invariant. The sequence $\mathcal{D}_\rho^k, k \in \mathbb{N}$ of algebras is increasing such that $\cup_{k \in \mathbb{N}} \mathcal{D}_\rho^k$ is dense in $\mathcal{D}_\rho$. We may find $k \in \mathbb{N}$ such that $I_\rho \cap \mathcal{D}_\rho^k \neq 0$. It is easy to see that $\lambda_\rho^k(\mathcal{D}_\rho^k) \subset \mathcal{A}$ so that there exists a nonzero positive element $x \in I_\rho \cap \mathcal{D}_\rho^k$ such that $\lambda_\rho^k(x) \in I_\rho \cap \mathcal{A}$. Hence $I_\rho \cap \mathcal{A}$ is a nonzero $\lambda_\rho$-invariant ideal of $\mathcal{A}$. By the hypothesis that $(\mathcal{A}, \rho, \Sigma)$ is irreducible, we have a contradiction.

For a faithful invariant state $\tau$ on $\mathcal{A}$, we will write the canonical extension $\tilde{\tau}$ of $\tau$ to $\mathcal{D}_\rho$ as still $\tau$. Define a unital endomorphism $\phi_\rho : \mathcal{D}_\rho \longrightarrow \mathcal{D}_\rho$ by setting

$$\phi_\rho(y) = \sum_{\alpha \in \Sigma} S(0) S(1), \quad y \in \mathcal{D}_\rho. \quad (4.8)$$

It induces a unital endomorphism on the enveloping von Neumann algebra $\mathcal{D}_\rho^{**}$ of $\mathcal{D}_\rho$, which we still denote by $\phi_\rho$. The restriction of the positive map $\lambda_\rho$ on $\mathcal{F}_\rho$ to $\mathcal{D}_\rho$ similarly induces a positive map on $\mathcal{D}_\rho^{**}$. We then need the following lemma for further discussions.
Lemma 4.7. The equality
\[ \lambda_\rho(x\phi_\rho(y)) = \lambda_\rho(x)y, \quad x, y \in \mathcal{D}_\rho^{**} \] (4.9)
holds.

Proof. Since \( \mathcal{D}_\rho \) is dense in \( \mathcal{D}_\rho^{**} \) under \( \sigma(\mathcal{D}_\rho^{**}, \mathcal{D}_\rho^*) \)-topology, it suffices to show the equality (4.9) for \( x, y \in \mathcal{D}_\rho \). One has
\[
\lambda_\rho(x\phi_\rho(y)) = \sum_{\alpha, \gamma \in \Sigma} S_\alpha^* x S_\gamma y S_\alpha^* S_\alpha = \sum_{\alpha \in \Sigma} S_\alpha^* x S_\alpha y = \lambda_\rho(x)y.
\]
\[\square\]

Recall that for a continuous linear functional \( \psi \) on a \( C^* \)-algebra \( B \) there exist a partial isometry \( v \in B^{**} \) and a positive linear functional \( |\psi| \in B^* \) in a unique way such that
\[ v^*v = s(|\psi|), \quad \psi(x) = |\psi|(xv) \quad \text{for} \quad x \in B, \] (4.10)
where \( s(|\psi|) \) denotes the support projection of \( |\psi| \) (cf. [43, Theorem 4.2]). The decomposition (4.10) is called the polar decomposition of \( \psi \). The linear functional \( \psi : x \rightarrow |\psi|(xv) \) is denoted by \( v|\psi| \).

Lemma 4.8. Let \( \beta = re^{i\theta} \in \mathbb{C} \) be \( r, \theta \in \mathbb{R} \) with \( r > 1 \). For a (not necessarily positive) linear functional \( \varphi \in \mathcal{E}_\beta(\rho) \) on \( \mathcal{A} \), let \( \tilde{\varphi} \) be the extension on \( \mathcal{D}_\rho^{**} \) satisfying (4.3). Suppose that the linear functional \( \tilde{\varphi} \) extends to a continuous linear functional on \( \mathcal{D}_\rho^{**} \). Denote by \( \varphi = v|\tilde{\varphi}| \) its polar decomposition for a partial isometry \( v \in \mathcal{D}_\rho^{**} \) and a positive linear functional \( |\tilde{\varphi}| \) on \( \mathcal{D}_\rho \) such that \( v^*v = s(|\tilde{\varphi}|) \). Then we have
\[ \phi_\rho(v) = e^{i\theta}v, \quad |\tilde{\varphi}|(\lambda_\rho(x)) = r|\tilde{\varphi}|(x) \quad \text{for} \quad x \in \mathcal{D}_\rho. \]
Hence the restriction of \( |\tilde{\varphi}| \) to \( \mathcal{A} \) belongs to \( \mathcal{E}_r(\rho) \) and \( |\tilde{\varphi}| \) satisfies
\[ |\tilde{\varphi}|(S_\alpha a S_\mu^*) = \frac{1}{r^{|\mu|}} |\tilde{\varphi}|(a\rho_\mu(1)), \quad a \in \mathcal{A}, \mu \in B_*(\Lambda). \]

Proof. Put a positive linear functional \( \psi \) on \( \mathcal{D}_\rho \) and a partial isometry \( u \) in \( \mathcal{D}_\rho^{**} \) by setting
\[ \psi(x) = \frac{1}{r} |\tilde{\varphi}|(\lambda_\rho(x)) \quad \text{for} \quad x \in \mathcal{D}_\rho \quad \text{and} \quad u = e^{-i\theta} \phi_\rho(v). \]
As \( \lambda_\rho(xu) = e^{-i\theta} \lambda_\rho(x)v \) for \( x \in \mathcal{D}_\rho \) by Lemma 4.7. It follows that for \( x \in \mathcal{D}_\rho \)
\[ (u\psi)(x) = \frac{1}{r} |\tilde{\varphi}|(\lambda_\rho(xu)) = \frac{1}{r} |\tilde{\varphi}|(\lambda_\rho(x)v) = \tilde{\varphi}(x). \]
Hence we have
\[ \tilde{\varphi} = u\psi \quad \text{on} \quad \mathcal{D}_\rho. \]
We will next show that $s(\psi) = u^*u$. For $y \in D_\rho$, we have by Lemma 4.7
\[
\psi(yu^*u) = \frac{1}{r} |\tilde{\phi}(\lambda(yu^*u))| = \frac{1}{r} |\tilde{\phi}(\lambda(\phi(yv^*v)))| = \frac{1}{r} |\tilde{\phi}(\lambda(\phi(v^*v)))| = \psi(y).
\]
Hence we have $u^*u \geq s(\psi)$. On the other hand, suppose that a projection $p \in D_\rho^{**}$ satisfies
\[
\psi(yp) = \psi(y) \quad \text{for } y \in D_\rho.
\]
We then have $|\tilde{\phi}(\lambda(y(1-p)))| = 0$ for all $y \in D_\rho$. For $y = S_\alpha S^*_\alpha$, $\alpha \in \Sigma$, one has $|\tilde{\phi}(S_\alpha^*(1-p)S_\alpha)| = 0$. As $S_\alpha^*(1-p)S_\alpha$ is a projection in $D_\rho$, one obtains that $S_\alpha^*(1-p)S_\alpha \leq 1 - v^*v$ so that $1 - p \leq 1 - \phi(v^*v)$. This implies that $u^*u \leq p$. Therefore we have $u^*u \leq s(\psi)$ and hence
\[
u^*u = s(\nu).
\]
By the uniqueness of the polar decomposition, we conclude that
\[
u = u \quad \text{and} \quad |\tilde{\nu}| = \psi \quad \text{on } D_\rho
\]
so that
\[
\phi_\rho(v) = e^{i\theta}v, \quad |\tilde{\phi}|(\lambda(x)) = r|\tilde{\phi}|(x) \quad \text{for } x \in D_\rho.
\]
Therefore we have

**Theorem 4.9.** Suppose that $(A, \rho, \Sigma)$ is irreducible and power-bounded. For $\beta \in C$ with $|\beta| > 1$, a (not necessarily positive) linear functional $\phi \in E_\beta(\rho)$ on $A$ extends to $D_\rho$, as a continuous linear functional $\tilde{\phi}$ satisfying
\[
\tilde{\phi}(S_\rho a S^*_\rho) = \frac{1}{|\beta|} |\phi(a\rho(1))|, \quad a \in A, \mu \in B_\rho(\Lambda)
\]
if $|\beta| = r_\rho$. If in particular, $(A, \rho, \Sigma)$ is mean ergodic, the converse implication holds.

**Proof.** The first part of the assertions is direct from Lemma 4.4. Under the condition that $(A, \rho, \Sigma)$ is mean ergodic, assume that the canonical extension $\tilde{\phi}$ is continuous on $D_\rho$. The preceding lemma says that the positive linear functional $|\tilde{\phi}|$ belongs to $E_{|\beta|}(\rho)$. Since the mean ergodicity implies (FP), by Lemma 3.6 (i) we see that $|\beta| = r_\rho$. 

Let us now assume that $(A, \rho, \Sigma)$ is irreducible and satisfies $\dim E_{r_\rho}(\rho) = 1$, and hence it is uniquely ergodic. Take a unique invariant state $\tau$ on $A$ and denote still by $\tau$ its canonical extension on $D_\rho$. Denote by $p_\tau \in D_\rho^{**}$ its support projection.

**Lemma 4.10.** Let $w \in D_\rho^{**}$ be a partial isometry satisfying
\[
w^*w = p_\tau \quad \text{and} \quad \phi_\rho(w) = w.
\]
Then $w$ is a scalar multiple of the projection $p_\tau$. 

**Documenta Mathematica 16 (2011) 133–175**
Proof. Put \( w\tau(x) = \tau(xw) \) for \( x \in \mathcal{D}_\rho \) and hence \( w\tau \in \mathcal{D}_\rho^{**} \). Since \( \lambda_\rho(x)w = \lambda_\rho(x\phi_\rho(w)) = \lambda_\rho(xw) \) by Lemma 4.7, it follows that for \( x \in \mathcal{D}_\rho \)

\[
(1) = \tau(w\tau(x)) = \tau(\lambda_\rho(xw)) = r_\rho \tau(xw) = r_\rho w\tau(x).
\]

In particular, we have \( w\tau \in \mathcal{E}_{\tau_\rho}(\rho) \). As \( \dim \mathcal{E}_{\tau_\rho}(\rho) = 1 \) by hypothesis, \( w\tau \) is a scalar multiple of \( \tau \). Hence there exists \( c \in \mathbb{C} \) such that \( \tau(xw) = c\tau(x) \) for \( x \in \mathcal{A} \). Since \( w\tau \) is the canonical extension of \( \tau(\cdot w) = w\tau \) on \( \mathcal{A} \) to \( \mathcal{D}_\rho \) and the canonical extension is unique, one has \( \tau(xw) = c\tau(x) \) for \( x \in \mathcal{D}_\rho \) so that

\[
\tau(xw) = \tau(xc\rho) \quad \text{for } x \in \mathcal{D}_\rho.
\]

(4.12)

As \( c = c\tau(1) = \tau(w) \), one has

\[
1 = \tau(p_\tau) = \tau(w^*w) = c\tau(w^*) = c\tau(w) = c\bar{c}
\]

so that

\[
(c\rho_\tau)^*(c\rho_\tau) = p_\tau = w^*w.
\]

By the uniqueness of the polar decomposition, we have by (4.12) \( w = c\rho_\tau \). \( \Box \)

Proposition 4.11. Suppose that \((\mathcal{A}, \rho, \Sigma)\) is irreducible and satisfies\n
\[
\dim \mathcal{E}_\beta(\rho) = 1. \quad \text{Then } \dim \mathcal{E}_\beta(\rho) \leq 1 \text{ for } \beta \in \mathbb{C} \text{ with } |\beta| > r_\rho > 1.
\]

Proof. Let \( |\beta| = r_\rho > 1 \). Take an arbitrary linear functional \( \varphi \in \mathcal{E}_\beta(\rho) \) with \( \varphi \neq 0 \). Its canonical extension \( \hat{\varphi} \) to \( \mathcal{D}_\rho \) is continuous. Denote by \( \check{\varphi} = v_\beta|\hat{\varphi}| \) the polar decomposition in \( \mathcal{D}_\rho^{**} \) where \( v_\beta \) is a partial isometry in \( \mathcal{D}_\rho^{**} \). By Lemma 4.7, the restriction of \( |\hat{\varphi}| \) to \( \mathcal{A} \) is a positive linear functional belonging to \( \mathcal{E}_{\tau_\rho}(\rho) \). Since \((\mathcal{A}, \rho, \Sigma)\) is uniquely ergodic, by putting \( c_\varphi = |\hat{\varphi}|(1) \) one has \( |\hat{\varphi}| = c_\varphi \tau \) as a positive linear functional on \( \mathcal{A} \). The canonical extension to \( \mathcal{D}_\rho \) which satisfies (4.3) is unique and determined by its behavior on \( \mathcal{A} \). Hence the equality \( |\hat{\varphi}| = c_\varphi \tau \) holds as a positive linear functional on \( \mathcal{D}_\rho \) so that we have \( \text{supp}(|\hat{\varphi}|) = \text{supp}(\tau) \) and hence \( v_\varphi^*v_\varphi = p_\tau \). For another linear functional \( \psi \in \mathcal{E}_\beta(\rho) \) with \( \psi \neq 0 \), we have similar decompositions

\[
\tilde{\psi} = v_\varphi|\tilde{\psi}|, \quad |\tilde{\psi}| = c_\varphi \tau, \quad v_\varphi^*v_\varphi = p_\tau.
\]

Put a partial isometry \( w = v_\varphi^*v_\varphi \in \mathcal{D}_\rho^{**} \) so that \( w^*w = p_\tau \). By Lemma 4.8, one has \( \phi_\rho(w) = w \). Lemma 4.10 implies \( w = c\rho_\tau \) for some \( c \in \mathbb{C} \) with \( |c| = 1 \) so that \( v_\varphi^* = cv_\varphi \). Therefore we have

\[
\tilde{\psi} = v_\varphi|\tilde{\psi}| = cv_\varphi c_\varphi \tau = c\frac{c_\varphi}{c_\varphi} \tilde{\varphi}
\]

on \( \mathcal{D}_\rho \). In particular we have \( \psi = c\frac{c_\varphi}{c_\varphi} \varphi \) on \( \mathcal{A} \) so that \( \dim \mathcal{E}_\beta(\rho) \leq 1 \). \( \Box \)

Corollary 4.12. Suppose that \((\mathcal{A}, \rho, \Sigma)\) is irreducible and mean ergodic. Then for \( \beta \in \mathbb{C} \) with \( |\beta| > 1 \), we have \( \dim \mathcal{E}_\beta(\rho) \leq 1 \) if \( |\beta| = r_\rho \), otherwise \( \mathcal{E}_\beta(\rho) = \{0\} \).
Suppose that \((A, \rho, \Sigma)\) is irreducible and mean ergodic. Hence \((A, \rho, \Sigma)\) is uniquely ergodic with a unique faithful invariant state \(\tau \in \mathcal{E}_\rho(\rho)\). Denote by \(p_\tau \in \mathcal{D}_\rho^{**}\) the support projection of the canonical extension of \(\tau\) to \(\mathcal{D}_\rho\), where the extension is still denoted by \(\tau\). For \(\beta = re^{i\theta} \in \mathbb{C}\) with \(r = r_\rho > 1\), we set

\[P_\beta(\mathcal{D}_\rho, \tau) = \{v \in \mathcal{D}_\rho^{**} \mid \phi_\rho(v) = e^{i\theta}v, \ v^*v = p_\tau\}.
\]

Denote by \(R_+\) the set of all nonnegative real numbers. For \(\varphi \in \mathcal{E}_{\beta}(\rho)\) denote by \(\tilde{\varphi}\) its canonical extension to \(\mathcal{D}_\rho\). As \(|\beta| = r_\rho\), \(\tilde{\varphi}\) is continuous and has a unique polar decomposition \(\tilde{\varphi} = v_\tilde{\varphi}|\tilde{\varphi}|\) for some \(v_\tilde{\varphi} \in \mathcal{D}_\rho^{**}\) and positive linear functional \(|\tilde{\varphi}| \in \mathcal{D}_\rho^*\). By Lemma 4.8, we know the structure of the eigenspace \(\mathcal{E}_{\beta}(\rho)\) as in the following way:

**Proposition 4.13.** Suppose that \((A, \rho, \Sigma)\) is irreducible and mean ergodic.

There exists a bijective correspondence between the eigenspace \(\mathcal{E}_{\beta}(\rho)\) and the product set \(P_\beta(\mathcal{D}_\rho, \tau) \times R_+\) through the correspondences

\[
\varphi \in \mathcal{E}_{\beta}(\rho) \mapsto (v_\tilde{\varphi}, |\tilde{\varphi}|(1)) \in P_\beta(\mathcal{D}_\rho, \tau) \times R_+,
\]

\[
ct(\cdot, v) \in \mathcal{E}_{\beta}(\rho) \longleftarrow (v, c) \in P_\beta(\mathcal{D}_\rho, \tau) \times R_+.
\]

### 5. Extension to \(O_\rho\) and KMS condition

In [9], Enomoto-Fujii-Watatani have proved that KMS states for gauge action on the Cuntz-Krieger algebra \(O_A\) exist if and only if its inverse temperature is \(\log r_A\), where \(r_A\) is the Perron-Frobenius eigenvalue for the irreducible matrix \(A\). They have showed that the KMS states bijectively correspond to the normalized positive eigenvectors of \(A\) for the eigenvalue \(r_A\).

In this section, we will study KMS conditions for linear functionals without assuming its positivity at inverse temperature taking complex numbers. The extended notation is needed to study eigenvector spaces for \(C^*\)-symbolic dynamical systems.

Following after [3], KMS states for one-parameter group action \(\alpha\) on a \(C^*\)-algebra \(B\) is defined as follows: For a positive real number \(\gamma \in \mathbb{R}\), a state \(\psi\) on \(B\) is a KMS state at inverse temperature \(\gamma\) if \(\psi\) satisfies

\[
\psi(y\alpha_\gamma(x)) = \psi(xy), \quad x \in B^a, y \in B
\]

where \(B^a\) is the set of analytic elements of the action \(\alpha : \mathbb{R} \longrightarrow \text{Aut}(B)\) (cf.[3]). The equation (5.1) for \(\psi\) is called the KMS condition with respect to the action \(\alpha\).

In what follows, we restrict our interest to periodic actions so as to extend KMS condition to (not necessarily positive) linear functionals at inverse temperature taking complex numbers. We assume that an action \(\alpha\) of \(\mathbb{R}\) has its period \(2\pi\) so that \(\alpha\) is regarded as an action of one-dimensional torus group \(T = \mathbb{R}/2\pi\mathbb{Z}\). Let \(B\) be a \(C^*\)-algebra and \(\alpha : T \longrightarrow \text{Aut}(B)\) a continuous action of \(T\) to the automorphism group \(\text{Aut}(B)\). We write a complex number \(\beta \in \mathbb{C}\) as \(\beta = re^{i\theta}\) where \(r, \theta \in \mathbb{R}\) with \(r > 1\).
DEFINITION. A continuous linear functional \( \varphi \in \mathcal{B}^* \) on \( \mathcal{B} \) is said to satisfy the KMS condition at \( \log \beta \) if \( \varphi \) satisfies the following condition

\[
\varphi(y_{\alpha \log r}(x)) = \varphi(\alpha(x)y), \quad x \in \mathcal{B}^a, y \in \mathcal{B}.
\]

(5.2)

REMARK.

(i) As \( \alpha \theta(x) = \alpha \theta + 2\pi(x) \), the right hand side \( \varphi(\alpha(x)y) \) of (5.2) does not depend on the choice of \( \theta \in \mathbb{R} \) as long as \( \beta = re^{i\theta} \).

(ii) The above KMS condition (5.2) is equivalent to the following condition:

\[
\varphi(y_{\alpha \theta}(x)y) = \varphi(\alpha(x)y), \quad x \in \mathcal{B}^a, y \in \mathcal{B}, \ \zeta \in \mathbb{C}
\]

(5.3)

(iii) In case of \( \theta = 0 \), the above definition of KMS condition coincides with the original definition of KMS condition for states.

(iv) The above equality (5.2) can be written formally as

\[
\varphi(y_{\alpha \log \beta}(x)) = \varphi(xy), \quad x \in \mathcal{B}^a, y \in \mathcal{B},
\]

(5.4)

if we denote \( \log \beta = \log r + i\theta \).

We will present some examples of linear functionals satisfying the extended KMS conditions.

EXAMPLES.

(i) Let \( \alpha : \mathbb{T} \rightarrow \text{Aut}(\mathcal{B}) \) be an action of \( \mathbb{T} \) to a \( \mathcal{C}^* \)-algebra \( \mathcal{B} \) such that there exists a projection \( H \in \mathcal{B} \) satisfying \( \alpha_t(a) = e^{itH}ae^{-itH}, a \in \mathcal{B}, t \in \mathbb{T} \). Assume that there exists an \( \alpha \)-invariant tracial state \( \text{tr} \) on \( \mathcal{B} \).

Put

\[
\varphi(x) = \frac{\text{tr}(e^{-Log \beta H}x)}{\text{tr}(e^{-Log \beta H})}, \quad x \in \mathcal{B},
\]

where \( Log \beta = \log r + i\theta \). Then \( \varphi \) satisfies KMS condition at \( Log \beta \).

(iii) Let \( \mathcal{B} = \otimes_{k=1}^{\infty} M_2 \) be the UHF-algebra of type \( 2^\infty \) and \( \alpha : \mathbb{T} \rightarrow \text{Aut}(\mathcal{B}) \) an action of \( \mathbb{T} \) to \( \mathcal{B} \) defined by

\[
\alpha_t = \otimes_{k=1}^{\infty} \text{Ad}(1_{0 \ 0 \ 0 \ 1}e^{it}), \quad t \in \mathbb{T}.
\]

Put

\[
\mathcal{B}_n = \otimes_{k=1}^{n} M_2 = M_2 \otimes \cdots \otimes M_2,
\]

\[
u_t^n = \otimes_{k=1}^{n} [1_0 0 0 e^{it}] = [1_0 0 e^{it}] \otimes \cdots \otimes [1_0 0 e^{it}] \in \mathcal{B}_n,
\]

\[
\alpha_t^n = \text{Ad}(\nu_t^n) \in \text{Aut}(\mathcal{B}_n), \quad t \in \mathbb{T}.
\]

Let \( \beta = re^{i\theta} \in \mathbb{C} \) be \( r > 1 \). Put

\[
H = [0_0 0 0 0 0 0 1] \in M_2, \quad h_n = \otimes_{k=1}^{n} [1_0 0 0 0 0 0 1] \in \mathcal{B}_n
\]
and hence $h_n = \otimes_{k=1}^n e^{-\log \beta H} \alpha_t^n = \otimes_{k=1}^n \text{Ad}(e^{i\theta H}), t \in T$. It is straightforward to see that
\[ \text{tr}(e^{-\log \beta H} b \alpha_t \log r(a)) = \text{tr}(e^{-\log \beta H} \alpha_t(a)b), \quad a,b \in M_2. \]

Put
\[ \varphi_n(x) = \otimes_{k=1}^n \text{tr}(x h_n) \quad \text{for } x \in \mathcal{B}_n \]
so that we have
\[ \varphi_n(y \alpha_t \log r(x)) = \varphi_n(\alpha_t(x)y), \quad x,y \in \mathcal{B}_n. \]

As $\|h_n\| = 1$, $\varphi_n$ extends to a continuous linear functional on $\mathcal{B}$, which we denote by $\varphi$. Then $\varphi$ satisfies KMS condition at $\log \beta$,
\[ \varphi(y \alpha_t \log r(x)) = \varphi_n(\alpha_t(x)y), \quad x \in \mathcal{B}^\ast, y \in \mathcal{B}. \]

We see the following two propositions whose proofs are similar to the case of usual KMS states.

**Proposition 5.1** (cf. [39, 8.12.3]). Let $\alpha : T \longrightarrow \text{Aut}(\mathcal{B})$ be a continuous action of $T$ to the automorphism group $\text{Aut}(\mathcal{B})$ of a $C^*$-algebra $\mathcal{B}$ and $\beta$ a complex number with $\beta = re^{i\theta}, r > 1$. The following conditions for a continuous linear functional $\varphi$ on $\mathcal{B}$ are equivalent:

(i) $\varphi$ satisfies the KMS condition at $\log \beta$.

(ii) $\varphi$ satisfies the equality (5.2) for just a dense set of elements in $\mathcal{B}^\ast$.

(iii) For all $x,y \in \mathcal{B}$, there is a bounded continuous function $f$ on the strip
\[ \Omega_{\log r} = \{ \zeta \in \mathbb{C} \mid 0 \leq \text{Im} \zeta \leq \log r \} \]

such that $f$ is holomorphic in the interior of $\Omega_{\log r}$ and
\[ f(t) = \varphi(y \alpha_t(x)), \quad f(t + i \log r) = \varphi(\alpha_t(x)y), \quad t \in \mathbb{R}. \]

**Proposition 5.2** (cf. [39, 8.12.4]). Let $\mathcal{B}$ be a $C^*$-algebra and $\alpha : T \longrightarrow \text{Aut}(\mathcal{B})$ be a continuous action of $T$ to the automorphism group $\text{Aut}(\mathcal{B})$. Let $\varphi$ be a continuous linear functional on $\mathcal{B}$. If $\varphi$ satisfies KMS condition at $\log \beta$ for some complex number $\beta$ with $\beta = re^{i\theta}$ with $r > 1$, then $\varphi$ is $\alpha$-invariant, that is,
\[ \varphi \circ \alpha_t = \varphi, \quad t \in T. \]

We henceforth go back to our previous situations. Let $(\mathcal{A}, \rho, \Sigma)$ be a $C^*$-symbolic dynamical system. Recall that the positive operator $\lambda_\rho$ on $\mathcal{A}$ extends to $\mathcal{F}_\rho$ by setting $\lambda_\rho(x) = \sum_{\alpha \in \Sigma} S^\rho_\alpha x S^\rho_\alpha, x \in \mathcal{F}_\rho$. For $\beta \in \mathbb{C}$ with $\beta \neq 0$, we set
\[ E^D_\beta(\rho) = \{ \varphi \in \mathcal{D}_\rho^\ast \mid \varphi(\lambda_\rho(x)) = \beta \varphi(x), x \in \mathcal{D}_\rho \}, \]
\[ E^F_\beta(\rho) = \{ \phi \in \mathcal{F}_\rho^\ast \mid \phi(\lambda_\rho(x)) = \beta \phi(x), x \in \mathcal{F}_\rho, \phi \text{ is tracial on } \mathcal{F}_\rho \}. \]

It is possible that both $E^D_\beta(\rho)$ and $E^F_\beta(\rho)$ are $\{ 0 \}$. Recall that $E_\mathcal{D} : \mathcal{F}_\rho \longrightarrow \mathcal{D}_\rho$ is the canonical expectation satisfying by $E_\mathcal{D}(S_\mu a S^\rho_\nu) = \delta_{\mu,\nu} a S^\rho_\nu$ for $a \in \mathcal{A}$ with $\mu, \nu \in \mathcal{B}, (\mathcal{A}), |\mu| = |\nu|$. By composing it to a given linear functional $\varphi \in E^D_\beta(\rho)$ on $\mathcal{D}_\rho$, $\varphi$ extends to $\mathcal{F}_\rho$. 

*Documenta Mathematica 16 (2011) 133–175*
Lemma 5.3. Let $\beta \in \mathbb{C}$ with $|\beta| > 1$. A (not necessarily positive) continuous linear functional $\varphi \in \mathcal{E}_\beta^F(\rho)$ on $\mathcal{D}_\rho$ uniquely extends to $\mathcal{F}_\rho$ as a tracial continuous linear functional $\psi = \varphi \circ \mathcal{E}_D$ such that

$$\phi(S_\mu x S_\nu^*) = \delta_{\mu,\nu} \frac{1}{|\beta|^{\mu}} \phi(x S_\mu S_\nu^*), \quad x \in \mathcal{F}_\rho, \mu, \nu \in B_\epsilon(\Lambda) \text{ with } |\mu| = |\nu|. \quad (5.7)$$

Hence the sets $\mathcal{E}_\beta^F(\rho)$ and $\mathcal{E}_\beta^D(\rho)$ bijectively correspond to each other.

Proof. For $\varphi \in \mathcal{E}_\beta^D(\rho)$, as in the proof of Lemma 4.3 (i) $\Rightarrow$ (ii), the equality

$$\varphi(S_\mu a S_\nu^*) = \frac{1}{|\beta|^{\mu}} \varphi(a \rho_\mu(1)), \quad a \in \mathcal{A}, \mu, \nu \in B_\epsilon(\Lambda)$$

holds so that

$$\phi(S_\mu x S_\nu^*) = \delta_{\mu,\nu} \frac{1}{|\beta|^{\mu}} \varphi(a \rho_\mu(1)), \quad a \in \mathcal{A}, \mu, \nu \in B_\epsilon(\Lambda) \text{ with } |\mu| = |\nu|. \quad (5.8)$$

By Lemma 4.3 (iii) $\Rightarrow$ (i), $\phi$ belongs to $\mathcal{E}_\beta^F(\rho)$. \hfill $\Box$

Recall that $E_\rho : \mathcal{O}_\rho \longrightarrow \mathcal{O}_\rho^\hat{\rho} = \mathcal{F}_\rho$ denotes the conditional expectation defined by (2.3).

Proposition 5.4. For any tracial continuous linear functional $\phi \in \mathcal{E}_\beta^F(\rho)$, the composition $\psi = \phi \circ E_\rho$ is a continuous linear functional on $\mathcal{O}_\rho$ which satisfies KMS condition at $\log \beta$ for gauge action $\hat{\rho}$ of $\mathcal{T}$.

Proof. Let $\mathcal{P}_\rho$ be the dense $*$-subalgebra of $\mathcal{O}_\rho$ generated algebraically by $S_\alpha, \alpha \in \Sigma$ and $a \in \mathcal{A}$. It is clear that for each element $x \in \mathcal{P}_\rho$ the function $t \in \mathcal{T} = \mathbb{R}/2\pi \mathbb{R} \rightarrow \hat{\rho}_t(x) \in \mathcal{O}_\rho$ extends to an entire analytic function on $\mathbb{C}$. Put $\psi = \phi \circ E_\rho$. We will show that the equality (5.2) holds for $\psi$. Elements $x, y \in \mathcal{F}_\rho$ can be expanded as finite linear combinations

$$x = \sum x \cdot x_\nu S_\nu^*, \quad y = \sum y \cdot y_\nu S_\nu^* + y_0 + \sum S_\mu y_\mu \quad (5.8)$$

for some $x \cdot x_\nu, x_0, y \cdot y_\nu, y_0, y_\mu \in \mathcal{F}_\rho^{\text{alg}}$. As $\psi$ is a tracial linear functional on $\mathcal{F}_\rho$, it suffices to check the equality (5.2) for the following two cases

1. $x = S_\nu x_\nu, \quad y = y \cdot y_\nu S_\nu^*$,
2. $x = x \cdot x_\nu^*, \quad y = S_\nu y_\mu$.

Case (1):

$$\psi(y \hat{\rho}_t(x)) = \psi(y \cdot y_\nu S_\nu^* e^{-|\nu| \log \beta} S_\nu x_\nu)$$

$$= \frac{1}{|\beta|^{\nu}} \psi(e^{i|\nu| \theta} x_\nu y \cdot y_\nu S_\nu^*)$$

$$= \psi(e^{i|\nu| \theta} S_\nu x_\nu y \cdot y_\nu S_\nu^*)$$

$$= \psi(\hat{\rho}_t(x)y).$$
Case (2):

\[
\psi(y_\beta t \log r(x)) = \psi(S_\mu y_\mu e^{\frac{\mu}{\beta} x - \mu} S_\mu^* y_\mu)
\]
\[
= \frac{1}{\beta^{\mu}} \psi(y_\mu x - \mu S_\mu^* S_\mu)
\]
\[
= \psi(e^{\frac{i\lambda r}{\mu}} x - \mu S_\mu^* S_\mu y_\mu)
\]
\[
= \psi(\hat{\rho}_\theta(x) y).
\]

This completes the proof. \(\blacksquare\)

Conversely we have

**Lemma 5.5.** If a continuous linear functional \(\psi\) on \(O_\rho\) satisfies KMS condition at \(\log \beta\) for some \(\beta \in \mathbb{C}\) with \(|\beta| > 1\), then the restriction \(\phi = \psi|_{F_\rho}\) to \(F_\rho\) belongs to \(\mathcal{E}_\beta^F(\rho)\) and satisfies the equality \(\psi = \phi \circ E_\rho\).

**Proof.** Let \(\beta = re^{i\theta}\) with \(r > 1\). For any \(x \in F_\rho, \mu \in B_\rho(A),\) we see

\[
\psi(S_\mu x) = \frac{1}{\beta^{\mu}} \psi(x S_\mu) = \frac{1}{\beta^{\mu}} \psi(S_\mu \hat{\rho}_t \log r(\alpha_{\theta}(x))) = \frac{1}{\beta^{\mu}} \psi(S_\mu x)
\]

so that \(\psi(S_\mu x) = 0\) because \(|\beta| > 1\). We similarly have \(\psi(x S_\mu^* y) = 0\). Since \(\alpha_{\theta}\) belongs to \(P_\rho\) and can be expanded as in (5.8), we get \(\psi(y) = \phi \circ E_\rho(y)\) for \(y \in P_\rho\). We will next show that \(\phi\) belongs to \(\mathcal{E}_\beta^F(\rho)\). For \(x, y \in F_\rho,\) one sees

\[
\hat{\rho}_t \log r(x) = \hat{\rho}_{t \theta}(x) = x
\]

so that \(\psi(xy) = \psi(yx)\). Hence \(\psi\) gives rise to a tracial linear functional \(\phi\) on \(F_\rho\). By KMS condition, we get for any \(x \in F_\rho, \mu \in B_\rho(A),\)

\[
\psi(S_\mu x S_\mu^*) = \psi(x S_\mu^* \hat{\rho}_t \log r(\hat{\rho}_{t \theta}(S_\mu))) = \frac{1}{\beta^{\mu}} \psi(x S_\mu^* S_\mu)
\]

Thus by Lemma 4.3, we know \(\phi \in \mathcal{E}_\beta^F(\rho)\). \(\blacksquare\)

We set for \(\beta \in \mathbb{C}\) with \(|\beta| > 1\),

\[
KMS_\beta(O_\rho) = \{\psi \in O_\rho^* \mid \psi\text{ satisfies KMS condition at }\log \beta \text{ for gauge action}\}
\]

and

\[
Sp(\rho) = \{\beta \in \mathbb{C} \mid \varphi \circ \lambda_\rho = \beta \varphi \text{ for some } \varphi \in A^* \text{ with } \varphi \neq 0\}.
\]

By Proposition 5.4 and Lemma 5.5, we have

**Proposition 5.6.** Let \((A, \rho, \Sigma)\) be an irreducible \(C^*\)-symbolic dynamical system. Assume that \((A, \rho, \Sigma)\) is power-bounded. Let \(\beta \in \mathbb{C}\) be a complex number with \(|\beta| > 1\). If \(|\beta| = r_\rho\) and \(\beta \in Sp(\rho)\), we have \(KMS_\beta(O_\rho) \neq \{0\}\). If in particular, \((A, \rho, \Sigma)\) is mean ergodic, \(KMS_\beta(O_\rho) \neq \{0\}\) if and only if \(|\beta| = r_\rho\) and \(\beta \in Sp(\rho)\).

**Proof.** Under the assumption that \((A, \rho, \Sigma)\) is power-bounded, any continuous linear functional \(\varphi \in \mathcal{E}_\beta(\rho)\) on \(A\) can uniquely extend to a continuous linear functional \(\hat{\varphi}\) on \(D_\rho\), that belongs to \(\mathcal{E}_\beta^D(\rho)\) if \(|\beta| = r_\rho\). By Proposition 5.4,
\[ \tilde{\varphi} \circ E_D \in \mathcal{E}_D^\beta(\rho) \] has an extension on \( \mathcal{O}_\rho \) as a continuous linear functional that satisfies KMS condition at \( \log \beta \).

Conversely, the restriction of a continuous linear functional \( KMS_\beta(\mathcal{O}_\rho) \) to the subalgebra \( \mathcal{A} \) yields a nonzero element of \( \mathcal{E}_\beta(\rho) \) which has continuous extension to \( D_\rho \). If in particular, \( (\mathcal{A}, \rho, \Sigma) \) is mean ergodic, \( |\beta| \) must be \( r_\rho \) by Theorem 4.9.

Therefore we conclude

**Theorem 5.7.** Let \( (\mathcal{A}, \rho, \Sigma) \) be an irreducible \( C^* \)-symbolic dynamical system. Let \( \beta \in \mathbb{C} \) be a complex number with \( |\beta| = r_\rho > 1 \).

(i) Suppose that \( (\mathcal{A}, \rho, \Sigma) \) is power-bounded. Then there exist linear isomorphisms among the four spaces \( \mathcal{E}_\beta(\rho), \mathcal{E}_D^\beta(\rho), \mathcal{E}_F^\beta(\rho) \) and \( KMS_\beta(\mathcal{O}_\rho) \) through the correspondences \( \varphi \in \mathcal{E}_\beta(\rho), \tilde{\varphi} \in \mathcal{E}_D^\beta(\rho), \tilde{\varphi} \circ E_D \in \mathcal{E}_F^\beta(\rho), \tilde{\varphi} \circ E_D \circ E_\rho \in KMS_\beta(\mathcal{O}_\rho) \) respectively. In particular, there exists a bijective correspondence between the set \( \mathcal{E}_\beta(\rho) \) of eigenvectors of \( \lambda_\rho^* \) for eigenvalue \( \beta \) consisting of continuous linear functionals on \( \mathcal{A} \) and the set \( KMS_\beta(\mathcal{O}_\rho) \) of continuous linear functionals on \( \mathcal{O}_\rho \) satisfying KMS condition at \( \log \beta \).

(ii) Suppose that \( (\mathcal{A}, \rho, \Sigma) \) is mean ergodic. Then the dimension \( \dim KMS_\beta(\mathcal{O}_\rho) \) of the space of continuous linear functionals on \( \mathcal{O}_\rho \) satisfying KMS condition at \( \log \beta \) is one if there exists a nonzero eigenvector of \( \lambda_\rho^* \) on \( \mathcal{A}^* \) for the eigenvalue \( \beta \). In particular there uniquely exists a faithful KMS state on \( \mathcal{O}_\rho \) at \( \log r_\rho \).

The following corollary is a generalization of [9, Theorem 6].

**Corollary 5.8.** Suppose that \( A \) is an irreducible matrix with entries in \{0, 1\} with its period \( p_A \). Let \( \beta \) be a complex number with \( |\beta| > 1 \).

(i) There exists a nonzero continuous linear functional on the Cantz-Krieger algebra \( \mathcal{O}_A \) satisfying KMS condition for gauge action at \( \log \beta \) if and only if \( \beta \) is a \( p_A \)-th root of the Perron-Frobenius eigenvalue \( r_A \) of \( A \).

(ii) The space of admitted continuous linear functionals on \( \mathcal{O}_A \) satisfying KMS condition for gauge action at \( \log \beta \) is of one-dimensional.

(iii) If in particular \( \beta = r_A \), the space of admitted continuous linear functionals on \( \mathcal{O}_A \) satisfying KMS condition for gauge action at \( \log r_A \) is the scalar multiples of a unique KMS state.

### 6. KMS states and invariant measures

In this section, we will study a relationship between KMS states on \( \mathcal{O}_\rho \) and invariant measures on \( D_\rho \) under \( \phi_\rho \). In what follows we assume that \( (\mathcal{A}, \rho, \Sigma) \) is irreducible and fix a faithful invariant state \( \tau \) on \( \mathcal{A} \).
We denote by $\|a\|_2$ the $L^2$-norm $\tau(a^*a)^{1/2}$ for $a \in A$, and by $\mathcal{H}_\tau$ the completion of $A$ by the norm $\| \cdot \|_2$. By the inequalities for $n \in \mathbb{N}$, $a \in A$

\[
\tau(\lambda^n_p(a^*)^*\lambda^n_p(a)) \leq \|\lambda^n_p\|\tau(\lambda^n_p(a^*a)) = \|\lambda^n_p\|\|\tau(a^*a)\| \leq \|\lambda^n_p\|^2\|a\|^2, 
\]

(6.1)

the operators $T^n_p$, $n \in \mathbb{N}$ induce bounded linear operators on $\mathcal{H}_\tau$. The induced operators on $\mathcal{H}_\tau$, which we also denote by $T^n_p$, $n \in \mathbb{N}$, are uniformly bounded in the operator norm on $\mathcal{H}_\tau$, if $(A, \rho, \Sigma)$ is power-bounded. We provide the following lemma, which shows power-boundedness of $(A, \rho, \Sigma)$ induces an ordinary mean ergodicity on $\mathcal{H}_\tau$, is a direct consequence from [22, p.73, Theorem 1.2]. We give a proof for the sake of completeness.

**Lemma 6.1.** Suppose that $(A, \rho, \Sigma)$ is irreducible and power-bounded. Then

\[
\lim_{n \to \infty} \frac{1 + T_\rho + T^2_\rho + \cdots + T^{n-1}_\rho}{n}
\]

converges to an idempotent $P_\rho$ on $\mathcal{H}_\tau$ under strong operator topology in $B(\mathcal{H}_\tau)$. The subspace $P_\rho \mathcal{H}_\tau$ consists of the vectors of $\mathcal{H}_\tau$ fixed under $T_\rho$.

**Proof.** The mean operators $M_n$, $n \in \mathbb{N}$ on $A$ defined by (3.1) naturally act on $\mathcal{H}_\tau$. Since $(A, \rho, \Sigma)$ is power-bounded, there exists a positive number $c > 0$ such that $\|T^n_\rho\| < c$ for all $n \in \mathbb{N}$. As $\|M_n\| < 1 + c$, $n \in \mathbb{N}$, the sequence $M_nv \in \mathcal{H}_\tau$, $n \in \mathbb{N}$ for a vector $v \in \mathcal{H}_\tau$ has a cluster point $v_0$ under the weak topology of $\mathcal{H}_\tau$. The identities

\[
(I - T_\rho)M_n = M_n(I - T_\rho) = \frac{1}{n}(I - T^n_\rho)
\]

imply the inequalities

\[
\|\left(\frac{1}{n}(I - T^n_\rho)\right)\| = \|M_n(I - T_\rho)\| = \frac{1}{n}\|I - T^n_\rho\| < \frac{1}{n}(1 + c). \quad (6.2)
\]

Hence we have $T_\rho v_0 = v_0$. Put

\[
Q_n = \frac{1}{n}\{I + T_\rho + (I + T_\rho + T^2_\rho) + \cdots + (I + T_\rho + \cdots + T^{n-2}_\rho)\}.
\]

Then we have $v - M_nv = (I - T_\rho)Q_nv$, $n \in \mathbb{N}$. Hence $v - v_0$ belongs to the weak closure $\mathcal{K}_\tau$ of the subspace $(I - T_\rho)\mathcal{H}_\tau$. The weak closure $\mathcal{K}_\tau$ is also the norm closure of the subspace $(I - T_\rho)\mathcal{H}_\tau$. For $w \in \mathcal{K}_\tau$, take $w_j \in (I - T_\rho)\mathcal{H}_\tau$ such that $\|w - w_j\|_2 \to 0$ and $w_j = (I - T_\rho)x_j$ for some $x_j \in \mathcal{H}_\tau$. Then we have by (6.2)

\[
\|M_nw\|_2 \leq \|M_n\|\|w - w_j\|_2 + \|M_n(I - T_\rho)x_j\|_2
\]

\[
\leq (1 + c)\|w - w_j\|_2 + \frac{1}{n}(1 + c)\|x_j\|_2
\]

so that $\lim_{n \to \infty} \|M_nw\|_2 = 0$. Since $M_nv - v_0 = M_nv - v_0$ and $v - v_0 \in \mathcal{K}_\tau$, one has

\[
\lim_{n \to \infty} \|M_nv - v_0\|_2 = 0.
\]
Put $P_\rho v = v_0$. The inequality

$$
\|M_n v - T_\rho M_n v\|_2 \leq \|(I - T_\rho)M_n v\|_2 < \frac{1}{n}(1 + c)\|v\|_2
$$

implies that $P_\rho = T_\rho P_\rho$ that is equal to $P_\rho T_\rho$. Therefore $P_\rho = M_n P_\rho = P_\rho M_n$ and hence $P_\rho = P_\rho^2$.

**Remark.** Under the same assumption above, one may prove that the limit

$$
\lim_{r \downarrow r_\rho} (r - r_\rho)R(r)
$$

for the resolvent $R(r) = (r - \lambda_\rho)^{-1}$ with $r > r_\rho$ converges to the idempotent $P_\rho$ on $H_r$ under strong operator topology in $B(H_r)$. Hence the equality

$$
\lim_{r \downarrow r_\rho} (r - r_\rho)R(r) = \lim_{n \to \infty} \frac{1 + T_\rho + T_\rho^2 + \cdots + T_\rho^{n-1}}{n}
$$

holds. We will give a proof of the equality (6.3). It is enough to consider the limit $\lim_{n \to \infty} \frac{1}{n}R(r_\rho + \frac{1}{n})$ instead of $\lim_{r \downarrow r_\rho} (r - r_\rho)R(r)$. As in the above proof, there exists $c > 0$ such that $\|T_\rho^k(a)\|_2 \leq c\|a\|_2$ for all $a \in A, k \in \mathbb{N}$. Put $R_n = \frac{1}{n}R(r_\rho + \frac{1}{n})$. Since for $y \in A$

$$
R(r_\rho + \frac{1}{n})y = \sum_{k=0}^{\infty} \frac{\lambda_\rho^k(y)}{(r_\rho + \frac{1}{n})^{k+1}}
$$

one has

$$
\|R(r_\rho + \frac{1}{n})y\|_2 \leq \sum_{k=0}^{\infty} \|T_\rho^k(y)\| \frac{r_\rho^k}{(r_\rho + \frac{1}{n})^{k+1}} \leq nc\|y\|_2
$$

and hence $\|R_n\| \leq c$ for $n \in \mathbb{N}$. The identities

$$(I - T_\rho)R_n = R_n(I - T_\rho) = \frac{1}{n r_\rho}(R_n - I)$$

hold so that we have

$$
\|(I - T_\rho)R_n\| = \|R_n(I - T_\rho)\| \leq \frac{1}{n r_\rho}(1 + c).
$$

A similar argument to the proof of Lemma 6.1 works so that for $u \in H_r$ by taking a cluster point $u_0$ of the sequence $R_n u, n \in \mathbb{N}$ under the weak topology of $H_r$ we have

$$
\lim_{n \to \infty} \|R_n u - u_0\|_2 = 0.
$$

Put $\hat{P}_\rho u = u_0$. The inequality $\|R_n u - T_\rho R_n u\|_2 \leq \frac{1}{n r_\rho}(1 + c)\|u\|_2$ implies that $\hat{P}_\rho = T_\rho \hat{P}_\rho$ that is equal to $\hat{P}_\rho T_\rho$. Hence $\hat{P}_\rho = R_n \hat{P}_\rho$ and $\hat{P}_\rho = \hat{P}_\rho^2$. The equality $\hat{P}_\rho = T_\rho \hat{P}_\rho$ implies $\hat{P}_\rho = M_n \hat{P}_\rho$ for all $n \in \mathbb{N}$ so that $\hat{P}_\rho = P_\rho \hat{P}_\rho$. Similarly the equalities $P_\rho = T_\rho P_\rho$ and $R_n = \sum_{k=0}^{\infty} T_\rho^k \frac{r_\rho^k}{(r_\rho + \frac{1}{n})^{k+1}}$ imply $P_\rho = R_n P_\rho$ for all $n \in \mathbb{N}$ so that $P_\rho = \hat{P}_\rho P_\rho$. As $\hat{P}_\rho \hat{P}_\rho = \hat{P}_\rho P_\rho$, one has $P_\rho = \hat{P}_\rho$. 

*Documenta Mathematica* 16 (2011) 133–175
We denote by \( \|a\| \) the \( L^1 \)-norm \( \tau(|a|) \) of \( a \in \mathcal{A} \), and by \( L^1(\mathcal{A}, \tau) \) the completion of \( \mathcal{A} \) by the norm \( \| \cdot \| \). The positive operators \( \lambda_\rho, T_\rho : \mathcal{A} \to \mathcal{A} \) and the state \( \tau : \mathcal{A} \to \mathbb{C} \) extend to \( L^1(\mathcal{A}, \tau) \) in natural way, that are also denoted by \( \lambda_\rho, T_\rho \) and \( \tau \) respectively.

**Lemma 6.2.** Suppose that \((\mathcal{A}, \rho, \Sigma)\) is uniquely ergodic and power-bounded. Then for \( a \in \mathcal{A} \) the limit \( \lim_{n \to \infty} M_n(a) \) converges in \( L^1(\mathcal{A}, \tau) \) under \( \| \cdot \|_1 \)-topology. In particular \( \lim_{n \to \infty} M_n(1) = x_\rho \) exists in \( L^1(\mathcal{A}, \tau) \) and satisfies the equalities

\[
\tau(x_\rho) = 1 \quad \text{and} \quad \lim_{n \to \infty} M_n(a) = \tau(a)x_\rho \quad \text{for} \ a \in \mathcal{A}. \tag{6.4}
\]

**Proof.** Since \((\mathcal{A}, \rho, \Sigma)\) is irreducible and power-bounded, \( \lim_{n \to \infty} M_n(a) \) for \( a \in \mathcal{A} \) converges in \( \mathcal{H}_\tau = L^2(\mathcal{A}, \tau) \) under \( \| \cdot \|_2 \)-norm by the previous lemma. By the inequality

\[
\| M_n(a) - M_m(a) \| \leq \| M_n(a) - M_m(a) \|_2, \quad a \in \mathcal{A}
\]

the limit \( \lim_{n \to \infty} M_n(a) \) exists in \( L^1(\mathcal{A}, \tau) \) under \( \| \cdot \|_1 \)-norm. We denote it by \( \Phi_1(a) \). Hence \( x_\rho = \Phi_1(1) \). We will show that \( \tau(f(\Phi_1(a) - \tau(a)x_\rho)) = 0 \) for \( f \in \mathcal{A} \). It suffices to show that \( \tau(b\Phi_1(a)b^*) = \tau(a)\tau(bx_\rho b^*) \) for \( b \in \mathcal{A} \). One may assume that \( a \geq 0 \). The inequality \( a \leq \|a\|1 \) and hence \( M_n(a) \leq \|a\|M_n(1) \) implies \( \tau(b\Phi_1(a)b^*) \leq \|a\|\tau(bx_\rho b^*) \) so that we have \( 0 \leq \tau(b\Phi_1(a)b^*) \leq \|a\|\tau(bx_\rho b^*) \). Hence \( \tau(bx_\rho b^*) = 0 \) implies \( \tau(b\Phi_1(a)b^*) = 0 \). We may assume that \( \tau(bx_\rho b^*) \neq 0 \). Put \( \omega(a) = \frac{\tau(b\Phi_1(a)b^*)}{\tau(bx_\rho b^*)}, a \in \mathcal{A} \). As \( \Phi_1 \circ T_\rho(a) = \Phi_1(a) \), one sees that \( \omega \) is an invariant state on \( \mathcal{A} \). Hence we have \( \omega = \tau \) by the unique ergodicity of \((\mathcal{A}, \rho, \Sigma)\). Therefore we have \( \tau(b\Phi_1(a)b^*) = \tau(a)\tau(bx_\rho b^*) \) for \( b \in \mathcal{A} \).

The equality \( \tau(x_\rho) = 1 \) is clear. \[
\blacksquare
\]

**Lemma 6.3.** Keep the above assumptions and notations. The limit \( \lim_{n \to \infty} M_n(f) \) for \( f \in L^1(\mathcal{A}, \tau) \) converges in \( L^1(\mathcal{A}, \tau) \) under \( \| \cdot \|_1 \)-topology and satisfies the equality

\[
\lim_{n \to \infty} M_n(f) = \tau(f)x_\rho \quad \text{for} \ f \in L^1(\mathcal{A}, \tau).
\]

**Proof.** Since for \( f \in L^1(\mathcal{A}, \tau) \) the inequality \( |\lambda_\rho(f)| \leq \lambda_\rho(|f|) \) holds, one has \( |T_\rho(f)| \leq T_\rho(|f|) \) and hence \( \|M_n(f)\|_1 \leq \|f\|_1 \). Take \( a_k \in \mathcal{A} \) such that \( |f - a_k|_1 \to 0 \) as \( k \to \infty \). It then follows that

\[
\|M_n(f) - \tau(f)x_\rho\|_1 \leq \|M_n(f) - M_n(a_k)\|_1 + \|M_n(a_k) - \tau(a_k)x_\rho\|_1 + \|\tau(a_k)x_\rho - \tau(f)x_\rho\|_1
\]

and hence \( \lim_{n \to \infty} \|M_n(f) - \tau(f)x_\rho\|_1 = 0 \) by the preceding lemma. \[
\blacksquare
\]

**Proposition 6.4.** Keep the above assumptions and notations. If \( f \in L^1(\mathcal{A}, \tau) \) satisfies \( T_\rho(f) = f \) and \( \tau(f) = 1 \), then \( f = x_\rho \). Namely the space of the fixed elements in \( L^1(\mathcal{A}, \tau) \) under \( T_\rho \) is one-dimensional.

\[
\text{Documenta Mathematica 16 (2011) 133–175}
\]
Proof. By the preceding lemma, we have for $f \in L^1(A, \tau)$ $\lim_{n \to \infty} M_n(f) = \tau(f)x_\rho$ in $\|\cdot\|_1$-topology. By the condition $T_\rho(f) = f$, we have $M_n(f) = f$ with $\tau(f) = 1$ and hence $f = x_\rho$. 

Let us define the space $L^1(D_\rho, \tau)$ in a similar way to $L^1(A, \tau)$. The operators $\lambda_\rho, T_\rho : D_\rho \to D_\rho$ and the state $\tau : D_\rho \to C$ naturally act on $L^1(D_\rho, \tau)$. The inclusion relation $A \subset D_\rho$ induces the inclusion relation $L^1(A, \tau) \subset L^1(D_\rho, \tau)$.

**Lemma 6.5.** Keep the above assumptions and notations. Let $x$ be an element of $L^1(D_\rho, \tau)$ such that $T_\rho(x) = x$. Then $x$ belongs to $L^1(A, \tau)$.

Proof. Take $x_n \in D_\rho \setminus \{0\}$ such that $\|x_n - x\|_1 \to 0$ as $n \to \infty$. As $|\lambda_\rho(y)| \leq \lambda_\rho(|y|)$, $y \in D_\rho$, it then follows that

$$
\|\lambda_\rho(x_n) - \lambda_\rho(x)\|_1 = \tau(|\lambda_\rho(x_n - x)|) \leq \tau(\lambda_\rho(|x_n - x|)) = r_\rho\|x_n - x\|_1
$$

so that $\|T_\rho(x_n) - T_\rho(x)\|_1 \leq \|x_n - x\|_1$. The element $x$ is fixed by $T_\rho$ so that

$$
\|T_\rho^n(x_n) - x\|_1 \leq \|x_n - x\|_1, \quad n \in \mathbb{N}, \quad k \in \mathbb{N}.
$$

Since $x_n \in D_\rho \setminus \{0\}$, there exists $k_n \in \mathbb{N}$ such that $T_\rho^{k_n}(x_n) \in A$. Hence $x$ belongs to $L^1(A, \tau)$. 

**Definition.** A state $\mu$ on $D_\rho$ is called a $\phi_\rho$-invariant measure if it satisfies

$$
\mu(y) = \mu(\phi_\rho(y)), \quad y \in D_\rho.
$$

If the probability measure for a state $\mu$ on $D_\rho$ is absolutely continuous with respect to the probability measure for the state $\tau$ on $D_\rho$, we write it as $\mu \ll \tau$.

**Proposition 6.6.** Assume that $(A, \rho, \Sigma)$ is irreducible and uniquely ergodic. For a fixed positive element $x \in L^1(A, \tau)$ by $T_\rho$ satisfying $\tau(x) = 1$, the state $\mu_x$ on $D_\rho$ defined by

$$
\mu_x(y) = \tau(yx), \quad y \in D_\rho
$$

is a $\phi_\rho$-invariant measure on $D_\rho$ such that $\mu \ll \tau$. Conversely, for any $\phi_\rho$-invariant measure $\mu$ on $D_\rho$ such that $\mu \ll \tau$, there exists a fixed positive element $x_\mu \in L^1(A, \tau)$ by $T_\rho$ satisfying $\tau(x_\mu) = 1$ such that

$$
\mu(y) = \tau(y x_\mu), \quad y \in D_\rho.
$$

Proof. Let $x \in L^1(A, \tau)$ be a fixed positive element by $T_\rho$ satisfying $\tau(x) = 1$. As $\lambda_\rho(x) = r_\rho x$, it follows that from Lemma 4.7

$$
\mu_x(\phi_\rho(y)) = \frac{1}{r_\rho} \tau(\lambda_\rho(\phi_\rho(y)x)) = \frac{1}{r_\rho} \tau(y \lambda_\rho(x)) = \mu_x(y), \quad y \in D_\rho
$$

so that the state $\mu_x$ is a $\phi_\rho$-invariant measure on $D_\rho$ such that $\mu_x \ll \tau$. Conversely for a $\phi_\rho$-invariant measure $\mu$ on $D_\rho$ such that $\mu \ll \tau$, there exists a Radon-Nikodym derivative $x_\mu \in L^1(D_\rho, \tau)$ such that $x_\mu \geq 0, \tau(x_\mu) = 1$ and

$$
\mu(y) = \tau(y x_\mu), \quad y \in D_\rho.
$$
By the equality $\tau(\phi_\rho(y)x_\mu) = \tau(yT_\rho(x_\mu)), y \in \mathcal{D}_\rho$, one sees that $\tau(yx_\mu) = \tau(yT_\rho(x_\mu)), y \in \mathcal{D}_\rho$ so that $T_\rho(x_\mu) = x_\mu$. Hence $x_\mu$ is regarded as an element of $L^1(\mathcal{A}, \tau)$ by the preceding lemma. This completes the proof. 

Especially the measure $\mu_\rho$ defined by $\mu_\rho(y) = \tau(yx_\mu), y \in \mathcal{D}_\rho$ is a $\phi_\rho$-invariant measure on $\mathcal{D}_\rho$ such that $\mu_\rho \ll \tau$.

Therefore we have

**Theorem 6.7.** Assume that $(\mathcal{A}, \rho, \Sigma)$ is irreducible, uniquely ergodic and power-bounded. Then a $\phi_\rho$-invariant measure on $\mathcal{D}_\rho$ absolutely continuous with respect to $\tau$ is unique and is of the form

$$\mu_\rho(y) = \tau(yx_\rho), \quad y \in \mathcal{D}_\rho.$$ (6.5)

**Proof.** Let $\mu$ be a $\phi_\rho$-invariant measure on $\mathcal{D}_\rho$. By the preceding proposition there exists a fixed positive element $x_\mu \in L^1(\mathcal{A}, \tau)$ under $T_\rho$ satisfying $\tau(x_\mu) = 1$ such that

$$\mu(y) = \tau(yx_\mu), \quad y \in \mathcal{D}_\rho.$$ By Proposition 6.4 we have $x_\mu = x_\rho$. For $x, y \in \mathcal{D}_\rho$, the equality

$$\lambda_\rho^k(\phi_\rho^k(y)xx_\rho) = y\lambda_\rho^k(xx_\rho)$$

holds by Lemma 4.7 so that

$$\frac{1}{n} \sum_{k=0}^{n-1} \mu_\rho(\phi_\rho^k(y)x) = \frac{1}{n} \sum_{k=0}^{n-1} \tau(\phi_\rho^k(y)xx_\rho)$$

$$= \frac{1}{n} \sum_{k=0}^{n-1} \frac{1}{\rho^k} \tau(\lambda_\rho^k(\phi_\rho^k(y)xx_\rho))$$

$$= \frac{1}{n} \sum_{k=0}^{n-1} \frac{1}{\rho^k} \tau(y\lambda_\rho^k(xx_\rho))$$

$$= \tau(yM_n(xx_\rho)).$$

Since

$$\| \cdot \|_1 - \lim_{n \to \infty} M_n(xx_\rho) = \tau(xx_\rho)x_\rho = \mu_\rho(x)x_\rho,$$

we have

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu_\rho(\phi_\rho^k(y)x) = \tau(y\mu_\rho(x)x_\rho)) = \mu_\rho(y)\mu_\rho(x).$$

□

**Corollary 6.8.** Assume that $(\mathcal{A}, \rho, \Sigma)$ is irreducible and mean ergodic.
Proof. (i) Under the assumption that \((A, \rho, \Sigma)\) is irreducible. Mean ergodicity implies unique ergodicity and (FP), which implies power-boundedness. Therefore the assertion is immediate.

(ii) By the mean ergodicity, the fixed element \(x_\rho\) belongs to \(A\) and is strictly positive by Lemma 3.5 (ii). Hence we have \(\tau(y) = \mu_\rho(yx_\rho^{-1}), y \in \mathcal{D}_\rho\) so that \(\tau \ll \mu_\rho\).

\[
\begin{array}{c}
\text{7. Examples} \\
\text{We will present examples of continuous linear functionals satisfying KMS conditions on some C}^*\text{-symbolic dynamical systems.}
\end{array}
\]

1. Finite directed graphs

Let \(A = [A(i, j)]_{i,j=1,...,N}\) be an \(N \times N\) matrix with entries in nonnegative integers. Denote by \(G_A = (V_A, E_A)\) the associated finite directed graph with vertex set \(V = \{v_1, \ldots, v_N\}\) and edge set \(E_A\). Let \(\mathcal{O}_{A[\Sigma]}\) be the Cuntz-Krieger algebra such that the generating partial isometries \(S_e, e \in E_A\) indexed by the edges in \(G_A\) satisfy

\[
\sum_{f \in E_A} S_f S_f^* = 1, \quad S_e^* S_e = \sum_{f \in E_A} A^{[2]}(e, f) S_f S_f^*, \quad e \in E_A,
\]

where \(A^{[2]}(e, f)\) is defined to be one if the edge \(f\) follows the edge \(e\), otherwise zero. Put \(\mathcal{A}_{G_A}\) the \(C^*\)-subalgebra of \(\mathcal{O}_{A[\Sigma]}\) generated by the projections \(S_e^* S_e, e \in E_A\). Denote by \(\rho^A_e\) for \(e \in E_A\) the endomorphism \(\mathcal{A}_{G_A}\) defined by \(\rho^A_e(a) = S_e^* a S_e\). Consider the \(C^*\)-symbolic dynamical system \((\mathcal{A}_{G_A}, \rho^A_e, E_A)\). Its associated \(C^*\)-algebra \(\mathcal{O}_{\rho_A}\) is nothing but the Cuntz-Krieger algebra \(\mathcal{O}_{A[\Sigma]}\). The finite directed graphs \(G_A\) is naturally considered to be a finite labeled graph by regarding an edge itself as its label. Hence this example will be contained in the following examples.

2. Finite labeled graphs

Let \(G = (G, \lambda)\) be a left-resolving finite labeled graph over \(\Sigma\) with underlying finite directed graph \(G = (V, E)\) and labeling map \(\lambda : E \rightarrow \Sigma\). Suppose that the graph \(G\) is irreducible. Let \(\{v_1, \ldots, v_N\}\) be the vertex set \(V\). As in Section 2, we have a \(C^*\)-symbolic dynamical system \((\mathcal{A}_G, \rho^G, \Sigma)\) such that \(\mathcal{A}_G = CE_1 \oplus \cdots \oplus CE_N\) and \(\rho^G_{e}(E_i) = \sum_{j=1}^N A^G(i, \alpha, j) E_j\) for \(i = 1, \ldots, N, \alpha \in \Sigma, \) where the \(N \times N\)-matrix \([A^G(i, \alpha, j)]_{i,j=1,...,N}\) for \(\alpha \in \Sigma\) is defined by (2.1).

Put \(A_G(i, j) = \sum_{\alpha \in \Sigma} A^G(i, \alpha, j)\) for \(i = 1, \ldots, N\). Then the matrix \(A_G = [A_G(i, j)]_{i,j=1}^N\) is irreducible. Let \(r_G\) denote the Perron-Frobenius eigenvalue of the matrix \(A_G\). It is easy to see that \(r_G\) is equal to the spectral radius \(r_{\rho^G}\) of

\begin{thebibliography}{9}
\end{thebibliography}
the positive operator $\lambda_{\rho^\phi}(x) = \sum_{n \in \Sigma} \rho_n^\phi(x)$, $x \in \mathcal{A}_G$. As

$$\lambda_{\rho^\phi}(E_i) = \sum_{j=1}^{N} A_G(i,j)E_j, \quad i = 1, \ldots, N,$$

by identifying $x = \sum_{i=1}^{N} x_i E_i \in \mathcal{A}_G$ with the vector $[x_i]_{i=1}^{N} \in \mathbb{C}^N$, one may regard the operator $\lambda_{\rho^\phi}$ as the transposed matrix $A_G^t$ of $\mathcal{A}_G$. For a complex number $\beta \in \mathbb{C}$ with $|\beta| > 1$, let $\phi \in \mathcal{A}_G^*$ be a continuous linear functional belonging to $\mathcal{E}_\beta(\rho^\phi)$. The equality $\phi \circ \lambda_{\rho^\phi}(E_i) = \beta \phi(E_i)$ implies

$$\sum_{j=1}^{N} A_G(i,j)\phi(E_j) = \beta \phi(E_i), \quad i = 1, \ldots, N,$$

so that the vector $[\phi(E_j)]_{j=1}^{N}$ is an eigenvector of $A_G$ for eigenvalue $\beta$. Conversely an eigenvector $[u_i]_{i=1}^{N} \in \mathbb{C}$ of the matrix $A_G$ for an eigenvalue $\beta$ gives rise to a continuous linear functional $\phi$ on $\mathcal{A}_G$ by setting $\phi(E_i) = u_i$, $i = 1, \ldots, N$ so that $\phi \in \mathcal{E}_\beta(\rho^\phi)$. Hence the space $\mathcal{E}_\beta(\rho^\phi)$ is identified with the eigenspace of the matrix $A_G$ for eigenvalue $\beta$. Especially a faithful invariant state $\tau$ on $\mathcal{A}_G$ is the positive normalized eigenvector of $A_G$ for eigenvalue $r_G$. Similarly an element $x = \sum_{j=1}^{N} x_j E_j \in \mathcal{A}_G$ is fixed by $T_{r_{\phi^\tau}}$ if and only if the vector $[x_j]_{j=1}^{N}$ is an eigenvector of $A_G^t$ for the eigenvalue $r_G$. The ordinary Perron-Frobenius theorem for nonnegative matrices asserts that $(A_G, \rho^\phi, \Sigma)$ is mean ergodic if $A_G$ is irreducible. The following proposition comes from the ordinary Perron-Frobenius theorem for irreducible nonnegative matrices, which is a special case of Theorem 3.13, and Corollary 6.8.

**Proposition 7.1.** Suppose that the adjacency matrix $A_G = [A_G(i,j)]_{i,j=1}^{N}$ is irreducible. Let $[r_i]_{i=1}^{N}$ and $[x_i]_{i=1}^{N}$ be right and left Perron eigenvector of $A_G$ respectively, then is

$$A_G[r_i]_{i=1}^{N} = r_G[r_i]_{i=1}^{N}, \quad A_G^t[x_i]_{i=1}^{N} = r_G[x_i]_{i=1}^{N},$$

such that $\sum_{i=1}^{N} r_i = 1$ and $\sum_{i=1}^{N} r_i x_i = 1$. Put $x_{\phi^\tau} = \sum_{i=1}^{N} x_i E_i \in \mathcal{A}_G$ and $\tau(a) = \sum_{i=1}^{N} r_i a_i$ for $a = \sum_{i=1}^{N} a_i E_i \in \mathcal{A}_G$. Then $\tau$ is a unique faithful invariant state on $\mathcal{A}_G$ such that the following equalities hold:

$$\lim_{n \to \infty} M_n(a) = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} T_{r^\phi}^k(a) = \tau(a) x_{\phi^\tau}.$$

Furthermore the measure $\mu_{\phi^\tau}$ on $\mathcal{D}_{\rho^\phi}$ defined $\mu_{\phi^\tau}(y) = \tau(y x_{\phi^\tau})$ for $y \in \mathcal{D}_{\rho^\phi}$ is a unique $\phi_{\rho^\phi}$-invariant measure equivalent to the measure $\tau$ on $\mathcal{D}_{\rho^\phi}$.

**Remark.** Let $X_G$ be the right one-sided sofic shift presented by $G$. The commutative $C^*$-algebra $C(X_G)$ on $X_G$ is naturally regarded as a $C^*$-subalgebra of $\mathcal{D}_{\rho^\phi}$ through the correspondence

$$\chi_\nu \in C(X_G) \longrightarrow S_{\rho^\phi} \in \mathcal{D}_{\rho^\phi}, \quad \nu \in B_k(\mathcal{A}_G).$$
where \(\chi_v\) is the characteristic function for the cylinder
\[
U_v = \{(x_i)_{i \in \mathbb{N}} \in X_G \mid x_1 = v_1, \ldots, x_k = x_k\}.
\]
The restriction of the \(\phi_{\rho}\)-invariant measure \(\mu_{\rho}\) on \(D_{\rho}\) to the subalgebra \(C(X_G)\) is nothing but a shift-invariant measure on \(X_G\) (cf. [21]).

We will next find continuous linear functionals on \(O_{\rho}G\) satisfying KMS conditions in concrete way. Now suppose that the irreducible matrix \(A_G\) has its period \(p_G\) and put
\[
N_G(i,j) = \{n \in \mathbb{Z}_+ \mid A_G^n(i,j) > 0\}.
\]
It is well-known that for \(n, m \in N_G(i,j)\) one has \(n \equiv m \pmod{p_G}\). Then for an eigenvalue \(\beta \in \mathbb{C}\) of \(A_G\) with \(|\beta| = r_G\), \(\beta r_G^{-1}\) is a \(p_G\)-th root of unity. We fix a vertex \(v_1\) and for \(k \in \{1, 2, \ldots, N\}\) take \(u_k \in N_G(1, k)\). We set
\[
u_k = (\beta r_G^{-1})^{u_k} \tau(E_k).
\]
Then \(u_k\) does not depend on the choice of \(n_k\) as long as \(n_k \in N_G(1, k)\).

**Lemma 7.2.** \(\sum_{j=1}^{N} A_G(i,j)u_j = \beta u_i, \ i = 1, \ldots, N\).

**Proof.** If \(A_G(i,j) \neq 0\), one sees \(n_1 + 1 \in N(1, j)\) so that
\[
A_G(i,j)u_j = \frac{\beta}{r_G} (\beta r_G^{-1})^{u_i} A_G(i,j) \tau(E_j) = \frac{\beta}{r_G} \frac{u_i}{\tau(E_j)} A_G(i,j) \tau(E_j).
\]
It follows that
\[
\sum_{j=1}^{N} A_G(i,j)u_j = \frac{\beta}{r_G} \frac{u_i}{\tau(E_i)} \sum_{j=1}^{N} A_G(i,j) \tau(E_j) = \frac{\beta}{r_G} \frac{u_i}{\tau(E_i)} \tau(E_i) = \beta u_i.
\]
Hence \(u = [u_k]_{k=1}^{N}\) yields a nonzero eigenvector of \(A_G\). Define a nonzero continuous linear functional \(\varphi\) on \(A_G\) by setting
\[
\varphi(E_k) = u_k, \ k = 1, \ldots, N
\]
so that the equality \(\varphi \circ \lambda_G = \beta \varphi\) on \(A_G\) holds. Put \(v_\varphi = \sum_{i=1}^{N} \tau(E_i)E_i \in A_G\).

It is easy to see that \(v_\varphi\) is a partial isometry such that \(\varphi(E_j) = \tau(E_j v_\varphi), j = 1, \ldots, N\) so that
\[
\varphi(x) = \tau(xv_\varphi), \ x \in A_G
\]
holds. Therefore we have the following proposition.

**Proposition 7.3.** Let \(G = (G, \lambda)\) be a left-resolving finite labeled graph with underlying finite directed graph \(G = (V, E)\) and labeling map \(\lambda : E \to \Sigma\). Denote by \(\{v_1, \ldots, v_N\}\) the vertex set \(V\). Assume that \(G\) is irreducible. Consider the \(N\)-dimensional commutative \(C^*\)-algebra \(A_G = C E_1 \oplus \cdots \oplus C E_N\) where each minimal projection \(E_i\) corresponds to the vertex \(v_i\) for \(i = 1, \ldots, N\). Define an
$N \times N$-nonnegative matrix $A_G = [A_G(i, j)]_{i,j=1}^{N}$ by $A_G(i, j) = \sum_{\alpha \in \Sigma} A_G(i, \alpha, j)$ where for $\alpha \in \Sigma$ and $i, j = 1, \ldots, N$

\[
A_G(i, \alpha, j) = \begin{cases} 
1 & \text{if there exists an edge } e \text{ from } v_i \text{ to } v_j \text{ with } \lambda(e) = \alpha, \\
0 & \text{otherwise.}
\end{cases}
\]

Let $O_{A_G}$ be the associated Cuntz-Krieger algebra and $\tau$ be the unique KMS state on $O_{A_G}$ for gauge action. Let $\beta \in \mathbb{C}$ be an eigenvalue of $A_G$ such that $|\beta| = r_G$ the Perron-Frobenius eigenvalue of the matrix $A_G$. Then a continuous linear functional on $O_{A_G}$ satisfying KMS condition at $\log \beta$ is a scalar multiple of $\varphi \in O_{A_G}^*$ giving by for $k = 1, \ldots, N$

\[
\varphi(E_k) = \left(\frac{\beta}{r_G}\right)^{n_k} \tau(E_k) \quad \text{where } n_k \text{ satisfies } A_G^{n_k}(1, k) \neq 0.
\]

Consider a finite labeled graph $G$ whose adjacency matrix $A$ is

\[
A = \begin{bmatrix}
0 & 0 & 1 & 1 & 0 \\
0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 \\
1 & 1 & 0 & 0 & 0
\end{bmatrix}.
\]

As

\[
A^2 = \begin{bmatrix}
2 & 2 & 0 & 0 & 0 \\
2 & 2 & 0 & 0 & 0 \\
0 & 0 & 2 & 2 & 0 \\
0 & 0 & 2 & 4 & 0 \\
0 & 0 & 0 & 0 & 4
\end{bmatrix},
\]

the period of the matrix is 3. The characteristic polynomial of $A$ is $\det(t - A) = t^2(t^3 - 4)$ so that $\text{Sp}(A) = \{\sqrt[3]{4}, \sqrt[3]{4} e^{\frac{2\pi i}{3}}, \sqrt[3]{4} e^{\frac{4\pi i}{3}}, 0\}$ and $r_A = \sqrt[3]{4}$. Hence $\beta \in \text{Sp}(A)$ satisfying $|\beta| = \sqrt[3]{4}$ are $\sqrt[3]{4}, \sqrt[3]{4} e^{\frac{2\pi i}{3}}, \sqrt[3]{4} e^{\frac{4\pi i}{3}}$.

Therefore the Cuntz-Krieger algebra $O_A$ has three continuous linear functionals satisfying KMS condition for gauge action at inverse temperatures $\frac{1}{3} \log 4$, $\frac{1}{3} \log 4 + \frac{2\pi i}{3}$, $\frac{1}{3} \log 4 + \frac{4\pi i}{3}$ respectively.

3. Dyck shifts

We consider the Dyck shift $D_N$ for a fixed natural number $N > 1$ with alphabet $\Sigma = \Sigma^- \cup \Sigma^+$ where $\Sigma^- = \{\alpha_1, \ldots, \alpha_N\}$, $\Sigma^+ = \{\beta_1, \ldots, \beta_N\}$. The symbols $\alpha_i, \beta_i$ correspond to the brackets $(,)$, respectively. The Dyck inverse monoid has the relations

\[
\alpha_i \beta_j = \begin{cases} 
1 & \text{if } i = j, \\
0 & \text{otherwise}
\end{cases}
\] (7.1)
for \(i, j = 1, \ldots, N\) (cf. [23],[26]). A word \(\omega_1 \cdots \omega_n\) of \(\Sigma\) is admissible for \(D_N\) precisely if \(\prod_{m=1}^{n} \omega_m \neq 0\). For a word \(\omega = \omega_1 \cdots \omega_n\) of \(\Sigma\), we denote by \(\tilde{\omega}\) its reduced form. Namely \(\tilde{\omega}\) is a word of \(\Sigma \cup \{0, 1\}\) obtained after the operations (7.1). Hence a word \(\omega\) of \(\Sigma\) is forbidden for \(D_N\) if and only if \(\tilde{\omega} = 0\).

In [26], an irreducible \(\lambda\)-graph system presenting \(D_N\) called the Cantor horizon \(\lambda\)-graph system has been introduced. It is a minimal irreducible component of the canonical \(\lambda\)-graph system \(\mathfrak{L}^C(D_N)\) and written as \(\mathfrak{L}^{\text{Ch}(D_N)}\). Let us describe the Cantor horizon \(\lambda\)-graph system \(\mathfrak{L}^{\text{Ch}(D_N)}\) of \(D_N\). Let \(\Sigma_N\) be the full \(N\)-shift \(\{1, \ldots, N\}^\mathbb{Z}\). We denote by \(B_l(D_N)\) and \(B_l(\Sigma_N)\) the set of admissible words of length \(l\) of \(D_N\) and that of \(\Sigma_N\) respectively. The vertices \(V_l\) of \(\mathfrak{L}^{\text{Ch}(D_N)}\) at level \(l\) are given by the words of length \(l\) consisting of the symbols of \(\Sigma^+\). That is,

\[
V_l = \{ (\beta_{\mu_1} \cdots \beta_{\mu_l}) \in B_l(D_N) \mid \mu_1 \cdots \mu_l \in B_l(\Sigma_N) \}.
\]

Hence the cardinal number of \(V_l\) is \(N^l\). The mapping \(\iota = \iota_{l,l+1} : V_{l+1} \to V_l\) deletes the rightmost symbol of a word in \(B_l(\Sigma_N)\) such as

\[
\iota((\beta_{\mu_1} \cdots \beta_{\mu_{l+1}})) = (\beta_{\mu_1} \cdots \beta_{\mu_l}), \quad (\beta_{\mu_1} \cdots \beta_{\mu_{l+1}}) \in V_{l+1}.
\]

There exists an edge labeled \(\alpha_j\) from \((\beta_{\mu_1} \cdots \beta_{\mu_l}) \in V_l\) to \((\beta_{\mu_0} \beta_{\mu_1} \cdots \beta_{\mu_l}) \in V_{l+1}\) precisely if \(\mu_0 = j\), and there exists an edge labeled \(\beta_j\) from \((\beta_{\mu_1} \cdots \beta_{\mu_l}) \in V_l\) to \((\beta_{\mu_1} \cdots \beta_{\mu_{l+1}}) \in V_{l+1}\). The resulting labeled Bratteli diagram with \(\iota\)-map becomes a \(\lambda\)-graph system over \(\Sigma\), denoted by \(\mathfrak{L}^{\text{Ch}(D_N)}\), that presents the Dyck shift \(D_N\) ([26]). It gives rise to a purely infinite simple \(C^*\)-algebra \(\mathcal{O}_{\mathfrak{L}^{\text{Ch}(D_N)}}\) ([32]) such that

\[
K_0(\mathcal{O}_{\mathfrak{L}^{\text{Ch}(D_N)}}) \cong \mathbb{Z}/N\mathbb{Z} \oplus \mathbb{C}(\mathbb{R}, \mathbb{Z}), \quad K_1(\mathcal{O}_{\mathfrak{L}^{\text{Ch}(D_N)}}) \cong 0.
\]

Let us denote by \((A^D_N, \rho^D_N, \Sigma)\) the \(C^*\)-symbolic dynamical system associated to the \(\lambda\)-graph system \(\mathfrak{L}^{\text{Ch}(D_N)}\) as in Section 2. Since the vertex set \(V_l\) is indexed by the set \(B_l(\Sigma_N)\) of words, the family of projections denoted by \(E_{\mu_1 \cdots \mu_l}\) for \(\mu_1 \cdots \mu_l \in B_l(\Sigma_N)\) in the \(C^*\)-algebra \(A^D_N\) forms the minimal projectins of \(A_l = C(V_l)\) such that

\[
\sum_{\mu_1 \cdots \mu_l \in B_l(\Sigma_N)} E_{\mu_1 \cdots \mu_l} = 1, \quad E_{\mu_1 \cdots \mu_l} = \sum_{\mu_{l+1}=1}^{N} E_{\mu_1 \cdots \mu_{l+1}}.
\]

As the algebra \(A_l\) is embedded into \(A_{l+1}\), the \(C^*\)-algebra \(A^D_N\) is a commutative AF-algebra generated by the subalgebras \(A_l, l \in \mathbb{N}\). The endomorphisms \(\rho^D_N : A^D_N \to A^D_N\) for \(\gamma \in \Sigma\) are defined by

\[
\rho^D_N(\epsilon_{\mu_1 \cdots \mu_l}) = \epsilon_{j\mu_1 \cdots \mu_l}, \quad \rho^D_N(\epsilon_{\mu_1 \cdots \mu_{l+1}}) = \sum_{\mu_{l+2}=1}^{N} \epsilon_{\mu_1 \cdots \mu_{l+2}}.
\]
for $\mu_1 \ldots \mu_N \in B_1(\Sigma_N)$ and $j = 1, \ldots, N$. It then follows that
\[
\lambda_{\rho_D}(1) = \sum_{j=1}^{N} \rho_{\alpha_j}^D(1) + \sum_{j=1}^{N} \rho_{\beta_j}^D(1)
\]
\[
= \sum_{j=1}^{N} \sum_{\mu_1 \cdot \mu_j \in B_1(\Sigma_N)} E_{\mu_1 \cdot \mu_j} + \sum_{j=1}^{N} \sum_{\mu_2 \cdot \mu_1 \in B_1(\Sigma_N)} \sum_{\mu_2 \cdot \mu_1 = 1}^N E_{\mu_2 \cdot \mu_1} + \sum_{j=1}^{N} E_{\mu_j}
\]
\[
= 1 + N
\]
so that we have $\|\lambda_{\rho_D} - \lambda_{\rho_D}(1)\| = 1 + N$. Hence we obtain
\[
r_{\rho_D} = 1 + N, \quad T_{\rho_D}(1) = 1.
\]
This implies that $1$ is a fixed element by $T_{\rho_D}$ and hence $(A^D, \rho_D, \Sigma)$ satisfies (FP). As in [32], $(A^D, \rho_D, \Sigma)$ is irreducible and uniquely ergodic, so that it is mean ergodic. One then sees that there exists a KMS state at inverse temperature $\log \beta$ if and only if $\beta = 1 + N$. The admitted KMS state is unique ([32, Theorem 1.2]).

4. $\beta$-shifts
Let $\beta > 1$ be an arbitrary real number. Take a natural number $N$ with $N - 1 < \beta \leq N$. Put $\Sigma = \{0, 1, \ldots, N - 1\}$. For a nonnegative real number $t$, we denote by $[t]$ the integer part of $t$. Let $f_\beta : [0, 1] \to [0, 1]$ be the mapping defined by
\[
f_\beta(x) = \beta x - \lfloor \beta x \rfloor, \quad x \in [0, 1]
\]
that is called the $\beta$-transformation ([38], [42]). The $\beta$-expansion of $x \in [0, 1]$ is a sequence $\{d_i(x, \beta)\}_{i \in \mathbb{N}}$ of integers of $\Sigma$ determined by
\[
d_i(x, \beta) = \lfloor \beta f_\beta^{-1}(x) \rfloor, \quad i \in \mathbb{N}.
\]
By this sequence, we can write $x$ as
\[
x = \sum_{i=1}^{\infty} \frac{d_i(x, \beta)}{\beta^i}.
\]
We endow the infinite product $\Sigma^\mathbb{N}$ with the product topology and the lexicographical order. Put $\zeta_{\beta} = \sup_{x \in [0, 1]}(d_i(x, \beta))_{i \in \Sigma^\mathbb{N}}$. We define the shift-invariant compact subset $X_\beta$ of $\Sigma^\mathbb{N}$ by
\[
X_\beta = \{ \omega \in \Sigma^\mathbb{N} | \sigma^i(\omega) \leq \zeta_{\beta}, \ i = 0, 1, 2, \ldots \},
\]
where $\sigma$ denotes the shift $\sigma((\omega_i)_{i \in \mathbb{N}}) = (\omega_{i+1})_{i \in \mathbb{N}}$. The one-sided subshift $(X_\beta, \sigma)$ is called the right one-sided $\beta$-shift (cf. [38], [42]). Its (two-sided) subshift
\[
A_\beta = \{ (\omega_i)_{i \in \mathbb{Z}} \in \Sigma^\mathbb{Z} | (\omega_{i-k})_{i \in \mathbb{Z}} = X_\beta, k = 0, 1, 2, \ldots \}
\]
is called the $\beta$-shift. In [17], the C*-algebra $O_\beta$ associated with the $\beta$-shift has been introduced and studied. It is simple and purely infinite for every $\beta > 1$ and generated by $N - 1$ isometries $S_0, S_1, \ldots, S_{N-2}$ and one partial isometry $S_{N-1}$ having certain operator relations (see [17]). The family $O_\beta$, $1 < \beta \in \mathbb{R}$.
exists a positive constant $M > 1$.

We note that $(\tau_\beta)$ is irreducible, uniquely ergodic and power-bounded. 

**LEMMA 7.4.** The spectral radius $r_\beta$ of the positive operator $\lambda_\beta$ on $A_\beta$ is $\beta$.

**Proof.** Denote by $\theta_k$ the cardinal number of the admissible words $B_k(\Lambda_\beta)$ of length $k$. Then we have

$$\|\lambda_\beta^k\| = \|\lambda_\beta^k(1)\| \leq \sum_{\mu \in B_k(\Lambda_\beta)} \|S_\mu^*S_\mu\| = \theta_k.$$

As in [44, p. 179], $\lim_{k \to \infty} \frac{\theta_k}{\beta^k}$ converges to a positive real number so that there exists a positive constant $M > 0$ such that $\|\lambda_\beta^k\|^{\frac{1}{k}} < M$ for all $k \in \mathbb{N}$. Hence $\lim_{k \to \infty} \|\lambda_\beta^k\|^{\frac{1}{k}} < \beta$ so that $r_\beta \leq \beta$. As in [17], there exists a state $\tau$ on $A_\beta$ satisfying $\tau \circ \lambda_\beta = \beta \tau$. This implies $\beta \in \text{Sp}(\lambda_\beta)$ so that $r_\beta = \beta$. 

**PROPOSITION 7.5.** $(A_\beta, \rho^\beta, \Sigma)$ is irreducible, uniquely ergodic and power-bounded.

**Proof.** It has been proved in [17] that there is no nontrivial ideal of $A_\beta$ invariant under $\lambda_\beta$ and there exists a unique state $\tau$ on $A_\beta$ satisfying $\tau \circ \lambda_\beta = r_\beta \tau$. Hence $(A_\beta, \rho^\beta, \Sigma)$ is irreducible, uniquely ergodic. As in the proof of the above lemma, there exists a positive constant $M > 0$ such that $\|\lambda_\beta^k\|^{\frac{1}{k}} < M$ for all $k \in \mathbb{N}$. This means that $(A_\beta, \rho^\beta, \Sigma)$ is power-bounded.

By the above proposition, one knows that $(A_\beta, \rho^\beta, \Sigma)$ satisfies the hypothesis of Theorem 6.7 so that there uniquely exists a $\phi_{\rho^\beta}$-invariant measure on $D_{\rho^\beta}$ absolutely continuous with respect to the restriction of the unique KMS-state $\tau$ to $D_{\rho^\beta}$. We note that $C(X_\beta)$ is a $C^*$-subalgebra of $D_{\rho^\beta}$ and the restriction of $\phi_{\rho^\beta}$ to $C(X_\beta)$ comes from the shift transformation $\sigma$. As in [17], the restriction of the KMS-state $\tau$ to $D_{\rho^\beta}$ corresponds to the Lebesgue measure on $[0, 1]$ in translating the $\beta$-shift to the $\beta$-transformation. Hence the uniqueness of the $\phi_{\rho^\beta}$-invariant measure on $D_{\rho^\beta}$ absolutely continuous with respect to $\tau$ exactly corresponds to the uniqueness of the invariant measure on $[0, 1]$ under the $\beta$-transformation absolutely continuous with respect to the Lebesgue
measure studied in [14], [38] and [42]. In fact, the density function $h_\beta$ appeared in [14], [38] and [42] of the invariant measure for the $\beta$-transformation with respect to the Lebesgue measure is the element $x_{\rho,\beta}$ realized as the mean

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \frac{\lambda^k(1)}{\beta^k}$$

in Theorem 6.7.

Acknowledgments: The author would like to thank Yasuo Watatani who pointed out inaccuracy in the proof of Lemma 3.4 in the first draft of the paper. The author is also deeply indebted to the referee whose many useful suggestions, guidences and comments improved the paper. This work was supported by JSPS Grant-in-Aid for Scientific Research ((C), No 20540215).

References

[1] T. Bates and D. Pask, The $C^*$-algebras of labelled graphs, J. Operator Theory 57 (2007), pp. 207–226.
[2] R. Bowen, Equilibrium states and the ergodic theory of Anosov diffeomorphisms, Lecture Notes in Math. 470, Springer-Verlag, New York, Berlin, Heidelberg (1975).
[3] O. Bratteli and D. W. Robinson, Operator algebras and quantum statistical mechanics II, Springer-Verlag, New York, Berlin, Heidelberg (1981).
[4] J. Cuntz, Simple $C^*$-algebras generated by isometries, Comm. Math. Phys. 57 (1977), pp. 173–185.
[5] J. Cuntz, A class of $C^*$-algebras and topological Markov chains II: reducible chains and the Ext-functor for $C^*$-algebras, Invent. Math. 63 (1980), pp. 25–40.
[6] J. Cuntz and W. Krieger, A class of $C^*$-algebras and topological Markov chains, Invent. Math. 56 (1980), pp. 251–268.
[7] J. Cuntz and G. K. Pedersen, Equivalence and KMS states on periodic $C^*$-dynamical systems, J. Funct. Anal. 34 (1979), pp. 79–86.
[8] N. Dunford and J. T. Schwartz, Linear operators, Part I : General Theory, Interscience Publishers, New York (1966).
[9] M. Enomoto, M. Fujii and Y. Watatani, KMS states for gauge action on $O_A$, Math. Japon. 29 (1984), pp. 607–619.
[10] R. Exel, KMS states for generalized gauge actions on Cuntz–Krieger algebras, An application of the Ruelle–Perron–Frobenius Theorem, Bull. Braz. Math. Soc., New Series 35 (2004), pp. 1–12.
[11] R. Exel and M. Laca, Partial dynamical systems and the KMS condition, Commun. Math. Phys. 232 (2002), pp. 223–277.
[12] R. Fischer, Sofic systems and graphs, Monats. für Math. 80 (1975), pp. 179–186.
[13] T. Furuta, *Invitation to Linear Operators, from matrices to bounded linear operators on a Hilbert space*, Taylor & Francis, London (2001).

[14] F. Hofbauer, *β-shifts have unique maximal measure*, Monatsh. Math. 85(1978), pp. 189–198.

[15] M. Izumi, T. Kajiwara and Y. Watatani, *KMS states and branched points*, Ergodic Theory Dynam. Systems 27(2007), pp. 1887–1918.

[16] T. Kajiwara, C. Pinzari and Y. Watatani, *Ideal structure and simplicity of the C*-algebras generated by Hilbert modules*, J. Funct. Anal. 159(1998), pp. 295–322.

[17] Y. Katayama, K. Matsumoto and Y. Watatani, *Simple C*-algebras arising from β-expansion of real numbers*, Ergodic Theory Dynam. Systems 18(1998), pp. 937–962.

[18] D. Kerr and C. Pinzari, *Noncommutative pressure and the variational principle in Cuntz-Krieger-type C*-algebras*, J. Funct. Anal. 188(2002), pp. 156–215.

[19] M. Kessebohmer, M. Stadlbauer and B. O. Stratmann, *Lyapunov spectra for KMS states on Cuntz-Krieger algebras*, Math. Z. 256(2007), pp. 871–893.

[20] B. P. Kitchens, *Symbolic dynamics*, Springer-Verlag, New York, Berlin, Heidelberg (1998).

[21] B. P. Kitchens and S. Tuncel, *Finitary measures for subshifts of finite type and sofic systems*, Memoirs of Amer. Math. Soc. 58(1985), No 338.

[22] U. Krengel, *Ergodic Theorems*, de Gruyter Studies in Mathematics 6, Walter de Gruyter, Berlin, New York (1985).

[23] W. Krieger, *On the uniqueness of the equilibrium state*, Math. Systems Theory 8(1974), pp. 97–104.

[24] W. Krieger, *On sofic systems I*, Israel J. Math. 48(1984), pp. 305–330.

[25] W. Krieger, *On sofic systems II*, Israel J. Math. 60(1987), pp. 167–176.

[26] W. Krieger and K. Matsumoto, *A lambda-graph system for the Dyck shift and its K-groups*, Documenta Math. 8(2003), pp. 79–96.

[27] M. Laca and S. Neshveyev, *KMS-states of quasi-free dynamics on Pimsner algebras*, J. Funct. Anal. 211(2004), pp. 457–482.

[28] D. Lind and B. Marcus, *An introduction to symbolic dynamics and coding*, Cambridge University Press, Cambridge (1995).

[29] K. Matsumoto, *Presentations of subshifts and their topological conjugacy invariants*, Documenta Math. 4(1999), pp. 285–340.

[30] K. Matsumoto, *C*-algebra associated with presentations of subshifts*, Documenta Math. 7(2002), pp. 1–30.
Ergodic Properties and KMS Conditions

[31] K. Matsumoto, *Actions of symbolic dynamical systems on C*-algebras*, J. Reine Angew. Math. 605(2007), pp. 23–49.

[32] K. Matsumoto, *On the simple C*-algebras arising from Dyck systems*, J. Operator Theory 58(2007), pp. 205–226.

[33] K. Matsumoto, *Orbit equivalence in C*-algebras defined by actions of symbolic dynamical systems*, Contemporary Math. 503(2009), pp. 121–140.

[34] K. Matsumoto, *Actions of symbolic dynamical systems on C*-algebras II, Simplicity and some examples*, Math. Z. 265(2010), pp. 735–760.

[35] K. Matsumoto, Y. Watatani and M. Yoshida, *KMS-states for gauge actions on C*-algebras associated with subshifts*, Math. Z. 228(1998), pp. 489–509.

[36] R. Okayasu, *Type III factors arising from Cuntz-Krieger algebras*, Proc. Amer. Math. Soc. 131(2003), pp. 2145–2153.

[37] D. Olesen and G. K. Pedersen, *Some C*-dynamical system with a single KMS state*, Math. Scand. 42(1978), pp. 111–118.

[38] W. Parry, *On the β-expansion of real numbers*, Acta Math. Acad. Sci. Hungar. 11(1960), pp. 401–416.

[39] G. K. Pedersen, *C*-algebras and their automorphism groups*, Academic Press, London, New York, San Francisco (1979).

[40] M. V. Pimsner, *A class of C*-algebras generalizing both Cuntz-Krieger algebras and crossed products by Z*, in Free Probability Theory, Fields Institute Communications 12(1996), pp. 189–212.

[41] C. Pinzari, Y. Watatani and K. Yonetani, *KMS-states, entropy and the variational principle in full C*-dynamical systems*, Comm. Math. Phys. 213(2000), pp. 231–379.

[42] A. Rényi, *Representations for real numbers and their ergodic properties*, Acta Math. Acad. Sci. Hungar. 8(1957), pp. 477–493.

[43] M. Takesaki, *Theory of Operator Algebras I*, Springer-Verlag, New York, Berlin, Heidelberg (1979).

[44] P. Walters, *An Introduction to Ergodic Theory*, Graduate Texts in Mathematics 79, Springer-Verlag, New York, Berlin, Heidelberg (1975).

[45] B. Weiss, *Subshifts of finite type and sofic systems*, Monats. Math. 77, (1973), pp. 462–474.

Kengo Matsumoto
Department of Mathematics,
Joetsu University of Education,
Joetsu 943-8512, Japan
kengo@juen.ac.jp
AN INFINITE LEVEL ATOM COUPLED TO A HEAT BATH

MARTIN KOENENBERG

Received: January 13, 2011

Communicated by Heinz Siedentop

Abstract. We consider a $W^*$-dynamical system $(M_\beta, \tau)$, which models finitely many particles coupled to an infinitely extended heat bath. The energy of the particles can be described by an unbounded operator, which has infinitely many energy levels. We show existence of the dynamics $\tau$ and existence of a $(\beta, \tau)$-KMS state under very explicit conditions on the strength of the interaction and on the inverse temperature $\beta$.

2010 Mathematics Subject Classification: Primary 81V10; Secondary 37N20, 47N50.

Keywords and Phrases: KMS-state, thermal equilibrium, $W^*$-dynamical system, open quantum system.

1 Introduction

In this paper, we study a $W^*$-dynamical system $(M_\beta, \tau)$ which describes a system of finitely many particles interacting with an infinitely extended bosonic reservoir or heat bath at inverse temperature $\beta$. Here, $M_\beta$ denotes the $W^*$-algebra of observables and $\tau$ is an automorphism-group on $M_\beta$, which is defined by

$$\tau_t(X) := e^{it L_Q} X e^{-it L_Q}, \ X \in M_\beta, \ t \in \mathbb{R}.$$ (1)

In this context, $t$ is the time parameter. $L_Q$ is the Liouvillean of the dynamical system at inverse temperature $\beta$. $Q$ describes the interaction between particles and heat bath. On the one hand the choice of $L_Q$ is motivated by heuristic arguments, which allow to derive the Liouvillean $L_Q$ from the Hamiltonian $H$ of the joint system of particles and bosons at temperature zero. On the other hand we ensure that $L_Q$ anti-commutes with a certain anti-linear conjugation $J$, that will be introduced later on. The Hamiltonian, which represents the interaction

---

1Supported by the DFG (SFB/TR 12)
with a bosonic gas at temperature zero, can be the Standard Hamiltonian of the non-relativistic QED, (see or instance [2]), or the Pauli-Fierz operator, which is defined in [7, 2], or the Hamiltonian of Nelson’s Model. We give the definition of these Hamiltonians in the sequel of Definition 11.

Our first result is the following:

**Theorem 1.1.** $L_Q$, defined in (16), has a unique self-adjoint realization and $\tau_t(X) \in \mathfrak{M}_\beta$ for all $t \in \mathbb{R}$ and all $X \in \mathfrak{M}_\beta$.

The proof follows from Theorem 4.2 and Lemma 5.2. The main difficulty in the proof is, that $L_Q$ is not semi-bounded, and that one has to define a suitable auxiliary operator in order to apply Nelson’s commutator theorem.

Partly, we assume that the isolated system of finitely many particles is confined in space. This is reflected in Hypothesis 1, where we assume that the particle Hamiltonian $H_{el}$ possesses a Gibbs state. In the case where $H_{el}$ is a Schrödinger-operator, we give in Remark 2.1 a sufficient condition on the external potential $V$ to ensure the existence of a Gibbs state for $H_{el}$. Our second theorem is

**Theorem 1.2.** Assume Hypothesis 1 and that $\Omega^\beta_0 \in \text{dom}(e^{-\beta/L_0})$. Then there exists a $(\beta, \tau)$-KMS state $\omega^\beta$ on $\mathfrak{M}_\beta$.

This theorem ensures the existence of an equilibrium state on $\mathfrak{M}_\beta$ for the dynamical system $(\mathfrak{M}_\beta, \tau)$. Its proof is part of Theorem 5.3 below. Here, $L_0$ denotes the Liouvillean for the joint system of particles and bosons, where the interaction part is omitted. $\Omega^\beta_0$ is the vector representative of the $(\beta, \tau)$-KMS state for the system without interaction. In a third theorem we study the condition $\Omega^\beta_0 \in \text{dom}(e^{-\beta/L_0})$:

**Theorem 1.3.** Assume Hypothesis 1 is fulfilled. Then there are two cases,

1. If $0 \leq \gamma < 1/2$ and $\eta_1 (1 + \beta) \ll 1$, then $\Omega^\beta_0 \in \text{dom}(e^{-\beta/L_0})$.
2. If $\gamma = 1/2$ and $(1 + \beta)(\eta_1 + \eta_2) \ll 1$, then $\Omega^\beta_0 \in \text{dom}(e^{-\beta/L_0})$.

Here, $\gamma \in [0, 1/2)$ is a parameter of the model, see (32) and $\eta_1, \eta_2$ are parameters, which describe the strength of the interaction, see (32). In a last theorem we consider the case where $H_{el} = -\Delta_q + \Theta q^2$ and the interaction Hamiltonian is $\lambda q \Phi(f)$ at temperature zero for $\lambda \neq 0$. Then,

**Theorem 1.4.** $\Omega^\beta_0$ is in $\text{dom}(e^{-\beta/L_0})$ for all $\beta \in (0, \infty)$, whenever

$$|2\Theta^{-1} \lambda| |\eta| \Gamma^{-1/2} f |_{\eta_{ph}} < 1.$$ 

Furthermore, we show that our strategy can not be improved to obtain a result, which ensures existence for all values of $\lambda$, see (60).

In the last decade there appeared a large number of mathematical contributions to the theory of open quantum system. Here we only want to mention
some of them [3, 6, 8, 9, 10, 13, 14, 15], which consider a related model, in which the particle Hamilton \( H_{el} \) is represented as a finite symmetric matrix and the interaction part of the Hamiltonian is linear in annihilation and creation operators. In this case one can prove existence of a \( \beta \)-KMS without any restriction to the strength of the coupling. (In this case we can apply Theorem 1.3 with \( \gamma = 0 \) and \( \eta_1 = 0 \)). We can show existence of KMS-states for an infinite level atom coupled to a heat bath. Furthermore, in [6] there is a general theorem, which ensures existence of a \((\beta, \tau)\)-KMS state under the assumption, that \( \Omega_0^\beta \in \text{dom}(e^{-(\beta/2)Q}) \), which implies \( \Omega_0^\beta \in \text{dom}(e^{-(\beta/2)(L_0+Q)}) \). In Remark 7.3 we verify that this condition implies the existence of a \((\beta, \tau)\)-KMS state in the case of a harmonic oscillator with dipole interaction \( \lambda q \cdot \Phi(f) \), whenever \((1 + \beta)\lambda\|((1 + |k|^{-1/2})f)|| \ll 1 \).

2 Mathematical Preliminaries

2.1 Fock Space, Field-Operators and Second Quantization

We start our mathematical introduction with the description of the joint system of particles and bosons at temperature zero. The Hilbert space describing bosons at temperature zero is the bosonic Fock space \( \mathcal{F}_b \), where

\[
\mathcal{F}_b := \mathcal{F}_b[\mathcal{H}_{ph}] := \mathbb{C} \oplus \bigoplus_{n=1}^{\infty} \mathcal{H}^{(n)}_{ph}, \quad \mathcal{H}^{(n)}_{ph} := \bigotimes_{\text{sym}}^{n} \mathcal{H}_{ph}.
\]

\( \mathcal{H}_{ph} \) is either a closed subspace of \( L^2(\mathbb{R}^3) \) or \( L^2(\mathbb{R}^3 \times \{\pm\}) \), being invariant under complex conjugation. If phonons are considered we choose \( \mathcal{H}_{ph} = L^2(\mathbb{R}^3) \), if photons are considered we choose \( \mathcal{H}_{ph} = L^2(\mathbb{R}^3 \times \{\pm\}) \). In the latter case "+" or "-" labels the polarization of the photon. However, we will write \( \langle f | g \rangle_{\mathcal{H}_{ph}} := \int \overline{f(k)} g(k) \, dk \) for the scalar product in both cases. This is an abbreviation for \( \sum_{p = \pm} \int \overline{f(k, p)} g(k, p) \, dk \) in the case of photons. \( \mathcal{H}^{(n)}_{ph} \) is the \( n \)-fold symmetric tensor product of \( \mathcal{H}_{ph} \), that is, it contains all square integrable functions \( f_n \) being invariant under permutations \( \pi \) of the variables, i.e., \( f_n(k_1, \ldots, k_n) = f_n(k_{\pi(1)}, \ldots, k_{\pi(n)}) \). For phonons we have \( k_j \in \mathbb{R}^3 \) and \( k_j \in \mathbb{R}^3 \times \{\pm\} \) for photons. The wave functions in \( \mathcal{H}^{(n)}_{ph} \) are states of \( n \) bosons.

The vector \( \Omega := (1, 0, \ldots) \in \mathcal{F}_b \) is called the vacuum. Furthermore we denote the subspace \( \mathcal{F}_b \) of finite sequences with \( \mathcal{F}_b^{\text{fin}} \). On \( \mathcal{F}_b^{\text{fin}} \) the creation and
annihilation operators, $a^*(h)$ and $a(h)$, are defined for $h \in \mathcal{H}_{ph}$ by

\begin{align}
(a^*(h) f_n)(k_1, \ldots, k_{n+1}) &= (n+1)^{-1/2} \sum_{i=1}^{n+1} h(k_i) f_n(k_1, \ldots, k_{i-1}, k_{i+1}, \ldots, k_{n+1}), \\
(a(h) f_{n+1})(k_1, \ldots, k_n) &= (n+1)^{1/2} \int h(k_{n+1}) f_{n+1}(k_1, \ldots, k_{n+1}) dk_{n+1},
\end{align}

and $a^*(h) \Omega = h$, $a(h) \Omega = 0$. Since $a^*(h) \subset (a(h))^*$ and $a(h) \subset (a^*(h))^*$, the operators $a^*(h)$ and $a(h)$ are closable. Moreover, the canonical commutation relations (CCR) hold true, i.e.,

$[a(h), a(\tilde{h})] = [a^*(h), a^*(\tilde{h})] = 0$, $[a(h), a^*(\tilde{h})] = (h | \tilde{h})_{\mathcal{H}_{ph}}$.

Furthermore we define field operator by

$\Phi(h) := 2^{-1/2} (a(h) + a^*(h))$, $h \in \mathcal{H}_{ph}$.

It is a straightforward calculation to check that the vectors in $\mathcal{F}_b^{fin}$ are analytic for $\Phi(h)$. Thus, $\Phi(h)$ is essentially self-adjoint on $\mathcal{F}_b^{fin}$. In the sequel, we will identify $a^*(h)$, $a(h)$ and $\Phi(h)$ with their closures. The Weyl operators $W(h)$ are given by $W(h) = \exp(i \Phi(h))$. They fulfill the CCR-relation for the Weyl operators, i.e.,

$W(h) W(g) = \exp(i/2 \text{Im} (h | g)_{\mathcal{H}_{ph}}) W(g + h),$

which follows from explicit calculations on $\mathcal{F}_b^{fin}$. The Weyl algebra $W(f)$ over a subspace $\mathfrak{f}$ of $\mathcal{H}_{ph}$ is defined by

$W(f) := \text{cl} LH\{ W(g) \in \mathcal{B}(\mathcal{F}_b) : g \in \mathfrak{f} \}.$

Here, $\text{cl}$ denotes the closure with respect to the norm of $\mathcal{B}(\mathcal{F}_b)$, and "LH" denotes the linear hull.

Let $\alpha : \mathbb{R}^3 \to [0, \infty)$ be a locally bounded Borel function and $\text{dom}(\alpha) := \{ f \in \mathcal{H}_{ph} : \alpha f \in \mathcal{H}_{ph} \}$. Note, that $(\alpha f)(k)$ is given by $\alpha(k) f(k, p)$ for photons. If $\text{dom}(\alpha)$ is dense subspace of $\mathcal{H}_{ph}$, $\alpha$ defines a self-adjoint multiplication operator on $\mathcal{H}_{ph}$. In this case, the second quantization $d\Gamma(\alpha)$ of $\alpha$ is defined by

$(d\Gamma(\alpha) f_n)(k_1, \ldots, k_n) := (\alpha(k_1) + \alpha(k_2) + \ldots + \alpha(k_n)) f_n(k_1, \ldots, k_n)$

and $d\Gamma(\alpha) \Omega = 0$ on its maximal domain.
2.2 Hilbert space and Hamiltonian for the particles

Let $\mathcal{H}_{el}$ be a closed, separable subspace of $L^2(X, d\mu)$, that is invariant under complex conjugation. The Hamiltonian $H_{el}$ for the particle is a self-adjoint operator on $\mathcal{H}_{el}$ being bounded from below. We set $H_{el,+} := H_{el} - \inf \sigma(H_{el}) + 1$. Partly, we need the assumption

**Hypothesis 1.** Let $\beta > 0$. There exists a small positive constant $\epsilon > 0$, and

$$\text{Tr}_{\mathcal{H}_{el}} \{ e^{-(\beta - \epsilon) H_{el}} \} < \infty.$$ 

The condition implies the existence of a Gibbs state

$$\omega_\beta^\epsilon(A) = Z^{-1} \text{Tr}_{\mathcal{H}_{el}} \{ e^{-\beta H_{el}} A \}, \quad A \in \mathcal{B}(\mathcal{H}_{el}),$$

for $Z = \text{Tr}_{\mathcal{H}_{el}} \{ e^{-\beta H_{el}} \}$.

**Remark 2.1.** Let $\mathcal{H}_{el} = L^2(\mathbb{R}^n, d^n x)$ and $H_{el} = -\Delta_x + V_1 + V_2$, where $V_1$ is a $-\Delta_x$-bounded potential with relative bound $a < 1$ and $V_2$ is in $L^1_{loc}(\mathbb{R}^n, d^n x)$. Thus $H_{el}$ is essentially self-adjoint on $C_\infty^\infty(\mathbb{R}^n)$. Moreover, if additionally

$$\int e^{-(\beta - \epsilon) V_2(x)} d^n x < \infty,$$

then one can show, using the Golden-Thompson-inequality, that Hypothesis 1 is satisfied.

2.3 Hilbert space and Hamiltonian for the interacting system

The Hilbert space for the joint system is $\mathcal{H} := \mathcal{H}_{el} \otimes \mathcal{F}_b$. The vectors in $\mathcal{H}$ are sequences $f = (f_n)_{n \in \mathbb{N}_0}$ of wave functions, $f_n \in \mathcal{H}_{el} \otimes \mathcal{H}_{ph}^{(n)}$, obeying

$$\vec{k}_n \mapsto f_n(x, \vec{k}_n) \in \mathcal{H}_{ph}^{(n)} \quad \text{for } \mu_\nu \text{- almost every } x,$$

$$x \mapsto f_n(x, \vec{k}_n) \in \mathcal{H}_{el} \quad \text{for } \text{Lebesgue- almost every } \vec{k}_n,$$

where $\vec{k}_n = (k_1, \ldots, k_n)$. The complex conjugate vector is $\vec{f} := (\overline{f_n})_{n \in \mathbb{N}_0}$.

Let $G^j := \{ G_k^j \}_{k \in \mathbb{R}^3}$, $H^j := \{ H_k^j \}_{k \in \mathbb{R}^3}$ and $F := \{ F_k \}_{k \in \mathbb{R}^3}$ be families of closed operators on $\mathcal{H}_{el}$ for $j = 1, \ldots, r$. We assume, that dom($F_k^*$), dom($F_k^\dagger$) $\supset$ dom($H_{el,+}^{1/2}$) and that

$$k \mapsto G_k^j, \ (H_k^j), \ F_k H_{el,+}^{-1/2}, \ (F_k)^* H_{el,+}^{-1/2} \in \mathcal{B}(\mathcal{H}_{el})$$

are weakly (Lebesgue-)measurable. For $\phi \in \text{dom}(H_{el,+}^{1/2})$ we assume that

$$k \mapsto (G_k^j \phi)(x), \ (H_k^j \phi)(x), \ (F_k \phi)(x) \in \mathcal{H}_{ph},$$

$$k \mapsto ((G_k^j)^* \phi)(x), \ ((H_k^j)^* \phi)(x), \ ((F_k)^* \phi)(x) \in \mathcal{H}_{ph}, \text{ for } x \in X.$$
Moreover we assume for $\vec{G} = (G^1, \ldots, G^r)$, $\vec{H} := (H^1, \ldots, H^r)$ and $F$, that

$$\|\vec{G}\|_w < \infty, \|\vec{H}\|_w < \infty, \|F\|_{w,1/2} < \infty,$$

where

$$\|G_j\|_w^2 := \int (\alpha(k) + \alpha(k)^{-1}) \left( \|G_j^*\|^2_{B(H_{el})} + \|G_j\|^2_{B(H_{el})} \right) dk$$

$$\|\vec{G}\|_w^2 := \sum_{j=1}^r \|G_j\|_w^2, \quad \|F\|_{w,1/2} := \|FH_{el,+}^{-1/2}\|_w^2 + \|F^*H_{el,+}^{-1/2}\|_w^2.$$ 

We define for $f = (f_n)_{n=0}^\infty \in \text{dom}(H_{el,+}^{1/2}) \otimes F_{b}^{fin}$ the (generalized) creation operator

$$(a^*(F) f_n)(x, k_1, \ldots, k_n+1)$$

$$:= (n+1)^{-1/2} \sum_{i=1}^{n+1} (F_k, f_n)(x, k_1, \ldots, k_{i-1}, k_{i+1}, \ldots, k_{n+1})$$

and $a(F) f_0(x) = 0$. The (generalized) annihilation operator is

$$(a(F) f_n)(x, k_1, \ldots, k_n)$$

$$:= (n+1)^{1/2} \int (F_{k_{n+1}}, f_{n+1})(x, k_1, \ldots, k_n, k_{n+1}) dk_{n+1}.$$ 

Moreover, the corresponding (generalized) field operator is $\Phi(F) := 2^{-1/2}(a(F) + a^*(F))$. $\Phi(F)$ is symmetric on $\text{dom}(H_{el,+}^{1/2}) \otimes F_{b}^{fin}$. The bounds follow directly from Equations (8) and (9).

$$\|a(F)H_{el,+}^{-1/2} f\|_H^2 \lesssim \int |\alpha(k)|^{-1} \|F_k^*H_{el,+}^{-1/2}\|_{B(H_{el})} dk \cdot \|d\Gamma(|\alpha|)^{1/2} f\|_H^2 (0)$$

$$\|a^*(F)H_{el,+}^{-1/2} f\|_H^2 \lesssim \int |\alpha(k)|^{-1} \|F_k H_{el,+}^{-1/2}\|_{B(H_{el})} dk \cdot \|d\Gamma(|\alpha|)^{1/2} f\|_H^2 + \int \|F_k H_{el,+}^{-1/2}\|_{B(H_{el})}^2 dk \cdot \|f\|_H^2.$$ 

For $(G_k)^j, (H_k)^j \in B(H_{el})$, the factor $H_{el,+}^{-1/2}$ can be omitted. The Hamiltonians for the non-interacting, resp. interacting model are

**Definition 2.2.** On $\text{dom}(H_{el}) \otimes \text{dom}(d\Gamma(|\alpha|)) \cap F_{b}^{fin}$ we define

$$H_0 := H_{el} \otimes 1 + 1 \otimes d\Gamma(|\alpha|), \quad H := H_0 + W,$$

where $W := \Phi(\vec{G}) \Phi(\vec{H}) + \text{h.c.} + \Phi(F)$ and $\Phi(\vec{G}) \Phi(\vec{H}) := \sum_{j=1}^r \Phi(G^j) \Phi(H^j)$. The abbreviation "h.c." means the formal adjoint operator of $\Phi(\vec{G}) \Phi(\vec{H})$.
We give examples for possible configurations:

Let $\gamma \in \mathbb{R}$ be a small coupling parameter.

- The Nelson Model:
  
  $\mathcal{H}_{cl} \subset L^2(\mathbb{R}^3), H_{cl} := -\Delta + V$, $\mathcal{H}_{ph} = L^2(\mathbb{R}^3)$ and $\alpha(k) = |k|$. The form factor is $F_k = \gamma \sum_{\nu=1}^N e^{-ikx_\nu} |k|^{-1/2} 1[|k| \leq \kappa]$, $x_\nu \in \mathbb{R}^3$ and $H^3$, $G^{\alpha} = 0$.

- The Standard Model of Nonrelativistic QED:
  
  $\mathcal{H}_{el} \subset L^2(\mathbb{R}^3)^N$, $H_{el} := -\Delta + V$, $\mathcal{H}_{ph} = L^2(\mathbb{R}^3 \times \{\pm\})$ and $\alpha(k) = |k|$. The form factors are
  
  \begin{align*}
  F_k &= 4\gamma^{3/2} \pi^{-1/2} \sum_{\nu=1}^N (-i\nabla_{x_\nu} \cdot e(k,p)) e^{-i\gamma^{1/2} k x_\nu} (2|k|)^{-1/2} 1[|k| \leq \kappa] + \text{h. c.}, \\
  G^{\alpha}_{i,k} &= H^{\alpha}_{i,k} = 2\gamma^{3/2} \pi^{-1/2} \epsilon_i(k,p) e^{-i\gamma^{1/2} k x_\nu} (2|k|)^{-1/2} 1[|k| \leq \kappa]
  \end{align*}

  for $i = 1, 2, 3, \nu = 1, \ldots, N, x_\nu \in \mathbb{R}^3$ and $k = (k,p) \in \mathbb{R}^3 \times \{\pm\}$. $\epsilon_i(k, \pm) \in \mathbb{R}^3$ are polarization vectors.

- The Pauli-Fierz-Model:
  
  $\mathcal{H}_{el} \subset L^2(\mathbb{R}^3)^N$, $H_{el} := -\Delta + V$, $\mathcal{H}_{ph} = L^2(\mathbb{R}^3)$ or $\mathcal{H}_{ph} = L^2(\mathbb{R}^3 \times \{\pm\})$, and $\alpha(k) = |k|$. The form factor is $F_k = \gamma \sum_{\nu=1}^N 1[|k| \leq \kappa] k \cdot x_\nu$ and $G^{\alpha}_{i,k} = H^{\alpha}_{i,k} = 0$.

3 The Representation $\pi$

In order to describe the particle system at inverse temperature $\beta$ we introduce the algebraic setting. For $\mathfrak{f} = \{f \in \mathcal{H}_{ph} : \alpha^{-1/2} f \in \mathcal{H}_{ph}\}$ we define the algebra of observables by

\[ \mathfrak{A} = \mathcal{B}(\mathcal{H}_{el}) \otimes \mathcal{W}(f). \]

For elements $A \in \mathfrak{A}$ we define $\pi^0(A) := e^{itH_0} A e^{-itH_0}$ and $\pi^i(A) := e^{itH} A e^{-itH}$. We first discuss the model without interaction.

3.1 The Representation $\pi_f$

The time-evolution for the Weyl operators is given by

\[ e^{itH} \mathcal{W}(f) e^{-itH} = \mathcal{W}(e^{it\alpha} f). \]

For this time-evolution an equilibrium state $\omega^\beta_f$ is defined by

\[ \omega^\beta_f(\mathcal{W}(f)) = \langle f | (1 + 2 \varrho^\beta) f \rangle_{\mathcal{H}_{ph}}, \]

where $\varrho^\beta(k) = \left( \exp(\beta \alpha(k)) - 1 \right)^{-1}$. It describes an infinitely extended gas of bosons with momentum density $\varrho^\beta$ at temperature $\beta$. Since $\omega^\beta_f$ is a quasi-free state on the Weyl algebra, the definition of $\omega^\beta_f$ extends to polynomials of
creation and annihilation operators. One has
\[
\omega_f^\beta(a(f)) = \omega_f^\beta(a^*(f)) = \omega_f^\beta(a(f)a(g)) = \omega_f^\beta(a^*(f)a^*(g)) = 0,
\]
\[
\omega_f^\beta(a^*(f)a(g)) = \langle g | \varrho_\beta f \rangle_{H_{ph}}.
\]
For polynomials of higher degree one can apply Wick’s theorem for quasi-free states, i.e.,
\[
\omega_f^\beta(a^{\sigma_2m}(f_{2m}) \cdots a^{\sigma_i}(f_1)) = \sum_{P \in \mathbb{Z}_2} \prod_{i > j \in P} \omega_f^\beta(a^{\sigma_i}(f_i)a^{\sigma_j}(f_j)),
\]
where \(a^{\sigma_k} = a^*\) or \(a^{\sigma_k} = a\) for \(k = 1, \ldots, 2m\). \(\mathbb{Z}_2\) are the pairings, that is \(P \in \mathbb{Z}_2\) if \(P = \{Q_1, \ldots, Q_m\}, \#Q_i = 2\) and \(\bigcup_{i=1}^m Q_i = \{1, \ldots, 2m\}\).

The Araki-Woods isomorphism \(\pi_f : W(f) \rightarrow B(F_0 \otimes F_0)\) is defined by
\[
\pi_f[W(f)] := \mathcal{W}_\beta(f) := \exp(i \Phi_\beta(f)),
\]
\[
\Phi_\beta(f) := \Phi((1 + \varrho_\beta)^{1/2} f) \otimes 1 + 1 \otimes \Phi(\varrho_\beta^{1/2} f).
\]
The vector \(\Omega_f^\beta := \Omega \otimes \Omega\) fulfills
\[
\omega_f^\beta(W(f)) = \langle \Omega_f^\beta | \pi_f[W(f)] | \Omega_f^\beta \rangle. \tag{13}
\]

3.2 The representation \(\pi_{cl}\)

The particle system without interaction has the observables \(B(\mathcal{H}_{el})\), the states are defined by density operators \(\rho\), i.e., \(\rho \in B(\mathcal{H}_{el})\), \(0 \leq \rho\), \(\text{Tr}\{\rho\} = 1\). The expectation of \(A \in B(\mathcal{H}_{el})\) in \(\rho\) at time \(t\) is
\[
\text{Tr}\{ \rho e^{it\mathcal{H}_{el}} A e^{-it\mathcal{H}_{el}} \}.
\]

Since \(\rho\) is a compact, self-adjoint operator, there is an ONB \((\phi_n)_n\) of eigenvectors, with corresponding (positive) eigenvalues \((p_n)_n\). Let
\[
\sigma(x, y) = \sum_{n=1}^{\infty} p_n^{1/2} \phi_n(x) \overline{\phi_n(y)} \in \mathcal{H}_{el} \otimes \mathcal{H}_{el}. \tag{14}
\]

For \(\psi \in \mathcal{H}_{el}\) we define \(\sigma \psi := \int \sigma(x, y) \psi(y) \, d\mu(y)\). Obviously, \(\sigma\) is an operator of Hilbert-Schmidt class. Note, \(\overline{\sigma \psi} := \overline{\sigma} \overline{\psi}\) has the integral kernel \(\overline{\sigma(x, y)}\). It is a straightforward calculation to verify that
\[
\text{Tr}\{ \rho e^{it\mathcal{L}_{el}} A e^{-it\mathcal{L}_{el}} \} = (e^{-it\mathcal{L}_{el}} \sigma | (A \otimes 1) e^{-it\mathcal{L}_{el}} \sigma )_{\mathcal{H}_{el} \otimes \mathcal{H}_{el}},
\]
where \(\mathcal{L}_{el} = \mathcal{H}_{el} \otimes 1 - 1 \otimes \overline{\mathcal{H}_{el}}\). This suggests the definition of the representation
\[
\pi_{cl} : B(\mathcal{H}_{el}) \rightarrow B(\mathcal{H}_{el} \otimes \mathcal{H}_{el}), \quad A \mapsto A \otimes 1.
\]
Now, we define the representation map for the joint system by
\[ \pi : \mathcal{A} \rightarrow \mathcal{B}(\mathcal{K}), \quad \pi := \pi_{el} \otimes \pi_f, \]
where \( \mathcal{K} := \mathcal{H}_{el} \otimes \mathcal{H}_{el} \otimes \mathcal{F}_b \otimes \mathcal{F}_b \). Let \( \mathfrak{M}_\beta := \pi[\mathfrak{A}]'' \) be the enveloping \( W^* \)-algebra, here \( \pi[\mathfrak{A}]'' \) denotes the commutant of \( \pi[\mathfrak{A}] \), and \( \pi[\mathfrak{A}]'' \) the bicommutant.

We set \( \mathcal{D} := \cup_n \mathcal{H}_{fin} \otimes \mathcal{F}_b \), with compact support, and \( \mathcal{U}_1 := \cup_n \mathcal{H}_{el} \otimes \mathcal{F}_b \). On \( \mathcal{D} \) the operator \( L_0 \),
\[
L_0 := L_{el} \otimes 1 + 1 \otimes L_f, \quad \text{on} \mathcal{K},
\]
\[
L_f := d\Gamma(\alpha) \otimes 1 - 1 \otimes d\Gamma(\alpha), \quad \text{on} \mathcal{F}_b \otimes \mathcal{F}_b,
\]
is essentially self-adjoint and we can define
\[
\tau_t^0(X) := e^{itL_0}Xe^{-itL_0} \in \mathfrak{M}_\beta, \quad X \in \mathfrak{M}_\beta, \quad t \in \mathbb{R},
\]
which is not hard to see, that
\[
\pi[\tilde{\tau}_t^0(A)] = \tau_t^0(\pi[A]), \quad A \in \mathfrak{A}, \quad t \in \mathbb{R}
\]
On \( \mathcal{K} \) we introduce a conjugation by
\[
\mathcal{J}(\phi_1 \otimes \phi_2 \otimes \psi_1 \otimes \psi_2) = \bar{\phi_2} \otimes \bar{\phi_1} \otimes \bar{\psi_2} \otimes \bar{\psi_1}.
\]
It is easily seen, that \( \mathcal{J} L_0 = -L_0 \mathcal{J} \). In this context one has \( \mathfrak{M}_\beta' = \mathcal{J} \mathfrak{M}_\beta \mathcal{J} \), see for example [4]. In the case, where \( \mathcal{H}_{el} \) fulfills Hypothesis 1, we define the vector representative \( \Omega_{el} \in \mathcal{H}_{el} \otimes \mathcal{H}_{el} \) of the Gibbs state \( \omega_{el} \) as in (14)
\[
X \Omega_{el}^\beta \in \text{dom}(e^{-\beta/2L_0}), \quad \mathcal{J} X \Omega_{el}^\beta = e^{-\beta/2L_0}X^* \Omega_{el}^\beta
\]
for all \( X \in \mathfrak{M}_\beta \) and \( L_0 \Omega_{el}^\beta = 0 \). Moreover,
\[
\omega^\beta(X) := (\Omega_{el}^\beta | X \Omega_{el}^\beta)_\mathcal{K}, \quad X \in \mathfrak{M}_\beta
\]
is a \((\tau^0, \beta)\)-KMS-state for \( \mathfrak{M}_\beta \), i.e., for all \( X, Y \in \mathfrak{M}_\beta \) exists \( F_{\beta}(X, Y, \cdot) \), analytic in the strip \( S_{\beta} = \{ z \in \mathbb{C} : 0 < \text{Im}z < \beta \} \), continuous on the closure and taking the boundary conditions
\[
F_{\beta}(X, Y, t) = \omega^\beta_0(X \tau_t^0(Y))
\]
\[
F_{\beta}(X, Y, t + i\beta) = \omega^\beta_0(\tau_t^0(Y)X)
\]
For a proof see [14].
4 The Liouvillean $\mathcal{L}_Q$

In this and the next section we will introduce the Standard Liouvillean $\mathcal{L}_Q$ for a dynamics $\tau$ on $\mathfrak{M}_\beta$, describing the interaction between particles and bosons at inverse temperature $\beta$. The label $Q$ denotes the interaction part of the Liouvillean, it can be deduced from the interaction part $W$ of the corresponding Hamiltonian by means of formal arguments, which we will not give here. In a first step we prove self-adjointness of $\mathcal{L}_Q$ and of other Liouvilleans. A main difficulty stems from the fact, that $\mathcal{L}_Q$ and the other Liouvilleans, mentioned before, are not bounded from below. The proof of self-adjointness is given in Theorem 4.2, it uses Nelson’s commutator theorem and auxiliary operators which are constructed in Lemma 4.1. The proof, that $\mathcal{L}_Q$ is self-adjoint and affiliated with $\mathcal{H}_\beta$ by means of formal arguments, which we will not give here. In a first step we prove self-adjointness of $\mathcal{L}_Q$ and of other Liouvilleans. A main difficulty stems from the fact, that $\mathcal{L}_Q$ and the other Liouvilleans, mentioned before, are not bounded from below. The proof of self-adjointness is given in Theorem 5.3.

For each family $K = \{K_k\}_k$ of closed operators on $\mathcal{H}_{el}$ with $\|K\|_{w,1/2} < \infty$ we set

$$Q := \Phi_\beta(\hat{G}) \Phi_\beta(\hat{H}) + \text{h.c.} + \Phi_\beta(F), \quad \Phi_\beta(\hat{G}) \Phi_\beta(\hat{H}) := \sum_{j=1}^r \Phi_\beta(G^j) \Phi_\beta(H^j).$$

For each family $K = \{K_k\}_k$ of closed operators on $\mathcal{H}_{el}$ with $\|K\|_{w,1/2} < \infty$ we set

$$\Phi_\beta(K) := (a^*(1 + \varrho_\beta)^{1/2} K) \otimes 1 + 1 \otimes a^*(\varrho_\beta^{1/2} K^*) + \text{h.c.}$$

Here, $K_k$ acts as $K_k \otimes 1$ on $\mathcal{H}_{el} \otimes \mathcal{H}_{el}$. A Liouvillean, that describes the dynamics of the joint system of particles and bosons is the so-called Standard Liouvillean

$$\mathcal{L}_Q \phi := (\mathcal{L}_0 + Q - Q^\mathcal{J}) \phi, \quad \phi \in \mathcal{D},$$

which is distinguished by $\mathcal{J} \mathcal{L}_Q = -\mathcal{L}_Q \mathcal{J}$. For an operator $A$, acting on $\mathcal{K}$, the symbol $A^\mathcal{J}$ is an abbreviation for $\mathcal{J} A \mathcal{J}$. An important observation is, that $[Q, Q^\mathcal{J}] = 0$ on $\mathcal{D}$. Next, we define four auxiliary operators on $\mathcal{D}$

$$\mathcal{L}^{(1)}_a := (H_{el,\pm} \otimes 1 + 1 \otimes H_{el,\pm}) \otimes 1 + 1 \otimes L_{f,a} + 1$$

(17)

$$\mathcal{L}^{(2)}_a := H_{el,\pm}^{Q^\mathcal{J}} + (H_{el,\pm}^Q)^\mathcal{J} + c_1 1 \otimes L_{f,a} + c_2$$

$$\mathcal{L}^{(3)}_a := H_{el,\pm}^{Q^\mathcal{J}} + (H_{el,\pm}^Q)^\mathcal{J} + c_1 1 \otimes L_{f,a} + c_2$$

$$\mathcal{L}^{(4)}_a := H_{el,\pm} \otimes 1 + (H_{el,\pm}^{Q^\mathcal{J}})^\mathcal{J} + c_1 1 \otimes L_{f,a} + c_2,$$
where \( L_{f,a} \) is an operator on \( \mathcal{F}_b \otimes \mathcal{F}_b \) and \( H^Q_{\text{el}+} \) acts on \( K \). Furthermore,

\[
L_{f,a} = d \Gamma (1 + \alpha) \otimes 1 + 1 \otimes d \Gamma (1 + \alpha) + 1,
\]

\[
L_{\text{el},a} = H_{\text{el},+} \otimes 1 + 1 \otimes H^Q_{\text{el},+} =: H_{\text{el},+} \otimes 1 + Q.
\]

Obviously, \( L^{(i)}_a, i = 1, 2, 3, 4 \) are symmetric operators on \( D \).

**Lemma 4.1.** For sufficiently large values of \( c_1, c_2 \geq 0 \) we have that \( L^{(i)}_a, i = 1, 2, 3, 4 \) are essentially self-adjoint and positive. Moreover, there is a constant \( c_3 > 0 \) such that

\[
c_3^{-1} \| L^{(1)}_a \phi \| \leq \| L^{(i)}_a \phi \| \leq c_3 \| L^{(1)}_a \phi \|, \quad \phi \in \text{dom}(L^{(1)}_a).
\]

**Proof.** Let \( a, a' \in \{l, r\} \) and \( K_i, i = 1, 2 \) be families of bounded operators with \( \| K_i \| \omega < \infty \). Let \( \Phi_i(K_i) = \Phi(K_i) \otimes 1 \) and \( \Phi_f(K_i) := 1 \otimes \Phi(K_i) \). We have for \( \phi \in D \)

\[
\| \Phi_a(\eta K_i) \Phi_{a'}(\eta' K_2) \phi \| \leq \text{const} \| L_{f,a} \phi \|
\]

\[
\| \Phi_a(\eta F) \phi \| \leq \text{const} \| (L_{\text{el},a})^{1/2} (L_{f,a})^{1/2} \phi \|
\]

where \( \eta, \eta' \in \{(1 + a_2) \sqrt{2}, a_2 \sqrt{2}\} \). Note, that the estimates hold true, if \( \Phi_a(\eta K_i) \) or \( \Phi_a(\eta F) \) are replaced by \( \Phi_a(\eta K_i)' \) or \( \Phi_a(\eta F)' \). Thus, we obtain for sufficiently large \( c_1 \gg 1 \), depending on the form-factors, that

\[
\| Q \phi \| + \| Q^J \phi \| \leq 1/2 \| (L_{\text{el},a} + c_1 L_{f,a}) \phi \|.
\]

By the Kato-Rellich-Theorem ( [17], Thm. X.12) we deduce that \( L^{(i)}_a \) is self-adjoint on \( \text{dom}(L_{\text{el},a} + c_1 L_{f,a}) \), bounded from below and that \( L_{\text{el},a} + c_1 L_{f,a} \) is \( L^{(i)}_a \)-bounded for every \( c_2 \gg 0 \) and \( i = 2, 3, 4 \). In particular, \( D \) is a core of \( L^{(i)}_a \). The proof follows now from \( \| L^{(i)}_a \phi \| \leq \| (L_{\text{el},a} + c_1 L_{f,a}) \phi \| \leq c_1 \| L^{(i)}_a \phi \| \) for \( \phi \in D \).

**Theorem 4.2.** The operators

\[
L_0, \quad L_Q = L_0 + Q - Q^J, \quad L_0 + Q, \quad L_0 - Q^J,
\]

defined on \( D \), are essentially self-adjoint. Every core of \( L^{(i)}_a \) is a core of the operators in line (21).

**Proof.** We restrict ourselves to the case of \( L_Q \). We check the assumptions of Nelson's commutator theorem ([17], Thm. X.37). By Lemma 4.1 it suffices to show \( \| L_Q \phi \| \leq \text{const} \| L^{(i)}_a \phi \| \) and \( |(L_Q \phi) L^{(2)}_a \phi) - (L^{(2)}_a \phi) L_Q \phi) | \leq \text{const} \| (L^{(i)}_a)^{1/2} \phi \|^2 \) for \( \phi \in D \). The first inequality follows from Equation (20).
To verify the second inequality we observe

$$
\left| \langle \mathcal{L}_Q \phi | \mathcal{L}_a^{(2)} \phi \rangle - \langle \mathcal{L}_a^{(2)} \phi | \mathcal{L}_Q \phi \rangle \right| \leq c_1 \left| \langle \mathcal{L}_Q \phi | \mathcal{L}_{f,a} \phi \rangle - \langle \mathcal{L}_{f,a} \phi | \mathcal{L}_Q \phi \rangle \right| + c_1 \left| \langle \mathcal{L}_Q \phi | \mathcal{L}_{f,a} \phi \rangle - \langle \mathcal{L}_{f,a} \phi | \mathcal{L}_Q \phi \rangle \right|
$$

where we used, that $$\left[ H_{\alpha}^{(2)}, (H_{\alpha}^{(2)})^T \right] = 0$$. Let $$K_i \in \{G_j, H_j\}$$ and $$\eta, \eta' \in \{0^{1/2}, (1+\varphi)^{1/2}\}$$. We remark, that

$$
\begin{align*}
[\Phi_a(\eta K_1) \Phi_a(\eta' K_2), \mathcal{L}_{f,a}] &= i \Phi_a(i(1+\alpha)\eta K_1) \Phi_a(\eta' K_2) \\
&\quad + i \Phi_a(i(1+\alpha)\eta K_1) \Phi_a(i(1+\alpha)\eta' K_2) \\
[\Phi_a(\eta F), \mathcal{L}_{f,a}] &= i \Phi_a(i(1+\alpha)\eta F).
\end{align*}
$$

Hence, for $$\phi \in \text{dom}(\mathcal{L}_a^{(2)})$$, we have by means of (10) that

$$
\left| \langle \phi | [\Phi_a(\eta K_1) \Phi_a(\eta' K_2), \mathcal{L}_{f,a}] \phi \rangle \right| \leq \text{const} \|\mathcal{L}_{f,a} \phi\|^2
$$

for

$$
\left| \langle \phi | [\Phi_a(\eta F), \mathcal{L}_{f,a}] \phi \rangle \right| \leq \text{const} \|\mathcal{L}_{f,a} \phi\| \|\mathcal{L}_{\alpha}^{(1)} \phi\|^{1/2}.
$$

Thus, (24) is bounded by a constant times $$\|\mathcal{L}_{f,a}^{(1)} \phi\|^2$$. The essential self-adjointness of $$\mathcal{L}_Q$$ follows now from estimates analog to (23) and (24), where $$\mathcal{L}_{f,a}$$ is replaced by $$\mathcal{L}_f$$ in (23) and in the left side of (24). For $$\mathcal{L}_Q + \mathcal{L}_Q^T$$ one has to consider the commutator with $$\mathcal{L}_a^{(3)}$$ and $$\mathcal{L}_a^{(4)}$$, respectively.

**Remark 4.3.** In the same way one can show, that $$H$$ is essentially self-adjoint on any core of $$H_1 := H_{\alpha} + d\Gamma(1+\alpha)$$, even if $$H$$ is not bounded from below.

5 Regularized Interaction and Standard Form of $$\mathcal{M}_\beta$$

In this subsection a regularized interaction $$Q_N$$ is introduced:

$$
Q_N := \left\{ \Phi_\beta(\tilde{G}_N) \Phi_\beta(\tilde{H}_N) + \text{h.c.} \right\} + \Phi_\beta(F_N).
$$

The regularized form factors $$\tilde{G}_N, \tilde{H}_N, F_N$$ are obtained by multiplying the finite rank projection $$P_N := 1[\alpha \leq N]$$ from the left and the right. Moreover, an additional ultraviolet cut-off $$1[\alpha \leq N]$$, considered as a spectral projection, is added. The regularized form factors are

$$
\begin{align*}
\tilde{G}_N(k) := 1[\alpha \leq N] P_N \tilde{G}(k) P_N, \\
\tilde{H}_N(k) := 1[\alpha \leq N] P_N \tilde{H}(k) P_N, \\
F_N(k) := 1[\alpha \leq N] P_N F(k) P_N.
\end{align*}
$$
Lemma 5.1. i) $Q_N$ is essentially self-adjoint on $\mathcal{D} \subset \text{dom}(Q_N)$. $Q_N$ is affiliated with $\mathcal{M}_\beta$, i.e., $Q_N$ is closed and

$$X'Q_N \subset Q_N X', \quad \forall X' \in \mathcal{M}_\beta',$$

ii) $\mathcal{L}_0 + Q_N, \mathcal{L}_0 - \mathcal{J}Q_N\mathcal{J}$ and $\mathcal{L}_0 + Q_N - \mathcal{J}Q_N\mathcal{J}$ converges in the strong resolvent sense to $\mathcal{L}_0 + Q, \mathcal{L}_0 - \mathcal{J}Q\mathcal{J}$ and $\mathcal{L}_0 + Q - \mathcal{J}Q\mathcal{J}$, respectively.

Proof. Let $Q_N$ be defined on $\mathcal{D}$. With the same arguments as in the proof of Theorem 4.2 we obtain

$$\|Q_N\phi\| \leq C\|L_{f,a}\phi\|, \quad |\langle Q_N\phi, L_{f,a}\phi \rangle - \langle L_{f,a}\phi, Q_N\phi \rangle| \leq C\|(L_{f,a})^{1/2}\phi\|^2,$$

for $\phi \in \mathcal{D}$ and some constant $C > 0$, where we have used that $\|F_N\|_w < \infty$.

Thus, from Theorem 4.2 and Nelson’s commutator theorem we obtain that $\mathcal{D}$ is a common core for $Q_N, \mathcal{L}_0 + Q_N, \mathcal{L}_0 - Q_N\mathcal{J}$, and for the operators in line (21). A straightforward calculation yields

$$\lim_{N \to \infty} Q_N\phi = Q\phi, \quad \lim_{N \to \infty} \mathcal{J}Q_N\mathcal{J}\phi = \mathcal{J}Q\mathcal{J}\phi \quad \forall \phi \in \mathcal{D}.$$

Thus statement ii) follows, since it suffices to check strong convergence on the common core $\mathcal{D}$, see [16, Theorem VIII.25 a)].

Let $N_f := d\Gamma(1) \otimes 1 + 1 \otimes d\Gamma(1)$ be the number-operator. Since $\text{dom}(N_f) \supset \mathcal{D}$ and $W_\beta(f)^\mathcal{J} : \text{dom}(N_f) \to \text{dom}(N_f)$, see [4], we obtain

$$Q_N(A \otimes 1 \otimes W_\beta(f))^{\mathcal{J}}\phi \equiv (A \otimes 1 \otimes W_\beta(f))^{\mathcal{J}}Q_N\phi \quad (26)$$

for $A \in \mathcal{B}(H_{\mathcal{D}}), f \in \mathcal{F}$ and $\phi \in \mathcal{D}$. By closedness of $Q_N$ and density arguments the equality holds for $\phi \in \text{dom}(Q_N)$ and $X \in \mathcal{M}_\beta$ instead of $A \otimes 1 \otimes W_\beta(f)$. Thus $Q_N$ is affiliated with $\mathcal{M}_\beta$ and therefore $e^{itQ_N} \in \mathcal{M}_\beta$ for $t \in \mathbb{R}$. \hfill $\square$

Lemma 5.2. We have for $X \in \mathcal{M}_\beta$ and $t \in \mathbb{R}$

$$\tau_t(X) = e^{it(L_0 + Q)}Xe^{it(L_0 + Q)}, \quad \tau_t^0(X) = e^{it(L_0 - Q^\mathcal{J})}Xe^{it(L_0 - Q^\mathcal{J})} \quad (27)$$

Moreover, $\tau_t(X) \in \mathcal{M}_\beta$ for all $X \in \mathcal{M}_\beta$ and $t \in \mathbb{R}$, such as

$$E_Q(t) := e^{it(L_0 + Q)}e^{-itL_0} = e^{itL_0}e^{-it(L_0 - Q^\mathcal{J})} \in \mathcal{M}_\beta.$$

Proof. First, we prove the statement for $Q_N$, since $Q_N$ is affiliated with $\mathcal{M}_\beta$ and therefore $e^{itQ_N} \in \mathcal{M}_\beta$. We set

$$\hat{\tau}_t^{Q_N}(X) = e^{it(L_0 + Q_N)}Xe^{-it(L_0 + Q_N)}, \quad \hat{\tau}_t(X) = e^{it(L_0 + Q)}Xe^{-it(L_0 + Q)} \quad (28)$$
On account of Lemma 5.1 and Theorem 4.2 we can apply the Trotter product formula to obtain
\[
\hat{\tau}_t^N (X) = \lim_{n \to \infty} (e^{it\hat{Q} N} e^{it\delta N})^n X (e^{-it\hat{Q} N} e^{-it\delta N})^n.
\]

Since \( e^{it\delta N} X \in \mathfrak{M}_\beta \) and since \( \tau^0 \) leaves \( \mathfrak{M}_\beta \) invariant, \( \hat{\tau}_t^N (X) \) is the weak limit of elements of \( \mathfrak{M}_\beta \), and hence \( \hat{\tau}_t^N (X) \in \mathfrak{M}_\beta \). Moreover,
\[
\hat{\tau}_t (X) = \lim_{N \to \infty} \hat{\tau}_t^N (X) \in \mathfrak{M}_\beta.
\]

For \( E_N(t) := e^{it(L_0 + Q_N)} e^{-itL_0} \in \mathcal{B}(\mathcal{K}) \) we obtain
\[
e^{it(L_0 + Q_N)} e^{-itL_0} = \lim_{n \to \infty} (e^{itL_0} e^{it\delta N})^n e^{-itL_0} = \lim_{n \to \infty} \tau_n^0 (e^{it\delta N}) \tau_n^0 (e^{it\delta N}) \cdots \tau_n^0 (e^{it\delta N}) \in \mathfrak{M}_\beta.
\]

By virtue of Lemma 5.1 we get \( E_Q(t) := e^{it(L_0 + Q)} e^{-itL_0} \in \mathfrak{M}_\beta \). Since \( \mathcal{J} \) leaves \( \mathcal{D} \) invariant and thanks to Lemma 5.1, we deduce, that \( \mathcal{D} \) is a core of \( \mathcal{J} Q \mathcal{J} \). Moreover, we have \( e^{-itQ^N_\beta} = \mathcal{J} e^{-itQ_N} \mathcal{J} \in \mathfrak{M}_\beta \). Since we have shown, that \( \hat{\tau}_t \) leaves \( \mathfrak{M}_\beta \) invariant, we get
\[
\hat{\tau}_t^N (X) = \lim_{n \to \infty} (e^{it(L_0 + Q_N)} e^{i\hat{Q} N})^n X (e^{-it(L_0 + Q_N)} e^{-i\hat{Q} N})^n = \lim_{n \to \infty} \hat{\tau}_t^N (X).
\]

Thanks to Lemma 5.1 we also have
\[
\tau_t (X) = \lim_{n \to \infty} \tau_t^N (X) = \lim_{N \to \infty} \hat{\tau}_t^N (X) = \hat{\tau}_t (X). \tag{29}
\]

The proof of \( \hat{\tau}_t^N (X) = e^{it(L_0-Q^\beta)} X e^{it(L_0-Q^\beta)} \) follows analogously. Using the Trotter product formula we obtain
\[
e^{it(L_0 + Q_N)} e^{-itL_0} = \lim_{n \to \infty} (e^{itL_0} e^{i\hat{Q} N})^n e^{-itL_0} = \lim_{n \to \infty} \tau_n^0 (e^{i\hat{Q} N}) \tau_n^0 (e^{i\hat{Q} N}) \cdots \tau_n^0 (e^{i\hat{Q} N}) = \lim_{n \to \infty} (e^{i\hat{Q} N} e^{-it(L_0-Q^\beta)}) e^{-it(L_0-Q^\beta)}.
\]

By strong resolvent convergence we may deduce \( E(t) = e^{itL_0} e^{-it(L_0-Q^\beta)} \).

Let \( \mathcal{C} \) be the natural positive cone associated with \( \mathcal{J} \) and \( \Omega^\beta_0 \) and let \( \mathfrak{M}_\beta^\text{ana} \) be the \( \tau \)-analytic elements of \( \mathfrak{M}_\beta \), (see [4]).
THEOREM 5.3. Assume Hypothesis 1 and $\Omega_0^β \in \text{dom}(e^{-β/2(L_0+Q)})$. Let $\Omega^β := e^{-β/2(L_0+Q)} \Omega_0^β$. Then

$$
\mathcal{J} \Omega^β = \Omega^β, \quad \Omega^β = e^{β/2(L_0-Q^β)} \Omega_0^β,
$$

(30)

$\mathcal{L}_Q \Omega^β = 0, \quad \mathcal{J} X^* \Omega^β = e^{-β/2 L Q} X \Omega^β, \quad \forall X \in \mathfrak{M}_β$

Furthermore, $\Omega^β$ is separating and cyclic for $\mathfrak{M}_β$, and $\Omega^β \in C$. The state $ω^β$ is defined by

$$
ω^β(X) := \|Ω^β\|^2 \langle Ω^β | X Ω^β \rangle, \quad X \in \mathfrak{M}_β
$$

is a $(τ, β)$-KMS state on $\mathfrak{M}_β$.

Proof. First, we define $Ω(z) = e^{-z(L_0+Q)} \Omega_0^β$ for $z ∈ \mathbb{C}$ with $0 ≤ \text{Re} z ≤ β/2$. Since $Ω^β_0 \in \text{dom}(e^{-β/2(L_0+Q)})$, $Ω(z)$ is analytic on $S_{β/2} := \{z ∈ \mathbb{C} : 0 < \text{Re} (z) < α\}$ and continuous on the closure of $S_{β/2}$, see Lemma A.2 below.

Proof of $\mathcal{J} Ω(β/2) = Ω(β/2)$:

We pick $φ ∈ \bigcup_{n∈\mathbb{N}} \text{ran} 1[L_0 - Q^β] ≤ n$. Let $f(z) := \langle φ | J Ω(τ) \rangle$ and $g(z) := \langle e^{(β/2-τ) L_0} φ | e^{-z(L_0+Q)} \Omega_0^β \rangle$. Both $f$ and $g$ are analytic on $S_{β/2}$ and continuous on its closure. Thanks to Lemma 5.2 we have $E_Q(t) ∈ \mathfrak{M}_β$, and hence

$$
f(it) = \langle φ | J E_Q(t) \Omega_0^β \rangle = \langle φ | e^{-β/2 L_0} E_Q(t)^* Ω_0^β \rangle = g(it), \quad t ∈ \mathbb{R}.
$$

By Lemma A.1, $f$ and $g$ are equal, in particular in $z = β/2$. Note that $φ$ is any element of a dense subspace.

Proof of $Ω^β_0 \in \text{dom}(e^{β/2(L_0-Q^β)})$ and $Ω(β/2) = e^{β/2(L_0-Q^β)} ω_0^β$.

Let $φ ∈ \bigcup_{n∈\mathbb{N}} \text{ran} 1[L_0 - Q^β] ≤ n$. We set $g(z) := \langle e^{-z L_0} φ | e^{-z L_0} Ω_0^β \rangle$. Since $E_Q(t)^β = e^{it(L_0-Q^β)} e^{-it L_0}$, $g$ coincides for $z = it$ with $f(z) := \langle φ | J Ω(τ) \rangle$. Hence they are equal in $z = β/2$. The rest follows since $e^{-β/2(L_0-Q^β)}$ is self-adjoint.

Proof of $L_Q Ω(β/2) = 0$:

Choose $φ ∈ \bigcup_{n∈\mathbb{N}} \text{ran} 1[L_0] ≤ n$. We define $g(z) := \langle e^{-z L_0} φ | e^{-z L_0} Ω_0^β \rangle$ and $f(z) := \langle φ | Ω(z) \rangle$ for $z$ in the closure of $S_{β/2}$. Again both functions are equal on the line $z = it, \quad t ∈ \mathbb{R}$. Hence $f$ and $g$ are identical, and therefore $Ω(β/2) \in \text{dom}(e^{-β/2 L Q})$ and $e^{-β/2 L Q} Ω(β/2) = Ω(β/2)$. We conclude that $L_Q Ω(β/2) = 0$.

Proof of $\mathcal{J} X^* Ω(β/2) = e^{-β/2 L Q} X Ω(β/2), \quad ∀ X \in \mathfrak{M}_β$:

Fore $A ∈ \mathfrak{M}_β$, we have that

$$
\mathcal{J} A^* Ω(-it) = \mathcal{J} A^* E_Q(t) Ω_0^β = e^{-β/2 L_0} E_Q(t)^* A Ω_0^β
$$

$$
= e^{-(β/2-it) L_0} e^{-it(L_0+Q)} A Ω_0^β
$$

$$
= e^{-(β/2-it) L_0} τ_{-it}(A) e^{-it(L_0+Q)} Ω_0^β.
$$

DOCUMENTA MATHEMATICA 16 (2011) 177–208
Let \( \phi \in \bigcup_{n \in \mathbb{N}} \text{ran} \mathbf{1}[[\mathcal{L}_0^n]] \). We define \( f(z) = \langle \phi, \mathcal{J} A^* \Omega(\tau) \rangle \) and \( g(z) = \langle e^{-\beta z/2} - \mathcal{L}_0 \phi, \tau_2(A) \Omega(z) \rangle \). Since \( f \) and \( g \) are analytic and equal for \( z = it \), we have \( \mathcal{J} A^* \Omega(\beta/2) = \tau_{1\beta/2}(A) \Omega(\beta/2) \). To finish the proof we pick \( \phi \in \bigcup_{n \in \mathbb{N}} \text{ran} \mathbf{1}[[\mathcal{L}_Q^n]] \leq n \), and set \( f(z) := \langle \phi, \tau_{1\beta}(A) \Omega(\beta/2) \rangle \) and \( g(z) := \langle e^{-\beta z/2} \phi, A\Omega(\beta/2) \rangle \). For \( z = it \) we see

\[
\text{Hence } A \Omega(\beta/2) \in \text{dom}(e^{-\beta z/2} \mathcal{L}_0) \text{ and } \mathcal{J} A^* \Omega(\beta/2) = e^{-\beta z/2} \mathcal{L}_0 A\Omega(\beta/2). \]

Since \( \mathfrak{M}_\beta^{\text{ana}} \) is dense in the strong topology, the equality holds for all \( X \in \mathfrak{M}_\beta^\beta \).

\[ \mathbf{\text{Proof, that } \Omega^\beta \text{ is separating for } \mathfrak{M}_\beta^\beta} \]

Let \( A \in \mathfrak{M}_\beta^\beta \). We choose \( \phi \in \bigcup_{n \in \mathbb{N}} \text{ran} \mathbf{1}[[\mathcal{L}_Q + Q]] \leq n \). First, we have

\[ \mathcal{J} A^* \Omega(\beta/2) = \tau_{1\beta/2}(A) \Omega(\beta/2). \]

Let \( f_\phi(z) = \langle \phi, \tau_2(A) \Omega(\beta/2) \rangle \) and \( g_\phi(z) = \langle e^{-\tau_2(L_0 + Q)} \phi, A e^{-\beta z/2} \tau_2(L_0 + Q) \Omega_0^\beta \rangle \) for \( -\beta/2 \leq \text{Re } z \leq 0 \). Both functions are continuous and analytic if \( -\beta/2 < \text{Re } z < 0 \). Furthermore, \( f_\phi(it) = g_\phi(it) \) for \( t \in \mathbb{R} \). Hence \( f_\phi = g_\phi \) and for \( z = -\beta/2 \)

\[ \langle \phi, \mathcal{J} A^* \Omega(\beta/2) \rangle = \langle e^{-\beta/2(L_0 + Q)} \phi, A\Omega_0^\beta \rangle. \]

This equation extends to all \( A \in \mathfrak{M}_\beta \), we obtain \( A \Omega_0^\beta \in \text{dom}(e^{-\beta/2(L_0 + Q)}) \), such as \( e^{-\beta z/2(L_0 + Q)} A \Omega_0^\beta = \mathcal{J} A^* \Omega(\beta/2) \) for \( A \in \mathfrak{M}_\beta \). Assume \( A^* \Omega(\beta/2) = 0 \), then \( e^{-\beta/2(L_0 + Q)} A\Omega_0^\beta = 0 \) and hence \( A\Omega_0^\beta = 0 \). Since \( \Omega_0^\beta \) is separating, it follows that \( A = 0 \) and therefore \( A^* = 0 \).

\[ \mathbf{\text{Proof of } \Omega^\beta \in \mathcal{C}, \text{ and that } \Omega^\beta \text{ is cyclic for } \mathfrak{M}_\beta^\beta} \]

To prove that \( \phi \in \mathcal{C} \) it is sufficient to check that \( \langle \phi, A \mathcal{J} A \Omega_0^\beta \rangle \geq 0 \) for all \( A \in \mathfrak{M}_\beta \). We have

\[ \langle \Omega(\beta/2), A \mathcal{J} A \Omega_0^\beta \rangle = \langle \mathcal{J} A^* \Omega(\beta/2), A\Omega_0^\beta \rangle \]

\[ = \langle e^{-\beta/2(L_0 + Q)} A\Omega_0^\beta, A\Omega_0^\beta \rangle \geq 0. \]

The proof follows, since every separating element of \( \mathcal{C} \) is cyclic.

\[ \mathbf{\text{Proof, that } \omega^\beta \text{ is a } (\tau, \beta)\text{-KMS state}} \]

For \( A, B \in \mathfrak{M}_\beta \) and \( z \in S_\beta \) we define

\[ F_\beta(A, B, z) = c \langle e^{-\tau/2 \mathcal{L}_0} A^* \Omega^\beta | e^{iz/2 \mathcal{L}_0} B \Omega^\beta \rangle, \]

where \( c := \|\Omega^\beta\|^{-2} \). First, we observe

\[ F_\beta(A, B, t) = c \langle e^{-it/2 \mathcal{L}_0} A^* \Omega^\beta | e^{it/2 \mathcal{L}_0} B \Omega^\beta \rangle = c \langle \Omega^\beta | A\tau_t(B) \Omega^\beta \rangle \]

\[ = \omega^\beta(A \tau_t(B)) \]

\[ \text{Documenta Mathematica 16 (2011) 177–208} \]
and

\[
\omega^\beta(\tau_1(B)A) = c(\langle \tau_1(B^* )\Omega^\beta | A\Omega^\beta \rangle ) = c(\langle J A\Omega^\beta | J \tau_1(B^*)\Omega^\beta \rangle ) \\
= c(\langle e^{-\beta/2}\mathcal{L}_0 A^* \Omega^\beta | e^{-\beta/2}\mathcal{L}_0 \tau_1(B)\Omega^\beta \rangle ) \\
= c(\langle e^{-\beta((\beta+1)/2)}\mathcal{L}_0 A^* \Omega^\beta | e^{(i\beta+1)/2}\mathcal{L}_0 B\Omega^\beta \rangle ) \\
= F_\beta(A, B, t + i\beta).
\]

The requirements on the analyticity of \( F_\beta(A, B, \cdot) \) follow from Lemma A.2.

---

6 Proof of Theorem 1.3

For \( \underline{x}_n := (s_n, \ldots, s_1) \in \mathbb{R}^n \) we define

\[
Q_N(\underline{x}_n) := Q_N(s_n) \cdots Q_N(s_1), \quad Q_N(s) := e^{-s\mathcal{L}_0}Q_N e^s\mathcal{L}_0, \quad s \in \mathbb{R}
\]

(31)

At this point, we check that \( Q_N(\underline{x}_n)\Omega^\beta_0 \) is well defined, and that it is an analytic vector of \( \mathcal{L}_0 \), see Equation (25). The goal of Theorem 1.3 is to give explicit conditions on \( H_{el} \) and \( W \), which ensure \( \Omega^\beta_0 \in \text{dom}(e^{-\beta/2}(\mathcal{L}_0 + Q)) \). Let

\[
\eta_1 := \int (\| \tilde{G}(k) \|_{B(\mathcal{H}_{el})}^2 + \| \tilde{H}(k) \|_{B(\mathcal{H}_{el})}^2)(2 + 4\alpha(k)^{-1}) \, dk
\]

(32)

\[
\eta_2 := \int (\| F(k) H_{el}^\gamma \|_{B(\mathcal{H}_{el})}^2 + \| F(k)^* H_{el}^{-\gamma} \|_{B(\mathcal{H}_{el})})(2 + 4\alpha(k)^{-1}) \, dk
\]

The idea of the proof is the following. First, we expand \( e^{-\beta/2}(\mathcal{L}_0 + Q_N) e^{\mathcal{L}_0} \) in a Dyson-series, i.e.,

\[
e^{-\beta/2}(\mathcal{L}_0 + Q_N) e^{\mathcal{L}_0} = 1 + \sum_{n=1}^{\infty}(-1)^n \int_{\Delta_n^2} e^{-s_n\mathcal{L}_0}Q_N e^{s_n} e^{\mathcal{L}_0} d\underline{x}_n.
\]

(33)

Under the assumptions of Theorem 1.3 we obtain an upper bound, uniform in \( N \), for

\[
\langle \Omega^\beta_0 | e^{-\beta(\mathcal{L}_0 + Q_N)}\Omega^\beta_0 \rangle
\]

(34)

\[
= 1 + \sum_{n=1}^{\infty}(-1)^n \int_{\Delta_n^2} \langle \Omega^\beta_0 | e^{-s_n\mathcal{L}_0}Q_N e^{s_n} e^{\mathcal{L}_0} \rangle \Omega^\beta_0 d\underline{x}_n.
\]

This is proven in Lemma 6.4 below, which is the most important part of this section. In Lemma 6.1 and Lemma 6.2 we deduce from the upper bound for (34) an upper bound for \( \| e^{-\beta/2}(\mathcal{L}_0 + Q_N) \| \), which is uniform in \( N \). The proof of Theorem 1.3 follows now from Lemma 6.3, where we show that \( \Omega^\beta_0 \in \text{dom}(e^{-\beta/2}(\mathcal{L}_0 + Q)) \).
LEMMA 6.1. Assume

\[ \limsup_{n \to \infty} \sup_{0 \leq x \leq \beta/2} \left\| \int_{\Delta_n^2} Q_N(C_n) \Omega_0^\beta \, d\xi_n \right\|^{1/n} < 1. \]

for all \( N \in \mathbb{N} \). Then \( \Omega_0^\beta \in \text{dom}(e^{-x(C_0 + Q_N)}) \), \( 0 < x \leq \beta/2 \) and

\[
e^{-x(C_0 + Q_N)} \Omega_0^\beta = \Omega_0^\beta + \sum_{n=1}^{\infty} (-1)^n \int_{\Delta_n^2} Q_N(C_n) \Omega_0^\beta \, d\xi_n. \tag{35}\]

In this context \( \Delta_n^2 = \{(s_1, \ldots, s_n) \in \mathbb{R}^n : 0 \leq s_n \leq \ldots \leq s_1 \leq x\} \) is a simplex of dimension \( n \) and sidelength \( x \).

**Proof.** Let \( \phi \in \text{ran} 1||C_0 + Q_N|| \leq k \) and \( 0 \leq x \leq \beta/2 \) be fixed. An \( m \)-fold application of the fundamental theorem of calculus yields

\[
\langle e^{-x(C_0 + Q_N)} \phi | e^{x(C_0 + Q_N)} \Omega_0^\beta \rangle = \langle \phi | \Omega_0^\beta + \sum_{n=1}^{m} (-1)^n \int_{\Delta_n^2} Q_N(C_n) \Omega_0^\beta \, d\xi_n \rangle \\
+ (-1)^{m+1} \int_{\Delta_{m+1}^n} \langle e^{-s_{m+1}(C_0 + Q_N)} \phi | e^{s_{m+1}C_0} Q_N(C_{m+1}) \Omega_0^\beta \rangle \, d\gamma_{m+1}. \tag{36}\]

Since \( C_0 \Omega_0^\beta = 0 \) we have for \( r(C_{m+1}) := (s_m - s_{m+1}, \ldots, s_1 - s_{m+1}) \) that

\[
e^{s_{m+1}C_0} Q_N(C_{m+1}) \Omega_0^\beta = Q_N(r(C_{m+1})) \Omega_0^\beta.
\]

We turn now to the second expression on the right side of Equation (36), after a linear transformation depending on \( s_{m+1} \) we get

\[
(-1)^{m+1} \int_0^x \langle e^{-s_{m+1}(C_0 + Q_N)} \phi | Q_N \int_{\Delta_{r-s_{m+1}}^m} Q_N(C_m) \Omega_0^\beta \, d\xi_m \rangle \, ds_{m+1}.
\]

Since \( ||e^{-s_{m+1}(C_0 + Q_N)} \phi|| \leq e^\beta/2 ||\phi|| \), and using that \( Q_N(C_m) \Omega_0^\beta \) is a state with at most \( 2m \) bosons, we obtain the upper bound

\[
\text{const} \, ||\phi|| \sqrt{(2m)(2m+1)} \sup_{0 \leq x \leq \beta/2} \left\| \int_{\Delta_{r-s_{m+1}}^m} Q_N(C_m) \Omega_0^\beta \, d\xi_m \right\|.
\]

Hence, for \( m \to \infty \) we get

\[
\langle e^{-x(C_0 + Q_N)} \phi | \Omega_0^\beta \rangle = \langle \phi | \Omega_0^\beta + \sum_{n=1}^{\infty} (-1)^n \int_{\Delta_n^2} Q_N(C_n) \Omega_0^\beta \, d\xi_n \rangle.
\]

Since \( \bigcup_{k=1}^{\infty} \text{ran} 1||C_0 + Q_N|| \leq k \) is a core of \( e^{-x(C_0 + Q_N)} \), the proof follows from the self-adjointness of \( e^{-x(C_0 + Q_N)} \). \( \square \)
LEMMA 6.2. Let $0 < x \leq \beta/2$. We have the identity

$$
\int_{\Delta^m_{x/2}} \int_{\Delta^m_{y/2}} \langle Q_N(z_m) \Omega_0^\beta | Q_N(z_n) \Omega_0^\beta \rangle \, dz_m \, dz_n
= \int_{\Delta^m_{x/2}} \mathbf{1}[z_m \geq \beta - x \geq z_{m+1}] \langle \Omega_0^\beta | Q_N(z_{n+m}) \Omega_0^\beta \rangle \, dz_{n+m}.
$$

For $m = n$ it follows

$$
\bigg\| \int_{\Delta^m_{x/2}} Q_N(z_n) \Omega_0^\beta \, dz_n \bigg\|^2 \leq \int_{\Delta^m_{x/2}} \bigg| \langle \Omega_0^\beta | Q_N(z_{2n}) \Omega_0^\beta \rangle \bigg| \, dz_{2n}.
$$

Proof. Recall Theorem 3.1 and Lemma 5.1. Since $\mathcal{J}$ is a conjugation we have $\langle \phi | \psi \rangle = \langle \mathcal{J} \psi | \mathcal{J} \phi \rangle$, and for every operator $X$, that is affiliated with $\mathcal{M}_\beta$, we have $\mathcal{J} X \Omega_0^\beta = e^{-\beta/2 \mathcal{L}_0} X^* \Omega_0^\beta$. Thus,

$$
\int_{\Delta^m_{x/2}} \int_{\Delta^m_{y/2}} \langle Q_N(z_m) \Omega_0^\beta | Q_N(z_n) \Omega_0^\beta \rangle \, dz_m \, dz_n
= \int_{\Delta^m_{x/2}} \int_{\Delta^m_{y/2}} \langle e^{-\beta/2 \mathcal{L}_0} Q_N(z_m)^* \Omega_0^\beta | e^{-\beta/2 \mathcal{L}_0} Q_N(z_n)^* \Omega_0^\beta \rangle \, dz_m \, dz_n
$$

Since $\mathcal{L}_0 \Omega_0^\beta = 0$ we have

$$
e^{-\beta \mathcal{L}_0} Q_N(z_m)^* \Omega_0^\beta = Q_N(\beta - r_1) \cdots Q_N(\beta - r_m) \Omega_0^\beta.
$$

Next, we introduce new variables for $z$ namely $y_i := \beta - r_{m-i+1}$. Let $D_{x/2} := \{y_m \in \mathbb{R}^m : \beta - x \leq y_m \leq \ldots \leq y_1 \leq \beta\}$. Thus the right side of Equation (39) equals

$$
\int_{\Delta^m_{x/2}} \int_{D_{x/2}} \langle \Omega_0^\beta | Q_N(z_n) Q_N(y_m)^* \Omega_0^\beta \rangle \, dz_n \, dy_m
= \int_{\Delta^m_{x/2}} \mathbf{1}[z_m \geq \beta - x \geq z_{m+1}] \langle \Omega_0^\beta | Q_N(z_{n+m}) \Omega_0^\beta \rangle \, dz_{n+m}.
$$

The second statement of the Lemma follows by choosing $n = m$. \hfill $\square$

LEMMA 6.3. Assume $\sup_{N \in \mathbb{N}} \|e^{-x(\mathcal{L}_0 + Q_N)} \Omega_0^\beta\| < \infty$ then $\Omega_0^\beta \in \text{dom}(e^{-x(\mathcal{L}_0 + Q)})$ and

$$
\|e^{-x(\mathcal{L}_0 + Q)} \Omega_0^\beta\| \leq \sup_{N \in \mathbb{N}} \|e^{-x(\mathcal{L}_0 + Q_N)} \Omega_0^\beta\|
$$

Proof. For $f \in \mathcal{C}_b^\infty(\mathbb{R})$ and $\phi \in \mathcal{K}$ we define $\psi_N := f(\mathcal{L}_0 + Q_N) \phi$. Obviously, for $g(r) = e^{-x r} f(r) \in \mathcal{C}_b^\infty(\mathbb{R})$ we have $e^{-x(\mathcal{L}_0 + Q_N)} \psi_N = g(\mathcal{L}_0 + Q_N) \phi$. 

\begin{document}
\bibliographystyle{amsplain}
\bibliography{references}
\end{document}
Since $\mathcal{L}_0 + Q_N$ tends to $\mathcal{L}_0 + Q$ in the strong resolvent sense as $N \to \infty$, we know from [16] that $\lim_{N \to \infty} \psi_N = f(\mathcal{L}_0 + Q) =: \psi$ and

$$
\lim_{N \to \infty} e^{-x(\mathcal{L}_0 + Q_N)} \psi_N = \lim_{N \to \infty} g(\mathcal{L}_0 + Q_N) \phi = \psi.$$

Thus,

$$
|\langle e^{-x(\mathcal{L}_0 + Q)} \psi | \Omega^\beta_0 \rangle| = \lim_{N \to \infty} |\langle e^{-x(\mathcal{L}_0 + Q_N)} \psi_N | \Omega^\beta_0 \rangle| \\
\leq \sup_{N \in \mathbb{N}} \| e^{-x(\mathcal{L}_0 + Q)} \Omega^\beta_0 \| \| \psi \|
$$

Since $\{f(\mathcal{L}_0 + Q) \phi \in \mathcal{K} : \phi \in \mathcal{K}, f \in \mathcal{C}_0^\infty(\mathbb{R})\}$ is a core of $e^{-x(\mathcal{L}_0 + Q)}$, we obtain $\Omega^\beta_0 \in \text{dom}(e^{-x(\mathcal{L}_0 + Q)})$. \qed

**Lemma 6.4.** For some $C > 0$ we have

$$
\int_{\Delta^\beta_n} \left|\langle \Omega^\beta_0 | Q_N(\vec{s}_n) \Omega^\beta_0 \rangle \right| d\vec{s}_n \\
\leq \text{const} (n+1)^2 (1+\beta)^n \left( 8\eta_1 + \frac{(8C\eta_2)^{1/2}}{(n+1)(1-2\beta)/2} \right)^n,
$$

where $\eta_1$ and $\eta_2$ are defined in (32).

**Proof of 6.4.** First recall the definition of $Q_N$ and $Q_N(\vec{s}_n)$ in Equation (25) and Equation (31), respectively. Let

$$
\int_{\Delta^\beta_n} \left|\langle \Omega^\beta_0 | Q_N(\vec{s}_n) \Omega^\beta_0 \rangle \right| d\vec{s}_n =: \int_{\Delta^\beta_n} J_n(\beta, \vec{s}) d\vec{s}_n.
$$

The functions $J_n(\beta, \vec{s})$ clearly depends on $N$, but since we want to find an upper bound independent of $N$, we drop this index. Let $W_1 := \Phi(\vec{G}) \Phi(\vec{H}) + \text{h.c.}$, $W_2 := \Phi(F)$ and $W := W_1 + W_2$. By definition of $\omega^\beta_0$ in (3.1), see also (13), we obtain

$$
J_n(\beta, \vec{s}_n) = \omega^\beta_0 \left( e^{-\beta \kappa H_0} W e^{\beta \kappa H_0} W e^{-\beta H_0} \right) \cdots \left( e^{-\beta \kappa H_0} W e^{\beta H_0} \right)
$$

where $\eta_1 := \text{Tr}_H(\{ e^{-\beta \kappa H_0} W e^{\beta \kappa H_0} \}) \cdots \left( e^{-\beta H_0} \right)

By definition of $\omega^\beta_0$ it suffices to consider expressions with an even number of field operators. In the next step we sum over all expression, where $n_1$ times $W_1$ occurs and $2n_2$ times $W_2$. The sum of $n_1$ and $n_2$ is denoted by $m$. For fixed $n_1$ and $n_2$ the remaining expressions are all expectations in $\omega^\beta_0$ of $2m$ field

\cite{M.Koenenberg} Documenta Mathematica 16 (2011) 177–208
operators. In this case the expectations in $\omega_1^2$ can be expressed by an integral over $\mathbb{R}^{2m} \times \{-1, 1\}^{2m}$ with respect to $\nu$, which is defined in Lemma A.4 below.

To give a precise formula we define

$$M(m_1, m_2) = \{ \kappa \in \{1, 2\}^n : \# \kappa^{-1}\{i\} = m_i, \quad i = 1, 2 \}. $$

Thus we obtain

$$J_n (\beta, \mathcal{Z}_m) = (\mathcal{Z})^{-1} \sum_{(n_1, n_2) \in \mathbb{N}_0^2 m_0 = n} \sum_{\kappa \subseteq M(n_1, n_2)} \int \nu(d\mathcal{Z}_{2m} \otimes d\mathcal{Z}_{2m})$$

$$\text{Tr}_{\mathcal{H}_{el}} \left\{ e^{-\beta (s_1 - s_{2m})} H_{el} I_{2m} e^{-\beta (s_{2m-1} - s_{2m})} H_{el} \cdots e^{-\beta (s_1 - s_{2})} H_{el} I_1 \right\},$$

Of course $I_j$ depends on $\mathcal{Z}_{2m} \times \mathcal{Z}_{2m}$ namely for $\kappa(j) = 1, 2$ we have

$$I_j = \begin{cases} I_j(m, \tau, m', \tau'), & \kappa(j) = 1 \\ I_j(m, \tau), & \kappa(j) = 2, \end{cases}$$

where $(m, \tau, m', \tau') \in \{(k_j, \tau_j) : j = 1, \ldots, m \}$. For $\kappa(j) = 1$ we have that

$$I_j(m, +, m', -) = \tilde{G}(m) \tilde{H}(m') + \tilde{H}^*(m) \tilde{G}(m')$$

$$I_j(m, -, m', +) = \tilde{G}(m) \tilde{H}^*(m') + \tilde{H}(m) \tilde{G}^*(m')$$

$$I_j(m, +, m', +) = \tilde{G}^*(m) \tilde{H}^*(m') + \tilde{H}^*(m) \tilde{G}^*(m')$$

$$I_j(m, -, m', -) = \tilde{G}(m) \tilde{H}(m') + \tilde{H}(m) \tilde{G}(m')$$

and for $\kappa(j) = 2$ we have that

$$I_j(m, +) = F^*(m)$$

$$I_j(m, -) = F(m).$$

In the integral (40) we insert for $(m, \tau)$ and $(m', \tau')$ in the definition of $I_j$ from left to right $\mathcal{Z}_{2m}, \mathcal{Z}_{2m}, \ldots, \mathcal{Z}_{2m}.$

For fixed $(\mathcal{Z}_{2m}, \mathcal{Z}_{2m})$ the integrand of (40) is a trace of a product of $4m$ operators in $\mathcal{H}_{el}$. We will apply Hölder’s inequality for the trace, i.e.,

$$|\text{Tr}_{\mathcal{H}_{el}} \{ A_{2m} B_{2m} \cdots A_1 B_1 \} | \leq \prod_{j=1}^{2m} \| B_j \|_{\mathcal{B}(\mathcal{H}_{el})} \cdot \prod_{j=1}^{2m} \text{Tr}_{\mathcal{H}_{el}} \{ A_j^{p_j} \}^{p_j^{-1}}.$$

In our case $p_i := (s_{i-1} - s_i)^{-1}$ for $i = 2, \ldots, 2m$ and $p_1 := (1 - s_1 + s_{2m})^{-1}$ and

$$(A_j, B_j) := \begin{cases} \{ e^{-\beta p_j^{-1}} H_{el}, I_j(m, \tau, m', \tau') \}, & \kappa(j) = 1 \\ \{ e^{-\beta p_j^{-1}} H_{el}^\gamma, I_j(m, \tau) \}, & \kappa(j) = 2, \end{cases}$$
We define
\[ \eta_1(k) = \max \{ ||\tilde{G}(k)||_{l^2(H_{el})}, ||\tilde{H}(k)||_{l^2(H_{el})}^r \}, \]
\[ \eta_2(k) = \max \{ ||F(k)H_{el,+}^{-\gamma}||_{l^2(H_{el})}, ||F^*(k)H_{el,+}^{-\gamma}||_{l^2(H_{el})} \}. \]

By definition of \( B_j \) we have
\[ ||B_j||_{l^2(\mathcal{H}_{el})} \leq \begin{cases} \eta_1(m)\eta_1(m'), & \kappa(j) = 1 \\ \eta_2(m), & \kappa(j) = 2. \end{cases} \]

Furthermore,
\[ \text{Tr}_{\mathcal{H}_{el}} \{ A_i^p \}^p_j^{-1} = \text{Tr}_{\mathcal{H}_{el}} \{ e^{-\beta H_{el}} H_{el,+}^{p_j} \}^p_j^{-1} \]
\[ \leq ||e^{-\epsilon H_{el}} H_{el,+}^{p_j}||_{l^2(\mathcal{H}_{el})} \text{Tr}_{\mathcal{H}_{el}} \{ e^{-\beta H_{el}} \}^p_j^{-1}, \quad k(j) = 2. \]

Let \( E_{gs} := \inf \sigma(H_{el}). \) The spectral theorem for self-adjoint operators implies
\[ ||e^{-\epsilon H_{el}} H_{el,+}^{p_j}||_{l^2(\mathcal{H}_{el})} \leq \sup_{r \geq E_{gs}} e^{-\epsilon p_j^{-1}} r(r - E_{gs} + 1)^{\gamma} \leq e^{-\gamma} p_j e^{-\epsilon p_j^{-1}} (E_{gs} - 1). \]

Inserting this estimates we get
\[ \text{Tr}_{\mathcal{H}_{el}} \{ e^{-\beta - \beta(s_1 - s_2m)}H_{el} I_{2m} e^{-\beta (s_{2m-1} - s_{2m}) H_{el}} H_{el} \cdots e^{-\beta (s_1 - s_2) H_{el}} I_1 \} \]
\[ \leq C_\kappa(\mathcal{S}_n) \prod_{j=1}^{2m} ||B_j||_{l^2(\mathcal{H}_{el})} \]
where
\[ C_\kappa(\mathcal{S}_n) := (1 - s_1 + s_n)^{-\alpha_1} \prod_{i=1}^{n-1} (s_i - s_{i+1})^{-\alpha_i}. \]

and
\[ \alpha_i = \begin{cases} 0, & \kappa(i) = 1 \\ 1/2, & \kappa(i) = 2. \end{cases} \]

Now, we recall the definition of \( \nu. \) Roughly speaking, one picks a pair of variables \((k_i, k_j)\) and integrates over \( \delta k_i, k_j \coth(\beta/2\alpha(k_i)) \) \( dk_i dk_j. \) Subsequently one picks the next pair and so on. At the end one sums up all \( \frac{(2m)!}{m!} \) pairings and all \( 4^n \) combinations of \( \mathcal{S}_{2m}. \) Inserting Estimate (41) and that
\[ \int \eta_0(k)\eta_0(k) \coth(\beta/2\alpha(k)) \, dk \leq (1 + \beta^{-1}) \eta_0^{1/2} \eta_0^{1/2}, \]
we obtain
\[ |J_n(\beta, \mathcal{S})| \leq \frac{(1 + \beta^{-1})^n}{Z} \sum_{(n_1, n_2) \in \mathbb{N}_0^2} \sum_{n_1 + 2n_2 = n} (\eta_1)^{n_1} (C\eta_2)^{n_2} \frac{(2m)!2^n}{m!} C_\kappa(\mathcal{S}) \]
By Lemma A.3 below and since $(2m)!/(m!)^2 \leq 4^m$ we have

$$\int_{\Delta_n^\beta} \left| \langle \Omega_0^\beta | Q_N(\Omega_n) \Omega_0^\beta \rangle \right| d\Omega_n \leq \text{const}(1 + \beta)^n \sum_{(n_1, n_2) \in \mathbb{N}^2, n_1 + 2n_2 = n} \left( \frac{8n_1}{n_1 + 1} \frac{(8Cn_2)^{n_1}}{(1 - 2\gamma)^{n_2 - 2}} \right)$$

This completes the proof. \[\square\]

7 The Harmonic Oscillator

Let $L^2(X, d\mu) = L^2(\mathbb{R})$ and $H_{el} := -\Delta_q + \Theta^2 q^2$ be the one dimensional harmonic oscillator and $H_{ph} = L^2(\mathbb{R}^3)$. We define

$$H = H_{osc} + \Phi(F) + \check{\mathcal{H}}, \quad \check{\mathcal{H}} := d\Gamma(|k|), \quad (44)$$

where $\Phi(F) = q \cdot \Phi(f)$, with $\lambda(|k|^{-1/2} + |k|^{1/2}) f \in L^2(\mathbb{R}^3)$. $H_{osc}$ is the harmonic oscillator, the form-factor $F$ comes from the dipole approximation.

The Standard Liouvillean for this model is denoted by $\mathcal{L}_{osc}$. Now we prove Theorem 1.4.

**Proof.** We define the creation and annihilation operators for the electron.

$$A^* = \frac{\Theta^{1/2} q - i \Theta^{-1/2} p}{\sqrt{2}}, \quad A = \frac{\Theta^{1/2} q + i \Theta^{-1/2} p}{\sqrt{2}}, \quad p = -i \partial_x, \quad (45)$$

$$\Phi(c) = c_1 q + c_2 p, \quad \text{for } c = c_1 + ic_2 \in \mathbb{C}, \quad c_i \in \mathbb{R}. \quad (46)$$

These operators fulfill the CCR-relations and the harmonic-oscillator is the number-operator up to constants.

$$[A, A^*] = 1, \quad [A^*, A^*] = [A, A] = 0, \quad H_{osc} = \Theta A^* A + \Theta/2, \quad (47)$$

$$[H_{osc}, A] = -\Theta A, \quad [H_{osc}, A^*] = \Theta A^*. \quad (48)$$

The vector $\Omega := \left( \frac{\Theta}{2} \right)^{1/4} e^{-\Theta q^2/2}$ is called the vacuum vector. Note, that one can identify $\mathcal{F}_b[\mathbb{C}]$ with $L^2(\mathbb{R})$, since $LH\{(A^*)^n \Omega | n \in \mathbb{N}^0\}$ is dense in $L^2(\mathbb{R})$. It follows, that $\omega^{osc}_\beta$ is quasi-free, as a state over $W(\mathbb{C})$ and

$$\omega^{osc}_\beta(W(c)) = (Z)^{-1} \text{Tr}_{\mathcal{H}_{el}} \{ e^{-\beta H_{el}} W(c) \} = \exp \left( -\frac{1}{4} \coth(\beta/2) |c|^2 \right), \quad (49)$$

where $Z = \text{Tr}_{\mathcal{H}_{el}} \{ e^{-\beta H_{el}} \}$ is the partition function for $\mathcal{H}_{el}$.

First, we remark, that Equation (31) is defined for this model without regularization by $P_N := 1[H_{el} \leq N]$. Moreover we obtain from Lemma 6.2,
that
\[ \left\| \int_{\Delta_{\beta/2}^n} Q(s)\Omega_0^\beta ds_2n \right\|^2 \leq \int_{\Delta_{\beta}^n} \left| \langle \Omega_0^\beta | Q(s_2n)\Omega_0^\beta \rangle \right| ds_2n =: h_{2n}(\beta, \lambda). \] (50)

To show that \( \Omega^\beta \in \text{dom}(e^{-\beta/2(L_0 + Q)}) \) is suffices to prove, that \( \sum_{n=0}^{\infty} h_{2n}(\beta, \lambda)^{1/2} < \infty \). We have
\[ h_{2n}(\beta, \lambda) = \frac{(-\beta \lambda)^{2n}}{n!} \int_{\Delta_{\beta}^n} \omega^n_{\text{osc}} \left( (e^{-\beta s_{2n}}H_{el} q e^{\beta s_{2n}}H_{el}) - (e^{-\beta s_{2n}}H_{el} q e^{\beta s_{2n}}H_{el}) + \omega^f \left( (e^{-\beta s_{2n}}H \Phi(f) e^{\beta s_{2n}}H) \right) \right) ds_2n. \] (51)

Moreover, we have
\[ e^{-\beta s_{k}H_{el}} q e^{\beta s_{k}H_{el}} = (2\Theta)^{-1/2} \left( e^{-\beta \Theta s_{k} A^*} + e^{\beta \Theta s_{k} A} \right) \]
\[ e^{-\beta s_{k}H} \Phi(f) e^{\beta s_{k}H} = 2^{-1/2} \left( a^* (e^{-\beta s_{k}} |k| f) + a(e^{-\beta s_{k}} |k| f) \right). \] (52)

Inserting the identities of Equation (52) in Equation (51) and applying Wick’s theorem [5, p. 40] yields
\[ h_{2n}(\beta, \lambda) = (\lambda \beta)^{2n} \int_{\Delta_{\lambda}^{\beta n}} \sum_{P \in Z_2} \prod_{(i, j) \in P} K_{\text{osc}}(|i - j|, \beta) \]
\[ = (\lambda \beta)^{2n} \int_{[0, \lambda]^\Delta_{\beta n}} \sum_{P, P' \in Z_2} \prod_{(i, j) \in P} \prod_{(k, l) \in P'} K_{\text{osc}}(|i - j|, \beta) K_f(|k - l|, \beta) ds_2n, \] (53)

where for \( k < l \) and \( i < j \), such as
\[ K_f(|k - l|, \beta) := \omega^f \left( (e^{-\beta s_{k}H} \Phi(f) e^{\beta s_{k}H}) - (e^{-\beta s_{k}H} \Phi(f) e^{\beta s_{k}H}) \right) \]
\[ K_{\text{osc}}(|i - j|, \beta) := \omega^n_{\text{osc}} \left( e^{-\beta s_{k}H_{el}} q e^{\beta s_{k}H_{el}} e^{-\beta s_{k}H_{el}} q e^{\beta s_{k}H_{el}} \right). \]

The last equality in (53) holds, since the integrand is invariant with respect to a change of the axis of coordinates.

We interpret two pairings \( P \) and \( P' \) as an indirected graph \( G = G(P, P') \), where \( M_{2n} = \{1, \ldots, 2n\} \) is the set of points. Any graph in \( G \) has two kinds of lines, namely lines in \( L_{\text{osc}}(G) \), which belong to elements of \( P \) and lines in \( L_f(G) \), which belong to elements of \( P' \).

Let \( G(A) \) be the set of undirected graphs with points in \( A \subset M_{2n} \), such that for each point "i" in \( A \), there is exact one line in \( L_f(G) \), which begins in "i", and...
exact one line in \( L_{osc}(G) \), which begins with "if". \( \mathcal{G}_c(A) \) is the set of connected graphs. We do not distinguish, if points are connected by lines in \( L_f(G) \) or by lines in \( L_{osc}(G) \).

Let

\[
P_k := \left\{ P : P = \{ A_1, \ldots, A_k \}, \emptyset \neq A_i \subset M_{2n}, A_i \cap A_j = \emptyset \text{ for } i \neq j, \bigcup_{i=1}^k A_i = M_{2n} \right\}
\]

be the family of decompositions of \( M_{2n} \) in \( k \) disjoint sets. It follows

\[
J(G, A, \beta) := \int_{A_k} \prod_{(i,j) \in E_{osc}(G)} K_{osc}([s_i - s_j], \beta)K_f([s_k - s_l], \beta) \, \text{d}\mathcal{L}_n
\]

\begin{align}
(54)
\int_{A_k} \text{d}\mathcal{L}_n = \int_{j_1}^{1} \text{d}s_{j_1} \int_{j_2}^{1} \text{d}s_{j_2} \cdots \int_{j_m}^{1} \text{d}s_{j_m}, \quad \text{where } A_k = \{ j_1, \ldots, j_m \}
\end{align}

\[
(55)
\]

\[
J(G, A, \beta) := \int_{A_k} \prod_{(i,j) \in E_{osc}(G)} K_{osc}([s_i - s_j], \beta)K_f([s_k - s_l], \beta) \, \text{d}\mathcal{L}_n
\]

\[
(56)
\sum_{\substack{A_1, \ldots, A_k \subset M_{2n}, \#A_i = 2m_i \\{ A_1, \ldots, A_k \} \in P_k}} 1 = \frac{(2n)!}{(2m_1)! \cdots (2m_k)!}.
\]

**Documenta Mathematica** 16 (2011) 177–208
Let now be $A_n \subset M_{2n}$ with $\# A_n = 2m_n > 2$ fixed. In $G_n$ are $\# A_n$ lines in $L_{osc}(G_n)$, since such lines have no points in common, we have $\binom{2m_n}{m_n}^{2m_n}$ choices! Let now be the lines in $L_{osc}(G_n)$ fixed. We have now $(2m_n - 2)(2m_n - 4)\cdots 1$ choices for $m_n$ lines in $L_f(G_n)$, which yield a connected graph. Thus

$$\sum_{G_n \in G_n(A_n)} 1 = \frac{(2m_n)!}{m_n!2^{2m_n}}((2m_n - 2)(2m_n - 4)\cdots 1) = \frac{(2m_n)!}{2^{2m_n}}. \quad (57)$$

For $\# A_n = 2$ exists only one connected graph. We obtain for $h_{2n}$

$$h_{2n} = (\lambda)^{2n} \sum_{k=1}^{2n} \frac{1}{k!} \sum_{(m_1, \ldots, m_k) \in \mathbb{N}^k} \prod_{a=1}^{k} \frac{J(2m_a, \beta)(\beta^2)^{m_a}}{2m_a} \quad (58)$$

$$\leq (2\Theta^{-1} ||k|^{-1/2} f|| \lambda)^{2n} \sum_{k=1}^{2n} \frac{1}{k!} \sum_{(m_1, \ldots, m_k) \in \mathbb{N}^k} \prod_{a=1}^{k} \frac{(C \beta + 1)}{2m_a} \quad (59)$$

$$\leq (2\Theta^{-1} ||k|^{-1/2} f|| \lambda)^{2n} \sum_{k=1}^{2n} \left(\frac{(C \beta + 1)^2}{2} \frac{\sum_{m=1}^{n} 1}{m^k}\right).$$

Since the $\sum_{m=1}^{n} \frac{1}{m^k}$ can be considered as a lower Riemann sum for the integral $\int_1^{n+1} r^{-1} dr$, we have $\sum_{m=1}^{n} \frac{1}{m^k} \leq \ln(n + 1)$. Thus,

$$h_{2n} \leq (2\Theta^{-1} ||k|^{-1/2} f|| \lambda)^{2n} \sum_{k=1}^{2n} \left(\frac{(C \beta + 1)^2}{2} \frac{\sum_{m=1}^{n} 1}{m^k}\right).$$

Since $2||\lambda| ||k|^{1/2} f|| < \Theta$ the series $\sum_{n=0}^{\infty} h_{2n} \lambda^{1/2}$ converges absolutely for all $\beta > 0$. It follows, that

$$e^{-\beta/2(C_0 + Q)} \Omega_0^\beta = \Omega_0^\beta + \sum_{n=1}^{\infty} \int_{\Delta_{n/2}} Q(\xi_n) \Omega_0^\beta d\xi_n$$

exists. \hfill $\square$

Conversely, Equation (58) and Lemma 7.2 imply

$$h_{2n} \geq \lambda^{2n} \frac{J(2n, \beta^2)n}{2n} = \left(\Theta^{-1} \frac{\int \frac{\beta^2 f(k)^2}{\sinh(\beta \theta) / \sinh(\beta \theta / 2)} dk}{2n}\right)^{n}. \quad (60)$$

Hence for every $\beta > 0$ exists a $\lambda \in \mathbb{R}$, such that $h_{2n} (\beta, \lambda) \geq \frac{1}{2n}$. Thus $\sum_{n=1}^{\infty} h_{2n} (\beta, \lambda)^{1/2} = \infty$.

Remark 7.1. We can therefore not extended Theorem 1.4 to an existence proof for all $\lambda > 0$. 

Documenta Mathematica 16 (2011) 177–208
Lemma 7.2. Following statements are true.

\[ J(G, A, \beta) = J(#A, \beta), \ G \in \mathcal{G}_c(A) \]
\[ J(#A, \beta) \leq (2\|k^{-1/2}f\|_2(\Theta \beta)^{-1})^{#A} \cdot (C \beta + 1) \]
\[ J(#A, \beta) \geq \left( \Theta^{-1} \int \frac{|f(k)|^2}{\sinh(|k| \beta/2) \sinh(\Theta \beta/2)} dk \right)^{#A/2}, \]

where \( #A = 2m \) and \( C = (1/2) \frac{\|f\|^2}{\|k\|^2} \).

Proof of 7.2. A relabeling of the integration variables yields

\[ J(G, A, \beta) \leq \mathcal{K}_f \int_{[0,1]^{2m}} K_{osc}(|t_1 - t_2|, \beta) K_f(|t_2 - t_3|, \beta) \cdots \]
\[ \cdots K_{osc}(|t_{2m-1} - t_{2m}|, \beta) dt \]

for \( \mathcal{K}_f := \sup_{s \in [0,1]} K_f(s, \beta) \). We transform due to \( s_i := t_i - t_{i+1}, \ i \leq 2m-1 \) and \( s_{2m} = t_{2m} \), hence \(-1 \leq s_i \leq 1, \ i = 1, \ldots, 2m \), since integrating a positive function we obtain

\[ J(G, A, \beta) \leq \left( \int_{-1}^{1} K_{osc}(|s|, \beta) ds \right)^m \left( \int_{-1}^{1} K_f(|s|, \beta) ds \right)^{m-1} \]
\[ \cdot \sup_{s \in [0,1]} K_f(s, \beta). \]

We recall that

\[ \int_{-1}^{1} K_{osc}(|s|, \beta) ds = (2\Theta)^{-1} \int_{-1}^{1} \frac{\cosh(\Theta |s| \beta/2)}{\sinh(\Theta \beta/2)} ds = 2(\Theta^2 \beta)^{-1} \]

and

\[ \int_{-1}^{1} K_f(|s|, \beta) ds = \int_{-1}^{1} \int \frac{\cosh(\beta |s| |k| - \beta|k|/2) |f(k)|^2}{2 \sinh(\beta |k|/2)} dk ds \]
\[ = 2 \int \frac{|f(k)|^2}{\beta |k|} dk. \]

Using \( \coth(x) \leq 1 + 1/x \) and using convexity of \( \cosh \), we obtain

\[ \sup_{s \in [0,1]} K_f(s, \beta) \leq (1/2) \int |f(k)|^2 dk + \frac{1}{\beta} \int \frac{|f(k)|^2}{|k|} dk. \]

Due to the fact, that \( t \mapsto K_f(t, \beta) \) and \( t \mapsto K_{osc}(t, \beta) \) attain their minima at \( t = 1/2 \), we obtain the lower bound for \( J(#A, \beta) \).
Remark 7.3. In the literature there is one criterion for \( \Omega_{0}^{\beta} \in \text{dom}(e^{-\beta/2(L_{0}+Q)}) \), to our knowledge, that can be applied in this situation [6]. One has to show that \( \|e^{-\beta/2Q}\Omega_{0}^{\beta}\| < \infty \). If we consider the case, where the criterion holds for \( \pm \lambda \), then the expansion in \( \lambda \) converges,

\[
\|e^{-\beta/2Q}\Omega_{0}^{\beta}\|^{2} = \sum_{n=0}^{\infty} \frac{(\lambda \beta)^{2n}}{(2n)!} \omega_{\beta}^{2} q^{2n} \omega_{\beta}^{2}(\Phi(f)^{2n})
\]

\[
= \sum_{n=0}^{\infty} \frac{(\lambda \beta)^{2n}}{(2n)!} \left( \frac{2n!}{n! 2^{n}} \right)^{2} K_{osc}(0, \beta)^{n} K_{f}(0, \beta)^{n}
\]

\[
= \sum_{n=0}^{\infty} (\lambda \beta)^{2n} \Theta^{-n} \left( \frac{2n!}{n!} \right)^{2} K_{osc}(0, \beta)^{n} K_{f}(0, \beta)^{n}
\]

\[
\geq \sum_{n=0}^{\infty} (\lambda \beta)^{2n} (4 \Theta)^{-n} \left( \int |f(k)|^{2} \coth(\beta |k|/2) dk \right)^{n}.
\]

Obviously, for any value of \( \lambda \neq 0 \), there is a \( \beta > 0 \), for which \( \|e^{-\beta/2Q}\Omega_{0}^{\beta}\| < \infty \) is not fulfilled.

Acknowledgments:
This paper is part of the author’s PhD requirements. I am grateful to Volker Bach for many useful discussions and helpful advice. The main part of this work was done during the author’s stay at the Institut for Mathematics at the University of Mainz. The work has been partially supported by the DFG (SFB/TR 12).

Lemma A.1. Let \( f, g : \{ z \in \mathbb{C} : 0 \leq \Re(z) \leq \alpha \} \rightarrow \mathbb{C} \) continuous and analytic in the interior. Moreover, assume that \( f(t) = g(t) \) for \( t \in \mathbb{R} \). Then \( f = g \).

Proof of A.1. Let \( h : \{ z \in \mathbb{C} : \Im(z) < \alpha \} \rightarrow \mathbb{C} \) defined by

\[
h(z) := \begin{cases} f(z) - g(z), & \text{on } \{ z \in \mathbb{C} : 0 \leq \Im(z) < \alpha \} \\ \frac{f(\overline{z})}{g(\overline{z})}, & \text{on } \{ z \in \mathbb{C} : -\alpha < \Im(z) < 0 \} \end{cases}
\]

(61)

Thanks to the Schwarz reflection principle \( h \) is analytic. Since \( h(t) = 0 \) for all \( t \in \mathbb{R} \), we get \( h = 0 \). Hence \( f = g \) on \( \{ z \in \mathbb{C} : 0 \leq \Re(z) < \alpha \} \). Since both \( f \) and \( g \) are continuous, we infer that \( f = g \) on the whole domain. \( \square \)

Lemma A.2. Let \( H \) be some self-adjoint operator in \( \mathcal{H} \), \( \alpha > 0 \) and \( \phi \in \text{dom}(e^{\alpha H}) \). Then \( \phi \in \text{dom}(e^{z H}) \) for \( z \in \{ z \in \mathbb{C} : 0 \leq \Re(z) \leq \alpha \} \). \( z \mapsto e^{z H} \phi \) is continuous on \( \{ z \in \mathbb{C} : 0 \leq \Re(z) \leq \alpha \} \) and analytic in the interior.
Proof of A.2. Due to the spectral calculus we have
\[
\int e^{2\Re z s} d(\phi | E_s \phi) \leq \int (1 + e^{2\alpha s}) d(\phi | E_s \phi) =: C_1^2 < \infty.
\]

Thus \( \phi \in \text{dom}(e^{zH}) \). Let \( \psi \in \mathcal{H} \) and \( f(z) = (\psi | e^{zH} \phi) \). There is a sequence \( \{ \psi_n \} \) with \( \psi_n \in \bigcup_{m \in \mathbb{N}} \text{ran} \| H \| \leq m \) and \( \lim_{n \to \infty} \psi_n = \psi \). We set \( f_n(z) = (\psi_n | e^{zH} \phi) \). It is not hard to see that \( f_n \) is analytic, since \( \psi_n \) is an analytic vector for \( H \), and that \( |f_n(z)| \leq C_1 \| \psi_n \| \) and \( \lim_{n \to \infty} f_n(z) = f(z) \). Thus \( f \) is analytic and hence \( z \mapsto e^{zH} \phi \) is analytic. Thanks to the dominated convergence theorem the right side of
\[
\|e^{z_n H} \phi - e^{z H} \phi\|^2 \leq \int (e^{2\Re z_n s} + e^{2\Re z s} - e^{z_n s + z s} - e^{z s + z_n s}) d(\phi | E_s \phi) \quad (62)
\]
tends to zero for \( \lim_{n \to \infty} z_n = z \). This implies the continuity of \( z \mapsto e^{zH} \phi \). \( \square \)

Lemma A.3. We have for \( n_1 + n_2 \geq 1 \)
\[
\int_{\Delta_1^n} C_n(\underline{\alpha}) d\underline{\alpha}_n \leq \text{const } C^{n_2}_1 (n_1 + n_2)! (n_1 + 1)(1 - 2\gamma)_{n_2 - 2} \quad (63)
\]

Proof of A.3. We turn now to the integral
\[
\int_{\Delta_1^n} C_n(\underline{\alpha}) d\underline{\alpha}_n = \int_{\Delta_1^n} (1 - s_1 + s_n)^{-\alpha_1} \prod_{i=1}^{n-1} (s_i - s_{i+1})^{-\alpha_i} d\underline{s}_n. \quad (64)
\]

We define for \( k = 1, \ldots, 2n \), a change of coordinates by \( s_k = r_1 - \sum_{j=2}^{k} r_j \), the integral transforms to
\[
\int_{S^n} (1 - (r_2 + \cdots + r_n))^{-\alpha_1} \prod_{i=2}^{n} r_i^{-\alpha_i} d\underline{r}_n \quad (65)
\]
\[
= \int_{T^{n-1}} (1 - (r_2 + \cdots + r_n))^{1 - \alpha_1} \prod_{i=2}^{n} r_i^{-\alpha_i} d\underline{r}_{n-1}
\]
\[
= \frac{\Gamma(1 - \alpha_1)^{-1} \Gamma(1 - \gamma)^{2n_2}}{\Gamma(n_1 + 2n_2 (1 - \gamma))}
\]

where \( S^{2n} := \{ r \in \mathbb{R}^{2n} : 0 \leq r_i \leq 1, r_2 + \cdots + r_{2n} \leq r_1 \} \) and \( T^{2n-1} := \{ \underline{r} \in \mathbb{R}^{2n-1} : 0 \leq r_i \leq 1, r_2 + \cdots + r_{2n} \leq 1 \} \). From the first to the second formula we integrate over \( dr_1 \). The last equality follows from [11, Formula 4.635 (4)], here \( \Gamma \) denotes the Gamma-function.

From Stirling’s formula we obtain
\[
(2\pi)^{1/2} x^{x - 1/2} e^{-x} \leq \Gamma(x) \leq (2\pi)^{1/2} x^{x - 1/2} e^{-x + 1}, \quad x \geq 1. \quad (66)
\]
Since \( n_1 + n_2 \geq 1 \) get

\[
\frac{\Gamma(n_1 + n_2 + 1)}{\Gamma(n_1 + 2(1 - \gamma)n_2)} \leq (n + 1)^2 \left( \frac{n_1 + 2(1 - \gamma)n_2}{e} \right)^{-(1-2\gamma)n_2}. \tag{67}
\]

Note that \( \Gamma(n_1 + n_2 + 1) = (n_1 + n_2)! \).

**Lemma A.4.** Let \((1 + \alpha(k)^{-1/2}) f_1, \ldots, (1 + \alpha(k)^{-1/2}) f_{2m} \in \mathcal{H}_{ph} \) and \( \sigma \in \{+, -\}^{2m} \). Let \( a^+ = a^* \) and \( a^- = a \)

\[
\omega^\beta_j \left( a^{\sigma_{2m}}(e^{-\sigma_{2m} s_{2m} \alpha(k)} f_{2m}) \cdots a^{\sigma_1}(e^{-\sigma_1 s_1 \alpha(k)} f_1) \right) = \int \mathcal{F}^\beta_{2m}(k_{2m}, \tau_{2m}) \cdots \mathcal{F}^\sigma_1(k_1, \tau_1) \nu(dk_{2m} \otimes d\tau_{2m}),
\]

where \( \nu(dk_{2m} \otimes d\tau_{2m}) \) is a measure on \((\mathbb{R}^3)^{2m} \times \{+, -\}^{2m} \) for phonons, respectively on \((\mathbb{R}^3 \times \{\pm\}^{2m} \times \{+, -\}^{2m} \) for photons, and

\[
\nu(dk_{2m} \otimes d\tau_{2m}) \leq \sum_{P \in \mathcal{Z}_{2m}} \sum_{P \in \{+, -\}^{2m} \{i > j\} \in P} \prod_{k = 1} \left( \delta_{k_i, k_j} \coth(\beta \alpha(k_i)/2) \right) dk_{2m}.
\]

for \( f^+(k, \tau) := f(k) 1[\tau = +] \) and \( f^+(k, \tau) := f(k) 1[\tau = -] \).

**Proof of A.4.** Since \( \omega^\beta_j \) is quasi-free, we obtain with \( a^+ := a^* \) and \( a^- := a \)

\[
\omega^\beta_j (a^{\sigma_{2m}}(e^{-\sigma_{2m} s_{2m} \alpha(k)} f_{2m}) \cdots a^{\sigma_1}(e^{-\sigma_1 s_1 \alpha(k)} f_1)) = \sum_{P \in \mathcal{Z}_{2}} \prod_{i > j} \omega^\beta_j(a^{\sigma_i}(e^{-\sigma_i s_i \alpha(k)} f_i) a^{\sigma_j}(e^{-\sigma_j s_j \alpha(k)} f_j)),
\]

see Equation (12). For the expectation of the so called two point functions we obtain:

\[
\omega^\beta_j(a^+(e^{s_i \alpha(k)} f_i) a^+(e^{s_j \alpha(k)} f_j)) = 0 = \omega^\beta_j(a(e^{-s_i \alpha(k)} f_i) a(e^{-s_j \alpha(k)} f_j)),
\]

such as

\[
\omega^\beta_j(a^+(e^{s_i \alpha(k)} f_i) a^+(e^{-s_j \alpha(k)} f_j)) = \int f_i(k) f_j(k) \frac{e^{x(s_i, s_j \alpha(k))}}{e^{\beta \alpha(k) - 1}} dk
\]

\[
\omega^\beta_j(a^+(e^{s_i \alpha(k)} f_i) a^+(e^{s_j \alpha(k)} f_j)) = \int f_i(k) f_j(k) \frac{e^{(\beta+2s_j \alpha(k)) \alpha(k)}}{e^{\beta \alpha(k) - 1}} dk
\]

Hence it follows

\[
\omega^\beta_j (a^{\sigma_{2m}}(e^{-\sigma_{2m} s_{2m} \alpha(k)} f_{2m}) \cdots a^{\sigma_1}(e^{-\sigma_1 s_1 \alpha(k)} f_1)) = \int \mathcal{F}^\beta_{2m}(k_{2m}, \tau_{2m}) \cdots \mathcal{F}^\sigma_1(k_1, \tau_1) \nu(dk_{2m} \otimes d\tau_{2m}),
\]

**References:**

- [Documenta Mathematica 16 (2011) 177–208]
where $f^+(k, \tau) := f(k) 1[\tau = +]$ and $f^-(k, \tau) := \overline{f(k)} 1[\tau = -]$.

$\nu(d^{(2m)} k \otimes d^{2m} \tau)$ is a measure on $(\mathbb{R}^3)^{2m} \times \{+,-\}^{2m}$, which is defined by

$$\sum_{P \in \mathbb{Z}^{2m}} \sum_{\ell \in \{+,-\}^{2m}} \prod_{i > j \in P} \delta_{r_{ij}, -r} \delta_{k_i, k_j}$$

$$\left( \delta_{r, +} \frac{e^{\alpha(s_i - s_j) \alpha(k_i)}}{e^{\beta \alpha(k_i)} - 1} + \delta_{r, -} \frac{e^{\beta - (s_i - s_j) \alpha(k_i)}}{e^{\beta \alpha(k_i)} - 1} \right) \, dk_{2m}. \quad (69)$$

References

[1] H. Araki, E. Woods. Representations of the Canonical Commutation Relations describing a non-relativistic infinite free Bose Gas. *J. Math. Phys.*, 4:637–662, 1963.

[2] V. Bach, J. Fröhlich, and I. M. Sigal. Quantum electrodynamics of confined non-relativistic particles. *Adv. in Math.*, 137:299–395, 1998.

[3] V. Bach, J. Fröhlich, I. M. Sigal. Return to Equilibrium. *J. Math. Phys.*, 41:3985–4060, 2000.

[4] O. Bratteli, D. Robinson. Operator Algebras and Quantum Statistical Mechanics 1, Text and Monographs in Physics, Springer-Verlag, 1987.

[5] O. Bratteli, D. Robinson. Operator Algebras and Quantum Statistical Mechanics 2, Text and Monographs in Physics, Springer-Verlag, 1996.

[6] Dereziński J., Jaksic V., Pillet C.-A. Perturbation theory of $W^*$-dynamics, Liouvileans and KMS-states. *Rev. Math. Phys.*, 15: 447–489, 2003.

[7] J. Dereziński, V. Jaksic. Spectral theory of Pauli-Fierz operators. *J. Funct. Anal.*, 180: 243–327, 2001

[8] J. Dereziński, V. Jaksic. Return to Equilibrium for Pauli-Fierz Operators. *Ann. Henri Poincare*, 4: 739–793, 2003

[9] J. Fröhlich, M. Merkli. Another Return to Equilibrium *Commun. Math. Phys.*, 251: 235–262, 2004.

[10] J. Fröhlich, M. Merkli, I. M. Sigal. Ionization of Atoms in a Thermal Field. *Journal of Statistical Physics*, 116: 311–359, 2004. DOI 10.1023/B:JOSS.0000037226.16493.5e

[11] I.S. Gradstein, I.M. Ryzhik. Table of Integrals, Series, and Products. Academic Press, 1980, 4

[12] R. Haag, N. Hugenholtz, M. Winnink. On the Equilibrium States in Quantum Statistical Mechanics. *Commun. Math. Phys.*, 5: 215–236, 1967.
[13] V. Jakšić, C. A. Pillet. On a Model for Quantum Friction. II: Fermi’s
Golden Rule and Dynamics at Positive Temperature.  Commun. Math.
Phys., 176:619–643, 1996

[14] V. Jakšić, C. A. Pillet. On a Model for Quantum Friction III: Ergodic
Properties of the Spin-Boson System. Commun. Math. Phys., 178:627–
651, 1996

[15] M. Merkli. Positive Commutators in Non-Equilibrium Statistical Quantum
Mechanics, Commun. Math. Phys., 223: 327–362, 2001.

[16] M. Reed, B. Simon. Methods of Modern Mathematical Physics: I. Func-
tional Analysis Academic Press, 1980.

[17] M. Reed, B. Simon. Methods of Modern Mathematical Physics: II. Fourier
Analysis and Self-Adjointness Academic Press, 1980.
**Divisorial Cohomology Vanishing on Toric Varieties**

Markus Perling

Received: March 10, 2010  
Revised: February 24, 2011  
Communicated by Thomas Peternell

**Abstract.** This work discusses combinatorial and arithmetic aspects of cohomology vanishing for divisorial sheaves on toric varieties. We obtain a refined variant of the Kawamata-Viehweg theorem which is slightly stronger. Moreover, we prove a new vanishing theorem related to divisors whose inverse is nef and has small Iitaka dimension. Finally, we give a new criterion for divisorial sheaves for being maximal Cohen-Macaulay.

2010 Mathematics Subject Classification: 13C14, 14B15, 14M25  
Keywords and Phrases: toric varieties, cohomology, local cohomology, Weil divisors, Maximal Cohen-Macaualy modules

**Contents**

1 **Introduction** .... 210

2 **Toric Preliminaries** .... 216

3 **Discriminants and Combinatorial Aspects Cohomology Vanishing** .... 218
   3.1 Circuits .... 219
   3.2 Circuits and discriminantal arrangements .... 221
   3.3 Secondary Fans .... 225
   3.4 MCM sheaves, \(\mathbb{Q}\)-Cartier divisors and the toric Kawamata-Viehweg vanishing theorem .... 228

*Documenta Mathematica* 16 (2011) 209–251
1 Introduction

This work is motivated by numerical experiments [Per04] related to the conjecture of King [Kin97] concerning the derived category smooth complete toric varieties. These experiments led to the calculations of [HP06], where a counterexample to King’s conjecture was given. Our goal is to develop a more systematic approach to the combinatorial and arithmetic aspects of cohomology vanishing for divisorial sheaves on toric varieties and to better understand from these points of view some phenomena related to this problem.

Based on work of Bondal (see [Rud90], [Bon90]), it was conjectured [Kin97] that on every smooth complete toric variety \(X\) there exists a full strongly exceptional collection of line bundles. That is, a collection of line bundles \(L_1, \ldots, L_n\) on \(X\) which generates \(D^b(X)\) and has the property that \(\text{Ext}^k(L_i, L_j) = 0\) for all \(k > 0\) and all \(i, j\). Such a collection induces an equivalence of categories \(\text{RHom}\left(\bigoplus L_i, \cdot\right) : D^b(X) \to D^b(\text{End}(\bigoplus L_i) \mod)\). This possible generalization of Beilinson’s theorem (pending the existence of a full strongly exceptional collection) has attracted much interest, notably also in the context of the homological mirror conjecture [Kon95]. For line bundles, the problem of \(\text{Ext}\)-vanishing can be reformulated to a problem of cohomology vanishing for line bundles by the isomorphisms

\[
\text{Ext}^k(L_i, L_j) \cong H^k(X, L_i \otimes L_j) = 0 \quad \text{for all } k \geq 0 \text{ and all } i, j.
\]

So we are facing a quite peculiar cohomology vanishing problem: let \(n\) denote the rank of the Grothendieck group of \(X\), then we look for a certain constellation of \(n(n - 1)\) – not necessarily distinct – line bundles, all of which have vanishing higher cohomology groups. The strongest general vanishing theorems so far are of the Kawamata-Viehweg type (see [Mus02] and [Fuj07], and also [Mat02] for Bott type formulas for cohomologies of line bundles), but it can be seen from very easy examples, such as Hirzebruch surfaces, that these alone in general do not suffice to prove or disprove the existence of strongly exceptional collections by means of cohomology vanishing. In [HP06], on a certain toric surface \(X\), all line bundles \(\mathcal{L}\) with the property that \(H^i(X, \mathcal{L}) = H^i(X, \mathcal{L}^\vee) = 0\) for all \(i > 0\) were completely classified by making use of an explicit toric representation of the cohomology vanishing problem for line bundles. This approach exhibits quite complicated combinatorial as well as number theoretic conditions for cohomology vanishing which we are going to describe in general.

We will consider and partially answer the following more general problem. Let \(D\) be a Weil divisor on any toric variety \(X\) and \(V \subset X\) a torus invariant closed

---

Documenta Mathematica 16 (2011) 209–251
subscheme. Then what are necessary and sufficient conditions for the (global) local cohomology modules $H^i_V(X, \mathcal{O}_X(D))$ to vanish? Given this spectrum of cohomology vanishing problems, we have at one extreme the cohomology vanishing problem for line bundles, and at the other extreme the classification problem for maximal Cohen Macaulay (MCM) modules over semigroup rings: on an affine toric variety $X$, the sheaf $\mathcal{O}_X(D)$ is MCM if and only if the local cohomologies $H^i_x(X, \mathcal{O}_X(D))$ vanish for $i \neq \dim X$, where $x \in X$ is the torus fixed point. These local cohomologies have been studied by Stanley [Sta82], [Sta96] and Bruns and Gubeladze [BG03] showed that only finitely many sheaves in this class are MCM. MCM sheaves over affine toric varieties have only been classified for some special cases (see for instance [BGS87] and [Yos90]). Our contribution will be to give a more explicit combinatorial characterization of MCM modules of rank one over normal semigroup rings and their ties to the birational geometry of toric varieties.

Our main results. The first main result will be an integral version of the Kawamata-Viehweg vanishing theorem. Consider the nef cone $\text{nef}(X) \subset A_{d-1}(X)_{\mathbb{Q}}$, then the toric Kawamata-Viehweg vanishing theorem (see Theorem 3.29) can be interpreted such that if $D - K_X$ is contained in the interior of $\text{nef}(X)$, then $H^i(X, \mathcal{O}_X(D)) = 0$ for all $i > 0$. For our version we will define a set $\mathfrak{A}_{\text{nef}} \subset A_{d-1}(X)$, which we call the arithmetic core of $\text{nef}(X)$ (see definition 4.11). The set $\mathfrak{A}_{\text{nef}}$ has the property that it contains all integral Weil divisors which map to the interior of the cone $K_X + \text{nef}(X)$ in $A_{d-1}(X)_{\mathbb{Q}}$. But in general it is strictly larger, as in the example above. We can lift the cohomology
vanishing theorem for divisors in nef($X$) to $\mathcal{A}_{\text{nef}}$:

**Theorem (4.14):** Let $X$ be a complete toric variety and $D \in \mathcal{A}_{\text{nef}}$. Then $H^i(X, \mathcal{O}_X(D)) = 0$ for all $i > 0$.

One can consider Theorem 4.14 as an “augmentation” of the standard vanishing theorem for nef divisors to the subset $\mathcal{A}_{\text{nef}}$ of $A_{d-1}(X)$. In general, Theorem 4.14 is slightly stronger than the toric Kawamata-Viehweg vanishing theorem and yields refined arithmetic conditions.

However, the main goal of this paper is to find vanishing results which cannot directly be derived from known vanishing theorems. Let $D$ be a nef Cartier divisor whose Iitaka dimension is positive but smaller than $d$. This class of divisors is contained in nonzero faces of the nef cone of $X$ which are contained in the intersection of the nef cone with the boundary of the effective cone of $X$ (see Section 4.3). Let $F$ be such a face. Similarly as with $\mathcal{A}_{\text{nef}}$, we can define for the inverse cone $-F$ an arithmetic core $\mathcal{A}_{-F}$ (see 4.11) and associate to it a vanishing theorem, which may be considered as the principal result of this article:

**Theorem (4.17):** Let $X$ be a complete $d$-dimensional toric variety. Then $H^i(X, \mathcal{O}(-D)) = 0$ for every $i$ and all $D$ which are contained in some $\mathcal{A}_{-F}$, where $F$ is a face of nef($X$) which contains nef divisors of Iitaka dimension $0 < \kappa(D) < d$. If $\mathcal{A}_{-F}$ is nonempty, then it contains infinitely many divisor classes.

This theorem cannot be an augmentation of a vanishing theorem for $-F$, as it is not true in general that $H^i(X, \mathcal{O}_X(-D)) = 0$ for all $i$ for $D$ nef of Iitaka dimension smaller than $d$. In particular, the set of $\mathbb{Q}$-equivalence classes of elements in $\mathcal{A}_{-F}$ does not intersect $-F$.

For the case of a toric surface $X$ we show that above vanishing theorems combine to a nearly complete vanishing theorem for $X$. Recall that in the fan associated to a complete toric surface $X$ every pair of opposite rays by projection gives rise to a morphism from $X$ to $\mathbb{P}^1$ (e.g. such a pair does always exist if $X$ is smooth and $X \neq \mathbb{P}^2$). Correspondingly, we obtain a family of nef divisors of Iitaka dimension 1 on $X$ given by the pullbacks of the sheaves $\mathcal{O}_{\mathbb{P}^1}(i)$ for $i > 0$. We get:

**Theorem (4.21):** Let $X$ be a complete toric surface. Then there are only finitely many divisors $D$ with $H^i(X, \mathcal{O}_X(D)) = 0$ for all $i > 0$ which are not contained in $\mathcal{A}_{\text{nef}} \cup \bigcup_F \mathcal{A}_{-F}$, where the union ranges over all faces of nef($X$) which correspond to pairs of opposite rays in the fan associated to $X$.

Some more precise numerical characterizations on the sets $\mathcal{A}_{-F}$ will be given in subsection 4.3. The final result is a birational characterization of MCM-sheaves of rank one. This is a test case to see whether point of view of birational geometry might be useful for classifying more general MCM-sheaves. The idea for this comes from the investigation of MCM-sheaves over surface singularities in...
terms of resolutions in the context of the McKay correspondence (see [GSV83], [AV85], [EK85]). For an affine toric variety $X$, in general one cannot expect to find a similar nice correspondence. However, there is a set of preferred partial resolutions of singularities $\pi : \tilde{X} \rightarrow X$ which is parameterized by the secondary fan of $X$. Our result is a toric analog of a technical criterion of loc. cit.

**Theorem (4.36):** Let $X$ be a $d$-dimensional affine toric variety whose associated cone has simplicial facets and let $D \in A_{d-1}(X)$. If $R^i\pi_*\mathcal{O}_{\tilde{X}}(\pi^*D) = 0$ for every regular triangulation $\pi : \tilde{X} \rightarrow X$, then $\mathcal{O}_X(D)$ is MCM. For $d = 3$ the converse is also true.

Note that the facets of a 3-dimensional cone are always simplicial.

To prove our results we will require a lot of bookkeeping, combining various geometric, combinatorial and arithmetic aspects of toric varieties. This has the unfortunate effect that the exposition will be rather technical and incorporate many notions (though not much theory) coming from combinatorics. As this might be cumbersome to follow for the more geometrically inclined reader, we will give an overview of the key structures and explain how they fit together.

**The Circuit Geometry of a Toric Variety.** In order to compute the cohomology $H^i_V(X, \mathcal{O}_X(D))$ of a torus-invariant Weil divisor $D = \sum_{i=1}^n c_i D_i$ with respect to some torus-invariant support $V \subseteq X$, one uses the induced eigenspace decomposition

$$H^i_V(X, \mathcal{O}_X(D)) \cong \bigoplus_{m \in M} H^i_V(X, \mathcal{O}_X(D))_m.$$ 

By a well-known formula, we can compute every eigenspace by computing the relative cohomology of a certain simplicial complex:

$$H^i_V(X, \mathcal{O}_X(D))_m \cong H^{i-1}(\tilde{\Delta}_m, \tilde{\Delta}_{V,m}; k).$$

Here $\tilde{\Delta}$ denotes the simplicial model of $\Delta$, i.e. the abstract simplicial complex on the set $[n]$ such that any subset $I \subseteq [n]$ is in $\tilde{\Delta}$ if and only if there exists a cone $\sigma$ in $\Delta$ such that elements in $I$ are faces of $\sigma$. Similarly, $\tilde{\Delta}_V$ is a subcomplex of $\tilde{\Delta}$, generated by only those cones in $\Delta$ whose associated orbits in $X$ are not contained in $V$ (see also Section 2). For any character $m \in M$, $\tilde{\Delta}_m$ and $\tilde{\Delta}_{V,m}$ are the full subcomplexes which are supported on those $l_i$ with $l_i(m) < -c_i$ (see Theorem 2.1).
By this, for an invariant divisor \( D = \sum_{i=1}^{n} c_i D_i \), the eigenspaces \( H^*_V(X, \mathcal{O}_X(D)) \) depend on the simplicial complexes \( \Delta, \Delta_V \) as well as on the position of the characters \( m \) with respect to the hyperplanes \( H^*_c = \{ m \in M_\mathbb{Q} \mid l_i(m) = -c_i \} \), where \( M_\mathbb{Q} = M \otimes \mathbb{Q} \). The chamber decomposition of \( M_\mathbb{Q} \) induced by the \( H^*_c \) (or their intersection poset) can be interpreted as the combinatorial type of \( D \). Our strategy will be to consider the variations of combinatorial types depending on \( \zeta = (c_1, \ldots, c_n) \in \mathbb{Q}^n \). The solution to this discriminantal problem is given by the discriminantal arrangement associated to the vectors \( l_1, \ldots, l_n \), which has first been considered by Crapo [Cra84] and Manin and Schechtman [MS89]. The discriminantal arrangement is constructed as follows. Consider the standard short exact sequence associated to \( X \):

\[
0 \to M_\mathbb{Q} \xrightarrow{L} \mathbb{Q}^n \xrightarrow{D} A_\mathbb{Q} \to 0,
\]

where \( L \) is given by \( L(m) = (l_1(m), \ldots, l_n(m)) \), and \( A_\mathbb{Q} := A_{d-1}(X) \otimes \mathbb{Q} \) is the rational divisor class group of \( X \). The matrix \( D \) is called the Gale transform of \( L \), and its \( i \)-th column \( D_i \) is the Gale transform of \( l_i \). The most important property of the Gale transform is that the linear dependencies among \( l_i \) and among the \( D_i \) are inverted. That is, for any subset among the \( l_i \) which forms a basis, the complementary subset of the \( D_i \) forms a basis of \( A_\mathbb{Q} \), and vice versa. Moreover, for every circuit, i.e. a minimal linearly dependent subset, \( C \subset [n] \) the complementary set \( \{ D_i \mid l_i \notin C \} \) spans a hyperplane \( H_C \) in \( A_\mathbb{Q} \). Then the discriminantal arrangement is given by the hyperplane arrangement

\[
\{ H_C \mid C \subset [n] \text{ circuit} \}.
\]

The stratification of \( A_\mathbb{Q} \) by this arrangement then is in bijection with the combinatorial types of the arrangements given by the \( H^*_c \) under variation of \( \zeta \). As we will see, virtually all properties of \( X \) concerning its birational geometry and cohomology vanishing of divisorial sheaves on \( X \) depend on the discriminantal arrangement. In particular, (see Proposition 3.19), the discriminantal arrangement coincides with the hyperplane arrangement generated by the facets of the secondary fan. Ubiquitous standard constructions such as the effective cone, nef cone, and the Picard group can easily be identified as its substructures. Another interesting aspect is that the discriminantal arrangement by itself (or the associated matroid, respectively) represents a combinatorial invariant of the variety \( X \), which one can refer to as its circuit geometry. This circuit geometry refines the combinatorial information coming with the toric variety, that is, the fan \( \Delta \) and the matroid structure underlying the \( l_i \) (i.e. their linear dependencies). It depends only on the \( l_i \), and even for two combinatorially equivalent fans \( \Delta, \Delta' \) such that corresponding sets of primitive vectors \( l_1, \ldots, l_n \) and \( l'_1, \ldots, l'_n \) have the same underlying linear dependencies, their associated circuit geometries are different in general. This already is the case for surfaces, see, for instance, Crapo’s example of a plane tetrahedral line configuration ([Cra84], §4). Falk ([Fal94], Example 3.2) gives a 3-dimensional example.
Circuits and the Diophantine Frobenius problem. Circuits are also the building blocks for our arithmetic conditions on cohomology vanishing, which can easily be illustrated for the case of weighted projective spaces. Assume, for simplicity, that the first \( d + 1 \) primitive vectors \( l_1, \ldots, l_{d+1} \) generate \( N \) and form a circuit. Then we have a relation

\[
\sum_{i=1}^{d+1} \alpha_i l_i = 0
\]

where the \( \alpha_i \) are nonzero integers whose largest common divisor is one. This relation is unique up to sign and we assume for simplicity that \( \alpha_i > 0 \) for at least one \( i \). In the special case that all the \( \alpha_i \) are positive, \( l_1, \ldots, l_{d+1} \) generate the fan of a weighted projective space \( \mathbb{P}(\alpha_1, \ldots, \alpha_{d+1}) \). Denote \( D \) the unique \( \mathbb{Q} \)-effective generator of \( A_{d-1}(\mathbb{P}(\alpha_1, \ldots, \alpha_{d+1})) \). Then there is a standard construction for counting global sections

\[
\dim H^0(\mathbb{P}(\alpha_1, \ldots, \alpha_{d+1}), O_{\mathbb{P}(\alpha_1, \ldots, \alpha_{d+1})}(nD)) = \left| \left\{ (k_1, \ldots, k_{d+1}) \in \mathbb{N}^{d+1} \mid \sum_{i=1}^{d+1} k_i \alpha_i = n \right\} \right| =: \text{VP}_{\alpha_1, \ldots, \alpha_{d+1}}(n),
\]

for any \( n \in \mathbb{Z} \). Here, \( \text{VP}_{\alpha_1, \ldots, \alpha_{d+1}} \) is the so-called vector partition function (or denumerant function) with respect to the \( \alpha_i \). The problem of determining the zero set of \( \text{VP}_{\alpha_1, \ldots, \alpha_{d+1}} \) (or the maximum of this set) is quite famously known as the diophantine Frobenius problem. This problem is hard in general (though not necessarily so in specific cases) and there does not exist a general closed expression to determine the zero set (for a survey of the diophantine Frobenius problem we refer to the book [Ram05]). Analogously, one can write down similar functions for any circuit among the \( l_i \) (see subsection 4.1).

The basic idea now is to transport the discriminantal arrangement from \( A_{\mathbb{Q}} \) to some diophantine analog in \( A_{d-1}(X) \). For any circuit \( C \subset [n] \) there is a short exact sequence

\[
0 \rightarrow H_C \rightarrow A_{\mathbb{Q}} \rightarrow A_C \rightarrow 0.
\]

By lifting the surjection \( A_{\mathbb{Q}} \rightarrow A_C \) to its integral counterpart \( A_{d-1}(X) \rightarrow A_C \), we lift the zero set of the corresponding vector partition function on \( A_{\mathbb{Q}} \) to \( A_{d-1}(X) \). By doing this for every circuit \( C \), we construct in \( A_{d-1}(X) \) what we call the Frobenius discriminantal arrangement. One can consider the Frobenius discriminantal arrangement as an arithmetic thickening of the discriminantal arrangement. This thickening in general is just enough to enlarge the relevant strata in the discriminantal arrangement such that it encompasses the Kawamata-Viehweg-like theorems. To derive other vanishing results, our analysis will mostly be concerned with analyzing the birational geometry of \( X \) and its implications on the combinatorics of the discriminantal arrangement, and the transport of this analysis to the Frobenius arrangement.
OVERVIEW. Section 2 introduce some general notation and results related to toric varieties. In section 3 we survey discriminantal arrangements, secondary fans, and rational aspects of cohomology vanishing. Several technical facts will be collected which are important for the subsequent sections. Section 4 contains all the essential results of this work. In 4.3 we will prove our main arithmetic vanishing results. These will be applied in 4.4 to give a quite complete characterization of cohomology vanishing for toric surfaces. Section 4.5 is devoted to maximal Cohen-Macaulay modules.

Acknowledgments. Thanks to Laurent Bonavero, Michel Brion, Lutz Hille, Vic Reiner, and Jan Stienstra for discussion and useful hints.

2 Toric Preliminaries

In this section we first introduce notions from toric geometry which will be used throughout the rest of the paper. As general reference for toric varieties we use [Oda88], [Ful93]. We will always work over an algebraically closed field $k$.

Let $\Delta$ be a fan in the rational vector space $N_\mathbb{Q} := N \otimes \mathbb{Z} \otimes \mathbb{Q}$ over a lattice $N \cong \mathbb{Z}^d$. Let $M$ be the lattice dual to $N$, then the elements of $N$ represent linear forms on $M$ and we write $n(m)$ for the canonical pairing $N \times M \to \mathbb{Z}$, where $n \in N$ and $m \in M$. This pairing extends naturally over $M \otimes \mathbb{Q}$, $\mathbb{Q}$.

Elements of $M$ are denoted by $m$, $m'$, etc. if written additively, and by $\chi(m)$, $\chi(m')$, etc. if written multiplicatively, i.e. $\chi(m + m') = \chi(m) \chi(m')$.

The lattice $M$ is identified with the group of characters of the algebraic torus $T = \text{Hom}(M, k^*) \cong (k^*)^d$ which acts on the toric variety $X = X_\Delta$ associated to $\Delta$. Moreover, we will use the following notation:

- cones in $\Delta$ are denoted by small greek letters $\rho, \sigma, \tau, \ldots$, their natural partial order by $\prec$, i.e. $\rho \prec \tau$ if $\rho \subseteq \tau$;
- $|\Delta| := \bigcup_{\sigma \in \Delta} \sigma$ denotes the support of $\Delta$;
- for $0 \leq i \leq d$ we denote $\Delta(i) \subset \Delta$ the set of $i$-dimensional cones; for $\sigma \in \Delta$, we denote $\sigma(i)$ the set of $i$-dimensional faces of $\sigma$;
- $U_\sigma$ denotes the associated affine toric variety for any $\sigma \in \Delta$;
- $\sigma := \{m \in M_\mathbb{Q} \mid n(m) \geq 0 \text{ for all } n \in \sigma\}$ is the cone dual to $\sigma$;
- $\sigma^\perp = \{m \in M_\mathbb{Q} \mid n(m) = 0 \text{ for all } n \in \sigma\}$;
- $\sigma_M := \sigma \cap M$ is the submonoid of $M$ associated to $\sigma$.

We will mostly be interested in the structure of $\Delta$ as a combinatorial cellular complex. For this, we make a few convenient identifications. We always denote $n$ the cardinality of $\Delta(1)$. i.e. the number of 1-dimensional cones (rays) and...
[n] := \{1, \ldots, n\}. The primitive vectors along rays are denoted \(l_1, \ldots, l_n\), and, by abuse of notion, we will usually identify the sets \(\Delta(1)\), the set of primitive vectors, and \([n]\). Also, we will often identify \(\sigma \in \Delta\) with the set \(\sigma(1) \subset [n]\). With these identifications, and using the natural order of \([n]\), we obtain a combinatorial cellular complex with support \([n]\); we may consider this complex as a combinatorial model for \(\Delta\). In the case where \(\Delta\) is simplicial, this complex is just a combinatorial simplicial complex in the usual sense. If \(\Delta\) is not simplicial, we consider the simplicial cover \(\hat{\Delta}\) of \(\Delta\), modelled on \([n]\): some subset \(I \subset [n]\) is in \(\hat{\Delta}\) iff there exists some \(\sigma \in \Delta\) such that \(I \subset \sigma(1)\). The identity on \([n]\) then induces a surjective morphism \(\Delta \to \hat{\Delta}\) of combinatorial cellular complexes. This morphism has a natural representation in terms of fans. We can identify \(\hat{\Delta}\) with the fan in \(\mathbb{Q}^n\) which is defined as follows. Let \(e_1, \ldots, e_n\) be the standard basis of \(\mathbb{Q}^n\), then for any set \(I \subset [n]\), the vectors \(\{e_i\}_{i \in I}\) span a cone over \(\mathbb{Q}_{\geq 0}\) if and only if there exists \(\sigma \in \Delta\) with \(I \subset \sigma(1)\). The associated toric variety \(\hat{X}\) is open in \(\mathbb{A}^n\), and the vector space homomorphism defined by mapping \(e_i \mapsto l_i\) for \(i \in [n]\) induces a map of fans \(\Delta \to \hat{\Delta}\). The induced morphism \(\hat{X} \to X\) is the quotient presentation due to Cox [Cox95]. We will not make explicit use of this construction, but it may be useful to have it in mind.

An important fact used throughout this work is the following exact sequence which exists for any toric variety \(X\) with associated fan \(\Delta\):

\[
M \xrightarrow{L} \mathbb{Z}^n \to A_{d-1}(X) \to 0.
\]  

Here \(L(m) = (l_1(m), \ldots, l_n(m))\), i.e. as a matrix, the primitive vectors \(l_i\) represent the row vectors of \(L\). Note that \(L\) is injective iff \(\Delta\) is not contained in a proper subspace of \(\mathbb{N}\). The sequence follows from the fact that every Weil divisor \(D\) on \(X\) is rationally equivalent to a \(T\)-invariant Weil divisor, i.e. \(D \sim \sum_{i=1}^n c_i D_i\), where \(c = (c_1, \ldots, c_n) \in \mathbb{Z}^n\) and \(D_1, \ldots, D_n\), the \(T\)-invariant irreducible divisors of \(X\). Moreover, any two \(T\)-invariant divisors \(D, D'\) are rationally equivalent if and only if there exists \(m \in M\) such that \(D - D' = \sum_{i=1}^n l_i(m)D_i\). To every Weil divisor \(D\), one associates its divisorial sheaf \(\mathcal{O}_X(D) = \mathcal{O}(D)\) (we will omit the subscript \(X\) whenever there is no ambiguity), which is a reflexive sheaf of rank one and locally free if and only if \(D\) is Cartier. Rational equivalence classes of Weil divisors are in bijection with isomorphism classes of divisorial sheaves. If \(D\) is \(T\)-invariant, the sheaf \(\mathcal{O}(D)\) acquires a \(T\)-equivariant structure and the equivariant isomorphism classes of sheaves \(\mathcal{O}(D)\) are one-to-one with \(\mathbb{Z}^n\).

Consider a closed \(T\)-invariant subscheme \(V \subseteq X\). Then for any \(T\)-invariant Weil divisor \(D\) there are induced linear representations of \(T\) on the local cohomology groups \(H^*_T(X, \mathcal{O}(D))\). In particular, each such module has a natural eigenspace decomposition

\[
H^*_T(X, \mathcal{O}(D)) \cong \bigoplus_{m \in M} H^*_T(X, \mathcal{O}(D))_m.
\]

The eigenspaces \(H^*_T(X, \mathcal{O}(D))_m\) can be characterized by the relative cohomologies of certain simplicial complexes. For any \(I \subset [n]\) we denote \(\hat{\Delta}_I\) the maximal
subcomplex of $\hat{\Delta}$ which is supported on $I$. We denote $\hat{\Delta}_V$ the simplicial cover of the fan associated to the complement of the reduced subscheme underlying $V$ in $X$. Correspondingly, for $I \subset [n]$ we denote $\hat{\Delta}_{V,I}$ the maximal subcomplex of $\hat{\Delta}_V$ which is supported on $I$. If $c \in \mathbb{Z}^n$ is fixed, and $D = \sum_{i \in [n]} c_i D_i$, then every $m \in M$ determines a subset $I(m)$ of $[n]$ which is given by

$$I(m) = \{i \in [n] \mid l_i(m) < -c_i\}.$$ 

Then we will write $\hat{\Delta}_m$ and $\hat{\Delta}_{V,m}$ instead of $\hat{\Delta}_{I(m)}$ and $\hat{\Delta}_{V,I(m)}$, respectively. In the case where $\Delta$ is generated by just one cone $\sigma$, we will also write $\hat{\sigma}_m$, etc.

With respect to these notions we get:

**Theorem 2.1:** Let $D \in \mathbb{Z}^{\Delta(1)}$ be a $T$-invariant Weil divisor on $X$. Then for every $T$-invariant closed subscheme $V$ of $X$, every $i \geq 0$ and every $m \in M$ there exists an isomorphism of $k$-vector spaces

$$H^i_{\mathcal{V}}(X, \mathcal{O}(D))_m \cong H^{i-1}(\hat{\Delta}_m, \hat{\Delta}_{V,m}; k).$$

Note that here $H^{i-1}(\hat{\Delta}_m, \hat{\Delta}_{V,m})$ denotes the reduced relative cohomology group of the pair $(\hat{\Delta}_m, \hat{\Delta}_{V,m})$.

**Proof.** For $V = X$ it follows from [EMS00], §2 that $H^i(X, \mathcal{O}(D))_m \cong H^{i-1}(\hat{\Delta}_m; k)$ and $H^i(X \setminus V, \mathcal{O}(D))_m \cong H^{i-1}(\hat{\Delta}_{V,m}; k)$. Then the assertion follows from comparing the long exact relative cohomology sequence of the pair $(\hat{\Delta}_m, \hat{\Delta}_{V,m})$ with the long exact local cohomology sequence with respect to $X$ and $V$ in degree $m$. \qed

We mention a special case of this theorem, which follows from the long exact cohomology sequence.

**Corollary 2.2:** Let $X = U_\sigma$ and $V$ a $T$-invariant closed subvariety of $X$ and denote $\hat{\sigma}$ the simplicial model for the fan generated by $\sigma$. Then for every $m \in M$ and every $i \in \mathbb{Z}$:

$$H^i_{\mathcal{V}}(X, \mathcal{O}(D))_m = \begin{cases} 0 & \text{if } \hat{\sigma}_m = \emptyset, \\ H^{i-2}(\hat{\sigma}_{V,m}; k) & \text{else.} \end{cases}$$

### 3 Discriminants and combinatorial aspects

Cohomology vanishing

A toric variety $X$ is specified by the set of primitive vectors $l_1, \ldots, l_n \in \mathbb{N}$ and the fan $\Delta$ supported on these vectors. We can separate three properties which govern the geometry of $X$ and are relevant for cohomology vanishing problems:

(i) the linear algebra given by the vectors $l_1, \ldots, l_n$ and their linear dependencies as $\mathbb{Q}$-vectors;
(ii) arithmetic properties, which are also determined by the $l_i$, but considered
as integral vectors;

(iii) its combinatorics, which is given by the fan $\Delta$.

In this section we will have a closer look into the linear algebraic and com-
binatorial aspects. In subsection 3.1 we will introduce the notion of oriented
and non-oriented circuits associated to the vectors $l_i$. In subsection 3.2 we
consider circuits of the matrix $L$ and the induced stratification of $A_{d-1}(X)_\mathbb{Q}$.
In subsection 3.3 we will collect some well-known material on secondary fans
from [GKZ94], [OP91], and [BFS90] and explain their relation to discriminant-
al arrangements. Subsection 3.4 then applies this to certain statements about
the birational geometry of toric varieties and cohomology vanishing.

For this section and the following sections we will introduce the followin-
g conventions.

**Convention 3.1:** We will denote $L$ the matrix whose rows are given by the
$l_i$. For any subset $I$ of $[n]$ we will denote $L_I$ the submatrix of $L$ consis-
ting of the rows which are given by the $l_i$ with $i \in I$. In general, we will not distinguish
between \{l_i\}$_{i \in I}$ and $L_I$. Similarly, we will usually identify subsets $I \subset [n]$ with
the corresponding subsets of \{l_1,\ldots,l_n\}. If $\Delta$ is a fan in $\mathbb{N}_\mathbb{R}$ such that $\Delta(1)$ is
generated by some subset of the $l_i$, then we say that $\Delta$ is
supported on $L$ (resp. on $l_1,\ldots,l_n$).

Let $C$ be a subset of $[n]$ which is minimal with the property that the $l_i$ with
$i \in C$ are linearly dependent. Then the set \{l_i\}$_{i \in C}$ is called a circuit. By abuse
of notion we will also call $C$ itself a circuit.

3.1 Circuits

Let $C \subseteq [n]$ be a circuit. Then we have a relation

$$\sum_{i \in [n]} \alpha_i l_i = 0,$$

which is unique up to a common multiple of the $\alpha_i$, and the $\alpha_i$ are nonzero.
Without loss of generality, we will assume that the $\alpha_i$ are integral and
$\gcd\{|\alpha_i|\}_{i \in [n]} = 1$. To simplify the discussion, we will further assume that
$L_C$ generates a submodule $N_C$ of finite index in $N$. For a fixed choice of the
$\alpha_i$, we have a partition $C = C^+ \bigsqcup C^-$, where $C^\pm = \{i \in [n] \mid \pm \alpha_i > 0\}$. This
decomposition depends only on the signs of the $\alpha_i$; flipping the signs exchanges
$C^+$ and $C^-$. We want to keep track of these two possibilities and call the choice
of $C^+ \bigsqcup C^-$ the oriented circuit with underlying circuit $C$ (or simply an orientation of $C$), and $-C := -C^+ \bigsqcup -C^-$ its inverse, where $-C^\pm := C^\mp$.

**Definition 3.2:** We denote $C(L)$ the set of circuits of $L$ and $\mathcal{C}(L)$ the set of
oriented circuits of $L$, i.e. the set of all orientations $C$, $-C$ for $C \in C(L)$. 

**Documenta Mathematica 16 (2011) 209–251**
For a given circuit $C$, the primitive vectors $L_C$ can support at most two simplicial fans, each corresponding to an orientation of $C$. For fixed orientation $\mathcal{C}$, we denote $\Delta = \Delta_{\mathcal{C}}$ the fan whose maximal cones are generated by $\mathcal{C} \setminus \{i\}$, where $i$ runs over the elements of $\mathcal{C}^+$. The only exception for this procedure is the case where $\mathcal{C}^+$ is empty, which we leave undefined. The associated toric variety $X_{\Delta}$ is simplicial and quasi-projective.

**Definition 3.3:** We call a toric variety $X = X_{\Delta}$ associated to an oriented circuit $\mathcal{C}$ a *toric 1-circuit variety*.

Now let us assume that the sublattice $N_C$ of $N$ which is generated by $L_C$ is saturated. Then we have a short exact sequence

$$0 \to M \xrightarrow{L_C} \mathbb{Z}^n \xrightarrow{G_C} A \to 0,$$

such that $A \cong \mathbb{Z}$ and thus torsion free. Here, $L_C$ is considered as a tuple of linear forms on $M$, $A \cong \mathbb{Z}$ and $G_C = (\alpha_1, \ldots, \alpha_n)$ is a $(1 \times n)$-matrix, i.e. we can consider the $\alpha_i$ as the *Gale transform* of the $l_i$. Conversely, if the $\alpha_i$ are given, then the $l_i$ are determined up to a $\mathbb{Z}$-linear automorphism of $N$. We will make more extensive use of the Gale transform later on. For generalities we refer to [OP91] and [GKZ94].

If $N_C \subseteq N$, we can formally consider the inclusion of $N_C$ as the image of $N$ via an injective endomorphism $\xi$ of $N$. The inverse images of the $l_i$ with respect to $\xi$ satisfy the same relation as the $l_i$. Therefore, a general toric circuit variety is completely specified by $\xi$ and the integers $\alpha_i$. More precisely, a toric 1-circuit variety is specified by the Gale duals $l_i$ of the $\alpha_i$ and an injective endomorphism $\xi$ of $N$ with the property that $\xi(l_i)$ is primitive in $N$ for every $i \in [n]$. 

**Definition 3.4:** Let $\alpha = (\alpha_i \mid i \in \mathcal{C}) \in \mathbb{Z}^\mathcal{C}$ with $\alpha_i \neq 0$ for every $i$ and $\gcd(|\alpha_i|)_{i \in [n]} = 1$, $\mathcal{C}$ the associated oriented circuit with $\mathcal{C}^+ = \{i \mid \alpha_i > 0\}$, and $\xi : N \to N$ an injective endomorphism of $N$ which maps the Gale duals of the $\alpha_i$ to primitive elements $p_i$ in $N$. Then we denote $\mathbb{P}(\alpha, \xi)$ the toric 1-circuit variety associated to the fan $\Delta_{\mathcal{C}}$ spanned by the primitive vectors $p_i$.

The endomorphism $\xi$ translates into an isomorphism

$$\mathbb{P}(\alpha, \xi) \cong \mathbb{P}(\alpha, \text{id}_N)/H,$$

where $H \cong \text{spec} \mathbb{k}[N/N_C]$. Note that in positive characteristic, $H$ in general is a group scheme rather than a proper algebraic group. Moreover, in sequence (4) we can identify $A$ with the divisor class group $A_d-1(\mathbb{P}(\alpha, \text{id}_N))$. Similarly, we get $A_{d-1}(\mathbb{P}(\alpha, \xi)) \cong A \otimes N/N_C$ and the natural surjection from $A_{d-1}(\mathbb{P}(\alpha, \xi))$ onto $A_{d-1}(\mathbb{P}(\alpha, \text{id}_N))$ just projects away the torsion part.

**Remarks 3.5:** (i) In the case $\alpha_i > 0$ for all $i$ and $\xi = \text{id}_N$, we just recover the usual weighted projective spaces. In many respects, the spaces $\mathbb{P}(\alpha, \xi)$ can be treated the same way as has been done in the standard references for weighted...
projective spaces, see [Del75], [Dol82], [BR86]. In our setting there is the slight
simplification that we naturally can assume that \( \gcd\{\alpha_j\}_{j \neq i} = 1 \) for every
\( i \in [n] \), which eliminates the need to discuss reduced weights.

(ii) In the case that \( L \) spans a subspace \( N' \) of \( N_\mathbb{Q} \) of positive codimension \( r \),
then for some orientation \( C \) of \( C \) the variety \( X(\Delta_\mathbb{C}) \) is isomorphic to \( \mathbb{P}(\vec{a}, \xi) \times (k^*)^r \), where \( \mathbb{P}(\vec{a}, \xi) \) is defined as before with respect to \( N' \). Note that if
\( C^+ = C \), then the fan \( \Delta_\mathbb{C} \) is empty. By convention, in that case one can define
\( X(\Delta_\mathbb{C}) := (k^*)^r \) as the associated toric variety.

(iii) The spaces \( \mathbb{P}(\vec{a}, \xi) \) are building blocks for the birational geometry of general
toric varieties. In fact, to every extremal curve \( V(\tau) \) in some simplicial toric
variety \( X \), there is associated some variety \( \mathbb{P}(\vec{a}, \xi) \) whose fan \( \Delta \) is a subfan of
\( \Delta \) and which embeds as an open invariant subvariety of \( X \). If \( |C^+| \notin \{n, n-1\} \),
the primitive vectors \( l_i \) span a convex polyhedral cone, giving rise to an affine
toric variety \( Y \) and a canonical morphism \( \pi : \mathbb{P}(\vec{a}, \xi) \rightarrow Y \) which is a partial
resolution of singularities. Sign change \( \vec{a} \rightarrow -\vec{a} \) then encodes the transition
from \( C \) to \( -C \) and a birational map from \( \mathbb{P}(\vec{a}, \xi) \) to \( \mathbb{P}(\vec{a}, \xi) \), which provides
a local model for well-known combinatorial operation which called bistellar
operation [Rei99] or modification of a triangulation [GKZ94]:

\[
\begin{array}{ccc}
\mathbb{P}(\vec{a}, \xi) & \xrightarrow{\pi} & \mathbb{P}(\vec{a}, \xi) \\
Y & \xleftarrow{\pi'} & Y
\end{array}
\]

(for \( |C^+| = d - 1 \), one can identify \( \mathbb{P}(\vec{a}, \xi) \) with \( Y \) and one just obtains a
blow-down).

3.2 Circuits and discriminantal arrangements

Recall that for any torus invariant divisor \( D = \sum_{i \in [n]} c_i D_i \), the isotypical
components \( H^i_1(X, O(D))_m \) for some cohomology group depend on simplicial complexes \( \tilde{\Delta}_I \), where \( I = I(m) = \{ i \in [n] \mid l_i(m) < -c_i \} \). So, the set of
all possible subcomplexes \( \tilde{\Delta}_I \) depends on the chamber decomposition of \( M_\mathbb{Q} \)
which is induced by the hyperplane arrangement which is given by hyperplanes
\( H^*_1, \ldots, H^*_n \), where
\[
H^*_i := \{ m \in M_\mathbb{Q} \mid l_i(m) = -c_i \}.
\]
The set of all relevant \( I \subset [n] \) is determined by the map
\[
\varphi : M_\mathbb{Q} \rightarrow 2^{[n]}, \quad m \mapsto \{ i \in [n] \mid l_i(m) < -c_i \}.
\]

**Definition 3.6:** For \( m \in M_\mathbb{Q} \), we call \( \varphi \) the *signature of* \( m \). We call the image of \( M_\mathbb{Q} \) in \( 2^{[n]} \) the *combinatorial type* of \( \varphi \).

**Remark 3.7:** The combinatorial type encodes what in combinatorics is known as *oriented matroid* (see [BLS+93]). We will not make use of this kind of structure, but we will find it sometimes convenient to borrow some notions.
So, given $l_1, \ldots, l_n$, we would like to classify all possible combinatorial types, depending on $c \in \mathbb{Q}^n$. The natural parameter space for all hyperplane arrangements up to translation by some element $m \in M_\mathbb{Q}$ is given by the set $A_\mathbb{Q} \cong \mathbb{Q}^n/M_\mathbb{Q}$, which is given by the following short exact sequence:

$$0 \longrightarrow M_\mathbb{Q} \overset{L}{\longrightarrow} \mathbb{Q}^n \overset{D}{\longrightarrow} A_{d-1}(X)_\mathbb{Q} = A_\mathbb{Q} \longrightarrow 0.$$  

Then the $D_1, \ldots, D_n$ are the images of the standard basis vectors of $\mathbb{Q}^n$. This procedure of constructing the $D_i$ from the $l_i$ is often called Gale transformation, and the $D_i$ are the Gale duals of the $l_i$.

Now, a hyperplane arrangement $H_c$ for some $c \in \mathbb{Q}^n$, is considered in general position if the hyperplanes $H_c$ intersect in the smallest possible dimension. When varying $c$ and passing from one arrangement in general position to another one which has a different combinatorial type, this necessarily implies that has to take place some specialization for some $c \in \mathbb{Q}^n$, i.e. where the corresponding hyperplanes $H_c^\sharp$ do not intersect in the smallest possible dimension. So we see that the combinatorial types of hyperplane arrangements with fixed $L$ and varying induce a stratification of $A_\mathbb{Q}$, where the maximal strata correspond to hyperplane arrangements in general position. To determine this stratification is the discriminant problem for hyperplane arrangements. To be more precise, let $I \subseteq [n]$ and denote

$$H_I := \{c + M_\mathbb{Q} \in A_\mathbb{Q} \mid \bigcap_{i \in I} H_c^\sharp \neq \emptyset\},$$

i.e. $H_I$ represents the set of all hyperplane arrangements (up to translation) such that the hyperplanes $\{H_c^\sharp\}_{c \in I}$ have nonempty intersection. The sets $H_I$ can be described straightforwardly by the following commutative exact diagram:

$$
\begin{array}{ccc}
0 & \longrightarrow & M_\mathbb{Q} \\
\uparrow & & \uparrow \\
M_\mathbb{Q} & \overset{L_I}{\longrightarrow} & \mathbb{Q}^I \\
\downarrow & & \downarrow \\
0 & \longrightarrow & A_I. \\
\end{array}
$$

In particular, $H_I$ is a subvector space of $A_\mathbb{Q}$. Moreover, we immediately read off diagram (5):

**Lemma 3.8:** (i) $H_I$ is generated by the $D_i$ with $i \in [n] \setminus I$.

(ii) $\dim H_I = n - |I| - \dim(\ker L_I)$.

(iii) If $J \subseteq I$ then $H_I \subseteq H_J$.

(iv) Let $I, J \subseteq [n]$, then $H_{I \cup J} \subset H_I \cap H_J$. 

Documenta Mathematica 16 (2011) 209–251
Note that in (iv) the reverse inclusion in general is not true. It follows that the hyperplanes among the $H_I$ are determined by the formula:

$$|I| = \text{rk } L_I + 1.$$ 

By Lemma 3.8 (iii), we can always consider circuits fulfilling this condition. It turns out that the hyperplane $H_C$ suffice to completely describe the discriminants of $L$:

**Lemma 3.9:** Let $I \subset [n]$, then

$$H_I = \bigcap_{C \subset I \text{ circuit}} H_C,$$

where, by convention, the right hand side equals $A_Q$, if the $l_i$ with $i \in I$ are linearly independent.

Hence, the stratification of $A_Q$ which we were looking for is completely determined by the hyperplanes $H_C$.

**Definition 3.10:** We denote the set $\{ H_C \mid C \subset [n] \text{ a circuit} \}$ the discriminantal arrangement of $L$.

**Remark 3.11:** The discriminantal arrangement carries a natural matroid structure. This structure can be considered as another combinatorial invariant of $L$ (or the toric variety $X$, respectively), its circuit geometry. Discriminantal arrangements seem to have been appeared first in [Cra84], where the notion of 'circuit geometry' was coined. The notion of discriminantal arrangements stems from [MS89]. Otherwise, this subject seems to have been studied explicitly only in very few places, see for instance [Fal94], [BB97], [Ath99], [Rei99], [Coh86], [CV03], though it is at least implicit in the whole body of literature on secondary fans. Above references are mostly concerned with genericity properties of discriminantal arrangements. Unfortunately, in toric geometry, the most interesting cases (such as smooth projective toric varieties, for example) virtually never give rise to discriminantal arrangements in general position. Instead, we will focus on certain properties of nongeneric circuit geometries, though we will not undertake a thorough combinatorial study of these.

Virtually all problems related to cohomology vanishing on a toric variety $X$ must depend on the associated discriminantal arrangement and therefore on the circuits of $L$. In subsection 3.3 we will see that the discriminantal arrangement is tightly tied to the geometry of $X$.

As we have seen in section 3.1, to every circuit $C \subset [n]$ we can associate two oriented circuits. These correspond to the signature of the bounded chamber of the subarrangement in $M_Q$ given by the $H^+_i$ with $i \in C$ (or better to the bounded chamber in $M_Q/\ker L_I$, as we do no longer require that the $l_i$ with $i \in C$ span $M_Q$). Lifting this to $A_Q$, this corresponds to the half spaces in $A_Q$ which are bounded by $H_C$.
**Definition 3.12**: Let \( C \subset [n] \) be a circuit, then we denote \( H_C \) the half space in \( \mathbb{A}_Q \) bounded by \( H_C \) corresponding to the orientation \( \mathcal{C} \).

The following is straightforward to check:

**Lemma 3.13**: Let \( C \) be a circuit of \( L \) and \( \mathcal{C} \) an orientation of \( C \). Then the hyperplane \( H_C \) is separating, i.e. for every \( i \in [n] \) one of the following holds:

(i) \( i \in [n] \setminus C \) iff \( D_i \in H_C \);

(ii) if \( i \in \mathcal{C}^+ \), then \( D_i \in H_\mathcal{C} \setminus H_C \);

(iii) if \( i \in \mathcal{C}^- \), then \( D_i \in H_{-\mathcal{C}} \setminus H_C \).

Now we are going to borrow some terminology from combinatorics. Consider any subvector space \( U \subset \mathbb{A}_Q \) which is the intersection of some of the \( H_C \). Then the set \( \mathcal{F}_U \) of all \( C \in \mathcal{C}(L) \) such that \( H_C \) contains \( U \) is called a flat. The subvector space is uniquely determined by the flat and vice versa. We can do the same for the actual strata rather than for subvector spaces. For this, we just need to consider instead the oriented circuits and their associated half spaces in \( \mathbb{A}_Q \): any stratum \( S \) of the discriminantal arrangement uniquely determines a finite set \( \mathfrak{F}_S \) of oriented circuits \( \mathcal{C} \) such that \( S \subset H_\mathcal{C} \). From the set \( \mathfrak{F}_S \) we can reconstruct the closure of \( S \):

\[
\mathfrak{S} = \bigcap_{\mathcal{C} \in \mathfrak{F}_S} H_\mathcal{C},
\]

We give a formal definition:

**Definition 3.14**: For any subvector space \( U \subset \mathbb{A}_Q \) which is a union of strata of the discriminantal arrangement, we denote \( \mathcal{F}_U := \{ C \in \mathcal{C}(L) \mid U \subset H_C \} \) the associated flat. For any single stratum \( S \subset \mathbb{A}_Q \) of the discriminantal arrangement, we denote \( \mathfrak{F}_S := \{ \mathcal{C} \in \mathcal{C}(L) \mid U \subset H_\mathcal{C} \} \) the associated oriented flat.

The notion of flats gives us some flexibility in handling strata. Note that flats reverse inclusions, i.e. \( S \subset T \) iff \( \mathfrak{F}_T \subset \mathfrak{F}_S \). Moreover, if a stratum \( S \) is contained in some \( H_C \), then its oriented flat contains both \( H_\mathcal{C} \) and \( H_{-\mathcal{C}} \), and vice versa. So from the oriented flat we can reconstruct the subvector space of \( \mathbb{A}_Q \) generated by \( S \).

**Definition 3.15**: Let \( \mathcal{S} := \{ S_1, \ldots, S_k \} \) be a collection of strata of the discriminantal arrangement. We call

\[
\mathfrak{F}_\mathcal{S} := \bigcap_{i=1}^k \mathfrak{F}_{S_i}
\]

the **discriminantal hull** of \( \mathcal{S} \).
The discriminantal hull defines a closed cone in $A_\mathbb{Q}$ which is given by the intersection $\bigcap_{c \in \mathcal{S}} H_c$. This cone contains the union of the closures $\overline{S}_i$, but is bigger in general.

**Lemma 3.16:**

(i) Let $\mathcal{S} = \{S_1, \ldots, S_k\}$ be a collection of discriminantal strata whose union is a closed cone in $A_\mathbb{Q}$. Then $\mathcal{S} = \bigcap_{i=1}^k S_i$.

(ii) Let $\mathcal{S} = \{S_1, \ldots, S_k\}$ be a collection of discriminantal strata and $U$ the subvector space of $A_\mathbb{Q}$ generated by the $S_i$. Then the forgetful map $\mathcal{S} \rightarrow \mathcal{F}_U$ is surjective iff $\mathcal{S} = S_i$ for some $i$.

**Proof.** For (i) just note that because $\bigcup_{i=1}^k S_i$ is a closed cone, it must be an intersection of some $H_c$. For (ii): the set $\bigcap_{c \in \mathcal{S}} H_c$ is a cone which contains the convex hull of all the $S_i$. If some $C$ is not in the image of the forgetful map, then the hyperplane $H_C$ must intersect the relative interior of this cone. So the assertion follows.

### 3.3 Secondary Fans

For any $c \in \mathbb{Q}^n$ the arrangement $H_c$ induces a chamber decomposition of $M_\mathbb{Q}$, where the closures of the chambers are given by

$$P^I_c := \{m \in M_\mathbb{Q} \mid l_i(m) \leq -c_i \text{ for } i \in I \text{ and } l_i(m) \geq -c_i \text{ for } i \notin I\}$$

for every $I \subset [n]$ which belongs to the combinatorial type of $c$. In particular, $c$ represents an element $D \in A_\mathbb{Q}$ with

$$D \in \bigcap_{I \in \Sigma(M_\mathbb{Q})} C_I,$$

where $C_I$ is the cone in $A_\mathbb{Q}$ which is generated by the $-D_i$ for $i \in I$ and the $D_i$ with $i \notin I$ for some $I \subset [n]$. For an invariant divisor $D = \sum_{i \in [n]} c_i D_i$ we will also write $P^I_c$ instead of $P^I_c$. If $I = \emptyset$, we will occasionally omit the index $I$.

The faces of the $C_I$ can be read off directly from the signature:

**Proposition 3.17:** Let $I \subset [n]$, then $C_I$ is an nonredundant intersection of the $H_c$ with $\mathcal{C}^- \subset I$ and $\mathcal{C}^+ \cap I = \emptyset$.

**Proof.** First of all, it is clear that $C_I$ coincides with the intersection of half spaces

$$C_I = \bigcap_{\mathcal{C}^- \subset I \mathcal{C}^+ \cap I = \emptyset} H_c.$$

For any $H_c$ in the intersection let $H_c$ its boundary. Then $H_c$ contains a cone of codimension 1 in $A_\mathbb{Q}$ which is spanned by $D_i$ with $i \in [n] \setminus (\mathcal{C} \cup I)$ and by $-D_i$ with $i \in I \setminus \mathcal{C}$ which thus forms a proper facet of $C_I$.
Recall that the secondary fan of \( L \) is a fan in \( A_Q \) whose maximal cones are in one-to-one correspondence with the regular simplicial fans which are supported on the \( l_i \). That is, if \( \mathbf{c} \) is chosen sufficiently general, then the polyhedron \( P^\mathbf{c}_L \) is simplicial and its inner normal fan is a simplicial fan which is supported on the \( l_i \). Wall crossing in the secondary fan then corresponds locally to a transition \( \Delta_{\mathbf{c}} \to \Delta_{-\mathbf{c}} \) as in section 3.1. Clearly, the secondary fan is a substructure of the discriminantal arrangement in the sense that its cones are unions of strata of the discriminantal arrangements. However, the secondary fan in general is much coarser than the discriminantal arrangement, as it only keeps track of the particular chamber \( P^\emptyset_L \). In particular, the secondary fan is only supported on \( C^\emptyset \) which in general does not coincide with \( A_Q \). Of course, there is no reason to consider only one particular type of chamber — we can consider secondary fans for every \( I \subset [n] \) and every type of chamber \( P^I_L \). For this, observe first that, if \( \mathcal{B} \) is a subset of \([n]\) such that the \( l_i \) with \( i \in \mathcal{B} \) form a basis of \( M_Q \), then the complementary Gale duals \( \{D_i\}_{i \in \mathcal{B}} \) form a basis of \( A_Q \). Then we set:

**Definition 3.18:** Let \( I \subset [n] \) and \( \mathcal{B} \subset [n] \) such that the \( l_i \) with \( i \in \mathcal{B} \) form a basis of \( M_Q \), then we denote \( K^\emptyset_B^I \) the cone in \( A_Q \) which is generated by \( -D_i \) for \( i \in I \setminus \mathcal{B} \) and by \( D_i \) for \( i \in [n] \setminus (I \cup \mathcal{B}) \). The secondary fan \( \text{SF}(L, I) \) of \( L \) with respect to \( I \) is the fan whose cones are the intersections of the \( K^\emptyset_B^I \), where \( \mathcal{B} \) runs over all bases of \( L \).

Note that \( \text{SF}(L, \emptyset) \) is just the secondary fan as usually defined. Clearly, the chamber structure of the discriminantal arrangement still refines the chamber structure induced by all secondary arrangements. But now we have sufficient data to even get equality:

**Proposition 3.19:** The following induce identical chamber decompositions of \( A_Q \):

(i) the discriminantal arrangement,

(ii) the intersection of all secondary fans \( \text{SF}(L, I) \),

(iii) the intersection of the \( C_I \) for all \( I \subset [n] \).

**Proof.** Clearly, the facets of every orthant \( C_I \) span a hyperplane which is part of the discriminantal arrangement, so the chamber decomposition induced by the secondary fan is a refinement of the intersection of the \( C_I \)’s. The \( C_I \) induce a refinement of the secondary fans as follows. Without loss of generality, it suffices to show that every \( K^\emptyset_B^I \) is the intersection of some \( C_I \). We have

\[
K^\emptyset_B^I \subseteq \bigcap_{I \subseteq \mathcal{B}} C_I.
\]

On the other hand, for every facet of \( K^\emptyset_B^I \), we choose \( I \) such that \( C_I \) shares this face and \( K^\emptyset_B^I \) is contained in \( C_I \). This can always be achieved by choosing \( I \) so
that every generator of $C_I$ is in the same half space as $K_B^0$. The intersection of these $C_I$ then is contained in $K_B^0$.

Now it remains to show that the intersection of the secondary fans refines the discriminantal arrangement. This actually follows from the fact, that for every hyperplane $H_C$, one can choose a minimal generating set which we can complete to a basis of $A_\mathbb{Q}$ from the $D_i$, where $i \notin C$. By varying the signs of this generating set, we always get a simplicial cone whose generators are contained in some secondary fan, and this way $H_C$ is covered by a set of facets of secondary cones.

The maximal cones in the secondary fan $SF(L, \emptyset)$ correspond to regular simplicial fans supported on $l_1, \ldots, l_n$. More precisely, if $\Delta$ denotes such a fan, then the corresponding cone is given by $\bigcap_B K_B^0$, where $B$ runs over all bases among the $l_i$ which span a maximal cone in $\Delta$. With respect to a simplicial model $\hat{\Delta}$ for $\Delta$, we define:

**Definition 3.20:** Let $\Delta$ be a fan supported on $L$, then we set:

$$ \text{nef}(\Delta) := \bigcap_{B \in \hat{\Delta} \text{ basis in } L} K_B^0 $$

and denote $\mathfrak{F}_{\text{nef}} = \mathfrak{F}_{\text{nef}(\Delta)}$ the discriminantal hull of $\text{nef}(\Delta)$.

Note that by our conventions we identify $B \in \hat{\Delta}$ with the set of corresponding primitive vectors, or the corresponding rows of $L$, respectively. Of course, $\text{nef}(\Delta)$ is just the nef cone of the toric variety associated to $\Delta$.

**Proposition 3.21:** We have:

$$ \text{nef}(\Delta) = \bigcap_{\Delta \cap (\Delta_\varepsilon)_{max} \neq \emptyset} H_\varepsilon. $$

**Proof.** For some basis $B \subset [n]$, the cone $K_B^0$ is simplicial, and for every $i \in [n] \setminus B$, the facet of $K_B^0$ which is spanned by the $D_j$ with $j \notin B \cup \{i\}$, spans a hyperplane $H_C$ in $P$. This hyperplane corresponds to the unique circuit $C \subset B \cup \{i\}$. As we have seen before, a maximal cone in $\Delta_\varepsilon$ is of the form $C \setminus \{j\}$ for some $j \in \varepsilon^+$. So we have immediately:

$$ K_B^0 = \bigcap_{\exists F \in (\Delta_\varepsilon)_{max} \text{ with } F \subset B} H_\varepsilon $$

and the assertion follows. 

**Remark 3.22:** If $\Delta = \hat{\Delta}$ is a regular simplicial fan, then $\text{nef}(\Delta)$ is a maximal cone in the secondary fan. Let $\varepsilon$ be an oriented circuit such that $\Delta$ is supported on $\Delta_\varepsilon$ in the sense of [GKZ94], §7, Def. 2.9, and denote $\Delta'$ the fan resulting in the bistellar operation by changing $\Delta_\varepsilon$ to $\Delta_{-\varepsilon}$. Then, by [GKZ94], §7, Thm. 2.10, the hyperplane $H_\varepsilon$ is a proper wall of $\text{nef}(\Delta)$ iff $\Delta'$ is regular, too.
3.4 MCM sheaves, \(\mathbb{Q}\)-Cartier divisors and the toric Kawamata-Viehweg vanishing theorem

Recall that a \(\mathbb{Q}\)-divisor on \(X\) is \(\mathbb{Q}\)-Cartier if an integral multiple is Cartier in the usual sense. A torus invariant Weil divisor \(D = \sum_{i \in [n]} c_i D_i\) is \(\mathbb{Q}\)-Cartier iff for every \(\sigma \in \Delta\) there exists some \(m_\sigma \in M_\mathbb{Q}\) such that \(c_i = l_i(m)\) for all \(i \in \sigma(1)\). A result of Bruns and Gubeladze [BG03] states that every toric \(\mathbb{Q}\)-Cartier divisor is maximal Cohen-Macaulay. The MCM property is useful, as it allows to replace the Ext-groups by cohomologies in Serre duality:

**Proposition 3.23:** Let \(X\) be a normal variety with dualizing sheaf \(\omega_X\) and \(F\) a coherent sheaf on \(X\) such that for every \(x \in X\), the stalk \(F_x\) is MCM over \(\mathcal{O}_{X,x}\). Then for every \(i \in \mathbb{Z}\) there exists an isomorphism

\[
\text{Ext}_X^i(F, \omega_X) \cong H^i(X, \text{Hom}(F, \omega_X)).
\]

**Proof.** For any two \(\mathcal{O}_X\)-modules \(F, G\) there exists the following spectral sequence

\[
E_2^{pq} = H^p(X, \text{Ext}^q_{\mathcal{O}_X}(F, G)) \Rightarrow \text{Ext}^{p+q}_{\mathcal{O}_X}(F, G).
\]

We apply this spectral sequence to the case \(G = \omega_X\). For every closed point \(x \in X\) we have the following identity of stalks:

\[
\text{Ext}^q_{\mathcal{O}_X}(F, \omega_X)_x \cong \text{Ext}^q_{\mathcal{O}_{X,x}}(F_x, \omega_{X,x}).
\]

As \(F\) is maximal Cohen-Macaulay, the latter vanishes for all \(q > 0\), and thus the sheaf \(\text{Ext}^q_{\mathcal{O}_X}(F, \omega_X)\) is the zero sheaf for all \(q > 0\). So the above spectral sequence degenerates and we obtain an isomorphism

\[
H^p(X, \text{Hom}(F, \omega_X)) \cong \text{Ext}_X^p(F, \omega_X)
\]

for every \(p \in \mathbb{Z}\).

In the case where \(X\) a toric variety, we have \(\omega_X \cong \mathcal{O}(K_X)\), where \(K_X = -\sum_{i \in [n]} D_i\). Then, if \(F = \mathcal{O}(D)\) for some \(D \in A\), we can identify \(\text{Hom}(\mathcal{O}(D), \omega_X)\) with \(\mathcal{O}(K_X - D)\):

**Corollary 3.24:** Let \(X\) be a toric variety and \(D\) a Weil divisor such that \(\mathcal{O}(D)\) is an MCM sheaf. Then there is an isomorphism:

\[
\text{Ext}_X^i(\mathcal{O}(D), \omega_X) \cong H^i(X, \mathcal{O}_X(K_X - D)).
\]

And by Grothendieck-Serre duality:

**Corollary 3.25:** If \(X\) is a complete toric variety and \(D\) a Weil divisor such that \(\mathcal{O}(D)\) is an MCM sheaf, then

\[
H^i(X, \mathcal{O}(D)) \cong H^{d-i}(X, \mathcal{O}(K_X - D))^*.
\]
For any Cartier divisor $D$ on some normal variety $X$ denote $N(X,D) := \{ k \in \mathbb{N} \mid H^0(X,\mathcal{O}(kD)) \neq 0 \}$. Then the Iitaka dimension of $D$ is defined as

$$\kappa(D) := \max_{k \in N(X,D)} \{ \dim \phi_k(X) \},$$

where $\phi_k : X \to \mathbb{P}[kD]$ is the family of morphisms given by the linear series $|kD|$.

In the case where $X$ is a toric variety and $D = \sum_{i \in [n]} c_i D_i$ invariant, the Iitaka dimension of $D$ is just the dimension of $P_{kD}$ for $k \gg 0$. For a $\mathbb{Q}$-Cartier divisor $D$, we define its Iitaka dimension by $\kappa(D) := \kappa(rD)$ for $r > 0$ such that $rD$ is Cartier.

If $D$ is a nef divisor, then $\phi : X \to \mathbb{P}[D]$ is torus equivariant, its image is a projective toric variety of dimension $\kappa(D)$ whose associated fan is the inner normal fan of $P_D$. If $\kappa(D) < d$, then necessarily $D$ is contained in some hyperplane $H_C$ such that $\mathcal{C}^+ = \mathcal{C}$ for some orientation $\mathcal{C}$ of $\mathcal{C}$. The toric variety $X_{\Delta_C}$ is isomorphic to a finite cover of a weighted projective space. This kind of circuit will play an important role later on, so that we will give it a distinguished name:

**Definition 3.26:** We call a circuit $\mathcal{C}$ such that $\mathcal{C} = \mathcal{C}^+$ for one of its orientations, fibrational. For $D \in A_{d-1}(X)_\mathbb{Q}$ we denote $\text{fib}(D) \subset \mathcal{C}(L)$ the set of fibrational circuits such that $D \in H_C$.

By Proposition 3.17, such a divisor $D$ is contained in the intersection of nef$(X)$ with the effective cone of $X$, which we identify with $C_\mathcal{Q}$. More precisely, it follows from linear algebra that $D$ is contained in all $H_C$ where $\mathcal{C}$ is fibrational and $l_i(P_D) = 0$ for all $i \in \mathcal{C}$.

The fibrational circuits of a nef divisor $D$ tell us immediately about its Iitaka dimension:

**Proposition 3.27:** Let $D$ be a nef $\mathbb{Q}$-Cartier divisor. Then $\kappa(D) = d - \text{rk}\ L_T$, where $T := \bigcup_{\mathcal{C} \in \text{fib}(D)} \mathcal{C}$.

**Proof.** We just remark that $\text{rk}\ L_T$ is the dimension of the subvector space of $M_\mathbb{Q}$ which is generated by the $l_i$ which are contained in a fibrational circuit.

**Proposition 3.28:** Let $X$ be a complete toric variety and $D$ a nef divisor, then $H^i(X,\mathcal{O}(D)) = 0$ for $i \neq \kappa(D)$.

**Proof.** Consider the hyperplane arrangement given by the $H_i^\mathbb{Z}$ in $M_\mathbb{Q}$. Let $m \in M_\mathbb{Q}$ and $I = \varphi(m)$. Then the simplicial complex $\Delta_I$ can be characterized as follows. Consider $Q \subset P_D$ the union of the set faces of $P_D$ which are contained in any $H_i^\mathbb{Z}$ with $i \in I$. This is precisely the portion of $P_D$, which the the point $m$ “sees”, and therefore contractible, where the convex hull of $Q$ and $m$ provides the homotopy between $Q$ and $m$. Therefore, every $\Delta_I$ is contractible with an exception for $I = \emptyset$, because $\Delta_\emptyset = \emptyset$, which is not acyclic with respect.
to reduced cohomology. Now we pass to the inverse, i.e. we consider the signature of \(-m\) with respect to \(H_t^{-\Delta}\). Then for any such \(-m\) which does not sit in the relative interior of the polytope \(P^{[n]}_{\mathbb{Z}}\), there exists \(m' \in M_\mathbb{Q}\) with signature \(\hat{\mathcal{A}}(m') = J\) such that \(\hat{\Delta}_J\) is contractible and \(\hat{\mathcal{A}}(m) = [n] \setminus J\). As \(\hat{\Delta}\) is homotopic to a \(d-1\)-sphere, we can apply Alexander duality and thus the simplicial complex \(\hat{\Delta}_{[n] \setminus J}\) is acyclic. Thus there remain only the elements in the relative interior of \(P^{[n]}_{\mathbb{Z}}\). Let \(m\) be such an element with signature \(I\), then \(\hat{\Delta}_I\) is isomorphic to a \(d-\kappa(D)-1\)-sphere, and the assertion follows.

This proposition implies the toric Kodaira and Kawamata-Viehweg vanishing theorems (see also [Mus02]):

**Theorem 3.29 (Kodaira & Kawamata-Viehweg):** Let \(X\) be a complete toric variety and \(D, E \in \mathbb{Q}\)-divisors with \(D\) nef and \(E = \sum_{i \in [n]} c_i D_i\) with \(-1 < c_i < 0\) for all \(i \in [n]\). Then:

(i) if \(D\) is integral, then \(H^i(X, \mathcal{O}(D + K_X)) = 0\) for all \(i \neq 0, d - \kappa(D)\);

(ii) if \(D + E\) is a Weil divisor, then \(H^i(X, \mathcal{O}(D + E)) = 0\) for all \(i > 0\).

*Proof.* Because a toric \(\mathbb{Q}\)-Cartier divisor is MCM, we can apply Serre duality (Corollary 3.25) and obtain \(H^i(X, \mathcal{O}(D + K_X)) \cong H^{d-i}(X, \mathcal{O}(-D))\) and (i) follows from Proposition 3.28. For (ii): \(D + E\) is contained the interior of every half space \(K_X + H_\mathbb{Q}\) for \(\mathcal{C} \in \mathfrak{S}_{\text{nef}}\), and the result follows.

**4 Arithmetic aspects of cohomology vanishing**

In this section we want to generalize classical vanishing results for integral divisors which cannot directly be derived from the setting of \(\mathbb{Q}\)-divisors as in section 3.4. From now on we assume that the \(l_i\) are integral. Recall that for any integral divisor \(D = \sum_{i \in [n]} c_i D_i\) and any torus invariant closed subvariety \(V\) of \(X\), vanishing of \(H_t^{-1}(X, \mathcal{O}(D))\) depends on two things:

(i) whether the set of signatures \(\mathcal{A}(M_\mathbb{Q})\) consists of \(I \subset [n]\) such that the relative cohomology groups \(H^{d-1}(\Delta_I, \hat{\mathcal{A}}_{V,I}; k)\) vanish, and,

(ii) if \(H^{d-1}(\Delta_I, \hat{\mathcal{A}}_{V,I}; k)\) for one such \(I\), whether the corresponding polytope \(P^I_\mathbb{Z}\) contain lattice points \(m\) with \(\mathcal{A}(m) = I\).

In the Gale dual picture, the signature \(\mathcal{A}(M_\mathbb{Q})\) coincides with the set of \(I \subset [n]\) such that the class of \(D\) in \(A_{d-1}(X)_\mathbb{Q}\) is contained in \(C_I\). For fixed \(I\), the classes of divisors \(D\) in \(A_{d-1}(X)\) such that the equation \(l_i(m) < -c_i\) for \(i \in I\) and \(l_i(m) \geq -c_i\) for \(i \notin I\) is satisfied, is counted by the *generalized partition function*. That is, by the function

\[
D \mapsto \{(k_1, \ldots, k_n) \in \mathbb{N}^n \mid \sum_{i \in [n] \setminus I} k_i D_i - \sum_{i \in I} k_i D_i = D \text{ where } k_i > 0 \text{ for } i \in I\}.
\]
So, in the most general picture, we are looking for $D$ lying in the common zero set of the vector partition function for all relevant signatures $I$ of $D$. In general, this is a difficult problem to determine these zero sets, and it is hardly necessary for practical purposes.

Vector partition functions play an important role in the combinatorial theory of rational polytopes and have been considered, e.g. in [Stu95], [BV97] (see also references therein). In [BV97] closed expressions in terms of residue formulas have been obtained. Moreover it was shown that the vector partition function is a piecewise quasipolynomial function, where the domains of quasipolynomiality are chambers (or possibly unions of chambers) of the secondary fan. In particular, for if $P^\circ$ is a rational bounded polytope, then the values of the vector partition function for $P^\circ_k$ for $k \geq 0$, is just the Ehrhart quasipolynomial.

A special case which we will consider in subsection 4.1 is where the vectors $l_1, \ldots, l_n$ form circuit. In this form, the computation of generalized partition functions is essentially equivalent to the classically known diophantine Frobenius problem (also known as money change problem or denumerant problem). We refer to the book [Ram05] for a general overview.

4.1 Arithmetic cohomology vanishing for circuits

In this subsection we assume that $n = d + 1$ and $C = [n]$ forms a circuit. In light of Theorem 2.1, for cohomology vanishing on a toric 1-circuit variety, we have to consider the reduced cohomology of simplicial complexes associated to its fan:

**Lemma 4.1:** Let $I \subset [n]$, such that $I \neq C^+$, then $H^i((\hat{\Delta}_C) I; k) = 0$ for all $i$. Moreover,

$$(\hat{\Delta}_C)^{C^+}_+ \cong S|C^+| - 2 \quad \text{and} \quad (\hat{\Delta}_C)^{C^-}_- \cong B^{|C^-| - 1},$$

where $B^k$ is the $k$-ball, with $B^{-1} := 0$.

**Proof.** It is easy to see that $(\hat{\Delta}_C)^{C^+}_+$ corresponds to the boundary of the $(|C^+| - 1)$-simplex, so it is homeomorphic to $S|C^+| - 2$. Similarly, $\{l_i\}_{i \in C^+}$ span a simplicial cone in $\Delta_C$ and thus $(\hat{\Delta}_C)^{C^-}_- \cong B^{|C^-| - 1}$. Now assume there exists $i \in C^+ \setminus I$, then $I$ is a face of the cone $\sigma_i$ and $(\hat{\Delta}_C)^{\sigma_i}_I$ is contractible. On the other hand, if $C^+$ is a proper subset of $I$, the set $I \cap C^-$ spans a cone $\tau$ in $\Delta_C$. The simplicial complex $\Delta_I$ then is homeomorphic to a simplicial decomposition of the $(|C^+| - 1)$-ball with center $\tau$ and boundary $(\hat{\Delta}_C)^{C^+}_+$. \qed

In this special situation, the chamber decomposition of $M_Q$ by hyperplanes $H^s$ as in subsection 3.2 contains at most one bounded chamber. In fact, if $D$ is a rational divisor, all maximal chambers are unbounded. If $D \not\sim 0$, we have precisely one bounded chamber for whose signatures there are precisely two possibilities. Namely, we either have for every $m$ in this chamber that $l_i(m) < -c_i$ for every $i \in C^-$ and $l_i(m) \geq -c_i$ for every $i \in C^+$, or vice versa. The set of rational divisor classes in $A_{d-1}(\mathbb{P}(\alpha, \xi))^Q \cong \mathbb{Q}$ corresponding
to torus invariant divisors whose associated bounded chamber has signature either $\mathcal{C}^+$ or $\mathcal{C}^-$ corresponds precisely to the two open intervals $(-\infty, 0)$ and $(0, \infty)$, respectively, in $A_{d-1}(\mathbb{P}(\alpha, \xi))^\vee$.

To count lattice points in the bounded chamber we can use a special case of the generalized partition function, i.e. the number of lattice points $m$ such that $l_i(m) \geq -c_i$ for $i \in \mathcal{C}^+$ and $l_i(m) \leq -c_i$ for $i \in \mathcal{C}^-$ coincides with the cardinality of the following set:

$$\{(k_1, \ldots, k_{d+1}) \in \mathbb{N}^{d+1} | k_i > 0 \text{ for } i \in \mathcal{C}^- \text{ and } \sum_{i \in \mathcal{C}^+} k_i D_i - \sum_{i \in \mathcal{C}^-} k_i D_i = D\}.$$  

For the integral case, this leads to arithmetic thickenings of the intervals $(-\infty, 0)$ and $(0, \infty)$ as follows:

**Definition 4.2:** We denote $F_\mathcal{C} \subset A_{d-1}(\mathbb{P}(\alpha, \xi))$ the complement of the semigroup of the form $\sum_{i \in \mathcal{C}^-} c_i D_i - \sum_{i \in \mathcal{C}^+} c_i D_i$, where $c_i \in \mathbb{N}$ for all $i$ with $c_i > 0$ for $i \in \mathcal{C}^+$.

The set $F_\mathcal{C}$ is the complement of the set of classes whose associated chamber has signature $\mathcal{C}^-$ and contains a lattice point. With this we can give a complete characterization of global cohomology vanishing:

**Proposition 4.3:** Let $\mathbb{P}(\alpha, \xi)$ be as before with associated fan $\Delta_\mathcal{C}$ and $D \in A_{d-1}(\mathbb{P}(\alpha, \xi))$, then:

(i) $H^i(\mathbb{P}(\alpha, \xi), \mathcal{O}(D)) = 0$ for $i \neq 0, |\mathcal{C}^+| - 1$;

(ii) $H^{|\mathcal{C}^+| - 1}(\mathbb{P}(\alpha, \xi), \mathcal{O}(D)) = 0$ iff $D \in F_\mathcal{C}$;

(iii) if $\mathcal{C}^+ \neq \mathcal{C}$, then $H^0(\mathbb{P}(\alpha, \xi), \mathcal{O}(D)) \neq 0$;

(iv) if $\mathcal{C}^+ = \mathcal{C}$, then $H^0(\mathbb{P}(\alpha, \xi), \mathcal{O}(D)) = 0$ iff $D \in F_{-\mathcal{C}}$.

**Proof.** The proof is immediate. Just observe that the simplicial complex $(\Delta_\mathcal{C})_m$, for $m$ an element in the bounded chamber, coincides either with $(\Delta_\mathcal{C})_{\mathcal{C}^+}$ or $(\Delta_\mathcal{C})_{\mathcal{C}^-}$.

Another case of interest is where $\mathcal{C}^+ \neq \mathcal{C}$ and $V = V(\tau)$, where $\tau$ is the cone spanned by the $l_i$ with $i \in \mathcal{C}^-$, i.e. $V$ is the unique maximal complete torus invariant subvariety of $\mathbb{P}(\alpha, \xi)$.

**Proposition 4.4:** Consider $\mathbb{P}(\alpha, \xi)$ such that $\alpha_i < 0$ for at least one $i$, $D \in A_{d-1}(\mathbb{P}(\alpha, \xi))$ and $V$ the maximal complete torus invariant subvariety of $\mathbb{P}(\alpha, \xi)$, then:

(i) $H^0(\mathbb{P}(\alpha, \xi), \mathcal{O}(D)) \neq 0$;

(ii) $H^{|\mathcal{C}^-|}(\mathbb{P}(\alpha, \xi), \mathcal{O}(D)) = 0$ iff $D \in F_\mathcal{C}$. 

Documenta Mathematica 16 (2011) 209–251
(iii) \(H^i((\mathbb{P}(\mathcal{C}, \xi), \mathcal{O}(D))) = 0\) for all \(i \neq d, |\mathcal{C}|\).

Proof. Consider first \(I = \mathcal{C}\), then \((\hat{\Delta}_\mathcal{C})_I = \hat{\Delta}_\mathcal{C} \cong B^{d-1}\) and \((\hat{\Delta}_\mathcal{C})_{V,I} = (\hat{\Delta}_\mathcal{C})_V \cong S^{d-2}\). It follows that \(H^i(\hat{\Delta}_\mathcal{C}; k) = 0\) for all \(i\) and \(H^{i-1}(\hat{\Delta}_\mathcal{C}, (\hat{\Delta}_\mathcal{C})_V; k) \cong H^{i-2}((\hat{\Delta}_\mathcal{C})_V; k)\). As by assumption, \(\mathcal{C}^+ \neq \mathcal{C}\), so the associated hyperplane arrangement contains an unbounded chamber such that \(l_i(m) \geq -c_i\) for all \(i \in \mathcal{C}\) and all \(m\) in this chamber. Hence (i) follows. As in the proof of lemma 4.1, it follows that \(\hat{\Delta}_I\) is contractible whenever \(\mathcal{C}^+ \cap I \neq \emptyset\) and \(\mathcal{C}^- \cap I \neq \emptyset\). So in that case \(H^i(\hat{\Delta}_I) = 0\) for all \(i\) and \(H^{i-1}(\hat{\Delta}_I, (\hat{\Delta}_\mathcal{C})_{V,I}; k) \cong H^{i-2}(\hat{\Delta}_{V,I}; k)\) for all \(i\).

Now let \(I = \mathcal{C}^+\); then \((\hat{\Delta}_\mathcal{C})_I = (\hat{\Delta}_\mathcal{C})_{V,I} \cong S^{d-2}, \) so \(H^i((\hat{\Delta}_\mathcal{C})_I, (\hat{\Delta}_\mathcal{C})_{V,I}; k) = 0\) for all \(i\). For \(I = \mathcal{C}^-\), then \((\hat{\Delta}_\mathcal{C})_I \cong B^{d-1}\) and \((\hat{\Delta}_\mathcal{C})_{V,I} \cong S^{d-2}\), the former by Lemma 4.1, the latter by Lemma 4.1 and the fact that \((\hat{\Delta}_\mathcal{C})_{V,I}\) has empty intersection with \(\text{star}(\tau)\). This implies (ii) and consequently (iii).

4.2 Arithmetic Kawamata-Viehweg vanishing

A first — trivial — approximation is given by the observation that the divisors \(D\) where the vector partition function takes a nontrivial value map to the cone \(C_I\), shifted by the offset \(e_I := -\sum_{i \in I} e_i\).

**Definition 4.5:** We denote \(\mathcal{O}'(L, I)\) the saturation of the cone generated the \(-D_i\) for \(i \in I\) and the \(D_i\), for \(i \not\in I\), and \(\mathcal{O}(L, I) := e_I + \mathcal{O}'(L, I)\). Moreover, we denote \(\Omega(L, I)\) the zero set in \(\mathcal{O}(L, I)\) of the vector partition function as defined above.

In the next step we want to approximate the sets \(\Omega(L, I)\) by reducing to the classical diophantine Frobenius problem. For this, fix some \(I \subset [n]\) and consider some polytope \(P^I_L\). It follows from Proposition 3.17 that \(D\) is contained in the intersection of half spaces \(H_\mathcal{C}\) for \(\mathcal{C} \in \mathcal{E}(L)\) such that \(\mathcal{C}^- = \mathcal{C} \cap I\). In the polytope picture, we can interpret this as follows. For every \(\mathcal{C}\) and its underlying circuit \(\mathcal{C}\), we set

\[
P^\mathcal{C}_L := \{ m \in M_\mathbb{Q} \mid l_i(m) \leq -c_i \text{ for } i \in \mathcal{C}^- \text{ and } l_i(m) \geq -c_i \text{ for } i \in \mathcal{C}^+ \}.
\]

Consequently, we get

\[
P^I_L = \bigcap_{\mathcal{C}} P^\mathcal{C}_L,
\]

where the intersection runs over all \(\mathcal{C} \in \mathcal{E}(L)\) with \(\mathcal{C}^- = \mathcal{C} \cap I\). It follows that if there exists a compatible oriented circuit \(\mathcal{C}\) such that \(P^\mathcal{C}_L\) does not contain a lattice point, then \(P^I_L\) also does not contain a lattice point. We want to capture this by considering an arithmetic analogue of the discriminantal arrangement in \(A_{d-1}(X)\) rather than in \(A_{d-1}(X)_\mathbb{Q}\). Using the integral pendant to diagram (5):
Consider the morphism $\eta_I : A_{d-1}(X) \to A_I$. Then we denote $Z_I$ its kernel. For $I = C$ and $C$ some orientation of $C$ we denote by $F_C$ the preimage in $A_{d-1}(X)$ of the complement of the semigroup consisting of elements $\sum_{i \in C^-} c_i D_i - \sum_{i \in C^+} c_i D_i$, where $c_i \geq 0$ for $i \in C^-$ and $c_i > 0$ for $i \in C^+$. We set $F_C := F_C \cap F_{-C}$.

So, there are two candidates for a discriminantal arrangement in $A_{d-1}(X)$, the $Z_C$ on the one hand, and the $F_C$ on the other.

**Definition 4.7:** We denote:
- $\{Z_C\}_{C \in C(L)}$ the *integral* discriminantal arrangement, and
- $\{F_C\}_{C \in C(L)}$ the *Frobenius* discriminantal arrangement.

The integral discriminantal arrangement has similar properties as the $H_I$, as they give a solution to the integral discriminant problem (compare Lemma 3.9):

**Lemma 4.8:** Let $I \subset [n]$, then

$$Z_I = \bigcap_{C \in C(L)} Z_C.$$  

We can now locate both the rational as well as the integral Picard group in $A_{d-1}(X)_\mathbb{Q}$ and $A_{d-1}(X)$, respectively:

**Theorem 4.9 (see [Eik92], Theorem 3.2):** Let $X$ be any toric variety, then:

(i) $\text{Pic}(X)_\mathbb{Q} = \bigcap_{\sigma \in \Delta_{\max}} H_\sigma(1) = \bigcap_{C \in C(L)} H_C$. 

(ii) $\text{Pic}(X) = \bigcap_{\sigma \in \Delta_{\max}} Z_\sigma(1) = \bigcap_{C \in C(L)} Z_C$.

**Proof.** (i) As remarked in subsection 3.4, a $\mathbb{Q}$-Cartier divisor is specified by a collection $\{m_\sigma\}_{\sigma \in \Delta} \subset \mathbb{M}_\mathbb{Q}$. In particular, all for every $\sigma \in \Delta$, the hyperplanes $H^\sigma_i$ with $i \in \sigma(1)$ have nonempty intersection. So the first equality follows. The second equality follows from Lemma 3.9.

(ii) A Cartier divisor is specified by a collection $\{m_\sigma\}_{\sigma \in \Delta} \subset \mathbb{M}$ such that the hyperplanes $H^\sigma_i$ with $i \in \sigma(1)$ intersect in integral points. So the first equality follows. The second equality follows from Lemma 4.8.

The Frobenius discriminantal arrangement is not as straightforward. First, we note the following properties:

**Lemma 4.10:** Let $C \in C(L)$, then:

(i) $F_C$ is nonempty;

(ii) the saturation of $Z_C$ in $A_{d-1}(X)$ is contained in $F_C$ iff $C$ is not fibrational.
Proof. The first assertion follows because $F_C$ contains all elements which map to the open interval $(K_C, K_C)$ in $A_C$, where $K_C = -\sum_{i \in C}^+ D_i$. For the second assertion, note that the set $\{ m \in M \mid l_i(m) = 0 \text{ for all } i \in C \}$ is in $F_C$ iff $C^+ \neq C$ for either orientation $C$.

Lemma 4.10 shows that the $F_C$ are thickenings of the $Z_C$ with one notable exception, where $C$ is fibrational. In this case, $F_C$ can be considered as parallel to, but slightly shifted away from $Z_C$. In the sequel we will not make any explicit use of the $Z_C$ anymore, but these facts should be kept in mind.

Regarding the Frobenius discriminantal arrangement, we want also to consider integral versions of the discriminantal strata:

**Definition 4.11**: Let $C \in C(L)$ and let $\mathfrak{F}_S$ be a discriminantal hull of $S = \{S_1, \ldots, S_k\}$, then we denote

$$\mathfrak{A}_S := \bigcap_{C \in \mathfrak{F}_S} F_C,$$

the arithmetic core of $\mathfrak{F}_S$. In the special case $\mathfrak{F}_S = \mathfrak{F}_{\text{nef}}$ we write $\mathfrak{A}_{\text{nef}}$.

**Remark 4.12**: The notion core refers to the fact that we consider all $F_C$, instead of a non-redundant subset describing the set $S$ as a convex cone.

We will use arithmetic cores to derive arithmetic versions of known vanishing theorems formulated in the setting of $\mathbb{Q}$-divisors and to get refined conditions on cohomology vanishing. This principle is reflected in the following theorem:

**Theorem 4.13**: Let $V$ be a $T$-invariant closed subscheme of $X$ and $S$ a discriminantal stratum in $A_{d-1}(X)$. If $H^i_V(X, O(D)) = 0$ for some $i$ and for all integral divisors $D \in S$, then also $H^i_V(X, O(D)) = 0$ for all $D \in \mathfrak{A}_S$.

**Proof.** Without loss of generality we can assume that $\dim S > 0$. Consider some nonempty $P_I$ for some $I \subset [n]$. Then for any such $I$, we can choose some multiple of $kD$ such that $P'_I$ contains a lattice point. But if $H^i_V(X, O(D)) = 0$, then also $H^i_V(X, O(kD)) = 0$, hence $H^{i-1}(\Delta_I, \Delta_V; k) = 0$. Now, any divisor $D' \in \mathfrak{A}_S$ which does not map to $S$, is contained in $F_C$ for all $C \in \mathfrak{F}_S$ and therefore for any $I$ which is in the signature for $D'$ but not for $D$, the equations $l_i(m) < -c'_i$ for $i \in I$ and $l_i(m) \geq -c'_i$ for $i \notin I$ cannot have any integral solution.

We apply Theorem 4.13 to $\mathfrak{A}_{\text{nef}}$:

**Theorem 4.14** (Arithmetic version of Kawamata-Viehweg vanishing): Let $X$ be a complete toric variety. Then $H^i(X, O(D)) = 0$ for all $i > 0$ and all $D \in \mathfrak{A}_{\text{nef}}$.  

_DOCUMENTA MATHEMATICA 16 (2011) 209–251_
Proof. We know that the assertion is true if $D$ is nef. Therefore we can apply Theorem 4.13 to the maximal strata $S_1, \ldots, S_k$ of nef($X$). Therefore the assertion is true for $D \in \bigcap_{i=1}^{k} \mathfrak{A}_{S_i}$. To prove the theorem, we have to get rid of the $F_{C}$, where $H_C$ intersects the relative interior of a face of nef($X$). Let $C$ be such a circuit and $R$ the face. Without loss of generality, $\dim R > 0$. Then we can choose elements $D'$ in $R$ at an arbitrary distance from $H_C$, i.e. such that the polytope $P^C$ becomes arbitrarily big and finally contains a lattice point. Now, if we move outside nef($X$), but stay inside $\mathfrak{A}_{\text{nef}}$, the lattice points of $P^C$ cannot acquire any cohomology and the assertion follows.

One can imagine an analog of the set $\mathfrak{A}_{S}$ in $A_{d-1}(X)\mathbb{Q}$ as the intersection of shifted half spaces

$$\bigcap_{C \in \mathfrak{S}} \left( - \sum_{i \in \mathcal{C}^+} D_i + H_C \right).$$

The main difference here is that one would picture the proper facets of this convex polyhedral set as “smooth”, whereas the proper “walls” of $\mathfrak{A}_{S}$ have “ripples”, which arise both from the fact that the groups $A_C$ may have torsion, and that we use Frobenius conditions to determine the augmentations of our half spaces.

In general, the set $\mathfrak{S}$ is highly redundant when it comes to determine $\mathfrak{S}$, which implies that above intersection does not yield a cone but rather a polyhedron, whose recession cone corresponds to $\mathfrak{S}$. In the integral situation we do not quite have a recession cone, but a similar property holds:

**Proposition 4.15:** Let $V \subset X$ be a closed invariant subscheme and $S = \{S_1, \ldots, S_k\}$ a collection of discriminantal strata different from zero. Then for any nonzero face of its discriminantal hull $\mathfrak{S}$ there exists the class of an integral divisor $D' \in \mathfrak{S}$ such that the intersection of the half line $D + rD'$ for $0 \leq r \in \mathbb{Q}$ with $\mathfrak{A}_{S}$ contains infinitely many classes of integral divisors.

**Proof.** Let $R \subset \mathfrak{S}$ be any face of $S$, then the vector space spanned by $R$ is given by an intersection $\bigcap_{C \in K} H_C$ for a certain subset $K \subset \mathfrak{S}$. We assume that $K$ is maximal with this property. The intersection $\bigcap_{C \in K} F_C$ is invariant with respect to translations along certain (though not necessarily all) $D' \in R$. This implies that the line (or any half line, respectively), generated by $D'$ intersects $\bigcap_{C \in K} F_C$ in infinitely many points. As $K$ is maximal, there is no other $C \in \mathfrak{S}$ parallel to $R$ and the assertion follows.

The property of Proposition 4.15 is necessary for elements in $\mathfrak{A}_{S}$, but not sufficient. This leads to the following definition:

**Definition 4.16:** Let $S = \{S_1, \ldots, S_k\}$ be a collection of discriminantal strata and $D \in A_{d-1}(X)$ such that the property of Proposition 4.15 holds. If $D$ is not contained in $\mathfrak{A}_{S}$, then we call $D$ $\mathfrak{A}_{S}$-residual. We call $D$ 0-residual if it is in the complement of $\mathfrak{A}_{0} = \bigcap_{C \in \mathcal{C}(\mathbb{L})} F_C$. 

*Documenta Mathematica* 16 (2011) 209–251
In the next subsections we will consider several special cases of interest for cohomology vanishing, which are not directly related to Kawamata-Viehweg vanishing theorems. In subsection 4.3 we will consider global cohomology vanishing for divisors in the inverse nef cone. In subsection 4.4 we will present a more explicit determination of this type of cohomology vanishing for toric surfaces. Finally, in subsection 4.5, we will give a geometric criterion for determining maximally Cohen-Macaulay modules.

4.3 Nonstandard Cohomology Vanishing

In this subsection we want to give a qualitative description of cohomology vanishing which is related to divisors which are inverse to nef divisors of Iitaka dimension $0 < \kappa(D) < d$. We show the following theorem:

**Theorem 4.17:** Let $X$ be a complete $d$-dimensional toric variety. Then $H^i(X, \mathcal{O}(D)) = 0$ for every $i$ and all $D$ which are contained in some $\mathfrak{A}_{-F}$, where $F$ is a face of nef($X$) which contains nef divisors of Iitaka dimension $0 < \kappa(D) < d$. If $\mathfrak{A}_{-F}$ is nonempty, then it contains infinitely many divisor classes.

**Proof.** Recall that such a divisor, as a $\mathbb{Q}$-divisor, is contained in the intersection $\bigcap_{C \in \text{fib}(D)} H_C$ and therefore it is in the intersection of the nef cone with the boundary of the effective cone of $X$ by Proposition 3.17. Denote this intersection by $F$. Then we claim that $H^i(X, \mathcal{O}(D')) = 0$ for all $D' \in \mathfrak{A}_{-F}$. By Corollary 3.28 we know that $H^i(X, \mathcal{O}(E)) = 0$ for $0 \leq i < d$ for any divisor $E$ in the interior of the inverse nef cone. This implies that $H^i(X, \mathcal{O}(E)) = 0$ for any $E \in \mathfrak{A}_{-nef}$ and hence $H^i(X, \mathcal{O}(D')) = 0$ for any $D' \in \mathfrak{A}_{-F}$, because $\mathfrak{A}_{-F} \subset \mathfrak{A}_{-nef}$. The latter assertion follows from the fact that the assumption on the Iitaka dimension implies that the face $F$ has positive dimension.

Note that criterion is not very strong, as it is not clear in general whether the set $\mathfrak{A}_{-F}$ is nonempty. However, this is the case in a few interesting cases, in particular for toric surfaces, as we will see in the next subsection. The following remark shows that our condition indeed is rather weak in general:

**Remark 4.18:** The inverse of any big and nef divisor $D$ with the property that $P_D$ does not contain any lattice point in its interior has the property that $H^i(X, \mathcal{O}(D)) = 0$ for all $i$. This follows directly from the standard fact in toric geometry that the Euler characteristics $\chi(-D)$ counts the inner lattice points of the lattice polytope $P_D$.

4.4 The case of complete toric surfaces.

Let $X$ be a complete toric surface. We assume that the $l_i$ are cyclically ordered. We consider the integers $[n]$ as system of representatives for $\mathbb{Z}/n\mathbb{Z}$, i.e. for some $i \in [n]$ and $k \in \mathbb{Z}$, the sum $i + k$ denotes the unique element in $[n]$ modulo $n$. 
Proposition 4.19: Let $X$ be a complete toric surface. Then $\text{nef}(X) = \overline{S}$, where $S$ is a single stratum of maximal dimension of the discriminant arrangement.

Proof. $X$ is simplicial and projective and therefore $\text{nef}(X)$ is a cone of maximal dimension in $A_1(X)_\mathbb{Q}$. We show that no hyperplane $H_C$ intersects the interior of $\text{nef}(X)$. By Proposition 3.17 we can at once exclude fibrational circuits. This leaves us with non-fibrational circuits $C$ with cardinality three, having orientation $\mathcal{C}$ with $|\mathcal{C}^+| = 2$. Assume that $D$ is contained in the interior of $H_{-\mathcal{C}}$. Then there exists $m \in M_{\mathbb{Q}}$ such that $\mathcal{C}^+ \subset s^D(m)$, which implies that the hyperplane $H_{-\mathcal{C}}^p$ for $[i] = \mathcal{C}^-$ does not intersect $P_D$, and thus $D$ cannot be nef. It follows that $\text{nef}(X) \subset H_{\mathcal{C}}$.

Now assume there exist $p, q \in [n]$ such that $l_q = -l_p$, i.e. $l_p$ and $l_q$ represent a one-dimensional fibrational circuit of $L$. Then for any nef divisors $D$ which is contained in $H_{p,q}$, the associated polytope $P_D$ is one-dimensional. The only possible variation for $P_D$ is its length in terms of lattice units. So we can conclude that $\text{nef}(X) \cap H_{p,q}$ is a one-dimensional face of $\text{nef}(X)$.

Definition 4.20: Let $X$ be a complete toric surface and $\mathcal{C} = \{p, q\}$ such that $l_p = -l_q$. Then we denote $S_{p,q}$ the relative interior of $-\text{nef}(X) \cap H_{\mathcal{C}}$. Moreover, we denote $\mathfrak{X}_{p,q}$ the arithmetic core of $S_{p,q}$.

Our aim in this subsection is to prove the following:

Theorem 4.21: Let $X$ be a complete toric surface. Then there are only finitely many divisors $D$ with $H^i(X, \mathcal{O}(D)) = 0$ for all $i > 0$ which are not contained in $\mathfrak{X}_{\text{nef}} \cup \bigcup \mathfrak{X}_{p,q}$, where the union ranges over all pairs $\{p, q\}$ such that $l_p = -l_q$.

We will prove this theorem in several steps. First we show that the interiors of the $C_l$ such that $H^0(\Delta_l; k) \neq 0$ cover all of $A_1(X)_\mathbb{Q}$ except $\text{nef}(X)$ and $-\text{nef}(X)$.

Proposition 4.22: Let $D = \sum_{i \in [n]} c_i D_i$ be a Weil divisor which is not contained in $\text{nef}(X)$ or $-\text{nef}(X)$, then the corresponding arrangement $\text{H}_{\mathcal{C}}^p$ in $M_{\mathbb{Q}}$ has a two-dimensional chamber $P_{\mathcal{C}}^l$ such that complex $\Delta_l$ has at least two components.

Proof. Recall that $\text{nef}(X) = \bigcap H_{\mathcal{C}}$, where the intersection runs over all oriented circuits which are associated to extremal curves of $X$. As the statement is well-known for the case where $X$ is either a 1-circuit toric variety or a Hirzebruch surface, we can assume without loss of generality, that the extremal curves belong to blow-downs, i.e. the associated oriented circuits are of the form $\mathcal{C}^+ = \{i-1, i+1\}$, $\mathcal{C}^- = \{i\}$ for any $i \in [n]$. Now assume that $D$ is in the interior of $H_{\mathcal{C}}$ for such an oriented circuit $\mathcal{C}$. Then there exists a bounded chamber $P_\mathcal{C}^l$ in $M_{\mathbb{Q}}$ such that $\mathcal{C}^- = \mathcal{C} \cap \mathfrak{s}(m)$. In order for $\Delta_{\mathfrak{s}(m)}$ to be acyclic, it is necessary that $\mathfrak{s}(m) \cap ([n] \setminus \mathcal{C}) = \emptyset$. Let $\{j, k, l\} =:\ D \subset [n]$.
represent any other circuit such that $\mathcal{D}^+ = \{j, l\}$ for some orientation $\mathcal{D}$ of $\mathcal{D}$. The hyperplane arrangement given by the three hyperplanes $H_{n-1}^n, H_{n-2}^n, H_{n-3}^n$ has six unbounded regions, whose signatures contain any subset of $\{j, k, l\}$ except $\{j, l\}$ and $\{k\}$. In the cases $j = i - 2, k = i - 1, l = 1$ or $j = i, k = i + 1, l = i + 2$, $P_{i}^j$ must be contained in the region with signature $\{i\}$. In every other case $P_{i}^j$ must be contained in the region with signature $\emptyset$. In the case, say, $\{j, k, l\} = \{i - 2, i - 1, i\}$, the hyperplane $H_{n-2}^n$ should not cross the bounded chamber related to the subarrangement given by the hyperplanes $H_{n-1}^n, H_{n-2}^n, H_{n-3}^n$, as else we obtain a chamber whose signature contains $\{i - 1, i + 1\}$, but not $\{i - 2, i\}$. Then the associated subcomplex of $\Delta$ can never be acyclic. This implies that, if $D$ is in the interior of $H_\epsilon$, then $D \in H_\mathcal{D}$, where either $D = \{i - 2, i - 1, i\}$ or $\mathcal{D} = \{i, i + 1, i + 2\}$. By iterating for every extremal (i.e. every invariant) curve, we conclude that $D \in \cap_{i \in \mathbb{N}} H_\epsilon = \text{nef}(X)$. Analogously, we conclude for $D \in -\text{nef}(X)$, and the statement follows.

Let $\{p, q\} \subset [n]$ such that $l_p = -l_q$. Then these two primitive vectors span a 1-dimensional subvector space of $N_\mathbb{Q}$, which naturally separates the set $[n] \setminus \{p, q\}$ into two subsets.

**Definition 4.23:** Let $\{p, q\} \subset [n]$ such that $l_p = -l_q$. Then we denote $A_{p, q}^1, A_{p, q}^2 \subset [n]$ the two subsets of $[n] \setminus \{p, q\}$ separated by the line spanned by $l_p, l_q$.

For some fibrational circuit $\{p, q\}$, the closure $\overline{\mathfrak{A}}_{p, q}$ is a one-dimensional cone in $A_1(X)_{\mathbb{Q}}$ which has a unique primitive vector:

**Definition 4.24:** Consider $\{p, q\}$ as before. Then the closure $\overline{\mathfrak{A}}_{p, q}$ is a one-dimensional cone with primitive lattice vector $D_{p, q} := \sum_{i \in A_1^2} l_i(m) D_i$, where $m \in M$ the unique primitive vector on the ray in $M_\mathbb{Q}$ with $l_p(m) = l_q(m) = 0$ and $l_i(m) < 0$ for $i \in A_1^1$.

**Proposition 4.25:** Let $X$ be a complete toric surface. Then every $\mathfrak{A}_{p, q}$-residual divisor on $X$ is either contained in $\mathfrak{A}_{\text{nef}}$, or in some $\mathfrak{A}_{p, q}$, or is $\mathfrak{A}_{\text{nef}}$-residual.

**Proof.** For any nef divisor $D \in -\mathfrak{A}_{p, q}$, the polytope $P_D$ is a line segment such that all $H_\epsilon^n$ intersect this line segment in one of its two end points, depending on whether $i \in A_{p, q}^1$ or $i \in A_{p, q}^2$. This implies that the line spanned by $S_{p, q}$ is the intersection of all $H_\epsilon$, where $\mathcal{C} \subset A_{p, q}^1 \cup \{p, q\}$ or $\mathcal{C} \subset A_{p, q}^2 \cup \{p, q\}$. Let $D$ be $\mathfrak{A}_{p, q}$-residual and assume that $H^1(X, \mathcal{O}(D + rD_{p, q})) = 0$ for all $i$ and for infinitely many $r$. We first show that $D \in F_{\{p, q\}}$, i.e. that $c_p + c_q = -1$ for any torus invariant representative $D = \sum_{i \in [n]} c_i D_i$. Assume that $c_p + c_q > -1$. Then there exists $m \in M$ such that $p, q \notin \Theta(m)$. By adding sufficiently high multiples of $D_{p, q}$ such that $D + rD_{p, q} = \sum c'_i D_i$, we can even find such an $m$ such that $A_1 \cup A_2 \subset \Theta(m)$, hence $H^1(X, \mathcal{O}(D + rD_{p, q})) \neq 0$ for large $r$ and thus $D$ is not $\mathfrak{A}_{p, q}$-residual. If $c_p + c_q < -1$, there is an $m \in M$ with
\{p, q\} \subset \mathcal{A}(m)$, and by the same argument, we get $H^2(X, \mathcal{O}(D + rD_{p,q})) \neq 0$ for large $r$. Hence $c_p + c_q = -1$, i.e. $D \in F_{\{p,q\}}$. This implies that for every $m \in M$ either $p \in \mathcal{A}(m)$ and $q \notin \mathcal{A}(m)$, or $q \in \mathcal{A}(m)$ and $p \notin \mathcal{A}(m)$. Now assume that $D \notin F_{\mathcal{C}}$ for some $\mathcal{C} = \{i, j, k\} \subset A_1 \cup \{p, q\}$ such that $\mathcal{C}^+ = \{i, k\}$ for some orientation. Then there exists some $m \in M$ with $\{i, k\} \subset \mathcal{A}(m)$ or $\{j\} \subset \mathcal{A}(m)$. In the first case, as before we can simply add some multiple of $D_{p,q}$ such that $i \in \mathcal{A}(m)$ and $i \in A_2$, hence $\mathcal{A}(m)$ contains at least two $-i$-intervals. In the second case, we have either $p \notin \mathcal{A}(m)$ or $q \notin \mathcal{A}(m)$, thus at least two $-i$-intervals, too. Hence $D \in \mathfrak{A}_{p,q}$ and the assertion follows.

**Proposition 4.26:** Let $X$ be a complete toric surface. Then $X$ has only a finite number of $\mathfrak{A}_{\text{net}}$-residual divisors.

**Proof.** We can assume without loss of generality that $X$ is not $\mathbb{P}^2$ nor a Hirzebruch surface. Assume there is $D \in A_1(X)$ which is not contained in $F_{\mathcal{C}}$ for some circuit $\mathcal{C} = \{i-1, i, i+1\}$ corresponding to an extremal curve on $X$. Then there exists a chamber in the corresponding arrangement whose signature contains $\{i-1, i, i+1\}$. To have this signature to correspond to an acyclic subcomplex of $\Delta$, the rest of the signature must contain $[n] \setminus \mathcal{C}$. Now assume we have some integral vector $D_{\mathcal{C}} \in H_{\mathcal{C}}$, then we can add a multiple of $D_{\mathcal{C}}$ to $D$ such that $D$ is parallel translated to $\text{nef}(X)$. In this process necessarily at least one hyperplane passes the critical chamber and thus creates cohomology. Now, $D$ might be outside of $F_{\mathcal{D}}$ for some $\mathcal{D} \in \mathcal{E}(L)$ not corresponding to an extremal curve. If the underlying circuit is not fibrational, then $D$ being outside $F_{\mathcal{D}}$ implies outside $F_{\mathcal{E}}$ for some extremal circuit $\mathcal{E}$. If $\mathcal{D}$ is fibrational and $\mathcal{D} = \{p, q\}$, then we argue as in Proposition 4.25 that $D$ has cohomology. If $D$ is fibrational of cardinality three, the corresponding hypersurface $H_D$ is not parallel to any nonzero face of $\text{nef}(X)$ and there might be a finite number of divisors lying outside $F_{\mathcal{D}}$ but in the intersection of all $F_{\mathcal{E}}$, where $\mathcal{C}$ corresponds to an extremal curve.

**Proposition 4.27:** Let $X$ be a complete toric surface. Then $X$ has only a finite number of $0$-residual divisors.

**Proof.** Let us consider some vector partition function $VP(L, I) : \mathbb{N} \to \mathbb{N}$, for $I$ such that $C_I$ does not contain a nonzero subvector space. Let $D = \sum_{i \in [n]} c_i D_i \in \Omega(L, I)$ and let $P_D$ the polytope in $M_Q$ such that $m \in M_Q$ is in $P_D$ iff $l_i(m) < -c_i$ for $i \in I$ and $l_i(m) \geq -c_i$ for $i \in [n] \setminus I$. For any $J \subset [n]$ we denote $P_{D,J}$ the polytope defined by the same inequalities, but only for $i \in J$. Clearly, $P_D \subset P_{D,J}$. Let $J \subset [n]$ be maximal with respect to the property that $P_{D,J}$ does not contain any lattice points. If $J \neq [n]$, then we can freely move the hyperplanes given by $l_i(m) = -c_i$ for $i \in [n] \setminus I$ such that $P_{D,J}$ remains constant and thus lattice point free. This is equivalent to say that there exists a nonzero $D' \in \bigcap_{\mathcal{C} \in \mathcal{E}(L,I)} H_{\mathcal{C}}$ and for every such $D'$ the polytope $P_{D+JD'}$ does not contain any lattice point for any $j \in \mathbb{Q}_{>0}$. 

**Documenta Mathematica** 16 (2011) 209–251
Now assume that $J = [n]$. This implies that the defining inequalities of $P_D$ are irredundant and thus there exists a unique maximal chamber in $C_I$ which contains $D$ (if $I = \emptyset$ this would be the nef cone by 4.19) and thus the combinatorial type of $P_D$ is fixed. Now, clearly, the number of polygons of shape $P_D$ with parallel faces given by integral linear inequalities and which do not contain a lattice point is finite.

By applying this to all (and in fact finitely many) cones $O_I$ such that $C_I$ does not contain a nontrivial subvector space of $A\mathbb Q$, we see that there are only finitely many divisors $D$ which are not contained in $A_{\text{nef}}$ or $A_{p,q}$.

**Proof of theorem 4.21.** By 4.22, $\text{nef}(X)$ and the $S_{p,q}$ are indeed the only relevant strata, which by 4.25 and 4.26 admit only finitely many residual elements. Hence, we are left with the 0-residuals, of which exist only finitely many by 4.27.

**Example 4.28:** Figure 1 shows the cohomology free divisors on the Hirzebruch surface $\mathbb F_3$ which is given by four rays, say $l_1 = (1,0)$, $l_2 = (0,1)$, $l_3 = (-1,3)$, $l_4 = (0,-1)$ with respect to some choice of coordinates for $N$. In $\text{Pic}(\mathbb F_3) \cong \mathbb Z^2$ there are two cones such that $H^1(X, \mathcal O(D)) \neq 0$ for every $D$ which is contained in one of these cones. Moreover, there is one cone such that $H^2(X, \mathcal O(D)) \neq 0$ for every $D$; its tip is sitting at $K_{\mathbb F_3}$. The nef cone is indicated by the dashed lines.

The picture shows the divisors contained in $A_{\text{nef}}$ as black dots. The white dots indicate the divisors in $A_{2,4}$. There is one 0-residual divisor indicated by the grey dot.

The classification of smooth complete toric surfaces implies that every such surface which is not $\mathbb P^2$, has a fibrational circuit of rank one. Thus the theorem implies that on every such surface there exist families of line bundles with vanishing cohomology along the inverse nef cone. For a given toric surface $X$,
these families can be explicitly computed by checking for every \( C \subset A_1 \cup \{ p, q \} \) and every \( C \subset A_2 \cup \{ p, q \} \), respectively, whether the inequalities

\[
\begin{align*}
    c_i + l_i(m) & \geq 0 \quad \text{for } i \in \mathcal{C}^+ \\
    c_i + l_i(m) & < 0 \quad \text{for } i \in \mathcal{C}^-
\end{align*}
\]

have solutions \( m \in M \) for at least one of the two orientations \( \mathcal{C}, -\mathcal{C} \) of \( C \). This requires to deal with \( \binom{|A_1| + 2}{3} + \binom{|A_2| + 2}{3} \), i.e. of order \( \sim n^3 \), linear inequalities. We can reduce this number to order \( \sim n^2 \) as a corollary from our considerations above:

**Corollary 4.29:** Let \( C \subset A_i \) for \( i = 1 \) or \( i = 2 \). Then there exist \( \{ i, j \} \subset C \) such that \( F_{\{p,q\}} \cap F_{\mathcal{C}} \supseteq F_{\{p,q\}} \cap F_{\{i,j\}} \cap F_{\{i,j\}} \).

**Proof.** Assume first that there exists \( m \in M \) which for the orientation \( \mathcal{C} \) of \( C = \{i_1, i_2, i_3\} \) with \( \mathcal{C}^+ = \{i_1, i_3\} \) such that fulfills the inequalities \( l_{i_k}(m) + c_{i_k} \geq 0 \) for \( k = 1, 3 \) and \( l_{i_2}(m) + c_{i_2} < 0 \). This implies that \( H^1(X, \mathcal{O}(D)) \neq 0 \), independent of the configuration of the other hyperplanes, as long as \( c_p + c_q = -1 \). It is easy to see that we can choose \( i, j \in \mathcal{C} \) such that \( \{i, j, p\} \) and \( \{i, j, q\} \) form circuits. We can choose one of those such that \( m \) is contained in the triangle, fulfilling the respective inequalities, and which is not fibrational. For the inverse orientation \(-\mathcal{C}\), we can the same way replace one of the elements of \( C \) by one of \( p, q \). By adding a suitable positive multiple of \( D_{p,q} \), we can rearrange the hyperplanes such that \( H^1(X, \mathcal{O}(D + rD_{p,q})) \neq 0 \).

One should read the corollary the way that for any pair \( i, j \) in \( A_1 \) or in \( A_2 \), one has only to check whether a given divisor fulfills certain inequalities for triples \( \{i, j, q\} \) and \( \{i, j, p\} \). It seems that it is not possible to reduce further the number of equations in general. However, there is a criterion which gives a good reduction of cases for practical purposes:

**Corollary 4.30:** Let \( X \) be a smooth and complete toric surface and \( D = \sum_{i \in [n]} \alpha_i D_i \in \mathcal{A}_{p,q} \), then for every \( i \in A_1 \cup A_2 \), we have:

\[
c_i - 1 + c_{i+1} - \alpha_ic_i \in [-1, \alpha_i - 1],
\]

where the \( \alpha_i \) are the self-intersection numbers of the \( D_i \).

**Proof.** The circuit \( C = \{i - 1, i, i + 1\} \) comes with the integral relation \( l_{i-1} + l_{i+1} + \alpha_il_i = 0 \). So the Frobenius problem for these circuits is trivial and we have only to consider the offset part.

The following example shows that these equalities are necessary, but not sufficient in general:

**Example 4.31:** We choose some coordinates on \( N \cong \mathbb{Z}^2 \) and consider the complete toric surface defined by 8 rays \( l_1 = (0, -1), l_2 = (1, -2), l_3 = (1, -1), \ldots \).
$l_4 = (1, 0)$, $l_5 = (1, 1)$, $l_6 = (1, 2)$, $l_7 = (0, 1)$, $l_8 = (-1, 0)$. Then any divisor $D = c_1 D_1 + \cdots + c_8 D_8$ with $c = (-1, 1, 0, 0, 1, 0, -k)$ for some $k \gg 0$ has nontrivial $H^1$, though it fulfills the conditions of corollary 4.30.

An interesting and more restricting case is the additional requirement that also $H^i(X, \mathcal{O}(-D)) = 0$ for all $i > 0$. One may compare the following with the classification of bundles of type $B$ in [HP06].

**Corollary 4.32:** Let $X$ be a smooth and complete toric surface and $D \in \mathfrak{A}_{p,q}$ such that $H^i(X, \mathcal{O}(D)) = H^i(X, \mathcal{O}(-D)) = 0$ for all $i > 0$. Then for every $i \in A_1 \cup A_2$, we have:

$$c_{i-1} + c_{i+1} - a_i c_i \in \begin{cases} \{\pm 1, 0\} & \text{if } a_i < -1 \\ \{-1, 0\} & \text{if } a_i = -1, \end{cases}$$

where the $a_i$ are the self-intersection numbers of the $D_i$.

**Proof.** For $-D$, we have $c_q + c_l = 1$. Assume that there exists a circuit circuit $C$ with orientation $C_1$ and $C^- = \{i, j\}$, and moreover, some lattice point $m$ such that $\mathcal{C}(m) \cap C = C^-$. Then we get $s^{-2(-m)} \cap C = C^+$. This implies that $H^1(X, \mathcal{O}(-D)) \neq 0$. This implies the restriction $c_{i-1} + c_{i+1} - a_i c_i \in [-1, \min\{1, a_i - 1\}]$. \qed

Note that example 4.31 also fulfills these more restrictive conditions.

### 4.5 Maximal Cohen-Macaulay Modules of Rank One

The classification of maximal Cohen-Macaulay modules can sometimes be related to resolution of singularities, the most famous example for this being the McKay correspondence in the case of certain surface singularities ([GSV83], [AV85]; see also [EK85]). In the toric case, in general one cannot expect to arrive at such a nice picture, as there does not exist a canonical way to construct resolutions. However, there is a natural set of preferred partial resolutions, which is parameterized by the secondary fan.

Let $X$ be a $d$-dimensional affine toric variety whose associated convex polyhedral cone $\sigma$ has dimension $d$. Denote $x \in X$ torus fixed point. For any Weil divisor $D$ on $X$, the sheaf $\mathcal{O}_X(D)$ is MCM if and only if $H^0_i(X, \mathcal{O}_X(D))$ for all $i < d$. It was shown in [BG03] (see also [BG02]) that there exists only a finite number of such modules.

Now let $\tilde{X}$ be a toric variety given by some triangulation of $\sigma$. The natural map $\pi : \tilde{X} \to X$ is a partial resolution of the singularities of $X$ which is an isomorphism in codimension two and has at most quotient singularities. In particular, the map of fans is induced by the identity on $N$ and, in turn, induces a bijection on the set of torus invariant Weil divisors. This bijection induces a natural isomorphism $\pi^{-1} : A_{d-1}(X) \to A_{d-1}(\tilde{X})$ which can be represented by the identity morphism on the invariant divisor group $\mathbb{Z}^n$. This allows us to

Documenta Mathematica 16 (2011) 209–251
identify a torus invariant divisor $D$ on $X$ with its strict transform $\pi^{-1}D$ on $\tilde{X}$. Moreover, there are the natural isomorphisms

$$\pi_*\mathcal{O}_X(\pi^{-1}D) \cong \mathcal{O}_X(D) \quad \text{and} \quad \mathcal{O}_X(\pi^{-1}D) \cong (\pi^*\mathcal{O}_X(D))^\sim.$$ 

Our aim is to compare local cohomology and global cohomology, i.e. $H^i_X(X, \mathcal{O}_X(D))$ and $H^i(\tilde{X}, \mathcal{O}_\tilde{X}(D))$.

Probably the easiest class of cones $\sigma$ which one can consider is where the primitive vectors $l_1, \ldots, l_n$ form a circuit $C \equiv [n]$. Associated to this data are two small resolutions of singularities $\pi : P(\underline{\omega}, \xi) \rightarrow X$ and $\pi' : P(-\underline{\omega}, \xi) \rightarrow X$ which are induced by triangulations $\Delta_\xi$ and $\Delta_{-\xi}$, respectively.

Now, the question whether $\mathcal{O}(D)$ is a maximal Cohen-Macaulay sheaf can be decided directly on $Y$ or, equivalently, on the resolutions:

**Theorem 4.33:** Let $X$ be an affine toric variety whose associated cone $\sigma$ is spanned by a circuit $C$ and denote $P(\underline{\omega}, \xi)$ and $P(-\underline{\omega}, \xi)$ the two canonical small toric resolutions of singularities. Then the sheaf $\mathcal{O}(D)$ is maximal Cohen-Macaulay if and only if $R^i\pi_*\mathcal{O}(\pi^{-1}D) = R^i\pi'_*\mathcal{O}((\pi')^{-1}D) = 0$ for all $i > 0$.

**Proof.** This toric variety corresponds to the toric subvariety of $Y$ which is the complement of its unique fixed point, which we denote $y$. We have to show that $H^i_y(Y, \mathcal{O}(D)) = 0$ for all $i < d$. By Corollary 2.2, we have

$$H^i_y(Y, \mathcal{O}(D)) = H^{i-2}(\tilde{\sigma}_y,\mathbb{P}^1;k)$$ 

for every $m \in M$, where $\tilde{\sigma}_y$ denotes the simplicial model for the fan associated to $Y \setminus \{y\}$. Denote $\tau$ and $\tau'$ the cones corresponding to the minimal orbits of $P(\underline{\omega}, \xi)$ and $P(-\underline{\omega}, \xi)$, respectively. We observe that $(\Delta_\xi)_{V(\tau)} = (\Delta_\xi)_{V(\tau')}$ both coincide with the subfan of $\sigma$ generated by its facets. It follows that the simplicial complexes relevant for computing the isotypical decomposition of $H^i_y(Y, \mathcal{O}(D))$ coincide with the simplicial complexes relevant for computing the $H^i_V(P(\underline{\omega}, \xi), \mathcal{O}(\pi^{-1}D))$ and $H^i_{V'}(P(-\underline{\omega}, \xi), \mathcal{O}((\pi')^{-1}D))$, respectively, where $V, V'$ denote the exceptional sets of the morphisms $\pi$ and $\pi'$, respectively.

By Proposition 4.4 the corresponding cohomologies vanish for $i < d$ iff $D \in F_{\xi} \cap F_{-\xi}$. Now we observe that $\Gamma(Y, R^i\pi_*\mathcal{O}(\pi^{-1}D)) = H^i(P(\underline{\omega}, \xi), \mathcal{O}(\pi^{-1}D))$ and $\Gamma(Y, R^i\pi'_*\mathcal{O}((\pi')^{-1}D)) = H^i(P(-\underline{\omega}, \xi), \mathcal{O}((\pi')^{-1}D))$. By Proposition 4.3, both cohomologies vanish for $i > 0$ iff $D \in F_{\xi} \cap F_{-\xi}$.

**Remark 4.34:** The relation between maximal Cohen-Macaulay modules and the diophantine Frobenius problem has also been discussed in [Sta96]. See [Yos90] for a discussion of MCM-finiteness of toric 1-circuit varieties.

More generally, we have the following easy statement about general (i.e. non-regular) triangulations:

**Theorem 4.35:** Let $X$ be an affine toric variety of dimension $d$ and $D \in A_{d-1}(X)$. If $D$ is 0-essential, then $R^i\pi_*\mathcal{O}_X(\pi^*D) = 0$ for every triangulation $\pi : \tilde{X} \rightarrow X$. 

*Documenta Mathematica* 16 (2011) 209–251
Proof. If $D$ is 0-essential, then it is contained in the intersection of all $F_{\tau}$, where $C \in C(L)$, thus it represents a cohomology-free divisor. \qed

Note that the statement does hold for any triangulation and not only for regular triangulations. We have a refined statement for affine toric varieties whose associated cone $\sigma$ has simplicial facets:

**Theorem 4.36:** Let $X$ be a $d$-dimensional affine toric variety whose associated cone $\sigma$ has simplicial facets and let $D \in A_{d-1}(X)$. If $R^i\pi_*\mathcal{O}_X(\pi^*D) = 0$ for every regular triangulation $\pi : \breve{X} \to X$ then $\mathcal{O}_X(D)$ is MCM. For $d = 3$ the converse is also true.

**Proof.** Recall that $H^i_x(X, \mathcal{O}(D))|_m = H^{i-2}(\tilde{\sigma}_V, m; k)$ for some $m \in M$ and $D \in A$. We are going to show that for every subset $I \subseteq [n]$ there exists a regular triangulation $\Delta$ of $\sigma$ such that the simplicial complexes $\tilde{\sigma}_{V,I}$ and $\tilde{\Delta}_I$ coincide. This implies that if $H^i_\tau(X, \mathcal{O}_X(D))|_m \neq 0$ for some $m \in M$, then also $H^{i+1}(\breve{X}, \mathcal{O}_{\breve{X}}(D))|_m \neq 0$, i.e. if $\mathcal{O}_X(D)$ is not MCM, then $H^i(\breve{X}, \mathcal{O}_{\breve{X}}(D)) \neq 0$ for some $i > 0$.

For a given $I \subseteq [n]$ we get such a triangulation as follows. Let $i \in [n] \setminus I$ and consider the dual cone $\tilde{\sigma}$. Denote $\rho_i := \mathbb{Q}_{\geq 0} l_i$ and recall that $\rho_i$ is a halfspace which contains $\tilde{\sigma}$ and which defines a facet of $\tilde{\sigma}$ given by $\rho_i \cap \tilde{\sigma}$. Now we move $\rho_i$ to $m + \tilde{\rho}$, where $l_i(m) > 0$. So we obtain a new polytope $P := \tilde{\sigma} \cap (m + \tilde{\rho})$. As $\rho_i$ is not parallel to any face of $\tilde{\sigma}$, the hyperplane $m + \rho_i$ intersects every face of $\tilde{\sigma}$. This way the inner normal fan of $P$ is a triangulation $\tilde{\Delta}$ of $\sigma$ which has the property that every maximal cone is spanned by $\rho_i$ and some facet of $\sigma$. This implies $\tilde{\Delta}_I = \tilde{\sigma}_{V,I}$ and the first assertion follows.

For $d = 3$, a sheaf $\mathcal{O}(D)$ is MCM iff $H^2_x(X, \mathcal{O}(D)) = 0$, i.e. $H^0(\sigma_{V,m}; k) = 0$ for every $m \in M$. The latter is only possible if $\sigma_{V,m}$ represents an interval on $S^1$. To compare this with $H^2(\breve{X}, \mathcal{O}(D))$ for some regular triangulation $\breve{X}$, we must show that $H^2(\Delta_m; k) = 0$ for the corresponding complex $\Delta_m$. To see this, we consider some cross-section $\sigma \cap H$, where $H \subset N \otimes_{\mathbb{Z}} \mathbb{R}$ is some hyperplane which intersects $\sigma$ nontrivially and is not parallel to any of its faces. Then this cross-section can be considered as a planar polygon and $\sigma_{V,m}$ as some connected sequence of faces of this polygon. Now with respect to the triangulation $\Delta$ of this polygon, we can consider two vertices $p, q \in \sigma_{V,m}$ which are connected by a line belonging to the triangulation and going through the interior of the polygon. We assume that $p$ and $q$ have maximal distance in $\sigma_{V,m}$ with this property. Then it is easy to see that the triangulation of $\sigma$ induces a triangulation of the convex hull of the line segments connecting $p$ and $q$. Then $\Delta_m$ is just the union of this convex hulls with respect all such pairs $p, q$ and the remaining line segments and thus has trivial topology. Hence $H^2_x(X, \mathcal{O}(D)) = 0$ implies $H^2(\breve{X}, \mathcal{O}(D)) = 0$ for every triangulation $\Delta$ of $\sigma$. \qed

**Example 4.37:** Consider the 3-dimensional cone spanned over the primitive vectors $l_1 = (1,0,1)$, $l_2 = (0,1,1)$, $l_3 = (-1,0,1)$, $l_4 = (-1,-1,1)$, $l_5 = (1,-1,1)$, $l_6 = (1,-1,-1)$, $l_7 = (1,1,-1)$, $l_8 = (1,1,1)$, $l_9 = (0,0,1)$, $l_{10} = (0,0,0)$. Then the associated cone $\sigma$ is not 0-essential, but $\mathcal{O}_X(D)$ is MCM because of the fact that $H^1(\breve{X}, \mathcal{O}(D)) = 0$. On the other hand, $H^2_x(X, \mathcal{O}(D)) = 0$, i.e. $H^2(\breve{X}, \mathcal{O}(D)) = 0$, because $\sigma$ is not simplicial.

\textbf{Documenta Mathematica} 16 (2011) 209–251
(1, −1, 1). The corresponding toric variety $X$ is Gorenstein and its divisor class group is torsion free. For $A_2(X) \cong \mathbb{Z}^2$ we choose the basis $D_1 + D_2 + D_5$, $D_5$. In this basis, the Gale duals of the $l_i$ are $D_1 = (-1, -1)$, $D_2 = (2, 0)$, $D_3 = (-3, 1)$, $D_4 = (2, -1)$, $D_5 = (0, 1)$. Figure 2 shows the set of MCM modules in $A_2(X)$ which are indicated by circles which are supposed to sit on the lattice $A_2(X) \cong \mathbb{Z}^2$. The picture also indicates the cones $C_I$ with vertices $-e_I$, where $I \in \{(1, 3), (1, 4), (2, 4), (2, 5), (3, 5), (1, 2, 5), (1, 3, 4), (1, 3, 5), (2, 3, 5), (2, 4, 5)\}$. Note that the picture has a reflection symmetry, due to the fact that $X$ is Gorenstein. Altogether, there are 19 MCM modules of rank one, all of which are 0-essential. For $\mathcal{C} = \{l_1, l_3, l_4, l_5\}$, the group $A_2(X)_\mathcal{C} \cong \mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$ has torsion. The two white circles indicate modules which are contained in the $\mathbb{Q}$-hyperplanes $D_1 + D_4 + H_2$ and $D_2 + D_3 + D_5 + H_6$, respectively, but not in the sets $D_1 + D_4 + Z_6$ and $D_2 + D_3 + D_5 + Z_6$, respectively. Some of the $\Omega_I$ are not saturated, however, every divisor which is contained in some $(-e_I + C_1) \cap \Omega(L, I)$ is also contained in some $\Omega_{I'} \setminus \Omega(L, I')$ for some other $I' \neq I$. So for this example, the Frobenius arrangement gives a full description of MCM modules of rank one.

![Figure 2: The 19 MCM modules of example 4.37.](image)

**Example 4.38:** To give a counterexample to the reverse direction of theorem 4.36 for $d > 3$, we consider the four-dimensional cone spanned over the primitive vectors $l_1 = (0, -1, -1, 1)$, $l_2 = (-1, 0, 1, 1)$, $l_3 = (0, 1, 0, 1)$, $l_4 = (-1, 0, 0, 1)$, $l_5 = (-1, -1, 0, 1)$, $l_6 = (1, 0, 0, 1)$. The corresponding variety $X$ has 31 MCM modules of rank one which are shown in figure 3. Here, with basis $D_1$ and $D_6$, we have $D_1 = (1, 0)$, $D_2 = (1, 0)$, $D_3 = (-1, -2)$, $D_4 = (3, 1)$, $D_5 = (2, -2)$, $D_6 = (0, 1)$. There are six cohomology cones corresponding to $I \in \{(1, 2), (3, 5), (4, 6), (1, 2, 3, 5), (1, 2, 4, 6), (3, 4, 5, 6)\}$. The example features two modules which are not 0-essential, indicated by the grey dots sitting on the boundary of the cones $-e_I + C_1$, where $I \in \{(1, 2), (3, 5), (4, 6), (1, 2, 3, 5), (1, 2, 4, 6), (3, 4, 5, 6)\}$. The example features two modules which are not 0-essential, indicated by the grey dots sitting on the boundary of the cones $-e_I + C_1$, where $I \in \{(1, 2), (3, 5), (4, 6), (1, 2, 3, 5), (1, 2, 4, 6), (3, 4, 5, 6)\}$. For $A_2(X) \cong \mathbb{Z}^2$ we choose the basis $D_1 + D_2 + D_5$, $D_5$. In this basis, the Gale duals of the $l_i$ are $D_1 = (-1, -1)$, $D_2 = (2, 0)$, $D_3 = (-3, 1)$, $D_4 = (2, -1)$, $D_5 = (0, 1)$. Figure 2 shows the set of MCM modules in $A_2(X)$ which are indicated by circles which are supposed to sit on the lattice $A_2(X) \cong \mathbb{Z}^2$. The picture also indicates the cones $C_I$ with vertices $-e_I$, where $I \in \{(1, 3), (1, 4), (2, 4), (2, 5), (3, 5), (1, 2, 5), (1, 3, 4), (1, 3, 5), (2, 3, 5), (2, 4, 5)\}$. Note that the picture has a reflection symmetry, due to the fact that $X$ is Gorenstein. Altogether, there are 19 MCM modules of rank one, all of which are 0-essential. For $\mathcal{C} = \{l_1, l_3, l_4, l_5\}$, the group $A_2(X)_\mathcal{C} \cong \mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$ has torsion. The two white circles indicate modules which are contained in the $\mathbb{Q}$-hyperplanes $D_1 + D_4 + H_2$ and $D_2 + D_3 + D_5 + H_6$, respectively, but not in the sets $D_1 + D_4 + Z_6$ and $D_2 + D_3 + D_5 + Z_6$, respectively. Some of the $\Omega_I$ are not saturated; however, every divisor which is contained in some $(-e_I + C_1) \cap \Omega(L, I)$ is also contained in some $\Omega_{I'} \setminus \Omega(L, I')$ for some other $I' \neq I$. So for this example, the Frobenius arrangement gives a full description of MCM modules of rank one.

![Figure 2: The 19 MCM modules of example 4.37.](image)

**Example 4.38:** To give a counterexample to the reverse direction of theorem 4.36 for $d > 3$, we consider the four-dimensional cone spanned over the primitive vectors $l_1 = (0, -1, -1, 1)$, $l_2 = (-1, 0, 1, 1)$, $l_3 = (0, 1, 0, 1)$, $l_4 = (-1, 0, 0, 1)$, $l_5 = (-1, -1, 0, 1)$, $l_6 = (1, 0, 0, 1)$. The corresponding variety $X$ has 31 MCM modules of rank one which are shown in figure 3. Here, with basis $D_1$ and $D_6$, we have $D_1 = (1, 0)$, $D_2 = (1, 0)$, $D_3 = (-1, -2)$, $D_4 = (3, 1)$, $D_5 = (2, -2)$, $D_6 = (0, 1)$. There are six cohomology cones corresponding to $I \in \{(1, 2), (3, 5), (4, 6), (1, 2, 3, 5), (1, 2, 4, 6), (3, 4, 5, 6)\}$. The example features two modules which are not 0-essential, indicated by the grey dots sitting on the boundary of the cones $-e_I + C_1$, where $I \in \{(1, 2), (3, 5), (4, 6), (1, 2, 3, 5), (1, 2, 4, 6), (3, 4, 5, 6)\}$.
The white dots denote MCM divisors $D, -D$ such that there exists a triangulation of the cone of $X$ such that on the associated variety $\tilde{X}$ we have $H^i(\tilde{X}, \mathcal{O}(\pm D)) \neq 0$ for some $i > 0$. Namely, we consider the triangulation which is given by the maximal cones spanned by \{1, 2, 4, 5\}, \{1, 2, 4, 6\}, \{1, 2, 5, 6\}, \{1, 3, 4, 6\}, \{2, 3, 4, 6\}. Figure 4.38 indicates the two-dimensional faces of this triangulation via a three-dimensional cross-section of the cone.

We find that we have six cohomology cones corresponding to $I \in \{\{1, 2\}, \{3, 5\}, \{1, 2, 3\}, \{4, 5, 6\}, \{1, 2, 3, 5\}, \{3, 4, 5, 6\}\}$. In particular, we have non-vanishing $H^1$ for the points $-D_1 - D_2 - D_3$ and for $-D_4 - D_5 - D_6$, which correspond to $D$ and $-D$. 
References

[Ath99] A. Athanasiadis. The largest intersection lattice of a discriminantal arrangement. \textit{Beitr. Algebra Geom.}, 40(2):283–289, 1999.

[AV85] M. Artin and J.-L. Verdier. Reflexive modules over rational double points. \textit{Math. Ann.}, 270(1):79–82, 1985.

[BB97] M. Bayer and K. A. Brandt. Discriminantal arrangements, fiber polytopes and formality. \textit{J. Algebraic Combin.}, 6(3):229–246, 1997.

[BFS90] L. J. Billera, P. Filliman, and B. Sturmfels. Constructions and Complexity of Secondary Polytopes. \textit{Advances in Mathematics}, 83:155–179, 1990.

[BG02] W. Bruns and J. Gubeladze. Semigroup algebras and discrete geometry. In \textit{Séminaires et Congrès}, volume 6, pages 43–127. Soc. Math. France, Paris, 2002.

[BG03] W. Bruns and J. Gubeladze. Divisorial Linear Algebra of Normal Semigroup Rings. \textit{Algebras and Representation Theory}, 6:139–168, 2003.

[BGS87] R.-O. Buchweitz, G.-M. Greuel, and F.-O. Schreyer. Cohen-Macaulay modules on hypersurface singularities II. \textit{Invent. Math.}, 88:165–182, 1987.

[BLS+93] A. Björner, M. Las Vergnas, B. Sturmfels, N. White, and G. Ziegler. \textit{Oriented matroids}, volume 46 of \textit{Encyclopedia of Mathematics and its Applications}. Cambridge University Press, 1993.

[Bon90] A. I. Bondal. Representation of associative algebras and coherent sheaves. \textit{Math. USSR Izvestiya}, 34(1):23–42, 1990.

[BR86] M. Beltrametti and L. Robbiano. Introduction to the theory of weighted projective spaces. \textit{Expo. Math.}, 4:111–162, 1986.

[BV97] M. Brion and M. Vergne. Residue formulae, vector partition functions and lattice points in rational polytopes. \textit{J. Am. Math. Soc.}, 10(4):797–833, 1997.

[Coh86] D. C. Cohen. A Conjecture about Compact Quotients by Tori. In \textit{Arrangements — Tokyo 1998}, volume 27 of \textit{Adv. Studies in Pure Mathematics}, pages 59–68. Math. Soc. Japan, 1986.

[Cox95] D. A. Cox. The Homogeneous Coordinate Ring of a Toric Variety. \textit{J. Algebr. Geom.}, 4, 1:17–50, 1995.

[Cra84] H. Crapo. Concurrence geometries. \textit{Adv. Math.}, 54:278–301, 1984.
D. C. Cohen and A. N. Varchenko. Resonant local systems on comple-
ments of discriminantal arrangements and $\mathfrak{sl}_2$ representations. Geom. 
Dedicata, 101:217–233, 2003.

C. Delorme. Espaces projectifs anisotropes. Bull. Soc. Math. France, 
103:203–223, 1975.

I. Dolgachev. Weighted projective varieties. In Group actions and 
vector fields, Proc. Pol.-North Am. Semin., Vancouver 1981, pages 
34–71, 1982.

M. Eikelberg. The Picard group of a compact toric variety. Result. 
Math., 22(1–2):509–527, 1992.

H. Esnault and H. Knörrer. Reflexive modules over rational double 
points. Math. Ann., 272(4):545–548, 1985.

D. Eisenbud, M. Mustaţă, and M. Stillman. Cohomology on Toric 
Varieties and Local Cohomology with Monomial Supports. J. Symb. 
Comput., 29(4–5):583–600, 2000.

M. Falk. A note on discriminantal arrangements. Proc. AMS, 
122(4):1221–1227, 1994.

O. Fujino. Multiplication maps and vanishing theorems for toric va-
rieties. Math. Z., 257(3):631–641, 2007.

W. Fulton. Introduction to Toric Varieties. Princeton University 
Press, 1993.

I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky. Discriminants, 
resultants and multidimensional determinants. Math. Theory Appl. 
Birkhäuser, 1994.

G. González-Sprinberg and J.-L. Verdier. Construction géométrique 
de la correspondance de McKay. Ann. Sci. Ecole Norm. Sup., 16:409– 
449, 1983.

L. Hille and M. Perling. A Counterexample to King’s Conjecture. 
Comp. Math, 142(6):1507–1521, 2006.

A. King. Tilting bundles on some rational surfaces. Unpublished 
manuscript, 1997.

M. Kontsevich. Homological algebra of mirror symmetry. In Pro-
ceedings of the International Congress of Mathematicians, Vol. 1, 2 
(Zürich 1994), pages 120–139. Birkhäuser, 1995.

E. N. Materov. The Bott formula for toric varieties. Mosc. Math. J., 
2(1):161–182, 2002.
[MS89] Y. I. Manin and V. V. Schechtman. Arrangements of hyperplanes, higher braid groups and higher Bruhat orders. In J. Coates et al., editors, Algebraic number theory. Papers in honor of K. Iwasawa on the occasion of his 70th birthday on September 11, 1987, volume 17 of Advances Studies in Pure Mathematics, pages 67–84. Braunschweig: Vieweg, 1989.

[Mus02] M. Mustaţă. Vanishing Theorems on Toric Varieties. Tohoku Math. J., II. Ser., 54(3):451–470, 2002.

[Oda88] T. Oda. Convex Bodies and Algebraic Geometry, volume 15 of Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer, 1988.

[OP91] T. Oda and H. S. Park. Linear Gale transforms and Gelfand-Kapranov-Zelevinskij decomposition. Tohoku Math. J., II Ser., 43(3):375–399, 1991.

[Per04] M. Perling. TiltingSheaves — a program to compute strongly exceptional collections on toric varieties. http://www.mathematik.uni-kl.de/~perling/ts/ts.html, 2004.

[Ram05] J. L. Ramírez Alfonsín. The Diophantine Frobenius Problem, volume 30 of Oxford Lecture Series in Mathematics and its Applications. Oxford University Press, 2005.

[Rei99] V. Reiner. The generalized Baues problem. In L. J. Billera et al., editors, New perspectives in algebraic combinatorics, pages 293–336. Cambridge University Press, 1999.

[Rud90] A. N. Rudakov. Helices and vector bundles: seminar Rudakov, volume 148 of London Mathematical Society lecture note series. Cambridge University Press, 1990.

[Sta82] R. P. Stanley. Linear diophantine equations and local cohomology. Invent. Math., 68:175–193, 1982.

[Sta96] R. P. Stanley. Combinatorics and commutative algebra. Birkhäuser, 1996.

[Stu95] B. Sturmfels. On Vector Partition Functions. Journal of Combinatorial Theory, Series A, 72:302–309, 1995.

[Yos90] Y. Yoshino. Cohen-Macaulay modules over Cohen-Macaulay rings, volume 146 of London Mathematical Society lecture note series. Cambridge University Press, 1990.
Divisorial Cohomology Vanishing on Toric Varieties

Markus Perling
Ruhr-Universität Bochum
Fakultät für Mathematik
Universitätsstraße 150
44780 Bochum, Germany
Markus.Perling@rub.de
ALGEBRAIC ZIP DATA

RICHARD PINK, TORSTEN WEDHORN, AND PAUL ZIEGLER

Received: November 11, 2010
Communicated by Peter Schneider

Abstract. An algebraic zip datum is a tuple $Z = (G, P, Q, \varphi)$ consisting of a reductive group $G$ together with parabolic subgroups $P$ and $Q$ and an isogeny $\varphi: P/R_uP \to Q/R_uQ$. We study the action of the group $E_Z := \{ (p, q) \in P \times Q \mid \varphi(\pi_P(p)) = \pi_Q(q) \}$ on $G$ given by $((p, q), g) \mapsto pgq^{-1}$. We define certain smooth $E_Z$-invariant sub-varieties of $G$, show that they define a stratification of $G$. We determine their dimensions and their closures and give a description of the stabilizers of the $E_Z$-action on $G$. We also generalize all results to non-connected groups.

We show that for special choices of $Z$ the algebraic quotient stack $[E_Z \backslash G]$ is isomorphic to $[G \backslash Z]$ or to $[G \backslash Z']$, where $Z$ is a $G$-variety studied by Lusztig and He in the theory of character sheaves on spherical compactifications of $G$ and where $Z'$ has been defined by Moonen and the second author in their classification of $F$-zips. In these cases the $E_Z$-invariant subvarieties correspond to the so-called “$G$-stable pieces” of $Z$ defined by Lusztig (resp. the $G$-orbits of $Z'$).

2010 Mathematics Subject Classification: 14L30 (20G15, 20F55, 20G40)
Keywords and Phrases: Generalized $G$-stable pieces, zip data, $F$-zips

1 Introduction

1.1 Background

Let $G$ be a connected reductive linear algebraic group over an algebraically closed field $k$. Then $G \times G$ acts on $G$ via simultaneous left and right translation $((g_1, g_2), g) \mapsto g_1gg_2^{-1}$. In a series of papers, Lusztig ([Lus1], [Lus2]), He...
Richard Pink, Torsten Wedhorn, and Paul Ziegler

([He2], [He1], [He3]), and Springer ([Spr3]) studied a certain spherical $G \times G$-equivariant smooth compactification $\bar{G}$ of $G$. For $G$ semi-simple adjoint this is the so-called wonderful compactification from [DCP]. In general the $G \times G$-orbits $Z_I \subset \bar{G}$ are in natural bijection to the subsets $I$ of the set of simple reflections in the Weyl group of $G$. Lusztig and He defined and studied so-called $G$-stable pieces in $Z_I$, which are certain subvarieties that are invariant under the diagonally embedded subgroup $G \subset G \times G$. These $G$-stable pieces play an important role in their study of character sheaves on $\bar{G}$. Lusztig and He also consider non-connected groups, corresponding to twisted group actions. Other generalizations of these varieties have been considered by Lu and Yakimow ([LY2]). A further motivation to study $G$-stable pieces comes from Poisson geometry: It was proved by Evens and Lu ([EL]), that for certain Poisson structure, each $G$-orbit on $Z_I$ is a Poisson submanifold.

In [MW] Moonen and the second author studied the De Rham cohomology $H^*_\text{DR}(X/k)$ of a smooth proper scheme $X$ with degenerating Hodge spectral sequence over an algebraically closed field $k$ of positive characteristic. They showed that $H^*_\text{DR}(X/k)$ carries the structure of a so-called $F$-zip, namely: it is a finite-dimensional $k$-vector space together with two filtrations (the “Hodge” and the “conjugate” filtration) and a Frobenius linear isomorphism between the associated graded vector spaces (the “Cartier isomorphism”). They showed that the isomorphism classes of $F$-zips of fixed dimension $n$ and with fixed type of Hodge filtration are in natural bijection with the orbits under $G := \text{GL}_n(k)$ of a variant $Z'_I$ of the $G \times G$-orbit $Z_I$ studied by Lusztig. They studied the varieties $Z'_I$ for arbitrary reductive groups $G$ and determined the $G$-orbits in them as analogues of the $G$-stable pieces in $Z_I$. By specializing $G$ to classical groups they deduce from this a classification of $F$-zips with additional structure, e.g., with a non-degenerate symmetric or alternating form. They also consider non-connected groups. Moreover, the automorphism group of an $F$-zip is isomorphic to the stabilizer in $G$ of any corresponding point in $Z'_I$.

When $X$ varies in a smooth family over a base $S$, its relative De Rham cohomology forms a family of $F$-zips over $S$. The set of points of $S$ where the $F$-zip lies in a given isomorphism class is a natural generalization of an Ekedahl-Oort stratum in the Siegel moduli space. Information about the closure of such a stratum corresponds to information about how the isomorphism class of an $F$-zip can vary in a family, and that in turn is equivalent to determining which $G$-orbits in $Z'_I$ are contained in the closure of a given $G$-orbit.

In each of these cases one is interested in the classification of the $G$-stable pieces, the description of their closures, and the stabilizers of points in $G$. In this article we give a uniform approach to these questions that generalizes all the above situations.

1.2 Main results

The central definition in this article is the following:
DEFINITION 1.1. A connected algebraic zip datum is a tuple $Z = (G, P, Q, \varphi)$ consisting of a connected reductive linear algebraic group $G$ over $k$ together with parabolic subgroups $P$ and $Q$ and an isogeny $\varphi: P/R_uP \to Q/R_uQ$. The group
\[ E_Z := \{ (p, q) \in P \times Q \mid \varphi(\pi_P(p)) = \pi_Q(q) \} \]
is called the zip group associated to $Z$. It acts on $G$ through the map $(p, q, g) \mapsto pqg^{-1}$. The union of the $E_Z$-orbits of all elements of a subset $X \subset G$ is denoted by $\alpha_Z(X)$.

Fix such data $Z = (G, P, Q, \varphi)$. To apply the machinery of Weyl groups to $Z$ we choose a Borel subgroup $B$ of $G$, a maximal torus $T$ of $B$, and an element $g$ of $G$ such that $B \subset Q$, $gB \subset P$, $\varphi(\pi_P(gB)) = \pi_Q(B)$, and $\varphi(\pi_P(T)) = \pi_Q(T)$. Let $W$ denote the Weyl group of $G$ with respect to $T$, and $S \subset W$ the set of simple reflections corresponding to $T$. Let $I \subset S$ be the type of the parabolic $P$ and $W_I \subset W$ its Weyl group. Let $I^*W$ be the set of all $w \in W$ that have minimal length in their coset $W_Iw$. To each $w \in I^*W$ we associate the $E_Z$-invariant subset
\[ G^w = o_Z(gBwB) \] (1.2)
and prove (Theorems 5.10, 5.11 and 5.14):

**Theorem 1.3.** The $E_Z$-invariant subsets $G^w$ form a pairwise disjoint decomposition of $G$ into locally closed smooth subvarieties. The dimension of $G^w$ is $\dim P + \ell(w)$.

Next the isogeny $\varphi$ induces an isomorphism of Coxeter system $\psi: (W_I, I) \to (W_J, J)$ (see (3.11) for its precise definition), where $W_J \subset W$ and $J \subset S$ are the Weyl group and the type of the parabolic subgroup $Q$. Let $\leq$ denote the Bruhat order on $W$. We prove (Theorem 6.2):

**Theorem 1.4.** The closure of $G^w$ is the union of $G^{w'}$ for all $w' \in I^*W$ such that there exists $y \in W_I$ with $yw\psi(y)^{-1} \leq w$.

We call $Z$ orbitally finite if the number of $E_Z$-orbits in $G$ is finite. We give a necessary and sufficient criterion for this to happen (Proposition 7.1). In particular it happens when the differential of $\varphi$ at 1 vanishes, for instance if $\varphi$ is a Frobenius isogeny (Proposition 7.3). We prove (Theorem 7.5):

**Theorem 1.5.** If $Z$ is orbitally finite, then each $G^w$ is a single $E_Z$-orbit, and so the set $\{ gw \mid w \in I^*W \}$ is a set of representatives for the $E_Z$-orbits in $G$.

One can also consider the $E_Z$-orbit of $gw$ for any element $w \in W$ instead of just those in $I^*W$. It is then natural to ask when two such orbits lie in the same $E_Z$-invariant piece $G^w$. (For orbitally finite $Z$ this is equivalent to asking when the orbits are equal.) We give an explicit description of this equivalence relation on $W$ that depends only on the subgroup $W_I$ and the homomorphism $\psi$ (Theorem 9.17). We prove that all equivalence classes have the same cardinality $#W_I$, although they are in general no cosets of $W_I$ and we
do not know a simple combinatorial description for them. It is intriguing that we obtain analogous results for an abstract zip datum based on an arbitrary finitely generated Coxeter group (Theorem 9.11) or even an arbitrary abstract group (Theorem 9.6) in place of $W$.

Other results include information on point stabilizers and infinitesimal stabilizers (Section 8), the generalization of the main results to non-connected groups (Section 10), a dual parametrization by a set $W'$ in place of $W'$ (Section 11) and the relation with the varieties $Z_I$ studied by Lusztig and He and their generalizations $Z'_I$ (Section 12).

### 1.3 Applications

Let us explain why this theory of algebraic zip data is a generalization of the situations described in Subsection 1.1. In Section 12 we consider a connected reductive algebraic group $G$ over $k$, an isogeny $\varphi : G \to G$, a subset $I$ of the set of simple reflections associated to $G$, and an element $x$ in the Weyl group of $G$ satisfying certain technical conditions. To such data we associate a certain algebraic variety $X_{I,\varphi,x}$ with an action of $G$, a certain connected algebraic zip datum $Z$ with underlying group $G$, and morphisms

$$G \xleftarrow{\lambda} G \times G \xrightarrow{\rho} X_{I,\varphi,x} \quad (1.6)$$

In Theorem 12.8 we show that there is a closure preserving bijection between the $E_Z$-invariant subsets of $A \subset G$ and the $G$-invariant subsets of $B \subset X_{I,\varphi,x}$ given by $\lambda^{-1}(A) = \rho^{-1}(B)$. We also prove that the stabilizer in $E_Z$ of $g \in G$ is isomorphic to the stabilizer in $G$ of any point of the $G$-orbits in $X_{I,\varphi,x}$ corresponding to the orbit of $g$. These results can also be phrased in the language of algebraic stacks, see Theorem 12.7.

In the special case $\varphi = \text{id}_G$ the above $X_{I,\varphi,x}$ is the variety $Z_I$ defined by Lusztig. In Theorem 12.19 we verify that the subsets $G^{w} \subset G$ correspond to the $G$-stable pieces defined by him. Thus Theorem 1.4 translates to a description of the closure relation between these $G$-stable pieces, which had been proved before by He [He2].

If $\text{char}(k)$ is positive and $\varphi$ is the Frobenius isogeny associated to a model of $G$ over a finite field, the above $X_{I,\varphi,x}$ is the variety $Z'_I$ defined by Moonen and the second author. In this case the zip datum $Z$ is orbitally finite, and so we obtain the main classification result for the $G$-orbits in $Z'_I$ from [MW], the closure relation between these $G$-orbits, and the description of the stabilizers in $G$ of points in $Z'_I$. In this case the closure relation had been determined in the unpublished note [Wed], the ideas of which are reused in the present article. Meanwhile Viehmann [Vie] has given a different proof of the closure relation in this case using the theory of loop groups. For those cases which pertain to the study of modulo $p$ reductions of $F$-crystals with additional structure that show up in the study of special fibers of good integral models of Shimura varieties of Hodge type Moonen ([Moo]) and, more generally, Vasiu ([Vas]) have obtained...
similar classification results. In these cases Vasiu (loc. cit.) has also shown that the connected component of the stabilizers are unipotent.

For $G = \text{GL}_n$ (resp. a classical group) we therefore obtain a new proof of the classification of $F$-zips (resp. of $F$-zips with additional structure) from [MW]. We can also deduce how $F$-zips (possibly with additional structure) behave in families, and can describe their automorphism groups as the stabilizers in $E_Z$ of the corresponding points of $G$. This is applied in [VW] to the study of Ekedahl-Oort strata for Shimura varieties of PEL type.

1.4 Contents of the paper

In Section 2 we collect some results on algebraic groups and Coxeter systems that are used in the sequel. Algebraic zip data $Z$ are defined in Section 3, where we also establish basic properties of the triple $(B,T,g)$, called a frame of $Z$.

Section 4 is based on the observation that every $E_Z$-orbit is contained in the double coset $Pg\dot{x}Q$ for some $x \in W$ and meets the subset $g\dot{x}M$, where $M$ is a Levi subgroup of $Q$. In it we define another zip datum $Z_\dot{x}$ with underlying reductive group $M$ and establish a number of results relating the $E_Z$-orbits in $Pg\dot{x}Q$ to the $E_{Z_\dot{x}}$-orbits in $M$. This is the main induction step used in most of our results.

In Section 5 we give different descriptions of the $E_Z$-invariant subsets $G^w$ for $w \in I_W$ and prove Theorem 1.3. In Section 6 we determine the closure of $G^w$ and prove Theorem 1.4. Orbitally finite zip data are studied in Section 7, proving Theorem 1.5. Section 8 contains some results on point stabilizers and infinitesimal stabilizers. Abstract zip data are defined and studied in Section 9.

In Section 10 our main results are generalized to algebraic zip data based on non-connected groups.

In Section 11 we discuss a dual parametrization of the subsets $G^w$ by a subset $W^J$ of $W$ in place of $I_W$. Finally, in Section 12 we prove the results described in Subsection 1.3 above.

The paper is based on parts of the unpublished note [Wed] by the second author and the master thesis [Zie] by the third author, but goes beyond both.

After the referee pointed out to us the references [LY1] and [He3], we became aware that there Lu, Yakimov and He study a class of group actions which contains ours when $\varphi$ is an isomorphism. In this case, Theorems 1.3 and 1.4 were already proven in [loc. cit]. Also, many of the ideas we have used to study the decomposition of $G$ into $E_Z$-stable pieces are already present there.

**Acknowledgements.** We thank the referee for pointing out some references. The second author was partially supported by the SPP 1388 “Representation theory” of the DFG.
2 Preliminaries on algebraic groups and Coxeter groups

Throughout, the inner automorphism associated to an element \( h \) of a group \( G \) will be denoted \( \text{int}(h): G \to G, g \mapsto hg^{-1} \). Similarly, for any subset \( X \subset G \) we set \( hX := hXh^{-1} \).

2.1 General facts about linear algebraic groups

Throughout, we use the language of algebraic varieties over a fixed algebraically closed field \( k \). By an algebraic group \( G \) we always mean a linear algebraic group over \( k \). We let \( R_u G \) denote the unipotent radical of the identity component of \( G \) and \( \pi_G: G \to G/R_u G \) the canonical projection. An isogeny between two connected algebraic groups is a surjective homomorphism with finite kernel.

Consider an algebraic group \( G \), an algebraic subgroup \( H \) of \( G \), and a quasi-projective variety \( X \) with a left action of \( H \). Then we denote by \( G \times^H X \) the quotient of \( G \times X \) under the left action of \( H \) defined by \( h \cdot (g, x) = (gh^{-1}, h \cdot x) \), which exists by [Ser], Section 3.2. The action of \( G \) on \( G \times X \) by left multiplication on the first factor induces a left action of \( G \) on \( G \times^H X \). This is the pushout of \( X \) with respect to the inclusion \( H \hookrightarrow G \).

**Lemma 2.1.** For \( G, H \), and \( X \) as above, the morphism \( X \to G \times^H X \) which sends \( x \in X \) to the class of \((1, x)\) induces a closure-preserving bijection between the \( H \)-invariant subsets of \( X \) and the \( G \)-invariant subsets of \( G \times^H X \). If \( Y \subset X \) is an \( H \)-invariant subvariety of \( X \), then the corresponding \( G \)-invariant subset of \( G \times^H X \) is the subvariety \( G \times^H Y \) of \( G \times^H X \).

**Proof.** The morphism in question is the composite of the inclusion \( i: X \to G \times X, x \mapsto (1, x) \) and the projection \( \text{pr}: G \times X \to G \times^H X \). Let \((g, h) \in G \times H \) act on \( G \times X \) from the left by \((g', x) \mapsto (gg' h^{-1}, h \cdot x) \). Then the \( G \times H \)-invariant subsets of \( G \times X \) are the sets of the form \( G \times A \) for \( H \)-invariant subsets \( A \subset X \). Therefore \( i \) induces a closure-preserving bijection between the \( H \)-invariant subsets of \( X \) and the \( G \times H \)-invariant subsets of \( G \times X \). Furthermore, since \( G \times^H X \) carries the quotient topology with respect to \( \text{pr} \), the morphism \( \text{pr} \) induces a closure-preserving bijection between the \( G \times H \)-invariant subsets of \( G \times X \) and the \( G \)-invariant of \( G \times^H X \). Altogether this proves the claim.

**Lemma 2.2** (see [Slo], Lemma 3.7.4). Let \( G \) be an algebraic group with an algebraic subgroup \( H \). Let \( X \) be a variety with a left action of \( G \). Let \( f: X \to G/H \) be a \( G \)-equivariant morphism, and let \( Y \subset X \) be the fiber \( f^{-1}(H) \). Then \( Y \) is stabilized by \( H \), and the map \( G \times^H Y \to X \) sending the equivalence class of \((g, y)\) to \( g \cdot y \) defines an isomorphism of \( G \)-varieties.

**Lemma 2.3.** Let \( G \) be an algebraic group acting on an algebraic variety \( Z \) and let \( P \subset G \) be an algebraic subgroup such that \( G/P \) is proper. Then for any \( P \)-invariant subvariety \( Y \subset Z \) one has

\[
G \cdot Y = \overline{G \cdot Y}.
\]
Proof. Clearly we have
\[ G \cdot Y \subset G \cdot \bar{Y} \subset G \cdot Y \]
and therefore it suffices to show that \( G \cdot \bar{Y} \) is closed in \( Z \). The action \( \pi: G \times Z \to Z \) of \( G \) on \( Z \) induces a morphism \( \bar{\pi}: G \times^P Z \to Z \) which can be written as the composition
\[ G \times^P Z \xrightarrow{\sim} G/P \times Z \to Z. \]
Here the first morphism is the isomorphism given by \([g,z] \mapsto (gP, g \cdot z)\) and the second morphism is the projection. As \( G/P \) is proper, we deduce that the morphism \( \bar{\pi} \) is closed. Now \( Y \) is \( P \)-invariant and therefore \( G \times^P Y \) is defined, and it is a closed subscheme of \( G \times^P Z \). Therefore \( \bar{\pi}(G \times^P Y) = G \cdot \bar{Y} \) is closed in \( Z \).

The following statements concern images under twisted conjugation:

**Theorem 2.4** (Lang-Steinberg, see [Ste], Theorem 10.1). Let \( G \) be a connected algebraic group and \( \varphi: G \to G \) an isogeny with only a finite number of fixed points. Then the morphism \( G \to G, \ g \mapsto g\varphi(g)^{-1} \) is surjective.

**Proposition 2.5.** Let \( G \) be a connected reductive algebraic group with a Borel subgroup \( B \) and a maximal torus \( T \subset B \). Let \( \varphi: G \to G \) be an isogeny with \( \varphi(B) = B \). In (b) also assume that \( \varphi(T) = T \).

(a) The morphism \( G \times B \to G, \ (g, b) \mapsto gb\varphi(g)^{-1} \) is surjective.

(b) The morphism \( G \times T \to G, \ (g, t) \mapsto gt\varphi(g)^{-1} \) has dense image.

If \( G \) is simply connected semisimple and \( \varphi \) is an automorphism of \( G \), (b) has been shown by Springer ([Spr2] Lemma 4).

**Proof.** For (a) see [Ste], Lemma 7.3. Part (b) and its proof are a slight modification of this. Equivalently we may show that for some \( t_0 \in T \), the image of the morphism \( \tilde{\alpha}: G \times T \to G, \ (g, t) \mapsto gt_0\varphi(g)^{-1}t_0^{-1} \) is dense. For this it will suffice to show that the differential of \( \tilde{\alpha} \) at 1 is surjective. This differential is the linear map
\[ \text{Lie}(G) \times \text{Lie}(T) \to \text{Lie}(G) \]
\[ (X, Y) \mapsto X + Y - \text{Lie}(\varphi_{t_0})(X), \]
where \( \varphi_{t_0} := \text{int}(t_0) \circ \varphi \). This linear map has image
\[ \text{Lie}(T) + (1 - \text{Lie}(\varphi_{t_0})) \text{Lie}(G). \]
Let \( B^- \) be the Borel subgroup opposite to \( B \) with respect to \( T \). Since \( \varphi(B) = B \) and \( \varphi(T) = T \), the differential of \( \varphi_{t_0} \) at 1 preserves \( \text{Lie}(R_u B) \) and \( \text{Lie}(R_u B^-) \). If we find a \( t_0 \) such that \( \text{Lie}(\varphi_{t_0}) \) has no fixed points on \( \text{Lie}(R_u B) \) and \( \text{Lie}(R_u B^-) \) we will be done.
Let $\Phi$ be the set of roots of $G$ with respect to $T$. For each $\alpha \in \Phi$, let $x_\alpha$ be a basis vector of $\text{Lie}(U_\alpha)$, where $U_\alpha$ is the unipotent root subgroup of $G$ associated to $\alpha$. As the isogeny $\varphi$ sends $T$ to itself, it induces a bijection $\tilde{\varphi}: \Phi \rightarrow \Phi$ such that $\varphi(U_\alpha) = U_{\tilde{\varphi}(\alpha)}$. For each $\alpha \in \Phi$ there exists a $c(\alpha) \in k$ such that $\text{Lie}(\varphi)(x_\alpha) = c(\alpha)x_{\tilde{\varphi}(\alpha)}$. This implies $\text{Lie}(\varphi_{t_0})(x_\alpha) = \alpha(t_0)c(\alpha)x_{\tilde{\varphi}(\alpha)}$. Since $\varphi_{t_0}$ fixes $R_uB$ and $R_uB^-$, its differential permutes $\Phi^+$ and $\Phi^-$, where $\Phi^+$ (resp. $\Phi^-$) is the set of roots that are positive (resp. negative) with respect to $B$. Hence $\text{Lie}(\varphi_{t_0})$ can only have a fixed point in $\text{Lie}(R_uB)$ or $\text{Lie}(R_uB^-)$ if there exists a cycle $(\alpha_1, \cdots, \alpha_n)$ of the permutation $\tilde{\varphi}$ in $\Phi^+$ or $\Phi^-$ such that

$$
\prod_{i=1}^n \alpha_i(t_0)c(\alpha_i) = 1.
$$

This shows that for $t_0$ in some non-empty open subset of $T$, the differential $\text{Lie}(\varphi_{t_0})$ has no fixed points on $\text{Lie}(R_uB)$ and $\text{Lie}(R_uB^-)$.

### 2.2 Coset Representatives in Coxeter Groups

Here we collect some facts about Coxeter groups and root systems which we shall need in the sequel. Let $W$ be a Coxeter group and $S$ its generating set of simple reflections. Let $\ell$ denote the length function on $W$; thus $\ell(w)$ is the smallest integer $n > 0$ such that $w = s_1 \cdots s_n$ for suitable $s_i \in S$. Any such product with $\ell(w) = n$ is called a reduced expression for $w$.

Let $I$ be a subset of $S$. We denote by $W_I$ the subgroup of $W$ generated by $I$, which is a Coxeter group with set of simple reflections $I$. Also, we denote by $W^I$ (respectively $^IW$) the set of elements $w$ of $W$ which have minimal length in their coset $wW_I$ (respectively $W_Iw$). Then every $w \in W$ can be written uniquely as $w = w^I \cdot w_I = \tilde{w}_I \cdot ^Iw$ with $w_I, \tilde{w}_I \in W_I$ and $w^I \in W^I$ and $^Iw \in ^IW$. Moreover, these decompositions satisfy $\ell(w) = \ell(w_I) + \ell(w^I) = \ell(\tilde{w}_I) + \ell(^Iw)$ (see [DDPW], Proposition 4.16). In particular, $W^I$ and $^IW$ are systems of representatives for the quotients $W/W_I$ and $W_I/W$, respectively. The fact that $\ell(w) = \ell(w^{-1})$ for all $w \in W$ implies that $W^I = (^IW)^{-1}$.

Furthermore, if $J$ is a second subset of $S$, we let $^IW_J$ denote the set of $x \in W$ which have minimal length in the double coset $W_IxW_J$. Then $^IW_J = ^I(W \cap W_I)$, and it is a system of representatives for $W_I \backslash W/W_J$ (see [DDPW] (4.3.2)).

In the next propositions we take an element $x \in ^IW_J$, consider the conjugate subset $z^{-1}I \subset W$, and abbreviate $I_x := J \cap z^{-1}I \subset J$. Then $^I_zW_J$ is the set of elements $w_J$ of $W_J$ which have minimal length in their coset $W_Iw_J$. Likewise $W_I^{z^{-1}I}$ is the set of elements $w_I$ of $W_I$ which have minimal length in their coset $w_IW_I^{z^{-1}I}$.

**Proposition 2.6** (Kilmoyer, [DDPW], Proposition 4.17). For $x \in ^IW_J$ we have

$$
W_I \cap z^{-1}W_J = W_I^{z^{-1}I} \quad \text{and} \quad W_J \cap z^{-1}W_I = W_J^{z^{-1}I} = W_{I_x}.
$$
Proposition 2.7 (Howlett, [DDPW], Proposition 4.18). For any \( x \in I W^J \), every element \( w \) of the double coset \( W_I x W_J \) is uniquely expressible in the form \( w = w_I x w_J \) with \( w_I \in W_I \) and \( w_J \in I_s W_J \). Moreover, this decomposition satisfies
\[
\ell(w) = \ell(w_I x w_J) = \ell(w_I) + \ell(x) + \ell(w_J).
\]

Proposition 2.8. The set \( I W \) is the set of all \( xw_I \) for \( x \in I W^J \) and \( w_I \in I_s W_J \).

Proof. Take \( x \in I W^J \) and \( w_J \in I_s W_J \). Then for any \( w_I \in I W \), Proposition 2.7 applied to the product \( w_I x w_J \) implies that \( \ell(w_I x w_J) = \ell(w_I) + \ell(x) + \ell(w_J) \geq \ell(x) + \ell(w_J) = \ell(xw_J) \). This proves that \( xw_J \in I W \). Conversely take \( w \in I W \) and let \( w = w_I x w_J \) be its decomposition from Proposition 2.7. Then by the first part of the proof we have \( xw_J \in I W \). Since \( W_I w = W_I x w_J \), this implies that \( w = xw_J \).

Proposition 2.9. The set \( W^J \) is the set of all \( w_I x \) for \( x \in I W^J \) and \( w_I \in W_I^{I \setminus J} \).

Proof. Apply Proposition 2.8 with \( I \) and \( J \) interchanged and invert all elements of \( W^J \).

Next we recall the Bruhat order on \( W \), which we denote by \( \leq \) and \( < \). This natural partial order is characterized by the following property: For \( x, w \in W \) we have \( x \leq w \) if and only if for some (or, equivalently, any) reduced expression \( w = s_1 \cdots s_n \) as a product of simple reflections \( s_i \in S \), one gets a reduced expression for \( x \) by removing certain \( s_i \) from this product. More information about the Bruhat order can be found in [BB], Chapter 2.

Using this order, the set \( I W \) can be described as
\[
I W = \{w \in W \mid w < sw \text{ for all } s \in I\} \quad (2.10)
\]
(see [BB], Definition 2.4.2 and Corollary 2.4.5).

Assume in addition that \( W \) is the Weyl group of a root system \( \Phi \), with \( S \) corresponding to a basis of \( \Phi \). Denote the set of positive roots with respect to the given basis by \( \Phi^+ \) and the set of negative roots by \( \Phi^- \). For \( I \subset S \), let \( \Phi_I \) be the root system spanned by the basis elements corresponding to \( I \), and set \( \Phi_I^\pm := \Phi_I \cap \Phi^\pm \). Then by [Car], Proposition 2.3.3 we have
\[
W^I = \{w \in W \mid w\Phi_I^+ \subset \Phi^+\} = \{w \in W \mid w\Phi_I^- \subset \Phi^-\} \quad (2.11)
\]
Also, by [Car], Proposition 2.2.7, the length of any \( w \in W \) is
\[
\ell(w) = \#\{\alpha \in \Phi^+ \mid w\alpha \in \Phi^-\} \quad (2.12)
\]

Lemma 2.13. Let \( w \in I W \) and write \( w = xw_J \) with \( x \in I W^J \) and \( w_J \in W_J \). Then
\[
\ell(x) = \#\{\alpha \in \Phi^+ \setminus \Phi_J \mid w\alpha \in \Phi^- \setminus \Phi_I\}.
\]
Proof. First note that $\alpha \in \Phi^+$ and $w\alpha \in \Phi^-$ already imply $w\alpha \notin \Phi_I$, because otherwise we would have $\alpha \in w^{-1}\Phi_I$, which by (2.11) is contained in $\Phi^-$ because $w^{-1} \in W^I$. Thus the right hand side of the claim is $\#\{\alpha \in \Phi^+ \setminus \Phi_J \mid w\alpha \in \Phi^-\}$. Secondly, if $\alpha \in \Phi_J^+$, using again (2.11) and $x \in W^J$ we find that $w\alpha \in \Phi^-$ if and only if $w_J\alpha \in \Phi_J^-$. Thus with (2.12) we obtain

\[
\#\{\alpha \in \Phi^+ \setminus \Phi_J \mid w\alpha \in \Phi^-\} = \#\{\alpha \in \Phi^+ \mid w\alpha \in \Phi^-\} - \#\{\alpha \in \Phi_J^+ \mid w_J\alpha \in \Phi_J^-\} = \ell(w) - \ell(w_J) = \ell(x). \]

\[\square\]

2.3 Reductive groups, Weyl groups, and parabolics

Let $G$ be a connected reductive algebraic group, let $B$ be a Borel subgroup of $G$, and let $T$ be a maximal torus of $B$. Let $\Phi(G,T)$ denote the root system of $G$ with respect to $T$, let $W(G,T) := \text{Norm}_G(T) / T$ denote the associated Weyl group, and let $S(G,B,T) \subset W(G,T)$ denote the set of simple reflections defined by $B$. Then $W(G,T)$ is a Coxeter group with respect to the subset $S(G,B,T)$.

A priori this data depends on the pair $(B,T)$. However, any other such pair $(B',T')$ is obtained on conjugating $(B,T)$ by some element $g \in G$ which is unique up to right multiplication by $T$. Thus conjugation by $g$ induces isomorphisms $\Phi(G,T) \xrightarrow{\sim} \Phi(G,T')$ and $W(G,T) \xrightarrow{\sim} W(G,T')$ and $S(G,B,T) \xrightarrow{\sim} S(G,B',T')$ that are independent of $g$. Moreover, the isomorphisms associated to any three such pairs are compatible with each other. Thus $\Phi := \Phi(G,T)$ and $W := W(G,T)$ and $S := S(G,B,T)$ for any choice of $(B,T)$ can be viewed as instances of ‘the’ root system and ‘the’ Weyl group and ‘the’ set of simple reflections of $G$, in the sense that up to unique isomorphisms they depend only on $G$. It then also makes sense to say that the result of a construction (as in Subsection 5.2 below) depending on an element of $W$ is independent of $(B,T)$.

For any $w \in W(G,T)$ we fix a representative $\dot{w} \in \text{Norm}_G(T)$. By choosing representatives attached to a Chevalley system (see [DG] Exp. XXIII, §6) for all $w_1, w_2 \in W$ with $\ell(w_1w_2) = \ell(w_1) + \ell(w_2)$ we obtain

\[
\dot{w}_1 \dot{w}_2 = (w_1w_2). \tag{2.14}
\]

In particular the identity element $1 \in W$ is represented by the identity element $1 \in G$.

A parabolic subgroup of $G$ that contains $B$ is called a standard parabolic of $G$. Any standard parabolic possesses a unique Levi decomposition $P = R_uP \times L$ with $T \subset L$. Any such $L$ is called a standard Levi subgroup of $G$, and the set $I$ of simple reflections in the Weyl group of $L$ is called the type of $L$ or of $P$. In this way there is a unique standard parabolic $P_I$ of type $I$ for every subset $I \subset S$, and vice versa. The type of a general parabolic $P$ is by definition the type of the unique standard parabolic conjugate to $P$; it is independent of $(B,T)$ in the above sense. Any conjugate of a standard Levi subgroup of $G$ is called a Levi subgroup of $G$.  

\[\text{Documenta Mathematica 16 (2011) 253–300}\]
For any subset $I \subset S$ let $\text{Par}_I$ denote the set of all parabolics of $G$ of type $I$. Then there is a natural bijection $G/P_I \cong \text{Par}_I$, $gP_I \mapsto ^gP_I$. For any two subsets $I, J \subset S$ we let $G$ act by simultaneous conjugation on $\text{Par}_I \times \text{Par}_J$. As a consequence of the Bruhat decomposition (see [Spr1] 8.4.6 (3)), the $G$-orbit of any pair $(P, Q) \in \text{Par}_I \times \text{Par}_J$ contains a unique pair of the form $(P_I, {}^xP_J)$ with $x \in \text{I}W_j$. This element $x$ is called the relative position of $P$ and $Q$ and is denoted by $\text{relpos}(P, Q)$.

We will also use several standard facts about intersections of parabolics and/or Levi subgroups, for instance (see [Car], Proposition 2.8.9):

**Proposition 2.15.** Let $L$ be a Levi subgroup of $G$ and $T$ a maximal torus of $L$. Let $P$ be a parabolic subgroup of $G$ containing $T$ and $P = R_uP \rtimes H$ its Levi decomposition with $T \subset H$. Then $L \cap P$ is a parabolic subgroup of $L$ with Levi decomposition

$$L \cap P = (L \cap R_uP) \rtimes (L \cap H).$$

If $P$ is a Borel subgroup of $G$, then $L \cap P$ is a Borel subgroup of $L$.

### 3 Connected algebraic zip data

We now define the central technical notions of this article.

**Definition 3.1.** A connected algebraic zip datum is a tuple $Z = (G, P, Q, \varphi)$ consisting of a connected reductive group $G$ with parabolic subgroups $P$ and $Q$ and an isogeny $\varphi : P/R_uP \to Q/R_uQ$. The group

$$E_Z := \{(p, q) \in P \times Q \mid \varphi(\pi_P(p)) = \pi_Q(q)\}$$

(3.2)

is called the zip group associated to $Z$. It acts on $G$ by restriction of the left action

$$((p, q), g) \mapsto pgg^{-1}.$$  

(3.3)

For any subset $X \subset G$ we denote the union of the $E_Z$-orbits of all elements of $X$ by

$$o_Z(X).$$  

(3.4)

Note that if $X$ is a constructible subset of $G$, then so is $o_Z(X)$.

Throughout the following sections we fix a connected algebraic zip datum $Z = (G, P, Q, \varphi)$. We also abbreviate $U := R_uP$ and $V := R_uQ$, so that $\varphi$ is an isogeny $P/U \to Q/V$. Our aim is to study the orbit structure of the action of $E_Z$ on $G$.

**Example 3.5.** For dimension reasons we have $P = G$ if and only if $Q = G$. In that case the action of $E_Z = \text{graph}(\varphi)$ is equivalent to the action of $G$ on itself by twisted conjugation $(h, g) \mapsto hq\varphi(h)^{-1}$.

In order to work with $Z$ it is convenient to fix the following data.
Definition 3.6. A frame of $Z$ is a tuple $(B, T, g)$ consisting of a Borel subgroup $B$ of $G$, a maximal torus $T$ of $B$, and an element $g \in G$, such that

(a) $B \subset Q$,
(b) $^g B \subset P$,
(c) $\varphi(\pi_p(^g B)) = \pi_Q(B)$, and
(d) $\varphi(\pi_p(^g T)) = \pi_Q(T)$.

Proposition 3.7. Every connected algebraic zip datum possesses a frame.

Proof. Choose a Borel subgroup $B$ of $Q$ and a maximal torus $T$ of $B$. Let $T' \subset B' \subset P/U$ denote the respective identity components of $\varphi^{-1}(\pi_Q(T)) \subset \varphi^{-1}(\pi_Q(B)) \subset P/U$. Then $B'$ is a Borel subgroup of $P/U$, and $T'$ is a maximal torus of $B'$. Thus we have $B' = \pi_B(B')$ for a Borel subgroup $B'$ of $P$, and $T' = \pi_B(T')$ for some maximal torus $T'$ of $B'$. Finally take $g \in G$ such that $B' = ^g B$ and $T' = ^g T$. Then $(B, T, g)$ is a frame of $Z$.

Proposition 3.8. Let $(B, T, g)$ be a frame of $Z$. Then every frame of $Z$ has the form $(^q B, ^q T, pg t q^{-1})$ for $(p, q) \in E_Z$ and $t \in T$, and every tuple of this form is a frame of $Z$.

Proof. Let $(B', T', g')$ be another frame of $Z$. Since all Borel subgroups of $Q$ are conjugate, we have $B' = ^g B$ for some element $q \in Q$. Since all maximal tori of $B'$ are conjugate, after multiplying $q$ on the left by an element of $B'$ we may in addition assume that $T' = ^g T$. Similarly we can find an element $p \in P$ such that $^g B' = ^p g B$ and $^g T' = ^p g T$. Combining these equations with the defining properties of frames we find that

$$\varphi(\pi_p(p)) \pi_Q(B) = \varphi(\pi_p(p)) \varphi(\pi_p(^g B)) = \varphi(\pi_p(^p g B)) = \varphi(\pi_p(^p g T')) = \pi_Q(B') = \pi_Q(^p g B) = \pi_Q(q) \pi_Q(B),$$

and similarly $\varphi(\pi_p(p)) \pi_Q(T) = \pi_Q(q) \pi_Q(T)$. Thus $\varphi(\pi_p(p)) = \pi_Q(q) \cdot \pi_Q(t')$ for some element $t' \in T$. Since we may still replace $q$ by $g' t'$ without changing the above equations, we may without loss of generality assume that $\varphi(\pi_p(p)) = \pi_Q(q)$, so that $(p, q) \in E_Z$. On the other hand, the above equations imply that $B = g^{-1} p^{-1} g'^{-1} B$ and $T = g^{-1} p^{-1} g'^{-1} T$, so that $t := g^{-1} p^{-1} g' q \in T$ and hence $g' = pg t q^{-1}$. This proves the first assertion. The second involves a straightforward calculation that is left to the conscientious reader.

Throughout the following sections we fix a frame $(B, T, g)$ of $Z$. It determines unique Levi components $^g T \subset L \subset P$ and $T \subset M \subset Q$. Via the isomorphisms $L \cong P/U$ and $M \cong Q/V$ we can then identify $\varphi$ with an isogeny $\varphi: L \rightarrow M$. The zip group then becomes

$$E_Z = \{(u\ell, v\varphi(\ell)) \mid u \in U, v \in V, \ell \in L\} \quad (3.9)$$
and acts on \( G \) by \((u\ell , v\varphi (\ell )), g) \mapsto ulg\varphi (\ell )^{-1}v^{-1}\). Moreover, conditions 3.6 (c) and (d) are then equivalent to
\[
\varphi (gB \cap L) = B \cap M, \quad \text{and} \quad \varphi (gT) = T, \tag{3.10}
\]
which are a Borel subgroup and a maximal torus of \( M \), respectively.

Let \( \Phi \) be the root system, \( W \) the Weyl group, and \( S \subset W \) the set of simple reflections of \( G \) with respect to \((B, T)\). Let \( I \subset S \) be the type of \( g^{-1}P \) and \( J \subset S \) the type of \( Q \). Then \( M \) has root system \( \Phi_J \), Weyl group \( W_J \), and set of simple reflections \( J \subset W_J \). Similarly \( g^{-1}L \) has root system \( \Phi_I \), Weyl group \( W_I \), and set of simple reflections \( I \subset W_I \), and the inner automorphism \( \text{int}(g) \) identifies these with the corresponding objects associated to \( L \). Moreover, the equations (3.10) imply that \( \varphi \circ \text{int}(g) \) induces an isomorphism of Coxeter systems
\[
\psi: (W_I, I) \sim (W_J, J). \tag{3.11}
\]

Recall that \( \Phi, W, \) and \( S \) can be viewed as independent of the chosen frame, as explained in Subsection 2.3.

**Proposition 3.11.** The subsets \( I, J \) and the isomorphism \( \psi \) are independent of the frame.

**Proof.** Consider another frame \((qB, qT, pgqt^{-1})\) with \((p, q) \in E_Z \) and \( t \in T \), as in Proposition 3.8. Then we have a commutative diagram
\[
\begin{array}{c}
(g^{-1}L, B, T) \xrightarrow{\text{int}(g)} (L, gB, gT) \xrightarrow{\varphi} (M, B, T) \\
\text{int}(qt^{-1}) \downarrow \sim \downarrow \text{int}(p) \quad \text{int}(p) \downarrow \sim \downarrow \text{int}(q) \downarrow \\
(qg^{-1}L, qB, qT) \xrightarrow{\text{int}(pgqt^{-1})} (pL, pgB, pgT) \xrightarrow{\varphi} (qM, qB, qT),
\end{array}
\]
whose upper row contains the data inducing \( \psi \) for the old frame and whose lower row is the analogue for the new frame. Since the vertical arrows are inner automorphisms, they induce the identity on the abstract Coxeter system \((W, S)\) of \( G \) as explained in Subsection 2.3. Everything follows from this.

## 4 Induction step

We keep the notations of the preceding section. Since \( g^{-1}P \) and \( Q \) are parabolic subgroups containing the same Borel \( B \), by Bruhat (see [Spr1] 8.4.6 (3)) we have a disjoint decomposition
\[
G = \coprod_{x \in W^J} g^{-1}P \hat{x}Q.
\]
Left translation by \( g \) turns this into a disjoint decomposition
\[
G = \coprod_{x \in W^J} Pg \hat{x}Q. \tag{4.1}
\]
Here each component $Pg\dot{x}Q$ is an irreducible locally closed subvariety of $G$ that is invariant under the action of $E_Z$. In this section we fix an element $x \in \mathcal{W}^J$ and establish a bijection between the $E_Z$-orbits in $Pg\dot{x}Q$ and the orbits of another zip datum constructed from $Z$ and $\dot{x}$. This will allow us to prove facts about the orbits inductively. The base case of the induction occurs when the decomposition possesses just one piece, i.e., when $P = Q \cong G$.

**Lemma 4.2.** The stabilizer of $g\dot{x}Q \subset Pg\dot{x}Q$ in $E_Z$ is the subgroup

$$E_{Z,\dot{x}} := \{(p, q) \in E_Z \mid p \in P \cap g\dot{x}Q\},$$

and the action of $E_Z$ induces an $E_Z$-equivariant isomorphism

$$E_Z \times E_{Z,\dot{x}} \xrightarrow{\sim} Pg\dot{x}Q, \quad [(p, q), h] \mapsto phq^{-1}.$$

**Proof.** The action (3.3) of $(\dot{x})$ on $E_Z$ on $Pg\dot{x}Q$ induces the action on the quotient $Pg\dot{x}Q/Q$ by left multiplication with $p$. From (3.2) we see that the latter action is transitive, and the stabilizer of the point $g\dot{x}Q$ is $E_{Z,\dot{x}}$; hence there is an $E_Z$-equivariant isomorphism $Pg\dot{x}Q/Q \cong E_Z/E_{Z,\dot{x}}$. Thus everything follows by applying Lemma 2.2 to the projection morphism $Pg\dot{x}Q \to Pg\dot{x}Q/Q \cong E_Z/E_{Z,\dot{x}}$.

**Construction 4.3.** Consider the following subgroups of the connected reductive algebraic group $M$ (which are independent of the representative $\dot{x}$ of $x$):

$$P_x := M \cap \dot{x}^{-1}g^{-1}P, \quad Q_x := \varphi(L \cap g\dot{x}Q),$$
$$U_x := M \cap \dot{x}^{-1}g^{-1}U, \quad V_x := \varphi(L \cap g\dot{x}V),$$
$$L_x := M \cap \dot{x}^{-1}g^{-1}L, \quad M_x := \varphi(L \cap g\dot{x}M).$$

Proposition 2.15 shows that $P_x$ is a parabolic with unipotent radical $U_x$ and Levi component $L_x$, and that $Q_x$ is a parabolic with unipotent radical $V_x$ and Levi component $M_x$. Moreover, $\varphi \circ \text{int}(g\dot{x})$ induces an isogeny $\varphi_\dot{x} : L_x \to M_x$, or equivalently $P_x/U_x \to Q_x/V_x$. Thus we obtain a connected algebraic zip datum $Z_x := (M, P_x, Q_x, \varphi_\dot{x})$. By (3.9) its zip group is

$$E_{Z_x} = \{(u', v', \varphi_\dot{x}(\ell')) \mid u' \in U_x, v' \in V_x, \ell' \in L_x\}. \quad (4.4)$$

**Lemma 4.5.** There is a surjective homomorphism

$$E_{Z,\dot{x}} \twoheadrightarrow E_{Z_x}, \quad (p, q) \mapsto (m, \varphi(\ell)),

where $p = u\ell$ for $u \in U$ and $\ell \in L$, and $\dot{x}^{-1}g^{-1}p = vm$ for $v \in V$ and $m \in M$.

**Proof.** For ease of notation abbreviate $h := g\dot{x}$, so that $h^{-1}T = gT \subset L$ and therefore $T \subset h^{-1}L \subset h^{-1}P$. Thus $h^{-1}P$ and $Q$ are parabolics of $G$ with $T$-invariant Levi decompositions $h^{-1}P = h^{-1}U \times h^{-1}L$ and $Q = V \times M$. It follows...
(see [Car] Thm. 2.8.7) that any element of $h^{-1}P \cap Q$ can be written as a product $abu\ell'$ with unique $a \in h^{-1}U \cap V$, $b \in h^{-1}L \cap V$, $\ell' \in h^{-1}L \cap M = L_x$.

Consider $(p, q) \in E_{Z, x}$ with $p = ul$ and $h^{-1}p = vm$ as in the lemma. Then we can write the element $h^{-1}p = abu\ell' \in h^{-1}P \cap Q$ in the indicated fashion. Comparing the different factorizations yields the equations $v = ab, m = u\ell'$, $u = h(abu'b^{-1})$, and $\ell = h(b\ell')$. Thus $\varphi(\ell) = \varphi(h\ell') \varphi(u\ell') = v'\varphi_z(\ell')$ with $v' := \varphi(b) \in \varphi(L \cap hV) = V_x$. In view of (4.4) it follows that $(m, \varphi(\ell)) = (u\ell', v'\varphi_z(\ell'))$ lies in $E_{Z, x}$, and so the map in question is well-defined. Since $m$ and $\ell$ are obtained from $p$ by projection to Levi components, the map is a homomorphism. Conversely, every element of $E_{Z, x}$ can be obtained in this way from some element $p \in P \cap hQ$. By (3.9) we can then also find $q \in Q$ with $(p, q) \in E_{Z, x}$. Thus the map is surjective, and we are done.

**Lemma 4.6.** The surjective morphism

$$\pi: g\check{x}Q \rightarrow M, \quad g\check{m}\check{v} \mapsto \check{m}$$

for $\check{m} \in M$ and $\check{v} \in V$ is equivariant under the group $E_{Z, x}$, which acts on $g\check{x}Q$ as in Lemma 4.2 and on $M$ through the homomorphism from Lemma 4.5.

**Proof.** Take $(p, q) \in E_{Z, x}$ with $p = ul$ and $h^{-1}p = vm$ as in Lemma 4.5. Then (3.9) implies that $q = v_1\varphi(\ell)$ for some $v_1 \in V$. Thus the action of $(p, q)$ sends $g\check{m}\check{v} \in g\check{x}Q$ to the element

$$pg\check{x} \cdot \check{m}\check{v} \cdot q^{-1} = g\check{x}vm \cdot \check{m}\check{v} \cdot \varphi(\ell)^{-1}v_1^{-1} = g\check{x} \cdot \check{m}\check{v}\varphi(\ell)^{-1} \cdot (\text{an element of } V).$$

The morphism $\pi$ maps this element to $\check{m}\varphi(\ell)^{-1} \in M$. But this is also the image of $\check{m} = \pi(g\check{x}\check{m}\check{v})$ under the action of $(m, \varphi(\ell)) \in E_{Z, x}$. Thus the morphism is equivariant.

**Proposition 4.7.** There is a closure-preserving bijection between $E_{Z, x}$-invariant subsets $Y \subset M$ and $E_{Z, x}$-invariant subsets $X \subset Pg\check{x}Q$, defined by $Y = M \cap \check{x}^{-1}g^{-1}X$ and $X = o_Z(g\check{x}Y)$. Moreover, $Y$ is a subvariety if and only if $X$ is one, and in that case $X \cong E_{Z, x} \times E_{Z, x}^{-1}(Y)$.

**Proof.** From (3.2) and (3.3) we see that the subgroup $V \cong \{(1, v) \mid v \in V\} \subset E_{Z, x}$ acts by right translation on $g\check{x}Q$. Thus every $E_{Z, x}$-invariant subset of $g\check{x}Q$ is a union of cosets of $V$ and therefore of the form $Z = g\check{x}Q = \pi^{-1}(Y)$ for a subset $Y \subset M$, which moreover satisfies $Y = M \cap \check{x}^{-1}g^{-1}Z$. By Lemmas 4.5 and 4.6 this defines a bijection between $E_{Z, x}$-invariant subsets $Y \subset M$ and $E_{Z, x}$-invariant subsets $Z \subset g\check{x}Q$. On the other hand, Lemmas 2.1 and 4.2 yield a bijection between $E_{Z, x}$-invariant subsets $Z \subset g\check{x}Q$ and $E_{Z, x}$-invariant subsets $X \subset Pg\check{x}Q$ that is characterized by $Z = g\check{x}Q \cap X$ and $X = o_Z(Z)$. Together we
obtain the desired bijection with \( Y = M \cap \hat{x}^{-1}g^{-1}(g\hat{x}Q \cap X) = M \cap \hat{x}^{-1}g^{-1}X \) and \( X = o_Z(g\hat{x}YV) = o_Z(g\hat{x}Y) \).

The equations \( Z = \pi^{-1}(Y) \) and \( Y = M \cap \hat{x}^{-1}g^{-1}Z \) imply that the bijection between \( Y \) and \( Z \) preserves closures and maps subvarieties to subvarieties. The corresponding facts for the bijection between \( Z \) and \( X \) follow from Lemma 2.1, which also implies the last statement.

**Proposition 4.8.** If \( X \) and \( Y \) in Proposition 4.7 are subvarieties, then
\[
\dim X = \dim Y + \dim P - \dim P_x + \ell(x).
\]

**Proof.** By the definition of \( E_{Z,x} \) we have
\[
\dim E_Z - \dim E_{Z,x} = \dim P - \dim(P \cap g\hat{x}Q) = \dim P - \dim P_x - \dim(P \cap g\hat{x}V).
\]

With the last statement of Proposition 4.7 this implies that
\[
\dim X = \dim Y + \dim V + \dim P - \dim P_x - \dim(P \cap g\hat{x}V).
\]

From the decomposition of \( V \) into root subgroups it follows that \( \dim V - \dim(P \cap g\hat{x}V) = \dim V - \dim(V \cap \hat{x}^{-1}g^{-1}P) \) is the cardinality of the set
\[
\{ \alpha \in \Phi^+ \setminus \Phi_J \mid x\alpha \in \Phi^- \setminus \Phi_I \}.
\]

By Lemma 2.13 for \( w_J = 1 \) this cardinality is \( \ell(x) \).

**Lemma 4.9.** For any subset \( Y \subset M \) we have \( o_Z(g\hat{x}o_Z(Y)) = o_Z(g\hat{x}Y) \).

**Proof.** It suffices to show that \( g\hat{x}o_Z(Y) \subset o_Z(g\hat{x}Y) \), which follows from a straightforward calculation that is left to the reader. Alternatively the formula can be deduced from the formal properties stated in Proposition 4.7.

We can also give an inductive description of the stabilizers of points in \( P_g\hat{x}Q \). However, this does not give the scheme-theoretic stabilizers, which may in fact be non-reduced. Likewise, the following lemma does not describe the scheme-theoretic kernel:

**Lemma 4.10.** The kernel of the homomorphism from Lemma 4.5 is \((U \cap g\hat{x}V) \times V\).

**Proof.** Let \( p = u\ell \) and \( \hat{x}^{-1}g^{-1}p = vm \) be as in Lemma 4.5. Then \((p, q)\) is in the kernel if and only if \( m = 1 \) and \( \varphi(\ell) = 1 \). The first equation is equivalent to \( p = g\hat{x}v \in g\hat{x}V \), which implies that \( \ell \) is unipotent. Being in the kernel of the isogeny \( \varphi \) is then equivalent to \( \ell = 1 \). Thus the second equation is equivalent to \( p \in U \), and the two together are equivalent to \( p \in U \cap g\hat{x}V \). By (3.9) we then have \( q \in V \), and so we are done.

**Proposition 4.11.** For any \( m \in M \) there is a short exact sequence
\[
1 \rightarrow U \cap g\hat{x}V \rightarrow \text{Stab}_{E_Z}(g\hat{x}m) \xrightarrow{4.5} \text{Stab}_{E_{Z,x}}(m) \rightarrow 1.
\]
First, the assumptions \( M \) and \( \phi \) being \( \phi \) components. We can also directly deduce that \( \phi \) is a common maximal torus of \( B \), respectively a homomorphism \( \Phi \). In the same way one shows that \( \psi \). Hence \( \Phi \) implies that \( \psi \). They must then coincide, and therefore \( \Phi \) is the isomorphism induced by \( \phi \circ \text{int}(g) \). This proves the statement about the Levi components. We can also directly deduce that \( \Phi(T) = \phi(g^T) = T \). Next, as \( T \) is a common maximal torus of \( M \) and \( \phi \), Proposition 2.15 implies that \( M \cap B \) is a Borel subgroup of \( M \). Recall that \( M \) has the root system \( \Phi_J \), so that \( M \cap B \) corresponds to the subset \( \Phi_J^+ = \Phi_J \cap \Phi^+ \). For the same reasons \( M \cap \phi^{-1}B \) is a Borel subgroup of \( M \) corresponding to the subset \( \Phi_J \cap x^{-1}\Phi^+ \). But with (2.11) the assumption \( x \in W^J \subseteq W \) implies that \( x\Phi_J^+ \subseteq \Phi^+ \), and hence \( \Phi_J^+ \subseteq \Phi_J \cap x^{-1}\Phi^+ \). Since both subsets correspond to Borel subgroups, they must then coincide, and therefore \( M \cap B = M \cap \phi^{-1}B \). With the inclusion \( g^B \subseteq P \) from (3.6) we deduce that
\[
M \cap B = M \cap \phi^{-1}B \subset M \cap \phi^{-1}g^{-1}P \overset{4.3}{=} P_x.
\]
In the same way one shows that \( L \cap g^B = L \cap g^B \), which together with \( B \subseteq Q \) implies that
\[
M \cap B \overset{(3.10)}{=} \phi(L \cap g^B) = \varphi(L \cap g^B) \subseteq \phi(L \cap g^B) \overset{4.3}{=} Q_x.
\]
The equation \( M \cap B = \varphi(L \cap g^B) \) and Construction 4.3 also imply that
\[
\varphi_z((M \cap B) \cap L_x) = \varphi(g^M \cap g^B \cap L) \subset \varphi(L \cap g^M) \cap \varphi(L \cap g^B) = (M \cap B) \cap M_x.
\]
As both sides of this inclusion are Borel subgroups of \( M_x \), they must be equal. Thus \((M \cap B, T, 1)\) satisfies Definition 3.6 in the variant (3.10), as desired. 

Recall that \( M \) has the Weyl group \( W_J \) with the set of simple reflections \( J \), and that \( \psi: W_T \overset{\sim}{\to} W_J \) is the isomorphism induced by \( \phi \circ \text{int}(g) \).

\[\text{Proof.}\] The second half of Lemma 4.2 and Lemma 4.6 imply that we have an equality, respectively a homomorphism
\[
\text{Stab}_{E_x}(g^m) = \text{Stab}_{E_x}(g^m) \overset{4.5}{=} \text{Stab}_{E_x}(m).
\]
This homomorphism is surjective, because the subgroup \( V \triangleq \{ (1, v) \mid v \in V \} \) contained in the kernel of the surjection \( E_{\xi,x} \to E_{\xi,x} \) acts transitively on the fibers of \( \pi \). By Lemma 4.10 the kernel is the stabilizer of \( g^m \) in the group \( (U \cap g^2V) \times V \) acting by left and right translation. This stabilizer consists of \( (u, (g^m)^{-1}) \) for all \( u \in U \cap g^2V \), and we are done. \( \square \)

Finally, the assumption \( x \in W^J \) allows us to construct a frame of \( Z_\xi \):

**Proposition 4.12.** The tuple \((M \cap B, T, 1)\) is a frame of \( Z_\xi \), and the associated Levi components of \( P_x \) and \( Q_x \) are \( L_x \) and \( M_x \), respectively.

**Proof.** First, the assumptions \( T \subseteq M \) and \( g^2T = g^2T \subseteq L \) imply that \( T \subseteq M \cap g^{-1}g^{-1}L \), the latter being \( L_x \) by Construction 4.3. Together with the equation \( \varphi(\psi(T)) = T \) from (3.10) they also imply that \( T = \varphi(g^2T) \subseteq \varphi(L \cap g^2M) \), the latter being \( M_x \) by Construction 4.3. This proves the statement about the Levi components.
Proposition 4.13. (a) The type of the parabolic $P_x$ of $M$ is $I_x := J \cap \pi^{-1} I$.

(b) The type of the parabolic $Q_x$ of $M$ is $J_x := \psi(I \cap \pi J)$.

(c) The isomorphism $\psi_x : W_{I_x} \xrightarrow{\sim} W_{J_x}$ induced by $\varphi_x$ is the restriction of $\psi \circ \text{int}(x)$.

Proof. Proposition 2.6 implies that $L_x = M \cap \pi^{-1} g^{-1} L$ has the Weyl group $W_J \cap \pi^{-1} W_I = W_{I_x}$, which shows (a). Likewise $M_x = \varphi(L \cap g^2 M)$ has the Weyl group $\varphi(W_I \cap \pi^2 W_J) = W_{J_x}$, which implies (b). Finally, (c) follows from $\varphi_x = \varphi \circ \text{int}(g \tilde{x})$.

5 Decomposition of $G$

In this section we construct a natural decomposition of $G$ into finitely many $E_{\tilde{g}}$-invariant subvarieties $G^w$.

5.1 The Levi subgroup $H_w$

Fix an element $w \in W$. Note that we can compare any subgroup $H$ of $w^{-1} g^{-1} L$ with its image $\varphi \circ \text{int}(g \tilde{w})(H)$ in $M$, because both are subgroups of $G$. Moreover, the collection of all such $H$ satisfying $\varphi \circ \text{int}(g \tilde{w})(H) = H$ possesses a unique largest element, namely the subgroup generated by all such subgroups.

Definition 5.1. We let $H_w$ denote the unique largest subgroup of $w^{-1} g^{-1} L$ satisfying $\varphi \circ \text{int}(g \tilde{w})(H_w) = H_w$. We let $\varphi_w : H_w \to H_w$ denote the isogeny induced by $\varphi \circ \text{int}(g \tilde{w})$, and let $H_w$ act on itself from the left by the twisted conjugation $h \mapsto h^w \varphi_w(h)^{-1}$.

Remark 5.2. Since $\varphi \circ \text{int}(g \tilde{w})(T) = \varphi(\tilde{w} T) = T$ by (3.10), the defining property of $H_w$ implies that $T \subset H_w$. Thus $H_w$ does not depend on the choice of representative $\tilde{w}$ of $w$, justifying the notation $H_w$. Also, in the case that $w \in W^J$ observe that the $\varphi_w$ defined here is the restriction to $H_w$ of the isogeny $\varphi_x$ from Construction 4.3. Using the same notation for both is therefore only mildly abusive.

Example 5.3. In the case $P = Q = G$ from Example 3.5 we have $M = L = G$ and $I = J = \psi(J) = S$ and hence $W = \{1\}$ and $H_1 = G$.

To analyze $H_w$ in the general case we apply the induction step from Section 4. Let $w = x w_J$ be the decomposition from Proposition 2.8 with $x \in W^J$ and $w_J \in W^J$ for $I_x := J \cap \pi^{-1} I$. Since $W_J$ is the Weyl group of $M$, and $I_x$ is the type of the parabolic $P_x \subset M$ by Proposition 4.13 (a), we can also apply Definition 5.1 to the pair $(Z_x, w_J)$ in place of $(Z, w)$.

Lemma 5.4. The subgroup $H_w$ and the isogeny $\varphi_w$ associated to $(Z, w)$ in Definition 5.1 are equal to those associated to $(Z_x, w_J)$.
Proof. Since $\hat{w}_J \in M = \varphi(L)$, Definition 5.1 and Construction 4.3 imply that

$$H_w \subset M \cap \hat{w}_J^{-1}g^{-1}L = \hat{w}_J^{-1}(M \cap \hat{\gamma}^{-1}g^{-1}L) = \hat{w}_J^{-1}L_x$$

and that $\varphi_x \circ \int(\hat{w}_J)(H_w) = \varphi \circ \int(\hat{w}_J)(H_w) = H_w$. Since $H_w$ is the largest subgroup of $\hat{w}_J^{-1}g^{-1}L$ with this property, it is also the largest in $\hat{w}_J^{-1}L_x$. □

Remark 5.5. The preceding lemma implies that $H_w$ and $\varphi_w$ also remain the same if we repeat the induction step with $(Z, w, I)$ in place of $(Z, w)$, and so on. When the process becomes stationary, we have reached a pair consisting of a zip datum as in Example 5.3 and the Weyl group element $1$, whose underlying connected reductive group and isogeny are $H_w$ and $\varphi_w$. This induction process is the idea underlying many proofs throughout this section.

Proposition 5.6. The subgroup $H_w$ is the standard Levi subgroup of $G$ containing $T$ whose set of simple reflections is the unique largest subset $K_w$ of $w^{-1}I$ satisfying $\psi \circ \int(w)(K_w) = K_w$.

Proof. For any subset $K$ of $w^{-1}I$ the equality $\psi \circ \int(w)(K) = K$ makes sense, because both sides are subsets of $W$. The collection of all such $K$ satisfying that equality possesses a unique largest element $K_w$, namely the union of all of them. Then $K_w = \psi \circ \int(w)(K_w) \subset \psi(I) = J \subset S$, and so $K_w$ consists of simple reflections.

Let $H$ denote the standard Levi subgroup of $G$ containing $T$ with the set of simple reflections $K_w$. Then the isogeny $\varphi \circ \int(\hat{w}_J) : w^{-1}g^{-1}L \to M$ sends $T$ to itself by Remark 5.2, and the associated isomorphism of Weyl groups $\psi \circ \int(w) : w^{-1}W_J \to W_J$ sends $K_w$ to itself by construction. Together this implies that $\varphi \circ \int(\hat{w}_J)(H) = H$ and hence $H \subset H_w$.

We now prove the equality $H_w = H$ by induction on $\dim G$. In the base case $M = G$ we have $I = J = S$ and $w = 1$ and thus $K_1 = S$ and $H = G$, while $H_1 = G$ by Example 5.3; hence we are done. Otherwise write $w = xw_J$ as above. Then Lemma 5.4 and the induction hypothesis show that $H_w$ is a Levi subgroup of $M$ containing $T$ with a set of simple reflections $K \subset w^{-1}I_x$ satisfying $\varphi_x \circ \int(w_J)(K) = K$. But $w^{-1}_JL_x = w^{-1}_J(J \cap \hat{\gamma}^{-1}I) \subset w^{-1}I$ and $\varphi_x \circ \int(w_J)$ is the restriction of $\varphi \circ \int(x) \circ \int(w_J) = \psi \circ \int(w)$. By the maximality of $K_w$ we thus have $K \subset K_w$ and therefore $H_w \subset H$. Together with the earlier inequality $H \subset H_w$ we deduce that $H_w = H$, as desired. □

5.2 First description of $G^w$

Definition 5.7. For any $w \in I'W$ we set $G^w := \alpha_2(\hat{w}_JH_w)$.

Proposition 5.8. The set $G^w$ does not depend on the representative $\hat{w}$ of $w$ or the frame.
Proof. The independence of \( \dot{w} \) follows from the inclusion \( T \subset H_w \). For the rest note first that by Propositions 3.11 and 5.6 the set \( K_w \) is independent of the frame. Consider another frame \((T, p, g, q)\) for \( (p, q) \in E_Z \) and \( t \in T \), as in Proposition 3.8. Recall from Subsection 2.3 that the isomorphism \( W(G, T) \cong W(G, T) \) is induced by \( \text{int}(q) : \text{Norm}_G(T) \to \text{Norm}_G(T) \). It follows that \( w \in \mathcal{I}W \) as an element of the abstract Weyl group of \( G \) is represented by \( q \dot{w}q^{-1} \in \text{Norm}_G(T) \), and with Proposition 5.6 it follows that the Levi subgroup associated to \( w \) and the new frame is \( G_{w} \). Thus the right hand side in Definition 5.7 associated to the new frame is

\[
o_Z((pqtq^{-1})(q \dot{w}q^{-1})QH_w) = o_Z(pqt \dot{w}H_wq^{-1}) = o_Z(gt \dot{w}H_w) = o_Z(g \dot{w}H_w),
\]

where the second equation follows from \( (p, q) \in E_Z \) and the third from \( \dot{w}^{-1} t \dot{w} \in T \subset H_w \). Thus \( G^w \) is independent of the frame.

In Example 5.3 we have \( H_1 = G \) and hence \( G^1 = G \). Otherwise recall from Proposition 4.12 that \( Z_x \) has the frame \((M \cap B, T, 1)\). Thus by Lemma 5.4, the subset associated to \((Z_x, wJ)\) by Definition 5.7 is \( M^{wJ} := o_{Z_x} (wJH_w) \).

**Lemma 5.9.** Under the bijection of Proposition 4.7, the subset \( M^{wJ} \subset M \) corresponds to the subset \( G^{w} \subset PgzQ \). In particular \( G^{w} = o_Z(g \dot{z} M^{wJ}) \). Also, there is a bijection between the \( E_{Z_x} \)-orbits \( X' \subset M^{wJ} \) and the \( E_{Z_x} \)-orbits \( X \subset G^{w} \), defined by \( X = o_Z(g \dot{z} X') \).

**Proof.** Using, in this order, the definition of \( G^w \), the equation (2.14), Lemma 4.9, and the definition of \( M^{wJ} \) we find that

\[
G^w = o_Z(g \dot{z}H_w) = o_Z(g \dot{z}wJH_w) = o_Z(g \dot{z} o_{Z_x}(wJH_w)) = o_Z(g \dot{z} M^{wJ}).
\]

The other assertions follow from Proposition 4.7.

5.3 **Main properties of** \( G^w \)

**Theorem 5.10.** The \( G^w \) for all \( w \in \mathcal{I}W \) form a disjoint decomposition of \( G \).

**Proof.** We show this by induction on \( \dim G \). In the base case \( M = G \) we have \( \mathcal{I}W = \{ \} \) and \( H_1 = G = G^1 \) by Example 5.3; hence the theorem is trivially true. Otherwise take an element \( x \in \mathcal{I}W^J \). By the induction hypothesis applied to the zip datum \( E_{Z_x} \), the subsets \( M^{wJ} \) for \( wJ \in \mathcal{I}W^J \) form a disjoint decomposition of \( M \). Thus by Proposition 4.7 and Lemma 5.9, the subsets \( G^{xwJ} \) for \( xJ \in \mathcal{I}W^J \) form a disjoint decomposition of \( PgzQ \). Combining this with the Bruhat decomposition (4.1) it follows that the subsets \( G^{xwJ} \) for all \( x \) and \( wJ \) form a disjoint decomposition of \( G \). But by Proposition 2.8 these are precisely the subsets \( G^w \) for \( w \in \mathcal{I}W \), as desired.

**Theorem 5.11.** For any \( w \in \mathcal{I}W \) the subset \( G^w \) is a nonsingular subvariety of \( G \) of dimension \( \dim P + \ell(w) \).

*Documenta Mathematica 16 (2011) 253–300*
Proof. Again we proceed by induction on \( \dim G \). If \( M = G \), there is only one piece \( G^1 = G = P \) associated to \( w = 1 \), and the assertion is clear. Otherwise write \( w = xw_j \) as in Proposition 2.8. By the induction hypothesis the subset \( M_{\omega_j} \) is a nonsingular subvariety of \( M \) of dimension \( \dim P_x + \ell(w_j) \). Thus by Propositions 4.7 and 4.8 and Lemma 5.9 the subset \( G_{\omega} \) is a nonsingular subvariety of dimension

\[
[\dim P_x + \ell(w_j)] + \dim P - \dim P_x + \ell(x) = \dim P + \ell(x) + \ell(w_j).
\]

By Proposition 2.7 the last expression is equal to \( \dim P + \ell(w) \), as desired.

**Theorem 5.12.** For any \( w \in \text{I} W \), there is a bijection between the \( H_w \)-orbits \( Y \subset H_w \) and the \( E_Z \)-orbits \( X \subset G^w \), defined by \( X = o_Z(gY) \) and satisfying

\[
\text{codim}(X \subset G^w) = \text{codim}(Y \subset H_w).
\]

Proof. If \( M = G \), we have \( w = 1 \) and \( G = G^1 = H_1 \), and \( E_Z \cong G \) acts on itself by the twisted conjugation \( (h, h') \mapsto h \cdot h' \cdot \varphi (h)^{-1} \). Thus the \( E_Z \)-orbits \( X \subset G \) are precisely the cosets \( gY \) for \( H_1 \)-orbits \( Y \) according to Definition 5.1, which finishes that case.

If \( M \neq G \) write \( w = xw_j \) as in Proposition 2.8. Then \( Z_x \) has the frame \((M \cap B, T, 1)\) by Proposition 4.12, and so by Lemma 5.4 and the induction hypothesis there is a bijection between the \( H_w \)-orbits \( Y \subset H_w \) and the \( E_{Z_x} \)-orbits \( X' \subset M^{\omega_j} \), defined by \( X' = o_{Z_x}(\tilde{w}_j Y) \) and satisfying \( \text{codim}(X' \subset M^{\omega_j}) = \text{codim}(Y \subset H_w) \). By Proposition 4.7 and Lemma 5.9 there is a bijection between these \( X' \) and the \( E_Z \)-orbits \( X \subset G^w \), defined by \( X = o_Z(g\tilde{w} X') \). Moreover, since pushout and flat pullback preserve codimensions, the last statement in Proposition 4.7 implies that

\[
\text{codim}(X \subset G^w) = \text{codim}(o_Z(g\tilde{x} X') \subset o_Z(g\tilde{x} M^{\omega_j})) = \text{codim}(X' \subset M^{\omega_j}).
\]

Finally, since \( \tilde{w} = \tilde{x} \tilde{w}_j \) by (2.14), Lemma 4.9 shows that \( X = o_Z(g\tilde{x} \circ o_{Z_x}(\tilde{w}_j Y)) = o_Z(g\tilde{x} \tilde{w}_j Y) = o_Z(gY) \), finishing the induction step.

### 5.4 Other Descriptions of \( G^w \)

**Lemma 5.13.** For any element \( g' \in G \) we have

\[
o_Z(gBg') = o_Z(g' B) = o_Z(gBg').
\]

Proof. Take any element \( b \in B \). Then the condition 3.6 (b) implies that \( p := gbg^{-1} \in P \), and so there exists \( q \in Q \) such that \( (p, q) \in E_Z \). By the condition 3.6 (c) we then have \( q \in B \). It follows that \( gbg'B = pgq' Bq^{-1} \subset o_Z(gg'B) \). Since \( b \) was arbitrary, this shows that \( gBg' B \subset o_Z(gg'B) \), whence the first equality. A similar argument proves the second equality.

**Theorem 5.14.** For any \( w \in \text{I} W \) we have

\[
G^w = o_Z(g\tilde{w} H_w) = o_Z(g\tilde{w}(H_w \cap B)) = o_Z(\tilde{g}w B) = o_Z(gB\tilde{w}) = o_Z(gBwB).
\]
Proof. The first equation is Definition 5.7 of $G^w$, and the last two equations are cases of Lemma 5.13. The remaining two equations are proved by induction on $\dim G$. In the base case $M = G$ we have $w = 1$ and $H_1 = G$; hence the second term is $o_Z(gG) = G$, and the third and fourth terms are both equal to $o_Z(gB)$. By Proposition 2.5 applied to the isogeny $\varphi \circ \text{int}(g)$ the latter is equal to $G$, as desired.

In the case $M \neq G$ write $w = xw_J$ as in Proposition 2.8. Then $Z_x$ has the frame $(M \cap B, T, 1)$ by Proposition 4.12, and so by Lemma 5.4 and the induction hypothesis we have

$$o_{Z_x}(\hat{w}_J H_w) = o_{Z_x}(\hat{w}_J (H_w \cap B)) = o_{Z_x}(\hat{w}_J (M \cap B)).$$

Using Lemma 4.9 this implies that

$$o_Z(g\hat{w}_J H_w) = o_Z(g\hat{w}_J (H_w \cap B)) = o_Z(g\hat{w}_J (M \cap B)).$$

By (2.14) we may replace $\hat{xw}_J$ by $\hat{w}$ in these equations. Moreover, (3.3) and (3.9) show that $g\hat{w}B = g\hat{w}(M \cap B)V \subset o_Z(g\hat{w}(M \cap B))$ and so $o_Z(g\hat{w}B) = o_Z(g\hat{w}(M \cap B))$. Thus both equations follow.

Example 5.15. If $P$ is a Borel subgroup, then so is $Q$, and we have $\check{I}W = W$. The last equation in Theorem 5.14 then implies that $G^w = gB\hat{w}B$ for all $w \in W$.

For a further equivalent description of $G^w$ see Subsection 11.1.

6 Closure relation

In this section, we determine the closure of $G^w$ in $G$ for any $w \in \check{I}W$. To formulate a precise result recall that $\leq$ denotes the Bruhat order on $W$.

Definition 6.1. For $w, w' \in \check{I}W$ we write $w' \preceq w$ if and only if there exists $y \in W_I$ such that $gw' \psi(y)^{-1} \leq w$.

Theorem 6.2. For any $w \in \check{I}W$ we have

$$G^w = \bigcap_{w' \preceq w, w' \in \check{I}W} G^{w'}.$$

A direct consequence of this is:

Corollary 6.3. The relation $\preceq$ is a partial order on $\check{I}W$.

Remark 6.4. The relation $\preceq$ has been introduced by He in [He2] for a somewhat more special class of isomorphisms $\psi: W_I \rightarrow W_J$. He gives a direct combinatorial proof that $\preceq$ is a partial order (Proposition 3.13 of loc. cit.), which can be adapted to our more general setting (see [Wed], Section 4).
The rest of this section is devoted to proving Theorem 6.2. We will exploit the fact that the closure relation for the Bruhat decomposition of $G$ is known. Namely, for any $w \in W$ we have by [Spr1], Proposition 8.5.5:

$$BuB = \prod_{w' \leq w} Bu'B. \quad (6.5)$$

**Lemma 6.6.** For any $w \in W$ we have

$$o_Z(gBuB) = \bigcup_{w' \leq w} o_Z(gu'B).$$

**Proof.** Let $B_Z \subset E_Z$ denote the subgroup of all elements $(u\ell, v\varphi(\ell))$ with $u \in U$, $v \in V$, and $\ell \in L \cap B$. Then $E_Z/B_Z \cong L/(L \cap B)$ is proper, and $gBuB \subset G$ is a $B_Z$-invariant subvariety. Thus Lemma 2.3 and (6.5) imply that

$$o_Z(gBuB) = o_Z(gBuB) = \bigcup_{w' \leq w} o_Z(gBu'B).$$

The desired equality then follows from Lemma 5.13. \hfill \square

**Lemma 6.7.** For any $w, v \in W$ and $b \in B$ there exists $u \in W$ such that $u \leq v$ and $wu' \in Bw'$.  

**Proof.** We prove the statement by induction on $\ell(v)$. If $v = 1$, we may take $u = 1$. For the induction step write $v = s'v$ for some simple reflection $s$ such that $\ell(v') = \ell(v) - 1$. By the induction hypothesis there exists $u' \leq v'$ such that $wu' \in Bw'B$. Hence $wu' \in Bw'Bs \subset Bw'Bs \cup Bw'$, so either $u = u'$ or $u = u'$ will have the required property. \hfill \square

**Lemma 6.8.** For any $z \in W$ and $w \in W$ and $v \in W_I$ such that $z \leq wv$, there exists $y \in W_I$ such that $yz\psi(y)^{-1} \leq vw$.

**Proof.** Choose reduced expressions for $w$ and $v$ as products of simple reflections. Since $\psi(I) = J$, this also yields a reduced expression for $\psi(v)$. Together this yields an expression for $wv\psi(v)$ as a product of simple reflections, which is not necessarily reduced. However, by [BB], Theorem 2.2.2 a reduced expression for $wv\psi(v)$ can be obtained from the given one by possibly deleting some factors. By the definition of the Bruhat order, the assumption $z \leq wv\psi(v)$ means that a reduced expression for $z$ is obtained from this by deleting further factors, if any. Let $y'$ denote the product of all factors remaining from $w$. Since all factors in the reduced expression for $v$ lie in $I$, the product of all factors remaining from $\psi(v)$ is equal to $\psi(y)$ for some $y \in W_I$. By construction we then have $z = y'\psi(y)$, and so $yz\psi(y)^{-1} = yy'$. But the assumptions on $w$ and $v$ imply that $\ell(vw) = \ell(v) + \ell(w)$; hence the product of the given reduced expressions for $v$ and $w$ is a reduced expression for $vw$. By construction $yy'$ is obtained from that product by possibly deleting some factors, so we deduce that $yy' \leq vw$, as desired. \hfill \square
Lemma 6.9. For any $w \in I^t W$ and $w' \in W$ and $b, b' \in B$ such that $o_Z(gu b) = o_Z(gu' b')$ there exists $y \in W_I$ such that $yw\psi(y)^{-1} \leq w'$. 

Proof. We proceed by induction on $\dim G$. In the base case $M = G$ we have $w = 1$ and may take $y = 1$. So assume that $M \neq G$. Write $w = xw_J$ as in Proposition 2.8 with $x \in I^t W^J$ and $w_J \in I^t W_J$. From $o_Z(gu b) = o_Z(gu' b')$ we deduce that $Pg\tilde{Q} = Pg\tilde{Q} = Pg\tilde{Q}$, which in view of (4.1) implies that $w' \in W_J x W_J$. Write $w' = v' x w'_J$ for $v' \in W_I$ and $w'_J \in I^t W_J$, as in Proposition 2.7.

Recall that $\varphi(gu' g^{-1}) \in \text{Norm}_M(T)$ is a representative of $\psi(v') \in W_J$. Thus by Lemma 6.7, there exists $u \in W$ such that $u \leq \psi(v')$ and $\hat{x}u'_J b' \varphi(gu' g^{-1}) \in B\hat{x}u'_J u B$. The first condition implies that $u \in W_J$, the Weyl group of $M$. The action of $E_Z$ and the second condition imply

$$o_Z(gu b) = o_Z(gu' b') = o_Z(gu' b') = o_Z(gu' b').$$

By the induction hypothesis there therefore exists $y_z \in W_I$ such that

$$y_z w_J \psi_z(y_z)^{-1} \leq w'_J u.$$ 

Now we work our way back up. Since both sides of the last relation lie in $W_J$, and since $x \in W_J$, we deduce that

$$z := xy z w_J \psi_z(y_z)^{-1} \leq x w'_J u.$$ 

Recall that $u \leq \psi(v')$, which implies that $u = \psi(u')$ for some $u' \in W_I$ satisfying $u' \leq v'$. Also, note that $xw'_J \in I^t W$ by Proposition 2.8. Thus by Lemma 6.8 there exists $y' \in W_I$ such that

$$y' z \psi(y')^{-1} \leq u' x w'_J.$$ 

As $u'$ and $v'$ lie in $W_I$, and $xw'_J \in I^t W$, we deduce that

$$y' z \psi(y')^{-1} \leq u' x w'_J \leq v' x w'_J = w'.$$

Finally, since $\psi_z = \psi \circ \text{int}(x)$, we have

$$y' z \psi(y')^{-1} = y' x y z w_J \psi(x y z x^{-1})^{-1} \psi(y')^{-1} = (y' x y z x^{-1}) x w_J \psi(y' x y z x^{-1})^{-1} = y w \psi(y)^{-1}$$

with $y := y' x y z x^{-1} \in W_I$. Thus $yw \psi(y)^{-1} \leq w'$, as desired. 

\hfill $\Box$
Lemma 6.10. For any $w \in I W$, the set $o_Z(g\dot wT)$ is dense in $G^w$.

Proof. Theorem 5.12 implies that $o_Z(g\dot wT) = o_Z(g\dot wY)$, where $Y \subset H_w$ is the orbit of $T$ under twisted conjugation by $H_w$. But Proposition 2.5 (b) asserts that $Y$ is dense in $H_w$. Thus $o_Z(g\dot wY)$ is dense in $o_Z(g\dot wH_w) = G^w$, as desired.

Proof of Theorem 6.2. Consider $w' \in I W$ such that $G^{v'} \cap \overline{G^w} \neq \emptyset$. Then by Theorem 5.14 and Lemma 6.6 there exist $b, b' \in B$ and $w'' \leq w$ and $o_Z(g\dot w'b) = o_Z(g\dot w'b')$. Lemma 6.9 then implies that $yw\psi(y)^{-1} \leq w''$ for some $y \in W_I$. Together it follows that $yw\psi(y)^{-1} \leq w$, and hence $w' \preceq w$, proving "$\supset$".

To prove "$\supset$" consider $w' \in I W$ with $w' \preceq w$. By definition there exists $y \in W_I$ such that $w'' := yw\psi(y)^{-1} \leq w$. Lemma 6.6 and Theorem 5.14 then show that $o_Z(g\dot w''T) \subset \overline{G^w}$. Therefore

$$o_Z(g\dot w'T) = o_Z(g\dot w'T\varphi(g\dot yg^{-1})^{-1}) =$$

$$= o_Z(g\dot w'\varphi(g\dot yg^{-1})^{-1}T) = o_Z(g\dot w''T) \subset \overline{G^w}.$$ With Lemma 6.10 for $o_Z(g\dot w'T)$ we conclude that $G^{v'} \subset \overline{G^w}$, as desired.

7 Orbitally finite zip data

Proposition 7.1. The following assertions are equivalent:

(a) For any $w \in I W$, the number of fixed points of the endomorphism $\varphi_w = \varphi \circ \text{int}(g\dot w)$ of $H_w$ from Definition 5.1 is finite.

(b) For any $w \in I W$ the $E_Z$-invariant subvariety $G^w$ is a single orbit under $E_Z$.

(c) The number of orbits of $E_Z$ on $G$ is finite.

Proof. If (a) holds, the Lang-Steinberg Theorem 2.4 shows that the orbit of $1 \in H_w$ under twisted conjugation is all of $H_w$, and by Theorem 5.12 this implies (b). The implication (b)$\Rightarrow$(c) is trivial. So assume (c). Then again by Theorem 5.12, the number of orbits in $H_w$ under twisted conjugation by $\varphi_w$ is finite for any $w \in I W$. In particular there exists an open orbit; let $h$ be an element thereof. Then for dimension reasons its stabilizer is finite. But

$$\text{Stab}_{H_w}(h) = \{ h' \in H_w \mid h'h\varphi_w(h')^{-1} = h \} = \{ h' \in H_w \mid h' = h\varphi_w(h')h^{-1} \}$$

is also the set of fixed points of the endomorphism $\text{int}(h) \circ \varphi_w$ of $H_w$. Thus the Lang-Steinberg Theorem 2.4 implies that $\{ h'h\varphi_w(h')^{-1}h^{-1} \mid h' \in H_w \} = H_w$. After right multiplication by $h$ this shows that the orbit of $h$ is all of $H_w$. We may thus repeat the argument with the identity element in place of $h$, and deduce that the set of fixed points of $\varphi_w$ on $H_w$ is finite, proving (a).
Definition 7.2. We call $Z$ orbitally finite if the conditions in Proposition 7.1 are met.

Proposition 7.3. If the differential of $\varphi$ at 1 vanishes, then $Z$ is orbitally finite.

Proof. If the differential of $\varphi$ vanishes, then so does the differential of $\varphi_\dot{w} = \varphi \circ \operatorname{int}(gw) \vert H_w$ for any $w \in I^W$. Let $H^1_w$ denote the fixed point locus of $\varphi_\dot{w}$, which is a closed algebraic subgroup. Then the restriction $\varphi_\dot{w} \vert H^1_w$ is the identity and its differential is zero. This is possible only when $\dim H^1_w = 0$, that is, when $H^1_w$ is finite. \hfill $\square$

Remark 7.4. In particular Proposition 7.3 applies when the base field has characteristic $p > 0$ and the isogeny $\varphi$ is a relative Frobenius $L \to L^{(p)} \simeq M$.

Since $gw \in G^w$ by Definition 5.7, we can now rephrase condition 7.1 (b) and Theorems 5.10, 5.11, and 6.2 as follows:

Theorem 7.5. Assume that $Z$ is orbitally finite. Then:

(a) For any $w \in I^W$ we have $G^w = \alpha_Z(gw)$.

(b) The elements $gw$ for $w \in I^W$ form a set of representatives for the $E_Z$-orbits in $G$.

(c) For any $w \in I^W$ the orbit $\alpha_Z(gw)$ has dimension $\dim P + \ell(w)$.

(d) For any $w \in I^W$ the closure of $\alpha_Z(gw)$ is the union of $\alpha_Z(gw')$ for all $w' \in I^W$ with $w' \lesssim w$.

8 Point stabilizers

In this section we study the stabilizer in $E_Z$ of an arbitrary element $g' \in G$. Take $w \in I^W$ such that $g' \in G^w$. Then Theorem 5.12 shows that $g'$ is conjugate to $gw h$ for some $h \in H_w$. Thus it suffices to consider the stabilizer of $gw h$.

Recall from Definition 5.1 that $H_w$ acts on itself by twisted conjugation with the isogeny $\varphi_\dot{w}$, which is defined as the restriction of $\varphi \circ \operatorname{int}(gw)$.

Theorem 8.1. For any $w \in I^W$ and $h \in H_w$ the stabilizer $\operatorname{Stab}_{E_Z}(gw h)$ is the semi-direct product of a connected unipotent normal subgroup with the subgroup

$$\{(\operatorname{int}(gw)(h'), \varphi(\operatorname{int}(gw)(h')) \mid h' \in \operatorname{Stab}_{H_w}(h)\}. \quad (8.2)$$

Proof. For any $h' \in H_w$ we have $\operatorname{int}(gw)(h') \cdot gw h \cdot \varphi(\operatorname{int}(gw)(h'))^{-1} = gw h$ if and only if $h' \varphi(\operatorname{int}(gw)(h'))^{-1} = h$ if and only if $h' \in \operatorname{Stab}_{H_w}(h)$. This implies that (8.2) is a subgroup of $\operatorname{Stab}_{E_Z}(gw h)$.

For the rest we proceed by induction on $\dim G$. If $M = G$, we have $w = 1$ and $\dot{w} = 1$ and $G = H_1$, and $(g', \varphi(g')) \in E_Z$ acts on $G$ by the twisted conjugation $g' \mapsto g' g'' \varphi(g')^{-1}$. Under left translation by $g$ this corresponds to the action...
of $H_w$ on itself, so that $\text{Stab}_{E_2}(gw\hat{h})$ is precisely the subgroup (8.2) and the normal subgroup is trivial. If $M \neq G$ write $w = xw_J$ as in Proposition 2.8. Then $\hat{w} = \hat{x}w_J$ and $Z_{\hat{h}}$ has the frame $(M \cap B, T, 1)$, and Proposition 4.11 shows that $\text{Stab}_{E_2}(gw\hat{h})$ is an extension of $\text{Stab}_{E_2}(\hat{w}_J,h)$ by a connected unipotent normal subgroup. Moreover, by the induction hypothesis $\text{Stab}_{E_2}(\hat{w}_J,h)$ is the semi-direct product of a connected unipotent normal subgroup with the subgroup

$$\{ (\text{int}(\hat{w}_J)(h'), \varphi_{x}(\text{int}(\hat{w}_J)(h')) ) \mid h' \in \text{Stab}_{H_w}(h) \}. \quad (8.3)$$

Furthermore a direct calculation shows that the projection in Proposition 4.11 sends the subgroup (8.2) isomorphically to the subgroup (8.3). Since any extension of connected unipotent groups is again connected unipotent, the theorem follows.

**Remark 8.4.** For the stabilizer, a similar result was obtained by Evens and Lu ([EL] Theorem 3.13).

If the differential of $\varphi$ at 1 vanishes, we can also describe the infinitesimal stabilizer in the Lie algebra. Since in that case the zip datum is orbitally finite by Proposition 7.3, it suffices to consider the stabilizer of $gw$.

**Theorem 8.5.** Assume that the differential of $\varphi$ at 1 vanishes. For any $w \in \mathfrak{I}W$ let $w = xw_J$ be the decomposition from Proposition 2.8. Then the infinitesimal stabilizer of $gw$ in the Lie algebra of $E_2$ has dimension $\dim V - \ell(x)$.

**Proof.** Since $d\varphi = 0$, we have $\text{Lie}E_2 = \text{Lie}P \times \text{Lie}V \subset \text{Lie}(P \times Q)$. Thus an arbitrary tangent vector of $E_2$ at 1 has the form $(1+dp, 1+dv)$ for $dp \in \text{Lie}P$ and $dv \in \text{Lie}V$, viewed as infinitesimal elements of $P$ and $V$ in Leibniz’s sense. That element stabilizes $gw$ if and only if $(1+dp)gw(1+dv)^{-1} = gw$. This condition is equivalent to $dp \cdot gw - gw \cdot dv = 0$, or again to $dp = \text{Ad}_{gw}(dv)$. The dimension is therefore $\dim(\text{Lie}P \cap \text{Ad}_{gw}(\text{Lie}V)) = \dim(\text{Lie}g^{-1}P \cap \text{Lie}^wV)$. As both $g^{-1}P$ and $^wV$ are normalized by $T$, the dimension is just the number of root spaces in the last intersection. This number is

$$\#[(\Phi^+ \cup \Phi_I) \cap w(\Phi^+ \setminus \Phi_J)] = \#\{\alpha \in \Phi^+ \setminus \Phi_J \mid w\alpha \in \Phi^+ \cup \Phi_I\} = \dim V - \#\{\alpha \in \Phi^+ \setminus \Phi_J \mid w\alpha \in \Phi^- \setminus \Phi_I\}. $$

By Lemma 2.13 it is therefore $\dim V - \ell(x)$, as desired.

**Remark 8.6.** The dimension in Theorem 8.5 depends only on the first factor of $w = xw_J$ and thus only on the Bruhat cell $Pg_{\hat{w}}Q$. Since that Bruhat cell is an irreducible variety and in general composed of more than one $E_2$-orbit, these orbits have different dimensions. Thus the corresponding point stabilizers in $E_2$ have different dimension, while the dimension of their Lie algebra stabilizer is constant. Therefore the scheme-theoretic stabilizer of $gw$ is in general not reduced.
9 Abstract zip data

By Theorem 5.10 the subsets $G^w$ for all $w \in W$ form a disjoint decomposition of $G$ satisfying $g\psi \in G^w$. It is natural to ask which other elements of the form $g\psi'$ for $w' \in W$ are contained in a given $G^w$. When $Z$ is orbitally finite, by Theorem 7.5 this question is equivalent to asking which elements $g\psi'$ for $w' \in W$ lie in the same $E_Z$-orbit. This problem turns out to depend only on the groups $W_I \subset W$ and the homomorphism $\psi$ and can therefore be studied abstractly. We return to this situation at the end of this section.

9.1 Abstract groups

**Definition 9.1.** An abstract zip datum is a tuple $A = (\Gamma, \Delta, \psi)$ consisting of a group $\Gamma$, a subgroup $\Delta$, and a homomorphism $\psi: \Delta \to \Gamma$.

Fix such an abstract zip datum $A$. For any $\gamma \in \Gamma$, the collection of subgroups $E$ of $\gamma^{-1}\Delta$ satisfying $\psi \circ \text{int}(\gamma)(E) = E$ possesses a unique largest element, namely the subgroup generated by all such subgroups.

**Definition 9.2.** For any $\gamma \in \Gamma$ we let $E_\gamma$ denote the unique largest subgroup of $\gamma^{-1}\Delta$ satisfying $\psi \circ \text{int}(\gamma)(E_\gamma) = E_\gamma$.

**Lemma 9.3.** For any $\gamma \in \Gamma$ and $\delta \in \Delta$ and $\varepsilon \in E_\gamma$, we have $E_{\delta\gamma\varepsilon(\psi)^{-1}} = \psi(\delta)E_\gamma$.

**Proof.** Abbreviate $\gamma' := \delta\gamma\varepsilon(\psi)^{-1}$. Then the calculation $\psi(\gamma'(\psi(\delta)E_\gamma)) = \psi(\delta\gamma\varepsilon(\psi)^{-1}) = \psi(\delta\gamma\varepsilon) = \gamma\varepsilon = \psi(\delta)E_\gamma$ and the definition of $E_{\gamma'}$ imply that $\psi(\delta)E_\gamma \subset E_{\gamma'}$. In particular, $\varepsilon' := \psi(\delta)^{-1}$ is an element of $E_{\gamma'}$. Since $\gamma = \delta\gamma'\varepsilon'\psi(\delta')^{-1}$ with $\delta' := \delta^{-1} \in \Delta$, a calculation like the first shows that $\psi(\delta')E_{\gamma'} \subset E_{\gamma'}$. Together it follows that $E_{\gamma'} = \psi(\delta)E_\gamma$, as desired.

**Definition 9.4.** For any $\gamma, \gamma' \in \Gamma$ we write $\gamma' \sim \gamma$ if and only if there exist $\delta \in \Delta$ and $\varepsilon \in E_\gamma$ such that $\gamma' = \delta\gamma\varepsilon\psi(\delta)^{-1}$. For any $\gamma \in \Gamma$ we abbreviate $o_A(\gamma) := \{\gamma' \in \Gamma \mid \gamma' \sim \gamma\}$.

**Lemma 9.5.** This is an equivalence relation.

**Proof.** Reflexivity is clear, and symmetry was shown already in the proof of Lemma 9.3. To prove transitivity, suppose that $\gamma' = \delta\gamma\varepsilon\psi(\delta)^{-1}$ for $\delta \in \Delta$ and $\varepsilon \in E_\gamma$ and $\gamma'' = \delta'\gamma'\varepsilon'\psi(\delta')^{-1}$ for $\delta' \in \Delta$ and $\varepsilon' \in E_{\gamma'}$. Then $\psi(\delta)^{-1}\varepsilon' \in E_{\gamma'}$ by Lemma 9.3, and so $\gamma'' = \delta'\gamma\varepsilon\psi(\delta)^{-1}\varepsilon'\psi(\delta')^{-1} = \delta''\gamma\varepsilon''\psi(\delta'')^{-1}$ for $\delta'' := \delta\delta' \in \Delta$ and $\varepsilon'' := \varepsilon \psi(\delta)^{-1}\varepsilon' \in E_\gamma$, as desired.

**Theorem 9.6.** If $\Delta$ is finite, each equivalence class in $\Gamma$ has cardinality $\#\Delta$ and the number of equivalence classes is $[\Gamma : \Delta]$. 
Proof. Take any $\gamma \in \Gamma$; then the group $E_\gamma \subset \gamma^{-1}\Delta$ is finite, too. Consider the surjective map $\Delta \times E_\gamma \rightarrow o_\mathcal{A}(\gamma)$, $(\delta, \varepsilon) \mapsto \delta \gamma \varepsilon \psi(\delta)^{-1}$. Two elements $(\delta, \varepsilon)$, $(\delta', \varepsilon') \in \Delta \times E_\gamma$ lie in the same fiber if and only if $\varepsilon \psi(\delta) = \varepsilon' \psi(\delta')^{-1}$, and then only if $\varepsilon \psi(\delta^{-1}) = \gamma^{-1}(\delta^{-1}) \gamma \varepsilon'$. With $\varepsilon'' := \gamma^{-1}(\delta^{-1}) \gamma \varepsilon'$, this is equivalent to $\varepsilon \psi(\gamma \varepsilon'') = \varepsilon'' \varepsilon'$. Since $\varepsilon, \varepsilon' \in E_\gamma$, this equation implies that the subgroup generated by $E_\gamma$ and $\varepsilon''$ is mapped onto itself under $\psi \circ \text{int}(\gamma)$. By maximality it is therefore equal to $E_\gamma$, and so $\varepsilon'' \in E_\gamma$. Together we find that the elements in the same fiber as $(\delta, \varepsilon)$ are precisely the elements $(\delta', \varepsilon')$ with $\delta' = \delta \gamma \varepsilon''$ and $\varepsilon' = (\varepsilon'')^{-1} \varepsilon \psi(\gamma \varepsilon'')$ for some $\varepsilon'' \in E_\gamma$. Thus each fiber has cardinality $\#E_\gamma$, and so the image has cardinality $\#\Delta$, proving the first assertion. The second assertion is a direct consequence of the first.

We can also perform an induction step as in Section 4 for abstract zip data, obtaining analogues of Lemma 5.4 and Proposition 4.7. For this fix an element $\xi \in \Gamma$, say in a set of representatives for the double quotient $\Delta \setminus \Gamma / \psi(\Delta)$. Then Definitions 9.2 and 9.4 imply that the equivalence class of any $\gamma \in \Delta \psi(\Delta)$ is again contained in $\Delta \psi(\Delta)$.

Construction 9.7. Set $\Gamma_\xi := \psi(\Delta)$ and $\Delta_\xi := \psi(\Delta) \cap \xi^{-1}\Delta$, and let $\psi_\xi : \Delta_\xi \rightarrow \Gamma_\xi$ denote the restriction of $\psi \circ \text{int}(\xi)$. This defines a new, possibly smaller, abstract zip datum

$$\mathcal{A}_\xi := (\Gamma_\xi, \Delta_\xi, \psi_\xi).$$

Lemma 9.8. For any $\gamma \in \Gamma_\xi$, the group $E_{\xi \gamma}$ associated by Definition 9.2 to the pair $(\mathcal{A}, \xi \gamma)$ is equal to the group associated to the pair $(\mathcal{A}_\xi, \gamma)$.

Proof. Since $\gamma \in \Gamma_\xi = \psi(\Delta)$, Definition 9.2 implies that

$$E_{\xi \gamma} \subset \psi(\Delta) \cap \gamma^{-1}\xi^{-1}\Delta = \gamma^{-1}(\psi(\Delta) \cap \xi^{-1}\Delta) = \gamma^{-1}\Delta_\xi$$

and that $E_{\xi \gamma} = \psi \circ \text{int}(\gamma)(E_\xi) = \psi_\xi \circ \text{int}(\gamma)(E_\xi)$. Since $E_\xi$ is the largest subgroup of $\gamma^{-1}\xi^{-1}\Delta$ with this property, it is also the largest in $\gamma^{-1}\Delta_\xi$.

Proposition 9.9. There is a bijection between $\mathcal{A}_\xi$-equivalence classes in $\Gamma_\xi$ and $\mathcal{A}$-equivalence classes in $\Delta \psi(\Delta)$, defined by $o_{\mathcal{A}_\xi}(\gamma) \mapsto o_{\mathcal{A}}(\xi \gamma)$ and $o_{\mathcal{A}_\xi}(\gamma) = \Gamma_\xi \cap \xi^{-1}o_{\mathcal{A}}(\xi \gamma)$.

Proof. Take any $\gamma, \gamma' \in \Gamma_\xi$. Then $\gamma' \in \xi^{-1}o_{\mathcal{A}}(\xi \gamma)$ if and only if $\xi \gamma' = \delta \xi \gamma \varepsilon \psi(\delta)^{-1}$ for some $\delta \in \Delta$ and $\varepsilon \in E_\xi$. Writing $\delta = \delta' \gamma \varepsilon \psi(\delta')^{-1}$ for $\delta' \in \xi^{-1}\Delta$ and $\varepsilon \in E_\xi$, this is equivalent to $\gamma' = \delta' \gamma \varepsilon \psi(\gamma \varepsilon')$ by assumption, and so does $\varepsilon \in E_\xi \subset \psi(\Delta)$ by Definition 9.2. Thus the equation requires that $\delta'$ lies in $\psi(\Delta)$, and so a fortiori in $\psi(\Delta) \cap \xi^{-1}\Delta = \Delta_\xi$. In view of Lemma 9.8 the condition is thus equivalent to $\gamma' \in o_{\mathcal{A}_\xi}(\gamma)$, proving the equation at the end of the proposition.

That equation implies that the map $o_{\mathcal{A}_\xi}(\gamma) \mapsto o_{\mathcal{A}}(\xi \gamma)$ from $\mathcal{A}_\xi$-equivalence classes in $\Gamma_\xi$ to $\mathcal{A}$-equivalence classes in $\Delta \psi(\Delta)$ is well-defined and injective.
But any element of $\Delta \xi \psi(\Delta)$ has the form $\delta \xi \gamma$ for $\delta \in \Delta$ and $\gamma \in \Gamma_\xi$ and is therefore equivalent to $\xi \gamma \psi(\delta) \in \xi \Gamma_\xi$. Thus the map is also surjective, and we are done. \qed

9.2 Coxeter groups

**Definition 9.10.** Let $W$ be a Coxeter group with a finite set of simple reflections $S$. Let $\psi : W_I \to W_J \subset W$ be an isomorphism of Coxeter groups with $\psi(I) = J$ for subsets $I, J \subset S$. Then $\mathcal{A} := (W, W_I, \psi)$ is an abstract zip datum that we call of Coxeter type.

Fix such an abstract zip datum of Coxeter type $\mathcal{A}$. Recall that $^I W^J$ is a set of representatives for the double quotient $W_I \backslash W / W_J$. We will apply the induction step from Proposition 9.9 to $x \in ^I W^J$. As in Proposition 4.13 set $I_x := J \cap ^x I$ and $J_x := \psi(I \cap ^x J)$, which are both subsets of $J$. Then $W_J = \psi(W_I)$, and $W_L = \psi(W_I) \cap ^x W_J$ by Proposition 2.6, and $\psi_x := \psi \circ \text{int}(x)$ induces an isomorphism $\psi_x : W_I \to W_L$ such that $\psi_x(I_x) = J_x$. Thus the new abstract zip datum from Construction 9.7 is $\mathcal{A}_x := (W_I, W_L, \psi_x)$ and hence again of Coxeter type. Using this we obtain the following analogue of Theorem 5.10, which also has been previously proved by He ([He3] Corollary 2.6).

**Theorem 9.11.** For $\mathcal{A}$ of Coxeter type $^I W$ is a set of representatives for the equivalence classes in $W$.

**Proof.** We prove this by induction on $\#S$. If $I = S$, we have $W_I = W_J = W$ and so $E_w = W$ for every $w \in W$. Then there is exactly one equivalence class, represented by the unique element of $^I W = \{1\}$, and the assertion holds.

Otherwise we have $\#I < \#S$. Take any $x \in ^I W^J$. Then by the induction hypothesis $^I W_J$ is a set of representatives for the $\mathcal{A}_x$-equivalence classes in $W_J$. Thus Proposition 9.9 implies that $x^{-1} W_J$ is a set of representatives for the $\mathcal{A}$-equivalence classes in $W_J x W_J$. Varying $x$, Proposition 2.8 implies that $^I W$ is a set of representatives for the equivalence classes in $W$, as desired. \qed

For use in Section 11 we include the following results.

**Lemma 9.12.** (a) For any $w \in ^I W$ there exists $y \in W_I$ such that $w' := y w \psi(y)^{-1} \in W^J$.

(b) The element $w'$ in (a) is independent of $y$.

**Proof.** For (a) we use induction on $\#S$. If $I = S$, we have $^I W = \{1\}$ and $w = 1$, and so $y = 1$ does the job. Otherwise $\#I < \#S$. Write $w = x w_J$ as in Proposition 2.8 with $x \in ^I W^J$ and $w_J \in ^{I_x} W_J$. Then by the induction hypothesis applied to $\mathcal{A}_x$ there exists $y' \in W_L$ such that

$$w'_J := y' w_J \psi_x(y')^{-1} \in W_J = W^{\psi(\cap x J)} = \psi(W^{\psi(\cap x J)})$$
Setting $y := \psi^{-1}(y' w_J) \in W_f$ and using the definition of $\psi_x$ we deduce that
\[
w' := yw\psi(y)^{-1} = \psi^{-1}(y' w_J) \cdot xw_J \cdot (y' w_J)^{-1} \\
= \psi^{-1}(y' w_J) \cdot x y'^{-1} x^{-1} \cdot x \\
= \psi^{-1}(y' w_J \psi_x(y')) \cdot x \\
= \psi^{-1}(w'_J) \cdot x \in W_f \cap w_J \cdot x.
\]

By Proposition 2.9 the right hand side is contained in $W_J$, showing (a).

To prove (b) consider another element $y' \in W_f$ such that $w'' := y' w\psi(y')^{-1} \in W_J$. Then with $\tilde{y} := \psi(y'y^{-1}) \in W_f$ we have $w'' = y'y^{-1} w' \psi(y') \psi(y')^{-1} = \psi^{-1}(\tilde{y}) w' \tilde{y}^{-1}$ and hence $w'' = \tilde{y} w'' \tilde{y}^{-1} = \tilde{y} w'' \tilde{y}^{-1}$. Now observe that on replacing $(I, J, \psi)$ by $(J, I, \psi^{-1})$ we obtain another abstract zip datum $A' := (W, W_f, \psi^{-1})$ dual to $A$. The last equality then shows that $w''$ and $w''$ are equivalent according to Definition 9.4 for $A'$. Since these elements also lie in $W_f$, Theorem 9.11 applied to $A'$ shows that they are equal. Therefore $w'' = w'$, as desired.

**Proposition 9.13.** There exists a unique bijection $\sigma : W \to W_f$ with the property that for any $w \in W$ there exists $y \in W_f$ such that $\sigma(w) = yw\psi(y)^{-1}$.

**Proof.** The existence of a unique map $\sigma : W \to W_f$ with the stated property is equivalent to Lemma 9.12. By applying the same lemma to the abstract zip datum $A := (W, W_f, \psi^{-1})$ in place of $A$ we find that for any $w' \in W_f$ there exists $y' \in W_f$ such that $w := y' w \psi(y')^{-1} \in W_f$, and the element $w'$ is independent of $y'$. After replacing $(w, w')$ by $(w^{-1}, w^{-1})$ this means that for any $w' \in W_f$ there exists $y' \in W_f$ such that $w := \psi^{-1}(y') w y'^{-1} \in W_f$, and the element $w$ is independent of $y'$. But with $y := \psi^{-1}(y')^{-1} \in W_f$ the last equation is equivalent to $w = yw\psi(y)^{-1}$, and so for any $w' \in W_f$ there exists a unique $w \in W$ with $w' = \sigma(w)$. In other words the map is bijective, as desired.

**Proposition 9.14.** The bijection in Proposition 9.13 satisfies $\ell(w) = \ell(\sigma(w))$ for all $w \in W_f$.

**Proof.** Write the defining relation in the form $yw = \sigma(w) \psi(y)$. Here $y \in W_f$ and $w \in W_f$ imply that $\ell(yw) = \ell(y) + \ell(w)$, and similarly $\sigma(w) \in W_f$ and $\psi(y) \in W_f$ imply that $\ell(\sigma(w) \psi(y)) = \ell(\sigma(w)) + \ell(\psi(y))$. Moreover, since $\psi$ sends simple reflections to simple reflections, it satisfies $\ell(\psi(y)) = \ell(y)$. Together it follows that $\ell(w) = \ell(\sigma(w))$.

**Lemma 9.15.** Let $\sigma : W_f \to W_f$ be the bijection from Proposition 9.13. For any $x \in W_f$ let $\sigma_x : W_f \to W_f$ denote the bijection obtained by applying Proposition 9.13 to $A_x$. Then for all $w_J \in W_f$ we have $\sigma(xw_J) = \psi^{-1}(\sigma_x(w_J)) \cdot x$.

**Proof.** The proof of Lemma 9.12 (a) shows that $\sigma(w) = w' = \psi^{-1}(w'_J) \cdot x$ where $w'_J = \sigma_x(w_J)$, as desired.
Remark 9.16. Propositions 9.13 and 9.14 can also be deduced from more general results of He ([He3] Proposition 4.3).

9.3 Back to algebraic groups

Now we return to the situation and the notations of the preceding sections. Clearly the connected algebraic zip datum \( Z \) gives rise to an abstract zip datum of Coxeter type \( A := (W, W_f, \psi) \), which by Proposition 3.11 is independent of the frame, up to unique isomorphism. Theorem 5.10 implies that for any \( w' \in W \) the element \( gw' \) lies in \( G_w \) for a unique \( w \in \hat{I}W \).

Theorem 9.17. For any \( w' \in W \) and \( w \in \hat{I}W \) we have \( gw' \in G_w \) if and only if \( w' \sim w \) with respect to \( A \).

Proof. We prove this by induction on \#S. If \( J = S \), there is exactly one \( G_w \) for \( w = 1 \) and exactly one \( A \)-equivalence class in \( W \), so the assertion holds. Otherwise we have \#J < \#S. Write \( w = xwJ \) with \( x \in \hat{I}W^J \) and \( wJ \in \hat{I}_xW \), as in Proposition 2.8. Then by (4.1) and Lemma 5.9 the condition \( gw' \in G_w \) requires that \( w' \in W_xW_J \), and so does the condition \( w' \sim w \) by the remarks in Subsection 9.2. It therefore suffices to consider \( w' = yxwJ' \) with \( y \in W_I \) and \( wJ' \in \hat{I}_yW \), as in Proposition 2.7. But then \( w' \sim xwJ' \psi(y) \) with respect to \( A \), and \( gw' = g\hat{x}wJ' \) is in the same \( E_Z \)-orbit as \( \hat{g}\hat{x}wJ' \psi(y) \). After replacing \( w' \) by \( xwJ' \psi(y) \) we may thus assume that \( w' = xwJ' \) for some \( wJ' \in W_J \). Then Proposition 4.7 and Lemma 5.9 show that \( gw' \in G_w \) if and only if \( wJ' \in \hat{M}wJ \).

By the induction hypothesis this is equivalent to \( wJ' \sim wJ \) with respect to \( A \).

By Proposition 9.9 this in turn is equivalent to \( w' \sim w \) with respect to \( A \), as desired.

Combining Theorems 7.5 and 9.17 we deduce:

Corollary 9.18. If \( Z \) is orbitally finite, then for any \( w, w' \in W \) the elements \( gw \) and \( gw' \) lie in the same \( E_Z \)-orbit if and only if \( w \sim w' \) with respect to \( A \).

10 Non-connected algebraic zip data

In this section we generalize the main results of Sections 5 and 6 to non-connected groups. Throughout we denote a not necessarily connected linear algebraic group by \( \hat{G} \), its identity component by \( G \), and its finite group of connected components by \( \pi_0(\hat{G}) := G/G \); and similarly for other letters of the alphabet. Note that the unipotent radical \( R_uG \) is a normal subgroup of \( \hat{G} \). Any homomorphism \( \hat{\varphi}: \hat{G} \to \hat{H} \) restricts to a homomorphism \( \varphi: G \to H \).

Definition 10.1. An algebraic zip datum is a tuple \( \hat{Z} = (\hat{G}, \hat{P}, \hat{Q}, \hat{\varphi}) \) consisting of a linear algebraic group \( \hat{G} \) with subgroups \( \hat{P} \) and \( \hat{Q} \) and a homomorphism \( \hat{\varphi}: \hat{P}/R_uP \to \hat{Q}/R_uQ \), such that \( Z := (G, P, Q, \varphi) \) is a connected algebraic zip datum. The zip group \( E_{\hat{Z}} \subseteq \hat{P} \times \hat{Q} \), its action on \( \hat{G} \), and the orbit \( o_{\hat{Z}}(X) \) of a subset \( X \subset \hat{G} \) are defined in exact analogy to (3.2), (3.3), and (3.4).

Documenta Mathematica 16 (2011) 253–300
Throughout this section we fix an algebraic zip datum \( \mathcal{Z} = (\hat{G}, \hat{P}, \hat{Q}, \hat{\varphi}) \) with associated connected algebraic zip datum \( Z = (G, P, Q, \varphi) \). We fix a frame \((B, T, g)\) of \( Z \) and use the other pertaining notations from Sections 3 through 5.

We also define

\[ \hat{W} := \text{Norm}_{\hat{L}}(T)/T \quad \text{and} \quad \Omega := (\text{Norm}_{\hat{G}}(B) \cap \text{Norm}_{\hat{G}}(T))/T, \]

so that \( \Omega \cong \pi_0(\hat{G}) \) and \( \hat{W} = W \times \Omega \). For each \( \omega \in \Omega \) we fix a representative \( \hat{\omega} \in \text{Norm}_{\hat{G}}(B) \cap \text{Norm}_{\hat{G}}(T) \), and for \( \hat{w} = \hat{w} \omega \in \hat{W} \) with \( w \in W \) and \( \omega \in \Omega \) we set \( \hat{w} := \hat{w} \omega \in \text{Norm}_{\hat{G}}(T) \).

Note that by definition \( E_\mathcal{Z} \) is the identity component of \( E_\mathcal{Z} \). Thus to study the \( E_\mathcal{Z} \)-orbits in \( \hat{G} \), we first study the orbits under \( E_\mathcal{Z} \) and then the action of \( E_\mathcal{Z}/E_\mathcal{Z} \) on them.

**Lemma 10.2.** For any \( \omega \in \Omega \) the conjugate connected algebraic zip datum

\[ \hat{\omega}_Z := (G, P, \hat{\omega} Q, \text{int}(\hat{\omega}) \circ \varphi) \]

has zip group \( E_{\omega Z} = \{(p, \hat{\omega} q) \mid (p, q) \in E_\mathcal{Z}\} \) and frame \((B, T, g)\), and the isomorphism of varieties \( G \rightarrow G \hat{\omega} \), \( g \mapsto g \hat{\omega} \) induces a bijection from the \( E_{\omega Z} \)-orbits in \( G \) to the \( E_\mathcal{Z} \)-orbits in \( G \hat{\omega} \).

**Proof.** Direct calculation.

**Lemma 10.3.** The subsets \( o_{\omega}(gB\hat{w}B) \) for all \( \hat{w} \in \hat{W} \Omega \) form a disjoint decomposition of \( G \).

**Proof.** Take any \( \omega \in \Omega \). Then by Theorems 5.10 and 5.14 the subsets \( o_{\omega}(gB\hat{w}B) \) for all \( w \in W \) form a disjoint decomposition of \( G \). Thus by Lemma 10.2 the subsets \( o_{\omega}(gB\hat{w}B) \) for all \( w \in W \) form a disjoint decomposition of \( G \hat{\omega} \). Since \( \hat{\omega} \in \text{Norm}_{\hat{G}}(B) \) by assumption, the latter subset is equal to \( o_{\omega}(gB\hat{w}B) \). By varying \( \omega \) the proposition follows.

Next define \( \hat{L} := \text{Norm}_{\hat{L}}(L) \) and \( \hat{M} := \text{Norm}_{\hat{M}}(M) \), so that \( \hat{P} = U \times \hat{L} \) and \( \hat{Q} = V \times \hat{M} \), and \( \hat{\varphi} \) can be identified with a homomorphism \( \hat{L} \rightarrow \hat{M} \). Set

\[ \hat{W}_I := \text{Norm}_{\hat{L}}(T)/T, \quad \Omega_I := (\text{Norm}_{\hat{L}}(B) \cap \text{Norm}_{\hat{L}}(T))/T, \]

\[ \hat{W}_J := \text{Norm}_{\hat{M}}(T)/T, \quad \Omega_J := (\text{Norm}_{\hat{M}}(B) \cap \text{Norm}_{\hat{M}}(T))/T. \]

These groups are subgroups of \( \hat{W} \) and satisfy

\[ \hat{W}_I = W_I \times \Omega_I, \quad \Omega_I \cong \pi_0(\hat{L}) \cong \pi_0(\hat{P}), \]

\[ \hat{W}_J = W_J \times \Omega_J, \quad \Omega_J \cong \pi_0(\hat{M}) \cong \pi_0(\hat{Q}). \]

Also \( \hat{\varphi} \circ \text{int}(g) \) induces a homomorphism \( \hat{\varphi} : \hat{W}_I \rightarrow \hat{W}_J \) extending \( \psi : W_I \rightarrow W_J \) and sending \( \Omega_I \) to \( \Omega_J \). Moreover, the elements \( (\hat{\omega}, \hat{\varphi}(\hat{\omega})) \) for all \( \omega \in \Omega_I \) are representatives of the connected components of \( E_\mathcal{Z} \).
Lemma 10.4. (a) The map \((v, \dot{w}) \mapsto v \dot{w} \check{\psi}(v)^{-1}\) defines a left action of \(\Omega_I\) on \(\check{I}W\).

(b) Take any \(v \in \Omega_I\) and \(\dot{w} \in \check{I}W\) and abbreviate \(\dot{w}' := v \dot{w} \check{\psi}(v)^{-1} \in \check{I}W\).
Then the element \((\check{v}' \check{g}, \check{\psi}(\check{v}'))\) \(\in E_\check{x}\) sends \(o_\check{w}(gB\dot{w}B)\) to \(o_\check{w}(gB\dot{w}'B)\).

Proof. Conjugation by \(\Omega_I\) preserves the set of simple reflections \(I\) and thus the subset \(\check{I}W \subset \check{W}\). In (a) we therefore have \(v \dot{w} \check{\psi}(v)^{-1} = \check{v}' \dot{w} \cdot \check{\psi}(\check{v})^{-1} \in \check{I}W\Omega \cdot \check{W} \check{W} = \check{I}W\Omega\), as desired. In (b) the elements \(\check{v}'\) and \(\check{\psi}(\check{v}')\) normalize \(B\); hence the image is

\[\check{v}' \cdot o_\check{z}(gB\dot{w}B) \cdot \check{\psi}(\check{v}')^{-1} = o_\check{z}(gB\dot{w}gB\check{\psi}(\check{v})^{-1}) = o_\check{z}(gB\dot{w}g\check{\psi}(\check{v})^{-1}B).\]

As \(\dot{v} \check{\psi}(\check{v})^{-1}\) differs from \(\dot{w}'\) by an element of \(T\), this proves (b). \(\square\)

For any \(\dot{w} \in \check{I}W\Omega\) we now define

\[\check{\mathcal{G}}^{\dot{w}} := o_\check{w}(gB\dot{w}B),\quad (10.5)\]

which is independent of the representative \(\dot{w}\). Lemma 10.4 implies that \(\check{\mathcal{G}}^{\dot{w}}\) is the union of \(o_\check{w}(gB\dot{w}'B)\) for all \(\dot{w}'\) in the \(\Omega_I\)-orbit of \(\dot{w}\) under the action in 10.4 (a). Thus \(\check{\mathcal{G}}^{\dot{w}}\) depends only on \(\dot{w}\) modulo \(\Omega_I\), and with Lemma 10.3 we conclude:

Theorem 10.6. The subsets \(\check{\mathcal{G}}^{\dot{w}}\) for all \(\dot{w} \in \check{I}W\Omega\) modulo the action of \(\Omega_I\) from 10.4 (a) form a disjoint decomposition of \(\check{G}\).

To describe the closure relation between the subsets \(\check{\mathcal{G}}^{\dot{w}}\) we define analogues of the Bruhat order \(\leq\) on \(\check{W} = \check{W}\Omega\) and of the relation \(\preceq\) from Definition 6.1 on \(\check{I}W\Omega\):

Definition 10.7. For \(\dot{w} = w\omega\) and \(\dot{w}' = w'\omega'\) with \(w, w' \in \check{W}\) and \(\omega, \omega' \in \Omega\) we write \(\dot{w}' \preceq \dot{w}\) if and only if \(w' \leq w\) and \(\omega' = \omega\).

Definition 10.8. For \(\dot{w}, \dot{w}' \in \check{I}W\Omega\) we write \(\dot{w}' \preceq \dot{w}\) if and only if there exists \(\dot{y} \in \check{W}_I\) such that \(\dot{y}\dot{w}'\check{\psi}(\check{y})^{-1} \preceq \dot{w}\).

Theorem 10.9. For any \(\dot{w} \in \check{I}W\Omega\) we have

\[\overline{\check{\mathcal{G}}^{\dot{w}}} = \bigcup_{\dot{w}' \in \check{I}W\Omega} \check{\mathcal{G}}^{\dot{w}'}_{\dot{w}' \preceq \dot{w}}.\]

Proof. Write \(\dot{w} = w\omega\) with \(w \in \check{I}W\) and \(\omega \in \Omega\). Then the conjugate zip datum \(\check{\omega} \check{Z}\) has the isogeny \(\text{int}(\check{\omega}) \circ \check{\varphi} : L \rightarrow \check{\omega}M\) and hence the induced isomorphism of Weyl groups \(\text{int}(\omega) \circ \psi : W_I \rightarrow \check{\omega}W = W_{\check{\omega}}\). Thus Theorems 5.14 and 6.2 and Definition 6.1 imply that

\[o_{\check{\omega}}(gB\dot{w}B) = \bigcup_{\dot{w}'} o_{\check{\omega}}(gB\dot{w}'B),\]

Documenta Mathematica 16 (2011) 253–300
where the union ranges over all \( w' \in iW \) such that \( yw' \psi(y)^{-1} \leq w \) for some \( y \in W_I \). Note that this inequality is equivalent to \( yw' \psi(y)^{-1} \leq w \) by Definition 10.7. Thus with Lemma 10.2 we deduce that

\[
\hat{G} \hat{w} = o_Z(gB \dot{\hat{w}}B) = \bigcup_{w'} o_Z(gB \dot{w'} \hat{B}) = \bigcup_{\hat{w}'} o_Z(gB \dot{\hat{w}}' \hat{B}),
\]

where the last union ranges over all \( \hat{w}' \in iW \) such that \( \hat{y} \hat{w}' \psi(y)^{-1} \leq \hat{w} \) for some \( y \in W_I \). By taking the union of conjugates of this under \( \hat{g} \in E \hat{Z} \) for all \( \hat{\psi} \in \Omega \) we obtain the closure of \( \hat{G} \hat{w} \). By Lemma 10.4 the right hand side then yields the union of \( o_Z(gB \dot{w} \hat{B}) \) for all \( \hat{w}'' = \hat{w}' \hat{\psi}(v)^{-1} \) with \( \hat{y} \hat{w}' \psi(y)^{-1} \leq \hat{w} \) for some \( v \in \Omega_I \) and \( y \in W_I \). But here \( \hat{y} := yu^{-1} \) runs through the group \( \hat{W} \hat{I} = W_I \) and the inequality is equivalent to

\[
\hat{y} \hat{w}'' \hat{\psi}(\hat{y})^{-1} = yu^{-1} \hat{w}' \psi(v) \psi(y)^{-1} \leq \hat{w}.
\]

By Definition 10.8 these \( \hat{w}'' \) are precisely the elements of \( iW \) satisfying \( \hat{w}'' \preceq \hat{w} \).

Finally, let us call \( \hat{Z} \) orbitally finite if the conjugates \( \hat{Z} \) are orbitally finite for all \( \hat{\psi} \in \Omega \). This holds in particular when the differential of \( \hat{\psi} \) at 1 vanishes, because then we can apply Proposition 7.3 to \( \hat{Z} \). Combining Theorem 7.5 with the remarks leading up to Theorem 10.6 we deduce:

**Theorem 10.10.** Assume that \( \hat{Z} \) is orbitally finite. Then:

(a) For any \( \hat{w} \in iW \) we have \( \hat{G} \hat{w} = o_Z(g\dot{\hat{w}}) \).

(b) If \( \hat{w} \in iW \) runs through a system of representatives for the action of \( \Omega_I \) from 10.4 (a), then \( g\dot{\hat{w}} \) runs through a set of representatives for the \( E \hat{Z} \)-orbits in \( \hat{G} \).

### 11 Dual parametrization

The decomposition of \( G \) from Theorem 5.10 is parametrized in a natural way by elements of \( iW \). In this section we translate that parametrization into an equally natural parametrization by elements of \( W^J \), which was used by Lusztig and He (see Section 12). We also carry out the corresponding translation in the non-connected case.

#### 11.1 The connected case

For any \( w \in W_I \) we set

\[
G^w := o_Z(gB \dot{w}B).
\]

Note that this does not depend on the representative \( \dot{w} \) of \( w \) and conforms to Definition 5.7 by Theorem 5.14. In Proposition 9.13 we have already established a natural bijection \( \sigma : iW \rightarrow W^J \).
Theorem 11.2. For any $w \in \mathcal{W}$ we have $G^w = G^\sigma(w)$.

Proof. If $I = J = S$, we have $\mathcal{W} = \mathcal{W}^J = \{1\}$ and so $w = \sigma(w) = 1$; hence the assertion holds trivially. Otherwise $\#I < \#S$. Write $w = xw_I$ as in Proposition 2.8 with $x \in \mathcal{W}^J$ and $w_I \in \mathcal{W}^I$, and let $\sigma_x: \mathcal{W}^I \rightarrow \mathcal{W}^I$ denote the bijection obtained by applying Proposition 9.13 to $A$. Using (3.3) and (3.9) for the action of $\mathcal{W}$, we have

\[
M_w = M^{\sigma(w)} = o_Z((M \cap B)\tilde{\sigma}_x(w_I)(M \cap B)).
\]

By Lemma 5.13 this is equal to $o_Z((M \cap B)\tilde{\sigma}_x(w_I))$, and so by Lemmas 5.9 and 4.9 we have

\[
G^w = o_Z(g\hat{x}M^w) = o_Z(g \circ o_Z((M \cap B)\tilde{\sigma}_x(w_I))) = o_Z(g \circ (M \cap B)\tilde{\sigma}_x(w_I)).
\]

Recall from Lemma 9.15 that $\sigma(w) = w_Ix$ with $w_I := \psi^{-1}(\sigma_x(w_I)) \in W_I$. It follows that $\tilde{\sigma}_x(w_I) = \psi(w_I)$ and therefore $\tilde{\sigma}_x(w_I) \in T : \varphi(\tilde{w}_I)$ and $\tilde{\sigma}(w_I) \in T : \tilde{w}_I$. Since $T \subset M \cap B$, using the action (3.3) of $E_Z$ we deduce that

\[
G^w = o_Z(g \circ (M \cap B)\varphi(\tilde{w}_I)) = o_Z(g \circ (M \cap B)) = o_Z(g \circ (M \cap B)) = G^w.
\]

Using (3.3) and (3.9) for the action of $V$, respectively Lemma 5.13, we conclude that

\[
G^w = o_Z(g \circ (M \cap B)) = o_Z(g \circ (M \cap B)) = G^w,
\]

as desired. \qed

Theorem 11.3. The $G^w$ for all $w \in \mathcal{W}^I$ form a disjoint decomposition of $G$ by nonsingular subvarieties of dimension $\dim P + \ell(w)$.

Proof. Combine Theorems 5.10, 5.11, 11.2 and Proposition 9.14. \qed

Next, in analogy to Definition 6.1 we define:

Definition 11.4. For $w, w' \in \mathcal{W}^I$ we write $w' \preceq w$ if and only if there exists $y \in W_I$ such that $yw' \psi(y)^{-1} \leq w$.

Theorem 11.5. For any $w \in \mathcal{W}^I$ we have

\[
\overline{G^w} = \bigcup_{w' \preceq w} G^{w'}.
\]

Proof. By combining Theorems 11.2 and 6.2 we already know that $\overline{G^w}$ is the disjoint union of $G^{w'}$ for certain $w' \in \mathcal{W}^I$; it only remains to determine which. First consider $w' \in \mathcal{W}^I$ with $G^{w'} \subset \overline{G^w}$. Then $\tilde{w}' \in \overline{G^w}$, and so by Lemma 6.6 there exist $b \in B$ and $w'' \in \mathcal{W}$ such that $w'' \leq w$ and $o_Z(gw') = o_Z(gw''b)$. Set $\tilde{w}' := \sigma^{-1}(w') \in \mathcal{W}$ and take $y \in W_I$ satisfying $w' = (y\tilde{w}') \psi(y)^{-1}$. Then $w' = (y\tilde{w}')_t\varphi(t\tilde{y})^{-1}$ for some $t \in T$, and thus $o_Z(gw') = o_Z(gw''t)$. Therefore
Again this does not depend on the representative \( \hat{w} \).

Conversely consider \( w' \in W_J \) with \( w' \leq w \), and take \( y \in W_I \) such that \( w'' := yw'\psi(y)^{-1} \leq w \). Lemma 6.6 then shows that \( o_Z(\hat{y}w'T) \subset G^{w''} \). Therefore

\[
o_Z(\hat{y}w'T) = o_Z(g\hat{y}w'T\varphi(g\hat{y}g^{-1})^{-1}) = o_Z(g\hat{y}w'T\varphi(g\hat{y}^{-1})^{-1}T) = o_Z(\hat{y}w''T) \subset G^{\hat{w}''}.
\]

Since also \( o_Z(\hat{y}w'T) \subset G^{w''} \), this with the preliminary remark on \( G^{w''} \) shows that \( G^{\hat{w}''} \subset G^{\hat{w}''} \), proving “\( \subset \)”. □

Remark 11.6. In Definitions 5.7 and 11.1 we have introduced the subsets \( G^w := o_Z(gB\hat{w}B) \) only for \( w \in \{
\}
\( W \cup W_J \), not for arbitrary \( w \in W \). Our results do not say anything directly about the latter. Note that in case \( \varphi \) is an isomorphism their closures have been determined in [LY1] Theorem 5.2 and [He3] Proposition 5.8.

11.2 The non-connected case

Now we return to the notations from Section 10. We begin with an analogue of Proposition 9.13:

Proposition 11.7. There exists a unique bijection \( \hat{\sigma}: I^1 W\Omega \to \Omega^W \) with the property that for any \( \hat{w} \in I^1 W\Omega \) there exists \( y \in W_I \) such that \( \hat{\sigma}(\hat{w}) = yw\psi(y)^{-1} \).

Proof. The equation requires that \( \hat{\sigma}(\hat{w}) \in \hat{W} \) lie in the same \( W \)-coset as \( \hat{w} \). Thus for any fixed \( \omega \in \Omega \), we need a unique bijection \( I^1 W\omega \to \omega^W \) sending \( \omega w \) to an element of the form \( yw\omega\psi(y)^{-1} \) for some \( y \in W_I \). Multiplying both elements on the right by \( \omega^{-1} \) this amounts to a unique bijection \( I^1 W \to \omega W_J \omega^{-1} = W^{-J} \) sending \( w \) to an element of the form \( yw\omega\psi(y)^{-1}\omega^{-1} \) for some \( y \in W_I \). But \( \text{int}(\omega) \circ \psi: W_I \to \omega W_J \omega^{-1} = W^{-J} \) is precisely the isomorphism associated to the conjugate connected algebraic zip datum \( \hat{w} \Omega \) from Lemma 10.2. Thus a unique bijection with that property exists by Proposition 9.13 applied to \( \hat{w} \Omega \).

For any \( \hat{w} \in \Omega W^J \) we now define

\[
\hat{G}^{\hat{w}} := o_Z(gB\hat{w}B).
\]

(11.8)

Again this does not depend on the representative \( \hat{w} \) of \( \hat{w} \) and conforms to Definition (10.5).

Theorem 11.9. For any \( \hat{w} \in I W\Omega \) we have \( \hat{G}^{\hat{w}} = \hat{G}^{\hat{\sigma}(\hat{w})} \).
Proof. Write \( \hat{w} = w\omega \) with \( w \in W \) and \( \omega \in \Omega \). In the proof of Proposition 11.7 we have seen that \( \hat{\sigma}(\hat{w}) = w'\omega \), where \( w' \in W^{-J} \) is the image of \( w \) under the isomorphism given by Proposition 9.13 applied to \( ^wZ \). Thus by Theorem 11.2 we have \( o_Z(gB\hat{w}B) = o_Z(gBwB) \) inside \( G \). On multiplying on the right by \( \hat{\omega} \) and applying Lemma 10.2 to both sides we deduce that
\[
o_Z(gB\hat{w}B) = o_Z(gB\hat{w}\hat{B}\hat{\omega}) = o_Z(gB\hat{w}'B\hat{\omega}) = o_Z(gB\hat{\sigma}(\hat{w})B).
\]
The desired equality follows from this by applying \( o_Z \).

Lemma 11.10. (a) The map \( (v, \hat{w}) \mapsto v\hat{w}\hat{\psi}(v)^{-1} \) defines a left action of \( \Omega_I \) on \( \Omega W^J \).

(b) The bijection \( \hat{\sigma} : ^I W \Omega \to \Omega W^J \) from Proposition 11.7 is \( \Omega_I \)-equivariant.

Proof. Take \( v \in \Omega_I \) and \( \hat{w} \in \Omega W^J \). To prove (a) observe that conjugation by \( \hat{\psi}(v) \in \Omega_I \) preserves the set of simple reflections \( J \) and thus the subset \( W^J \subset W \). We therefore have \( v\hat{w}\hat{\psi}(v)^{-1} = v\hat{\psi}(v)^{-1}, \hat{\psi}(v)\hat{w} \in \Omega \cdot \Omega W^J = \Omega W^J \), as desired. In (b) write \( \hat{\sigma}(\hat{w}) = y\hat{w}\hat{\psi}(y)^{-1} \) with \( y \in W_I \). Then
\[
v\hat{\sigma}(\hat{w})\hat{\psi}(v)^{-1} = (vy^{-1})(v\hat{w}\hat{\psi}(v)^{-1})\hat{\psi}(vy^{-1})^{-1} = \hat{\sigma}(v\hat{w}\hat{\psi}(v)^{-1)},
\]
because the left hand side is in \( \Omega W^J \) and \( v y v^{-1} \in W_I \). This proves (b).

Theorem 11.11. The subsets \( \hat{G}w \) for all \( \hat{w} \in \Omega W^J \) modulo the action of \( \Omega_I \) from 11.10 (a) form a disjoint decomposition of \( \hat{G} \).

Proof. Combine Theorems 10.6 and 11.9 with Lemma 11.10.

Definition 11.12. For \( \hat{w}, \hat{w}' \in \Omega W^J \) we write \( \hat{w} \preceq \hat{w}' \) if and only if there exists \( \hat{y} \in W_I \) such that \( \hat{y}\hat{w}'\hat{\psi}(\hat{y})^{-1} \preceq \hat{w} \).

Theorem 11.13. For any \( \hat{w} \in \Omega W^J \) we have
\[
\overline{G\hat{w}} = \bigcup_{\hat{w}' \in \Omega W^J, \hat{w} \preceq \hat{w}'} \hat{G}\hat{w}'
\]

Proof. Write \( \hat{w} = w\omega \) with \( \omega \in \Omega \) and \( w \in W^{-J} \). Applying Theorem 11.5 to the conjugate zip datum \( \hat{w} \hat{Z} \) shows that \( o_Z(gB\hat{w}B) \) is the union of the subsets \( o_Z(gBwB) \) for all \( w' \in W^{-J} \) such that \( yw'\hat{\omega}(y)^{-1} \omega^{-1} \leq w \) for some \( y \in W_I \). On multiplying on the right by \( \hat{\omega} \) and applying Lemma 10.2 to everything we deduce that \( o_Z(gB\hat{w}B) = o_Z(gB\hat{w}B\hat{\omega}) = o_Z(gB\hat{w}'B\hat{\omega}) = o_Z(gB\hat{w}'\hat{\omega}B) \) for the same elements \( w' \). Writing \( \hat{w}' = w'\omega \) this is equal to the union of the subsets \( o_Z(gB\hat{w}'B) \) for all \( \hat{w}' \in \Omega W^J \) such that \( yw'\hat{\omega}(y)^{-1} \leq w \) for some \( y \in W_I \). The theorem follows from this by applying \( o_Z \).
12 Generalization of certain varieties of Lusztig

In this section we consider a certain type of algebraic variety with an action of a reductive group $G$ whose orbit structure is closely related to the structure of the $E_Z$-orbits in $G$ for an algebraic zip datum $Z$. Special cases of such varieties have been defined by Lusztig ([Lus2]) and by Moonen and the second author in [MW].

12.1 The coset variety of an algebraic zip datum

Remark 12.1. To keep notations simple, we restrict ourselves to connected zip data, although everything in this section directly extends to non-connected ones by putting $\hat{}$ in the appropriate places.

In this section we use only the definition of algebraic zip data and the action of the associated zip group from Section 3, but none of the other theory or notations from the preceding sections, not even the concept of a frame. Fix a connected algebraic zip datum $Z = (G, P, Q, \varphi)$. Recall that $E_Z$ is a subgroup of $P \times Q$ and hence of $G \times G$. We also consider the image of $G$ under the diagonal embedding $\Delta: G \hookrightarrow G \times G$, $g \mapsto (g, g)$. We are interested in the left quotient $\Delta(G) \setminus (G \times G)$ and the right quotient $(G \times G)/E_Z$.

The first is isomorphic to $G$ via the projection morphism $\lambda: G \times G \rightarrow G$, $(g, h) \mapsto g^{-1}h$. (12.2)

Turn the right action of $E_Z$ on $G \times G$ into a left action by letting $(p, q) \in E_Z$ act by right translation with $(p, q)^{-1}$. Then with $E_Z$ acting on $G$ as in the definition of algebraic zip data, a direct calculation shows that $\lambda$ is $E_Z$-equivariant.

To describe the second quotient recall that $\varphi$ is a homomorphism $P/U \rightarrow Q/V$, where $U$ and $V$ denote the unipotent radicals of $P$ and $Q$. Consider a left $P$-coset $X \subset G$ and a left $Q$-coset $Y \subset G$. Then $X/U$ is a right torsor over $P/U$, and $Y/V$ is a right torsor over $Q/V$. By a $P/U$-equivariant morphism $\Phi: X/U \rightarrow Y/V$ we mean a morphism satisfying $\Phi(\bar{x} \bar{p}) = \Phi(\bar{x}) \varphi(\bar{p})$ for all $\bar{x} \in X/U$ and $\bar{p} \in P/U$.

Definition 12.3. The coset space of $Z$ is the set $C_Z$ of all triples $(X, Y, \Phi)$ consisting of a left $P$-coset $X \subset G$, a left $Q$-coset $Y \subset G$, and a $P/U$-equivariant morphism $\Phi: X/U \rightarrow Y/V$.

For any $X$, $Y$ as above and any $(g, h) \in G \times G$, left multiplication by $g$ induces an isomorphism $\ell_g: X/U \sim \rightarrow gX/U$, and left multiplication by $h$ induces an isomorphism $\ell_h: Y/V \sim \rightarrow hY/V$. Therefore $(X, Y, \Phi) \mapsto (gX, hY, \ell_h \circ \Phi \circ \ell_g^{-1})$ defines a left action of $G \times G$ on $C_Z$. By applying this action to the canonical base point $(P, Q, \varphi) \in C_Z$ we obtain a morphism

$$\rho: G \times G \rightarrow C_Z, \ (g, h) \mapsto (gP, hQ, \ell_h \circ \varphi \circ \ell_g^{-1}). \quad (12.4)$$

Clearly this morphism is equivariant under the left action of $G \times G$ and hence under the subgroup $\Delta(G)$. 
Lemma 12.5. There is a unique structure of algebraic variety on $C_Z$ such that $\rho$ identifies $C_Z$ with the quotient variety $(G \times G)/E_Z$.

Proof. The action of $G \times G$ is obviously transitive on the set of all pairs $(X,Y)$. Moreover, any $P/U$-equivariant morphism of right torsors $P/U \rightarrow Q/V$ has the form $\bar{p} \mapsto \pi_Q(q)\varphi(\bar{p}) = \ell_q \circ \varphi(\bar{p})$ for some $q \in Q$. Thus the subgroup $1 \times Q$ acts transitively on the set of all triples of the form $(P,Q,\Phi)$. Together it follows that the action of $G \times G$ on $C_Z$ is transitive.

On the other hand $(g,h)$ lies in the stabilizer of $(P,Q,\Phi)$ if and only if $g \in P$ and $h \in Q$ and $\ell_h \circ \varphi \circ \ell^{-1}_g = \varphi$. But under the first two of these conditions, we have for all $\bar{p} \in P/U$

$\ell_h \circ \varphi \circ \ell^{-1}_g(\bar{p}) = \pi_Q(h)\varphi(\pi_P(g)^{-1} \bar{p}) = \pi_Q(h)\varphi(\pi_P(g))^{-1} \varphi(\bar{p}),$

and so the third condition is equivalent to $\varphi(\pi_P(g)) = \pi_Q(h)$. Together this means precisely that $(g,h) \in E_Z$, which is therefore the stabilizer of $(P,Q,\Phi)$. It follows that $\rho$ induces a bijection $(G \times G)/E_Z \rightarrow C_Z$. Since the quotient variety exists by [Ser], Section 3.2, this yields the unique structure of algebraic variety on $C_Z$.

Following Lemma 12.5 we call $C_Z$ also the coset variety of $Z$. Recall from [Ser] Prop. 2.5.3 that the quotient of an algebraic group by an algebraic subgroup is always a torsor. To summarize we have therefore constructed morphisms with the following properties:

Recall that the actions of $\Delta(G)$ and $E_Z$ on $G \times G$ commute and thus combine to an action of $\Delta(G) \times E_Z$. Therefore (12.6) directly implies:

Theorem 12.7. There are natural isomorphisms of algebraic stacks

$$[E_Z\backslash G] \xrightarrow{[\lambda]} [(\Delta(G) \times E_Z) \backslash (G \times G)] \xrightarrow{[\rho]} [\Delta(G)\backslash C_Z].$$

Even without stacks, we can deduce:

Theorem 12.8. (a) There is a closure-preserving bijection between $E_Z$-invariant subsets $A \subset C_Z$ and $\Delta(G)$-invariant subsets $B \subset C_Z$, defined by $A = \lambda(\rho^{-1}(B))$ and $B = \rho(\lambda^{-1}(A))$.

(b) The subset $A$ in (a) is a subvariety, resp. a nonsingular subvariety, if and only if $B$ is one. In that case we also have $\dim A = \dim B$. 
(c) In particular (a) induces a bijection between $E_\mathbb{Z}$-orbits in $G$ and $\Delta(G)$-orbits in $C_\mathbb{Z}$.

(d) For any $g \in G$ and $(X,Y,\Phi) \in C_\mathbb{Z}$ whose orbits correspond, there is an isomorphism

$$\text{Stab}_{E_\mathbb{Z}}(g) \cong \text{Stab}_{\Delta(G)}((X,Y,\Phi)).$$

**Proof.** By (12.6) any $\Delta(G) \times E_\mathbb{Z}$-invariant subset of $G \times G$ must be simultaneously of the form $\lambda^{-1}(A)$ for an $E_\mathbb{Z}$-invariant subset $A \subset G$ and of the form $\rho^{-1}(B)$ for a $\Delta(G)$-invariant subset $B \subset C_\mathbb{Z}$. Then $A = \lambda(\rho^{-1}(B))$ and $B = \rho(\lambda^{-1}(A))$, giving the bijection in (a). The bijection preserves closures because $\lambda$ and $\rho$ are smooth. This proves (a), the first sentence in (b), and the special case (c). In (b) it also proves that $\dim A + \dim G = \dim B + \dim E_\mathbb{Z}$. But $\dim G = 2\dim U + \dim L = 2\dim V + \dim M$ and $\dim L = \dim M$ imply that $\dim U = \dim V$, and thus using (3.9) that $\dim E_\mathbb{Z} = \dim U + \dim L + \dim V = \dim G$. Therefore $\dim A = \dim B$, proving the rest of (b).

In (c) by assumption there exists a point $x \in G \times G$ such that $\lambda(x)$ lies in the $E_\mathbb{Z}$-orbit of $g$ and $\rho(x)$ lies in the $\Delta(G)$-orbit of $(X,Y,\Phi)$. Thus after replacing $x$ by a suitable translate under $\Delta(G) \times E_\mathbb{Z}$ we may assume that $\lambda(x) = g$ and $\rho(x) = (X,Y,\Phi)$. Then the fact that $\lambda$ and $\rho$ are torsors implies that the two projection morphisms

$$\xymatrix{ \text{Stab}_{E_\mathbb{Z}}(g) \ar[r] & \text{Stab}_{\Delta(G) \times E_\mathbb{Z}}(x) \ar[r] & \text{Stab}_{\Delta(G)}((X,Y,\Phi)) }$$

are isomorphisms, proving (c). (The isomorphism may depend on the choice of $x$.)

With Theorem 12.8 we can translate many results about the $E_\mathbb{Z}$-action on $G$ from the preceding sections to the $\Delta(G)$-action on $C_\mathbb{Z}$, in particular Theorems 5.10, 5.11, 6.2, 7.5, 8.1, and their counterparts from Sections 10 and 11.

### 12.2 Algebraic Zip Data Associated to an Isogeny of $G$

In this subsection we consider algebraic zip data whose isogeny extends to an isogeny on all of $G$. (Not every connected algebraic zip datum has that property, for instance, if $L$ and $M$ have root system $A_1$ associated to long and short roots, respectively, and the square of the ratio of the root lengths is different from the characteristic of $k$.)

Fix a connected reductive algebraic group $G$ over $k$ and an isogeny $\varphi: G \to G$. Choose a Borel subgroup $B \subset G$ and a maximal torus $T \subset B$, and let $W$ be the corresponding Weyl group of $G$ and $S$ its set of simple reflections. Choose an element $\gamma \in G$ such that $\varphi(\gamma B) = B$ and $\varphi(\gamma T) = T$. Then $\varphi \circ \text{int}(\gamma): \text{Norm}_G(T) \to \text{Norm}_G(T)$ induces an isomorphism of Coxeter systems

$$\varphi: (W,S) \cong (W,S).$$

For any subset $I \subset S$ recall from Subsection 2.3 that $P_I$ denotes the standard parabolic of type $I$. Thus the choices imply that $\varphi(P_I) = P_{\varphi(I)}$. We denote
the unipotent radicals of arbitrary parabolics $P$, $Q$, $P'$, $Q'$ by $U$, $V$, $U'$, $V'$, respectively.

Let $\check{G}$ be a linear algebraic group over $k$ having identity component $G$, and let $\check{G}$ be an arbitrary connected component of $G$. Choose an element $g_1 \in \text{Norm}_{\check{G}'}(B) \cap \text{Norm}_{\check{G}'}(T)$. Then $\text{int}(g_1)$ induces an automorphism of $G$ that we use to twist $\varphi$. Let $\delta: (W, S) \to (W, S)$ be the isomorphism of Coxeter systems induced by $\text{int}(g_1)$. Then for any subset $I \subset S$ we have $g_1 P_I = P_{\delta(I)}$.

Fix subsets $I, J \subset S$ and an element $x \in J W^\delta(I)$ with $J = \delta \varphi(I)$. Set $y := (\delta \varphi)^{-1}(x) \in W$.

**Lemma 12.9.**

(a) $x \Phi_{\delta \varphi(I)} = \Phi_J$.

(b) $x \Phi^+_{\delta \varphi(I)} = \Phi^+_J$.

**Proof.** Part (a) follows from $J = \delta \varphi(I)$. By (2.11) the fact that $x \in J W^\delta(I) \subset W^\delta(I)$ implies $x \Phi^+_{\delta \varphi(I)} \subset \Phi^+$. Together with (a) this implies (b). □

**Construction 12.10.** Set $Q := P_J$ and $P := \gamma \varphi P_I$ and let $L$ be the Levi component of $P$ containing $\gamma \varphi T$. Then $\gamma \varphi(P) = \bar{\varphi}(\gamma \varphi(P_I)) = \bar{\varphi} P_{\delta \varphi(I)}$ and $Q = P_I$ have relative position $x$. Set $M := \gamma \varphi(L)$; this is a Levi component of $\gamma \varphi(P)$ containing $\gamma \varphi(T) = \bar{\varphi}(\gamma \varphi(T)) = \bar{\varphi}(T)$. Since the root system of $M$ is $x \Phi_{\delta \varphi(I)}$, Lemma 12.9 shows that it is also the Levi component of $Q$ containing $T$. Let $\gamma \varphi: P/U \to Q/V$ denote the isogeny corresponding to $\text{int}(g_1) \circ \varphi|L: L \to M$. Then we obtain a connected algebraic zip datum $\mathcal{Z} := (G, P, Q, \gamma \varphi)$.

**Lemma 12.11.** The triple $(B, T, \gamma \varphi)$ is a frame of $\mathcal{Z}$, and the Levi components determined by it are $M \subset Q$ and $L \subset P$.

**Proof.** The statements about $M$ and $L$ follow from the inclusions $T \subset M$ and $\gamma \varphi T \subset L$. They also imply that the isogeny $L \to M$ corresponding to $\gamma \varphi$ is simply the restriction of $\text{int}(g_1) \circ \varphi$. Conditions (a) and (b) in Definition 3.6 assert that $B \subset Q$ and $\gamma \varphi B \subset P$, which hold by the construction of $Q$ and $P$. Condition (d) translates to $\gamma \varphi(\gamma \varphi T) = \bar{\varphi}(T)$, which was already shown in 12.10.

To prove (c) note first that by Lemma 12.9 we have $x \Phi^+_{\gamma \varphi(I)} = \Phi^+_J$ and therefore $x B \cap M = B \cap M$. The definition of $y$ implies that $\gamma \varphi(\gamma \varphi^{-1}y) \in xT$ and hence

$$
\gamma \varphi(\gamma \varphi^{-1}y B) = \gamma \varphi(\gamma \varphi^{-1}y \gamma \varphi B) = \gamma \varphi(\gamma \varphi^{-1}B) = \bar{\varphi} B.
$$

From this we can deduce that

$$
\gamma \varphi(\gamma \varphi^{-1}y B \cap L) = \gamma \varphi(\gamma \varphi^{-1}B) \cap \gamma \varphi(L) = \bar{\varphi} B \cap M = B \cap M,
$$

proving the remaining condition (c). □

The automorphism $\psi$ defined in (3.11) for the algebraic zip datum $\mathcal{Z}$ is given by

$$
\psi := \delta \circ \bar{\varphi} \circ \text{int}(y) = \text{int}(x) \circ \delta \circ \bar{\varphi}: (W_I, I) \overset{\sim}{\to} (W_J, J) \quad (12.7)
$$
DEFINITION 12.12. Let $X_{I,\varphi,x}$ be the set of all triples $(P',Q',[g'])$ consisting of parabolic subgroups $P'$, $Q'$ of $G$ of type $I$, $J$ and a double coset $[g'] := V'g'\varphi(U') \subset G^1$ of an element $g' \in G^1$ such that

$$\text{relos}(Q',g'\varphi(P')) = x.$$ 

One readily verifies that the condition on the relative position depends only on $[g']$, and that $((g,h),(P',Q',[g'])) \mapsto (P',hQ',[hg\varphi(g)^{-1}])$ defines a left action of $G \times G$ on $X_{I,\varphi,x}$. We also have a standard base point $(P,Q,[g_1]) \in X_{I,\varphi,x}$. One can use the definition of $X_{I,\varphi,x}$ to endow it with the structure of an algebraic variety over $k$, but in the interest of brevity we define that structure using the following isomorphism:

PROPOSITION 12.13. There is a natural $G \times G$-equivariant isomorphism

$$Z \overset{\sim}{\longrightarrow} X_{I,\varphi,x}, \quad (gP,hQ,\ell_h \circ g_1 \varphi \circ \ell_g^{-1}) \mapsto (P,hQ,[h_1g_1\varphi(g)^{-1}]).$$

Proof. In view of Lemma 12.5 the assertion is equivalent to saying that the action of $G \times G$ on $X_{I,\varphi,x}$ is transitive and the stabilizer of $(P,Q,[g_1])$ is $E_Z$. The transitivity follows directly from the definition of the action. For the stabilizer note that $(P,hQ,[h_1g_1\varphi(g)^{-1}]) = (P,Q,[g_1])$ if and only if $g \in P$ and $h \in Q$ and $Vh_1g_1\varphi(g)^{-1}(U) = Vg_1\varphi(U)$. Write $g = ul$ for $u \in U$, $l \in L$ and $h = vm$ for $v \in V$, $m \in M$. Then the last condition is equivalent to $Vmg_1\varphi(l)^{-1}\varphi(U) = Vg_1\varphi(U)$, or again to $m.\varphi(l)^{-1} \subset V.\varphi(U) \cap M$. But for every element $v'.\varphi(u') = m' \in V.\varphi(U) \cap M$ we have $\varphi(u') = v'^{-1}m' \in \varphi(U) \cap VM$, and since $M$ is also a Levi component of $\varphi(P)$, it follows that $\varphi(U) \cap VM = \varphi(U) \cap V$ and hence $m' = 1$. The last condition is therefore equivalent to $m = \varphi(l)$. Together this shows that the stabilizer is $E_Z$, as desired.

LEMMA 12.14. For any $w \in I^W \cup W^J$ the subset $G^w \subset G$ corresponds via Theorem 12.8 (a) and Proposition 12.13 to the subset

$$X^w_{I,\varphi,x} := \{(P_1,gwP,J,\gamma\varphi(g)^{-1}) \mid g \in G, \ b \in B\} \subset X_{I,\varphi,x},$$

which is a nonsingular variety of dimension $\dim P + \ell(w)$.

Proof. Since $(B,T,\gamma\hat{y})$ is a frame of $Z$ by Lemma 12.11, Theorem 5.14 for $w \in I^W$, respectively (11.1) and Lemma 5.13 for $w \in W^J$, show that $G^w = o_Z(\gamma\hat{y}B\hat{w})$. In other words $G^w$ is the union of the $E_Z$-orbits of $\gamma\hat{y}b\hat{w}$ for all $b \in B$. But by (12.2) and (12.4) we have

$$\lambda(\gamma\hat{y}^{-1},\hat{w}) = \gamma\hat{y}b\hat{w}, \quad \text{and}$$

$$\rho(\gamma\hat{y}^{-1},\hat{w}) = ((\gamma\hat{y}^{-1})P,\hat{w}Q,\ell_{\varphi} \circ \ell_{\gamma\hat{y}^{-1}}^{-1}),$$

and so the $E_Z$-orbit of the former corresponds to the $\Delta(G)$-orbit of the latter under the correspondence from Theorem 12.8. Moreover, under the isomorphism from Proposition 12.13 the latter corresponds to the triple

$$(\gamma\hat{y}^{-1}P,\hat{w}Q,[\hat{w}g_1\varphi((\gamma\hat{y}^{-1})^{-1})]).$$
The definitions of $P$ and $Q$ show that $(\gamma y b)^{-1} P = b^{-1} P_I = P_I$ and $\hat{w} Q = \hat{w} P_J$. The definition of $y$ means that $g_1 \varphi(\gamma y b) = \hat{x} t$ for some $t \in T$; hence

\[ \hat{w} g_1 \varphi((\gamma y b)^{-1})^{-1} = \hat{w} g_1 \varphi(\gamma y b) = \hat{w} \cdot \hat{x} t \cdot g_1 \cdot \varphi(\gamma b y^{-1}) \cdot \varphi(\gamma) = \hat{w} \cdot \hat{x} t \cdot g_1 \cdot \varphi(\gamma y b) \cdot \varphi(\gamma). \]

Since $\varphi(\gamma B) = B$, the factor $b' := g_1^{-1} t \varphi(\gamma b y^{-1})$ runs through $B$ while $b$ runs through $B$. Thus altogether it follows that $G^w$ corresponds to the union of the $\Delta(G)$-orbits of the triples

\[ (P_I, \hat{w} P_J, [\hat{w} \hat{x} g_1 b' \varphi(\gamma)]) \]

for all $b' \in B$. This union is just the set $X_{I,\varphi,x}$ in the lemma. The rest follows from Theorems 5.11, 11.3, and 12.8.

Combining this with Theorems 5.10 and 6.2 and 12.8 we conclude:

**Theorem 12.15.** (a) The $X_{I,\varphi,x}^w$ for all $w \in I W$ form a disjoint decomposition of $X_{I,\varphi,x}$ by nonsingular subvarieties of dimension $\dim P + \ell(w)$.

(b) For any $w \in IW$ we have

\[ X_{I,\varphi,x}^w = \bigsqcup_{w' \in IW} X_{I,\varphi,x}^{w'}. \]

Analogously, using Theorems 11.3 and 11.5 and 12.8 we obtain:

**Theorem 12.16.** (a) The $X_{I,\varphi,x}^w$ for all $WJ$ form a disjoint decomposition of $X_{I,\varphi,x}$ by nonsingular subvarieties of dimension $\dim P + \ell(w)$.

(b) For any $w \in WJ$ we have

\[ X_{I,\varphi,x}^w = \bigsqcup_{w' \in WJ} X_{I,\varphi,x}^{w'}. \]

### 12.3 Frobenius

Keeping the notations of the preceding subsection, we now assume that $k$ has positive characteristic and that $\varphi: G \to G$ is the Frobenius isogeny coming from a model $G_0$ of $G$ over a finite subfield $F_q \subset k$ of cardinality $q$. Then $G_0$ is quasi-split; hence we may, and do, assume that $B$ and $T$ come from subgroups of $G_0$ defined over $F_q$ and therefore satisfy $\varphi(B) = B$ and $\varphi(T) = T$. We can thus take $\gamma := 1$.

In this case, our varieties $X_{I,\varphi,x}$ coincide with the varieties $Z_I$ used in [MW] to study $F$-zips with additional structures. The isogeny $g_1 \tilde{\varphi}$ in the connected algebraic zip datum $Z$ then has vanishing differential; hence $Z$ is orbitally finite by Proposition 7.3. Thus by Theorem 7.5 each $G^w$ is a single $E_Z$-orbit, and so by Theorem 12.8 and Theorem 12.15 we deduce:
Theorem 12.17. (a) If $\varphi$ is the Frobenius isogeny associated to a model of $G$ over a finite field, each $X_{I,\varphi,x}^w$ in Theorem 12.15 is a single $\Delta(G)$-orbit. In particular the set

$$\{(P_I, \bar{w}P_J, [\bar{w}\xi g_1]) \mid w \in \mathcal{I}W\}$$

is a system of representatives for the action of $\Delta(G)$ on $X_{I,\varphi,x}$.

(b) For any $w \in \mathcal{I}W$, the closure of the orbit of $(P_I, \bar{w}P_J, [\bar{w}\xi g_1])$ is the union of the orbits of $(P_I, \bar{w}'P_J, [\bar{w}'\xi g_1])$ for those $w' \in \mathcal{I}W$ satisfying $w' \preccurlyeq w$.

Theorem 12.17 (a) was proved in [MW], Theorem 3 and (b) answers the question of the closure relation that was left open there.

12.4 Lusztig’s Varieties

Now we apply the results of Subsection 12.2 to the special case $\varphi = \text{id}$. In this case we can choose $\gamma := 1$ and obtain $\bar{\varphi} = \text{id}$. Then our varieties $X_{I,\varphi,x}$ coincide with the varieties $Z_{I,x,\delta}$ defined and studied by Lusztig in [Lus2]. There he defines a decomposition of $X_{I,\varphi,x}$ into a certain family of $\Delta(G)$-invariant subvarieties. In [He2], he shows how to parametrize this family by the set $W^{\delta(I)}$. We will denote the piece corresponding to $w \in W^{\delta(I)}$ in this parametrization by $\bar{X}_{I,\varphi,x}^w$. (In [He2], he denotes $X_{I,\varphi,x}$ by $Z_{I,x,\delta}$ and $X_{I,\varphi,x}$ by $\tilde{Z}_{I,x,\delta}$.) We will show that this decomposition is the same as ours from Theorem 12.16 up to a different parametrization.

Lemma 12.18. The map $w \mapsto wx$ induces a bijection $W^J \xrightarrow{\sim} W^{\delta(I)}$.

Proof. Take any $w \in W^J$. Using Lemma 12.9 and (2.11) we get $wx\Phi_{\delta(I)}^+ = w\Phi_j^+ \subset \Phi^+$. By (2.11) this shows that $wx \in W^{\delta(I)}$. A similar argument shows that $wx^{-1} \in W^J$ for any $w \in W^{\delta(I)}$, which finishes the proof.

Theorem 12.19. For any $w \in W^J$ we have $X_{I,\varphi,x}^w = \bar{X}_{I,\varphi,x}^{wx}$.

Proof. The statement makes sense by Lemma 12.18. Let $w \in W^J$ and $w' := wx \in W^{\delta(I)}$. In [He2], Proposition 1.7, he shows that

$$\bar{X}_{I,\varphi,x}^{w'} = \Delta(G) \cdot \{(P_I, b\bar{w}'x^{-1}P_J, [b\bar{w}'g_1b']) \mid b, b' \in B\}.$$

(In [He2], it is assumed that $G$ is semi-simple and adjoint. But this assumption is not needed for the proof of Proposition 1.7 in [loc. cit.].) By acting on such a point $(P_I, b\bar{w}'x^{-1}P_J, [b\bar{w}'g_1b'])$ with $\Delta(b^{-1})$ and using $w = w'x^{-1}$ we get

$$\bar{X}_{I,\varphi,x}^{w'} = \Delta(G) \cdot \{(P_I, \bar{w}P_J, [\bar{w}\xi g_1b']) \mid b' \in B\}.$$

Since $\gamma = 1$, comparison with Lemma 12.14 proves the claim.

From Theorem 12.16 we can now deduce the closure relation between the $\bar{X}_{I,\varphi,x}^w$. 

Documenta Mathematica 16 (2011) 253–300
Theorem 12.20. For any \( w \in W^{\delta(I)} \) we have

\[
\tilde{X}_{I, \varphi, x}^w = \coprod_{w' \in W^{\delta(I)}} \tilde{X}_{I, \varphi, x}^{w'}.
\]

In the special case \( x = 1 \) this result is due to He (see [He2], Proposition 4.6).

References

[BB] Anders Björner and Francesco Brenti. *Combinatorics of Coxeter groups*, volume 231 of *Graduate Texts in Mathematics*. Springer, New York, 2005.

[Car] Roger W. Carter. *Finite groups of Lie type*. Wiley Classics Library. John Wiley & Sons Ltd., Chichester, 1993. Conjugacy classes and complex characters, Reprint of the 1985 original, A Wiley-Interscience Publication.

[DCP] C. De Concini and C. Procesi. Complete symmetric varieties. In *Invariant theory (Montecatini, 1982)*, volume 996 of *Lecture Notes in Math.*, pages 1–44. Springer, Berlin, 1983.

[DDPW] Bangming Deng, Jie Du, Brian Parshall, and Jianpan Wang. *Finite dimensional Algebras and Quantum Groups*, volume 150 of *Mathematical Surveys and Monographs*. AMS, 2008.

[DG] Michel Demazure and Alexandre Grothendieck. *Schémas en groupes. III: Structure des schémas en groupes réductifs*. Sémaire de Géometrie Algébrique du Bois Marie 1962/64 (SGA 3). Dirigé par M. Demazure et A. Grothendieck. Lecture Notes in Mathematics, Vol. 153. Springer-Verlag, Berlin, 1962/1964.

[EL] Sam Evens and Jiang-Hua Lu. On the variety of Lagrangian subalgebras. II. *Ann. Sci. École Norm. Sup. (4)*, 39(2):347–379, 2006.

[He1] Xuhua He. The character sheaves on the group compactification. *Adv. in Math.*, 207:805–827, 2006.

[He2] Xuhua He. The \( G \)-stable pieces of the wonderful compactification. *Trans. Amer. Math. Soc.*, 359(7):3005–3024 (electronic), 2007.

[He3] Xuhua He. Minimal length elements in some double cosets of Coxeter groups. *Adv. in Math.*, 215:469–503, 2007.

[Lus1] George Lusztig. Parabolic character sheaves. I. *Mosc. Math. J.*, 4(1):153–179, 311, 2004.

[Lus2] George Lusztig. Parabolic character sheaves. II. *Mosc. Math. J.*, 4(4):869–896, 981, 2004.

[LY1] Jiang-Hua Lu and Milen Yakimov. Partitions of the wonderful group compactification. *Transform. Groups*, 12(4):695–723, 2007.
Jiang-Hua Lu and Milen Yakimov. Group orbits and regular partitions of poisson manifolds. *Comm. Math. Phys.*, 283(3):729–748, 2008.

Ben Moonen. Group schemes with additional structures and Weyl group cosets. In *Moduli of abelian varieties (Texel Island, 1999)*, volume 195 of *Progr. Math.*, pages 255–298. Birkhäuser, Basel, 2001.

Ben Moonen and Torsten Wedhorn. Discrete invariants of varieties in positive characteristic. *Int. Math. Res. Not.*, (72):3855–3903, 2004.

Jean-Pierre Serre. Espaces fibrés algébriques. *Séminaire Claude Chevalley*, 3:1–37, 1958.

Peter Slodowy. *Simple singularities and simple algebraic groups*, volume 815 of *Lecture Notes in Mathematics*. Springer, Berlin, 1980.

T. A. Springer. *Linear algebraic groups*, volume 9 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, second edition, 1998.

T. A. Springer. Twisted conjugacy in simply connected groups. *Transform. Groups*, 11(3):539–545, 2006.

T. A. Springer. An extension of Bruhat’s lemma. *J. Algebra*, 313(1):417–427, 2007.

Robert Steinberg, *Endomorphisms of linear algebraic groups*. Memoirs of the American Mathematical Society, No. 80. American Mathematical Society, Providence, R.I., 1968.

A. Vasiu. Mod $p$ classification of Shimura $F$-crystals. *Math. Nachr.*, 283(8):1068–1113, 2010.

Eva Viehmann. Truncations of level 1 of elements in the loop group of a reductive group. arXiv:math/0907.2331, 2009.

Eva Viehmann and Torsten Wedhorn. Ekedahl-Oort and Newton strata for Shimura varieties of PEL type. arXiv:1011.3230v1, 2010.

Torsten Wedhorn. Specialization of $F$-Zips. arXiv:math/0507175v1, 2005.

Paul Ziegler. Zip data. Master’s thesis, ETH Zürich, 2010.
Euler Characteristics of Categories
and Homotopy Colimits

THOMAS M. FIORE, WOLFGANG LÜCK, AND ROMAN SAUER

Received: October 20, 2010

Communicated by Stephan Schwede

Abstract. In a previous article, we introduced notions of finiteness
obstruction, Euler characteristic, and $L^2$-Euler characteristic for wide
classes of categories. In this sequel, we prove the compatibility of
those notions with homotopy colimits of $I$-indexed categories where
$I$ is any small category admitting a finite $I$-CW-model for its $I$-
classifying space. Special cases of our Homotopy Colimit Formula
include formulas for products, homotopy pushouts, homotopy orbits,
and transport groupoids. We also apply our formulas to Haefliger
complexes of groups, which extend Bass–Serre graphs of groups to
higher dimensions. In particular, we obtain necessary conditions for
developability of a finite complex of groups from an action of a finite
group on a finite category without loops.

2010 Mathematics Subject Classification: Primary: 18F30, 19J05;
Secondary: 18G10, 19A49, 55U35, 19A22, 46L10.

Keywords and Phrases: finiteness obstruction, Euler characteristic
of a category, $L^2$-Euler characteristic, projective class group, homo-
topy colimits of categories, Grothendieck construction, spaces over a
category, Grothendieck fibration, complex of groups, small category
without loops.

0. Introduction and Statement of Results

In our previous paper [16], we presented a unified conceptual framework for
Euler characteristics of categories in terms of finiteness obstructions and pro-
jective class groups. Many excellent properties of our invariants stem from
the homological origins of our approach: the theory of modules over categories
and the dimension theory of modules over von Neumann algebras provide us
with an array of tools and techniques. In the present paper, we additionally
draw upon the homotopy theory of diagrams to prove the compatibility of our invariants with homotopy colimits.

If \( C : \mathcal{I} \to \text{CAT} \) is a diagram of categories (or more generally a pseudo functor into the 2-category of small categories), then our invariants of the homotopy colimit can be computed in terms of the invariants of the vertex categories \( C(i) \).

In particular, our Homotopy Colimit Formula, Theorem 4.1 states

\[
\chi(\text{hocolim}_\mathcal{I} C; R) = \sum_{n \geq 0} (-1)^n \cdot \sum_{\lambda \in \Lambda_n} \chi(C(i_\lambda); R)
\]

under certain hypotheses. The set \( \Lambda_n \) indexes the \( \mathcal{I} \)-\( n \)-cells of a finite \( \mathcal{I} \)-CW-model \( E\mathcal{I} \) for the \( \mathcal{I} \)-classifying space of \( \mathcal{I} \), that is, we have a functor \( E\mathcal{I} : \mathcal{I}^{op} \to \text{SPACES} \) which is inductively built by gluing finitely many cells of the form \( \text{mor}_\mathcal{I}(-, i_\lambda) \times D^n \) for \( \lambda \in \Lambda_n \), and moreover \( E\mathcal{I}(i) \simeq * \) for all objects \( i \) of \( \mathcal{I} \). Similar formulas hold for the finiteness obstruction, the functorial Euler characteristic, the functorial \( L^2 \)-Euler characteristic, and the \( L^2 \)-Euler characteristic.

Motivation for such a formula is provided by the classical Inclusion-Exclusion Principle: if \( A, B, \) and \( A \cap B \) are finite simplicial complexes, then one has

\[
\chi(A \cup B) = \chi(A) + \chi(B) - \chi(A \cap B).
\]

However, one cannot expect the Euler characteristic to be compatible with pushouts, even in the simplest cases. The pushout in \( \text{CAT} \) of the discrete categories

\[
\{ * \} \leftarrow \{ y, z \} \rightarrow \{ *' \}
\]

is a point, but \( \chi(\text{point}) \neq 1 + 1 - 2 \). On the other hand, their homotopy pushout in \( \text{CAT} \) is the category whose objects and nontrivial morphisms are pictured below.

\[
\begin{array}{c}
y \\
\downarrow \\
* \\
\uparrow \\
z
\end{array} \quad \quad \begin{array}{c}
s' \\
\end{array}
\]

The classifying space of this category has the homotopy type of \( S^1 \), so that

\[
\chi(\text{homotopy pushout}) = \chi(\{ * \}) + \chi(\{ *' \}) - \chi(\{ y, z \})
\]

is true. In fact, the formula for homotopy pushouts is a special case of (0.1): the category \( \mathcal{I} = \{ 1 \leftarrow 0 \rightarrow 2 \} \) admits a finite model with \( \Lambda_0 = \{ 1, 2 \} \) and \( \Lambda_1 = \{ 0 \} \), as constructed in Example 2.6. See Example 5.4 for the homotopy pushout formulas of the other invariants.

The Homotopy Colimit Formula in Theorem 4.1 has many applications beyond homotopy pushouts. Other special cases are formulas for Euler characteristics of products, homotopy orbits, and transport groupoids. Our formulas also have ramifications for the developability of Haefliger’s complexes of groups in geometric group theory. If a group \( G \) acts on an \( M_c \)-polyhedral complex by isometries preserving cell structure, and if each \( g \in G \) fixes each cell pointwise that \( g \) fixes setwise, then the quotient space is also an \( M_c \)-polyhedral complex.
Euler Characteristics and Homotopy Colimits

303

see Bridson–Haefliger [11, page 534]. Let us call the quotient $M_\kappa$-polyhedral complex $Q$. To each face $\sigma$ of $Q$, one can assign the stabilizer $G_\sigma$ of a chosen representative cell $\sigma$. This assignment, along with the various conjugated inclusions of groups obtained from face inclusions, is called the complex of groups associated to the group action. It is a pseudo functor from the poset of faces of $Q$ into groups. In the finite case, the Euler characteristic and $L^2$-Euler characteristic of the homotopy colimit can be computed in terms of the original complex and the order of the group. We prove this in Theorem 8.30. Homotopy colimits of complexes of groups play a special role in Haefliger’s theory, see the discussion after Definition 8.9.

In Section 1, we review the notions and results from [16] that we need in this sequel. Explanations of the finiteness obstruction, the functorial Euler characteristic, the Euler characteristic, the functorial $L^2$-Euler characteristic, and the necessary theorems are all contained in Section 1 in order to make the present paper self-contained. Section 2 is dedicated to an assumption in the Homotopy Colimit Formula, namely the requirement that a finite $I$-CW-model exists for the $I$-classifying space of $I$. We recall the notion of $I$-CW-complex, present various examples, and prove that finite models are preserved under equivalences of categories. Homotopy colimits of diagrams of categories are recalled in Section 3. The homotopy colimit construction in CAT is the same as the Grothendieck construction, or the category of elements. Thomason proved that the homotopy colimit construction has the expected properties. We prove our main theorem, the Homotopy Colimit Formula, in Section 4, work out various examples in Section 5, and derive the generalized Inclusion-Exclusion Principle in Section 6. We review the groupoid cardinality of Baez–Dolan and the Euler characteristic of Leinster in Section 7 and compare our Homotopy Colimit Formula with Leinster’s compatibility with Grothendieck fibrations in terms of weightings. We apply our results to Haefliger complexes of groups in Section 8 to prove Theorems 8.30 and 8.35, which express Euler characteristics of complexes of groups associated to group actions in terms of the initial data.

Acknowledgements. All three authors were supported by the Sonderforschungsbereich 878 – Groups, Geometry and Actions – and the Leibniz-Preis of Wolfgang Lück.

Thomas M. Fiore was supported at the University of Chicago by NSF Grant DMS-0501208. At the Universitat Autònoma de Barcelona he was supported by grant SB2006-0085 of the Spanish Ministerio de Educación y Ciencia under the Programa Nacional de ayudas para la movilidad de profesores de universidad e investigadores españoles y extranjeros. Thomas M. Fiore also thanks the Centre de Recerca Matemàtica in Bellaterra (Barcelona) for its hospitality during the CRM Research Program on Higher Categories and Homotopy Theory in 2007-2008, where he heard Tom Leinster speak about Euler characteristics. He also thanks the Max Planck Insitut für Mathematik for its hospitality and support during his stay in Summer 2010.
1. The Finiteness Obstruction and Euler Characteristics

We quickly recall the main definitions and results needed from our first paper [16] in order to make this article as self-contained as possible. See [16] for proofs and more detail.

Throughout this paper, let $\Gamma$ be a category and $R$ an associative, commutative ring with identity. The first ingredient we need is the theory of modules over categories developed by Lück [23], and recalled in [16]. An $R\Gamma$-module is a contravariant functor from $\Gamma$ into the category of left $R$-modules. For example, if $\Gamma$ is a group $G$ viewed as a one-object category, then an $R\Gamma$-module is the same as a right module over the group ring $RG$. An $R\Gamma$-module $P$ is projective if it is projective in the usual sense of homological algebra, that is, for every surjective $R\Gamma$-morphism $p: M \to N$ and every $R\Gamma$-morphism $f: P \to N$ there exists an $R\Gamma$-morphism $\overline{f}: P \to M$ such that $p \circ \overline{f} = f$. An $R\Gamma$-module $M$ is finitely generated if there is a surjective $R\Gamma$-morphism $B(C) \to M$ from an $R\Gamma$-module $B(C)$ that is free on a collection $C$ of sets indexed by $\text{ob}(\Gamma)$ such that $\coprod_{x \in \text{ob}(\Gamma)} C_x$ is finite. Explicitly, the free $R\Gamma$-module on the $\text{ob}(\Gamma)$-set $C$ is

$$B(C) := \bigoplus_{x \in \text{ob}(\Gamma)} \bigoplus_{C_x} R\text{mor}_\Gamma(?;x).$$

A contravariant $R\Gamma$-module may be tensored with a covariant $R\Gamma$-module to obtain an $R$-module: if $M: \Gamma^{\text{op}} \to R\text{-MOD}$ and $N: \Gamma \to R\text{-MOD}$ are functors, then the tensor product $M \otimes_{R\Gamma} N$ is the quotient of the $R$-module

$$\bigoplus_{x \in \text{ob}(\Gamma)} M(x) \otimes_R N(x)$$

**Reference:**

Thomas M. Fiore, Wolfgang Lück, and Roman Sauer

Documenta Mathematica 16 (2011) 301–354
by the $R$-submodule generated by elements of the form
\[(M(f)m) \otimes n - m \otimes (N(f)n)\]
where $f : x \to y$ is a morphism in $\Gamma$, $m \in M(y)$, and $n \in N(x)$.

Finite projective resolutions of the constant $RT$-module $R$ play a special role in our theory of Euler characteristic for categories. A resolution $P_n$ of an $RT$-module $M$ is said to be finite projective if it has finite length and each $P_n$ is finitely generated and projective. We say that a category $\Gamma$ is of type $(FP_R)$ if the constant $RT$-module $R$: $\Gamma^\text{op} \to R\text{-MOD}$ with value $R$ admits a finite projective resolution. Categories in which every endomorphism is an isomorphism, the so-called EI-categories, provide important examples. Finite EI-categories in which $|\text{aut}(x)|$ is invertible in $R$ for each object $x$ are of type $(FP_R)$. Further examples of categories of type $(FP_R)$ include categories $\Gamma$ which admit a finite $\Gamma$-CW-model for the classifying $\Gamma$-space $E\Gamma$ (see Section 2 and Examples 2.4, 2.5, 2.6, and 2.7). In fact, such categories $\Gamma$ are even of type $(FF_R)$: the cellular chains on a finite $\Gamma$-CW-model for $E\Gamma$ provide a finite free resolution of $R$. In general, if a category is of type $(FF_R)$, then it is of type $(FF_R)$ for any associative, commutative ring $R$ with identity.

A home for the finiteness obstruction of a category $\Gamma$ is provided by the projective class group $K_0(R\Gamma)$. The generators of this abelian group are the isomorphism classes of finitely generated projective $R\Gamma$-modules and the relations are given by expressions $[P_0] - [P_1] + [P_2] = 0$ for every exact sequence $0 \to P_0 \to P_1 \to P_2 \to 0$ of finitely generated projective $R\Gamma$-modules.

**Definition 1.2** (Finiteness obstruction of a category). Let $\Gamma$ be a category of type $(FP_R)$ and $P_n$ a finite projective resolution of the constant $RT$-module $R$. The finiteness obstruction of $\Gamma$ with coefficients in $R$ is
\[o(\Gamma; R) := \sum_{n \geq 0} (-1)^n \cdot [P_n] \in K_0(R\Gamma).\]
We also use the notation $[R]$, or simply $[R]$, to denote the finiteness obstruction $o(\Gamma; R)$. The finiteness obstruction, when it exists, does not depend on the choice $P_n$ of finite projective resolution of $R$.

The finiteness obstruction is compatible with most everything one could hope for. If $F: \Gamma_1 \to \Gamma_2$ is a right adjoint, and $\Gamma_1$ is of type $(FP_R)$, then $\Gamma_2$ is of type $(FP_R)$ and $F_* o(\Gamma_1; R) = o(\Gamma_2; R)$ (here the group homomorphism $F_*$ is induced by induction with $F$). Since an equivalence of categories is a right adjoint (and also a left adjoint), a particular instance of the previous sentence is: if $F: \Gamma_1 \to \Gamma_2$ is an equivalence of categories, then $\Gamma_1$ is of type $(FP_R)$ if and only if $\Gamma_2$ is, and in this case $F_* o(\Gamma_1; R) = o(\Gamma_2; R)$. The finiteness obstruction is also compatible with finite coproducts of categories, finite products of categories, restriction along admissible functors, and homotopy colimits, as we prove in Theorem 4.1. If $G$ is a finitely presented group of type $(FP_Z)$, then Wall’s finiteness obstruction $o(BG)$ is the same as $o(\hat{G}; Z)$, which is the finiteness obstruction of $\hat{G}$ viewed as a one-object category $\hat{G}$ with morphisms $G$. 

**Documenta Mathematica 16 (2011) 301–354**
The finiteness obstruction in Definition 1.2 is a special case of the finiteness obstruction of a finitely dominated $R\Gamma$-chain complex $C$, denoted $o(C) \in K_0(R\Gamma)$.

The image of $o(C)$ in the reduced $K$-theory $K_\theta(R\Gamma)$ vanishes if and only if $C$ is $R\Gamma$-homotopy equivalent to a finite free $R\Gamma$-chain complex, see [23, Chapter 11].

We will occasionally work with directly finite categories. A category is called directly finite if for any two objects $x$ and $y$ and morphisms $u: x \to y$ and $v: y \to x$ the implication $vu = id_x \implies uv = id_y$ holds. If $\Gamma_1$ and $\Gamma_2$ are equivalent categories, then $\Gamma_1$ is directly finite if and only if $\Gamma_2$ is directly finite.

Examples of directly finite categories include groupoids, and more generally EI-categories.

A key result in the theory of modules over an EI-category is Lück’s splitting of the projective class group of $\Gamma$ into the projective class groups of the automorphism groups $\text{aut}_\Gamma(x)$, one ach isomorphism class of objects. We next recall the relevant maps and notation. For $x \in \text{ob}(\Gamma)$, we denote $R\text{aut}_\Gamma(x)$ by $R[x]$ for simplicity. The splitting functor at $x \in \text{ob}(\Gamma)$

$$S_x: \text{MOD}-R\Gamma \to \text{MOD}-R[x],$$

maps an $R\Gamma$-module $M$ to the quotient of the $R$-module $M(x)$ by the $R$-submodule generated by all images of $R$-module homomorphisms $M(u): M(y) \to M(x)$ induced by all non-invertible morphisms $u: x \to y$ in $\Gamma$. The right $R[x]$-module structure on $M(x)$ induces a right $R[x]$-module structure on $S_x M$. Note that $S_x M$ is an $R[x]$-module, not an $R\Gamma$-module. The functor $S_x$ respects direct sums, sends epimorphisms to epimorphisms, and sends free modules to free modules. If $\Gamma$ is directly finite, then $S_x$ also preserves finitely generated and projective. The extension functor at $x \in \text{ob}(\Gamma)$

$$E_x: \text{MOD}-R[x] \to \text{MOD}-R\Gamma$$

maps an $R[x]$-module $N$ to the $R\Gamma$-module $N \otimes_{R[x]} R\text{mor}_\Gamma(?,x)$. The functor $E_x$ respects direct sums, sends epimorphisms to epimorphisms, sends free modules to free modules, and preserves finitely generated and projective. If $\Gamma$ is directly finite, and $P$ is a projective $R[x]$-module, then there is a natural isomorphism $P \cong S_x E_x P$ compatible with direct sums.

**Theorem 1.5** (Splitting of $K_0(R\Gamma)$ for EI-categories, Theorem 10.34 on page 196 of Lück [23]). If $\Gamma$ is an EI-category, then the group homomorphisms

$$K_0(R\Gamma) \xrightarrow{S} \text{Split } K_0(R\Gamma) := \bigoplus_{\varpi \in \text{iso}(\Gamma)} K_0(\text{Raut}_\Gamma(x))$$

defined by

$$S[P] = \{[S_x P] \mid \varpi \in \text{iso}(\Gamma)\}$$

and

$$E\{[Q_x] \mid \varpi \in \text{iso}(\Gamma)\} = \sum_{\varpi \in \text{iso}(\Gamma)} [E_x Q_x],$$
are isomorphisms and inverse to one another. They are covariantly natural with respect to functors between EI-categories.

**Remark 1.6.** If \( \Gamma \) is not an EI-category, then the splitting homomorphism \( S : K_0(R\Gamma) \to \text{Split} K_0(R\Gamma) \) may not be an isomorphism. However, \( S \) is covariantly natural with respect to functors between directly finite categories, see [16, Lemma 3.15].

The splitting functors \( S_x \) allow us to define the notion of \( R\Gamma \)-rank \( \text{rk}_R \) for finitely generated \( R\Gamma \)-modules, which in turn allows the definition of the functorial Euler characteristic, as we explain next. We assume a fixed notion of a rank \( \text{rk}_R(N) \in \mathbb{Z} \) for finitely generated \( R \)-modules \( N \) such that \( \text{rk}_R(R) = 1 \) and \( \text{rk}_R(N_1) = \text{rk}_R(N_0) + \text{rk}_R(N_2) \) for any sequence \( 0 \to N_0 \to N_1 \to N_2 \to 0 \) of finitely generated \( R \)-modules. If \( R \) is a commutative principal ideal domain, we use \( \text{rk}_R(N) := \dim_F(F \otimes_R N) \), where \( F \) is the quotient field of \( R \).

**Definition 1.7 (Rank of a finitely generated \( R\Gamma \)-module).** If \( M \) is a finitely generated \( R\Gamma \)-module \( M \), then its \( R\Gamma \)-rank is

\[
\text{rk}_R(M) := \left\{ \text{rk}_R(S_x M \otimes_{R[x]} R) \mid x \in \text{iso}(\Gamma) \right\} \in U(\Gamma).
\]

**Definition 1.8 (The (functorial) Euler characteristic of a category).** Suppose that \( \Gamma \) is of type (FP\(_R\)). The functorial Euler characteristic of \( \Gamma \) with coefficients in \( R \) is the image of the finiteness obstruction \( o(\Gamma; R) \in K_0(R\Gamma) \) under the homomorphism \( \text{rk}_R : K_0(R\Gamma) \to U(\Gamma) \), that is

\[
\chi_f(\Gamma; R) := \text{rk}_R o(\Gamma; R) = \left\{ \sum_{n \geq 0} (-1)^n \text{rk}_R(S_x P_n \otimes_{R[x]} R) \mid x \in \text{iso}(\Gamma) \right\} \in U(\Gamma),
\]

where \( P_* \) is any finite projective \( R\Gamma \)-resolution of the constant \( R\Gamma \)-module \( R \).

The Euler characteristic of \( \Gamma \) with coefficients in \( R \) is the sum of the components of the functorial Euler characteristic, that is,

\[
\chi(\Gamma; R) := \epsilon(\chi_f(\Gamma; R)) = \sum_{x \in \text{iso}(\Gamma)} \sum_{n \geq 0} (-1)^n \text{rk}_R(S_x P_n \otimes_{R[x]} R).
\]

For example, if \( \mathcal{G} \) is a finite groupoid, then \( \chi_f(\mathcal{G}) \in U(\mathcal{G}) \) is \( (1, 1, \ldots, 1) \), and \( \chi(\mathcal{G}) \) counts the isomorphism classes of objects, or equivalently the connected components, of \( \mathcal{G} \).

**Theorem 1.9** (Theorem 4.20 of Fiore–Lück–Sauer [16]). Let \( R \) be a Noetherian ring and \( \Gamma \) a directly finite category of type (FP\(_R\)). Then the Euler characteristic and topological Euler characteristic of \( \Gamma \) agree. That is, \( H_n(B\Gamma; R) \) is a finitely generated \( R \)-module for every \( n \geq 0 \), there exists a natural number \( d \)
with $H_n(B\Gamma; R) = 0$ for all $n > d$, and
\[
\chi(\Gamma; R) = \chi(B\Gamma; R) = \sum_{n \geq 0} (-1)^n \cdot \text{rk}_R(H_n(B\Gamma; R)) \in \mathbb{Z}.
\]

Here $\chi(\Gamma; R)$ is defined in Definition 1.8 and $B\Gamma$ denotes the geometric realization of the nerve of $\Gamma$.

The functorial Euler characteristic and Euler characteristic have many desirable properties. They are invariant under equivalence of categories and are compatible with finite products and finite coproducts. As we prove in Theorem 4.1, they are also compatible with homotopy colimits.

The $L^2$-Euler characteristic, which is in some sense the better invariant, can be defined similarly by taking $R = \mathbb{C}$ and using the $L^2$-rank $\text{rk}^{(2)}_\Gamma$ rather than the $R\Gamma$-rank. For this we need group von Neumann algebras and their dimension theory from Lück [24] and [25], as recalled in our first paper [16] for the purpose of Euler characteristics. If $G$ is a group, its group von Neumann algebra
\[
N(G) = B(l^2(G))^G
\]
is the algebra of bounded operators on $l^2(G)$ that are equivariant with respect to the right $G$-action. If $G$ is finite, $N(G)$ is the group ring $\mathbb{C}[G]$. In any case, the group ring $\mathbb{C}[G]$ embeds as a subring of $N(G)$ by sending $g \in G$ to the isometric $G$-equivariant operator $l^2(G) \to l^2(G)$ given by left multiplication with $g$. In particular, we can view $N(G)$ as a $\mathbb{C}[G]$-$N(G)$-bimodule and tensor $\mathbb{C}[G]$-modules on the right with $N(G)$. If $G$ is the automorphism group of an object in $\Gamma$, then we write $N(x)$ for $N(\text{aut}_\Gamma(x))$.

The von Neumann dimension, $\text{dim}_{N(G)}$, is a function that assigns to every right $N(G)$-module $M$ a non-negative real number of $\infty$. It is the unique such function which satisfies Hattori-Stallings rank, additivity, cofinality, and continuity. If $G$ is a finite group, then $N(G) = \mathbb{C}[G]$ and we get for a $\mathbb{C}[G]$-module $M$ the von Neumann dimension
\[
\text{dim}_{N(G)}(M) = \frac{1}{|G|} \cdot \text{dim}_\mathbb{C}(M),
\]
where $\text{dim}_\mathbb{C}$ is the dimension of $M$ viewed as a complex vector space. A category $\Gamma$ is said to be of type $(L^2)$ if for one (and hence every) projective $\mathbb{C}[\Gamma]$-resolution $P_*$ of the constant $\mathbb{C}[\Gamma]$-module $\mathbb{C}$ we have
\[
\sum_{x \in \text{iso}(\Gamma)} \sum_{n \geq 0} \text{dim}_{N(x)} H_n(S_x P_\ast \otimes_{\mathbb{C}[x]} N(x)) < \infty.
\]

Note that the projective resolution $P_\ast$ of $\mathbb{C}$ is not required to be of finite length, nor finitely generated. Examples of categories of type $(L^2)$ include finite EI-categories, in particular finite posets and finite groupoids. Infinite categories can also be of type $(L^2)$, for example any (small) groupoid with finite automorphism groups such that
\[
\sum_{x \in \text{iso}(\Gamma)} \frac{1}{|\text{aut}_\Gamma(x)|} < \infty
\]

(1.10)
holds is of type \((L^2)\). The condition of type \((L^2)\) is weaker than \((\text{FP}_C)\), since any directly finite category of type \((\text{FP}_C)\) is also of type \((L^2)\).

**Definition 1.11** (The (functorial) \(L^2\)-Euler characteristic of a category). Suppose that \(\Gamma\) is of type \((L^2)\). Define

\[
U^{(1)}(\Gamma) := \left\{ \sum_{x \in \text{iso}(\Gamma)} r_x \cdot \overline{x} \, \middle| \, r_x \in \mathbb{R}, \sum_{x \in \text{iso}(\Gamma)} |r_x| < \infty \right\} \subseteq \prod_{x \in \text{iso}(\Gamma)} \mathbb{R}.
\]

The functorial \(L^2\)-Euler characteristic of \(\Gamma\) is

\[
\chi^{(2)}_f(\Gamma) := \left\{ \sum_{n \geq 0} (-1)^n \dim_{\mathcal{N}(x)} H_n(S_x P_\ast \otimes_{\mathbb{C}[x]} \mathcal{N}(x)) \mid \bar{x} \in \text{iso}(\Gamma) \right\} \subseteq U^{(1)}(\Gamma),
\]

where \(P_\ast\) is any projective \(\mathbb{C}\Gamma\)-resolution of the constant \(\mathbb{C}\Gamma\)-module \(\mathbb{C}\). The \(L^2\)-Euler characteristic of \(\Gamma\) is the sum over \(\bar{x} \in \text{iso}(\Gamma)\) of the components of the functorial Euler characteristic, that is,

\[
\chi^{(2)}(\Gamma) := \sum_{x \in \text{iso}(\Gamma)} \sum_{n \geq 0} (-1)^n \dim_{\mathcal{N}(x)} H_n(S_x P_\ast \otimes_{\mathbb{C}[x]} \mathcal{N}(x)).
\]

If \(\mathcal{G}\) is a groupoid such that \((1.10)\) holds, then the functorial \(L^2\)-Euler characteristic \(\chi^{(2)}_f(\mathcal{G}) \in \prod_{\bar{\xi} \in \text{iso}(\mathcal{G})} \mathbb{R}\) has at \(\bar{x} \in \text{iso}(\mathcal{G})\) the value \(1/|\text{aut}_\mathcal{G}(x)|\). The \(L^2\)-Euler characteristic is

\[
\chi^{(2)}(\mathcal{G}) = \sum_{\bar{x} \in \text{iso}(\mathcal{G})} \frac{1}{|\text{aut}_\mathcal{G}(x)|}.
\]

See Lemma \(7.5\) for an explicit formula for \(\chi^{(2)}(\Gamma)\) in the case of a finite, skeletal EI-category \(\Gamma\) in which the left \(\text{aut}_\Gamma(y)\)-action on \(\text{mor}_\Gamma(x, y)\) is free for every two objects \(x, y \in \text{ob}(\Gamma)\).

**Definition 1.13** \((L^2\)-rank of a finitely generated \(\mathbb{C}\Gamma\)-module). Let \(M\) be a finitely generated \(\mathbb{C}\Gamma\)-module \(M\). Its \(L^2\)-rank is

\[
\text{rk}^{(2)}_\Gamma(M) := \{ \dim_{\mathcal{N}(x)}(S_x M \otimes_{\mathbb{C}[x]} \mathcal{N}(x)) \mid \bar{x} \in \text{iso}(\Gamma) \} \subseteq U(\Gamma) \otimes_{\mathbb{Z}} \mathbb{R} = \bigoplus_{\text{iso}(\Gamma)} \mathbb{R}.
\]

**Theorem 1.14** (Relating the finiteness obstruction and the \(L^2\)-Euler characteristic, Theorem 5.22 of Fiore–Lück–Sauer \([10]\)). Suppose that \(\Gamma\) is a directly finite category of type \((\text{FP}_C)\). Then \(\Gamma\) is of type \((L^2)\) and the image of the finiteness obstruction \(o(\Gamma; C)\) (see Definition 1.2) under the homomorphism

\[
\text{rk}^{(2)}_\Gamma : K_0(\mathbb{C}\Gamma) \to U(\Gamma) \otimes_{\mathbb{Z}} \mathbb{R} = \bigoplus_{\text{iso}(\Gamma)} \mathbb{R}
\]

is the functorial \(L^2\)-Euler characteristic \(\chi^{(2)}_f(\Gamma)\).

The \(L^2\)-Euler characteristic agrees with the groupoid cardinality of Baez–Dolan \([4]\) and the Euler characteristic of Leinster \([21]\) in certain cases, see Lemma \(7.3\) and Section \(7\). In particular, the Baez–Dolan groupoid cardinality of a groupoid
1.10. However, the Baez–Dolan groupoid cardinality and Leinster’s Euler characteristic \( \chi_\Gamma(\Gamma) \) only depend on the underlying graph of \( \Gamma \), whereas our invariants truly depend on the category structure. For instance, \( \chi_\Gamma \) is \( \frac{1}{2} \) for both the two-element monoid \( \langle \mathbb{Z}/2, \times \rangle \) and the two-element group \( \langle \mathbb{Z}/2, + \rangle \), whereas \( \chi^{(2)} \) is 1 respectively \( \frac{1}{2} \). The distinction can already be seen on the level of the finiteness obstructions. The Euler characteristic \( \chi(\cdot) \) and topological Euler characteristic \( \chi(B\cdot) \) can also distinguish categories with the same underlying directed graph as in the following example. For \( S = \{1, 2, 3, 4\} \), \( G_1 = \langle (1234) \rangle \), \( G_2 = \langle (12), (34) \rangle \), and \( k = 1, 2 \), let \( \Gamma_k \) be the EI-category with objects \( x \) and \( y \) and \( \text{mor}(x, y) := S \), \( \text{mor}(x, x) := \{\text{id}_x\} \), \( \text{mor}(y, y) := G_k \), and \( \text{mor}(y, x) = \emptyset \). Composition in \( \Gamma_k \) is the composition in \( G_k \) and the left \( G_k \)-action on \( S \), that is, \( \Gamma_k \) is the EI-category associated to the respective \( G_k \{1\}\)-biset \( S \) as in Subsection 6.4 of Fiore–Lück–Sauer [10]. Then \( \Gamma_1 \) and \( \Gamma_2 \) have the same underlying directed graph but \( \chi(\Gamma_1; Q) = \chi(B\Gamma_1; Q) = 1 \) and \( \chi(\Gamma_2; Q) = \chi(B\Gamma_2; Q) = 0 \) by Theorem 6.23 (iii) of Fiore–Lück–Sauer [10]. An infinite example of categories with the same underlying graph but different Euler characteristics is provided by the groups \( \mathbb{Z} \) and \( \mathbb{Z} \star \mathbb{Z} \), each of which admits a finite \( \Gamma \)-CW-model for its respective \( \Gamma \)-classifying space. The categories \( \mathbb{Z} \) and \( \mathbb{Z} \star \mathbb{Z} \) have the same underlying directed graph, but we have \( \chi^{(2)}(\mathbb{Z}) = 0 \neq \chi^{(2)}(\mathbb{Z} \star \mathbb{Z}) \), and similarly for \( \chi \). Typically, the Euler characteristic of a category \( \Gamma_{\text{free}} \) free on a directed graph \( (V, E) \) is the same as the Euler characteristic of the directed graph \( (V, E) \). For the topological Euler characteristic this is clearly true, since \( B\Gamma_{\text{free}} \) is homotopy equivalent to the topological realization \( [(V, E)] \). If \( \Gamma_{\text{free}} \) is directly finite and \( R \) is Noetherian, then we also have \( \chi(\Gamma_{\text{free}}) = \chi([(V, E)]) \) by Theorem 1.9. For example for the directed graph with one vertex and one arrow we have \( \chi(\mathbb{N}) = 0 = \chi(S^1) \). The functorial \( L^2 \)-Euler characteristic and the \( L^2 \)-Euler characteristic have many desirable properties. They are invariant under equivalence of categories and are compatible with finite products, finite coproducts, and isofibrations and coverings between finite groupoids. We prove in Theorem 1.11 the compatibility with homotopy colimits. In the case of a group \( G \), the \( L^2 \)-Euler characteristic of \( \hat{G} \) coincides with the classical \( L^2 \)-Euler characteristic of \( G \), which is \( 1/|G| \) when \( G \) is finite. The \( L^2 \)-Euler characteristic is also closely related to the geometry and topology of the classifying space for proper \( G \)-actions, namely the functorial \( L^2 \)-Euler characteristic of the proper orbit category \( \text{Or} (G) \) is equal to the equivariant Euler characteristic of the classifying space \( E\hat{G} \) for proper \( G \)-actions, whenever \( E\hat{G} \) admits a finite \( G \)-CW-model.

The question arises: what are sufficient conditions for the Euler characteristic and \( L^2 \)-Euler characteristic to coincide with the Euler characteristic of the classifying space? This is answered in the following Theorem.

**Theorem 1.15 (Invariants agree for directly finite and type (FF2), Theorem 5.25 of Fiore–Lück–Sauer [10].)** Suppose \( \Gamma \) is directly finite and of type (FF2). Then the functorial \( L^2 \)-Euler characteristic of Definition 1.11 coincides with the functorial Euler characteristic of Definition 1.3 for any associative,
commutative ring $R$ with identity

$$\chi_f^{(2)}(\Gamma) = \chi_f(\Gamma; R) \in U(\Gamma) \subseteq U^{(1)}(\Gamma),$$

and thus $\chi^{(2)}(\Gamma) = \chi(\Gamma; R)$ in Definition 1.14 and Definition 1.15.

If $R$ is additionally Noetherian, then

$$\chi(B\Gamma; R) = \chi(\Gamma; R) = \chi^{(2)}(\Gamma).$$

Moreover, if $\Gamma$ is merely of type $(FF_C)$ rather than $(FF_Z)$, then equation (1.10) holds for any Noetherian ring $R$ containing $C$.

Any category $\Gamma$ which admits a finite $\Gamma$-CW-model in the sense of Section 2 is of type $(FF_R)$ for any ring $R$, by an application of the cellular $R$-chain functor. Thus, Theorem 1.15 applies to any directly finite category $\Gamma$ which admits a finite $\Gamma$-CW-model. For example, finite categories without loops are directly finite and admit finite models (Lemma 5.4 and Theorem 5.5), so equation (1.16) holds for instance for $\{ j \equiv k \}$, $\{ k \leftarrow j \rightarrow \ell \}$, and finite posets. The monoid $\mathbb{N}$ and group $\mathbb{Z}$, viewed as one-object categories $\hat{\mathbb{N}}$ and $\hat{\mathbb{Z}}$, are also directly finite and admit finite models (see Example 2.3), so we have

$$0 = \chi(S^1; R) = \chi(B\hat{\mathbb{N}}; R) = \chi(\hat{\mathbb{N}}; R) = \chi^{(2)}(\hat{\mathbb{N}})$$

and

$$0 = \chi(S^1; R) = \chi(B\hat{\mathbb{Z}}; R) = \chi(\hat{\mathbb{Z}}; R) = \chi^{(2)}(\hat{\mathbb{Z}})$$

($B\hat{\mathbb{N}} \to B\hat{\mathbb{Z}} \simeq S^1$ is a homotopy equivalence by Quillen’s Theorem A, see Rabrenović [35] Proposition 10). The equations $\chi(\hat{\mathbb{N}}; R) = 0 = \chi^{(2)}(\hat{\mathbb{N}})$ and $\chi(\hat{\mathbb{Z}}; R) = 0 = \chi^{(2)}(\hat{\mathbb{Z}})$ also follow from Example 5.3 since the finite models for $\hat{\mathbb{N}}$ and $\hat{\mathbb{Z}}$ in Example 2.3 each have one $I$-0-cell and one $I$-1-cell.

We may use Theorem 1.15 to obtain an explicit formula for Euler characteristics of finite categories without loops as follows. Let $\Gamma$ be a finite category without loops, and choose a skeleton $\Gamma'$. Let $c_n(\Gamma')$ denote the number of paths

$$i_0 \to i_1 \to i_2 \to \cdots \to i_n$$

doing of $n$-many non-identity morphisms in $\Gamma'$. Then $c_n(\Gamma')$ is the number of $n$-cells in the CW-complex $B\Gamma'$, and we have

$$\chi(\Gamma; R) = \chi^{(2)}(\Gamma) = \chi(B\Gamma; R) = \chi(B\Gamma'; R) = \sum_{n \geq 0} (-1)^n c_n(\Gamma').$$

See [21] Corollary 1.5] for a different derivation of this formula for Leinster’s Euler characteristic $\chi_L(\Gamma)$ in the case $\Gamma$ was already skeletal. See also Examples 5.3 and 5.7 where skeletality of $\hat{\mathbb{Z}}$ is not required.

Remark 1.18 (Homotopy Invariance). If $F: \Gamma_1 \to \Gamma_2$ is a functor such that $BF$ is a homotopy equivalence, and both $\Gamma_1$ and $\Gamma_2$ are of type $(FP_R)$, and if

$$\chi(\Gamma_1; R) = \chi(B\Gamma_1; R) \quad \text{and} \quad \chi(\Gamma_2; R) = \chi(B\Gamma_2; R),$$

then clearly $\chi(\Gamma_1; R) = \chi(\Gamma_2; R)$. However, it is possible for two categories to be homotopy equivalent, one of which is $(FP_R)$ and the other is not, so
that one has a notion of Euler characteristic and the other does not. In Section 10 of Fiore–Lück–Sauer \[16\] such an example is discussed.

2. Spaces over a Category

An important hypothesis in our Homotopy Colimit Formula involves the idea of a space over a category, see Davis–Lück \[14\]. Namely, we assume that the indexing category \(I\) for the diagram \(C\) of categories admits a finite \(I\)-CW-model for its \(I\)-classifying space. Essentially this means it is possible to functorially assign a contractible \(CW\)-complex \(E_I(i)\) to each \(i \in \text{ob}(I)\), and moreover, these local \(CW\)-complexes are constructed globally by gluing \(I\)-n-cells of the form \(\text{mor}_I(−, i_λ) \times D^n\) onto the already globally constructed \((n−1)\)-skeleton \(E_I\). The Homotopy Colimit Formula then expresses the invariants of the homotopy colimit of \(C\) in terms of the invariants of the categories \(C(i_λ)\) at the base objects \(i_λ\) for \(E_I\).

The gluing described above takes place in the more general category of \(I\)-spaces. A (contravariant) \(I\)-space is a contravariant functor from \(I\) to the category \(\text{SPACES}\) of (compactly generated) topological spaces. As usual, we will always work in the category of compactly generated spaces (see Steenrod \[39\]). A map between \(I\)-spaces is a natural transformation. Given an object \(i \in \text{ob}(I)\), we obtain an \(I\)-space \(\text{mor}_I(?, i)\) which assigns to an object \(j\) the discrete space \(\text{mor}_I(j, i)\).

The next definition is taken from Davis–Lück \[14\], Definition 3.2, where an \(I\)-CW-complex is called a free \(I\)-CW-complex and we will omit the word free here. The more general notion of \(I\)-CW-complex was defined by Dror Farjoun \[15\], 1.16 and 2.1. See also Piacenza \[34\].

**Definition 2.1 (\(I\)-CW-complex).** A (contravariant) \(I\)-CW-complex \(X\) is a contravariant \(I\)-space \(X\) together with a filtration

\[
\emptyset = X_{−1} \subset X_0 \subset X_1 \subset X_2 \subset \ldots \subset X_n \subset \ldots \subset X = \bigcup_{n \geq 0} X_n
\]

such that \(X = \text{colim}_{n \to \infty} X_n\) and for any \(n \geq 0\) the \(n\)-skeleton \(X_n\) is obtained from the \((n−1)\)-skeleton \(X_{n−1}\) by attaching \(I\)-\(n\)-cells, i.e., there exists a pushout of \(I\)-spaces of the form

\[
\begin{array}{ccc}
\prod_{\lambda \in \Lambda_n} \text{mor}_I(−, i_λ) \times S^{n−1} & \longrightarrow & X_{n−1} \\
\downarrow & & \downarrow \\
\prod_{\lambda \in \Lambda_n} \text{mor}_I(−, i_λ) \times D^n & \longrightarrow & X_n
\end{array}
\]

where the vertical maps are inclusions, \(\Lambda_n\) is an index set, and the \(i_λ\)-s are objects of \(I\). In particular, \(X_0 = \prod_{\lambda \in \Lambda_0} \text{mor}_I(−, i_λ)\).

We refer to the inclusion functor \(\text{mor}_I(−, i_λ) \times (D^n − S^{n−1}) \to X\) as an \(I\)-\(n\)-cell based at \(i_λ\).
An $I$-CW-complex has dimension $\leq n$ if $X = X_n$. We call $X$ finite dimensional if there exists an integer $n$ with $X = X_n$. It is called finite if it is finite dimensional and $A_n$ is finite for every $n \geq 0$.

The definition of a covariant $I$-CW-complex is analogous.

**Definition 2.2** (Classifying $I$-space). A model for the classifying $I$-space $EI$ is an $I$-CW-complex $EI$ such that $EI(i)$ is contractible for all objects $i$.

The universal property of $EI$ is that for any $I$-CW-complex $X$ there is up to homotopy precisely one map of $I$-spaces from $X$ to $EI$. In particular two models for $EI$ are $I$-homotopy equivalent (see Davis–Lück [14, Theorem 3.4]). A model for the usual classifying space $BI$ is given by $EI \otimes_I \{\bullet\}$ (see [14, Definition 3.10]), where $\{\bullet\}$ is the constant covariant $I$-space with value the one point space and $\otimes_I$ denotes the tensor product of a contravariant and a covariant $I$-space as follows (see [14, Definition 1.4]).

**Definition 2.3** (Tensor product of a contravariant and a covariant $I$-space). Let $X$ be a contravariant $I$-space and $Y$ a covariant $I$-space. The tensor product of $X$ and $Y$ is

$$X \otimes_I Y = \left( \prod_{i \in I} X(i) \times Y(i) \right) / \sim$$

where $(X(\phi)(x), y) \sim (x, Y(\phi)y)$ for all morphisms $\phi : i \to j$ in $I$ and points $x \in X(j)$ and $y \in Y(i)$.

We present some examples of classifying $I$-spaces for various categories $I$.

**Example 2.4.** If $I$ has a terminal object $t$, then a finite model for the classifying $I$-space $EI$ is simply $\text{mor}_I(-, t)$.

**Example 2.5.** Let $I = \{j \Rightarrow k\}$ be the category consisting of two objects and a single pair of parallel arrows between them. All other morphisms are identity morphisms. We obtain a finite model $X$ for the classifying $I$-space $EI$ as follows. The $I$-CW-space $X$ has a single $I$-0-cell based at $k$ and a single $I$-1-cell based at $j$. The gluing map $\text{mor}_I(-, j) \times S^0 \to \text{mor}_I(-, k)$ is induced by the two parallel arrows $j \Rightarrow k$. Then $X(j) = D^1 \simeq *$ and $X(k) = *$.

**Example 2.6.** Let $I = \{k \leftarrow j \to \ell\}$ be the category with objects $j$, $k$ and $\ell$, and precisely one morphism from $j$ to $k$ and one morphism from $j$ to $\ell$. All other morphisms are identity morphisms. A finite model for $EI$ is given by the $I$-CW-complex with precisely two $I$-0-cells $\text{mor}_I(?, k)$ and $\text{mor}_I(?, \ell)$ and precisely one $I$-1-cell $\text{mor}_I(?, j) \times D^1$ whose attaching map $\text{mor}_I(?, j) \times S^0 \to \text{mor}_I(?, k)$ is the disjoint union of the canonical maps $\text{mor}_I(?, j) \to \text{mor}_I(?, k)$ and $\text{mor}_I(?, j) \to \text{mor}_I(?, \ell)$. The value of this 1-dimensional $I$-CW-complex at the objects $k$ and $\ell$ is a point and at the object $j$ is $D^1$. Hence it is a finite model for $EI$.

**Example 2.7.** Let $I$ be the category with objects the non-empty subsets of $[q] = \{0, 1, \ldots, q\}$ and a unique arrow $J \to K$ if and only if $K \subseteq J$. In
Infinite categories may also admit finite models. Let
\[ \mathcal{I} \]
be the opposite of the poset of non-empty subsets of \([q]\). Then \(\mathcal{I}\) admits a finite \(\mathcal{I}\)-CW-model \(X\) for the classifying \(\mathcal{I}\)-space \(E\mathcal{I}\) as follows. The functor \(X: \mathcal{I}^{\text{op}} \to \text{SPACES}\) assigns to \(L\) the space \(|\Delta[L]|\), which is the geometric realization of the simplicial set which maps \([m]\) to the set of weakly order preserving maps \([m] \to L\). The space \(|\Delta[L]|\) is homeomorphic to the standard simplex with \(\text{card}(L)\) vertices. The \(n\)-skeleton \(X_n\) of \(X\) sends each \(L\) to the \(n\)-skeleton of \(|\Delta[L]|\). The \(\mathcal{I}\)-cells of \(X\) are attached globally in the following way. The 0-skeleton is
\[ X_0 = \prod_{J \subseteq [q] \mid |J| = 1} \text{mor}_\mathcal{I}(-, J). \]
For \(n \leq q\), we construct \(X_n\) out of \(X_{n-1}\) as the pushout
\[ \prod_{J} \text{mor}_\mathcal{I}(-, J) \times |\partial \Delta[n]| \to X_{n-1} \]
\[ \downarrow \]
\[ \prod_{J} \text{mor}_\mathcal{I}(-, J) \times |\Delta[n]| \to X_n. \]
The disjoint unions are over all \(J \subseteq [q]\) with \(|J| = n + 1\). The \(J\)-component of the gluing map is induced by the \((n-1)\)-face inclusion
\[ |\Delta[K]| \to |\partial \Delta[J]| \cong |\partial \Delta[n]| \]
for all \(K \subseteq J\) with \(|K| = n\). Clearly \(X\) is a finite \(\mathcal{I}\)-CW–complex. For each object \(L\) of \(\mathcal{I}\), we have \(X(L) = |\Delta[L]| \cong *\), so that \(X\) is a finite model for \(E\mathcal{I}\).

**Example 2.8.** Infinite categories may also admit finite models. Let \(\mathcal{I} = \widehat{\mathbb{N}}\) be the monoid of natural numbers \(\mathbb{N}\) viewed as a one-object category. A finite model \(X\) for the \(\widehat{\mathbb{N}}\)-classifying space has \(X_0(*) = \text{mor}_{\widehat{\mathbb{N}}}(*, *) = \mathbb{N}\) and \(X_1(*) = [0, \infty)\). This model has a single \(\widehat{\mathbb{N}}\)-0-cell \(\text{mor}_{\widehat{\mathbb{N}}}(-, *)\) and a single \(\widehat{\mathbb{N}}\)-1-cell \(\text{mor}_{\widehat{\mathbb{N}}}(-, *) \times D^1\). The gluing map \(\mathbb{N} \times S^0 \to \mathbb{N}\) sends \((n, -1)\) and \((n, 1)\) to \(n\) and \(n + 1\) respectively. Similarly, the group of integers \(\mathbb{Z}\) viewed as a one object category admits a finite model \(Y\) with one \(\widehat{\mathbb{Z}}\)-0-cell and one \(\widehat{\mathbb{Z}}\)-1-cell, so that \(Y_0(*) = \mathbb{Z}\) and \(Y_1(*) = \mathbb{R}\).

**Remark 2.9.** Suppose a category \(\mathcal{I}\) admits a finite \(\mathcal{I}\)-CW-model for \(E\mathcal{I}\). Then the cellular \(R\)-chains of a finite model provide a finite free resolution of the constant \(R\mathcal{I}\)-module \(\underline{R}\), so \(\mathcal{I}\) is of type \((\text{FF}_R)\). If \(\mathcal{I}\) is additionally directly finite and \(R\) is Noetherian, then \(\chi(B\mathcal{I}; R) = \chi(\mathcal{I}; R) = \chi^{(2)}(\mathcal{I})\) by Theorem 1.12.

**Remark 2.10** (Bar construction of classifying \(\mathcal{I}\)-space). There exists a functorial construction \(E^{\text{bar}}\mathcal{I}\) of \(E\mathcal{I}\) by a kind of bar construction. Namely, the contravariant functor \(E^{\text{bar}}\mathcal{I}: \mathcal{I} \to \text{SPACES}\) sends an object \(i\) to the space \(B^{\text{bar}}(i \downarrow \mathcal{I})\), which is the geometric realization of the nerve of the category of objects under \(i\) (see Davis–Lück [14] page 230 and also Bousfield–Kan [10] page 327). An equivalent definition of the bar construction in terms of the tensor product in Definition 2.3 is
\[
E^{\text{bar}}\mathcal{I} = \{*\} \otimes_{\mathcal{I}} B^{\text{bar}}(? \downarrow \mathcal{I} \downarrow ??),
\]
from which we prove that \(E^\text{bar} \mathcal{I}\) is an \(\mathcal{I}\)-CW-complex. The \(\mathcal{I} \times \mathcal{I}\)-space \(B^\text{bar}(\downarrow \mathcal{I} \downarrow \mathcal{J})\) is an \(\mathcal{I} \times \mathcal{I}\)-CW-complex (see [13, page 228]). For each path 
\[i_0 \to i_1 \to i_2 \to \cdots \to i_n\]
of \(n\)-many non-identity morphisms in \(\mathcal{I}\), \(B^\text{bar}(\downarrow \mathcal{I} \downarrow \mathcal{J})\) has an \(n\)-cell based at \((i_0, i_n)\), that is a cell of the form \(\text{mor}_\mathcal{I}(?, i_0) \times \text{mor}_\mathcal{I}(i_n, ?) \times D^n\). By [13, Lemma 3.19 (2)], the tensor product \(E^\text{bar} \mathcal{I} \otimes \mathcal{J}\) is an \(\mathcal{I}\)-CW-complex: an \((m + n)\)-cell based at \(i\) is an \(n\)-cell of \(B^\text{bar}(\downarrow \mathcal{I} \downarrow \mathcal{J})\) based at \((i, j)\) and an \(m\)-cell of the \(\mathcal{E}\)-complex \(\circ(j)\) (see [13, page 229]). More explicitly, for each path of \(n\)-many non-identity morphisms
\[(2.12)\]
the \(\mathcal{I}\)-CW-complex \(E^\text{bar} \mathcal{I}\) has an \(n\)-cell based at \(i_0\).

Though the bar construction is in general not a finite \(\mathcal{I}\)-CW-complex, it is in certain cases. For example, if \(\mathcal{I}\) has only finitely many morphisms, no nontrivial isomorphisms, and no nontrivial endomorphisms, then there are only finitely many paths as in \((2.12)\), and hence only finitely many \(\mathcal{I}\)-cells in \(E^\text{bar} \mathcal{I}\).

The bar construction is also compatible with induction. Given a functor \(\alpha: \mathcal{I} \to \mathcal{D}\), we obtain a map of \(\mathcal{D}\)-spaces
\[E^\text{bar} \alpha: \alpha_* E^\text{bar} \mathcal{I} \to E^\text{bar} \mathcal{D},\]
where \(\alpha_*\) denotes induction with the functor \(\alpha\) (see [13, Definition 1.8]). If \(T: \alpha \to \beta\) is a natural transformation of functors \(\mathcal{I} \to \mathcal{D}\), we obtain for any \(\mathcal{I}\)-space \(X\) a natural transformation \(T_*: \alpha_* X \to \beta_* X\) which comes from the map of \(\mathcal{I}\)-\(\mathcal{D}\)-spaces \(\text{mor}_\mathcal{D}(?, \alpha(?) \to \text{mor}_\mathcal{D}(?, \beta(?)\) sending \(g: ? \to \alpha(?)\) to \(T(?) \circ g: ? \to \beta(?)\).

**Lemma 2.13 (Invariance of finite models under equivalence of categories).** Suppose \(\mathcal{I}\) and \(\mathcal{J}\) are equivalent categories. Then \(\mathcal{I}\) admits a finite \(\mathcal{I}\)-CW-model for \(E\mathcal{I}\) if and only if \(\mathcal{J}\) admits a finite \(\mathcal{J}\)-CW-model for \(E\mathcal{J}\). More precisely, if \(F: \mathcal{I} \to \mathcal{J}\) is an equivalence of categories and \(Y\) is a finite \(\mathcal{J}\)-CW-model for \(E\mathcal{J}\), then the restriction \(\text{res}_F Y\) is a finite \(\mathcal{I}\)-CW-model for \(E\mathcal{I}\).

**Proof.** For any functor \(F: \mathcal{I} \to \mathcal{J}\), we have an adjunction
\[\text{ind}_F: \mathcal{I}\text{-SPACES} \rightleftarrows \mathcal{J}\text{-SPACES}: \text{res}_F\]
defined by
\[\text{ind}_F(X) := X(?) \otimes \mathcal{I} \text{mor}_\mathcal{J}(?, F(?))\]
\[\text{res}_F(Y) := Y \circ F(?).\]
The \(\mathcal{I}\)-space \(\text{res}_F(Y)\) is naturally homeomorphic to \(Y(?) \otimes \mathcal{J} \text{mor}_\mathcal{J}(F(?), ?)\). But since we are assuming \(F\) is an equivalence of categories, it is a left adjoint in an adjoint equivalence \((F, G)\), and we have natural homeomorphisms of \(\mathcal{I}\)-spaces
\[\text{res}_F(Y) \cong Y(?) \otimes \mathcal{J} \text{mor}_\mathcal{J}(F(?), ?)\]
\[\cong Y(?) \otimes \mathcal{J} \text{mor}_\mathcal{J}(?, G(?))\]
\[\cong \text{ind}_G(Y).\]
Since \( \text{ind}_G \) is a left adjoint, so is \( \text{res}_F \), and \( \text{res}_F \) therefore preserves pushouts. Note also

\[
\text{res}_F \circ \text{mor}_J(?, j) = \text{mor}_J(F(?, j)) \cong \text{mor}_I(?, G(j)).
\]

If \( Y \) is a finite \( J \)-CW-model for \( E_J \) with \( n \)-skeleton

\[
\prod_{\lambda \in \Lambda_n} \text{mor}_J(-, j_\lambda) \times S^{n-1} \longrightarrow Y_{n-1} \quad \text{and} \quad \prod_{\lambda \in \Lambda_n} \text{mor}_J(-, d_\lambda) \times D^n \longrightarrow Y_n,
\]

then \( X := \text{res}_F Y \) is a finite \( I \)-CW-complex with \( n \)-skeleton

\[
\prod_{\lambda \in \Lambda_n} \text{mor}_I(-, G(j_\lambda)) \times S^{n-1} \longrightarrow X_{n-1} \quad \text{and} \quad \prod_{\lambda \in \Lambda_n} \text{mor}_I(-, G(d_\lambda)) \times D^n \longrightarrow X_n.
\]

Clearly, \( \text{res}_F Y \) is contractible at each object \( i \), since \( \text{res}_F Y(i) = Y(F(i)) \simeq \ast \). ☐

### 3. Homotopy Colimits of Categories

**Definition 3.1 (Homotopy colimit for categories).** Let \( \mathcal{C} : \mathcal{I} \to \text{CAT} \) be a covariant functor from some (small) index category \( \mathcal{I} \) to the category of small categories. Its **homotopy colimit**

\[
\text{hocolim}_\mathcal{I} \mathcal{C}
\]

is the following category. Objects are pairs \((i, c)\), where \( i \in \text{ob}(\mathcal{I}) \) and \( c \in \text{ob}(\mathcal{C}(i)) \). A morphism from \((i, c)\) to \((j, d)\) is a pair \((u, f)\), where \( u : i \to j \) is a morphism in \( \mathcal{I} \) and \( f : \mathcal{C}(u)(c) \to d \) is a morphism in \( \mathcal{C}(j) \). The composition of the morphisms \((u, f) : (i, c) \to (j, d)\) and \((v, g) : (j, d) \to (k, e)\) is the morphism

\[
(v, g) \circ (u, f) = (v \circ u, g \circ \mathcal{C}(v)(f)) : (i, c) \to (k, e).
\]

The identity of \((i, c)\) is given by \((\text{id}_i, \text{id}_c)\).

This homotopy colimit construction for functors is often called the **Grothendieck construction** or the category of **elements**.

In which sense is \( \text{hocolim}_\mathcal{I} \mathcal{C} \) a homotopy colimit? First, recall from [20] that the nerve functor induces an equivalence of categories \( \text{Ho CAT} \to \text{Ho SSET} \), where \( \text{Ho CAT} \) denotes the localization of \( \text{CAT} \) with respect to nerve weak equivalences and \( \text{Ho SSET} \) denotes the localization of \( \text{SSET} \) with respect to the usual weak equivalences. In [40], Thomason proved that \( \text{hocolim}_\mathcal{I} \mathcal{C} \) in \( \text{CAT} \) corresponds to the Bousfield–Kan construction in \( \text{SSET} \) under this equivalence of categories. Consequently, \( \text{hocolim}_\mathcal{I} \mathcal{C} \) has a universal property in the form of a bijection

\[
\text{Ho CAT}(\text{hocolim}_\mathcal{I} \mathcal{C}, \Gamma) \cong \text{Ho CAT}^\mathcal{I}(\mathcal{C}, \Gamma).
\]
for any category \( \Gamma \). Here \( \Gamma \) indicates the \( I \)-diagram that is constant \( \Gamma \). In \cite{Thomason}, Thomason proved that \( \text{CAT} \) admits a cofibrantly generated model structure in which the weak equivalences are the nerve weak equivalences, so that the associated projective model structure on \( \text{CAT}^I \) exists. The model-theoretic construction of a homotopy colimit of the \( I \)-diagram \( \mathcal{C} \) in \( \text{CAT} \) as a colimit of a cofibrant replacement of \( \mathcal{C} \) in the projective model structure therefore works. This model-theoretic construction also has the universal property in \( (3.2) \), so is isomorphic to \( \text{hocolim}^I \mathcal{C} \) in \( \text{Ho CAT} \), i.e. weakly equivalent to \( \text{hocolim}^I \mathcal{C} \) in \( \text{CAT} \). A direct proof that \( \text{hocolim}^I \mathcal{C} \) satisfies the universal property \( (3.2) \) is in Grothendieck’s letter \cite{Grothendieck}, see the article of Maltsiniotis \cite{Maltsiniotis}, Section 3.1.

**Remark 3.3.** If \( \mathcal{C} \) is merely a pseudo functor, then it of course still has a homotopy colimit. A pseudo functor \( \mathcal{C} : \mathcal{I} \to \text{CAT} \) is like an ordinary functor, but only preserves composition and unit up to specified coherent natural isomorphisms \( \mathcal{C}_{v,u} : \mathcal{C}(v) \circ \mathcal{C}(u) \Rightarrow \mathcal{C}(v \circ u) \) and \( \mathcal{C}_i : 1_{\mathcal{C}(i)} \Rightarrow \mathcal{C}(\text{id}_i) \). Moreover, \( \mathcal{C}_{v,u} \) is required to be natural in \( v \) and \( u \). The objects and morphisms of the homotopy colimit \( \text{hocolim}^I \mathcal{C} \) are defined as in the strict case of Definition 3.1. The composition in \( \text{hocolim}^I \mathcal{C} \) is defined by the modified rule

\[
(v, g) \circ (u, f) = (v \circ u, g \circ (\mathcal{C}(v)(f)) \circ \mathcal{C}_{v,u}^{-1}(c))
\]

while the identity of the object \((i, c)\) is given by

\[
(\text{id}_i, \mathcal{C}_i^{-1}(c)).
\]

The homotopy colimit of a pseudo functor \( \mathcal{C} : \mathcal{I} \to \text{CAT} \) is an ordinary 1-category with strictly associative and strictly unital composition.

**Remark 3.4.** For a fixed category \( \mathcal{I} \), the homotopy colimit construction \( \text{hocolim}^I \mathcal{C} \) is a strict 2-functor from the strict 2-category of pseudo functors \( \mathcal{I} \to \text{CAT} \), pseudo natural transformations, and modifications into the strict 2-category \( \text{CAT} \).

**Example 3.5 (Homotopy colimit of a constant functor).** If \( \mathcal{C} : \mathcal{I} \to \text{CAT} \) is a constant functor, say constantly a category also called \( \mathcal{C} \), then \( \text{hocolim}^I \mathcal{C} = \mathcal{I} \times \mathcal{C} \).

**Example 3.6 (Homotopy colimit for \( \mathcal{I} \) with a terminal object).** Suppose \( \mathcal{I} \) has a terminal object \( t \) and \( \mathcal{C} : \mathcal{I} \to \text{CAT} \) is a strict covariant functor. Then \( \text{hocolim}^I \mathcal{C} \) is homotopy equivalent to \( \mathcal{C}(t) \) as follows. This is analogous to the familiar fact that \( \mathcal{C}(t) \) is a colimit of \( \mathcal{C} \). The components of the universal cocone

\[
\pi : \mathcal{C} \Rightarrow \Delta_{\mathcal{C}(t)}
\]

\(^1\)We thank George Maltsiniotis for clarifying these points about homotopy colimits in \( \text{CAT} \).
are $C(i \to t)$. Applying $\text{hocolim}_I = \text{hocolim}_I C$ and composing with the projection gives us a functor $F$

$$\text{hocolim}_I C \xrightarrow{\pi} \mathcal{I} \times C(t) \xrightarrow{\text{pr}_{C(t)}} C(t)$$

$(i, c) \xrightarrow{F} C(i \to t)(c)$.

The functor $G: C(t) \to \text{hocolim}_I C$, $G(c) = (t, c)$ is a homotopy inverse, since $F \circ G = \text{id}_{C(t)}$ and we have a natural transformation $\text{id}_{\text{hocolim}_I C} \Rightarrow G \circ F$ with components

$$(i \to t, \text{id}_{C(i \to t)}): (i, c) \xrightarrow{F} (t, C(i \to t)c).$$

Let $\mathcal{H}$ denote the homotopy colimit of the $\mathcal{I}$-diagram of categories $C$. We now construct an $\mathcal{I}$-diagram of $\mathcal{H}$-spaces $E^\mathcal{H}$ with the property that its tensor product with $E\mathcal{I}$ is $\mathcal{H}$-homotopy equivalent to a classifying $\mathcal{H}$-space for $\mathcal{H}$. This $\mathcal{I}$-diagram of $\mathcal{H}$-spaces $E^\mathcal{H}$ will play an important role in the inductive proof of the Homotopy Colimit Formula Theorem [4, 11].

**Construction 3.8 (Construction of $E^\mathcal{H}$).** Let $C: \mathcal{I} \to \text{CAT}$ be a strict covariant functor, and abbreviate $\mathcal{H} = \text{hocolim}_I C$. Define a functor

$$E^\mathcal{H}: \mathcal{I} \to \mathcal{H}\text{-SPACES}$$

as follows. Given an object $i \in \mathcal{I}$, we have the functor

$$(\alpha)\circ (i): \mathcal{C}(i) \to \mathcal{H}$$

sending an object $c$ to the object $(i, c)$ and a morphism $f: c \to d$ to the morphism $(\text{id}_i, f)$. We define

$$E^\mathcal{H}(i) = \alpha(i)_* E^\text{bar}(\mathcal{C}(i)).$$

Consider a morphism $u: i \to j$ in $\mathcal{I}$. It induces a natural transformation $T(u): \alpha(i) \to \alpha(j) \circ \mathcal{C}(u)$ from the functor $\alpha(i): \mathcal{C}(i) \to \mathcal{H}$ to the functor $\alpha(j) \circ \mathcal{C}(u): \mathcal{C}(i) \to \mathcal{H}$ by assigning to an object $c$ in $\mathcal{C}(i)$ the morphism

$$(u, \text{id}_{\mathcal{C}(u)(c)}): \alpha(i)(c) = (i, c) \to \alpha(j) \circ \mathcal{C}(u)(c) = (j, \mathcal{C}(u)(c)).$$

From Remark 2.10 we obtain a map of $\mathcal{H}$-spaces

$$T(u)_*: \alpha(i)_* E^\text{bar}(\mathcal{C}(i)) \to \alpha(j)_* \mathcal{C}(u)_* E^\text{bar}(\mathcal{C}(i))$$

and a map of $\mathcal{C}(j)$-spaces

$$E^\text{bar}(\mathcal{C}(u))): \mathcal{C}(u)_* E^\text{bar}(\mathcal{C}(i)) \to E^\text{bar}(\mathcal{C}(j)).$$

Finally, for the morphism $u$ in $\mathcal{I}$, we define $E^\mathcal{H}(u): E^\mathcal{H}(i) \to E^\mathcal{H}(j)$ by the composite of the two maps below.

$$\begin{align*}
\alpha(i)_* E^\text{bar}(\mathcal{C}(i)) & \xrightarrow{T(u)_*} \alpha(j)_* \mathcal{C}(u)_* E^\text{bar}(\mathcal{C}(i)) \\
\alpha(j)_* \mathcal{C}(u)_* E^\text{bar}(\mathcal{C}(i)) & \xrightarrow{\alpha(j)_* (E^\text{bar}(\mathcal{C}(u)))} \alpha(j)_* E^\text{bar}(\mathcal{C}(j))
\end{align*}$$

**Documenta Mathematica 16 (2011) 301–354**
Define the homotopy colimit of the covariant functor $E^H$ of (3.9) to be the contravariant $H$-space

$$
hocolim_I E^H := (i, c) \mapsto E I \otimes_I (E^H(i, c)).$$

(3.11)

**Lemma 3.12.** Consider any model $E I$ for the classifying $I$-space of the category $I$. Then the contravariant $H$-space $E I \otimes_I E^H$ of (3.11) is $H$-homotopy equivalent to the classifying $H$-space $E H$ of the category $H := hocolim_I C$.

**Proof.** We first show that for any object $(i, c)$ in $H$ the space $E I \otimes_I (E^H(i, c))$ is contractible. The covariant functor $E^H(i, c): I \to \text{SPACES}$ sends an object $j$ to

$$\alpha(j) \ast (E^H^\text{bar}(j))(i, c)$$

$$= \alpha(j) \ast (E^H^\text{bar}(j))(? \otimes_H \text{mor}_H((i, c), ?))$$

$$= (E^H^\text{bar}(j))(? \otimes_{C(j)} \text{mor}_{C(j)}((i, c), (j, ?)))$$

$$= \prod_{u \in \text{mor}_I(i, j)} (E^H^\text{bar}(j))(?(u) \otimes_{C(u)} \text{mor}_{C(u)}(C(u(c)), ?))$$

$$= \prod_{u \in \text{mor}_I(i, j)} (E^H^\text{bar}(j))(C(u(c))) .$$

Since $(E^H^\text{bar}(j))(C(u(c)))$ is contractible, the projection

$$\prod_{u \in \text{mor}_I(i, j)} (E^H^\text{bar}(j))(C(u(c))) \to \text{mor}_I(i, j)$$

is a homotopy equivalence. Hence the collection of these projections for $j \in \text{ob}(I)$ induces a map of $I$-spaces

$$\text{pr}: E^H(i, c) \to \text{mor}_I(i, ?)$$

whose evaluation at each object $j$ in $\text{ob}(I)$ is a homotopy equivalence. We conclude from Davis–Lück [14, Theorem 3.11] that $E I \otimes_I \text{pr}: E I \otimes_I E^H(i, c) \simeq E I \otimes_I \text{mor}_I(i, ?).$ is a homotopy equivalence. Since $E I \otimes_I \text{mor}_I(i, ?) = E I(i)$ is contractible, this implies that for any object $(i, c)$ in $H$ the space $E I \otimes_I (E^H(i, c))$ is contractible, as we initially claimed.

It remains to show that $E I \otimes_I E^H$ has the $H$-homotopy type of an $H$-$CW$-complex. It is actually an $H$-$CW$-complex. The following argument, that $E I \otimes_I E^H$ has the homotopy type of an $H$-$CW$-complex, will be used again later.

---

2This is a well-known standard argument, which we present only so that the reader easily sees that it works in the setting of $H$-spaces.
We have a filtration of $EI$
\[ \emptyset = EI_{-1} \subseteq EI_0 \subseteq EI_1 \subseteq \cdots \subseteq EI_n \subseteq \cdots \subseteq EI = \bigcup_{n \geq 0} EI_n \]
such that
\[ EI = \text{colim}_{n \to \infty} EI_n \]
and for every $n \geq 0$ there exists a pushout of $I$-spaces
\[
\begin{array}{c}
\coprod_{\lambda \in \Lambda_n} \text{mor}_I (-, i_{\lambda}) \times S^{n-1} \xrightarrow{f_{n-1}} EI_{n-1} \\
\downarrow \\
\coprod_{\lambda \in \Lambda_n} \text{mor}_I (-, i_{\lambda}) \times D^n \xrightarrow{g_{n-1}} EI_n
\end{array}
\]
(3.13)
Since $- \otimes_I Z$ has a right adjoint [14, Lemma 1.9] we get
\[ EI \otimes_I E^H = \text{colim}_{n \to \infty} EI_n \otimes_I E^H \]
as a colimit of $H$-spaces. After an application of $- \otimes_I E^H$ to (3.13), we obtain pushouts of $H$-spaces
\[
\begin{array}{c}
\coprod_{\lambda \in \Lambda_n} E^H(i_{\lambda}) \times S^{n-1} \xrightarrow{f_{n-1}} EI_{n-1} \otimes_I E^H \\
\downarrow \\
\coprod_{\lambda \in \Lambda_n} E^H(i_{\lambda}) \times D^n \xrightarrow{g_{n-1}} EI_n \otimes_I E^H
\end{array}
\]
(3.14)
where the left vertical arrow and hence the right vertical arrow are cofibrations of $H$-spaces. By induction we may assume that $EI_{n-1} \otimes_I E^H$ has the homotopy type of an $H$-CW-complex. Since the vertical maps are cofibrations, by replacing it with a homotopy equivalent $H$-CW-complex we do not change the homotopy type of the pushout (the usual proof for spaces goes through for $H$-spaces). Hence we may assume that $EI_{n-1} \otimes_I E^H$ is a $H$-CW-complex. We may also assume that $f_{n-1}$ is cellular: since the vertical maps are cofibrations, by replacing $f_{n-1}$ by a homotopic cellular map, which exists by Davis–Lück [14, cf. Theorem 3.7], we also do not change the homotopy type of the pushout. See Selick [33, Theorem 7.1.8] for a proof of this statement for spaces which translates verbatim to the setting of $H$-spaces. If $f_{n-1}$ is cellular, diagram (3.14) is a cellular pushout. Hence we completed the induction step, showing that $EI_n \otimes_I E^H$ has the homotopy type of an $H$-CW-complex.

It remains to show that $EI \otimes_I E^H$ has the homotopy type of a $H$-CW-complex: choose $H$-CW-complexes $Z_n$ and $H$-homotopy equivalences $g_n : Z_n \to EI_n \otimes_I E^H$. By iteratively replacing $Z_n$ by the mapping cylinder of
\[ Z_{n-1} \xrightarrow{g_{n-1}} EI_{n-1} \otimes_I E^H \to EI_n \otimes_I E^H \xrightarrow{g_n} Z_n, \]
where $\bar{g}_n$ is a homotopy inverse of $g_n$, one finds a new sequence of homotopy equivalences $g'_n : Z_n \to EI_n \otimes_I E^H$ (with the modified $H$-CW-complexes $Z_n$) such that $g'_n|_{Z_{n-1}} = g'_{n-1}$. $\square$
4. Homotopy Colimit Formula for Finiteness Obstructions and Euler Characteristics

In this section we prove the main theorem of this paper: the Homotopy Colimit Formula. It expresses the finiteness obstruction, the Euler characteristic, and the $L^2$-Euler characteristic of the homotopy colimit of a diagram in CAT in terms of the respective invariants for the diagram entries at the base objects for cells in a finite model for the $I$-classifying space of $I$. Analogous formulas for the functorial counterparts of the Euler characteristic and $L^2$-Euler characteristic are included. The Homotopy Colimit Formula is initially stated and proved for strict functors $C : I \to \text{CAT}$, but we prove that it also holds for pseudo functors $D : I \to \text{CAT}$ in Corollary 4.2. The full generality of pseudo functors is needed for the applications to complexes of groups in Section 5.

4.1. Homotopy Colimit Formula.

Theorem 4.1 (Homotopy Colimit Formula). Let $I$ be a small category such that there exists a finite $I$-CW-model for its classifying $I$-space. Fix such a finite $I$-CW-model $EI$. Denote by $\Lambda_n$ the finite set of $n$-cells $\lambda = \text{mor}(?, i_\lambda) \times D^n$ of $EI$. Let $C : I \to \text{CAT}$ be a covariant functor. Abbreviate $H = \text{hocolim}_C$. Then:

(i) If $I$ is directly finite, and $C(i)$ is directly finite for every object $i \in \text{ob}(I)$, then the category $H$ is directly finite;

(ii) If $I$ is an EI-category, $C(i)$ is an EI-category for every object $i \in \text{ob}(I)$, and for every automorphism $u : i \overset{=}{\to} i$ the map $\text{iso}(C(i)) \to \text{iso}(C(i))$, $x \mapsto C(u)(x)$ is the identity, then the category $H$ is an EI-category;

(iii) If for every object $i$ the category $C(i)$ is of type $(FP_R)$, then the category $\text{hocolim}_C$ is of type $(FP_R)$;

(iv) If for every object $i$ the category $C(i)$ is of type $(FF_R)$, then the category $\text{hocolim}_C$ is of type $(FF_R)$;

(v) If for every object $i$ the category $C(i)$ is of type $(FP_R)$, then we obtain for the finiteness obstruction

$$o(H; R) = \sum_{n \geq 0} (-1)^n \cdot \sum_{\lambda \in \Lambda_n} \alpha(i_\lambda)_* (o(C(i_\lambda); R)),$$

where $\alpha(i_\lambda)_* : K_0(\text{RC}(i_\lambda)) \to K_0(\text{RH})$ is the homomorphism induced by the canonical functor $\alpha(i_\lambda)_* : C(i_\lambda) \to H$ defined in (3.10);

(vi) Suppose that $I$ is directly finite and $C(i)$ is directly finite for every object $i \in \text{ob}(I)$. If for every object $i$ the category $C(i)$ is additionally of type $(FP_R)$ then we obtain for the functorial Euler characteristic

$$\chi_f(H; R) = \sum_{n \geq 0} (-1)^n \cdot \sum_{\lambda \in \Lambda_n} \alpha(i_\lambda)_* (\chi_f(C(i_\lambda); R)),$$

where $\alpha(i_\lambda)_* : U(C(i_\lambda)) \to U(H)$ is the homomorphism induced by the canonical functor $\alpha(i_\lambda)_* : C(i_\lambda) \to H$ defined in (3.10). Summing up,
we also have
\[ \chi(H; R) = \sum_{n \geq 0} (-1)^n \cdot \sum_{\lambda \in \Lambda_n} \chi(C(i_\lambda); R). \]

If \( R \) is Noetherian, in addition to the direct finiteness and \((FP_R)\) hypotheses, we obtain for the Euler characteristics of the classifying spaces
\[ \chi(BH; R) = \sum_{n \geq 0} (-1)^n \cdot \sum_{\lambda \in \Lambda_n} \chi(BC(i_\lambda); R); \]

(iii) Suppose that \( I \) is directly finite and \( C(i) \) is directly finite for every object \( i \in \text{ob}(I) \). If for every object \( i \) the category \( C(i) \) is additionally of type \((L^2)\), then \( H \) is of type \((L^2)\) and we obtain for the functorial \( L^2\)-Euler characteristic
\[ \chi^{(2)}_f(H) = \sum_{n \geq 0} (-1)^n \cdot \sum_{\lambda \in \Lambda_n} \alpha(i_\lambda)_* \chi^{(2)}(C(i_\lambda)), \]

where \( \alpha(i_\lambda)_*: U^{(1)}(C(i_\lambda)) \to U^{(1)}(H) \) is the homomorphism induced by the canonical functor \( \alpha(i_\lambda): C(i_\lambda) \to H \) defined in \([3,10]\), and we obtain for the \( L^2\)-Euler characteristic
\[ \chi^{(2)}(H) = \sum_{n \geq 0} (-1)^n \cdot \sum_{\lambda \in \Lambda_n} \chi^{(2)}(C(i_\lambda)). \]

**Proof.** (ii) Consider morphisms \((u, f): (i, c) \to (j, d)\) and \((v, g): (j, d) \to (i, c)\) in \( H \) with \((v, g) \circ (u, f) = \text{id}_{(i, c)}\). This implies \(vu = id_i\) and \(g \circ C(v)(f) = id_c\). Since \( I \) and \( C(i) \) are by assumption directly finite, we conclude \(uv = id_j\) and \(C(v)(f) \circ g = id_{C(v)(d)}\). Hence
\[ (u, f) \circ (v, g) = (uv, f \circ C(u)(g)) = (uv, C(u)(f) \circ g) = (uv, C(u)(\text{id}_{C(v)(d)})) = (\text{id}_j, \text{id}_d). \]

(iii) Consider an endomorphism \((u, f): (i, c) \to (i, c)\) in \( H \). Since \( I \) is an EI-category, \( u: i \to i \) is an automorphism. Since \( C(u)(c) = T \) by assumption, we can choose an isomorphism \( g: c \xrightarrow{\cong} C(u)(c) \). Hence \( fg \) is an endomorphism in \( C(i) \). Since \( C(i) \) is an EI-category, and \( g \) is an isomorphism, \( f \) is also an isomorphism. Since \( u \) and \( f \) are isomorphisms, \((u, f)\) is an isomorphism.
for any choice $P_\ast$ of finite projective approximation. This is independent of the choice of $P_\ast$ and the basic properties of it were studied by Lück [23, Chapter 11]. If $0[\hat{R}]$ is the $R\mathcal{H}$-chain complex concentrated in dimension zero and given there by the constant $R\mathcal{H}$-module $\hat{R}$, then $\mathcal{H}$ is of type $(FP_\ast)$ if and only if $0[\hat{R}]$ is of type $(FP_\ast)$ and in this case
\[ o(\mathcal{H}; R) = o(0[\hat{R}]) \in K_0(R\mathcal{H}). \]

Consider a finite $\mathcal{I}$-$CW$-complex $X$. We want to show by induction over the dimension of $X$ that the $R\mathcal{H}$-chain complex $C_\ast(X \otimes \mathcal{I} E^{\mathcal{H}})$ is of type $(FP_\ast)$ and satisfies
\[ o(C_\ast(X \otimes \mathcal{I} E^{\mathcal{H}})) = \sum_{n \geq 0} (-1)^n \cdot \sum_{\lambda \in A_n} \alpha(i_\lambda)_\ast (o(C(i_\lambda); R)), \]
where $A_n$ denotes the set of $n$-cells of $X$ and $i_\lambda$ is the object at which the $n$-cell $\lambda = \text{mor}(\mathcal{I}, i_\lambda) \times D^n$ of $X$ is based.

The induction beginning, where $X$ is the empty set, is obviously true. The induction step is done as follows. Let $d$ be the dimension of $X$. Then $X_d$ is obtained from $X_{d-1}$ by a pushout of $\mathcal{I}$-spaces
\[ \prod_{\lambda \in A_d} \text{mor}_C(-, i_\lambda) \times S^{d-1} \longrightarrow X_{d-1} \]
\[ \prod_{\lambda \in A_d} \text{mor}_C(-, i_\lambda) \times D^d \longrightarrow X = X_d. \]

Applying $- \otimes \mathcal{I} E^{\mathcal{H}}$ to it yields, because $E^{\mathcal{H}}(i) = \alpha(i)_\ast E^{\text{bar}}(C(i))$, a pushout of $\mathcal{H}$-spaces with a cofibration as left vertical arrow
\[ \prod_{\lambda \in A_d} \alpha(i_\lambda)_\ast E^{\text{bar}}(C(i_\lambda)) \times S^{d-1} \longrightarrow X_{d-1} \otimes \mathcal{I} E^{\mathcal{H}} \]
\[ \prod_{\lambda \in A_d} \alpha(i_\lambda)_\ast E^{\text{bar}}(C(i_\lambda)) \times D^d \longrightarrow X \otimes \mathcal{I} E^{\mathcal{H}}. \]

In the sequel we can assume without loss of generality that $X_{d-1} \otimes \mathcal{I} E^{\mathcal{H}}$ and $X \otimes \mathcal{I} E^{\mathcal{H}}$ are $\mathcal{H}$-$CW$-complexes and the diagram above is a pushout of $\mathcal{H}$-$CW$-complexes, since this can be arranged by replacing them by homotopy equivalent $\mathcal{H}$-$CW$-complexes (see the proof of Lemma 3.12). We obtain an exact sequence of $R\mathcal{H}$-chain complexes
\[ 0 \rightarrow C_\ast(X_{d-1} \otimes \mathcal{I} E^{\mathcal{H}}) \rightarrow C_\ast(X \otimes \mathcal{I} E^{\mathcal{H}}) \rightarrow \bigoplus_{\lambda \in A_d} \Sigma^d C_\ast(\alpha(i_\lambda)_\ast E^{\text{bar}} C(i_\lambda)) \rightarrow 0. \]

Consider $\lambda \in A_d$. Since $C(i_\lambda)$ is of type $(FP_\ast)$, we can find a finite projective $R\mathcal{C}(i_\lambda)$-chain complex $P_\ast$ whose homology is concentrated in dimension zero and given there by the constant $R\mathcal{C}(i_\lambda)$-module $\hat{R}$. Since $\bar{C}_\ast(E^{\text{bar}} C(i_\lambda))$ is a projective $R\mathcal{C}(i_\lambda)$-chain complex with the same homology, there is an $R\mathcal{C}(i_\lambda)$-chain
homotopy equivalence \( f_*: P_* \xrightarrow{\sim} C_* (E^{\bar{\text{bar}}} C(i)) \) (see Lück [23, Lemma 11.3 on page 213]) and

\[
o(C(i_\lambda); R) = o(P_*) = \sum_{n \geq 0} (-1)^n \cdot [P_n] \in K_0 (RC(i_\lambda)).\]

Obviously

\[
o(i_\lambda)_* f_*: \alpha(i_\lambda)_* P_* \xrightarrow{\sim} \alpha(i_\lambda)_* C_* (E^{\bar{\text{bar}}} (C(i_\lambda))) = C_* (\alpha(i_\lambda)_* E^{\bar{\text{bar}}} C(i_\lambda))\]

is an \( RH \)-chain homotopy equivalence. Hence \( C_* (\alpha(i_\lambda)_* E^{\bar{\text{bar}}} C(i_\lambda)) \) and, by the induction hypothesis, \( C_* (X_{d-1} \otimes I E^H) \) are \( RH \)-chain complexes of type \( (\text{FP}_R) \).

We conclude from Lück [23, Lemma 11.3 on page 213] that \( C_* (X \otimes I E^H) \) is of type \( (\text{FP}_R) \) and

\[
o(C_* (X \otimes I E^H)) = o(C_* (X_{d-1} \otimes I E^H)) + \sum_{\lambda \in \Lambda_d} o(\Sigma^d o(i_\lambda)_* C_* (E^{\bar{\text{bar}}} C(i_\lambda))).\]

This implies together with the induction hypothesis applied to \( X_{d-1} \)

\[
o(C_* (X \otimes I E^H)) = \sum_{n=0}^{d-1} (-1)^n \cdot \sum_{\lambda \in \Lambda_n} o(i_\lambda)_* (o(C(i_\lambda); R)) + \sum_{\lambda \in \Lambda_d} (-1)^d \cdot \alpha(i_\lambda)_* (o(C(i_\lambda); R))\]

\[
= \sum_{n=0}^{d} (-1)^n \cdot \sum_{\lambda \in \Lambda_n} o(i_\lambda)_* (o(C(i_\lambda); R)).\]

This finishes the induction step.

Assertions (iii) and (v) follow by taking \( X = ET \).

(iv) This proof is analogous to that of assertion (iii)

(vi) By (i) and (iii), the category \( H \) is directly finite and of type \( (\text{FP}_R) \). Then an application of \( rk_{R,H} \) to the formula for \( o(H; R) \) in (vi) yields the formula for \( \chi_f (H; R) \) by the naturality of \( rk_{R,-} \) with respect to the functors \( o(i_\lambda) \) between directly finite categories, see Fiore–Lück–Sauer [16, Lemma 4.9].

An application of the augmentation homomorphism \( \epsilon: U(H) \to \mathbb{Z} \) to the formula for \( \chi_f (H; R) \) yields the formula for \( \chi(H; R) \). We also use the naturality of the augmentation homomorphism, that is, the commutativity of diagram (4.5) in [10] for \( F = o(i_\lambda) \).

If \( R \) is additionally Noetherian, then Theorem 4.10 applies, and the Euler characteristics of the categories agree with the Euler characteristics of the classifying spaces.

(vii) The proofs for the functorial \( L^2 \)-Euler characteristic and the \( L^2 \)-Euler characteristic are somewhat more complicated since the property \( (L^2) \) is more general than \( (\text{FP}_R) \), and the \( L^2 \)-Euler characteristic comes from the finiteness obstruction only in the case \( (\text{FP}_R) \). The proofs are variations of the proofs for assertions (iii) and (v) instead of using Lück [23, Lemma 11.3 on page 213], we now use the basic properties of \( L^2 \)-Euler characteristics for chain complexes of modules over group von Neumann algebras [16, Lemma 5.7]. For example,
we use \[16\] Lemma 5.7 (iv), which says for any injective group homomorphism \(i: H \to G\) and \(N(H)\)-chain complex \(C_\ast\), we have \(\chi^{(2)}(C_\ast) = \chi^{(2)}(\text{ind}_i C_\ast)\), provided the sum of the \(L^2\)-Betti numbers of \(C_\ast\) is finite. The injectivity hypothesis is easily verified: for every object \(i \in \text{ob}(I)\) and object \(x \in C(i)\) the functor \(\alpha(i): C(i) \to \mathcal{H}\) clearly induces an injection \(\text{aut}_{C(i)}(x) \to \text{aut}_H(i, x)\). This finishes the proof of Theorem 4.1. \(\square\)

**Corollary 4.2.** Theorem 4.1 on homotopy colimits holds for pseudo functors \(D: \mathcal{I} \to \text{CAT}\).

**Proof.** We first remark that the pseudo functor \(D: \mathcal{I} \to \text{CAT}\) is equivalent to a strict functor \(\mathcal{C}: \mathcal{I} \to \text{CAT}\) in the following sense. As usual, we denote by \(\text{Hom}(\mathcal{I}, \text{CAT})\) the strict 2-category of pseudo functors \(\mathcal{I} \to \text{CAT}\), pseudo natural transformations between them, and modifications. The pseudo functor \(D\) is equivalent to a strict functor \(\mathcal{C}\) as objects of the 2-category \(\text{Hom}(\mathcal{I}, \text{CAT})\).

For example, we may take \(\mathcal{C}\) to be the strict functor

\[ i \mapsto \text{mor}_{\text{Hom}(\mathcal{I}, \text{CAT})}(\mathcal{I}(i, -), D). \]

The equivalence between \(\mathcal{C}\) and \(D\) in \(\text{Hom}(\mathcal{I}, \text{CAT})\) has two useful consequences. Since

\[ \text{hocolim}_\mathcal{I}: \text{Hom}(\mathcal{I}, \text{CAT}) \to \text{CAT} \]

is a 2-functor, it sends any equivalence between \(\mathcal{C}\) and \(D\) to an equivalence in \(\text{CAT}\) between the categories \(\text{hocolim}_\mathcal{I} \mathcal{C}\) and \(\text{hocolim}_\mathcal{I} \mathcal{D}\). Another consequence of the equivalence between \(\mathcal{C}\) and \(D\) is that for every \(i \in \mathcal{I}\), the categories \(\mathcal{C}(i)\) and \(\mathcal{D}(i)\) are equivalent. With these observations we reduce Corollary 4.1 to Theorem 4.1.

**(i)** Suppose \(\mathcal{D}(i)\) is directly finite for every \(i \in \text{ob}(\mathcal{I})\) and \(\mathcal{I}\) is directly finite. Since direct finiteness is preserved under equivalence of categories by Fiore–Lück–Sauer \[16\] Lemma 3.2, and \(\mathcal{C}(i)\) is equivalent to \(\mathcal{D}(i)\), we see that \(\mathcal{C}(i)\) is directly finite for every \(i \in \text{ob}(\mathcal{I})\). Hence \(\text{hocolim}_\mathcal{I} \mathcal{C}\) is directly finite by Theorem 4.1(i). Since \(\text{hocolim}_\mathcal{I} \mathcal{D}\) is equivalent to \(\text{hocolim}_\mathcal{I} \mathcal{C}\), it is also directly finite, again by \[16\] Lemma 3.2.

**(ii)** Suppose that \(\mathcal{I}\) is an EI-category, \(\mathcal{D}(i)\) is an EI-category for every \(i \in \text{ob}(\mathcal{I})\), and for every automorphism \(u: i \xrightarrow{\cong} i\) the map \(\text{iso}(\mathcal{D}(i)) \to \text{iso}(\mathcal{D}(i))\), \(\overline{y} \mapsto \overline{D(u)(y)}\) is the identity. Since EI is preserved under equivalence of categories \[16\] Lemma 3.11, and \(\mathcal{C}(i)\) is equivalent to \(\mathcal{D}(i)\), we see \(\mathcal{D}(i)\) is an EI-category. We claim that for each automorphism \(u\), the functor \(\mathcal{C}(u)\) also induces the identity on isomorphism classes of objects of \(\mathcal{C}(i)\). Let \(\phi: \mathcal{D} \to \mathcal{C}\) be a pseudo equivalence, that is, an equivalence in the 2-category \(\text{Hom}(\mathcal{I}, \text{CAT})\). For \(x \in \mathcal{C}(i)\), there is a \(y \in \mathcal{D}(i)\) and an isomorphism \(x \cong \phi_i(y)\). We have isomorphisms

\[ \mathcal{C}(u)(x) \cong \mathcal{C}(u)\phi_i(y) \cong \phi_i \mathcal{D}(u)(y) \cong \phi_i(y) \cong x, \]

and \(\mathcal{C}(u)\) induces the identity on isomorphism classes of objects of \(\mathcal{C}(i)\). Then \(\text{hocolim}_\mathcal{I} \mathcal{C}\) is an EI-category by Theorem 4.1(ii) and so is \(\text{hocolim}_\mathcal{I} \mathcal{D}\), again by \[16\] Lemma 3.11.
(iii) and (iv) similarly follow from Theorem 4.1(iii) and (iv) since property (FPₚ), property (FFₚ), and the finiteness obstruction are all invariant under equivalence of categories [16, Theorem 2.8].

(v) Suppose D(i) is of type (FPₚ) for every i ∈ ob(I). Then every C(i) is also of type (FPₚ), since property (FPₚ) is invariant under equivalence of categories [16, Theorem 2.8]. As in (3.10), we have for each i ∈ I the functor

\[ \alpha^{D}(i) : D(i) \to \text{hocolim}_{I} D \]

which sends an object d to the object \((i, d)\) and a morphism \(g : d \to d'\) to the morphism \((\text{id}_i, g \circ D^{-1}(d))\). From a pseudo equivalence \(\psi : C \to D\) we obtain a strictly commutative diagram

\[
\begin{array}{ccc}
C(i) & \xrightarrow{\alpha^{C}(i)} & \text{hocolim}_{I} C \\
\downarrow{\psi} & & \downarrow{\text{hocolim}_{I} \psi} \\
D(i) & \xrightarrow{\alpha^{D}(i)} & \text{hocolim}_{I} D
\end{array}
\]

for each \(i \in \text{ob}(I)\). Since the finiteness obstruction is invariant under equivalence of categories [16, Theorem 2.8], we may use Theorem 4.1(v) for C to obtain

\[ o(\text{hocolim}_{I} D; R) = (\text{hocolim}_{I} \psi)_{*}(o(\text{hocolim}_{I} C; R)) \]

\[ = (\text{hocolim}_{I} \psi)_{*} \left( \sum_{n \geq 0} (-1)^{n} \cdot \sum_{\lambda \in \Lambda_{n}} \alpha^{C}(i_{\lambda})_{*}(o(C(i_{\lambda}); R)) \right) \]

\[ = \sum_{n \geq 0} (-1)^{n} \cdot \sum_{\lambda \in \Lambda_{n}} (\text{hocolim}_{I} \psi)_{*} \circ \alpha^{C}(i_{\lambda})_{*}(o(C(i_{\lambda}); R)) \]

\[ = \sum_{n \geq 0} (-1)^{n} \cdot \sum_{\lambda \in \Lambda_{n}} \alpha^{D}(i_{\lambda})_{*} \circ (\psi_{i_{\lambda}})_{*}(o(D(i_{\lambda}); R)) \]

\[ = \sum_{n \geq 0} (-1)^{n} \cdot \sum_{\lambda \in \Lambda_{n}} \alpha^{D}(i_{\lambda})_{*}(o(D(i_{\lambda}); R)). \]

(vi) follows from (i), (iii) and (v) in the same way that Theorem 4.1(vi) follows from Theorem 4.1(iii) and (v).

(vii) Suppose that I is directly finite and D(i) is directly finite for every object \(i \in \text{ob}(I)\). Suppose also for every object \(i \in I\) the category D(i) is of type \((L^{2})\). By the proof of Corollary 4.2 above, the values of the strict functor C are directly finite categories. If \(\Gamma_{1}\) and \(\Gamma_{2}\) are equivalent categories, then \(\Gamma_{1}\) is both directly finite and of type \((L^{2})\) if and only if \(\Gamma_{2}\) is both directly finite and of type \((L^{2})\) [16, Lemma 5.15 (i)]. Since each D(i) is directly finite, of type \((L^{2})\), and equivalent to C(i), we see that each C(i) is also directly finite and of type \((L^{2})\). So we may now apply Theorem 4.1(i) and (vii) to \(\Gamma\) and conclude that \(\text{hocolim}_{I} C\) is directly finite and of type \((L^{2})\). Again using the preservation of the direct finiteness and \((L^{2})\) under equivalence [16, Lemma 5.15 (i)], and the
equivalence of \( \text{hocolim}_\mathcal{I} \mathcal{C} \) with \( \text{hocolim}_\mathcal{I} \mathcal{D} \), we see \( \text{hocolim}_\mathcal{I} \mathcal{D} \) is both directly finite and of type \((L^2)\).

To prove the formulas for \( \chi_f^{(2)} \) and \( \chi^{(2)} \), we use [15] Lemma 5.15 (iii), which says: if \( F : \Gamma_1 \rightarrow \Gamma_2 \) is an equivalence of categories, and both \( \Gamma_1 \) and \( \Gamma_2 \) are both directly finite and of type \((L^2)\), then \( U^{(1)}(F) \chi_f^{(2)}(\Gamma_1) = \chi_f^{(2)}(\Gamma_2) \) and \( \chi^{(2)}(\Gamma_1) = \chi^{(2)}(\Gamma_2) \). We apply this to the equivalences \( \psi_i \) and \( \text{hocolim}_\mathcal{I} \psi_i \), and use the commutativity of diagram (1.3). For readability, we write \( (\text{hocolim}_\mathcal{I} \psi) \), for \( U(\text{hocolim}_\mathcal{I} \psi) \) and \( \alpha(i\lambda) \), for \( U^{(1)}(\alpha(i\lambda)) \), et cetera.

\[
\chi_f^{(2)}(\text{hocolim}_\mathcal{I} \mathcal{D}) = (\text{hocolim}_\mathcal{I} \psi)_* \chi_f^{(2)} (\text{hocolim}_\mathcal{I} \mathcal{C})
\]

\[
= (\text{hocolim}_\mathcal{I} \psi)_* \sum_{n \geq 0} (-1)^n \sum_{\lambda \in \Lambda_n} \alpha(i\lambda)_* (\chi_f^{(2)}(\mathcal{C}(i\lambda)))
\]

\[
= \sum_{n \geq 0} (-1)^n \sum_{\lambda \in \Lambda_n} \alpha^{\mathcal{D}}(i\lambda)_* \circ (\psi_i)_* (\chi_f^{(2)}(\mathcal{C}(i\lambda)))
\]

\[
= \sum_{n \geq 0} (-1)^n \sum_{\lambda \in \Lambda_n} \alpha^{\mathcal{D}}(i\lambda)_* (\chi_f^{(2)}(\mathcal{D}(i\lambda))).
\]

The formula for \( \chi^{(2)} \) follows by summing up the components of the functorial \( L^2 \)-Euler characteristics.

### 4.2. The Case of an Indexing Category of Type \((FP^4)\)

The Homotopy Colimit Formula of Theorem [11] can be extended to the case, where \( \mathcal{I} \) is of type \((FP^4)\) and not necessarily of type \((FF^2)\) as follows (recall that the existence of a finite \( \mathcal{I} \)-CW-model for \( E\mathcal{I} \) implies \( \mathcal{I} \) is of type \((FP^4)\), since cellular chains then provide a finite free resolution of \( \mathcal{R}\)). The evaluation of the covariant functor

\[ E^\mathcal{R} : \mathcal{I} \rightarrow \mathcal{H}\text{-SPACES} \]

of \( \mathcal{H}\) at every object \( i \in \mathcal{I} \) is an \( \mathcal{H}\)-CW-complex. Composing it with the cellular chain complex functor yields a covariant functor

\[ C_* (E^\mathcal{H}) : \mathcal{I} \rightarrow \mathcal{RH}\text{-CHCOM} \]

whose evaluation at every object in \( \mathcal{I} \) is a free \( \mathcal{RH}\)-chain complexes. Since by assumption \( \mathcal{C}(i) \) is of type \((FP^4)\), \( C_* (E^\mathcal{H})(i) \) is \( \mathcal{RH}\)-chain homotopy equivalent to a finite projective \( \mathcal{RH}\)-chain complex for every object \( i \in \mathcal{I} \). Since \( R \text{mor}_{\mathcal{I}}(\cdot, i) \otimes_{\mathcal{R}^2} C_* (E^\mathcal{H}) \) is \( \mathcal{RH}\)-isomorphic to \( C_* (E^\mathcal{H}) \), we conclude for every finitely generated projective \( \mathcal{RH}\)-module \( P \) that \( P \otimes_{\mathcal{R}^2} C_* (E^\mathcal{H}) \) is \( \mathcal{RH}\)-chain homotopy equivalent to finite projective \( \mathcal{RH}\)-chain complex and in particular possesses a finiteness obstruction \( o(P \otimes_{\mathcal{R}^2} C_* (E^\mathcal{H})) \in K_0(\mathcal{RH}) \) (see Lück [23] Theorem 11.2 on page 212). Because of Lück [23] Theorem 11.2 on page 212 we obtain a homomorphism

\[ o_C : K_0(\mathcal{R}^2) \rightarrow K_0(\mathcal{RH}), \quad [P] \mapsto o(P \otimes_{\mathcal{R}^2} C_* (E^\mathcal{H})). \]
The chain complex version of the proof of Lemma 5.12 shows that the $RH$-chain complex $C_*\mathcal{I}$ of a terminal object $\mathcal{I}$ is a projective $RH$-module. Choose a finite projective $RH$-module $P_\alpha$ and an $RH$-chain homotopy equivalence $f_\alpha : P_\alpha \xrightarrow{\sim} C_*\mathcal{I}$. Then $f_* \otimes_{RH} \text{id} : P_\alpha \otimes_{RH} C_*\mathcal{I} \rightarrow C_*\mathcal{I}$ is an $RH$-chain homotopy equivalence of $RH$-chain complexes and $P_\alpha \otimes_{RH} C_*\mathcal{I}$ is isomorphism to finite projective $RH$-chain complex by Lück [23, Theorem 11.2 on page 212]. This implies

$$o(\Gamma; R) = o\left(P_\alpha \otimes_{RH} C_*\mathcal{I}\right).$$

We conclude from [23, Theorem 11.2 on page 212]

$$o\left(P_\alpha \otimes_{RH} C_*\mathcal{I}\right) = \sum_{n \geq 0} (-1)^n \cdot o\left(P_n \otimes_{RH} C_*\mathcal{I}\right)$$

Since $o(I; R)$ is $\sum_{n \geq 0} (-1)^n \cdot |P_n|$, this implies

**Theorem 4.4** (The Homotopy Colimit Formula for an indexing category of type $FP_R$). We obtain under the conditions above

$$\alpha_C(o(I; R)) = o(H; R).$$

**Remark 4.5.** See Section 7 for a comparison with Leinster’s Euler characteristic and his results.

5. Examples of the Homotopy Colimit Formula

We now present several examples of the Homotopy Colimit Formula Theorem 4.1. These include the cases: $\mathcal{I}$ with a terminal object, the constant functor, the trivial functor, homotopy pushouts, homotopy orbits, and the transport groupoid. For the transport groupoid in the finite case, see also Example 5.3.

**Example 5.1** (Homotopy Colimit Formula for $I$ with a terminal object). Suppose that $\mathcal{I}$ has a terminal object $t$ and $C : \mathcal{I} \rightarrow \text{CAT}$ is a functor. Then $\text{mor}_\mathcal{I}(\cdot, t)$ is a finite $\mathcal{I}$-CW model for $E\mathcal{I}$. If every category $C(i)$ is of type $FP_R$, then $o(H; R) = o(t)_o = o(C(t); R)$. If $\mathcal{I}$ and $C$ additionally satisfy the hypotheses of Theorem 4.1 then $\chi_f(H; R) = \chi_f(C(t); R)$ and $\chi(H; R) = \chi(C(t); R)$, as we anticipated in Example 5.3. Similar statements hold for $\chi^{(2)}$ and $\chi^{(2)}$ in the $L^2$ case.

**Example 5.2** (Homotopy Colimit Formula for a constant functor). Consider the situation of Theorem 4.1 in the special case where the covariant functor $C : \mathcal{I} \rightarrow \text{CAT}$ is constant $C \in \text{CAT}$. Suppose that $\mathcal{I}$ admits a finite $\mathcal{I}$-CW-model for $E\mathcal{I}$. Then we may draw various conclusions about the homotopy colimit $H = \mathcal{I} \times C$. If $\mathcal{I}$ and $C$ are of type $FP_R$, then so is $\mathcal{I} \times C$. If $\mathcal{I}$ and $C$ are of type $FP_R$, then so is $\mathcal{I} \times C$. The statements in Theorem 4.1 provide us with formulas in terms of $C$ for $o(\mathcal{I} \times C; R)$, $\chi_f(\mathcal{I} \times C; R)$, $\chi(\mathcal{I} \times C; R)$, $\chi^{(2)}(\mathcal{I} \times C)$, and $\chi^{(2)}(\mathcal{I} \times C)$. We recall that the invariants $o$, $\chi_f$, $\chi$, $\chi^{(2)}$, and $\chi^{(2)}$ are multiplicative, see Fiore–Lück–Sauer [16, Theorems 2.17, 4.22, and 5.17].

**Documenta Mathematica** 16 (2011) 301–354
Example 5.3 (Homotopy Colimit Formula for the trivial functor). Consider the situation of Theorem 4.1 in the special case where the covariant functor $C: \mathcal{I} \to \text{CAT}$ is constantly the terminal category, which consists of a single object and its identity morphism. Then $\text{hocolim}_I C$ agrees with $I$, as we see from Example 3.5. Obviously $C(i)$ is of type $(\text{FF}_R)$, its finiteness obstruction is $[R] \in K_0(R) = K_0(\text{RC}(i))$ and both its Euler characteristic and $L^2$-Euler characteristic equals 1. We obtain from Theorem 4.1

$$
\begin{align*}
o(I; R) &= \sum_{n \geq 0} (-1)^n \cdot \sum_{\lambda \in \Lambda_n} [R \text{mor}_I(? \rightarrow \lambda)] \quad \in K_0(\text{RT}); \\
\chi_f(I; R) &= \sum_{n \geq 0} (-1)^n \cdot \sum_{\lambda \in \Lambda_n} t_{\lambda} \quad \in U(\Gamma); \\
\chi(I; R) &= \sum_{n \geq 0} (-1)^n \cdot |\Lambda_n| \quad \in \mathbb{Z}; \\
\chi^{(2)}(I) &= \sum_{n \geq 0} (-1)^n \cdot \sum_{\lambda \in \Lambda_n} i_{\lambda} \quad \in U^{(1)}(\Gamma); \\
\chi^{(2)}(I) &= \sum_{n \geq 0} (-1)^n \cdot |\Lambda_n| \quad \in \mathbb{R}.
\end{align*}
$$

Example 5.4 (Homotopy pushout formula). Let $\mathcal{I}$ be the category with objects $j, k$ and $\ell$ such that there is precisely one morphism from $j$ to $k$ and from $j$ to $\ell$ and all other morphisms are identity morphisms.

$$
\mathcal{I} = \{ k \leftarrow j \rightarrow \ell \}
$$

By Example 2.6 the category $\mathcal{I}$ admits a finite model for the classifying $\mathcal{I}$-space $E\mathcal{I}$.

A covariant functor $C: \mathcal{I} \to \text{CAT}$ is the same as specifying three categories $C(j)$, $C(k)$ and $C(\ell)$ and two functors $C(g): C(j) \to C(k)$ and $C(h): C(j) \to C(\ell)$. Let $\mathcal{H} = \text{hocolim}\_I C$ be the homotopy colimit. Let $\alpha(i): C(i) \to \mathcal{H}$ be the canonical functor for $i = j, k, \ell$. Then we obtain a square of functors which commutes up to natural transformations

$$
\begin{array}{ccc}
C(j) & \xrightarrow{\alpha(j)} & C(k) \\
\downarrow & \alpha(k) & \downarrow \\
C(\ell) & \xrightarrow{\alpha(\ell)} & \mathcal{H}.
\end{array}
$$

It induces diagrams which do NOT commute in general

$$
\begin{array}{ccc}
K_0(\text{RC}(j)) & \xrightarrow{\alpha(j)_*} & K_0(\text{RC}(k)) \\
\downarrow & \alpha(k)_* & \downarrow \\
K_0(\text{RC}(\ell)) & \xrightarrow{\alpha(\ell)_*} & K_0(\mathcal{H})
\end{array}
$$

and

Documenta Mathematica 16 (2011) 301–354
Suppose that \( C(i) \) is of type \((\text{FP}_R)\) for \( i = j, k, \ell \). We conclude from Theorem 4.1 (iii) that \( H \) is of type \((\text{FP}_R)\) and

\[
\begin{align*}
o(H; R) &= \alpha(k)_* (\alpha(C(k); R)) + \alpha(\ell)_* (\alpha(C(\ell); R)) - \alpha(j)_* (\alpha(C(j); R)) \\
\chi_f(H; R) &= \alpha(k)_* (\chi_f(C(k); R)) + \alpha(\ell)_* (\chi_f(C(\ell); R)) - \alpha(j)_* (\chi_f(C(j); R)) \\
\chi(H; R) &= \chi(C(k); R) + \chi(C(\ell); R) - \chi(C(j); R) \\
\chi_f^2(H; R) &= \alpha(k)_* (\chi_f^2(C(k))) + \alpha(\ell)_* (\chi_f^2(C(\ell))) - \alpha(j)_* (\chi_f^2(C(j))) \\
\chi^2(H; R) &= \chi(C(k)) + \chi(C(\ell)) - \chi(C(j)) \in \mathbb{R}.
\end{align*}
\]

**Example 5.5 (Homotopy orbit formula).** Suppose that a group \( G \) acts on a category \( C \) from the left. This can be viewed as a covariant functor \( \tilde{G} \to \text{CAT} \) whose source is the groupoid \( \tilde{G} \) with one object and \( G \) as its automorphism group. Let \( H = \text{hocolim}_C \) be its homotopy colimit, also called the homotopy orbit. Notice that \( H \) and \( C \) have the same set of objects.

Suppose there is a finite model for \( BG \) of the group \( G \), or equivalently, a finite model for the \( \tilde{G} \)-classifying space \( E\tilde{G} \) of the category \( \tilde{G} \). Let \( \chi(BG) \in \mathbb{Z} \) be its Euler characteristic. Let \( \alpha: \tilde{C} \to H \) be the canonical inclusion. Suppose that \( C \) is of type \((\text{FP}_R)\). Then we conclude from Theorem 4.1 (iii) that \( H \) is of type \((\text{FP}_R)\) and we have

\[
\begin{align*}
o(H; R) &= \chi(BG) \cdot \alpha_* (o(C; R)) \in K_0(RH); \\
\chi_f(H; R) &= \chi(BG) \cdot \alpha_* (\chi_f(C; R)) \in U(H); \\
\chi(H; R) &= \chi(BG) \cdot \chi(C; R) \in \mathbb{Z}; \\
\chi_f^2(H; R) &= \chi(BG) \cdot \alpha_* (\chi_f^2(C; R)) \in U^1(H); \\
\chi^2(H; R) &= \chi(BG) \cdot \chi^2(C; R) \in \mathbb{R}.
\end{align*}
\]

**Example 5.6 (Transport groupoid).** Let \( G \) be a group and let \( S \) be a left \( G \)-set. Its transport groupoid \( G^G(S) \) has \( S \) as its set of objects. The set of morphisms from \( s_1 \) to \( s_2 \) is \( \{ g \in G \mid gs_1 = s_2 \} \). The composition is given by the multiplication in \( G \). Denote by \( S \) the category whose set of objects is \( S \) and which has no morphisms besides the identity morphisms. The group \( G \) acts from the left on \( S \). One easily checks that \( G^G(S) \) is the homotopy orbit of \( S \) defined in Example 5.5.

Recall from Fiore–Lück–Sauer [16, Lemma 6.15 (iv)]: if \( \Gamma \) is a quasi-finite EI-category and for any morphism \( f: x \to y \) in \( \Gamma \), the order of the finite group \( \{ g \in \text{aut}(x) \mid f \circ g = f \} \) is invertible in \( R \), then \( \Gamma \) is of type \((\text{FP}_R)\) if and only
if iso(Γ) is finite and for every object \( x \in \text{ob}(\Gamma) \) the trivial \( R[x] \)-module \( R \) is of type (FP\(_R\)). Thus, category \( \mathcal{S} \) is of type (FP\(_R\)) if and only if \( S \) is finite. Suppose that \( \mathcal{S} \) is of type (FP\(_R\)) and there is a finite model for \( BG \). Obviously \( o(\mathcal{S}; R) \) is given in \( K_0(\mathcal{S}) = \bigoplus_s K_0(R) \) by the collection \([ [R] \in K_0(R) \mid s \in S]\).

Suppose for simplicity that \( G \) acts transitively on \( S \). Fix an element \( s \in S \). Let \( G_s \) be its isotropy group. Since \( S \) is finite, \( G_s \) is a subgroup of \( G \) of finite index, namely \([G : G_s] = |S|\). The transport groupoid \( \mathcal{G}^G(S) \) is connected and the automorphism group of \( s \) is \( G_s \). Hence evaluation at \( s \) induces an isomorphism

\[
ev : K_0(R\mathcal{G}^G(S)) \xrightarrow{\cong} K_0(R[G_s]).
\]

The composition

\[
K_0(R\mathcal{S}) \xrightarrow{o} K_0(R\mathcal{G}^G(S)) \xrightarrow{\cong} K_0(R[G_s])
\]
sends \( o(\mathcal{S}; R) \) to \([S] \cdot [RG_s] \), where \( o : \mathcal{S} \to \mathcal{G}^G(S) \) is the obvious inclusion. Hence Example 5.5 implies

\[
ev(o(\mathcal{G}^G(S); R)) = \chi(BG) \cdot [S] \cdot [RG_s] \in K_0(RG_s).
\]

Since \( BG \) has a finite model, \( BG_s \) as a finite covering of \( BG \) has a finite model. The cellular \( RG_s \)-chain complex of \( EG_s \) yields a finite free resolution of the trivial \( RG_s \)-module \( R \). This implies

\[
ev(o(\mathcal{G}^G(S); R)) = \chi(BG_s) \cdot [RG_s] \in K_0(RG_s).
\]

Hence we obtain the equality in \( K_0(RG_s) \)

\[
\chi(BG_s) \cdot [RG_s] = \chi(BG) \cdot [S] \cdot [RG_s] = \chi(BG) \cdot [G : G_s] \cdot [RG_s].
\]

This is equivalent to the equality of integers

\[
\chi(BG_s) = \chi(BG) \cdot [G : G_s].
\]

This equation is compatible with the well-know fact that for a \( d \)-sheeted covering \( \overline{X} \to X \) of a finite CW-complex \( X \) the total space \( \overline{X} \) is again a finite CW-complex and we have \( \chi(\overline{X}) = d \cdot \chi(X) \).

For the transport groupoid in the finite case, see also Example 8.33.

**6. Combinatorial Illustrations of the Homotopy Colimit Formula**

The classical Inclusion-Exclusion Principle follows from the Homotopy Colimit Formula Theorem [11]. We can also easily calculate the cardinality of a coequalizer in \( \text{SETS} \) in certain cases. These are different proofs of Examples 3.4.d and 3.4.b of Leinster’s paper [21].

**Example 6.1 (Inclusion-Exclusion Principle).** Let \( X \) be a set and \( S_0, \ldots, S_q \) finite subsets of \( X \). Then

\[
|S_0 \cup S_1 \cup \cdots \cup S_q| = \sum_{\emptyset \neq J \subseteq [q]} (-1)^{|J|-1} \left| \bigcap_{j \in J} S_j \right|.
\]
Proof. Let $\mathcal{I}$ be the category in Example 2.7 and consider the finite $\mathcal{I}$-CW-model for its classifying $\mathcal{I}$-space constructed there. We define a functor $\mathcal{C} : \mathcal{I} \to \text{SETS}$ by $\mathcal{C}(J) := \bigcap_{j \in J} S_j$. The functor

$$\text{hocolim}_\mathcal{I} \mathcal{C} \to \text{colim}_\mathcal{I} \mathcal{C} = S_0 \cup S_1 \cup \cdots \cup S_q$$

is an equivalence of categories, since it is surjective on objects and fully faithful. We have

$$|S_0 \cup S_1 \cup \cdots \cup S_q| = \chi(S_0 \cup S_1 \cup \cdots \cup S_q)$$

$$= \chi(\text{hocolim}_\mathcal{I} \mathcal{C})$$

$$= \sum_{n \geq 0} (-1)^n \cdot \sum_{\lambda \in \Lambda_n} \chi(\mathcal{C}(i_\lambda))$$

$$= \sum_{n \geq 0} (-1)^n \cdot \sum_{J \subseteq [q] \text{ and } |J| = n+1} \chi(\mathcal{C}(J))$$

$$= \sum_{n \geq 0} (-1)^n \left( \sum_{J \subseteq [q] \text{ and } |J| = n+1} \left| \bigcap_{j \in J} S_j \right| \right)$$

$$= \sum_{\emptyset \neq J \subseteq [q]} (-1)^{|J|-1} \left| \bigcap_{j \in J} S_j \right| .$$

□

Example 6.2 (Cardinality of a Coequalizer). Let $\mathcal{I}$ be the category

$$a \xrightarrow{f} b \xleftarrow{g} c$$

and $\mathcal{C} : \mathcal{I} \to \text{SETS}$ a functor such that:

(i) the maps $\mathcal{C}f$ and $\mathcal{C}g$ are injective,

(ii) the images of the maps $\mathcal{C}f$ and $\mathcal{C}g$ are disjoint, and

(iii) the sets $\mathcal{C}a$ and $\mathcal{C}b$ are finite.

Then the coequalizer $\text{colim} \mathcal{C}$ has cardinality $|\mathcal{C}b| - |\mathcal{C}a|$. 

Proof. The assumptions that $\mathcal{C}f$ and $\mathcal{C}g$ are injective and have disjoint images imply that the functor

$$\text{hocolim}_\mathcal{I} \mathcal{C} \to \text{colim}_\mathcal{I} \mathcal{C}$$

is fully faithful. Clearly it is also surjective on objects, and hence an equivalence of categories. The category $\mathcal{I}$ has a finite $\mathcal{I}$-CW-model for its classifying $\mathcal{I}$-space, constructed explicitly in Example 2.7. By Theorem 4.1, we have

$$\chi(\text{colim}_\mathcal{I} \mathcal{C}) = \chi(\text{hocolim}_\mathcal{I} \mathcal{C})$$

$$= \sum_{n \geq 0} (-1)^n \cdot \sum_{\lambda \in \Lambda_n} \chi(\mathcal{C}(i_\lambda))$$

$$= \chi(\mathcal{C}b) - \chi(\mathcal{C}a)$$

$$= |\mathcal{C}b| - |\mathcal{C}a| .$$

Documenta Mathematica 16 (2011) 301–354
7. Comparison with Results of Baez–Dolan and Leinster

We recall Baez–Dolan’s groupoid cardinality \[3\] and Leinster’s Euler characteristic for certain finite categories \[21\], compare our Homotopy Colimit Formula with his result on compatibility with Grothendieck fibrations, prove an analogue for indexing categories \(I\) that admit finite \(I\)-CW-models for their classifying \(I\)-spaces, and finally mention a Homotopy Colimit Formula for Leinster’s invariant in a restricted case.

7.1. Review of Leinster’s Euler Characteristic. Let \(\Gamma\) be a category with finitely many objects and finitely many morphisms. A \textit{weighting} on \(\Gamma\) is a function \(q^\bullet : \text{ob}(\Gamma) \to \mathbb{Q}\) such that for all objects \(x \in \text{ob}(\Gamma)\), we have

\[
\sum_{y \in \text{ob}(\Gamma)} |\text{mor}_\Gamma(x, y)| \cdot q^y = 1.
\]

A \textit{coweighting} \(q^\bullet\) on \(\Gamma\) is a weighting on \(\Gamma^{\text{op}}\). If a finite category admits both a weighting \(q^\bullet\) and a coweighting \(q^\bullet\), then \(\sum_{y \in \text{ob}(\Gamma)} q^y = \sum_{x \in \text{ob}(\Gamma)} q^x\). For a discussion of which matrices have the form \((|\text{mor}_\Gamma(x, y)|)_{x, y \in \text{ob}(\Gamma)}\) for some finite category \(\Gamma\), see Allouch \[2\] and \[3\].

As proved in \[16\], free resolutions of the constant \(R\Gamma\)-module \(R\) give rise to weightings on \(\Gamma\).

\textbf{Theorem 7.1} (Weighting from a free resolution, Theorem 7.6 of Fiore–Lück–Sauer \[16\]). Let \(\Gamma\) be a finite category. Suppose that the constant \(R\Gamma\)-module \(R\) admits a finite free resolution \(P^\bullet\). If \(P_n\) is free on the finite \(\text{ob}(\Gamma)\)-set \(C_n\), that is

\[
P_n = B(C_n) = \bigoplus_{y \in \text{ob}(\Gamma)} C_n \otimes R \text{mor}_\Gamma(?; y),
\]

then the function \(q^\bullet : \text{ob}(\Gamma) \to \mathbb{Q}\) defined by

\[
q^y := \sum_{n \geq 0} (-1)^n \cdot |C^y_n|
\]

is a weighting on \(\Gamma\).

\textbf{Corollary 7.3} (Construction of a weighting from a finite \(I\)-CW-model for the classifying \(I\)-space, Corollary 7.8 of Fiore–Lück–Sauer \[16\]). Let \(I\) be a finite category. Suppose that \(I\) admits a finite \(I\)-CW-model \(X\) for the classifying \(I\)-space. Then the function \(q^\bullet : \text{ob}(I) \to \mathbb{Q}\) defined by

\[
q^y := \sum_{n \geq 0} (-1)^n (\text{number of } n\text{-cells of } X\text{ based at } y)
\]

is a weighting on \(I\).
As explained in Section 7.5 of [16], we use this Corollary to obtain several of Leinster’s weightings in [21] from \(\mathcal{I}\)-CW-models for \(\mathcal{I}\)-classifying spaces. If \(\mathcal{I}\) has a terminal object, then we obtain from the finite model in Example 2.4 the weighting which is 1 on the terminal object and 0 otherwise. The category \(\mathcal{I} = \{ j \rightrightarrows k \}\) in Example 2.5 has weighting \((q^j, q^k) = (-1,1)\). The category \(\mathcal{I} = \{ k \leftarrow j \rightarrow \ell \}\) in Example 2.6 has weighting \((q^j, q^k, q^\ell) = (-1,1,1)\). Lastly, the category in Example 2.7 has weighting \(q_J := (-1)^{|J|-1}\).

Weightings and coweightings play a key role in Leinster’s notion of Euler characteristic. See also Berger–Leinster [9].

**Definition 7.4 (Definition 2.2 of Leinster [21]).** A finite category \(\Gamma\) has an Euler characteristic in the sense of Leinster if it admits both a weighting and a coweighting. In this case, its Euler characteristic in the sense of Leinster is defined as

\[
\chi_L(\Gamma) := \sum_{y \in \text{ob}(\Gamma)} q^y = \sum_{x \in \text{ob}(\Gamma)} q_x
\]

for any choice of weighting \(q^y\) or coweighting \(q_x\).

The Euler characteristic of Leinster agrees with the groupoid cardinality of Baez–Dolan [4] in the case of a finite groupoid \(\mathcal{G}\), namely they are both

\[
\sum_{x \in \text{iso}(\mathcal{G})} \frac{1}{|\text{aut}_{\mathcal{G}}(x)|}.
\]

The Euler characteristic of Leinster agrees with our \(L^2\)-Euler characteristic in some cases, as in the following Lemma.

**Lemma 7.5 (Lemma 7.3 of Fiore–Lück–Sauer [16]).** Let \(\Gamma\) be a finite EI-category which is skeletal, i.e., if two objects are isomorphic, then they are equal. Suppose that the left \(\text{aut}_{\Gamma}(y)\)-action on \(\text{mor}_{\Gamma}(x,y)\) is free for every two objects \(x, y \in \text{ob}(\Gamma)\).

Then \(\Gamma\) is of type \((\text{FP}_\mathbb{C})\) and of type \((L^2)\), and has an Euler characteristic in the sense of Leinster. Furthermore, the \(L^2\)-Euler characteristic \(\chi^{(2)}(\Gamma)\) of Definition 7.4 coincides with Leinster’s Euler characteristic \(\chi_L(\Gamma)\) of Definition 7.4:

\[
\chi^{(2)}(\Gamma) = \chi_L(\Gamma).
\]

Moreover, these are both equal to

\[
\sum_{l \geq 0} (-1)^l \sum_{x_0, x_1 \in \text{ob}(\Gamma)} \sum_{|\text{aut}(x_0)| \cdot |\text{aut}(x_{l-1})| \cdots |\text{aut}(x_0)|} \frac{1}{|\text{aut}(x_1)| \cdot |\text{aut}(x_{l-1})| \cdots |\text{aut}(x_0)|},
\]

where the inner sum is over all paths \(x_0 \to x_1 \to \cdots \to x_l\) from \(x_0\) to \(x_l\) such that \(x_0, \ldots, x_l\) are all distinct [16 Example 6.33].

This concludes the review of Leinster’s and Baez–Dolan’s invariants and how they relate to our \(L^2\)-Euler characteristic. Next we turn to a comparison of homotopy colimit results.

*Documenta Mathematica* 16 (2011) 301–354
7.2. Comparison with Leinster’s Proposition 2.8. Leinster’s result on homotopy colimits, rephrased in our notation to make the comparison more apparent, is below.

**Theorem 7.6 (Proposition 2.8 of Leinster [21]).** Let $I$ be a category with finitely many objects and finitely many morphisms, and $C : I \to \text{CAT}$ a pseudo functor. Assume that $\text{hocolim}_I C$ has finitely many objects and finitely many morphisms. Let $q^\bullet$ be a weighting on $I$ and suppose that $\text{hocolim}_I C$ and all $C(i)$ have Euler characteristics. Then

$$\chi_L(\text{hocolim}_I C) = \sum_{i \in \text{ob}(I)} q^i \chi_L(C(i)).$$

For example, if $I = \{k \leftarrow j \to \ell\}$, then $I$ admits the weighting $(q^j, q^k, q^\ell) = (-1, 1, 1)$ as discussed above. If $C : I \to \text{CAT}$ is a pseudo functor, and the homotopy pushout has finitely many objects and finitely many morphisms, and $\text{hocolim}_I C$ and all $C(i)$ have Euler characteristics, then Leinster’s result says that the homotopy pushout has the Euler characteristic $\chi_L(C(k)) + \chi_L(C(\ell)) - \chi_L(C(j)).$

Leinster’s Proposition 2.8 tells us how the Euler characteristic is compatible with Grothendieck fibrations. We can remove the hypothesis of finite from that Proposition, at the expense of requiring a finite model, as in the following theorem for our invariants.

**Theorem 7.7.** Let $I$ be a finite category. Suppose that $I$ admits a finite $I$-CW-model $X$ for the classifying $I$-space of $I$. Let $q^\bullet : \text{ob}(I) \to \mathbb{Q}$ be the $I$-Euler characteristic of $X$, namely

$$q^i := \sum_{n \geq 0} (-1)^n (\text{number of } n\text{-cells of } X \text{ based at } i).$$

Let $C : I \to \text{CAT}$ be a functor such that for every object $i$ the category $C(i)$ is of type $(\text{FP}_R)$. Suppose that $I$ is directly finite and $C(i)$ is directly finite for all $i \in \text{ob}(I)$. Then

$$\chi(\text{hocolim}_I C; R) = \sum_{i \in \text{ob}(I)} q^i \chi(C(i); R).$$

If each $C(i)$ is of type $(L^2)$ rather than $(\text{FP}_R)$, we have

$$\chi^{(2)}(\text{hocolim}_I C) = \sum_{i \in \text{ob}(I)} q^i \chi^{(2)}(C(i)).$$
Proof. By Theorem 4.1 (vi) we have
\[
\chi(hocolim C; R) = \sum_{n \geq 0} (-1)^n \sum_{\lambda \in \Lambda_n} \chi(C(i_\lambda); R)
\]
\[
= \sum_{n \geq 0} (-1)^n \sum_{i \in ob(I)} (\text{number of } n\text{-cells of } X \text{ at } i) \chi(C(i); R)
\]
\[
= \sum_{i \in ob(I)} \sum_{n \geq 0} (-1)^n (\text{number of } n\text{-cells of } X \text{ at } i) \chi(C(i); R)
\]
\[
= \sum_{i \in ob(I)} q^i \chi(C(i); R).
\]
The statement for \(\chi(2)\) is proved similarly from Theorem 4.1 (vii). \(\square\)

Remark 7.8. Whenever \(\chi(colim C; R) = \chi(hocolim C; R)\), Theorem 4.1 and Theorem 7.7 can be used to calculate the Euler characteristic of a colimit. Indeed, the hypotheses of Examples 6.1 and 6.2 guaranteed the equivalence of the colimit and the homotopy colimit, and this equivalence was a crucial ingredient in those proofs. For example, under Leinster’s hypothesis of familial representability on \(C\), each connected component of hocolim \(I C\) has an initial object, so
\[
\chi(hocolim C; R) = \chi(colim C; R)
\]
(recall that colim \(C\) is the set of connected components of hocolim \(C\) whenever \(C\) takes values in \(SETS\)). This is the role of familial representability in his Examples 3.4.

As a corollary to our Homotopy Colimit Formula for the \(L^2\)-Euler characteristic, we have a Homotopy Colimit Formula for Leinster’s Euler characteristic when they agree.

Corollary 7.9 (Homotopy Colimit Formula for Leinster’s Euler characteristic). Let \(I\) be a skeletal, finite, EI-category such that the left aut\(_I(y)\)-action on mor\(_I(x, y)\) is free for every two objects \(x, y \in ob(I)\). Assume there exists a finite \(I\)-CW-model for the \(I\)-classifying space of \(I\). Let \(C: I \to \text{CAT}\) be a covariant functor such that for each \(i \in ob(I)\), the category \(C(i)\) is a skeletal, finite, EI and the left aut\(_{C(i)}(d)\)-action on mor\(_{C(i)}(c, d)\) is free for every two objects \(c, d \in ob(C(i))\). Assume for every object \(i \in ob(I)\), for each automorphism \(u: i \to i\) in \(I\), and each \(\bar{x} \in \text{iso}(C(i))\) we have \(C(u)(x) = \bar{x}\). Then \(\mathcal{H} := hocolim_{i \in I} C\) is again a skeletal, finite, EI-category such that the left aut\(_{\mathcal{H}}(h)\)-action on mor\(_{\mathcal{H}}(g, h)\) is free for every two objects \(g, h \in ob(hocolim_{i \in I} C)\), and
\[
\chi_{\mathcal{H}}(C(i)) = \sum_{n \geq 0} (-1)^n \sum_{\lambda \in \Lambda_n} \chi_{\mathcal{H}}(C(i_\lambda); R).
\]

Proof. The category \(\mathcal{H}\) is an EI-category by Theorem 4.1 (ii). Skeletality and finiteness of \(\mathcal{H}\) follow directly from the skeletality and finiteness of \(I\) and \(C(i)\), and the definition of \(\mathcal{H}\). The hypotheses on \(C(i)\) imply that...
Euler Characteristics and Homotopy Colimits

\[ \chi^{(2)}(C(i)) = \chi_L(C(i)) \] by Theorem 7.5, and similarly \( \chi^{(2)}(H) = \chi_L(H) \). Finally, Theorem 4.1 (vii), which is the Homotopy Colimit Formula for the \( L^2 \)-Euler characteristic \( \chi^{(2)} \), implies the formula is also true for Leinster’s Euler characteristic \( \chi_L \) in the special situation of the Corollary.

□

8. Scwols and Complexes of Groups

As an illustration of the Homotopy Colimit Formula, we consider Euler characteristics of small categories without loops (scwols) and complexes of groups in the sense of Haefliger [18, 19] and Bridson–Haefliger [11]. One-dimensional complexes of groups are the classical Bass–Serre graphs of groups [37]. For finite scwols, the Euler characteristic, \( L^2 \)-Euler characteristic, and Euler characteristic of the classifying space all coincide, essentially because finite scwols admit finite models for their classifying spaces. The Euler characteristic of a finite scwol is particularly easy to find: one simply chooses a skeleton, counts the paths of non-identity morphisms of length \( n \), and then computes the alternating sum of these numbers.

Scwols and complexes of groups are combinatorial models for polyhedral complexes and group actions on them. The poset of faces of a polyhedral complex is a scwol. Suppose a group \( G \) acts on an \( M_\kappa \)-polyhedral complex by isometries preserving cell structure, and suppose each group element \( g \in G \) fixes each cell pointwise that \( g \) fixes setwise. In this case, the quotient is also an \( M_\kappa \)-polyhedral complex, say \( Q \), and we obtain a pseudo functor from its scwol of faces into groups. Namely, to a face \( \sigma \) of \( Q \), one associates the stabilizer \( G_\sigma \) for a selected representative \( \sigma \) of \( \sigma \). Inclusions of subfaces of \( Q \) then correspond to inclusions of stabilizer subgroups up to conjugation. This pseudo functor is the complex of groups associated to the group action.

However, it is sometimes easier to work directly with the combinatorial model rather than with the original \( M_\kappa \)-polyhedral complex, and consider instead appropriate group actions on the associated scwol, as in Definition 8.11. Then the quotient category of a scwol is again a scwol, and the associated pseudo functor on the quotient scwol is called the complex of groups associated to the group action. Any group-valued pseudo functor on a scwol that arises in this way is called developable.

Our main results in this section concern the Euler characteristics of homotopy colimits of complexes of groups associated to group actions in the sense of Definition 8.11. Theorem 8.30 concludes that the Euler characteristic and \( L^2 \)-Euler characteristic of the homotopy colimit are \( \chi(\mathcal{X}/G) \) and \( \chi^{(2)}(\mathcal{X})/|G| \) respectively, \( G \) and \( \mathcal{X} \) are finite. These formulas provide necessary conditions for developability. That is, if \( F \) is a pseudo functor from a scwol \( \mathcal{Y} \) to groups, one may ask if there are a scwol \( \mathcal{X} \) and a group \( G \) such that \( \mathcal{Y} \) is isomorphic to \( \mathcal{X}/G \) and \( F \) is the associated complex of groups. To obtain conditions on \( \chi(\mathcal{X}), \chi^{(2)}(\mathcal{X}) \), and \( |G| \), one forms the homotopy colimit of \( F \), calculates its Euler characteristic and \( L^2 \)-Euler characteristic, and then compares with the
formulas of Theorem 8.30. A simple case is illustrated in Example 8.31. Another application of the formulas is the computation of the Euler characteristic and $L^2$-Euler characteristic for the transport groupoid of a finite left $G$-set, as in Example 8.33. We finish with Theorem 8.35, which extends Haefliger’s formula for the Euler characteristic of the classifying space of the homotopy colimit of a complex of groups in terms of Euler characteristics of lower links and groups.

One novel aspect of our approach is that we do not require scwols to be skeletal. We prove in Theorem 8.24 that any scwol with a $G$-action in the sense of Definition 8.11 can be replaced by a skeletal scwol with a $G$-action, and this process preserves quotients, stabilizers, complexes of groups, and homotopy colimits. Moreover, if the initial $G$-action was free on the object set, then so is the $G$-action on the object set of the skeletal replacement.

We begin by recalling the notions in Chapter III.C of Bridson–Haefliger [11], rephrased in the conceptual framework of 2-category theory.

**Notation 8.1 (2-Category of groups).** We denote by $\text{GROUPS}$ the 2-category of groups. Objects are groups and morphisms are group homomorphisms. The 2-cells are given by conjugation: a 2-cell $(g, a)$

\[
\begin{array}{ccc}
H & \xrightarrow{(g,a)} & G \\
\downarrow & & \downarrow \\
a' & \xleftarrow{g} & a
\end{array}
\]

is an element $g \in G$ such that $ga(h)g^{-1} = a'(h)$ for all $h \in H$. The vertical composition is $(g_2, a_2) \circ (g_1, a_1) = (g_2g_1, a_1)$ and the horizontal composition of

\[
\begin{array}{ccc}
H & \xrightarrow{(g,a)} & G \\
\downarrow & & \downarrow \\
a' & \xleftarrow{g} & a
\end{array}
\begin{array}{ccc}
G & \xrightarrow{(k,b)} & K \\
\downarrow & & \downarrow \\
b' & \xleftarrow{k} & b
\end{array}
\]

is $(kb(g), ba)$.

**Definition 8.2 (Scwol).** A scwol is a small category Without loops, that is, a small category in which every endomorphism is trivial.

**Example 8.3.** The categories $\{j \rightrightarrows k\}$ and $\mathcal{I} = \{k \leftleftarrows j \rightarrow \ell\}$ of Examples 2.5 and 2.6 are scwols. Every partially ordered set is a scwol, for example, the set of simplices of a simplicial complex, ordered by the face relation, is a scwol. The poset of non-empty subsets of $[q]$, and its opposite category in Example 2.7, are scwols. The opposite category of a scwol is also a scwol.

---

3Bridson–Haefliger additionally require scwols to be skeletal [11 page 574]. However, we do not require scwols to be skeletal, since we prove in Theorem 8.24 that general statements about scwols can be reduced to the skeletal case.

---

**Documenta Mathematica** 16 (2011) 301–354
Lemma 8.4. Every scwol is an EI-category and consequently also directly finite.

Proof. Every endomorphism in a scwol is trivial, and therefore an automorphism, so every scwol is an EI-category. By Fiore–Lück–Sauer [16] Lemma 3.13, every EI-category is also directly finite.

For a direct proof of direct finiteness: if \( u : x \to y \) and \( v : y \to x \) are morphisms in a scwol, then \( vu \) and \( uv \) are automorphisms, and hence both \( vu = \text{id}_x \) and \( uv = \text{id}_y \) hold automatically. □

Theorem 8.5 (Finite scwols admit finite models). Suppose \( \mathcal{I} \) is a finite scwol. Then \( \mathcal{I} \) admits a finite \( \mathcal{I} \)-CW-model for its \( \mathcal{I} \)-classifying space in the sense of Definition 2.3.

Proof. By Lemma 2.13 we may assume that \( \mathcal{I} \) is skeletal.

Since \( \mathcal{I} \) has only finitely many morphisms, no nontrivial isomorphisms, and no nontrivial endomorphisms, there are only finitely many paths of non-identity morphisms. Thus the bar construction of \( E^{\text{bar}} \mathcal{I} \) Remark 2.10 has only finitely many \( \mathcal{I} \)-cells. □

Corollary 8.6. Any finite scwol \( \mathcal{I} \) is of types \((\text{FF}_R)\) and \((\text{FP}_R)\) for every associative, commutative ring \( R \) with identity. Moreover, any finite scwol is also of type \((L^2)\).

Proof. The cellular \( R \)-chains of the finite model in Theorem 8.5 provide a finite, free resolution of the constant module \( R \). By Theorem 1.14 any directly finite category of type \((\text{FP}_C)\) is of type \((L^2)\). Scwols are directly finite by Lemma 8.4. □

Example 8.7 (Invariants coincide for finite scwols). Let \( \mathcal{I} \) be any finite scwol. Then by Corollary 8.6 it is of type \((\text{FF}_R)\) for any associative, commutative ring with identity, and by Theorems 1.9 and 1.15 we have

\[
\chi(\mathcal{I}; R) = \chi(B\mathcal{I}; R) = \chi^{(2)}(\mathcal{I}).
\]

If \( \Gamma \) is any skeleton of \( \mathcal{I} \), then by (1.14),

\[
(8.8) \quad \chi(\Gamma; R) = \sum_{n \geq 0} (-1)^n c_n(\Gamma),
\]

where \( c_n(\Gamma) \) is the number of paths of \( n \)-many non-identity morphisms in \( \Gamma \).

But by Fiore–Lück–Sauer [16] Theorem 2.8 and Corollary 4.19, type \((\text{FF}_R)\) and the Euler characteristic are invariant under equivalence of categories between directly finite categories, so \( \chi(\mathcal{I}; R) = \chi(\Gamma; R) \) and all three invariants \( \chi(\mathcal{I}; R), \chi(B\mathcal{I}; R), \chi^{(2)}(\mathcal{I}) \) are given by (8.8).

We now arrive at the main notion of this section: a complex of groups. We will apply our Homotopy Colimit Formula to complexes of groups.

Definition 8.9 (Complex of groups). Let \( \mathcal{Y} \) be a scwol. A complex of groups over \( \mathcal{Y} \) is a pseudo functor \( F : \mathcal{Y} \to \text{GROUPS} \) such that \( F(a) \) is injective for every morphism \( a \) in \( \mathcal{Y} \). For each object \( \sigma \) of \( \mathcal{Y} \), the group \( F(\sigma) \) is called the local group at \( \sigma \).
In 2.5 and 2.4 of [18] and [19] respectively, Haefliger denotes by $CG(X)$ the homotopy colimit of a complex of groups $G(X) : C(X) \to \text{GROUPS}$. Bridson–Haefliger use the notation $CG(Y)$ in [11, III.C.2.8]. The fundamental group of a complex of groups $G(X)$ in the sense of [11] Definition 3.5 on p. 548 equals the fundamental group of the geometric realization of $CG(X)$ [11, Appendix A.12 on p. 578 and Remark A.14 on p. 579]. Categories which are homotopy colimits of complexes of groups are characterized by Haefliger on page 283 of [19]. From the homotopy colimit $CG(X)$, Haefliger reconstructs the category $C(X)$ and the complex of groups $G(X)$ up to a coboundary on pages 282-283 of [19]. Every aspherical realization [19, Definition 3.3.4] of $G(X)$ has the homotopy type of the geometric realization of the homotopy colimit, denoted $BG(X)$ [19, page 296]. The homotopy colimit also plays a role in the homology and cohomology of complexes of groups [19, Section 4]; a left $G(X)$-module is a functor $CG(X) \to \text{ABELIAN-GROUPS}$.

We return to our recollection of complexes of groups and examples that arise from group actions.

**Definition 8.10 (Morphism from a complex of groups to a group).** A morphism from a complex of groups $F$ to a group $G$ is a pseudo natural transformation $F \Rightarrow \Delta_G$, where $\Delta_G$ indicates the constant 2-functor $Y \to \text{GROUPS}$ with value $G$.

A typical example of a complex of groups equipped with a morphism to a group $G$ arises from an action of a group $G$ on a scwol, as we now explain.

**Definition 8.11 (Group action on a scwol, 1.11 of Bridson–Haefliger [11]).** An action of a group $G$ on a scwol $X$ is a group homomorphism from $G$ into the group of strictly invertible functors $X \to X$ such that

(i) For every nontrivial morphism $a$ of $X$ and every $g \in G$, we have $gs(a) \neq t(a)$,

(ii) For every nontrivial morphism $a$ of $X$ and every $g \in G$, if $gs(a) = s(a)$, then $ga = a$.

**Example 8.12.** The group $G = \mathbb{Z}_2$ acts in the sense of Definition 8.11 on the scwol $X$ pictured below.

```
  x  h  z
 g   ̸= h'
 y
```

The group $\mathbb{Z}_2$ permutes respectively $x$ and $x'$, $g$ and $g'$, and $h$ and $h'$. The objects $y$ and $z$ are fixed. This action of $\mathbb{Z}_2$ on $X$ is a combinatorial model for a reflection action on $S^1$.

**Example 8.13.** Consider the scwol $X$ pictured below. The group $G = \{ \pm 1 \} \times \mathbb{Z}$ acts on $X$ in the sense of Definition 8.11 where $-1 \cdot m := -m$ and $n \cdot m := m + 2n$. 

```
  ... → -2 ← -1 ← 0 ← 1 → 2 ← ... 
```
This action of \( \{\pm 1\} \ltimes \mathbb{Z} \) on \( X \) is a combinatorial model for the reflection and translation action on \( \mathbb{R} \).

**Lemma 8.14** (Consequences of group action conditions). If a group \( G \) acts on a scwol \( X \) in the sense of Definition 8.11 then the following statements hold.

(i) If \( \sigma \) is an object of \( X \) and \( g, h \in G \), then \( g \sigma \cong h \sigma \) implies \( g \sigma = h \sigma \).

(ii) If \( a \) is a morphism in \( X \) and \( g, h \in G \), then \( gs(a) = hs(a) \) implies \( ga = ha \).

(iii) If \( \sigma \cong \tau \), then the stabilizers \( G_{\sigma} \) and \( G_{\tau} \) are equal.

**Proof.** For statement (i) \( g \sigma \cong h \sigma \) implies \( \sigma \cong (g^{-1}h)\sigma \), so \( \sigma = (g^{-1}h)\sigma \) by Definition 8.11 part (i) and \( ga = ha \).

For statement (ii) \( gs(a) = hs(a) \) implies \( (h^{-1}g)s(a) = s(a) \) and \( (h^{-1}g)a = a \) by Definition 8.11 part (ii), and finally \( ga = ha \).

For statement (iii) suppose \( \sigma \cong \tau \) and \( g \sigma = \sigma \). We have \( \tau \cong \sigma = g \sigma \cong g \tau \).

Then \( \tau = g \tau \) by (i) and \( G_{\sigma} \subseteq G_{\tau} \). The proof is symmetric, so we also have \( G_{\tau} \subseteq G_{\sigma} \). \qed 

**Definition 8.15** (Quotient of a scwol by a group action). If a scwol \( X \) is equipped with a \( G \)-action as above, then the quotient scwol \( X/G \) has objects and morphisms

\[
\text{ob}(X/G) := (\text{ob}(X))/G \\
\text{mor}(X/G) := (\text{mor}(X))/G.
\]

Composition and identities are induced by those of \( X \).

**Remark 8.16** (III.C.1.13 of Bridson–Haefliger [11]). The projection functor \( p: X \to X/G \) induces a bijection

\[
(8.17) \quad \{a \in \text{mor}(X)|sa = x\} \longrightarrow \{b \in \text{mor}(X/G)|sb = p(x)\}
\]

for each \( x \in X \). If \( G/X \) is connected and the action of \( G \) on \( \text{ob}(X) \) is free, then \( p \) is a covering of scwols. That is, in addition to the bijection (8.17), \( p \) induces a bijection

\[
(8.18) \quad \{a \in \text{mor}(X)|ta = x\} \longrightarrow \{b \in \text{mor}(X/G)|tb = p(x)\}
\]

for each \( x \in X \).

**Lemma 8.19** (Quotients of skeletal scwols are skeletal). If \( X \) is a skeletal scwol, and a group \( G \) acts on \( X \) in the sense of Definition 8.11 then the quotient scwol \( X/G \) is also skeletal.

**Proof.** Suppose \( \overline{\sigma} \) is isomorphic to \( \overline{\tau} \) in \( X/G \). We show \( \overline{\sigma} \) is actually equal to \( \overline{\tau} \). If \( \overline{\sigma} \to \overline{\tau} \) is an isomorphism with inverse \( \overline{\tau} \), then there are lifts \( a: \sigma \to \tau \) and \( b: \tau \to \sigma' \) in \( X \), and an element \( g \in G \) such that \( g(ba) = \text{id}_{\sigma} \). Since \( g \) fixes the source of \( ba \), the group element \( g \) fixes also \( ba \), so \( ba = \text{id}_{\sigma} \) and \( \sigma' = \sigma \). Since \( ab \) is an endomorphism of \( \tau \), it is therefore \( \text{id}_{\tau} \). By the skeletality of \( X \), we have \( \sigma = \tau \), and also \( \overline{\sigma} = \overline{\tau} \). \qed
LEMMA 8.20 (Quotient of path set is set of paths in quotient). Suppose \( \mathcal{X} \) is a scwol equipped with an action of a group \( G \) in the sense of Definition 8.11. Let \( \Lambda_n(\mathcal{X}) \) respectively \( \Lambda_n(\mathcal{X}/G) \) denote the set of paths of \( n \)-many non-identity composable morphisms in \( \mathcal{X} \) respectively \( \mathcal{X}/G \). Give \( \Lambda_n(\mathcal{X}) \) the induced \( G \)-action. Then the function

\[
\Lambda_n(\mathcal{X}) \rightarrow \Lambda_n(\mathcal{X}/G)
\]

\[
(a_1, \ldots, a_n) \mapsto (\overline{a}_1, \ldots, \overline{a}_n)
\]

induces a bijection \( \Lambda_n(\mathcal{X})/G \rightarrow \Lambda_n(\mathcal{X}/G) \).

Proof. Remark 8.19 implies that a path \( (a_1, \ldots, a_n) \) in \( \mathcal{X} \) consists entirely of non-identity morphisms if and only if the projection \( (\overline{a}_1, \ldots, \overline{a}_n) \) in \( \mathcal{X}/G \) consists entirely of non-identity morphisms, so from now on we work only with non-identity morphisms. Note

\[
(g_1a_1, g_2a_2, \ldots, g_na_n) = (g_1a_1, g_1a_2, \ldots, g_1a_n)
\]

by Definition 8.11(ii). For injectivity, we have \( (\overline{a}_1, \ldots, \overline{a}_n) = (\overline{b}_1, \ldots, \overline{b}_n) \) if and only if for some \( g_i \in G \)

\[
(g_1a_1, g_2a_2, \ldots, g_na_n) = (b_1, \ldots, b_n),
\]

which happens if and only if for some \( g \in G \)

\[
(ga_1, ga_2, \ldots, ga_n) = (b_1, \ldots, b_n),
\]

(take \( g = g_1 \)). For the surjectivity, we can lift any path \( (\overline{a}_1, \ldots, \overline{a}_n) \) by first lifting \( \overline{a}_1 \) to \( a_1 \), then \( \overline{a}_2 \) to \( a_2 \), and so on using Remark 8.19. \( \square \)

DEFINITION 8.21 (Complex of groups from a group action on a scwol, 2.9 of Bridson–Haefliger [11]). Let \( G \) be a group and \( \mathcal{X} \) a scwol upon which \( G \) acts in the sense of Definition 5.11. Let \( p: \mathcal{X} \rightarrow \mathcal{X}/G \) denote the quotient map. Haefliger and Bridson–Haefliger define a pseudo functor \( F: \mathcal{X}/G \rightarrow \text{GROUPS} \) as follows. In the procedure choices are made, but different choices lead to isomorphic complexes of groups. For each object \( \overline{\sigma} \) of \( \mathcal{X}/G \), choose an object \( \sigma \) of \( \mathcal{X} \) such that \( p(\sigma) = \overline{\sigma} \) (our overline convention is the opposite of that in [11]). Then \( F(\overline{\sigma}) \) is defined to be \( G_\sigma \), the isotropy group of \( \sigma \) under the \( G \)-action.

If \( \overline{\sigma}: \overline{\pi} \rightarrow \overline{\tau} \) is a morphism in \( \mathcal{X}/G \), then there exists a unique morphism \( a \) in \( \mathcal{X} \) such that \( p(a) = \overline{\sigma} \) and \( sa = \sigma \), as in 5.17. For \( \overline{\sigma} \) we choose an element \( h_{\overline{\sigma}} \in G \) such that \( h_{\overline{\sigma}} \cdot ta \) is the object \( \tau \) of \( \mathcal{X} \) chosen above so that \( p(\tau) = \overline{\tau} \).

An injective group homomorphism \( F(\overline{\pi}): G_\sigma \rightarrow G_\tau \) is defined by

\[
F(\overline{\pi})(g) := h_{\overline{\sigma}}gh_{\overline{\tau}}^{-1}.
\]

Suppose \( \overline{\pi} \) and \( \overline{b} \) are composable morphisms of \( \mathcal{X}/G \). We define a 2-cell in \( \text{GROUPS} \)

\[
F_{\overline{\pi}}: F(\overline{b}) \circ F(\overline{\pi}) \Rightarrow F(\overline{b} \circ \overline{\pi})
\]

to be \((h_{\overline{\sigma}}h_{\overline{\tau}}^{-1}h_{\overline{\tau}}^{-1}, F(\overline{b}) \circ F(\overline{\pi}))\) as in Notation 5.1. The pseudo functor \( F: \mathcal{X}/G \rightarrow \text{GROUPS} \) is called the complex of groups associated to the group action of \( G \) on the scwol \( \mathcal{X} \). This complex of groups
comes equipped with a morphism to the group $G$, that is, a pseudo natural transformation $F \Rightarrow \Delta G$. The inclusion of each isotropy group $F(\sigma) = G_{\sigma}$ into $G$ provides the components of the pseudo natural transformation.

**Example 8.22.** The quotient scwols for the actions in Examples 8.12 and 8.13 are both $\{k \leftarrow j \rightarrow \ell\}$, and the associated complexes of groups are both $\mathbb{Z}_2 \leftarrow \{0\} \rightarrow \mathbb{Z}_2$.

**Remark 8.23.** If a group $G$ acts on a scwel in the sense of Definition 8.11 each object stabilizer is finite, and the quotient scwel is finite, then the associated complex of groups $F: \mathcal{X}/G \rightarrow \text{GROUPS}$ satisfies all of the hypotheses of the Homotopy Colimit Formula in Theorem 4.1 (vii) and in Corollary 4.2 (vii). If, in addition, $R$ is a ring such that the order $|H|$ of each object stabilizer $H \subset G$ is invertible in $R$, then $F: \mathcal{X}/G \rightarrow \text{GROUPS}$ also satisfies all of the hypotheses of the Homotopy Colimit Formula in Theorem 4.1 (vi) and in Corollary 4.2 (vi). See Examples 8.12, 8.13, and 8.22.

Even without finiteness assumptions, it is possible to replace scwols with skeletal scwols and preserve much of the accompanying structure, as Theorem 8.24 explains.

**Theorem 8.24 (Reduction to skeletal case).** Let $G$ be a group acting on a scwel $\mathcal{X}$ in the sense of Definition 8.11. Let $\Gamma$ be any skeleton of $\mathcal{X}$, $i: \Gamma \rightarrow \mathcal{X}$ the inclusion, and $r: \mathcal{X} \rightarrow \Gamma$ a functor equipped with a natural isomorphism $ir \cong \text{id}_\mathcal{X}$, and satisfying $ri \cong \text{id}_\Gamma$. Then there is a $G$-action on the scwel $\Gamma$ in the sense of Definition 8.11 such that following hold.

(i) The functor $r$ is $G$-equivariant.

(ii) The induced functor $\overline{r}$ on quotient categories is an equivalence of categories compatible with the quotient maps, that is, the diagram below commutes.

\[
\begin{array}{ccc}
\mathcal{X} & \overset{r}{\longrightarrow} & \Gamma \\
p^{\mathcal{X}} \downarrow & & \downarrow p^\Gamma \\
\mathcal{X}/G & \overset{\tau}{\longrightarrow} & \Gamma/G
\end{array}
\]

(iii) The inclusion $i: \Gamma \rightarrow \mathcal{X}$ preserves stabilizers, that is $G_{\gamma} = G_{\overline{\gamma}}$ for all $\gamma \in \text{ob}(\Gamma)$. Note that the inclusion may not be $G$-equivariant.

(iv) Choices can be made in the definitions of $F^\mathcal{X}$ and $F^\Gamma$ (the complexes of groups associated to the $G$-actions on $\mathcal{X}$ and $\Gamma$ in Definition 8.21), so that the diagram below strictly commutes.

\[
\begin{array}{ccc}
\mathcal{X}/G & \overset{\tau}{\longrightarrow} & \Gamma/G \\
\downarrow F^{\mathcal{X}} & & \downarrow F^\Gamma \\
\text{GROUPS} & \overset{i}{\longrightarrow} & \text{GROUPS}
\end{array}
\]
The functor $(\mathcal{T}, \text{id})$ is an equivalence of categories

$$(\mathcal{T}, \text{id}): \text{hocolim}_X F_X \xrightarrow{\sim} \text{hocolim}_{\Gamma / G} F^\Gamma.$$
The vertical morphisms must be identities by skeletality of $\Gamma$ and the no loops condition, so $\varphi(g)r(f) = r(gf)$. Equivariance on objects then follows by taking $f = \text{id}_x$.

(ii) Diagram (8.25) commutes by definition of $\tau$. The functor $\tau$ is surjective on objects because $p^X$ and $p^\Gamma$ are. The functor $\tau$ is fully faithful since the equivariant bijection $r(x, y) \colon \text{mor}_X(x, y) \to \text{mor}_\Gamma(r(x), r(y))$ induces the equivariant bijection $\tau(p^Xx, p^\Gamma y)$.

(iii) Let $\gamma \in \text{ob}(\Gamma)$, and suppose $g_i\gamma = i\gamma$. Then
\[
\varphi(g)\gamma \quad \overset{\text{def}}{=} \quad r(g_i\gamma) \\
= \quad r(i\gamma) \\
= \quad \gamma
\]
and $G_{i\gamma} \subseteq G_{\gamma}$. Now suppose $\varphi(g)\gamma = \gamma$. Then $r(g_i\gamma) = \gamma$ by definition, and $g_i\gamma \cong i\gamma$ in $X$, which says $g \cdot i\gamma = i\gamma$ by Lemma 8.14(i) and $G_{\gamma} \subseteq G_{i\gamma}$.

(iv) We claim that choices can be made in the definitions of the associated complexes of groups $F^X$ and $F^\Gamma$ (see Definition 8.21) so that diagram (8.25) strictly commutes. First choose a skeleton $Q$ of the quotient $X/G$, define $F^X$ on object in the skeleton $Q$, and then extend to all objects in $X/G$. For every $\overline{q} \in \text{ob}(Q)$, select a $q \in \text{ob}(X)$ such that $p^X(q) = \overline{q}$ and define $F^X(\overline{q}) = G_q$. We remain with the choice of the selected preimage $q$ of $\overline{q}$ throughout. If $\overline{\sigma}, \overline{\tau} \in \text{ob}(X/G)$ and $\overline{\sigma}, \overline{\tau} \cong \overline{\sigma}$ is an isomorphism in $X/G$, then also define $F^X(\overline{\sigma}) = G_q$. This is allowed, since $\overline{\sigma}, \overline{\tau} \cong \overline{\sigma}$ implies existence of morphisms $a \colon q \to g_q\sigma$ and $b \colon \sigma \to g_qq$ in $X$, and the composite
\[
q \xrightarrow{a} g_q\sigma \xrightarrow{gb} g_qg_qq
\]
is trivial by Definition 8.14(i). The opposite composite is also trivial, as it is a loop, and we have $q \cong g_q\sigma$ in $X$. Then by Lemma 8.14(iii) $G_q = G_{g_q\sigma}$ and we may define $F^X(\overline{\sigma}) = G_q$ because $p^X(g_q\sigma) = \overline{\sigma}$. In particular, the selected preimage of $\overline{\tau}$ in $X$ is $g_q\sigma$ and we select $h_\tau = c_G$ for $\overline{\tau}, \overline{\tau} \cong \overline{\sigma}$ in Definition 8.21 so $F^\Gamma(\overline{\tau}) = \text{id}_{G_q}$. We remark that the isomorphism $\overline{\sigma}$ is the only morphism $\overline{\tau} \to \overline{\sigma}$ because there are no loops in $X/G$, so the element $g_q\sigma$ is uniquely defined as the target of the unique morphism $a$ with source $q$ and $p^X$-image $\overline{\sigma}$.

We next define $F^X$ on objects of $\Gamma/G$ using the equivalence $\overline{\tau}$ and the definition of $F^X$ on objects of $Q$. For $\overline{\tau} \in \text{ob}(Q)$, we also define $F^\Gamma(\overline{\tau}(\overline{\tau})) = G_q$. This is allowed: for $\overline{\tau}(\overline{\tau}) = r(\overline{\tau})$ we choose $r(q)$ as the selected preimage in $\text{ob}(\Gamma)$, and $r(q) \cong q$ in $X$, so $G_{r(q)} = G_{r(q)} = G_q$ by (iii) and Lemma 8.14(iii). Every $\overline{\tau} \in \text{ob}(\Gamma/G)$ is of the form $\overline{\tau}(\overline{\tau})$ for a unique $\overline{\tau} \in Q$, so $F^\Gamma$ is now defined on all objects of $\Gamma/G$, and we have $F^\Gamma \circ \tau = F^X$ on all objects of $X/G$.

We must now define $F^X$ and $F^\Gamma$ on morphisms so that $F^\Gamma \circ \tau = F^X$ for morphisms also. The idea is to first define $F^X$ on morphisms in the skeleton $Q$, then extend to all of $X/G$, and then define $F^\Gamma$ on morphisms of $\Gamma/G$. If $\overline{\tau} \colon \overline{\tau}_1 \to \overline{\tau}_2$ is a morphism in $Q$, then there is a unique morphism $a$ in $X$ with source $q_1$ and $p^X(a) = \overline{\tau}$. Select any $h_\tau$ such that $h_\tau a = q_2$. Then we define
an injective group homomorphism $F(\overline{\sigma}): G_{q_1} \to G_{q_2}$ by

$$F(\overline{\sigma})(g) := h_{\overline{\sigma}^{-1}} h_{\overline{\sigma}} g h_{\overline{\sigma}}^{-1}.$$ 

If $\overline{\tau}: \overline{\sigma}_1 \to \overline{\sigma}_2$ is any morphism in $\mathcal{X}/G$, then there exists a unique $\overline{\sigma}$ in $Q$ and a unique commutative diagram with vertical isomorphisms as below.

Then we choose $h_{\overline{\sigma}}$ to be $h_{\overline{\sigma}}$ and we consequently have $F(\overline{\sigma}) = F(\overline{\tau})$. If $\overline{\sigma}: \overline{\sigma}_1(\gamma) \to \overline{\sigma}_2(\gamma)$ is a morphism in $\Gamma/G$, then there is a unique $\overline{\sigma}: \overline{\sigma}_1 \to \overline{\sigma}_2$ in $Q$ with $\overline{\sigma}(\gamma) = \overline{\sigma}$ and we choose $h_{\overline{\sigma}}$ to be $h_{\overline{\sigma}}$. Manifestly, we have $F^\Gamma \circ \overline{\sigma} = F^\mathcal{X}$. The coherences of $F^\mathcal{X}$ and $F^\Gamma$ are also compatible, since they are determined by the $h_{\overline{\sigma}}$'s.

From (ii) we know $\overline{\sigma}$ is a surjective-on-objects equivalence of categories and from (iv) we have $F^\mathcal{X} = F^\Gamma \circ \overline{\sigma}$. From this, one sees

$$(\overline{\sigma}, \text{id}): \text{hocolim}_{\mathcal{X}/G} F^\mathcal{X} \to \text{hocolim}_{\Gamma/G} F^\Gamma$$

is an equivalence of categories.

If the action of $G$ on $\text{ob}(\mathcal{X})$ is free, then for each $\gamma \in \text{ob}(\Gamma)$, the group $G_\gamma = G_{\gamma}$ (see (iii)) is trivial, and $G$ acts freely on $\text{ob}(\Gamma)$.

**Remark 8.28.** In Theorem 8.24 it is even possible to select a skeleton so that the inclusion is $G$-equivariant, though we will not need this. See Section 9.

In [16] Theorems 5.30 and 5.37, we proved the compatibility of the $L^2$-Euler characteristic with coverings and isofibrations of finite connected groupoids. Theorem 8.29 is an analogue for scwols (see Remark 8.10).

**Theorem 8.29 (Compatibility with free actions on finite scwols).** Let $G$ be a finite group acting on a finite scwol $\mathcal{X}$. If $G$ acts freely on $\text{ob}(\mathcal{X})$, then

$$\chi(\mathcal{X}/G; R) = \frac{\chi(\mathcal{X}; R)}{|G|} \quad \text{and} \quad \chi^{(2)}(\mathcal{X}/G) = \frac{\chi^{(2)}(\mathcal{X})}{|G|}.$$ 

Recall $\chi(-; R)$ and $\chi^{(2)}(-)$ agree for finite scwols by Example 8.7.

**Proof.** By Theorem 8.24 (iii) and (vi) we may assume $\mathcal{X}$ is skeletal. A consequence of Definition 8.11 (ii) (independent of skeletonality) is that an element $g \in G$ fixes a path $a = (a_1, \ldots, a_n)$ in $\mathcal{X}$ if and only if $g$ fixes $sa_1$, so $G_{sa_1} = G_a$. Then $G$ acts freely on $\Lambda_a(\mathcal{X})$, since it acts freely on $\text{ob}(\mathcal{X})$. 

*Documenta Mathematica* 16 (2011) 301–354
The scwol $\mathcal{X}/G$ is skeletal by Lemma 8.19 and by Example 8.7 and Lemma 8.20 we have

$$\chi^2(\mathcal{X}/G) = \sum_{n \geq 0} (-1)^n c_n(\mathcal{X}/G)$$

$$= \sum_{n \geq 0} (-1)^n |\Lambda_n(\mathcal{X}/G)|$$

$$= \sum_{n \geq 0} (-1)^n |\Lambda_n(\mathcal{X})|$$

$$= \sum_{n \geq 0} (-1)^n \frac{|\Lambda_n(\mathcal{X})|}{|G|}$$

$$= \frac{1}{|G|} \sum_{n \geq 0} (-1)^n |\Lambda_n(\mathcal{X})|$$

$$= \frac{1}{|G|} \sum_{n \geq 0} (-1)^n c_n(\mathcal{X})$$

$$= \chi^2(\mathcal{X};\mathbb{C})$$

$$= \chi^2(\mathcal{X}/G;\mathbb{C})$$

A complex of groups is called developable if it is isomorphic to a complex of groups associated to a group action. A classical theorem of Bass–Serre says that every complex of groups on a scwol with maximal path length 1 is developable. The following gives a necessary condition of developability of a complex of groups from a scwol and group of specified Euler characteristics.

**Theorem 8.30 (Euler characteristics of associated complexes of groups).** Let $G$ be a finite group that acts on a finite scwol $\mathcal{X}$ in the sense of Definition 8.11. Let $F: \mathcal{X}/G \to \text{GROUPS}$ be the associated complex of groups. Then

$$\chi^2(\text{hocolim} \mathcal{X}/G F) = \chi^2(\mathcal{X}) \frac{\chi(\mathcal{X};\mathbb{C})}{|G|} = \chi^2(\mathcal{X}/G;\mathbb{C}) \frac{\chi(\mathcal{X}/G;\mathbb{C})}{|G|}.$$

If $R$ is a ring such that the orders of subgroups $H \subset G$ are invertible in $R$, then we also have

$$\chi(\text{hocolim} \mathcal{X}/G F; R) = \chi(\mathcal{X}/G; R).$$

**Proof.** By Theorem 8.24 [iii] [iv] and [v] we may assume $\mathcal{X}$ is skeletal. Then $\mathcal{X}/G$ is also skeletal by Lemma 8.19. Let $\Lambda_n(\mathcal{X})$ respectively $\Lambda_n(\mathcal{X}/G)$ denote the set of paths of $n$-many non-identity composable morphisms in $\mathcal{X}$ respectively $\mathcal{X}/G$. Then by Lemma 8.20 the sets $\Lambda_n(\mathcal{X}/G)$ and $\Lambda_n(\mathcal{X}/G)$ are in bijective correspondence.

We will also use the fact that an element $g \in G$ fixes a path $a = (a_1, \ldots, a_n)$ in $\mathcal{X}$ if and only if $g$ fixes $sa_1$, so $G_{sa_1} = G_a$. This is a consequence of Definition 8.11 [iii].
By Theorem 8.5, $E^\text{bar}_X$ and $E^\text{bar}_X(\mathcal X/G)$ are finite models for the skeletal scwols $\mathcal X$ and $\mathcal X/G$, and the $n$-cells are indexed by $\Lambda_n(\mathcal X)$ and $\Lambda_n(\mathcal X/G)$, respectively. For each path $(a_1, \ldots, a_n)$ in $\mathcal X$, there is an $n$-cell in $E^\text{bar}_X$ based at $s a_1$. A similar statement holds for $\mathcal X/G$ and $E^\text{bar}_X(\mathcal X/G)$.

Now we may apply the Homotopy Colimit Formula to the associated complex of groups $F: \mathcal X/G \to \text{GROUPS}$ by Remark 8.23. For the Euler characteristic, we have

$$
\chi(\text{hocolim}_{\mathcal X/G} F; R) = \sum_{n \geq 0} (-1)^n \left( \sum_{\pi \in \Lambda_n(\mathcal X/G)} \chi(F(s \pi_1); R) \right)
$$

$$
= \sum_{n \geq 0} (-1)^n \left( \sum_{\pi \in \Lambda_n(\mathcal X/G)} 1 \right)
$$

$$
= \sum_{n \geq 0} (-1)^n |\Lambda_n(\mathcal X/G)|
$$

$$
= \sum_{n \geq 0} (-1)^n c_n(\mathcal X/G)
$$

$$
= \chi(\mathcal X; R).
$$

For the $L^2$-Euler characteristic on the other hand, we have

$$
\chi^{(2)}(\text{hocolim}_{\mathcal X/G} F) = \sum_{n \geq 0} (-1)^n \left( \sum_{\pi \in \Lambda_n(\mathcal X/G)} \chi^{(2)}(F(s \pi_1)) \right)
$$

$$
= \sum_{n \geq 0} (-1)^n \left( \sum_{\pi \in \Lambda_n(\mathcal X/G)} \frac{1}{|G_{s a_1}|} \right)
$$

$$
= \sum_{n \geq 0} (-1)^n \left( \sum_{\pi \in \Lambda_n(\mathcal X/G)} \frac{1}{|G_{a_1}|} \right)
$$

$$
= \sum_{n \geq 0} (-1)^n \cdot \left( \sum_{\pi \in \Lambda_n(\mathcal X/G)} \frac{|\text{orbit}(a)|}{|G|} \right)
$$

$$
= \frac{1}{|G|} \sum_{n \geq 0} (-1)^n \left( \sum_{\pi \in \Lambda_n(\mathcal X/G)} |\text{orbit}(a)| \right)
$$

$$
= \frac{1}{|G|} \sum_{n \geq 0} (-1)^n |\Lambda_n(\mathcal X)|
$$

$$
= \frac{1}{|G|} \sum_{n \geq 0} (-1)^n c_n(\mathcal X)
$$

$$
= \frac{\chi^{(2)}(\mathcal X)}{|G|}.
$$
Example 8.31. By the classical theorem of Bass–Serre, any injective group homomorphism
\begin{equation}
G_0 \to G_1
\end{equation}
is a developable complex of groups. The $L^2$-Euler characteristic of the homotopy colimit of $G_0 \to G_1$ by Example 5.1. Theorem 8.30 then says we must have
\[
\frac{|G|}{|G_1|} = \chi^{(2)}(X) = \chi(BX; \mathbb{C})
\]
if $G_0 \to G_1$ is to be developable from a scwol $X$ by an action of $G$ in the sense of Definition 8.11. Thus $G_0 \to G_1$ is not developable from any scwol $X$ whose geometric realization has Euler characteristic 0, such as $\{j \to k\}$. Nor can $G_0 \to G_1$ be developed from any scwol $X$ with $\chi(BX; \mathbb{C})$ negative. The integer $|G|$ must also be divisible by $|G_1|$, since $\chi(BX; \mathbb{C})$ is always an integer. Moreover, the Euler characteristic of $X$ must be less than or equal to $|G|$. This trivial example illustrates how one can find necessary conditions on $X$ and $G$ if a given complex of groups is to be developable from $X$ and $G$.

Example 8.33 (Euler characteristics of transport groupoid in finite case). Let $X$ be a finite set and $G$ a finite group acting on $X$. Let $R$ be a ring such that the orders of subgroups of $G$ are invertible in $R$. Considering $X$ as a scwol, we clearly have an action in the sense of Definition 8.11. The associated complex of groups $F : X/G \to \text{GROUPS}$ assigns to orbit($\sigma$) the stabilizer $G_\sigma$. The homotopy colimit $\text{hocolim}_{X/G} F$ is equivalent to the transport groupoid $G^G(X)$ of Example 5.6, so
\[
\chi(G^G(X); R) = \chi(\text{hocolim}_{X/G} F; R) = \chi(X/G; R) = |X/G|.
\]
For the $L^2$-Euler characteristic, on the other hand, we have
\[
\chi^{(2)}(G^G(X)) = \chi^{(2)}(\text{hocolim}_{X/G} F) = \frac{\chi^{(2)}(X)}{|G|} = \frac{|X|}{|G|},
\]
a formula obtained by Baez–Dolan [4].

We also generalize the following formula of Haefliger for the Euler characteristic of the homotopy colimit of a (not necessarily developable) complex of groups.

Theorem 8.34 (Corollary 3.5.3 of Haefliger [19]). Let $G(X)$ be a complex of groups over a finite ordered simplicial cell complex $X$. Assume that each $G_\sigma$ is the fundamental group of a finite aspherical cell complex. Then $BG(X)$ has the homotopy type of a finite complex and its Euler-Poincaré characteristic is

\[
\chi(BG(X)) = \chi(\text{hocolim}_{X/G} F) = \frac{\chi^{(2)}(X)}{|G|} = \frac{|X|}{|G|}.
\]
given by

\[ \chi(BG(X)) = \sum_{\sigma \in \text{ob}(C(X))} (1 - \chi(Lk^\sigma))\chi(G_\sigma). \]

The terms in Haefliger’s theorem have the following meanings. An ordered simplicial cell complex \( X \) is by definition the nerve of a skeletal scwol, denoted \( C(X) \). The notation \( BG(X) \) signifies the geometric realization of the nerve of the homotopy colimit of the pseudo functor \( G(X) : C(X) \to \text{GROUPS} \). An aspherical cell complex is one for which all homotopy groups beyond the fundamental group vanish. The lower link \( Lk^\sigma \) of the object \( \sigma \) is the full subcategory of the scwol \( \downarrow C(X) \) on all objects except \( 1_\sigma \).

**Theorem 8.35 (Extension of Corollary 3.5.3 of Haefliger [19]).** Let \( I \) be a finite skeletal scwol and \( F : I \to \text{GROUPS} \) a complex of groups such that for each object \( i \) of \( I \), the group \( F(i) \) is of type \( \text{FF}_Z \). Then

\[ \chi(Bhocolim_I F) = \sum_{i \in \text{ob}(I)} (1 - \chi(BLk^i))\chi(BF(i)), \]

where \( B \) indicates geometric realization composed with the nerve functor.

**Proof.** All hypotheses of Theorem 4.1(vii) are satisfied. The skeletal scwol \( I \) is directly finite by Lemma 8.4 and admits a finite \( I \)-CW-model for its \( I \)-classifying space by Theorem 8.5. Each group \( C(i) \) is automatically directly finite, and assumed to be of type \( \text{FF}_Z \). The bar construction model \( E^\text{bar}I \) in Remark 2.10 has an \( n \)-cell based at \( i \) for each path of \( n \)-many non-identity morphisms in \( I \)

\[ i \to i_1 \to i_2 \to \cdots \to i_n. \]

Each such path in \( I \) corresponds uniquely to a path of \((n-1)\)-many non-identity morphisms in the scwol \( Lk^i \) beginning at the object \( i \to i_1 \). Thus

\[
1 - \chi(BLk^i) = 1 - \sum_{m \geq 0} (-1)^m \text{card}\{m + 1\text{-paths in } I \text{ beginning at } i\}
\]

\[
= 1 - \sum_{m \geq 0} (-1)^m \text{card}\{(m + 1)\text{-paths in } I \text{ beginning at } i\}
\]

\[
= 1 - \sum_{n \geq 1} (-1)^{n-1} \text{card}\{n\text{-paths in } I \text{ beginning at } i\}
\]

\[
= 1 + \sum_{n \geq 1} (-1)^n \text{card}\{n\text{-paths in } I \text{ beginning at } i\}
\]

\[
= \sum_{n \geq 0} (-1)^n \text{card}\{n\text{-paths in } I \text{ beginning at } i\}.
\]
Then by Theorem 4.1 (i) Theorem 4.1 (iv) Theorem 1.15 and Theorem 4.1 (vi) we have

\[
\chi(B \hocolim F) = \chi(\hocolim F) = \sum_{n \geq 0} (-1)^n \cdot \sum_{\lambda \in \Lambda_n} \chi(F(i_\lambda)) = \sum_{i \in \ob(I)} \left(1 - \chi(BLk^i)\right) \cdot \chi(F(i)).
\]

\[\square\]

Remark 8.36. The assumptions in our Theorem 8.35 on the groups \(F(i)\) are related to the assumptions in Theorem 8.34 on the groups \(G_\sigma\) in that any finitely presentable group of type \((\mathbb{F} \mathbb{F}_{\mathbb{Z}})\) admits a finite model for its classifying space.

9. Appendix

Let \(G\) be a group acting on a scwol \(\mathcal{X}\) in the sense of Definition 8.11 In connection with Theorem 8.24 we remark that it is possible to choose a skeleton \(\Gamma_0\) of \(\mathcal{X}\), a \(G\)-equivariant functor \(r: \mathcal{X} \to \Gamma_0\), and a natural isomorphism \(\eta: ir \cong \id \mathcal{X}\) so that

- the inclusion \(i_0: \Gamma_0 \to \mathcal{X}\) is \(G\)-equivariant,
- \(ri_0 = \id \Gamma_0\), and
- for every object \(x \in \ob(\mathcal{X})\) and each \(g \in G\), we have \(\eta_{gx} = g \eta_x\).

To prove this, we first choose the object set of \(\Gamma_0\) via an equivariant section of the projection \(\pi: \ob(\mathcal{X}) \to \iso(\mathcal{X})\), which assigns to each object of \(\mathcal{X}\) its isomorphism class of objects. Let \(\Theta\) denote the set of \(G\)-orbits of \(\iso(\mathcal{X})\). For each \(G\)-orbit \(\theta \in \Theta\), we use the axiom of choice to select an element \(x_\theta \in \theta\). For each \(\theta\), select then a \(\pi\)-preimage \(s(\pi(\theta)) := x_\theta\) of \(\theta\). On the orbit of each \(x_\theta\) we define the section \(s\) by \(s(gx_\theta) := gx_\theta\). This is well defined, for if \(g_1x_\theta = g_2x_\theta\), then \(g_1x_\theta \cong g_2x_\theta\), and \(g_1x_\theta = g_2x_\theta\) by Lemma 8.14 (i). Define the skeleton \(\Gamma_0\) to be the full subcategory of \(\mathcal{X}\) on the objects in the image of the equivariant section \(s: \iso(\mathcal{X}) \to \ob(\mathcal{X})\).

For each \(x_\theta\), and each \(x \in x_\theta\), choose an isomorphism \(\eta_x: x_\theta \to x\). For \(gx\), we define \(\eta_{gx}\) as \(g \eta_x\). Next, we define a functor \(r: \mathcal{X} \to \Gamma_0\) on objects \(x \in \ob(\mathcal{X})\) by \(r(x) := s \pi(x)\) and on morphisms \(f\) by \(r(f) := \eta_y \circ f \circ \eta_x^{-1}\). Then \(r\) is clearly a natural isomorphism, the inclusion \(i_0: \Gamma_0 \to \mathcal{X}\) is \(G\)-equivariant, and \(ri_0 = \id \Gamma_0\).

References

[1] M. Aigner. *Combinatorial theory*. Springer-Verlag, Berlin, 1979.
[2] S. Allouch. Sur l’existence d’une catégorie ayant une matrice strictement positive donnée. arXiv:0806.0060v1 [math.CT], 2008.
[3] S. Allouch. On the existence of a category with a given matrix. arXiv:1007.2884v1 [math.CT], 2010.

[4] J. Baez and J. Dolan. From finite sets to Feynman diagrams. In Mathematics Unlimited - 2001 and Beyond, vol. 1, eds. Björn Engquist and Wilfried Schmid, Springer, Berlin, 2001, pp. 29-50.

[5] A. Bartels and W. Lück. The Borel conjecture for hyperbolic and CAT(0)-groups. Preprintreihe SFB 478 — Geometrische Strukturen in der Mathematik, Münster, arXiv:0901.0442v1 [math.GT], 2009.

[6] A. Bartels, W. Lück, and H. Reich. The $K$-theoretic Farrell-Jones conjecture for hyperbolic groups. Invent. Math., 172(1):29–70, 2008.

[7] A. Bartels, W. Lück, and H. Reich. On the Farrell-Jones Conjecture and its applications. Journal of Topology, 1:57–86, 2008.

[8] P. Baum, A. Connes, and N. Higson. Classifying space for proper actions and $K$-theory of group $C^*$-algebras. In $C^*$-algebras: 1943–1993 (San Antonio, TX, 1993), pages 240–291. Amer. Math. Soc., Providence, RI, 1994.

[9] C. Berger and T. Leinster. The Euler characteristic of a category as the sum of a divergent series. Homology, Homotopy Appl., 10(1):41–51, 2008.

[10] A. K. Bousfield and D. M. Kan. Homotopy limits, completions and localizations. Springer-Verlag, Berlin, 1972. Lecture Notes in Mathematics, Vol. 304.

[11] M. R. Bridson and A. Haefliger. Metric spaces of non-positive curvature. Springer-Verlag, Berlin, 1999. Die Grundlehren der mathematischen Wissenschaften, Band 319.

[12] K. S. Brown. Cohomology of groups, volume 87 of Graduate Texts in Mathematics. Springer-Verlag, New York, 1982.

[13] J. Cheeger and M. Gromov. $L_2$-cohomology and group cohomology. Topology, 25(2):189–215, 1986.

[14] J. F. Davis and W. Lück. Spaces over a category and assembly maps in isomorphism conjectures in $K$- and $L$-theory. K-Theory, 15(3):201–252, 1998.

[15] E. Dror Farjoun. Homotopy and homology of diagrams of spaces. Algebraic topology, Proc. Workshop, Seattle/Wash. 1985, Lect. Notes Math. 1286, 93-134 (1987).

[16] T. M. Fiore, W. Lück, and R. Sauer. Finiteness obstructions and Euler characteristics of categories. arXiv, 2010. Accepted to Advances in Mathematics.

[17] A. Grothendieck. Pursuing Stacks. 1983.

[18] A. Haefliger. Complexes of groups and orbihedra. Group theory from a geometrical viewpoint (Trieste, 1990), 504–540, World Sci. Publ., River Edge, NJ, 1991.

[19] A. Haefliger. Extension of complexes of groups. Ann. Inst. Fourier (Grenoble), 42(1-2):275–311, 1992.

[20] L. Illusie. Complexes cotangent et déformations. II. Springer-Verlag, Berlin-New York, 1972.
[21] T. Leinster. The Euler characteristic of a category. *Doc. Math.*, 13:21–49, 2008.
[22] W. Lück. Equivariant Eilenberg-Mac Lane spaces $K(G, \mu, 1)$ for possibly non-connected or empty fixed point sets. *Manuscripta Math.*, 58:67–75, 1987.
[23] W. Lück. *Transformation groups and algebraic K-theory*. Springer-Verlag, Berlin, 1989. Mathematische Göttingensis.
[24] W. Lück. Dimension theory of arbitrary modules over finite von Neumann algebras and $L^2$-Betti numbers. I. Foundations. *J. Reine Angew. Math.*, 495:135–162, 1998.
[25] W. Lück. Dimension theory of arbitrary modules over finite von Neumann algebras and $L^2$-Betti numbers. II. Applications to Gromov-Hausdorff groups, $L^2$-Euler characteristics and Burnside groups. *J. Reine Angew. Math.*, 496:213–236, 1998.
[26] W. Lück. The universal functorial Lefschetz invariant. *Fund. Math.*, 161(1-2):167–215, 1999. Algebraic topology (Kazimierz Dolny, 1997).
[27] W. Lück. $L^2$-invariants: theory and applications to geometry and K-theory, volume 44 of Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]. Springer-Verlag, Berlin, 2002.
[28] W. Lück. $L^2$-invariants from the algebraic point of view. Preprintreihe SFB 478 — Geometrische Strukturen in der Mathematik, Heft 295, Münster, arXiv:math.GT/0310489 v1, 2003.
[29] W. Lück. Survey on classifying spaces for families of subgroups. In *Infinite groups: geometric, combinatorial and dynamical aspects*, volume 248 of *Progr. Math.*, pages 269–322. Birkhäuser, Basel, 2005.
[30] W. Lück and D. Meintrup. On the universal space for group actions with compact isotropy. In *Geometry and topology: Aarhus (1998)*, pages 293–305. Amer. Math. Soc., Providence, RI, 2000.
[31] W. Lück and H. Reich. The Baum-Connes and the Farrell-Jones conjectures in K- and L-theory. In *Handbook of K-theory. Vol. 1, 2*, pages 703–842. Springer, Berlin, 2005.
[32] G. Maltsiniotis. La théorie de l’homotopie de Grothendieck. *Astérisque* 301, 2005.
[33] D. Meintrup and T. Schick. A model for the universal space for proper actions of a hyperbolic group. *New York J. Math.*, 8:1–7 (electronic), 2002.
[34] R. J. Piacenza. Homotopy theory of diagrams and CW-complexes over a category. *Can. J. Math.*, 43(4):814–824, 1991.
[35] M. D. Rabrenović. Classifying spaces of monoids—applications in geometric combinatorics. *Publ. Inst. Math. (Beograd) (N.S.)*, 77(91):21–27, 2005.
[36] J.-P. Serre. *Linear representations of finite groups*. Springer-Verlag, New York, 1977. Translated from the second French edition by Leonard L. Scott, *Graduate Texts in Mathematics*, Vol. 42.
[37] J.-P. Serre. *Trees*. Springer-Verlag, Berlin, 1980. Translated from the French by J. Stillwell.
[38] P. Selick *Introduction to homotopy theory* AMS, Providence, 1997 *Fields Institute Monographs*.
[39] N. E. Steenrod. A convenient category of topological spaces. *Michigan Math. J.*, 14:133–152, 1967.
[40] R. W. Thomason. Homotopy colimits in the category of small categories. *Math. Proc. Camb. Philos. Soc.* 85:91-109, 1979.
[41] R. W. Thomason. Cat as a closed model category. *Cahiers Topologie Géom. Différentielle* 21(3):305-324, 1980.
[42] T. tom Dieck. *Transformation groups and representation theory*. Springer-Verlag, Berlin, 1979.
[43] T. tom Dieck. *Transformation groups*. volume 8 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1987.
[44] C. T. C. Wall. Finiteness conditions for CW-complexes. *Annals of Math. (2)*, 81:56–69, 1965.
[45] C. T. C. Wall. Finiteness conditions for CW-complexes. II. *Proc. Roy. Soc. Ser. A*, 295:129–139, 1966.

Thomas M. Fiore
Department of Mathematics and Statistics
University of Michigan-Dearborn
4901 Evergreen Road
Dearborn, MI 48128
U.S.A.
tmflore@umd.umich.edu
[http://www-personal.umd.umich.edu/~tmfiore/]

Wolfgang Lück
Mathematisches Institut der Universität Bonn
Endenicher Allee 60
53115 Bonn
Germany
wolfgang.lueck@him.uni-bonn.de
[http://www.him.uni-bonn.de/lueck]

Roman Sauer
Fakultät für Mathematik
Universität Regensburg
Universitätsstr. 31
93053 Regensburg
Germany
roman.sauer@mathematik.uni-regensburg.de
[http://www.mathematik.uni-regensburg.de/sauer/]

Documenta Mathematica 16 (2011) 301–354
Surfaces de Stein Associées
aux Surfaces de Kato Intermédiaires

Laurent Battisti ∗

Received: August 2, 2010
Revised: November 5, 2010
Communicated by Thomas Peternell

Abstract.
Soient $S$ une surface de Kato intermédiaire, $D$ le diviseur formé des courbes rationnelles de $S$, $	ilde{S}$ le revêtement universel de $S$ et $\tilde{D}$ la préimage de $D$ dans $\tilde{S}$. On donne deux résultats concernant la surface $\tilde{S} \setminus \tilde{D}$, à savoir qu'elle est de Stein (ce qui était connu dans le cas où $S$ est une surface d'Enoki ou d'Inoue-Hirzebruch) et on donne une condition nécessaire et suffisante pour que son fibré tangent holomorphe soit holomorphiquement trivialisable.

2010 Mathematics Subject Classification: 32E10, 32E30, 32J15, 32M25.
Keywords and Phrases: Variétés de Stein, surfaces de Kato.

1. Introduction
Les surfaces de la classe VII de Kodaira sont les surfaces complexes compactes dont le premier nombre de Betti vaut 1 ; on appelle surface de la classe VII₀ une surface de la classe VII qui est minimale. Le cas de ces surfaces dont le second nombre de Betti $b₂$ est nul est entièrement compris, il s’agit nécessairement d’une surface de Hopf ou d’une surface d’Inoue et le cas $b₂ > 0$ est toujours étudié actuellement ; il a été conjecturé qu’elles contiennent toutes une coquille sphérique globale. La preuve de ce résultat terminerait la classification des surfaces complexes compactes.
Les surfaces à coquille sphérique globale, qui nous intéressent ici, peuvent être obtenues selon un procédé dû à Kato (voir [11]), que l’on rappelle dans la section suivante. Ces surfaces se divisent en trois classes, les surfaces d’Enoki, d’Inoue-Hirzebruch et enfin les surfaces intermédiaires.

∗. Je tiens à remercier le referee pour ses remarques très importantes. Le financement de cette recherche est assuré par la Région Provence-Alpes-Côte d’Azur dans le cadre d’une bourse doctorale régionale.

Documenta Mathematica 16 (2011) 355–371
Étant donnés une surface minimale $S$ à coquille sphérique globale, $D$ le diviseur maximal de $S$ formé des $b_2(S)$ courbes rationnelles de $S$ et $\varpi : \tilde{S} \to S$ le revêtement universel de $S$, nous allons démontrer que $\tilde{S} \setminus \tilde{D}$ est une variété de Stein. Ce résultat était déjà connu pour les surfaces d’Enoki et d’Inoue-Hirzebruch ; nous allons le montrer dans le cas des surfaces intermédiaires. Dans la dernière partie et toujours dans le cas des surfaces intermédiaires, on donne une condition pour que le fibré tangent holomorphe de la variété $\tilde{S} \setminus \tilde{D}$ soit holomorphiquement trivialisable, à savoir que la surface $S$ soit d’indice 1.

2. Préliminaires

On dit qu’une surface compacte $S$ contient une coquille sphérique globale s’il existe une application qui envoie holomorphiquement un voisinage de la sphère $S^3 \subset \mathbb{C}^2 \setminus \{0\}$ dans $S$ et telle que le complémentaire dans $S$ de l’image de la sphère par cette application soit connexe.

Toute surface contenant une coquille sphérique globale peut être obtenue de la façon suivante : étant données une succession finie d’éclatements $\pi_1, \ldots, \pi_n$ de la boule unité $B$ de l’espace de dimension 4 au-dessus de 0 et $\pi := \pi_1 \circ \cdots \circ \pi_n : B^\pi \to B$ la composée de ces éclatements, ainsi qu’une application $\sigma : \overline{B} \to B^\pi$ holomorphe sur un voisinage de $\overline{B}$, on recolle les deux bords de $\text{Ann}(\pi, \sigma) := B^\pi \setminus \sigma(B)$ à l’aide de l’application $\sigma \circ \pi$.

La surface obtenue possède un groupe fondamental isomorphe à $\mathbb{Z}$ et son second nombre de Betti est égal à $n$ (voir [2]). Il s’agit d’une construction due à Kato [11]. Dans la suite, on appellera surface de Kato une surface complexe compacte minimale contenant une coquille sphérique globale, dont le second nombre de Betti est non nul.
Dans [2], Dloussky étudie le germe contractant d’application holomorphe \( \varphi = \pi \circ \sigma : B \to B \) associé à la construction précédente. Ce germe détermine à isomorphisme près la surface étudiée (proposition 3.16 loc. cit.).

Soit \( S \) une surface de Kato ; on note \( D \) le diviseur maximal de \( S \) formé des \( b_2(S) \) courbes de \( S \), \( \tilde{S} \) le revêtement universel de \( S \) et \( \hat{D} \) la préimage de \( D \) dans \( \tilde{S} \).

Suivant les notations de [2], on obtient la surface \( \tilde{S} \) en recollant une infinité d’anneaux \( A_i \) \((i \in \mathbb{Z})\) isomorphes à \( \text{Ann}(\pi, \sigma) \), en identifiant le bord pseudo-concave de \( A_i \) au bord pseudo-convexe de \( A_{i+1} \) via l’application \( \sigma \circ \pi \). La surface \( \tilde{S} \) possède deux bouts, notés \( 0 \) et \( \infty \), le bout \( 0 \) possédant une base de voisinages ouverts strictement pseudo-convexes (les \( \bigcup A_i \) pour \( j \in \mathbb{Z} \)) et le second une base de voisinages strictement pseudo-concaves (les \( \bigcup A_{i+1} \) pour \( j \in \mathbb{Z} \)). Enfin on définit un automorphisme \( G \) de \( \tilde{S} \) en posant \( G(z_i) = z_i + 1 \) où \( z_i \) et \( z_{i+1} \) sont les images dans \( A_i \) et \( A_{i+1} \) respectivement d’un même point \( z \in \text{Ann}(\pi, \sigma) \).

闭 Cylinder

\[ S \]

\[ \pi \]

\[ G(z_i) \]

Fixons une courbe compacte \( C \) de \( \tilde{S} \) avec \( C \subset A_0 \). On note \( (\tilde{S}_C, p_C) \) l’effondrement de \( \tilde{S} \) sur la courbe \( C \), c’est-à-dire la donnée d’une surface \( \tilde{S}_C \) n’ayant qu’un bout, d’une application holomorphe \( p_C \) de \( \tilde{S} \) dans \( \tilde{S}_C \), biholomorphe sur un voisinage du bout \( \infty \) dans \( \tilde{S} \) sur un voisinage du bout de \( \tilde{S}_C \), telles que \( \tilde{C} = p_C(C) \) soit une courbe d’auto-intersection \(-1\).

La proposition 3.4 de [2] nous assure l’existence d’une telle application \( p_C \) pour toute courbe compacte \( C \) de \( \tilde{S} \), et d’un point \( \hat{0}_C \in \tilde{C} \) tel que \( p_C \) soit également biholomorphe entre \( \tilde{S} \setminus p_C^{-1}(0_C) \) et \( \tilde{S}_C \setminus \{0_C\} \).

De plus, la restriction de \( p_C \) au complémentaire de \( \hat{D} \) est un biholomorphisme entre \( \tilde{S} \setminus \hat{D} \) et \( \tilde{S}_C \setminus p_C(D) \). Enfin, il existe une application holomorphe \( F_C \) de \( \tilde{S}_C \setminus \{0_C\} \) dans lui-même, contractante en \( \hat{0}_C \), conjuguée à \( \varphi \) et biholomorphe sur \( \tilde{S}_C \setminus p_C(D) \).
3. LA VARIÉTÉ $\tilde{S} \setminus \tilde{D}$ EST DE STEIN

Les surfaces de Kato se divisent en trois classes : les surfaces d’Enoki, d’Inoue-Hirzebruch et enfin les surfaces intermédiaires (voir [5]).

Dans le cas des surfaces d’Inoue-Hirzebruch et celles d’Enoki, le fait que $\tilde{S} \setminus \tilde{D}$ soit de Stein est déjà connu : pour une surface d’Inoue-Hirzebruch, la variété $\tilde{S} \setminus \tilde{D}$ est un domaine de Reinhardt holomorphiquement convexe (voir [13], proposition 2.2) tandis que pour une surface d’Enoki, on a $\tilde{S} \setminus \tilde{D} \cong \mathbb{C}^* \times \mathbb{C}$ qui sont bien dans chaque cas des variétés de Stein. Il reste donc à étudier le cas des surfaces intermédiaires.

Favre a donné dans [7] des formes normales pour les germes contractants d’applications holomorphes et on peut en particulier donner la forme du germe associé à une surface intermédiaire, à savoir qu’une telle surface est associée au gerne $\varphi$ de $(\mathbb{C}^2, 0) \rightarrow (\mathbb{C}^2, 0)$ donné par

$$(z, \zeta) \mapsto (\lambda^k z + P(\zeta) + c_{s_k}, \zeta^k)$$

où $\lambda \in \mathbb{C}^*$, $k, s \in \mathbb{N}$ avec $k > 1$ et $s > 0$, et $P(\zeta) = c_j \zeta^j + ... + c_s \zeta^s$ avec les conditions suivantes : $0 < j < k$, $j \leq s$, $c_j = 1$, $c_{s_k} = 0$ quand $\frac{m}{k} \notin \mathbb{Z}$ ou $\lambda \neq 1$ et enfin $\gcd\{k, m \mid c_m \neq 0\} = 1$. On trouve dans [12] une condition pour que deux tels germes soient conjugués (et déterminent donc deux surfaces isomorphes).

L’objectif de cette section est de démontrer, dans le cas de surfaces intermédiaires, le

**Théorème 3.1.** La surface $\tilde{S} \setminus \tilde{D}$ est de Stein.

Dans un premier temps (section 3.1), on montre qu’il est suffisant de se ramener à la situation du théorème 3.2 énoncé ci-dessous. Pour cela, nous allons écrire notre surface comme réunion croissante d’ouverts et nous verrons que seule une hypothèse manque a priori pour pouvoir effectivement appliquer ce théorème, à savoir que chaque paire constituée de deux tels ouverts consécutifs est de Runge. C’est dans la section 3.2 qu’on prouve que cette hypothèse est bien vérifiée.

3.1. Réduction du problème. Reprenons les notations précédentes et donnons-nous un germe de la forme (1). On regarde la surface intermédiaire $S$ associée et on choisit une courbe $C$ de $\tilde{S}$ donnée par la proposition 3.16 de [2] : quitte à numérotter les $A_i$ on suppose que $C \subset A_0$. Notre objectif est de prouver que la variété $\tilde{S} \setminus p_C(\tilde{D})$ est de Stein, en utilisant le théorème suivant (voir [10], théorème 10 p. 215) :

**Théorème 3.2.** Soient $X$ un espace analytique complexe et $(X_i)_{i \in \mathbb{N}}$ une suite croissante de sous-espaces de $X$ qui soit de Stein. Supposons que $X = \bigcup X_i$ et que chaque paire $(X_{i+1}, X_i)$ est de Runge, i.e. l’ensemble $\mathcal{O}(X_i)_{X_{i+1}}$ des
restrictions à $X_i$ des applications holomorphes sur $X_{i+1}$ est dense dans $\mathcal{O}(X_i)$. Alors $X$ est de Stein.

Notons :
- $\hat{A}_i := pc(A_i)$ pour tout $i \in \mathbb{Z}$ et
- $A_i := pc(\bigcup_{j \geq i} A_j)$ pour $i \leq 0$,
de sorte qu'on a $A_i \subset A_{i-1}$ et $\hat{S}_C \setminus pc(\vec{D}) = \bigcup_{i \leq 0} A_i \setminus pc(\vec{D})$.

Chaque $A_i \setminus pc(\vec{D})$ est strictement pseudo-convexe, donc de Stein. De plus, on a
$$ F_C(\hat{A}_i) = A_{i+1} $$ pour $i \leq -1$, car le diagramme
\[
\begin{array}{ccc}
\hat{S} & \xrightarrow{G} & \check{S} \\
\downarrow pc & & \downarrow pc \\
\hat{S}_C & \xrightarrow{F_C} & \check{S}_C
\end{array}
\]
est commutatif (c.f. [2], proposition 3.9). Ainsi, on a
\begin{equation}
F_C(A_{i-1} \setminus pc(\vec{D})) = A_{i} \setminus pc(\vec{D})
\end{equation}

Supposons établi le fait que la paire $(A_0 \setminus pc(\vec{D}), F_C(A_0 \setminus pc(\vec{D})))$ est de Runge. Alors la paire $(A_{i-1} \setminus pc(\vec{D}), A_{i} \setminus pc(\vec{D}))$ est automatiquement de Runge par l'égalité (2) ci-dessus, et par récurrence chaque paire $(A_{i-1} \setminus pc(\vec{D}), A_{i} \setminus pc(\vec{D}))$ est de Runge. Nous sommes alors en mesure d’appliquer le théorème 3.2 qui nous dit que la réunion des $A_i \setminus pc(\vec{D})$ est de Stein.

Le problème est donc ramené à montrer que le couple $(A_0 \setminus pc(\vec{D}), F_C(A_0 \setminus pc(\vec{D})))$ est de Runge.

**Remarque 3.3.** L’ensemble $A_0 \setminus pc(\vec{D})$ est biholomorphe à une boule ouverte centrée en 0 privée d’une droite complexe. En effet, on peut écrire $\varphi = \pi \circ \sigma$ où $\pi$ est une succession d’égalements de la boule au-dessus de 0 qui est de forme $C^2$, $\sigma : \overline{B} \to \pi^{-1}(B)$ est une application définie sur un voisinage de $B$ et biholomorphe sur son image, et $\varphi$ est de la forme normale (1). Par le choix de la courbe $C$, la proposition 3.16 p. 33 de [2] nous donne l’isomorphisme $A_0 \setminus pc(\vec{D}) \cong B \setminus \varphi^{-1}(0)$ et en utilisant la forme de $\varphi$, on voit que $\varphi^{-1}(0) = \{\zeta = 0\}$.

Finalement, démontrer que $(A_0 \setminus pc(\vec{D}), F_C(A_0 \setminus pc(\vec{D})))$ est de Runge revient à prouver que c’est le cas de la paire $(B \setminus \{\zeta = 0\}, \varphi(B \setminus \{\zeta = 0\}))$ pour une boule $B \subset \mathbb{C}^2$ centrée en 0 (en notant $(z, \zeta)$ les coordonnées de $\mathbb{C}^2$). C’est l’objet de la section suivante.

3.2. La paire $(B \setminus \{\zeta = 0\}, \varphi(B \setminus \{\zeta = 0\}))$ est de Runge. Etant donné un germe $\varphi$ de la forme (1), introduisons en premier lieu quelques notations :
1. Remarquons tout d’abord que chaque point de \( \mathbb{C} \times \Delta^* \) possède exactement \( k \) antécédents par \( \varphi \), où \( \Delta^* \) est le disque unité ouvert de \( \mathbb{C} \) privé de 0. Notons \( g \) l’automorphisme de \( \mathbb{C} \times \Delta^* \) suivant :
\[
g : (z, \zeta) \mapsto \left( e^{-s}z + \frac{P(\zeta) - P(e^s\zeta)}{\lambda e^s\zeta^s}, e^s\zeta \right)
\]
où \( e \) est une racine primitive \( k \)-ième de l’unité, de sorte que \( \varphi \circ g = \varphi \). Pour tout \( \ell \in \mathbb{Z} \), on a
\[
g^{\ell}(z, \zeta) = \left( (e^\ell)^{-s}z + \frac{P(\zeta) - P(e^\ell\zeta)}{\lambda(e^\ell)^s\zeta^s}, e^\ell\zeta \right)
\]
et \( \mathbb{Z} \cong \mathbb{Z}/k\mathbb{Z} \). L’automorphisme \( g \) permute les antécédents d’un même point de l’application \( \varphi \).

2. On notera également \( q(z, \zeta) \) le polynôme \( \prod_{\ell=1}^{k-1} a_\ell(z, \zeta)\zeta^{n_\ell} \) où \( a_\ell(z, \zeta) \) est la première composante de \( g^{\ell}(z, \zeta) \) et \( n_\ell = s - \min\{n|c_m(1 - (e^\ell)^n) \neq 0\} \), qui est bien défini et positif ou nul vu la dernière hypothèse sur les coefficients de \( P \), à savoir \( \text{pgcd}\{k, m | c_m \neq 0\} = 1 \). Le polynôme \( q(z, \zeta) \) est en particulier de la forme \( q(z, \zeta) = z(c + \epsilon(z, \zeta)) \) où \( \epsilon(z, \zeta) \xrightarrow{(z, \zeta)\to(0,0)} 0 \) et \( c \neq 0 \).

3. Pour \( \eta > 0 \), on note \( U_\eta \) l’ouvert \( \{(z, \zeta) \in \mathbb{C}^2 | |q(z, \zeta)| < \eta\} \). Soient \( a, b \) et \( c \) trois réels strictement positifs, on définit les ensembles
\[
K_{a,b} := \{(z, \zeta) \in \mathbb{C}^2 | |z|^2 + |\zeta|^2 \leq a^2, |\zeta| \geq b\} = B(0, a) \cap \{|\zeta| \geq b\}
\]
et
\[
L_{a,b,c} := \overline{B(0, a) \times \mathbb{A}_{b,c}}
\]
où \( \mathbb{A}_{b,c} \) est l’anneau ouvert centré en 0 de rayons \( b < c \). Enfin, on pose
\[
K_{a,b} := \bigcup_{\ell=0}^{k-1} g^{\ell}(K_{a,b})
\]
et
\[
L_{a,b,c} := \bigcup_{\ell=0}^{k-1} g^{\ell}(L_{a,b,c})
\]

**Remarque 3.4.** Pour \( a, b \) et \( c \) assez petits, les compacts \( g^j(L_{a,b,c}) \) (resp. \( g^j(K_{a,b}) \)) sont disjoints deux à deux : ceci est une conséquence du fait que la fonction \( \varphi \) est localement injective autour de l’origine de \( \mathbb{C}^2 \), ce qui est démontré, par exemple, dans [5], section 5. En particulier, les ensembles \( L_{a,b,c} \) et \( K_{a,b} \) possèdent chacun \( k \) composantes connexes.

D’autre part, on a \( K_{a',b'} \subset L_{a,b,c} \) pour \( b' \geq b \) et \( a' \leq \min\{a, c\} \), ce qui entraîne notamment \( K_{a',b'} \subset L_{a,b,c} \).

Enfin, pour \( \eta > 0 \) fixé, il existe \( A_\eta > 0 \) tel que pour tous réels \( t \) et \( \delta \) avec
0 < t < \delta < A_n$, on ait $L_{\delta,t,\delta} \subset U_\eta$ : en calculant $|q(z,\zeta)|$ pour $(z,\zeta) \in L_{\delta,t,\delta}$ on voit qu’il suffit de choisir $\delta$ assez petit pour avoir

$$|\delta| \prod_{\ell=1}^{k-1} (|\delta^{n\ell+1}| + 2(|c_{s-n\ell}| + |c_{s-n\ell+1}|\delta + ... + |c_s|\delta^n)/\lambda) < \eta.$$ 

On appelle $V_{\eta,\delta}$ l’ensemble $U_\eta \cap \{|\zeta| \leq \delta\}$.

**Proposition 3.5.** Pour $\delta > 0$ assez petit et pour tout $\varepsilon_1 \in [0,\delta[$, le compact $K_{\delta,\varepsilon_1}$ est holomorphiquement convexe.

**Preuve :** En premier lieu, remarquons que l’enveloppe holomorphiquement convexe de $V_{\eta,\delta}$ est l’adhérence $\overline{V}_{\eta,\delta}$ de cet ensemble. On note :

- $\hat{K}_{\delta,\varepsilon_1}$ l’enveloppe holomorphiquement convexe de $K_{\delta,\varepsilon_1}$,
- $\hat{K}_{\delta,\varepsilon_1}^\ell$ (resp. $\hat{V}_{\eta,\delta}^\ell$) la composante connexe de $\hat{K}_{\delta,\varepsilon_1}$ (resp. $\hat{V}_{\eta,\delta}^\ell$) qui contient $g^\ell(K_{\delta,\varepsilon_1})$, pour $\ell \in \{0, ..., k-1\}$.

**Étape 1 :** Montrons tout d’abord que pour $\eta$ et $\delta$ assez petits et pour tout $\varepsilon_1 < \delta$, on a $\overline{V}_{\eta,\delta}^0 \cap K_{\delta,\varepsilon_1} = K_{\delta,\varepsilon_1}$, autrement dit que la composante connexe de $\overline{V}_{\eta,\delta}$ qui contient $K_{\delta,\varepsilon_1}$ ne rencontre aucune autre composante de $K_{\delta,\varepsilon_1}$.

Soient $\delta > \varepsilon_1 > 0$. Pour $\ell \in \{1, ..., k-1\}$, on a

$$g^\ell(K_{\delta,\varepsilon_1}) \subset g^\ell(L_{\delta,\varepsilon_1,\delta}) = \{(z,\zeta) \in \mathbb{C}^2 \mid |a_{k-\ell}(z,\zeta)| \leq \delta, |\zeta| \in [\varepsilon_1,\delta]\}.$$ 

En particulier, pour $(z,\zeta) \in g^\ell(L_{\delta,\varepsilon_1,\delta})$, on a

$$z = \frac{P(\varepsilon^{k-\ell}\zeta) - P(\zeta)}{\lambda^\varepsilon} + w$$ 

où $|w| \leq \delta$. En développant, cette égalité devient

$$z = \lambda^{-1}\zeta^{-n_s} \left((e^{-\ell})^{s-n_s} - 1\right) + c_{s-n_s+1}(e^{-\ell})^{s-n_s+1} - 1)\zeta + ... + c_s((e^{-\ell})^s - 1)\zeta^s + w.$$
Autrement dit, $z$ est de la forme
\[ \lambda^{-1} \zeta^{-n_\iota} ((c_{s-n_\iota}((\varepsilon^{-\iota})^{s-n_\iota} - 1) + \zeta R_\iota(\zeta, w)) \]
où $R_\iota$ est un polynôme et par définition de $n_\iota$, le terme $c_{s-n_\iota}((\varepsilon^{-\iota})^{s-n_\iota} - 1)$ est non nul.

Grâce à cette dernière expression de $z$, on voit que lorsque $n_\iota > 0$, pour n'importe quelle constante $C > 0$ et lorsque $\delta$ est assez petit, tout élément $(z, \zeta) \in g'(L_{\delta, \varepsilon_1, \delta})$ vérifie $|z| > C$.

Dans le cas où $n_\iota = 0$ (donc $c_s \neq 0$), on a $|z| = |\lambda^{-1}(c_s((\varepsilon^{-\iota})^s - 1)) + w|$ est supérieur à une constante non nulle pour $\delta$ assez petit.

Posons alors $\alpha := \frac{1}{2} \min \{|c_s((\varepsilon^{-\iota})^s - 1)| \mid \ell s \neq 0[k]\}$ si $c_s \neq 0$ et $\alpha := 1$ sinon.

Par ce qui précède, il existe une constante $A > 0$ telle que pour tous $\delta < A$ et $\varepsilon_1 < \delta$ on ait, pour chaque $\ell \in \{1, \ldots, k-1\}$ et tout élément $(z, \zeta)$ de $g'(L_{\delta, \varepsilon_1, \delta})$, l’inégalité
\[ (4) \quad |z| \geq \alpha. \]

Fixons désormais $\eta > 0$ vérifiant les deux conditions suivantes :
1. $2\eta/|c| < \alpha$ (où $c \neq 0$ est le facteur de $z$ dans le développement limité de $q$
   en $(0, 0)$, à savoir $q(z, \zeta) = z(c + \varepsilon(z, \zeta))$, et
2. $|c + \varepsilon(z, \zeta)| > |c|/2$ pour tout $(z, \zeta) \in D(0, 3\eta/|c|) \times \overline{D(0, 3\eta/|c|)}$.

Choisissons maintenant $\delta < \min\{A, A_\eta, 2\eta/|c|\}$ et $\varepsilon_1 \in ]0, \delta[$. Alors on a $L_{\delta, \varepsilon_1, \delta} \subset U_\eta$ (remarque 3.4) et l’inégalité (4) ci-dessus est vérifiée.

Pour tout $|z| \leq \delta$ et $|\zeta| \in [\varepsilon_1, \delta]$ on a $(z, \zeta) \in K_{\delta, \varepsilon_1} \subset \overline{V}_{\eta, \delta}$. Soit maintenant $\ell \in \{1, \ldots, k-1\}$ et $(z, \zeta)$ un point de $g'(K)$, on a $|z| \geq \alpha$ et ceci entraîne que $(z, \zeta) \notin \overline{V}_{\eta, \delta}$.

En effet, supposons le contraire : la projection de $\overline{V}_{\eta, \delta}$ sur la première coordonnée étant connexe, et comme $\delta < 2\eta/|c| < \alpha$, il devrait exister un élément $(z', \zeta') \in \overline{V}_{\eta, \delta}$ avec $|z'| = 2\eta/|c|$, ce qui est impossible puisque dans ce cas
\[ |q(z', \zeta')| > \frac{2\eta}{|c|}|c|/2 = \eta. \]

Ainsi, la composante connexe $\overline{V}_{\eta, \delta}$ de $\overline{V}_{\eta, \delta}$ qui contient $K_{\delta, \varepsilon_1}$ ne rencontre aucune autre composante de $K_{\delta, \varepsilon_1}$, ce qu'il fallait démontrer.

À partir de maintenant, on omet les indices $\delta, \varepsilon_1$.

ÉTAPE 2 : Montrons à présent que $\hat{K}^0 = K$. Par l’étape 1, et comme $\hat{K}^0 \subset \overline{V}^0$, on sait que $\hat{K}^0$ ne rencontre pas d’autre composante de $K$ que l’ensemble $K$
lui-même.

Soit $(z_0, \zeta_0) \in \hat{K}^0 \setminus K$. On suppose que $|\zeta_0| \geq \varepsilon_1$ (sinon $(z_0, \zeta_0) \notin \hat{K}$), donc
nécessairement $|z_0|^2 + |ζ_0|^2 > δ^2$. Comme la boule fermée $B := \overline{B(0, δ)}$ est holomorphiquement convexe dans $\mathbb{C}^2$, il existe une fonction $h$ holomorphe sur $\mathbb{C}^2$ telle que $|h(z_0, ζ_0)| > \|h\|_F$. Notons respectivement $m_0$ et $m_∥$ les quantités $|h(z_0, ζ_0)|$ et $\|h\|_F$, ainsi que $m_\hat{K}$ la quantité $\|h\|_\hat{K}$, qui est finie puisque $\hat{K}$ est compact.

Considérons la fonction $χ_{\hat{K}_0}$ définie sur $\hat{K}$ valant 1 sur $\hat{K}_0$ (en particulier sur $K$) et 0 sur $\hat{K} \setminus \hat{K}_0$ (en particulier sur $g^ℓ(K)$ pour $ℓ ≠ 0[k]$).

Le théorème 6’ p. 213 de [10] nous dit que la fonction $χ_{\hat{K}_0}$ est limite uniforme sur $\hat{K}$ de fonctions holomorphes sur $\mathbb{C} \times Δ^*$. Soit donc $f$ une fonction holomorphe vérifiant $\|f - χ_{\hat{K}_0}\|_F < ε'$ avec $ε' < \min \left\{ \frac{m_0}{m_\hat{K} + m_0}, \frac{m_0 - m_∥}{m_0 + m_∥} \right\}$ et appelons $F$ l’application $(z, ζ) ↦ h(z, ζ)f(z, ζ)$.

Pour $(z_ℓ, ζ_ℓ) ∈ g^ℓ(K)$ (avec $ℓ ∈ \{1, ..., k - 1\}$), on a l’inégalité

$$|F(z_ℓ, ζ_ℓ)| ≤ \|F - hχ_{\hat{K}_0}\|_F + |h(z_ℓ, ζ_ℓ)χ_{\hat{K}_0}(z_ℓ, ζ_ℓ)|.$$  

Le second terme du membre de droite est nul ; quant au premier, il est majoré par $m_\hat{K}ε'$. De plus, on a $|F(z_0, ζ_0)| = m_0|f(z_0, ζ_0)| > m_0(1 - ε')$ d’une part, et pour tout $(z, ζ) ∈ K$ on a

$$|F(z, ζ)| ≤ |h(z, ζ)\left(f(z, ζ) - χ_{\hat{K}_0}(z, ζ)\right)| + |h(z, ζ)χ_{\hat{K}_0}(z, ζ)|$$  

donc $|F(z, ζ)| ≤ m_∥ε' + m_∥$ d’autre part. Le choix de $ε'$ nous assure que $\max\{m_\hat{K}ε', m_∥(ε' + 1)\} < m_0(1 - ε')$. Autrement dit, nous avons montré que $(z_0, ζ_0) ∉ \hat{K}$, d’où une contradiction. Ainsi, on a bien établi que $\hat{K}_0 = K$.

ÉTAPE 3 : Il nous reste à conclure. Remarquons que l’enveloppe holomorphe convexe $\hat{K}$ de $K$ est également stable par $g$ et supposons qu’il existe $ℓ_0 ∈ \{1, ..., k - 1\}$ et un point $(z_{ℓ_0}, ζ_{ℓ_0}) ∈ \hat{K}_{ℓ_0} \setminus g^ℓ_0(K)$. Alors on a les inclusions suivantes :

$$K ⊂ g^{-ℓ_0}(\hat{K}^{ℓ_0}) ⊂ \hat{K}_0,$$

la dernière inclusion provenant du fait que la continuité de $g$ entraîne la connexité de $g^{-ℓ_0}(\hat{K}^{ℓ_0})$. On a donc $g^{-ℓ_0}(z_{ℓ_0}, ζ_{ℓ_0}) ∈ \hat{K}_0 = K$, d’où une contradiction.

Finalement, on a établi que $\bigcup_{ℓ=0}^{k-1} \hat{K}^ℓ = K$. Comme $K$ est une réunion de composantes connexes de $\hat{K}$, c’est un sous-ensemble ouvert et fermé de $\hat{K}$, donc holomorphiquement convexe par le corollaire 8 p. 214 de [10].

Notons $\mathcal{O}(\mathbb{C} × Δ^*)$ l’algèbre des fonctions holomorphes sur $\mathbb{C} × Δ^*$ et $φ^∗(\mathcal{O}(\mathbb{C} × Δ^*))$ l’algèbre des éléments de $\mathcal{O}(\mathbb{C} × Δ^*)$ invariants par le groupe $g_\mathbb{C}$. Si $A$ est

Documenta Mathematica 16 (2011) 355–371
une algèbre de fonctions holomorphes, on note $K^A$ l’enveloppe de $K$ par rapport à l’algèbre $A$. On a montré que $K^{O(C \times \Delta^*)} = K$.

**Corollaire 3.6.** On a $\hat{K}^{\varphi'(O(C \times \Delta^*))} = K$.

**Preuve :** En effet, pour $x \notin K$, on a :

$$\left( g^2, x \right) \cap \bar{K} = (g^2, x) \cup K.$$

Ceci découle du fait que si $p \notin K$, pour $q \notin \{p\} \cup K$, il existe $f_1 \in O(C \times \Delta^*)$ telle que $\|f_1\|_K < f_1(q)$. Après avoir éventuellement multiplié $f_1$ par une constante, on peut supposer que $f_1(q) = 1$. Comme $p \neq q$, il existe également une fonction $f_2 \in O(C \times \Delta^*)$ qui vérifie $f_2(p) = 0$, $f_2(q) \neq 0$ et $\|f_2\|_K \leq 1/2$ ; quitte à remplacer $f_1$ par des puissances d’elle-même, on peut supposer que $\|f_1\|_K \leq |f_2(q)|$ et dans ce cas on a $\|f_1 f_2\|_{K \cup \{p\}} < |f_1(q)f_2(q)|$. Ainsi, on a

$$\left\{ p \right\} \cup \bar{K} = \{ p \} \cup K ;$$

par conséquent, en ajoutant un nombre fini de points à $K$ l’ensemble obtenu reste holomorphiquement convexe, et on a bien l’égalité (5).

On considère alors la fonction $f$ qui vaut $1$ sur $g^2, x$ et $0$ sur $K$, qui est holomorphe sur $(g^2, x) \cup K$. Alors (théorème 6' p. 213 de [10]) il existe une fonction $h \in O(C \times \Delta^*)$ telle que $\|f - h\| < 1/2$.

En définissant la fonction holomorphe $H := \frac{1}{k} \sum_{j=0}^{k-1} (h \circ g^j)$, il sort que l’on a

$$|H(x) - 1| < 1/2$$

tandis que pour tout $y \in K$, on a $|H(y)| < 1/2$, donc $x \notin \hat{K}^{\varphi'(O(C \times \Delta^*))}$.

**Corollaire 3.7.** Soit $\delta$ un réel positif donné par la proposition 3.5. Alors, la paire

$$(B(0, \delta) \setminus \{ \zeta = 0 \}, \varphi(B(0, \delta) \setminus \{ \zeta = 0 \}))$$

est de Runge.

**Preuve :** On se donne un compact $A$ de $\varphi(B(0, \delta) \setminus \{ \zeta = 0 \})$, il est inclus dans un certain $\varphi(K_{\delta-1/p,1/q})$ (pour $p$ et $q$ assez grands et avec $\delta > 1/p + 1/q$). L’enveloppe de $A$ par rapport à l’algèbre des fonctions holomorphes sur $B(0, \delta) \setminus \{ \zeta = 0 \}$ est incluse dans $\varphi(K_{\delta-1/p,1/q})$ par le corollaire 3.6, donc compacte. Ainsi $\varphi(B(0, \delta) \setminus \{ \zeta = 0 \})$ est holomorphiquement convexe par rapport aux fonctions holomorphes de $B(0, \delta) \setminus \{ \zeta = 0 \}$, ce qui nous donne la conclusion ([10], corollaire 9 p. 214).

### 3.3. Une généralisation.

Soit $\varphi$ un germe de $(C^3, 0)$ dans $(C^3, 0)$ donné par

$$\begin{align*}
(z, \zeta, \xi) &\mapsto (\lambda \zeta \zeta^\ell \zeta + P(\zeta, \xi), \zeta^k, \xi^\ell)
\end{align*}$$

364 L. Battisti
Surfaces de Stein Associées aux Surfaces de Kato . . .

\[ \lambda \in \mathbb{C}^*, k, \ell, r, s \in \mathbb{N} \text{ avec } k, \ell > 1, \text{ pgcd}(k, \ell) = 1 \text{ et } r, s > 0, \text{ et} \]

\[ P(\zeta, \xi) = \sum_{i_1=1}^r \sum_{i_2=1}^s c_{i_1,i_2} \zeta^{i_1} \xi^{i_2} \]

avec les conditions suivantes : 

\[ 0 < j_1 < k, \ 0 < j_2 < \ell, \ j_1 \leq r, \ j_2 \leq s \text{ et } c_{j_1,j_2} \neq 0. \]

Nous ajoutons une hypothèse supplémentaire, à savoir que pour tout \( \varepsilon \in \mathbb{U}_k \) (racines \( k \)-ièmes de l’unité) et \( \tau \in \mathbb{U}_\ell \) avec \( \varepsilon \tau \neq 1 \), il existe des entiers \( n \) et \( m \) et un polynôme \( Q \) avec \( Q(0,0) \neq 0 \), tels que l’on ait l’égalité :

\[ P(\zeta, \xi) - P(\varepsilon \zeta, \tau \xi) = \zeta^n \xi^m Q(\zeta, \xi). \]

Donnons quelques classes d’exemples de polynômes vérifiant cette dernière condition :

1. \( P(\zeta, \xi) = \sum_{p=1}^{\min(r,s)} a_{p} \zeta^{p} \xi^{p} \text{ avec ou bien pgcd}\{k, p \mid a_p \neq 0\} = 1, \text{ ou bien} \]
\[ \text{pgcd}\{\ell, p \mid a_p \neq 0\} = 1, \]
2. \( P(\zeta, \xi) = \zeta^{s'} \sum_{p=1}^{r} a_{p} \zeta^{p} \text{ avec pgcd}\{\ell, p \mid a_p \neq 0\} = 1 \text{ et } 1 \leq s' \leq s, \]
3. \( P \) de la forme précédente, mais en intervertissant les rôles de \( \zeta \) et \( \xi \).

Etant données \( \varepsilon_k \) et \( \tau_\ell \) deux racines primitives \( k \)-ième et \( \ell \)-ième de l’unité respectivement, notons \( g \) l’automorphisme de \( \mathbb{C} \times (\Delta^*)^2 \) qui à \( (z, \zeta, \xi) \) associe

\[ \underbrace{(\varepsilon_k^{-r} \tau_\ell^{-s} z + P(\zeta, \xi) - P(\varepsilon_k \zeta, \tau_\ell \xi)}_{a_{k,\ell}(z, \zeta, \xi)}, \zeta \zeta^s, \tau_\ell \tau^s \xi^s \]

et \( X \) l’ensemble \( B(0,1) \setminus \{\zeta \xi = 0\} \). La condition (7) permet d’adapter le raisonnement de la preuve de la proposition 3.5 et de ses deux corollaires dans cette situation, en posant cette fois-ci

\[ q(z, \zeta, \xi) = z^{k-1} \prod_{i=1}^{k-1} \prod_{j=1}^{\ell-1} a_{k,\ell}(z, \zeta, \xi) \zeta^{a_k} \xi^{a_m}. \]

Ainsi la paire \( (X, \varphi(X)) \) est de Runge. On obtient alors une variété de Stein en recollant une infinité dénombrable de copies de \( X \setminus \varphi(X) \) grâce à l’application \( \varphi \). Il est possible de généraliser cette dernière construction en prenant un germe de \( (\mathbb{C}^{n+1}, 0) \) dans lui-même, défini cette fois par \( (z, \zeta_1, ..., \zeta_n) \mapsto (\lambda \zeta_1^{e_1} ... \zeta_n^{e_n} z + P(\zeta_1, ..., \zeta_n), \zeta_1^{k_1} ..., \zeta_n^{k_n}) \) avec des conditions directement analogues à celles données ci-dessus.

† Comme \( k \) et \( \ell \) sont premiers entre eux, ceci revient à dire que \( \varepsilon \) et \( \tau \) ne sont pas simultanément égaux à 1.

Documenta Mathematica 16 (2011) 355–371
4. Invariants

Revenons à présent à notre situation de départ. On note désormais $X$ la variété $\tilde{S} \setminus \tilde{D}$. Étant donné un groupe $G$, on appelle espace $K(G, 1)$ tout espace topologique connexe dont le groupe fondamental est isomorphe à $G$ et qui possède un revêtement universel contractile.

Exemple 4.1. Le cercle unité $S^1$ est un espace $K(\mathbb{Z}, 1)$.

Remarquons tout d’abord que la variété $X$ est un espace $K(\mathbb{Z}[\frac{1}{2}], 1)$. En effet, $\pi_1(X) \cong \mathbb{Z}[\frac{1}{2}]$ et son revêtement universel $\mathbb{C} \times \mathbb{H}$ (c.f. [3] et [8]) est contractile. Le théorème I de [6] (pp. 482-483) nous dit alors que les groupes de cohomologie de $X$ sont isomorphes à ceux du groupe $\mathbb{Z}[\frac{1}{2}]$, c’est-à-dire que pour tout $n \in \mathbb{N}$ et pour tout groupe $G$, on a un isomorphisme entre $H^n(X,G)$ et $H^n(\mathbb{Z}[\frac{1}{2}],G)$. De plus, on sait (loc. cit. pp. 488-489) que le groupe $H^2(\mathbb{Z}[\frac{1}{2}],G)$ est isomorphe au groupe des extensions centrales de $\mathbb{Z}[\frac{1}{2}]$ par $G$. Une extension centrale est la donnée d’une extension de groupe

$$0 \to G \to E \xrightarrow{p} \mathbb{Z}[\frac{1}{2}] \to 0$$

où $E$ est un groupe avec $i(G) \subset Z(E)$, le centre de $E$.

Nous sommes maintenant en mesure de prouver la

Proposition 4.2. Le groupe $H^2(X, \mathbb{C})$ est trivial.

Preuve : Par ce qui précède, il suffit de montrer qu’une extension centrale $E$ de $\mathbb{Z}[\frac{1}{2}]$ par $\mathbb{C}$ est nécessairement triviale, i.e. isomorphe au produit cartésien $\mathbb{C} \times \mathbb{Z}[\frac{1}{2}]$. Soit donc $E$ une telle extension :

$$0 \to \mathbb{C} \to E \xrightarrow{p} \mathbb{Z}[\frac{1}{2}] \to 0.$$ 

Montrons que $E$ est abélien. Soient $x, y \in E$ et $a \in \mathbb{N}$ tels que $p(x)$ et $p(y)$ appartiennent tous deux à $\frac{1}{k} \mathbb{Z} := \{ \frac{n}{k}, n \in \mathbb{Z} \}$ qui est un sous-groupe de $\mathbb{Z}[\frac{1}{2}]$ isomorphe à $\mathbb{Z}$.

L’extension $E$ induit une extension $F := p^{-1}(\frac{1}{k} \mathbb{Z})$ de $\frac{1}{k} \mathbb{Z}$ par $\mathbb{C}$, donc une extension de $\mathbb{Z}$ par $\mathbb{C}$ :

$$0 \to \mathbb{C} \xrightarrow{i} F \xrightarrow{p'} \mathbb{Z} \to 0.$$ 

Il existe une section $s : \mathbb{Z} \to F$ (on choisisit $s(1) \in p'^{-1}(1)$ et on pose $s(n) = ns(1)$ pour $n \in \mathbb{Z}$) donc $F$ est produit semi-direct de $\mathbb{Z}$ par $\mathbb{C}$, donné par $\sigma \in \text{Hom}(\mathbb{Z}, \text{Aut}(\mathbb{C}))$. L’extension $F$ étant elle aussi centrale, $\sigma \equiv 1$ est l’unique possibilité, i.e. $F$ est abélien (il est isomorphe à $\mathbb{C} \times \mathbb{Z}$) donc $x$ et $y$ commutent. Ainsi, $E$ est abélien.

Il existe des sections $s : \mathbb{Z}[\frac{1}{2}] \to E$. Pour construire l’une d’elles, fixons $x_0 \in p^{-1}(1)$. Comme $\mathbb{Z}[\frac{1}{2}] \cong E/i(\mathbb{C})$, il existe $x'_1 \in p^{-1}(1/k)$ tel que $kx'_1 = x_0 + i(w)$ avec $w \in \mathbb{C}$. On pose alors $x_1 := x'_1 - i(w/k)$ et on a $kx_1 = x_1$; on définit ainsi par récurrence les $x_i \in p^{-1}(1/k^i)$ vérifiant $kx_{i+1} = x_i$, et notre section est donnée par $s(n/k^a) = nx_a$ pour $n \in \mathbb{Z}$ et $a \in \mathbb{N}$. L’existence d’une telle section nous dit que $E$ est isomorphe au produit.
semi-direct $\mathbb{C} \times \mathbb{Z}[\frac{1}{k}]$ donné par $\sigma \in \text{Hom}(\mathbb{Z}[\frac{1}{k}], \text{Aut}(\mathbb{C}))$. Le groupe $E$ étant abélien, on a nécessairement $\sigma \equiv 1$, i.e. $E$ est isomorphe au produit $\mathbb{C} \times \mathbb{Z}[\frac{1}{k}]$. \[\blacksquare\]

Étant donnés une surface intermédiaire $S$ et son germe associé sous la forme normale (1), on définit l’indice de $S$ comme le plus petit entier $m$ tel que $k-1$ divise $ms$ (voir [12]).

Il existe un feuilletage holomorphe $\mathcal{F}$ sur $X$ défini par la 1-forme holomorphe $d\zeta$, qui ne s’annule nulle part (c.f. [4]). De façon équivalente, les feuilles de ce feuilletage sont les ensembles $\{\zeta = \text{const.}\}$. Dans le cas où $S$ est d’indice 1, i.e. lorsque $k-1$ divise $s$, il existe un champ de vecteurs tangent à ce feuilletage qui ne s’annule nulle part, autrement dit on a le

**Lemme 4.3.** Lorsque la surface $S$ est d’indice 1, le fibré tangent au feuilletage $ TF $ est holomorphiquement trivialisable.

**Preuve :** Pour prouver cela, nous montrons qu’il suffit de considérer le champ de vecteurs $V$ sur $X$ induit par le champ de vecteurs $\widetilde{V} = \zeta^{\frac{n-1}{k-1}} \zeta^{\frac{i\pi}{k}}$ sur $\mathbb{C} \times \Delta^*$, tangent au feuilletage de $\mathbb{C} \times \Delta^*$ défini par $\omega$.

En effet, d’une part on remarque que $X$ est le quotient de $\mathbb{C} \times \Delta^*$ par $G$ où $G \cong \mathbb{Z}[\frac{1}{k}]/\mathbb{Z}$ est le groupe formé des automorphismes de $\mathbb{C} \times \Delta^*$ de la forme

$$ g_{k,n}(z,\zeta) = (ze^{-\frac{i\pi}{k}} + \sum_{i=0}^{n-1} e^{\frac{i\pi}{k(n-1)}}, \zeta), $$

pour $n \in \mathbb{N}, \ell \in \{0, \ldots, k^n - 1\}$ et avec $e_{k,n} = e^{\frac{i\pi}{k}}$. Ceci provient du fait que $X$ est le quotient de $\mathbb{C} \times \mathbb{H}_g$ par le groupe $\{\gamma^n \gamma^{-n} | n, \ell \in \mathbb{Z} \} \cong \mathbb{Z}[\frac{1}{k}]$ où $\mathbb{H}_g = \{w \in \mathbb{C} | \Re(w) < 0\}$, $\gamma(z, w) = (\lambda z e^{\bar{w}} + P(e^w), kw)$ et $\gamma_1(z, w) = (z, w + 2i\pi)$ (voir [4], proposition 2.3 et section 4). On considère alors le quotient par le sous-groupe $\{\gamma^n \gamma^{-n} | n, \ell \in \mathbb{Z} \} \cong \mathbb{Z}$ ce qui nous donne bien $X = (\mathbb{C} \times \Delta^*)/G$.

D’autre part, un champ de vecteurs $\widetilde{V}$ défini sur $\mathbb{C} \times \Delta^*$ induit un champ de vecteurs tangent à $X$ lorsqu’il est invariant par le groupe $G$, i.e. s’il vérifie :

$$ D(g_{k,n})_{z,\zeta}(\widetilde{V}(z,\zeta)) = \widetilde{V}(g_{k,n}(z,\zeta)). $$

Cette condition est bien vérifiée par $\widetilde{V}$ puisque l’on a l’égalité $e^{-\frac{i\pi}{k-1}} \zeta^{\frac{i\pi}{k}} = (e_{k,n}^{\frac{i\pi}{k}})^{\frac{i\pi}{k}}$. Comme $\omega(\widetilde{V}) = 0$, on a bien montré que $V \in H^0(X, TF)$. \[\blacksquare\]

**Remarque 4.4.** On peut montrer qu’un champ de vecteurs sur $X$ de la forme $f(z,\zeta) \frac{d}{dz}$ (où $f$ est une holomorphe ne s’annulant nulle part) existe bien si et seulement si la surface $S$ est d’indice 1, autrement dit on a une équivalence
dans le lemme précédent. C'est une conséquence de la condition (8) et le raisonnement est analogue à celui qui sera fait dans le lemme 4.7.

La trivialité du fibré $TF$ entraîne celle du fibré tangent $TX$, ce que nous voyons à présent.

**Lemme 4.5.** Lorsque le fibré $TF$ est holomorphiquement trivialisable, le fibré tangent holomorphe $TX$ de $X$ l’est aussi.

**Preuve :** Étant donné que nous avons une section holomorphe globale $V$ de $TF$, il nous suffit d’exhiber un deuxième champ de vecteurs global, linéairement indépendant de $V$ en chaque point. Par définition, on peut trouver un recouvrement de $X$ par des ouverts $U_i$ et sur chacun d’eux un champ de vecteurs $W_i$ qui soit linéairement indépendant de $V$ sur $U_i$. Quitte à remplacer $W_i$ par $W_i/\omega(W_i)$ on peut supposer que $\omega(W_i) \equiv 1$ sur $U_i$, de sorte que $\omega(W_{i,j}) = 0$ sur $U_{i,j} := U_i \cap U_j$, où l’on a posé $W_{i,j} := W_i - W_j$. La famille $(W_{i,j})$ forme donc un cocyle de $H^1(X, TF)$ qui est aussi un cobord par le théorème B de Cartan. Ainsi il existe un champ de vecteurs $Z_i$ sur chaque $U_i$ tel que $Z_i - Z_j = W_{i,j}$. Posons $Y_i := W_i - Z_i$, de sorte que $Y_i$ sur $U_{i,j}$, i.e. les $Y_i$ se recollent en une section holomorphe globale de $TX$. Nous avons deux champs de vecteurs $V$ et $Y$ vérifiant $\omega(V) \equiv 0$ et $\omega(Y) \equiv 1$, ce qui nous assure qu’ils sont linéairement indépendants en chaque point de la variété étudiée. 

Nous voulons à présent établir un lien entre le fait que $S$ soit d’indice 1 et la trivialité du fibré canonique de $X$.

On considère la suite exacte courte $0 \to Z \to C \to C^* \to 0$ de faisceaux, qui donne lieu à la suite exacte longue de cohomologie $\cdots \to H^1(X, Z) \to H^1(X, C) \to H^1(X, C^*) \to H^2(X, Z) \to H^2(X, C) \to \cdots$. Toujours d’après le théorème I de [6], on a les isomorphismes

$$H^1(X, Z) \cong \text{Hom}(\mathbb{Z}[\frac{1}{2}], \mathbb{Z}) = 0,$$

$$H^1(X, C^*) \cong \text{Hom}(\mathbb{Z}[\frac{1}{2}], C^*) = \mathbb{C}^*$$

e

et

$$H^1(X, C) \cong \text{Hom}(\mathbb{Z}[\frac{1}{2}], C) \cong \mathbb{C}.$$

D’autre part le groupe $H^2(X, C)$ est trivial, d’où l’on tire finalement la suite exacte courte

$$0 \to H^1(X, C) \xrightarrow{e^{2\pi i \rho}} H^1(X, C^*) \xrightarrow{c} H^2(X, Z) \to 0.$$

**Définition 4.6.** On dira qu’un élément $\rho$ de $H^1(X, C^*)$ admet un logarithme lorsqu’il existe un morphisme $\rho'$ de $\mathbb{Z}[\frac{1}{2}]$ dans $\mathbb{C}$ tel que $e^{2\pi i \rho'} = \rho$.

Ainsi, l’image par $c$ d’un élément $\rho \in H^1(X, C^*)$ est triviale dans $H^2(X, Z)$ si et seulement si $\rho$ admet un logarithme $\rho'$.

**Lemme 4.7.** Si le fibré canonique de $X$ est holomorphiquement trivialisable, alors la surface $S$ est d’indice 1.
Preuve : Le fibré canonique de $X$ est le fibré des $2$-formes holomorphes sur $X$ ; raisonons par l’absurde et supposons qu’il est holomorphiquement trivialisable et que la surface $S$ n’est pas d’indice $1$. Alors il existe une $2$-forme holomorphe globale sur $X$ qui ne s’annule nulle part. Une telle forme provient d’une $2$-forme holomorphe sur le revêtement $\mathbb{C} \times \Delta^*$ de $X$ donnée par $f(z, \zeta)dz \wedge d\zeta$ ( où $f$ est une fonction holomorphe sur $\mathbb{C} \times \Delta^*$ qui ne s’annule nulle part) qui soit stable par le groupe $G = \{g^n_\alpha \mid n \in \mathbb{N}, \alpha \in \{0, ..., k^n - 1\}\}$, i.e. vérifie l’équation

$$ (g^n_\alpha)^* (f(z, \zeta)dz \wedge d\zeta) = f(z, \zeta)dz \wedge d\zeta $$

(pour tout $n \in \mathbb{N}$ et $\alpha \in \{0, ..., k^n - 1\}$). Ceci donne la condition suivante sur la fonction $f$ :

$$ e^{2\pi i \frac{\alpha}{k^n}} f(z, \zeta) = f(g^n_\alpha(z, \zeta)). $$

Considérons l’homomorphisme de groupes

$$ \rho : \frac{\mathbb{Z}[\frac{1}{k}]}{\mathbb{Z}} \longrightarrow S^1, $$

$$ \frac{n}{k} \longmapsto e^{2\pi i \frac{n}{k^n}} $$

Il induit un fibré plat $L_\rho$ au-dessus de $X$, qui est holomorphiquement trivialisable si et seulement si $\rho$ admet un logarithme, puisque $H^1(X, O^*) \cong H^2(X, \mathbb{Z})$ car $X$ est de Stein. Étant donné que la fonction $f$ vérifie la condition $(9)$ ci-dessus, elle définit une section holomorphe du fibré plat $L_\rho$ au-dessus de $X$. Ainsi pour pouvoir aboutir à une contradiction, il nous reste à voir que $\rho$ n’admet pas de logarithme (et donc qu’une telle fonction $f$ n’existe pas).

Remarquons tout d’abord que l’application $\sigma : \frac{\alpha}{s} \longmapsto e^{2\pi i \frac{\alpha}{s}}$ est un homomorphisme de $\mathbb{Z}[\frac{1}{s}]$ dans $S^1$ qui admet un logarithme. Ainsi, $\rho$ admet un logarithme si et seulement si l’homomorphisme $\varphi := \rho/\sigma : \frac{\alpha}{k^n} \longmapsto e^{2\pi i \frac{\alpha}{k^n}}$ admet un logarithme.

Soit $m$ l’indice de la surface $S$. Comme $k - 1$ n’est pas un diviseur de $s$, le noyau de $\varphi$ est précisément $m\mathbb{Z}[\frac{1}{s}]$ et cet homomorphisme n’admet donc pas de logarithme. En effet, si un tel morphisme $\rho'$ existait, sa restriction à $m\mathbb{Z}[\frac{1}{s}]$ serait un homomorphisme à valeurs dans $\mathbb{Z}$, nécessairement trivial. On aurait alors $m.\rho \left(\frac{1}{k^n}\right) = 0$, i.e. $\rho \left(\frac{1}{k^n}\right) = 0$ pour tout $n \in \mathbb{N}$. ■

Remarque 4.8. Le groupe $H^2(X, \mathbb{Z})$ n’est pas trivial ; il contient des éléments de torsion et des éléments qui ne sont pas d’ordre fini. Le morphisme $\rho$ de la preuve du lemme 4.7 fournit un exemple d’élément de torsion, puisqu’on peut voir que $\rho^{k-1}$ admet un logarithme. Pour ce qui est des éléments qui ne sont pas d’ordre fini, donnons-en un exemple. Considérons le groupe $\mathbb{Z}[\frac{1}{p}]$. On a un isomorphisme de groupes $\varphi : \mathbb{Z}[\frac{1}{p}] / \mathbb{Z}[\frac{1}{p}] \cong \mathbb{Z}[\frac{1}{p}] / Z$ et une injection $i$ de ce groupe (le $3$-groupe de Prüfer) dans $S^1$. Alors on peut voir que $\rho := i \circ \varphi$ n’est pas d’ordre fini dans $H^2(X, \mathbb{Z})$. Ainsi, le groupe $H^2(X, \mathbb{Z})$ possède des éléments d’ordre infini dont l’image est nulle dans $H^2(X, \mathbb{Q})$, ceci est conséquence du fait que le groupe $H_1(X, \mathbb{Z}) \cong \mathbb{Z}[\frac{1}{p}]$ n’est pas finiment engendré (voir [1], théorème 4 p. 144).
Remarque 4.9. On a en fait une équivalence dans le lemme précédent. Lorsque
la surface $S$ est d’indice 1, on considère la forme $\zeta^{-1}(\zeta + 1)dz \wedge d\zeta$, qui trivialise
le fibré canonique.

Les trois lemmes précédents ont en particulier comme conséquence la

**Proposition 4.10.** Soient $S$ une surface intermédiaire et $X = \tilde{S} \setminus \tilde{D}$. Les trois
assertions suivantes sont équivalentes :
1. La surface $S$ est d’indice 1,
2. Le fibré tangent au feuilletage $TF$ de $X$ est holomorphiquement trivialisable,
3. Le fibré tangent holomorphe $TX$ de $X$ est holomorphiquement trivialisable.

**Preuve :** Vu les lemmes 4.3 et 4.5, il suffit de montrer que la troisième
assertion entraîne la première. C’est une conséquence du lemme 4.7, car si $S$
n’est pas d’indice 1, le fibré canonique de $X$ n’est pas holomorphiquement
trivialisable. Dans ce cas, le fibré cotangent de $X$ et donc le fibré tangent $TX$
ne le sont pas non plus. ■

Remarque 4.11. Le problème suivant demeure non résolu actuellement
(voir [9]) : une variété de Stein de dimension $n$ dont le fibré tangent holomorphe
est holomorphiquement trivialisable est-elle nécessairement un domaine de Rie-
mann au-dessus de $\mathbb{C}^n$ ? Nous ne connaissons pas la réponse pour les surfaces
de Stein que l’on vient de considérer.

**RÉFÉRENCES**

[1] R. Bott and J. Heitsch, *A remark on the integral cohomology of $B\Gamma_q$*, Topology 11 (1972), 141–146.

[2] G. Dloussky, *Structure des surfaces de Kato*, Mémoires de la S.M.F. 2e série 14 (1984).

[3] G. Dloussky and K. Oeljeklaus, *Vector fields and foliations on compact surfaces of class VII$_0$*, Ann. Inst. Fourier 49 (1999), no. 5, 1503–1545.

[4] G. Dloussky, K. Oeljeklaus, and M. Toma, *Surfaces de la classe VII$_0$ admettant un champ de vecteurs, ii*, Commentarii Mathematici Helvetici 76 (2001), no. 4, 640–664.

[5] ——, *Class VII$_0$ surfaces with $b_2$ curves*, Tohoku Math. Journal 55 (2003), 283–309.

[6] S. Eilenberg and S. MacLane, *Relations between homology and homotopy groups of spaces*, Annals of Mathematics 46 (1945), no. 3, 480–509.

[7] Ch. Favre, *Classification of 2-dimensional contracting rigid germs*, Journal de Mathématiques Pures et Appliquées 79 (2000), 475–514.

[8] ——, *Dynamique des applications rationnelles*, Ph.D. thesis, Université Paris-Sud, 2002.

[9] F. Forstnerič, *Noncritical holomorphic functions on Stein manifolds*, Acta Mathematica 191 (2003), no. 2, 143–189.
[10] R. C. Gunning and H. Rossi, *Analytic functions of several complex variables*, Prentice-Hall, 1965.

[11] M. Kato, *Compact complex manifolds containing “global” spherical shells*, Proceedings of the International Symposium on Algebraic Geometry (Kyoto Univ., Kyoto, 1977), Kinokuniya Book Store, Tokyo (1978).

[12] K. Oeljeklaus and M. Toma, *Logarithmic moduli spaces for surfaces of class VII*, Mathematische Annalen 341 (2008), 323–345.

[13] D. Zaffran, *Serre problem and Inoue-Hirzebruch surfaces*, Mathematische Annalen 319 (2001), 395–420.

Laurent Battisti
LATP-UMR(CNRS) 6632
CMI-Université d’Aix-Marseille I
39, rue Joliot-Curie
F-13453 Marseille Cedex 13
France
battisti@ci.univ-mrs.fr
Asymptotic Behavior of Word Metrics on Coxeter Groups

G. A. Noskov

Abstract. We study the geometry of tessellation defined by the walls in the Moussong complex $M_W$ of a Coxeter group $W$. It is proved that geodesics in $M_W$ can be approximated by geodesic galleries of the tessellation. A formula for the translation length of an element of $W$ is given. We prove that the restriction of the word metric on the $W$ to any free abelian subgroup $A$ is Hausdorff equivalent to a regular norm on $A$.

2000 Mathematics Subject Classification: 20F55, 51F15
Keywords and Phrases: Coxeter group, Moussong complex, translation length

Introduction

For any Coxeter system $(W, S)$, Moussong constructed a certain piecewise Euclidean complex $M_W$ on which $W$ acts properly and cocompactly by isometries [Mou88]. This complex is complete, contractible, has nonpositive curvature and the Cayley graph $C_W$ of $W$ (with respect to $S$) is isomorphic to the 1-skeleton of $M_W$. A wall in $M$ is the fixed-point set of a reflection in $W$. It turns out that the walls are totally geodesic subspaces in $M_W$ and each wall divides $M_W$ into two path components. The set of all walls defines a wall tessellation of $M$. The set of all tiles (= chambers) of this tessellation together with an appropriate adjacency relation is isomorphic to the Cayley graph $C_W$. We shall prove that geodesics in $M_W$ can be uniformly approximated by geodesic galleries of the wall tessellation (= geodesic paths in $C_W$) (Theorem 3.3.2). This approximation result allows us to prove that for any “generic” element $w \in W$ of infinite order there is a conjugate $v$ which is straight i.e., $\ell(v^n) = n\ell(v)$ for all $n \in \mathbb{N}$,
where \( \ell(v) \) is the word length on \( W \) (Theorem 4.1.5). There is a constant 
\( c = c(W) \), such that for any \( w \in W \) of infinite order there is a conjugate \( v \) of \( w^e \), which is straight (Theorem 4.1.6). The restriction of the word metric on \( W \) to any free abelian subgroup \( A \) is Hausdorff equivalent to a regular norm on \( A \) (Theorem 4.3.2).

\section*{Acknowledgment}
I would like to thank Mathematisches Institut der Heinrich-Heine-Universität for the hospitality during 1999. Thanks to H. Abels, B. Brink, M. Davis, M. Gromov, F. Grunewald, R. Howlett and E. B. Vinberg for their help. This research was supported by a DFG grant Gr 627–11 grant and by SFB 343, SFB 701 of the DFG at the University of Bielefeld.

\section{Preliminaries on Moussong complexes}

To any Coxeter system \((W, S)\) one can canonically associate the Moussong complex \( \mathcal{M} = \mathcal{M}_W \), which is a piecewise Euclidean complex with \( W \) as the set of vertices. Their cells are Euclidean polyhedra, which are the convex hulls of sets, naturally bijected with the spherical cosets of \( W \). In particular, the 1-cells of \( \mathcal{M} \) are in bijective correspondence with the sets \( \{w, ws\} \), where \( w \in W \) and \( s \in S \). Hence the 1-skeleton of \( \mathcal{M} \) is nothing but a modified Cayley graph of \( W \) with respect to \( S \) (the modification consists in identifying an edge \( w \rightarrow ws \) with its inverse \( ws \rightarrow w \)). \( W \) acts cellularly and isometrically on \( \mathcal{M}_W \) and this induces the standard \( W \)-action on the Cayley graph of \( W \). In the next subsections we carry out in detail the construction of \( \mathcal{M}_W \) following the thesis of D. Krammer [Kra94].

\subsection{Coxeter groups}

A Coxeter system is a pair \((W, S)\) where \( W \) is a group and where \( S \) is a finite set of involutions in \( W \) such that \( W \) has the following presentation:

\[ \langle s : s \in S \mid (ss')^{m_{ss'}} = 1 \text{ when } m_{ss'} < \infty \rangle, \]

where \( m_{ss'} \in \{1, 2, 3, \ldots, \infty\} \) is the order of \( ss' \), and \( m_{ss'} = 1 \) if and only if \( s = s' \). We refer to \( W \) itself as a Coxeter group when the presentation is understood. The number of elements of \( S \) is called its rank. The Coxeter system \((W, S)\) is called spherical if \( W \) has finite order. A subgroup of \( W \) is called special if it is generated by a subset of \( S \). For each \( T \subseteq S \), \( W_T \) denotes the special subgroup generated by \( T \). Any conjugate of such a subgroup will be called parabolic. A remarkable feature of Coxeter systems is that for any subset \( T \subseteq S \) the pair \((W_T, T)\) is a Coxeter system in its right and moreover a presentation of \( W_T \) is defined by the numbers \( m_{tt'}, t, t' \in T \). If \((W_S, S)\) is a Coxeter system of finite rank then we write \( V_S \) for the real vector space with a basis of elements \( (e_s) \) for \( s \in S \). Put a symmetric bilinear form \( B \) on \( V_S \) by requiring:

\[ B(e_s, e_{s'}) = -\cos(\pi/m_{ss'}). \]
(This expression is interpreted to be $-1$ in case $m(s, s') = \infty$.) Evidently $B(e_s, e_s) = 1$, while $B(e_s, e_{s'}) \leq 0$ if $s \neq s'$. Since $e_s$ is non-isotropic, the subspace $H_s = e_s^\perp$ orthogonal to $e_s$ is complementary to the line $\mathbb{R} e_s$. Associated to $s \in S$ is an automorphism $a_s$ of $B$ acting as the reflection $v \mapsto v - 2(v, e_s)e_s$ in the hyperplane $e_s^\perp$. The result by Tits asserts that the correspondence $s \mapsto a_s$ extends to a faithful representation of $W$ as a group of automorphisms of the form $B$. \(\text{cf. [Bou], Ch.V, s.4).}\)

1.2 Trading Coxeter cells

The Coxeter group $W$ is finite if and only if the form $B(e_s, e_{s'})$ is positive definite. We call a set $J \subseteq S$ spherical if $W_J$ is finite or, equivalently, the restriction of the form $B$ to the subspace $V_J = \sum_{j \in J} \mathbb{R} e_j$ is positive definite. Let $J \subseteq S$ be spherical. Since $V_J$ is non-degenerate, there exists a unique basis $\{f_j : j \in J\}$ of $V_J$ dual to $\{e_j : j \in J\}$ with respect to $B$. A space $V_J$ that comes equipped with a positive definite inner product $B|V_J$ will be denoted by $E_J$ and called the Euclidean space associated to $J$. Define the Coxeter cell $X_J$ to be the convex hull of the $W_J$-orbit:

$$X_J = \text{Ch}(W_J x_J)$$

where

$$x_J = \sum_{j \in J} f_j \in E_J.$$}

For convenience we define $W_\emptyset = \{1\}$ and $X_\emptyset = \{0\}$—the origin of $E_J$. More generally, for any spherical $K$ and any $J \subseteq K$ we consider the faces of the polyhedron $X_K = \text{Ch}(W_K x_K)$ of the form

$$X_{J|K} = \text{Ch}(W_J x_K).$$

We do not exclude the case $J = \emptyset$, where $X_{\emptyset|K} = \{x_K\}$. We call the extremal points of the cell $X_J$ the vertices.

For spherical $J \subseteq S$, let $p_J : V_S \to E_J$ denote the orthogonal projection. It is well defined since the quadratic form on $E_J$ is non-degenerate.

**Lemma 1.2.1** ([Kra94], B.2.2.) *The dimension of the cell $X_J$ equals the cardinality of $J$. For spherical subsets $J \subseteq K$ of $S$ we have $p_J x_K = x_J$. Moreover, $p_J X_{J|K} : X_{J|K} \to X_J$ is a $W_J$–equivariant isometry of cells. The nonempty faces of $X_K$ are precisely those of the form $w X_{J|K}$ ($J \subseteq K, w \in W_K$). In particular, the vertex set of $X_J$ is precisely $W_J x_J$.***

**Example 1.2.2** 1) If $J = \{j\}$ then $f^J_j = e_j$ and $X_J = \text{Ch}(e_j, -e_j)$ is a line segment. 2) Let $J = \{s, s'\}$ be spherical, so $w = ss'$ has finite order $m_{ss'}$. Set $V_{s,s'} = \mathbb{R} e_s + \mathbb{R} e_{s'}$. The restriction of $B$ to $V_{s,s'}$ is positive definite and both $s$ and $s'$ act as orthogonal reflections in the lines $\mathbb{R} f_s, \mathbb{R} f_{s'}$ respectively. Since $B(e_s, e_{s'}) = -\cos(\pi/m_{ss'}) = \cos(\pi - (\pi/m_{ss'}))$, the angle between the
rays $\mathbb{R}^+e_s$ and $\mathbb{R}^+e_{s'}$ is equal to $\pi - (\pi/m_{ss'})$, forcing the angle between the reflecting lines $\mathbb{R}f_s, \mathbb{R}f_{s'}$ to be equal $\pi/m_{ss'}$. The vectors $f_s, f_{s'}$ are of the same length, lie in the cone $\mathbb{R}^+e_s + \mathbb{R}^+e_{s'}$; moreover, $f_s + f_{s'}$ is a bisectrix between the reflecting lines $\mathbb{R}f_s, \mathbb{R}f_{s'}$ hence the convex hull of the orbit $W_J(f_s + f_{s'})$ is a regular $2m_{ss'}$-gon.

![Figure 1: The cell $X_J$ for $J = \{s, s'\}$, $m_{ss'} = 3$.](image.png)

1.3 Gluing the Moussong complex

Now we build the Moussong complex of $W = W_S$ as follows. Take the union

$$U = \bigcup \{(w, X_J) : w \in W, J \subseteq S \text{ spherical}\}.$$ 

Introduce an equivalence relation $R$ on $U$, generated by the following gluing relations:

1. $(wu, x) \sim (w, u^{-1}x)$, whenever $w \in W, u \in W_J, x \in X_J$,

2. The cells $(w, X_K), (w, X_L)$ are glued along the face $(w, X_J), J = K \cap L$, which is embedded into each of them (by the map $p_J$) as $(w, X_{JK})$ and $(w, X_{JL})$ respectively.

The quotient space of $U$ modulo $R$ is called the Moussong complex of $W$ and is denoted by $\mathcal{M}_W$. The group $W$ acts on $U$ by $u(w, x) = (uw, x)$. This action respects the relation $R$ and hence induces a cellular action of $W$ on $\mathcal{M}_W$. With some abuse in notation we will denote the natural image of $(1, X_J)$ in $\mathcal{M}$ by $X_J$, so any cell in $\mathcal{M}$ is of the form $wX_J$ for some $w \in W, J \subseteq S$. We call $J$ the type of the cell $wX_J$. There is a distinguished vertex $x_0 = X_\emptyset$ in $\mathcal{M}$. Note that $x_J = x_0$ for any spherical $J$ (by condition (2)).

It can be shown that the inclusion maps of the cells are injective, see [Kra94]. The canonical metric in each cell allows to measure the lengths of finite polygonal paths in $\mathcal{M}$. The path metric $d$ on $\mathcal{M}$ is defined by setting the distance
between \(x, y \in \mathcal{M}\) to be the infimum of the lengths of polygonal paths joining \(x\) to \(y\).

We summarize the main properties of \(\mathcal{M}\) in the following theorem.

**Theorem 1.3.1** ([Kra94],[Mou88]) Relative to the path metric \(\mathcal{M}\) is a contractible, complete, proper CAT(0) space. The Coxeter group \(W\) acts on \(\mathcal{M}\) cellularly and this action is isometric, proper and cocompact. This action is simply transitive on the set of vertices \(\mathcal{M}^{(0)}\) of \(\mathcal{M}\), in particular \(\mathcal{M}^{(0)}\) coincides with \(Wx_0\).

For the convenience of the reader we repeat the relevant definitions. A geodesic, or geodesic segment, in a metric space \((X,d)\) is a subset isometric to a closed interval of real numbers. Similarly, a loop \(S^1 \to X\) is a closed geodesic if it is an isometric embedding. (Here \(S^1\) denotes the standard circle equipped with its arc metric, possibly rescaled so that its length can be arbitrary). We say that \(X\) is a geodesic metric space if any two points of \(X\) can be connected by a geodesic. We denote by \([x,y]\) any geodesic joining \(x\) and \(y\). We will always parameterize \([x,y]\) by \(t \mapsto p_t(0 \leq t \leq 1)\), where \(d(x,p_t) = td(x,y)\) for all \(t\). Given three points \(x,y,z\) in \(X\), the triangle inequality implies that there is a comparison triangle in the Euclidean plane \(\mathbb{R}^2\), whose vertices \(\overline{x}, \overline{y}, \overline{z}\) have the same pairwise distances as \(x,y,z\). Given a geodesic \([x,y]\) and a point \(p = p_t \in [x,y]\), there is a corresponding point \(\overline{p} = \overline{p_t}\) on the line segment \([\overline{x},\overline{y}]\) in \(\mathbb{R}^2\). A geodesic metric space \(X\) is called a CAT(0) space if for any \(x,y\) in \(X\) there is a geodesic \([x,y]\) with the following property: For all \(p \in [x,y]\) and all \(z \in X\), we have

\[
d(z,p) \leq d_{\mathbb{R}^2}(\overline{z}, \overline{p}),
\]

with \(\overline{z}\) and \(\overline{p}\) as above. Let \(X\) be a CAT(0) space. Then there is a unique geodesic segment joining each pair of points \(x,y \in X\) and this geodesic segment varies continuously with its endpoints. Every local geodesic in \(X\) is a geodesic. For the proof see [BH99], Chapter II.1, Prop. 1.4.

**Examples 1.3.2** If \(W\) is a finite Coxeter group of rank \(n\) then \(\mathcal{M}_W\) is isometric to an \(n\)-dimensional convex polyhedron. If, for example, \(W\) is the dihedral group of order \(2m\), then \(\mathcal{M}\) is a regular \(2m\)-gon with the usual \(W\)-action. If \(W\) is an affine Coxeter group of rank \(n\) then \(\mathcal{M}_W\) is a tessellation of the \(n-1\)-dimensional Euclidean space \(E\). This tessellation is dual to the tessellation, given by the structure of a Coxeter complex on \(E\). Let, for example, \(W\) be an affine Coxeter group generated by the reflections \(s_1, s_2, s_3\) in the sides of an equilateral triangle \(C\) in the Euclidean plane. Then \(\mathcal{M}_W\) is the tessellation of the plane by hexagons, dual to the tessellation consisting of the images of \(C\) under \(W\). If \(W\) is a product of \(n \geq 2\) copies of \(\mathbb{Z}/2\) (that is \(m_{ss'} = \infty\) for \(s \neq s'\)), then \(\mathcal{M}_W\) is an infinite \(n\)-regular tree with edges of length \(2\).

**Lemma 1.3.3** Any cell of a CAT(0) piecewise Euclidean complex \(X\) is isometrically embedded into \(X\). In view of uniqueness of geodesics this is equivalent to the convexity of a cell.
We have to show is that for any two points \( a, b \) of a cell \( C \) the Euclidean arc \( \alpha \) in \( C \) between them is a global geodesic. We may assume that \( C \) is of minimal dimension. For any two points \( x \) and \( y \) in the interior of \( \alpha \) the closed subarc \( \beta \subset \alpha \) between \( x \) and \( y \) lies in the interior of \( C \). Clearly there is an \( \epsilon > 0 \), such that for any cell \( C' \), having \( C \) as a face, the distance from \( \beta \) to the set \( \partial C' \) is \( \geq \epsilon \). Let us cover \( \beta \) by intervals of radius \( \epsilon/2 \). Each such an interval is geodesic. Indeed, a geodesic \( \gamma \) connecting the points of the interval can not cross \( \partial C' \), hence it lies in the union \( U \) of cells, having \( C \) as a face. For any cell \( C' \), having \( C \) as a face, \( \gamma \) can not cross \( \partial C' - C \) since it has to pass a distance at least \( \epsilon \). Hence it lies in only one such \( C' \) and thus coincides with the interval. It follows from the considerations above that \( \beta \) is a local geodesic, and therefore a global geodesic since \( X \) is \( \text{CAT}(0) \).

Now let \( \gamma \) be a path in \( X \) joining \( a \) to \( b \). For any positive \( \epsilon < d_C(a, b)/2 \) we may choose points \( x \) and \( y \) in the interior of \( \alpha \) such that \( d(a, x) = \epsilon = d(y, b) \). A path from \( x \) to \( y \) obtained by traveling along \( \alpha \) to \( a \) then along \( \gamma \) to \( b \) has length \( \text{length}(\gamma) + 2\epsilon \), while a geodesic from \( x \) to \( y \) has length \( d_C(a, b) - 2\epsilon \), so \( d_C(a, b) \leq \text{length}(\gamma) + 4\epsilon \). Since this is true for any sufficiently small \( \epsilon > 0 \), we conclude that \( d_C(a, b) \leq \text{length}(\gamma) \), and so \( \alpha \) is a geodesic from \( a \) to \( b \). \( \Box \)

1.4 The action of reflections on cells

We refer to the notation of §1.3.

**Lemma 1.4.1** An element \( w \in W \) leaves the cell \( uX_K \) invariant if and only if \( u^{-1}wu \in W_K \). In the latter case \( w \) acts on the \( X_K \)-coordinate of \( ux \in uX_K \) as the element \( u^{-1}wu \in W_K \).

**Proof.** Indeed, the cell \( uX_K \) is uniquely determined by its set of vertices \( uW_Kx_0 \) and it is \( w \)-invariant if and only if \( uW_Kx_0 \) is \( w \)-invariant under left translation. The latter happens if and only if \( wuW_K = uW_K \Leftrightarrow u^{-1}wu \in W_K \).

The second assertion follows from the equality \( w(ux) = u(u^{-1}wu) \). \( \Box \)

**Lemma 1.4.2** (An "overcell" of invariant cell is invariant too.) If \( C \subseteq C' \) are cells and \( wC = C' \) for some \( w \in W \), then \( wC' = C' \).

**Proof.** Writing \( C = uX_J \) with \( w \in W, J \subseteq S \) we can represent \( C' \) in the form \( C' = uX_K, J \subseteq K \). By Lemma 1.4.1 \( wC = C' \) implies \( u^{-1}wu \in W_J \) and thus \( u^{-1}wu \in W_K \). Again by the same lemma \( wC' = C' \). \( \Box \)

**Definition 1.4.3** Let \((W; S)\) be a Coxeter system. A reflection in \( W \) is an element that is conjugate in \( W \) to an element of \( S \).

**Lemma 1.4.4** For any cell \( C \) of \( \mathcal{M} \) and any reflection \( w \in W \) either \( C \cap wC \) is empty or else \( w \) acts as a reflection on \( C \).

**Proof.** Suppose that the cell \( C \cap wC \) is nonempty. Then it is invariant under the action of \( w \). Since it is a face of \( C \), by Lemma 1.4.2 we conclude that \( wC = C \). Now by Lemma 1.4.1 \( w \in W \) acts as a reflection on \( C \). \( \Box \)
1.5 Angles and geodesics in $\mathcal{M}$

The notion of angle in an arbitrary piecewise Euclidean complex can be defined in terms of the link distance, see e.g. [BB97]. Namely, let $X$ be a piecewise Euclidean complex, $x \in X$ and let $A$ be a Euclidean cell of $X$ containing $x$. The link $lk_x A$ of is the set of unit tangent vectors $\xi$ at $x$ such that a nontrivial line segment with the initial direction $\xi$ is contained in $A$. We define the link $lk_x X$ by $lk_x X = \bigcup_{A \ni x} lk_x A$, where the union is taken over all closed cells containing $x$.

Recall that the CAT(0)– condition for $X$ is equivalent to the following (see e.g. [BB97]):

1. $X$ is 1-connected and
2. The length of any geodesic loop in the link of any vertex of $X$ is greater or equal to $2\pi$.

A path $\alpha: [a,b] \to X$ is geodesic if it is an isometric embedding: $d(\alpha(s), \alpha(t)) = |s - t|$, for any $s,t \in [a,b]$. Similarly, a loop $\alpha: S^1 \to X$ is a closed geodesic if it is an isometric embedding. Here $S^1$ denotes the standard circle equipped with its arc metric (possibly rescaled so that its length can be arbitrary). Angles in $lk_x X$ induce a natural length metric $d_x$ on $lk_x S^1$, which turns $lk_x S^1$ into a piecewise spherical complex. For $\xi, \eta \in lk_x X$ define $\angle(\xi, \eta) = \min(d_x(\xi, \eta), \pi)$.

Any two segments $\sigma_1, \sigma_2$ in $X$ with the same endpoint $x$ have the natural projection image in $lk_x X$ and we define $\angle_x(\sigma_1, \sigma_2)$ to be the angle between these two projections.

We will use the following criterion of geodesicity:

**Lemma 1.5.1 ([BB97])** Let $X$ be a piecewise Euclidean CAT(0)-complex. If each of the segments $\sigma_1, \sigma_2$ is contained in a cell and $\sigma_1 \cap \sigma_2 = \{x\}$, where $x$ is an endpoint of each of the segments, then the union $\sigma_1 \cup \sigma_2$ is geodesic if and only if $\angle_x(\sigma_1, \sigma_2) = \pi$.

An $m$-chain from $x$ to $y$ is an $(m+1)$-tuple $T = (x_0, x_1, \ldots, x_m)$ of points in $X$ such that $x = x_0, y = x_m$ and each consecutive pair of points is contained in a cell. Every $m$-chain determines a polygonal path in $X$, given by the concatenation of the line segments $[x_i, x_{i+1}], i = 0, \ldots, m$. An $m$-taut chain from $x$ to $y$ is an $m$-chain such that

1. there is no triple of consecutive points contained in a cell and
2. (2) the union of two subsequent segments is geodesic in the union of cells, containing these segments.

(The union is equipped with its path metric). Note that if a chain is taut then only its first and last entries lie in the interior of a top dimensional simplex of $X$. The result of M. Bridson asserts that if $X$ is a piecewise Euclidean complex then $X$ with its path metric is a geodesic space and the geodesics are the paths determined by taut chains [BH99, Theorem. 7.21].

Documenta Mathematica 16 (2011) 373–398
2 Walls in the Moussong complex

The notion of wall in the Moussong complex (as well as in the Coxeter complex) can be defined as the fixed-point set of reflection from the underlying Coxeter group. On the other hand they can be defined as the equivalence classes of "midplanes" (which are the fixed-point sets of stabilizers of cells). Both points of view are useful. Note that in contrast to the situation with Coxeter complexes, the walls in the Moussong complex are not subcomplexes.

2.1 Midplanes and blocks in cells

Let \((W_J, J)\) be a finite Coxeter group and \(V_J\) the Euclidean vector space on which \(W_J\) acts. We summarize here the basic properties of a Coxeter complex of \(W = W_J\). For more about them see [Hum90] or [Bro96]. We define a reflection in \(W_J\) to be a conjugate of element of \(J\). The reflecting hyperplanes \(H_w\) of reflection \(w \in W_J\) cut \(V_J\) into polyhedral pieces, which turn out to be cones over simplices. In this way one obtains a simplicial complex \(C = C(W)\) which triangulates the unit sphere in \(V_J\). This is called the Coxeter complex associated with \(W_J\). The group \(W_J\) acts simplicially on \(C\) and this action is simply transitive on the set of maximal simplices (=chambers). Moreover the closure of any chamber \(C\) is a fundamental domain of the action of \(W\) on \(C\), i.e., each \(x \in V\) is conjugated under \(W\) to one and only one point in \(C\). Two chambers are adjacent if they have a common codimension one face. For any two adjacent chambers there is a unique reflection in \(W_J\) interchanging these two chambers.

A similar picture we have for the Coxeter cell \(X_J\). By a midplane in \(X_J\) we mean the intersection \(H_w \cap X_J\), where \(w \in W_J\) is a reflection and \(H_w\) its reflecting hyperplane. We denote this midplane by \(M(J, w)\). By equivariance we define the notion of a midplane in any cell of \(M_W\). Each midplane \(M\) defines a unique cell in \(M_W\), the cell of least dimension in \(M_W\) which contains \(M\), and we will denote this by \(C(M)\).

**Lemma 2.1.1** Every cell \(X_J\) contains an open neighborhood of the origin of \(V_J\). In particular midplanes in \(X_J\) have dimension \(|J| - 1\) and there is one-to-one correspondence between reflecting hyperplanes and midplanes.

**Proof.** Note first that the ray \(\mathbb{R}^+x_J\) lies in the interior of the chamber \(C = \{x \in V_J : B(x, e_s) > 0 \ \forall s \in S\}\). Hence in each chamber \(wC, w \in W_J\) there is a vertex \(wx_J\) of \(X_J\). Now suppose that \(X_J\) does not contain the origin in the interior, then there is a hyperplane \(H\) through the origin such that \(X_J\) is contained in one of the closed half-spaces defined by \(H\), say in \(H_+\). This implies that each chamber has an interior point, lying in \(H_+\). Take an arbitrary closed chamber \(D\). If \(D\) lies entirely in \(H_+\) then \(-D\) lies in the opposite half-space \(H_-\) and hence there is no interior point in it belonging \(H_+\) -- contradiction. If \(D\) does not lie entirely in \(H_+\) then \(H\) separates some codimension one face \(F\) of \(D\) from the remaining vertex \(x\) of \(D\). Let \(D'\) be the chamber, adjacent to
Asymptotic Behavior of Word Metrics on Coxeter Groups

381

D in a face F, then D' lies entirely either in $H_+$ or in $H_-$ and the previous argument works.

Definitions 2.1.2 It follows from Lemma 2.1.1 that the midplanes $M(J, w)$ also cut $X_J$ into (relatively open) polyhedral pieces of dimension $|J| - \text{blocks}$. Two blocks are adjacent if they have a common codimension one face. There is a canonical one-to-one correspondence between blocks in $X_J$, chambers of the Coxeter complex $C(W_J)$ and vertices of $X_J$. This correspondence clearly preserves the adjacency relation. Each block contains a unique vertex of $X_J$ since a closed block $B$ is a fundamental domain of the action of $W$ on $X_J$, i.e., each $x \in X_J$ is conjugated under $W$ to one and only one point in $B$. The group $W_J$ acts on the set of blocks and this action is simply transitive. For a block $B$ the intersection of the closed block $B$ with a midplane is called by internal face of $B$.

Figure 2: The cell $X_J$ for $J = \{s, s'\}, m_{s,s'} = 3$ divided into blocks by midplanes.

Lemma 2.1.3 The only faces of a cell $X_K$ having nonempty intersection with midplane $M(K, s)$, $s \in S$ are those $wX_{JK}$ with $w^{-1}sw \in W_J$. In particular $M(K, s)$ contains no vertices of $X_K$. More generally a face of $X_K$ has nonempty intersection with midplane $M(K, usu^{-1})$, $s \in S, u \in W$ iff it is of the form $uwX_{JK}$ with $w^{-1}sw \in W_J$.

Proof. If $w^{-1}sw \in W_J$ then $swW_J = wW_J$, that is $s$ leaves the vertex set of $wX_{JK}$ invariant and hence it leaves invariant the cell itself and has a nonempty fixed-point set in this cell. Conversely if $M(K, s) \cap wX_{JK}$ is nonempty then there a face $F$ of the cell $wX_{JK}$ such that $M(K, s) \cap F$ contains an interior point of $F$. But then $s$ leaves $F$ invariant hence by Lemma 1.4.2 it also leaves any "overcell" invariant in particular $wX_{JK}$ and this implies that $w^{-1}sw \in W_J$.

To deduce the second statement from the first, one need only to note that $M(K, wsu^{-1}) \cap wX_{JK} = w(M(K, s) \cap X_{JK})$.

Documenta Mathematica 16 (2011) 373–398
Lemma 2.1.4 If \( w \in W_J \) leaves invariant some midplane \( M \) in \( X_J \) then it fixes this midplane pointwise.

Proof. Indeed, \( w \) leaves invariant the ambient face \( C \) and we can apply Lemma 1.4.1. \( \square \)

Lemma 2.1.1 For any cell \( X_K \) the following hold:
1. The intersection of a midplane of \( X_K \) with any of its face is again a midplane.
2. Any midplane of any face of \( X_K \) is an intersection with this face of a precisely one midplane of \( X_K \).

Proof. 1) We may assume that a given midplane \( M \) is of the form \( M(K,s) \) and the face of \( X_K \) is \( X_{JK} \), \( J \subseteq K \). Since \( s \) belongs to \( W_J \), it leaves \( X_{JK} \) invariant and its fixed-point set \( X_{sJK} \) bijects onto the fixed-point set \( X_{J} \) by a \( W_J \)-equivariant isometry \( p_J : X_{JK} \rightarrow X_J \). The general assertion follows by equivariance.
2) We may assume that the face is of the form \( X_{JK} \) for \( J \subseteq K \). Let \( M_{JK} \) be a midplane of \( X_{JK} \), then by definition \( M_{JK} = (p_J|X_{JK})^{-1}(M(J,w)) \) for some \( w \in W_J \). Hence, by \( W_J \)-equivariance, \( w \) is identical on \( M_{JK} \) thus \( M_{JK} = M(K,w) \cap X_{JK} \). Furthermore, \( w \in J \) by Lemma 2.1.3. Hence the segment \( \sigma = [wx_J,x_J] \) is an edge of the face \( X_{JK} \), flipped by \( w \). The intersection \( M_{JK} \cap \sigma = \{m\} \) is a midpoint of \( \sigma \) and \( M_{JK} \) is orthogonal to \( \sigma \). Now if \( M \) any midplane with the same intersection with \( X_{JK} \) as \( M_K \), then the reflection in \( M \) flips the edge \( \sigma \) and hence this edge is orthogonal to \( M \) and thus \( M = M_K \). \( \square \)

Lemma 2.1.5 1) For every \( x \in M(K,s) \cap X_{JK} \) there is a nondegenerate segment of the form \( [y,sy], y \in X_{JK} \) with \( x \) as a midpoint. 2) The segment \( [y,sy] \) is orthogonal to midplane \( M(K,s) \). 3) For any midplane \( M \) in \( X_K \) there is an edge of \( X_K \), intersected by \( M \) in the midpoint.

Proof. 1) Since \( M(K,s) \cap X_{JK} \) is nonempty, it follows from Lemma 2.1.3 that \( s \in J \). Let
\[
U = \{ u \in W_J; x_k \text{ and } ux_K \text{ are on the same side of } M(K,s) \}.
\]
Clearly \( W_J = U \cup sU, U \cap sU = \emptyset \) and the sets \( Ux_K, sUx_K \) lie entirely on the different sides of the midplane \( M(K,s) \). Since \( X_{JK} = \text{Ch}(W_Jx_K) \), we have
\[
x = \sum_{u \in U} (\lambda_u ux_K + \mu_usux_K),
\]
where \( \sum_{u \in U} (\lambda_u + \mu_u) = 1 \) and all coefficients are nonnegative. Since \( x \) is fixed by \( s \), applying \( s \) to both parts of the equality above we get
\[
x = \sum_{u \in U} (\mu_usux_K + \lambda_u ux_K),
\]
We conclude from these two equalities that $x = 1/2(y + sy)$, where $y = \sum_{u \in U}(\lambda_u + \mu_u)ux \in X_{JK}$.

2) The segment $[y, sy]$ is orthogonal to $M(K, s)$ since it is flipped by an orthogonal transformation $s$.

3) If $M = M(K, s)$, then the edge $[sx_K, x_K]$ of $X_K$ is intersected by $M$ in the midpoint. □

We will call the segment $[y, sy]$ from the lemma above to be a perpendicular to $M(K, s)$ in the point $x$.

**Lemma 2.1.6** Let $x \in M(K, s), z \in X_K, x \neq z$ and let $[y, sy]$ be a perpendicular to $M(K, s)$ in the point $x$. Then either $[x, z] \subset M(K, s)$ or one of the angles $\angle_x([x, z], [x, y]), \angle_x([x, z], [x, y])$ is strictly less than $\pi/2$.

**Proof.** It follows from the fact that the tangent space in $x$ is orthogonal sum of the tangent space of $M(K, s)$ and a tangent space of the segment $[y, sy]$.

□

### 2.2 Walls as equivalence classes of midplanes

We assume that $M = M_W$ is the Moussong complex of a Coxeter group $W$. The following definition mimics the definition of a hyperplane in a cube complex given in [NR98].

**Definitions 2.2.1** For midplanes $M_1$ and $M_2$ of the cells $C_1 = C(M_1)$ and $C_2 = C(M_2)$ respectively we write $M_1 \sim M_2$ if $M_1 \cap M_2$ is again a midplane (and then of course it is a midplane of $C_1 \cap C_2$). The transitive closure of this symmetric relation is an equivalence relation, and the union of all midplanes in an equivalence class is called a wall in $M$. Clearly the equivalences above are generated by those of the form $M_1 \sim M_2, C_1 \leq C_2$ or $C_2 \leq C_1$. Thus to prove some property $P$ for midplanes of a wall $H$ it is enough to prove this property for some midplane in $H$ and then show that the validity of $P$ is preserved under equivalences just mentioned. If $M$ is a midpoint of a 1-cell(=edge) in $M$ then the wall spanned by $M$ will be called a dual wall of $e$ and denoted by $H(e)$. We denote by $H_M$ the union of midplanes in the equivalence class of a midplane $M$.

It follows immediately from Lemma 2.1.5 that

**Lemma 2.2.2** Any wall $H$ of $M$ has the form $H(e)$ for some edge $e$.

Clearly $W$ acts on the set of midplanes, preserving the equivalence relation and hence acts on the set of walls. For any wall $H$ we denote by $\bar{H}$ the complex obtained from the disjoint union of midplanes in $H$ by gluing any two midplanes in $H$ along their common submidplane in $M$ (if such one exists). One can easily see that $\bar{H}$ is nonpositively curved, i.e. satisfies the link condition. Namely, the link of any cell $C$ of $M$ is isometric to the product $C \times [-\pi, \pi]$. 

**Documenta Mathematica 16 (2011) 373–398**
Lemma 2.2.3 Let \( p : \tilde{H} \to \mathcal{M} \) be the natural map which sends each midplane in \( \tilde{H} \) to its image in \( \mathcal{M} \). Then \( p \) is an isometry of \( \tilde{H} \) onto \( H \). As a consequence of the above walls are convex in \( \mathcal{M} \).

Proof. It is similar to the proof of lemma 2.6 in [NR98]. Clearly, \( p \) is an isometry on each midplane. By result of M. Gromov ([Gro87], Section 4) it is enough to show that \( p \) is a local isometry, that is if \( x \in H \), then there is a neighborhood \( U \) of \( x \) such that \( p|_U \) is an isometry. Clearly \( p \) bijects the star \( St(x) \) onto the union \( U \) of all midplanes, containing \( p(x) \). This union is the fixed-point set of some reflection from \( W \) (see Lemmas 1.4.1, 1.4.2, 1.4.4). Hence \( U \) is convex, and \( p \) maps \( St(x) \) isometrically onto \( U \).

Lemma 2.2.4 Each wall in \( \mathcal{M}_W \) is the fixed-point set of a precisely one reflection in \( W \). Conversely, the fixed-point set of a reflection in \( W \) is a wall.

Proof. Let \( H_M \) be the wall, spanned by a midplane \( M \) of the cell \( C \). From the description of cells and that of the action of \( W \) we know that \( M \) is the fixed point set of a reflection from the stabilizer \( S_C \) of \( C \) in \( W \). We will show that \( H_M \) coincides with the fixed-point set \( H_w \) of \( w \).

Any reflection \( w \) fixing a midplane \( M \) pointwise fixes also \( H_M \) pointwise, i.e., \( H_M \subseteq H_w \). We have to show that the claimed property is invariant under equivalence relation of midplanes, see §2.2.1. If \( M_1 \sim M_2 \) are midplanes in \( C_1 = C(M_1), C_2 = C(M_2) \) respectively, \( C_1 \leq C_2 \), and \( w \) fixes \( M_1 \) then \( w \) leaves \( C_1 \) invariant, hence by Lemma 1.4.2 it leaves \( C_2 \) invariant and by Lemma 2.1.1 it leaves \( M_2 \) invariant and finally by Lemma 2.1.4 it fixes \( M_2 \) pointwise. In case \( M_1 \sim M_2, C_1 \geq C_2 \), and \( w \) fixes \( M_1 \) pointwise it is clear that \( w \) fixes \( M_2 \) pointwise.

Every wall \( H \) is the fixed-point set of a unique reflection in \( W \). Write \( H \) as the dual wall \( H = H(e) \) of some edge \( e \) of \( \mathcal{M} \). If there were two reflections \( w, w' \) with the same reflection wall \( H \) then their difference \( w^{-1}w' \) would fix \( e \) pointwise. But \( W \) acts simply transitively on the vertices of \( \mathcal{M} \) hence \( w = w' \).

Now any \( w \in W \) fixing at least one cell pointwise is an identity. Indeed the set of cells fixing by \( w \) pointwise is nonempty and containing with each cell \( C \) every its "overcell" \( C' \supset C \) because by Lemma 1.4.2 \( wC' = C' \) and since the stabilizer of \( C' \) acts fixed point free on the cell we conclude that \( w = 1 \).

\( H_w \) coincides with \( H_M \). Suppose, to the contrary, that there is a \( w \)-fixed point \( x \) outside \( H_M \). Take any \( y \in H_M \), then \( w \) fixes the endpoints \( x, y \) of the geodesic \([x, y]\) hence, by uniqueness, it fixes the whole geodesic. Shortening \([x, y]\) if necessary we may assume that \([x, y]\) is outside \( H_M \). Take \( z \in [x, y], z \neq y \) such that the open segment \((z, y]\) is contained entirely in the interior of some cell \( C \). Since \( w \) fixes \((z, y]\), it leaves \( C \) invariant. As far as \( y \in H_M \cap C \), the point \( y \) is contained in some midplane \( M' \subset H_M \) of \( C \). Because \( w \) fixes \( M' \) and the segment \((z, y]\), lying entirely outside \( H_M \), we conclude that \( w \) fixes \( C \) pointwise - contradiction.

For the converse, let \( w \) be a reflection in \( W \). Note first that \( H_w \) contains at least one midplane. Indeed, since any reflection in \( W \) is conjugate to some \( s, s \in \mathbb{Z} \).
Asymptotic Behavior of Word Metrics on Coxeter Groups

S, we may assume that $w = s$. Take $J = \{s\}$, then the cell $X_J = Ch(x_J, x_{sJ})$ is a segment on which $s$ acts as a reflection thereby fixing its midpoint $M$.

We conclude that $H_w$ contains $H_M$ for some midplane $M$. Therefore, as was proved above, $H_w$ coincides with $H_M$. □

**Lemma 2.2.5** The edge path in $\mathcal{M}^{(1)}$ is geodesic if and only if it crosses each wall at most once.

**Proof.** If an edge path $p = e_1 e_2 \cdots e_k$ crosses a wall $H$ twice, say distinct edges $e_i, e_j, i < j$ cross $H$, then we delete the subpath $e_i \cdots e_j$, and instead insert the path $w(e_{i+1} \cdots e_{j-1})$, where $w$ is the reflection in the wall $H$. The resulting path is strictly shorter than $p$ but connects the same vertices. Conversely, suppose that an edge path $p$ from $x$ to $y$ crosses each wall at most once. Let $\mathcal{H}_H$ be the set of all walls crossing by $p$. Since $x$ and $y$ are at the different sides of each wall from $\mathcal{H}_H$, we conclude that any path from $x$ to $y$ should cross than that of $p$. □

Any wall in the Moussong complex is "totally geodesic" in the following sense

**Lemma 2.2.6** Any geodesic in $\mathcal{M}$ having nondegenerate piece in a wall $H$, lies entirely in $H$.

**Proof.** Suppose the lemma is false, then there are nondegenerate segments $\sigma_1 = [x, x_1], \sigma_2 = [x, x_2]$, cells $C_1, C_2$, and midplanes $M_1, M_2$ of $C_1, C_2$ respectively such that

1) $M_1 \sim M_2$,
2) $x \in M_1 \cap M_2$,
3) $\sigma_1 \subset M_1, x_2 \in C_2 - M_2$,
4) $\sigma_1 \cup \sigma_2$ is geodesic.

It follows from Lemma 2.1.5 that there is a reflection $w \in W$ and a segment $[y, wy]$ with $x$ as a midpoint and orthogonal to both $M_1$ and $M_2$. Write $\sigma = [x, y], \sigma' = [x, wy]$. Since, by 3), $x_2 \in (C_2 - M_2)$, it follows from Lemma 2.1.6 that one of the angles $\angle_x(\sigma_2, \sigma), \angle_x(\sigma_2, \sigma')$ is strictly less than $\pi/2$ and $\angle_x(\sigma_1, \sigma') = \angle_x(\sigma_2, \sigma) = \pi/2$. Hence the angle between the segments $\sigma_1, \sigma_2$ in the point $x$ is strictly less than $\pi$, thus $\sigma_1 \cup \sigma_2$ can not be geodesic by criterion of Lemma 1.5.1.

2.3 Separation properties

**Lemma 2.3.1** Every wall in $\mathcal{M}$ separates $\mathcal{M}$ into exactly two connected components.

**Proof.** First, we claim that $H$ separates $\mathcal{M}$ into at least two components. We know from Lemma 2.2.2 that $H = H(e)$ – the dual wall of some edge $e = [x, y]$. We will show that $x, y$ belong to different connectedness components. Suppose, to the contrary, that $x, y$ are in the same connectedness component. Then
there is a closed edge path $\alpha$ in $\mathcal{M}^{(1)}$ crossing $H$ only once. (Clearly any edge either intersects $H$ in a midpoint or does not intersects $H$ at all.) Since $\mathcal{M}$ is contractible this path can be contracted to a constant path by a sequence of combinatorial contractions in cells. By Lemma 2.1.1 any cell $C$ either has an empty intersection with $H_M$ or $H_M \cap C$ is a midplane of $C$. This implies that each combinatorial contraction of the edge path in the cell does not change the number of intersections with $H_M$ modulo 2. Since this number is 0 for the final constant path, it cannot be 1 for the initial path.

To prove that the $H$ cuts out $\mathcal{M}$ into exactly two components, we proceed as in [NR98], lemma 2.3 (preprint version.) Notice first that $H$ is 2-sided, that is there exists a neighborhood of $H$ in $\mathcal{M}$ which is homeomorphic $H \times I, I = [0,1]$. Indeed, by Lemma 2.1.5, in each cell there is a neighborhood which is fibered as $\mathcal{M} \times I$: the fibrations can be chosen to agree on face maps so this induces an $I$-bundle structure on some neighborhood $N$ over $H$.

Since $H$ itself is CAT(0) it is contractible so the bundle is trivial. It follows that $N$ has two disjoint components, $\{-1/2\} \times H$ and $\{1/2\} \times H$. Any point in the complement of $H$ can be joined to one of these boundary components by a path in the complement of $H$, and therefore $X - H$ has exactly 2 components as required.  

**Lemma 2.3.1** For any wall $H$ both components of the complement $\mathcal{M} - H$ are convex.

**Proof.** Suppose that $x_1, x_2$ lie on the same side of $H$, say $H^+$. We claim that $[x_1, x_2]$ lies entirely in $H^+$. Suppose the contrary, then by Lemma 2.2.6 the intersection $[x_1, x_2] \cap H$ consists of precisely one point, say $x$. Similar to the proof of Lemma 2.2.6 there are segments $\sigma_1 \subset [x, x_1], \sigma_2 \subset [x, x_2]$, cells $C_1, C_2$, and midplanes $M_1, M_2$ of $C_1, C_2$ respectively such that

1) $M_1, M_2 \subset H$,
2) $x \in \sigma_1 \cap \sigma_2$,
3) $\sigma_1 \subset C_1, \sigma_2 \subset C_2$,
4) $\sigma_1 \cup \sigma_2$ is geodesic.
5) The interiors of $\sigma_1, \sigma_2$ are contained entirely in $H^+$.

Then it follows from Lemma 2.1.5 that there exists a reflection $w \in W$ and a segment $[y, wy]$ such that the segment has $x$ as a midpoint and is orthogonal to both $M_1$ and $M_2$. By interchanging the roles of $y$ and $wy$ if necessary we may assume that $y \in H^+$. Denote $\sigma = [x, y], \sigma' = [x, wy]$. It follows from Lemma 2.1.6 that the angles $\angle_{2}(\sigma_2, \sigma), \angle_{2}(\sigma_2, \sigma')$ are both strictly less than $\pi/2$. But a small nonzero move of $x$ along $\sigma$ would strictly shorten the length of $\sigma_1 \cup \sigma_2$ contradicting the assumption 4) above.

□
3 Chambers and galleries

3.1 Chambers

Since the complex $\mathcal{M}$ is locally finite and there are only finite number of midplanes in each cell, we conclude that the set of all walls $\mathcal{H}$ in $\mathcal{M}$ is locally finite, in the sense that every point of $\mathcal{M}$ has a neighborhood which meets only finitely many $H \in \mathcal{H}$.

**Definition 3.1.1** By Lemma 2.3.1 the walls $H \in \mathcal{H}$ yield a partition of $\mathcal{M}$ into open convex sets, which are the connected components of the complement $\mathcal{M} - (\cup H)$. We call these sets chambers.

To distinguish chambers from cells, we will denote them by letter $D$, possibly with indices, dashes, etc.

**Lemma 3.1.2** For any two distinct chambers $D(x), D(y), x, y \in \mathcal{M}^{(0)}$ there is a wall $H$ separating them.

**Proof.** Consider a geodesic edge path $p = e_1 e_2 \cdots e_k$ from $x$ to $y$, then by Lemma 2.2.5 $H(e_1)$ separates $x$ from $y$ and hence separates $D(x)$ from $D(y)$.

**Lemma 3.1.3** Each chamber contains precisely one vertex of $\mathcal{M}$.

**Proof.** Since $W$ acts simply transitively on the set of vertices of $\mathcal{M}$ and each vertex is contained in some chamber we deduce that each chamber contains at least one vertex. Now, if $x, y$ are distinct vertices in a chamber $C$, we connect them by a geodesic path $p$ in $\mathcal{M}^{(1)}$. Then by criterion of geodesicity any wall crossed by $p$ separates $x$ from $y$, contradicting the definition of chamber.

In view of this lemma we will write $D(x)$ for the chamber containing the vertex $x$ of $\mathcal{M}$.

**Definitions 3.1.4** Recall from §2.1.2 that midplanes of any cell $C$ in $\mathcal{M}$ yield a partition of $C$ into convex (open) blocks. (Blocks are open in $C$, not in $\mathcal{M}$.) A maximal block is a block in a maximal cell. Two maximal blocks are adjacent if they are contained in the same maximal cell and share a codimension one face. Two chambers $D, D'$ are adjacent if there are maximal blocks $B \subset D, B' \subset D'$ which are adjacent. A wall $H$ is a wall of a chamber $D$ if there is a maximal cell $C$ such that $H \cap C$ contains a codimension one face $F$ of a maximal block $B$ of $D$.

**Lemma 3.1.5** 1) Every chamber is uniquely determined by any of its maximal blocks. 2) Every chamber is a union of maximal blocks, and it contains at most one maximal block from each maximal cell.
Proof. 1) Indeed, the interior of a maximal block is open in $M$ and does not intersect any wall, consequently there is only one chamber containing this block.

2) Since $M$ is a union of maximal cells, any chamber is a union of maximal blocks. Take a chamber $D$, then

$$D = \cup \{ D \cap C : C \text{ is a maximal Moussong cell} \}.$$  

The intersection $D \cap C$ is a union of maximal blocks because $D \cap C$ is an intersection of open half-cells in $M$. Next, if $D$ contains two maximal blocks $B_1, B_2$ from one cell, then there is a midplane $M$ separating $B_1$ from $B_2$ and the ambient wall $H$ also separates $B_1$ from $B_2$ contradicting the definition of $D$. □

Lemma 3.1.6 Let $B, B'$ be maximal adjacent blocks and let $D, D'$ be corresponding ambient chambers. Let $H$ be a wall separating $B$ from $B'$. Then $H$ is the only wall that separates $D$ from $D'$.

Proof. Let $C$ be a maximal cell containing $B, B'$, then $B, B'$ are adjacent in this cell and clearly there is only one midplane separating them. But the wall is uniquely determined by any of its midplanes, whence the lemma. □

Lemma 3.1.7 Let $D, D'$ be chambers such that their closures $\overline{D}, \overline{D'}$ have a nonempty intersection. Let $H$ be a wall, separating $D$ from $D'$. Then $H$ contains the intersection $\overline{D} \cap \overline{D'}$.

Proof. Suppose, to the contrary, that there is $b \in \overline{D} \cap \overline{D'}$ which is not contained in $H$. Since $H$ is closed a small neighborhood of $b$ does not intersect $H$. But this neighborhood contains points both from $D$ and $D'$, which thus belong to one halfspace of $H$, contradicting the separation hypothesis. □

Lemma 3.1.8 Two distinct chambers $D(x), D(y) \ (x, y \in \mathcal{M}^{(0)})$ are adjacent if and only if the vertices $x, y$ are adjacent in $\mathcal{M}^{(1)}$. For any two adjacent chambers there is a reflection in $W$, permuting these chambers and fixing the intersection of their closures pointwise.

Proof. The lemma is about Coxeter cell, thus it follows from the description of its structure as a Coxeter complex. □

Definition 3.1.9 The base chamber $D_0$ of $\mathcal{M}$ is the chamber, containing the base vertex $x_0$ of $\mathcal{M}$. For each $s$ from the generating set $S$ of $W$, we denote by $H_s^+$ those open halfspace of the wall $H_s$, which contains the base vertex $x_0$.

Lemma 3.1.10 $D_0 = \cap \{ H_s^+ : s \in S \}$. 

Documenta Mathematica 16 (2011) 373–398
Proof. Since $D = \cap \{H_s : s \in S\}$ contains $x_0$, it contains also $D_0$. Let $B_J$ be a block of a maximal cell $X_J$, containing $x_J = x_0$. Then $B_J \subset D_0$ — indeed it follows from the description of the chambers in the Coxeter complex that $B_J$ is bounded by the hyperplanes $H_s = e_s^+, s \in J$. Suppose now that $D$ strictly contains $D_0$ and let $x \in D - D_0$. Since $D$ is convex, the whole segment $[x, x_0]$ lies in $D$. Let $T = (x_0, x_1, \ldots, x_m)$ be a taut chain from $x_0$ to $x_m = x$. The first piece $[x_0, x_1]$ lies entirely in some maximal cell of the form $X_K$ and we know that the block $B_K = D_0 \cap X_K$ is the maximal block in $X_K$ and it is bounded by the hyperplanes $H_s^-, s \in K$.

If $x_1$ is a vertex of $X_K$, then it is separated by some $H_s, s \in K$ from $x_0$. If $x_1$ is not a vertex of $X_K$, then $x_1$ is the boundary point of $X_K$ and hence it is contained in the interior of some face $F$ of $X_K$. If $F$ contains $x_0$, then all three points $x_0, x_1, x_2$ lie in some cell contradicting to the choice. Hence $F$ does not contain $x_0$ and thus the open interval $(x_0, x_1)$ lies entirely in the interior of $X_J$ and hence crosses some wall $H_s^+, s \in J$ — contradiction.

3.2 Galleries

Definitions 3.2.1 A gallery is a sequence of chambers $\Gamma = D_1D_2\cdots D_k$ such that any two consecutive ones are adjacent.

Recall that the chambers are in one-to-one correspondence with the vertices of $\mathcal{M}$ and chambers are adjacent if and only if the correspondent vertices are adjacent in the 1-skeleton of $\mathcal{M}$. It follows immediately that the following lemma is true.

Lemma 3.2.2 1) Any two chambers $D, D'$ can be connected by a gallery of length $d(D, D')$. 2) A gallery is geodesic if and only if and only if it does not cross any wall more than once. 3) Given $s_1, \ldots, s_d \in S$, there is a gallery of the form $D_0(s_1D_0)(s_1s_2D_0)\cdots(s_1s_2\cdots s_dD_0)$. Conversely, any gallery starting at $C$ has this form. 4) The action of $W$ is simply transitive on the set of chambers.

Lemma 3.2.3 There is a constant $c(\mathcal{M})$ such that for any two distinct chambers $D, D'$ with nonempty intersection $\overline{D} \cap \overline{D'}$, there is a geodesic gallery $\Gamma = D_0D_1\cdots D_k$ from $D_0 = D$ to $D_k = D'$ whose length $k$ does not exceed $c(\mathcal{M})$.

Proof. Let $\mathcal{H}_0$ be the set of walls separating $D$ from $D'$. In view of Lemma 3.2.2, it is enough to bound the cardinality of $\mathcal{H}_0$. According to Lemma 3.1.7 each $H \in \mathcal{H}_0$ contains $\overline{D} \cap \overline{D'}$. Let $x \in \overline{D} \cap \overline{D'}$. Clearly the number of cells containing $x$ is uniformly bounded and for each such a cell $C$ the number of midplanes in $C$ containing $x$ is also uniformly bounded. Since a wall is uniquely determined by any of its midplanes, this proves the lemma.
3.3 Approximation property

Definition 3.3.1 Let \((X, G)\) be a pair consisting of a geodesic metric space \(X\) and a graph \(G\) embedded into \(X\). We say that \((X, G)\) satisfies the approximation property if \(X\)-geodesics between the vertices of \(G\) can be uniformly approximated by geodesics in \(G\). This means that there is a constant \(\delta\) such that for any \(X\)-geodesic \(\alpha_X\) between the vertices of \(G\) there is a \(G\)-geodesic \(\alpha_G\) between the same vertices such that both \(\alpha_X\) and \(\alpha_G\) lie entirely in the \(\delta\)-neighborhoods of each other. We will express this by saying that \(\alpha_X, \alpha_G\) are \(\delta\)-close to each other.

Of particular interest is the case when \(G\) is the embedded Cayley graph of a group acting on \(X\).

Theorem 3.3.2 Let \((W, S)\) be a Coxeter group and let \(M\) be its Moussong complex. Embed the Cayley graph \(C_W\) as an orbit \(W x_0\) for a point \(x_0\) in a base chamber \(D_0\) of \(M\). Then the pair \((M_W, C_W)\) satisfies the approximation property.

Proof. Let \(\sigma = [a, b]\) be a nondegenerate segment in \(M\) and \(H_\sigma\) be the set of walls having a nonempty intersection with the interior \((a, b)\). Since the family of all walls is locally finite and the walls are totally geodesic, we have a partition

\[ H_\sigma = H'_\sigma \cup H_1 \cup H_2 \cup \cdots \cup H_n, \]

where the walls from \(H'_\sigma\) contain \(\sigma\) and the walls from \(H_i\) cross \(\sigma\) precisely in the point \(a_i, i = 1, \ldots, n\), and \(a = a_0 < a_1 < \cdots < a_n < a_{n+1} = b\).

Now we define a gallery \(\Gamma\) along the geodesic \(\sigma = [a, b]\) as the gallery

\[ \Gamma = D_1\Gamma_1D_2\Gamma_2D_3\cdots D_n\Gamma_nD_{n+1} \]

such that

1) \(\overline{D_i} \cap [a, b] = [a_{i-1}, a_i]\) \((i = 1, 2, \ldots, n + 1)\),

2) Each spherical piece \(D_i\Gamma_iD_{i+1}\) is a geodesic gallery and the lengths of spherical pieces are bounded from above by the constant \(e(M)\) from Lemma 3.2.3,

3) Each spherical piece \(D_i\Gamma_iD_{i+1}, i = 1, \ldots, n\) crosses the walls only from the set \(H'_\sigma \cup H_i\).

Lemma 3.3.3 For any geodesic \(\sigma = [a, b]\) in \(M\) there is a geodesic gallery along \(\sigma\).

Proof of the lemma. By construction of the sequence \(\{a_i\}\), for each \(i = 1, \ldots, n + 1\) there is a chamber \(D_i\) such that \(\overline{D_i} \cap [a, b] = [a_{i-1}, a_i]\). The corresponding sequence of chambers \(D_1, D_2, \ldots, D_n, D_1 = D, D_n = D'\) is the first approximation to the required gallery. In general, this sequence is not a gallery, since two consecutive chambers are not necessarily adjacent. For each \(1 \leq i \leq n\), the intersection of neighbors \(\overline{D_i} \cap \overline{D_{i+1}}\) contains the point \(a_i\).
Application of Lemmas 3.1.7, 3.2.3 enables us to inscribe a spherical geodesic subgallery of bounded length between these neighbors and get a gallery

\[ \Gamma = D_1 \Gamma_1 D_2 \Gamma_2 \cdots D_{n-1} \Gamma_{n-1} D_n \]

such that the spherical pieces \( D_j \Gamma_j D_{j+1} \) are geodesic galleries of uniformly bounded length satisfying condition 3) from the definition above. We will show that \( \Gamma \) can be modified to a geodesic gallery along \([a, b]\). If \( \Gamma \) is not geodesic then by Lemma 3.2.2 it crosses some wall \( H \) at least twice. Clearly \( H \in \mathcal{H}' \) i.e., \( H \) contains \( \sigma \). Then there are indices \( i + 1 < j \) and subgalleries \( \Gamma_i, \Gamma_j \) each of length 1 such that

a) \( \Gamma_i, \Gamma_j \) belong to \( i \)-th and \( j \)-th spherical piece respectively,
b) \( \Gamma_i, \Gamma_j \) cross \( H \) and moreover there are no crossing subgalleries in between.

Let \( \Gamma_1 = DD', \Gamma_2 = D''D''' \). In particular the chambers \( D \) and \( D'' \) lie on the same side of \( H \), say \( H^- \), and the subgallery \( \Gamma' \) of \( \Gamma \), joining \( D' \) with \( D'' \) lies on the opposite side, say \( H^+ \).

Let \( w \in W \) be the reflection in the wall \( H \). If we modify \( \Gamma \) by applying \( w \) to the portion \( \tilde{\Gamma} \), we obtain the gallery from \( D \) to \( D'' \) that is strictly shorter than \( D' \Gamma' D'' \). Replacing \( D' \Gamma' D'' \) by \( w(\tilde{\Gamma}) \) we get the gallery \( \Gamma' \) that is strictly shorter than \( \Gamma \) but still is the gallery along \( \sigma \). Repeating the previous process will construct a geodesic gallery along \( \sigma \). This proves Lemma 3.3.3.

The theorem now follows easily from Lemma 3.3.3. Namely, given two chambers \( D, D' \) we take the points \( d, d' \) inside them and build a geodesic gallery \( \Gamma = D_1 \cdots D_n \) along \([d, d']\). \( \Gamma \) not necessarily joins \( D \) to \( D' \) but the intersections \( D_i \cap D_j, D_i' \cap D_j' \) are nonempty, so we can join \( D \) to \( D_1 \) and \( D' \) to \( D_2 \) respectively by the galleries of uniformly bounded length thereby getting the gallery joining \( D \) to \( D' \) and that is \( \delta(M) \) close to \( \sigma \) for some universal constant \( \delta(M) \). \( \square \)

![Figure 3: Gallery along geodesic. The spherical piece is \( D_4 D'D_5 \).](image)

4 Word length on abelian subgroups of a Coxeter group

4.1 Straightness

**Definition 4.1.1** Let \( G \) be a group with a fixed word metric \( x \mapsto \ell(x) \). We say that an element \( x \neq 0 \) is straight if \( \ell(x^n) = n\ell(x) \) for all natural \( n \).
Remark 4.1.2 Straight elements have been studied for Coxeter groups in [Kr94] and for small cancellation groups in [Kap97] (in the last paper they are called \textit{periodically geodesic}).

Example 4.1.3 (An element that is not straight.) Let $W$ be an affine Coxeter group generated by reflections $s_1, s_2, s_3$ in the sides of an equilateral triangle $C$ of a Euclidean plane. Let $L_1, L_2, L_3$ be the corresponding reflecting lines of this triangle. It is easily seen that there is nontrivial translation $u \in W$ with an axis $L_1$. We assert that nor $s_1 u$ neither any of it conjugates $v = w s_1 u w^{-1}$ are straight. Indeed, the length $|w s_1 u w^{-1}|$ is the length of a geodesic gallery $\Gamma$ from $C$ to $w s_1 u w^{-1} C$. Any such a gallery intersects the line $w L_1$. The concatenation $\Gamma(v \Gamma)$ is a gallery from $C$ to $v^2 C$ of length $2 |w s_1 u w^{-1}|$. But $\Gamma(v \Gamma)$ can not be geodesic, since it intersects $w L_1$ twice. Hence $|v^2| < 2 |v|$. \hfill $\Box$

Definition 4.1.4 Let $\mathcal{M}$ be the Moussong complex of a Coxeter group $W$. Recall that $\mathcal{M}$ is a proper complete CAT(0) space and $W$ acts properly and cocompactly on $\mathcal{M}$ by isometries. In particular, any element $w \in W$ of infinite order acts as an axial isometry i.e., there is a geodesic axis $A_w$ in $\mathcal{M}$, isometrical to $\mathbb{R}$, on which $w$ acts as a nonzero translation [Bal95]. We say that $w$ is generic if $A_w$ intersects any wall in at most one point. In view of Lemma 2.2.6, this is equivalent to saying that no nondegenerate segment of $A_w$ is contained in a wall.

Theorem 4.1.5 Let $(W, S)$ be a Coxeter system of finite type. For any generic element $w$ of $W$ of infinite order there is a conjugate $v$ which is straight, that is $\ell(v^n) = n \ell(v)$ for all $n \in \mathbb{N}$, where $\ell(v)$ is a word length in generators $S$.

Proof. We make use of the action of $W$ on the Moussong complex $\mathcal{M}$. Since the family of all walls is locally finite, there is a point $a$ on the axis $A_w$ such that $a$ does not belong to any wall of $\mathcal{M}$. Every point $w^i a (i \in \mathbb{Z})$ also does not belong to any wall of $\mathcal{M}$. Let $\mathcal{H}$ be the set of walls crossed by the segment $[a, wa]$ and let $a < a_1 < a_2 < \cdots < a_k < wa$ be the crossing points, so that $\mathcal{H}$ is a disjoint union $\mathcal{H} = \mathcal{H}_1 \cup \cdots \cup \mathcal{H}_k$ of subsets $\mathcal{H}_i$, crossing $[a, wa]$ in $a_i, i = 1, 2, \ldots, k$. There are the chambers $D_1, D_2, \ldots, D_k$ such that $D_i \cap [a, wa] = [a, a_1], D_i \cap [a, wa] = [a_{i-1}, a_i] (i = 1, 2, \ldots, k)$. Inscribe into the sequence $D_1, D_2, \ldots, D_k(w D_1)$ subgalleries $\Gamma_1, \ldots, \Gamma_k$, so that the concatenation $\Gamma = D_1 \Gamma_1 D_2 \cdots \Gamma_{k-1} D_k \Gamma_k(w D_1)$ is a gallery, crossing only the walls from $\mathcal{H}$ and crossing each wall precisely once. In particular this gallery is geodesic. Let $\Gamma_0 = D_1 \Gamma_1 D_2 \cdots \Gamma_{k-1} D_k \Gamma_k$. Translating by $w$ and concatenating, we get a gallery $\tilde{\Gamma} = \Gamma_0(w \Gamma_0)(w^2 \Gamma_0)\cdots(w^{n-1} \Gamma_0)w^n D_1$. The walls that it crosses are precisely those from the union $\mathcal{H} \cup w \mathcal{H} \cup w^2 \mathcal{H} \cup \cdots \cup w^{n-1} \mathcal{H}$, and each wall is crossed precisely once. Hence the gallery $\tilde{\Gamma}$ is geodesic. Now let $D_1 = u D_0, u \in W$, where $D_0$ is the base chamber. Being a geodesic path in the Cayley graph, the gallery $\tilde{\Gamma}$ joins the vertex $u$ to the vertex $w^n u = u(w^{-1} w^n u)$. Hence its length $\ell(\Gamma_0)$ equals the word length of the element $u^{-1} w^n u \in W$. 

Documenta Mathematica 16 (2011) 373–398
Proof. Let \( w \in W \) be of infinite order and let \( A_w \) be an axis of \( w \). Let \( H_w = H_w(A_w) \) denote the set of walls in the Moussong complex \( \mathcal{M}_W \), containing \( A_w \). It is easy to see that the cardinality of \( H_w \) is bounded by a constant depending only on \( W \) and we take \( c = c(W) \) to be the number

\[
2 \times \text{l.c.m.} \times (\text{card}\{H_w : w \in W \text{ is of infinite order }\}).
\]

Clearly \( A_w \) is an axis of \( w^c \) as well. Furthermore, \( w^c \) leaves invariant each wall \( H \in H_w \); moreover, it leaves invariant each of the two components of \( \mathcal{M}_W - H \). It follows that for any chamber \( D \), a geodesic gallery from \( D \) to \( w^cD \) does not cross a wall \( H \) from \( H_w \). Indeed, otherwise \( D \) and \( w^cD \) would lie in different components of \( \mathcal{M}_W - H \) implying that \( w^c \) interchanges these components, contradicting the property above. Take a chamber \( D \) such that \( \overline{D} \cap A_w \) is a nondegenerate segment and fix a point \( a \) in the interior of this segment. Let \( H \) denote the set of walls \( H \) that are crossed by the segment \([a, w^c a]\) but do not contain it. Clearly any \( H \in H \) separates \( D \) from \( w^c D \). And conversely, if \( H \) separates, then the points \( a, w^c a \) lie in different components of \( \mathcal{M}_W - H \) implying that \( H \) crosses the segment \([a, w^c a]\) in precisely one point. Let \( \Gamma \) be a geodesic gallery from \( D \) to \( w^c D \) then the walls that it crosses are precisely those from \( H \), and each wall \( H \in H \) is crossed by \( \Gamma \) precisely once. Iterating we obtain a gallery \( \Gamma = \Gamma(w \Gamma)(w^2 \Gamma) \cdots (w^{n-1} \Gamma)w^n D(n \in \mathbb{N}) \) of the length \( n \ell(\Gamma) \). This gallery crosses the walls only from (disjoint) union \( H \cup w^c H \cup w^2 c H \cup \cdots \cup w^{(n-1)} c H \), each precisely once. Hence the gallery \( \Gamma \) is geodesic. Now let \( D = uD_0, u \in W \), where \( D_0 \) is the base chamber. Being a geodesic path in the Cayley graph, the gallery \( \Gamma \) joins the vertex \( u \) to the vertex \( u^n a \). Hence its length \( n \ell(\Gamma) \) equals the word length of the element \( u^{-1} w^c a \) in \( W \). We conclude that for \( v = u^{-1} w^c a \) the equality \( |v^n| = n|v| \) holds for all \( n \in \mathbb{N} \). \( \square \)

For elements which are not necessarily generic we have the following

Lemma 4.1.7 Let \((W, S)\) be a Coxeter group of finite type and let \( w \in W \) be an element of infinite order. Fix an axis \( A_w \) of \( w \) in the Moussong complex \( \mathcal{M}_W \). There is a chamber \( D \) such that for all \( n \in \mathbb{Z} \)

\[
d(D, w^n D) = n d(D, wD) - n \text{ card} \ (u^{\infty} \mathcal{H}_w) + c_n,
\]

where \( |c_n| \) is bounded by a constant depending only on \( W \) and \( H_w \) is the set of all walls \( H \) in \( \mathcal{M}_W \), containing \( A_w \) and such that \( H \) separates \( wD \) from \( w^{i+1} D \) for some \( i \in \mathbb{Z} \).
PROOF. We follow the proof of Theorem 4.1.5. Take a chamber $D$, such $\overline{D}\cap A_w$ is a nondegenerate segment. Let $H$ be the set of walls, separating $D$ from $wD$ and do not containing $A_w$. By total geodesicity, any $H \in H$ crosses $A_w$ precisely in one point. Let $\Gamma$ be a geodesic gallery from $D$ to $wD$ then it crosses all $H \in H$, each precisely once, and some of the walls from $H_w$. Iterating we get the gallery $\overline{\Gamma} = \Gamma(w^2\Gamma)\cdots(w^{n\ell-1}\Gamma)w^nD$. This gallery crosses the walls from (disjoint) union $\mathcal{H} \cup w\mathcal{H} \cup w^2\mathcal{H} \cup \cdots \cup w^{n\ell-1}\mathcal{H}$, each precisely once. Also, it crosses some walls from $H_w$. Note that, whenever $\overline{\Gamma}$ crosses $H \in H_w$, it crosses it periodically with a period $r_H = \text{card } w^{2}H$. Hence, the integer part $\lfloor n/r_H \rfloor$ is the number of times the gallery $\overline{\Gamma}$ crosses each $H' \in w^2H$. Hence it crosses the walls from the orbit $w^2H$ approximately $n$ times, up to a universal constant. Hence, the number $d(D, w^nD)$ of walls, separating $D$ from $w^nD$, equals $n d(D, wD) - n \text{ card } (w^2\mathcal{H}_w) + c_n$, where $c_n$ is uniformly bounded. □

Theorem 4.1.8 If, under conditions of Lemma 4.1.7, $D = uD_0, u \in Z$, where $D_0$ is the base chamber, then $d(D, wD)$ is the word length of the conjugate $v = u^{-1}wu \in Z$ and we get the following formula

$$\ell(v^n) = n \ell(v) - \text{ card } (w^2\mathcal{H}_w) + c_n.$$ 

From this we get the following formula for a translation length $\|w\|$ of $w$:

$$\|w\| \overset{\text{def}}{=} \lim_{n \to \infty} \frac{\ell(v^n)}{n} = \lim_{n \to \infty} \frac{\ell(v^n)}{n} = \ell(v) - \text{ card } (w^2\mathcal{H}_w).$$

In particular, translation length of any element of $Z$ is rational (even integral).

Remark 4.1.9 The formula for translation length is similar to the one given in [Kra94], where it follows from the classification of roots. It seems unknown whether translation length is rational in an arbitrary ”semihyperbolic group”.

4.2 Norms and Burago’s Inequality

Let $A$ be a normed abelian group, so $A$ is equipped with a function $\ell : A \to \mathbb{R}$ satisfying (1) $\ell(a^{-1}) = \ell(a)$, (2) $\ell(ab) \leq \ell(a) + \ell(b)$, and (3) $\ell(a) \geq 0$ with $\ell(a) = 0$ if $a = 1$, for $a, b \in A$. If (3) is replaced by (3') $\ell(a) \geq 0$ for $a \in A$, we call $A$ a pseudonormed abelian group. Two pseudonorms $\ell$ and $\ell'$ on the abelian group $A$ are called Hausdorff equivalent if there is a constant $k > 0$ so that $|\ell(a) - \ell'(a)| \leq k$ for all $a \in A$. The (pseudo)norm $\ell$ on the abelian group $A$ is called regular if $\ell(a^n) = n\ell(a)$ for all $a \in A$ and all positive natural numbers $n$. Let $\ell$ be a norm on the abelian group $A$. We define the regularization $R\ell$ of $\ell$ by

$$R\ell(a) = \lim_{n \to \infty} \frac{\ell(a^n)}{n}.$$ 

By [PS78], p. 23, Exercise 99, this limit always exists, and it is an exercise to see that $R\ell$ is a regular pseudonorm.
Lemma 4.2.1 The norm \( \ell \) on the abelian group \( A \) is regular iff \( R\ell = \ell \).

Proof. If \( \ell \) is regular, then clearly \( R\ell = \ell \). Conversely, if \( \ell(a^n) < n\ell(a) \) for some positive number \( n \) and some \( a \in A \), then
\[
R\ell(a) = \lim_{m \to \infty} \frac{\ell(a^{mn})}{mn} \leq \frac{\ell(a^n)}{n} < \ell(a),
\]
thus the lemma. □

In general positivity of \( R\ell \) fails, so it is possible that \( R\ell(a) = 0 \) but \( a \neq 0 \). Also it may easily happen that \( R\ell \) is not Hausdorff equivalent to \( \ell \). We give a criterion for positivity and Hausdorff equivalence in terms of Burago’s inequality [Gro93].

Definition 4.2.2 We say that a norm \( \ell \) on an abelian group \( A \) satisfies the Burago’s inequality if there exists a constant \( c = c(A) > 0 \) such that
\[
\ell(a^2) \geq 2\ell(a) - c \quad \text{for all} \quad a \in A.
\]
The norm is discrete if for all \( n \in \mathbb{N} \) the ball \( B_n = \{ x \in A : \ell(x) \leq n \} \) is finite. For example any word metric, corresponding to a finite generating set, is discrete.

Lemma 4.2.3 If a discrete norm \( \ell \) on a torsionfree abelian group \( A \) satisfies Burago’s inequality then its regularization \( R\ell \) is a norm also and, furthermore, \( R\ell \) is Hausdorff equivalent to \( \ell \).

Proof. By induction from Burago’s inequality we deduce that \( \ell(a^{2^n}) \geq 2^n\ell(a) - (2^n - 1)c \), for all \( a \in A, n \in \mathbb{N} \). This implies that
\[
\ell(a) \geq R\ell(a) = \lim_{m \to \infty} \frac{\ell(a^{2^m})}{2^m} \geq \ell(a) - c
\]
for all \( a \in A \). Thus the regularization \( R\ell \) is Hausdorff equivalent to \( \ell \). As any regularization, this pseudonorm is regular. It remains to prove that \( R\ell \) is a norm on \( A \), i.e., it does not vanish on nonzero \( a \in A \). If \( a \in A \) is such that \( \ell(a) \geq 1 + c \), then
\[
R\ell(a) = \lim_{m \to \infty} \frac{\ell(a^{2^m})}{2^m} \geq \ell(a) - c \geq 1.
\]
Now suppose \( a \in A \) is arbitrary nonzero, then by the discreteness assumption \( \ell(a^n) \geq 1+c \) for sufficiently large \( n \), and since \( R\ell \) is regular, \( R\ell(a) = \frac{1}{n}R\ell(a^n) > 0 \). □
4.3 Approximation and Burago’s inequality

**Lemma 4.3.1** Let $\Gamma$ be a finitely generated group of isometries of a proper CAT(0) space $X$, acting cocompactly and properly on $X$. Suppose that $x_0 \in X$ has a trivial stabilizer so that the Cayley graph $C$ of $\Gamma$ can be considered as embedded into $X$ via the orbit map $g \mapsto gx_0 (g \in \Gamma)$. Suppose that the pair $(X, \Gamma x_0)$ satisfies the approximation property. Then the restriction of the word length $\ell$ on $\Gamma$ to any finitely generated free abelian subgroup $A$ satisfies the Burago’s inequality.

**Proof.** By assumption there is a $\delta > 0$ such that for any $g \in \Gamma$ the $X$-geodesic $\alpha_X$ from $x_0$ to $gx_0$ and some $C$–geodesic $\alpha_C$ from $x_0$ to $gx_0$ are $\delta$–close to each other. By the flat torus theorem [Bow95], [Bri95] there is a Euclidean subspace $F$ in $X$ on which $A$ acts by translation. Fix the point $y_0 \in F$ and let $a$ be an arbitrary nontrivial element in $A$. We will show that $ax_0$ is contained in a $c$-neighborhood of $\alpha_C$ for a suitable $c > 0$. Clearly $d_X(a^2x_0, a^2y_0) = d_X(x_0, y_0)$. Parameterize the segments $[x_0, a^2x_0], [y_0, a^2y_0]$ by the segment $[0, 1]$ proportionally to arc length. It follows from the convexity of $X$–metric that the corresponding points on the segments are distance at most $d_X(x_0, y_0)$ from each other. Let $u$ be the point on $[x_0, a^2x_0]$ corresponding to the point $ay_0$. By assumption $u$ is distance at most $\delta$ from some point $v$ on $\alpha_C$. Hence we have bounded the $X$-distance from $ax_0 \in C$ to $v \in C$. (This key observation is illustrated in Figure 4). Since the Cayley graph $C$ is quasimetric to $X$ this bounds the Cayley graph distance also. Thus, there is a constant $c = c(\Gamma, X) > 0$ such that $d_C(ax_0, v) \leq c$. We have $\ell(a) = d_C(x_0, v) + d_C(v, a^2x_0) \geq (d_C(x_0, ax_0) - c) + (d_C(ax_0, a^2x_0) - c) = (\ell(a) - c) + (\ell(a) - c) = 2\ell(a) - 2c$, that is the Burago’s inequality. □

![Figure 4: Lemma 4.3.1](image)

**Theorem 4.3.2** Let $(W, S)$ be a Coxeter group and let $\ell$ be the word length in generators $S$. Then the restriction of $\ell$ to any free abelian subgroup $A$ of $W$ is Hausdorff equivalent to a regular norm on $A$.

**Proof.** Consider the pair $(\mathcal{M}_W, C_W)$ where the Cayley graph $C_W$ is embedded into the Moussong complex as an orbit $Wx_0$. By Theorem 3.3.2 $(\mathcal{M}_W, C_W)$
satisfies the approximation property. Therefore by Lemma 4.3.1 the restriction of the word length $\ell$ on $W$ to any finitely generated free abelian subgroup $A$ satisfies the Burago’s inequality. Finally, by Lemma 4.2.3 $\ell$ is Hausdorff equivalent to its regularization $R\ell$ and thus $R\ell$ is the required norm on $A$.

I am grateful to Herbert Abels for asking the question that leads to the theorem above.

References

[Bal95] Ballmann, Werner. *Lectures on spaces of nonpositive curvature*. With an appendix by Misha Brin: Ergodicity of geodesic flows. DMV Seminar. Bd. 25. Basel: Birkhaeuser Verlag, 1995.

[BB97] Ballmann, Werner. *Rank rigidity of Euclidean polyhedra*. American Journal of Mathematics, 122(5):873–885, 2000.

[BH99] Martin Bridson and Andre Haefliger. *Metric Spaces of Non-Positive Curvature*, volume 319 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Berlin, Heidelberg, first edition, 1999.

[Bow95] B.H Bowditch. *Minkowskian subspaces of non-positively curved metric spaces*. Bull. London Math. Soc., 27(6):575–584, 1995.

[Bri95] M. Bridson. *On the existence of flat planes in spaces of non-positive curvature*. Proc. Am. Math. Soc., 123(1):223–235, 1995.

[Bro96] Kenneth S. Brown. *Buildings*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, Heidelberg, second edition, 1996.

[Gro87] M. Gromov. *Hyperbolic groups*. In *Essays in group theory*, volume 8 of *Math. Sci. Res. Inst. Publ.*, pages 75–263. Springer, New York, 1987.

[Gro93] M. Gromov. *Asymptotic invariants of infinite groups*. In *Geometric group theory, Vol. 2 (Sussex, 1991)*, volume 182 of *London Math. Soc. Lecture Note Ser.*, pages 1–295. Cambridge Univ. Press, Cambridge, 1993.

[Hum90] James E. Humphreys. *Reflection groups and Coxeter groups*. Cambridge Studies in Advanced Mathematics, 29. Cambridge University Press, 1990.

[Kap97] Ilya Kapovich. Small cancellation groups and translation numbers. *Trans. Amer. Math. Soc.*, 349(5):1851–1875, 1997.

[Kra94] Krammer Daan. *The conjugacy problem for Coxeter groups*. PhD thesis, Utrecht University, 1994.
[Mou88] Moussong Gabor. *Hyperbolic Coxeter groups*. PhD thesis, Ohio State University, 1988.

[NR98] G. A. Niblo and L. D. Reeves. *The geometry of cube complexes and the complexity of their fundamental groups*. Topology, 37(3):621–633, 1998.

[PS78] George Polia and Gabor Szege. *Problems and theorems in analysis. I*, volume 193 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1978. Series, integral calculus, theory of functions, Translated from the German by D. Aeppli, Corrected printing of the revised translation of the fourth German edition.

G. A. Noskov
Sobolev Institute of Mathematics,
Omsk 644099, Pevtsova 13,
Russia
and
Fakultät für Mathematik,
Universität Bielefeld,
Postfach 100131, D-33501
Bielefeld, Germany
Abstract. We show that there exists a fine moduli space for torsion-free sheaves on a projective surface which have a “good framing” on a big and nef divisor. This moduli space is a quasi-projective scheme. This is accomplished by showing that such framed sheaves may be considered as stable pairs in the sense of Huybrechts and Lehn. We characterize the obstruction to the smoothness of the moduli space and discuss some examples on rational surfaces.

2010 Mathematics Subject Classification: 14D20; 14D21; 14J60

Keywords and Phrases: Framed sheaves, moduli spaces, stable pairs, instantons

1. Introduction

There has been recently some interest in the moduli spaces of framed sheaves. One reason is that they are often smooth and provide desingularizations of the moduli spaces of ideal instantons, which in turn are singular [17, 19, 18]. For this reason, their equivariant cohomology under suitable toric actions is relevant to the computation of partition functions, and more generally expectation values of quantum observables in topological quantum field theory [20, 2, 19, 6, 3]. On the other hand, these moduli spaces can be regarded as higher-rank generalizations of Hilbert schemes of points, and as such they have interesting connections with integrable systems [12, 1], representation theory [26], etc.

While it is widely assumed that such moduli spaces exist and are well behaved, an explicit analysis, showing that they are quasi-projective schemes and are fine moduli spaces, is missing in the literature. In the present paper we provide such a construction for the case of framed sheaves on smooth projective surfaces under some mild conditions. We show that if $D$ is a big and nef curve in a smooth projective surface $X$, there is a fine quasi-projective moduli space for...
sheaves that have a “good framing” on $D$ (Theorem 3.1). The point here is that the sheaves under consideration are not assumed a priori to be semistable, and the basic idea is to show that there exists a stability condition making all of them stable, so that our moduli space is an open subscheme of the moduli space of stable pairs in the sense of Huybrechts and Lehn [8, 9].

In the papers [21, 22] T. Nevins constructed a scheme structure for these moduli spaces, however we obtain a stronger result, showing that these schemes are quasi-projective, and in particular are separated and of finite type. Moreover we compute the obstruction to the smoothness of these moduli spaces (Theorem 4.3). In fact, the tangent space is well known, but we provide a more precise description of the obstruction space than the one given by Lehn [14]. We show that it lies in the kernel of the trace map, thus extending a previous result of Lübke [15] to the non-locally free case.

In some cases there is another way to give the moduli spaces $\mathcal{M}(r,c,n)$ a structure of algebraic variety, namely, by using ADHM data. This was done for vector bundles on $\mathbb{P}^2$ by Donaldson [5], while (always in the locally free case) the case of the blowup of $\mathbb{P}^2$ at a point is studied in A. King’s thesis [13], and $\mathbb{P}^2$ blown-up at an arbitrary number of points was analyzed by Buchdahl [4]. The general case (i.e., including torsion-free sheaves) is studied by C. Rava for Hirzebruch surfaces [24] and A.A. Henni for multiple blowups of $\mathbb{P}^2$ at distinct points [7]. The equivalence between the two approaches follows from the fact that in both cases one has fine moduli spaces. On the ADHM side, this is shown by constructing a universal monad on the moduli space [23, 7, 25].

In the final section we discuss some examples, i.e. framed bundles on Hirzebruch surfaces with “minimal invariants”, and rank 2 framed bundles on the blowup of $\mathbb{P}^2$ at one point.

In the present article, all the schemes we consider are separated and are of finite type over $\mathbb{C}$, and “a variety” is a reduced irreducible scheme of finite type over $\mathbb{C}$. A “sheaf” is always coherent, the term “(semi)stable” always means “$\mu$-(semi)stable”, and the prefix $\mu$- will be omitted. Framed sheaves are always assumed to be torsion-free.

### 2. Framed sheaves

Let us characterize the objects that we shall study.

**Definition 2.1.** Let $X$ be a scheme over $\mathbb{C}$, $D \subset X$ an effective Weil divisor, and $E_D$ a sheaf on $D$. We say that a sheaf $E$ on $X$ is $(D, E_D)$-framable if $E$ is torsion-free and there is an epimorphism $E \rightarrow E_D$ of $O_X$-modules inducing an isomorphism $E|_D \xrightarrow{\sim} E_D$. An isomorphism $\phi: E|_D \xrightarrow{\sim} E_D$ will be called a $(D, E_D)$-framing of $E$. A framed sheaf is a pair $(E, \phi)$ consisting of a $(D, E_D)$-framable sheaf $E$ and a framing $\phi$. Two framed sheaves $(E, \phi)$ and $(E', \phi')$ are isomorphic if there is an isomorphism $f: E \rightarrow E'$ and a nonzero constant $\lambda \in \mathbb{C}$ such that $\phi' \circ f|_D = \lambda \phi$. 

**Documenta Mathematica 16 (2011) 399–410**
Let us remark that our notion of framing is the same as the one used in [14, 22, 21], but is more restrictive than that of [8], where a framing is any homomorphism $\alpha : E \to \mathcal{E}_D$ of $\mathcal{O}_X$-modules, not necessarily factoring through an isomorphism $\mathcal{E}_D \cong \mathcal{E}_D$. To distinguish between the two definitions, we will call such a pair $(\mathcal{E}, \alpha)$ a framed pair, whilst the term framed sheaf will refer to the notion introduced in Definition 2.1

Our strategy to show that framed sheaves on a projective variety make up "good" moduli spaces will consist in proving that, under some conditions, the framed sheaves $(\mathcal{E}, \phi)$ are stable according to a notion of stability introduced by Huybrechts and Lehn [8, 9]. The definition of stability for framed pairs depends on the choice of a polarization $H$ on $X$ and a positive real number $\delta$ (in our notation, $\delta$ is the leading coefficient of the polynomial $\delta$ in the definition of (semi)stability in [8]).

**Definition 2.2 ([8, 9]).** A framed pair $(\mathcal{E}, \alpha)$ on an $n$-dimensional projective variety $X$, consisting of a torsion-free sheaf $\mathcal{E}$ and its framing $\alpha : \mathcal{E} \to \mathcal{E}_D$, is said to be $(H, \delta)$-stable, if for any subsheaf $\mathcal{G} \subset \mathcal{E}$ with $0 \leq \text{rk} \mathcal{G} \leq \text{rk} \mathcal{E}$, the following inequalities hold:

1. $\frac{c_1(\mathcal{G}) \cdot H^{n-1}}{\text{rk}(\mathcal{G})} < \frac{c_1(\mathcal{E}) \cdot H^{n-1} - \delta}{\text{rk}(\mathcal{E})}$ when $\mathcal{G}$ is contained in $\ker \alpha$;
2. $\frac{c_1(\mathcal{G}) \cdot H^{n-1} - \delta}{\text{rk}(\mathcal{G})} < \frac{c_1(\mathcal{E}) \cdot H^{n-1} - \delta}{\text{rk}(\mathcal{E})}$ otherwise.

Remark, that according to this definition, any rank-1 framed sheaf is $(H, \delta)$-stable for any ample $H$ and any $0 \leq \delta < \deg D$.

For any sheaf $\mathcal{F}$ on $X$, $P^H_\mathcal{F}(k)$ denotes the Hilbert polynomial $P^H_\mathcal{F}(k) = \chi(\mathcal{F} \otimes \mathcal{O}_X(kH))$. For a non-torsion sheaf $\mathcal{F}$ on $X$, $\mu^H(\mathcal{F})$ denotes the slope of $\mathcal{F}$: $\mu^H(\mathcal{F}) = \frac{c_1(\mathcal{F}) \cdot H^{n-1}}{\text{rk} \mathcal{F}}$.

**Theorem 2.3 ([8, 9]).** Let $X$ be a smooth projective variety, $H$ an ample divisor on $X$ and $\delta$ a positive real number. Let $D \subset X$ be an effective divisor, and $\mathcal{E}_D$ a sheaf on $D$. Then there exists a fine moduli space $\mathcal{M} = \mathcal{M}_X^H(P)$ of $(H, \delta)$-stable $(D, \mathcal{E}_D)$-framed sheaves $(\mathcal{E}, \phi)$ with fixed Hilbert polynomial $P = P^H_D$, and this moduli space is a quasi-projective scheme.

Since we are using slope stability, and a more restrictive definition of framing with respect to that of [8, 9], our moduli space $\mathcal{M}_X^H(P)$ is actually an open subscheme of the moduli space constructed by Huybrechts and Lehn.

Another general result on framed sheaves we shall need is a boundedness theorem due to M. Lehn. Given $X$, $D$, $\mathcal{E}_D$ as above, a set $\mathcal{M}$ of $(D, \mathcal{E}_D)$-framed pairs $(\mathcal{E}, \phi)$ is bounded if there exists a scheme of finite type $S$ over $\mathbb{C}$ together with a family $(\mathcal{G}, \phi)$ of $(D, \mathcal{E}_D)$-framed pairs over $S$ such that for any $(\mathcal{E}, \phi) \in \mathcal{M}$, there exist $s \in S$ and an isomorphism $(\mathcal{G}_s, \phi_{s(D \times s)}) \cong (\mathcal{E}, \phi)$.

**Definition 2.4.** Let $X$ be a smooth projective variety. An effective divisor $D$ on $X$ is called a good framing divisor if we can write $D = \sum n_i D_i$, where
$D_i$ are prime divisors and $n_i > 0$, and there exists a nef and big divisor of the form $\sum a_i D_i$ with $a_i \geq 0$. For a sheaf $\mathcal{E}_D$ on $D$, we shall say that $\mathcal{E}_D$ is a good framing sheaf, if it is locally free and there exists a real number $A_0$, $0 \leq A_0 < \frac{1}{r} D^2 \cdot H^{n-2}$, such that for any locally free subsheaf $\mathcal{F} \subset \mathcal{E}_D$ of constant positive rank, $\frac{1}{rn} \deg c_1(\mathcal{F}) \leq \frac{1}{rn} \deg c_1(\mathcal{E}_D) + A_0$.

**Theorem 2.5.** Let $X$ be a smooth projective variety of dimension $n \geq 2$, $H$ an ample divisor on $X$, $D \subset X$ an effective divisor, and $\mathcal{E}_D$ a vector bundle on $D$. Assume that $D$ is a good framing divisor. Then for every polynomial $P$ with coefficients in $\mathbb{Q}$, the set of torsion-free sheaves $\mathcal{E}$ on $X$ that satisfy the conditions $P^H = P$ and $\mathcal{E}_D \simeq \mathcal{E}_D$ is bounded.

This is proved in [14], Theorem 3.2.4, for locally free sheaves, but the proof goes through also in the torsion-free case, provided that $\mathcal{E}_D$ is locally free, as we are assuming.

### 3. Quasi-projective moduli spaces

Using the notions introduced in the previous section, we now can state the main existence result for quasi-projective moduli spaces:

**Theorem 3.1.** Let $X$ be a smooth projective surface, $D \subset X$ a big and nef curve, and $\mathcal{E}_D$ a good framing sheaf on $D$. Then for any $c \in H^*(X, \mathbb{Q})$, there exists an ample divisor $H$ on $X$ and a real number $\delta > 0$ such that all the $(D, \mathcal{E}_D)$-framed sheaves $\mathcal{E}$ on $X$ with Chern character $c(\mathcal{E}) = c$ are $(H, \delta)$-stable, so that there exists a quasi-projective scheme $\mathcal{M}_X(c)$ which is a fine moduli space for these framed sheaves.

**Proof.** Let us fix an ample divisor $C$ on $X$. Set $\mathcal{O}_X(k) = \mathcal{O}_X(kC)$ and $\mathcal{E}(k) = \mathcal{E} \otimes \mathcal{O}_X(k)$ for any sheaf $\mathcal{E}$ on $X$ and for any $k \in \mathbb{Z}$. Recall that the Castelnuovo-Mumford regularity $\rho(\mathcal{E})$ of a sheaf $\mathcal{E}$ on $X$ is the minimal integer $m$ such that $h^i(X, \mathcal{E}(m - i)) = 0$ for all $i > 0$. According to Lehn’s Theorem (Theorem 2.5), the family $\mathcal{M}$ of all the sheaves $\mathcal{E}$ on $X$ with $\rho(\mathcal{E}) = c$ and $\mathcal{E}_D \simeq \mathcal{E}_D$ is bounded. Hence $\rho(\mathcal{E})$ is uniformly bounded over all $\mathcal{E} \in \mathcal{M}$. By Grothendieck’s Lemma (Lemma 1.7.9 in [10]), there exists $A_1 \geq 0$, depending only on $\mathcal{E}_D$, $c$ and $C$, such that $\mu(\mathcal{F}) \leq \mu(\mathcal{E}) + A_1$ for all $\mathcal{E} \in \mathcal{M}$ and for all nonzero subsheaves $\mathcal{F} \subset \mathcal{E}$.

For $n > 0$, denote by $H_n$ the ample divisor $C + nD$. We shall verify that there exists a positive integer $n$ such that the range of positive real numbers $\delta$, for which all the framed sheaves $\mathcal{E}$ from $\mathcal{M}$ are $(H_n, \delta)$-stable, is nonempty.

Let $\mathcal{F} \subset \mathcal{E}$, $0 < r' = \text{rk} \mathcal{F} \leq r = \text{rk} \mathcal{E}$. Assume first that $\mathcal{F} \notin \ker (\mathcal{E} \to \mathcal{E}_D)$. Then we may only consider the case $r' < r$, and the $(H_n, \delta)$-stability condition for $\mathcal{E}$ reads:

\begin{equation}
\mu^{H_n}(\mathcal{F}) < \mu^{H_n}(\mathcal{E}) + \left( \frac{1}{r'} - \frac{1}{r} \right) \delta.
\end{equation}
Saturating $F$, we make $\mu^{H_n}(F)$ bigger, so we may assume that $F$ is a saturated subsheaf of $E$, and hence that it is locally free. Then $F_{|D} \subset E_{|D}$ and we have:

$$\mu^{H_n}(F) = \frac{n}{r} \deg c_1(F_{|D}) + \mu^C(F) \leq \mu^{H_n}(E) + nA_0 + A_1.$$  

Thus we see that (2) implies (1) whenever

$$\frac{rr'}{r - r'}(nA_0 + A_1) < \delta.$$

Assume now that $F$ is a saturated, and hence a locally free subsheaf of $\ker(E \to E_{|D}) \simeq E(-D)$. Then the $(H_n, \delta)$-stability condition for $E$ is

$$\mu^{H_n}(F) < \frac{1}{r} \delta,$$

and the inclusion $F(D) \subset E$ yields:

$$\mu^{H_n}(F) < \mu^{H_n}(E) - H_nD + nA_0 + A_1 = \mu^{H_n}(E) - (D^2 - A_0)n + A_1 - DC.$$

We see that (5) implies (4) whenever

$$\delta < r[(D^2 - A_0)n - A_1 + DC].$$

The inequalities (3), (6) for all $r' = 1, \ldots, r - 1$ have a nonempty interval of common solutions $\delta$ if

$$n > \max \left\{ \frac{rA_1 - CD}{D^2 - rA_0}, 0 \right\}.$$ 

\[\square\]

**Remark 3.2.** Grothendieck’s Lemma is stated in [10] in terms of the so-called $\hat{\mu}$ slope. However, for torsion-free sheaves, the $\hat{\mu}$ slope and the usual slope differ by constants depending only on $(X, O_X(1))$, see Definition 1.6.8 in [10] and the following remark. \[\triangle\]

Note that up to isomorphism, the quasi-projective structure making $\mathfrak{M}_X(c)$ a fine moduli space is unique, which follows from the existence of a universal family of framed sheaves over it.

If $D$ is a smooth and irreducible curve and $D^2 > 0$, then our definition of a good framing sheaf with $A_0 = 0$ is just the definition of semistability. The following is thus an immediate consequence of the theorem:

**Corollary 3.3.** Let $X$ be a smooth projective surface, $D \subset X$ a smooth, irreducible, big and nef curve, and $E_D$ a semistable vector bundle on $D$. Then for any $c \in H^*(X, \mathbb{Q})$, there exists a quasi-projective scheme $\mathfrak{M}_X(c)$ which is a fine moduli space of $(D, E_D)$-framed sheaves on $X$ with Chern character $c$. 

**Documenta Mathematica 16 (2011) 399–410**
4. Infinitesimal study

Let $X$ be a smooth projective variety, $D$ an effective divisor on $X$, $\mathcal{E}_D$ a vector bundle on $D$. We shall consider sheaves $\mathcal{E}$ on $X$ framed to $\mathcal{E}_D$ on $D$. We recall the notion of a simplifying framing bundle introduced by Lehn.

**Definition 4.1.** $\mathcal{E}_D$ is simplifying if for any two vector bundles $\mathcal{E}$, $\mathcal{E}'$ on $X$ such that $\mathcal{E}|_D \simeq \mathcal{E}'|_D \simeq \mathcal{E}_D$, the group $H^0(X, \text{Hom}(\mathcal{E}, \mathcal{E}')(-D))$ vanishes.

An easy sufficient condition for $\mathcal{E}_D$ to be simplifying is

$$H^0(D, \text{End}(\mathcal{E}_D) \otimes O_X(-kD)|_D) = 0$$

for all $k > 0$.

Lehn [14] proved that if $D$ is good and $\mathcal{E}_D$ is simplifying, there exists a fine moduli space $\mathcal{M}$ of $(D, \mathcal{E}_D)$-framed vector bundles on $X$ in the category of separated algebraic spaces. Lübke [15] proved a similar result: if $X$ is a compact complex manifold, $D$ a smooth hypersurface (not necessarily “good”) and if $\mathcal{E}_D$ is simplifying, then the moduli space $\mathcal{M}$ of $(D, \mathcal{E}_D)$-framed vector bundles exists as a Hausdorff complex space. In both cases the tangent space $T_{[\mathcal{E}]}\mathcal{M}$ at a point representing the isomorphism class of a framed bundle $\mathcal{E}$ is naturally identified with $H^1(X, \text{End}(\mathcal{E})(-D))$, and the moduli space is smooth at $[\mathcal{E}]$ if $H^2(X, \text{End}(\mathcal{E})(-D)) = 0$. Lübke gives a more precise statement about smoothness: $[\mathcal{E}]$ is a smooth point of $\mathcal{M}$ if $H^2(X, \text{End}(\mathcal{E})(-D)) = 0$, where $\text{End}_0$ denotes the traceless endomorphisms. Huybrechts and Lehn in [9] define the tangent space and give a smoothness criterion for the moduli space of stable pairs that are more general objects than our framed sheaves. In this section, we adapt Lübke’s criterion to our moduli space $\mathcal{M}_X(c)$, parametrizing not only vector bundles, but also some non-locally-free sheaves. When we work with stable framed sheaves, we do not need the assumption that $\mathcal{E}_D$ is simplifying.

We shall use the notions of the trace map and traceless exts, see Definition 10.1.4 from [10]. Assuming $X$ is a smooth algebraic variety, $\mathcal{F}$ any (coherent) sheaf on it, and $\mathcal{N}$ a locally free sheaf (of finite rank), the trace map is defined

$$\text{tr} : \text{Ext}^i(\mathcal{F}, \mathcal{F} \otimes \mathcal{N}) \to H^i(X, \mathcal{N}), \quad i \in \mathbb{Z},$$

and the traceless part of the ext-group, denoted by $\text{Ext}^i(\mathcal{F}, \mathcal{F} \otimes \mathcal{N})_0$, is the kernel of this map.

We shall need the following property of the trace:

**Lemma 4.2.** Let $0 \to \mathcal{F} \xrightarrow{\alpha} \mathcal{G} \xrightarrow{\beta} \mathcal{E} \to 0$ be an exact triple of sheaves and $\mathcal{N}$ a locally free sheaf. Then there are two long exact sequences of ext-functors giving rise to the natural maps

$$\mu_i : \text{Ext}^i(\mathcal{F}, \mathcal{E} \otimes \mathcal{N}) \to \text{Ext}^{i+1}(\mathcal{E}, \mathcal{E} \otimes \mathcal{N})$$

and

$$\tau_i : \text{Ext}^i(\mathcal{F}, \mathcal{E} \otimes \mathcal{N}) \to \text{Ext}^{i+1}(\mathcal{F}, \mathcal{F} \otimes \mathcal{N}),$$

and we have $\text{tr} \circ \mu_i = (-1)^i \text{tr} \circ \tau_i$ as maps $\text{Ext}^i(\mathcal{F}, \mathcal{E} \otimes \mathcal{N}) \to H^{i+1}(X, \mathcal{N})$.

**Proof.** This is a particular case of the graded commutativity of the trace with respect to cup-products on Homs in the the derived category (see Section...
V.3.8 in [11]): if $\xi \in \text{Hom}(\mathcal{F}, \mathcal{E} \otimes N[i])$, $\eta \in \text{Hom}(\mathcal{F}, \mathcal{F}[j])$, then $\text{tr}(\xi \circ \eta) = (-1)^{ij} \text{tr}((\eta \circ \text{id}_N) \circ \xi)$. This should be applied to $\xi \in \text{Hom}(\mathcal{F}, \mathcal{E} \otimes N[i])$ and $\eta = \partial \in \text{Hom}(\mathcal{E}, \mathcal{F}[1])$, where $\partial$ is the connecting homomorphism in the distinguished triangle associated to the given exact triple:

$$\mathcal{E}[-1] \xrightarrow{-\partial} \mathcal{F} \cong \mathcal{G} \xrightarrow{\partial} \mathcal{E} \xrightarrow{\partial} \mathcal{F}[1].$$

\[ \Box \]

**Theorem 4.3.** Let $X$ be a smooth projective surface, $D \subset X$ an effective divisor, $\mathcal{E}_D$ a locally free sheaf on $D$, and $c \in H^1(X, \mathbb{Q})$ the Chern character of a $(D, \mathcal{E}_D)$-framed sheaf $\mathcal{E}$ on $X$. Assume that there exists an ample divisor $H$ on $X$ and a positive real number $\delta$ such that $\mathcal{E}$ is $(H, \delta)$-stable, and denote by $\mathfrak{M}_X(c)$ the moduli space of $(D, \mathcal{E}_D)$-framed sheaves on $X$ with Chern character $c$ which are $(H, \delta)$-stable. Then the tangent space to $\mathfrak{M}_X(c)$ is given by

$$T_{[c]}\mathfrak{M}_X(c) = \text{Ext}^1(\mathcal{E}, \mathcal{E} \otimes O_X(-D)),$$

and $\mathfrak{M}_X(c)$ is smooth at $[\mathcal{E}]$ if the traceless ext-group

$$\text{Ext}^2(\mathcal{E}, \mathcal{E} \otimes O_X(-D)) = \ker \left[ \text{tr} : \text{Ext}^2(\mathcal{E}, \mathcal{E} \otimes O_X(-D)) \to H^2(X, O(-D)) \right]$$

vanishes.

**Proof.** We prove this result by a combination of arguments of Huybrechts-Lehn and Mukai, so we just give a sketch, referring to [9, 16] for details. As in Section 4.iv) of [9], the smoothness of $\mathfrak{M} = \mathfrak{M}_X(c)$ follows from the $T^1$-lifting property for the complex $\mathcal{E} \to \mathcal{E}_D$.

Let $A_n = k[t]/(t^{n+1})$, $X_n = X \times \text{Spec} A_n$, $D_n = D \times \text{Spec} A_n$, $\mathcal{E}_{D_n} = \mathcal{E}_D \boxtimes A_n$, and let $\mathcal{E}_n \xrightarrow{\alpha_n} \mathcal{E}_{D_n}$ be an $A_n$-flat lifting of $\mathcal{E} \to \mathcal{E}_D$ to $X_n$. Then the infinitesimal deformations of $\alpha_n$ over $k[t]/(t^2)$ are classified by the hyper-ext $\text{Ext}^1(\mathcal{E}_n, \mathcal{E}_n \alpha_n \boxtimes \mathcal{E}_{D_n})$, and one says that the $T^1$-lifting property is verified for $\mathcal{E} \to \mathcal{E}_D$ if all the natural maps

$$T_n^1 : \text{Ext}^1(\mathcal{E}_n, \mathcal{E}_n \alpha_n \boxtimes \mathcal{E}_{D_n}) \to \text{Ext}^1(\mathcal{E}_{n-1}, \mathcal{E}_{n-1} \alpha_{n-1} \boxtimes \mathcal{E}_{D_{n-1}})$$

are surjective whenever $(\mathcal{E}_n, \alpha_n) \equiv (\mathcal{E}_{n-1}, \alpha_{n-1}) \mod (t^n)$. In loc. cit., the authors remark that there is an obstruction map $ob$ on the target of $T_n^1$ which embeds the cokernel of $T_n^1$ into $\text{Ext}^2(\mathcal{E}, \mathcal{E} \to \mathcal{E}_D)$, so that if the latter vanishes, the $T^1$-lifting property holds.

In our case, $\mathcal{E}$ is locally free along $D$, so the complex $\mathcal{E} \to \mathcal{E}_D$ is quasi-isomorphic to $\mathcal{E}(-D)$ and $\text{Ext}^1(\mathcal{E}, \mathcal{E} \to \mathcal{E}_D) = \text{Ext}^1(\mathcal{E}, \mathcal{E}(-D))$. It remains to prove that the image of $ob$ is contained in the traceless part of $\text{Ext}^2(\mathcal{E}, \mathcal{E}(-D))$. This is done by a modification of Mukai’s proof in the non-framed case.

First we assume that $\mathcal{E}$ is locally free. Then the elements of $\text{Ext}^1(\mathcal{E}_{n-1}, \mathcal{E}_{n-1}(-D_{n-1}))$ can be given by Čech 1-cocycles with values in $\mathcal{E}nd(\mathcal{E}_{n-1})(-D_{n-1})$ for some open covering of $X$, and the image of such a 1-cocycle $(a_{ij})$ under the obstruction map $\text{Ext}^1(\mathcal{E}_{n-1}, \mathcal{E}_{n-1}(-D_{n-1})) \to$
Ext²(ℰ, ℰ(−D)) is a Čech 2-cocycle (cijk) with values in \( End(ℰ)(−D) \). A direct calculation shows that \( tr(cijk) \) is a Čech 2-cocycle with values in \( \mathcal{O}_X(−D) \) which is the obstruction to the lifting of the infinitesimal deformation of the framed line bundle \( det \mathcal{E}_{n−1} \) from \( A_{n−1} \) to \( A_n \). As we know that the moduli space of line bundles, whether framed or not, is smooth, this obstruction vanishes, so the cocycle \( tr(cijk) \) is cohomologous to 0.

Now consider the case when \( \mathcal{E} \) is not locally free. Replacing \( \mathcal{E}, \mathcal{E}_D \) by their twists \( \mathcal{E}(n), \mathcal{E}_D(n) \) for some \( n > 0 \), we may assume that \( H^i(X, \mathcal{E}) = H^i(X, \mathcal{E}(−D)) = 0 \) for \( i = 1, 2 \) and that \( \mathcal{E} \) is generated by global sections. Then we get the exact triple of framed sheaves

\[
0 \to (G, \gamma) \to (H^0(X, \mathcal{E}) \otimes \mathcal{O}_X, \beta) \to (\mathcal{E}, \alpha) \to 0,
\]

where \( G \) is locally free (at this point it is essential that \( dim X = 2 \) and \( X \) is smooth). Then we verify the \( T^1 \)-lifting property for the exact triples

\[
0 \to (G_n, \gamma_n) \to (\mathcal{O}_{X_n}, \beta_n) \to (\mathcal{E}_n, \alpha_n) \to 0.
\]

The infinitesimal deformations of such exact triples are classified by \( Hom(G_n, \mathcal{E}_n(−D_n)) \), and the obstructions lie in \( Ext^1(G_n, \mathcal{E}_n(−D_n)) \). We have two connecting homomorphisms \( \mu_1 : Ext^1(G_n, \mathcal{E}_n(−D_n)) \to Ext^2(\mathcal{E}_n, \mathcal{E}_n(−D_n)) \) and \( \tau_1 : Ext^1(G_n, \mathcal{E}_n(−D_n)) \to Ext^2(G_n, \mathcal{E}_n(−D_n)) \). Our hypotheses on \( \mathcal{E} \) imply that: 1) every infinitesimal deformation of \( (\mathcal{E}_n, \alpha_n) \) lifts to that of the triple, and 2) \( \mu_1 \) is an isomorphism, that is, the infinitesimal deformation of \( \mathcal{E}_n \) is unobstructed if and only if that of the triple is. By Lemma 4.2, \( tr(\mu_1(\xi)) = − \text{tr}(\tau_1(\xi)) \) in \( H^2(X, \mathcal{O}_X(−D)) \). As in 1.10 of [16], \( \tau_1(\xi) \) is the obstruction \( ob(G_{n−1}, \gamma_{n−1}) \) to lifting \( (G_n, \gamma_n) \) from \( A_{n−1} \) to \( A_n \). As \( G_{n−1} \) is locally free, we can use the Čech cocycles as above and see that \( tr(\tau_1(\xi)) \in H^2(X, \mathcal{O}_X(−D)) \) is the obstruction to lifting \( (det G_{n−1}, det \gamma_{n−1}) \), hence it is zero and we are done. \( \Box \)

The following Corollary describes a situation where the moduli space \( \mathcal{M}_X(c) \) is smooth (hence, every connected component is a smooth quasi-projective variety).

**Corollary 4.4.** In addition to the hypothesis of Theorem 4.3, let us assume that \( D \) is irreducible, that \( (K_X + D) \cdot D < 0 \), and choose the framing bundle to be trivial. Then the moduli space \( \mathcal{M}_X(c) \) is smooth.

This happens for instance when \( X \) is a Hirzebruch surface, or the blowup of \( \mathbb{P}^2 \) at a number of distinct points, taking for \( D \) the inverse image of a generic line in \( \mathbb{P}^2 \) via the birational morphism \( X \to \mathbb{P}^2 \). In this case one can also compute the dimension of the moduli space, obtaining \( dim \mathcal{M}_X(c) = 2rn \), with \( r = rk(\mathcal{E}) \) and

\[
c_2(\mathcal{E}) - \frac{r−1}{2r}c_1(\mathcal{E})^2 = n \omega,
\]

where \( \omega \) is the fundamental class of \( X \). When \( X \) is the \( p \)-th Hirzebruch surface \( \mathbb{F}_p \) we shall denote this moduli space by \( \mathcal{M}^p(r, k, n) \) if \( c_1(\mathcal{E}) = kC \), where \( C \) is the unique curve in \( \mathbb{F}_p \) having negative self-intersection.
The next example shows that the moduli space may be nonsingular even if the group \( \text{Ext}^2(E, E \otimes \mathcal{O}_X(-D)) \) does not vanish.

**Example 4.5.** For \( r = 1 \) the moduli space \( \mathcal{M}(1, 0, n) \) is isomorphic to the Hilbert scheme \( X_0^{[n]} \) parametrizing length \( n \) 0-cycles in \( X_0 = X \setminus D \). Of course this space is a smooth quasi-projective variety of dimension \( 2n \). Indeed in this case the trace morphism \( \text{Ext}^2(E, E \otimes \mathcal{O}_X(-D)) \to H^2(X, \mathcal{O}(-D)) \) is an isomorphism.

5. **Examples**

5.1. **Bundles with small invariants on Hirzebruch surfaces.** Let \( X \) be the \( p \)-th Hirzebruch surface \( F_p \), and normalize the Chern character by twisting by powers of the line bundle \( \mathcal{O}_{F_p}(C) \) so that \( 0 \leq k \leq r - 1 \). It has been shown in [3] that the moduli space \( \mathcal{M}(r, k, n) \) is nonempty if and only if the bound

\[
 n \geq N = \frac{pk}{2r}(r - k)
\]

is satisfied. The moduli spaces \( \mathcal{M}(r, k, N) \) can be explicitly characterized: \( \mathcal{M}(r, k, N) \) is a rank \( k(r - k)(p - 1) \) vector bundle on the Grassmannian \( G(k, r) \) of \( k \)-planes in \( C^r \); in particular, \( \mathcal{M}(1, k, N) \simeq G(k, r) \), and \( \mathcal{M}(2, k, N) \) is isomorphic to the tangent bundle of \( G(k, r) \). This is consistent with instanton counting, which shows that the spaces \( \mathcal{M}(r, k, N) \) have the same Betti numbers as \( G(k, r) \) [3].

5.2. **Rank 2 vector bundles on \( F_1 \).** We study in some detail the moduli spaces \( \mathcal{M}(2, k, n) \). As [27] and [28] show, the non-locally free case turns out to be very complicated as soon as the value of \( n \) exceeds the rank. So we consider only locally free sheaves. To simplify notation we call this moduli space \( \hat{M}(k, n) \), where \( n \) denotes now the second Chern class. We normalize \( k \) so that it will assume only the values 0 and \(-1\). Moreover we shall denote by \( M(n) \) the moduli space of rank 2 bundles on \( \mathbb{P}^2 \), with second Chern class \( n \), that are framed on the “line at infinity” \( \ell_\infty \subset \mathbb{P}^2 \) (which we identify with the image of \( D \) via the blow-down morphism \( \pi: F_1 \to \mathbb{P}^2 \)).

Let us start with the case \( k = -1 \). We introduce a stratification on \( \hat{M}(-1, n) \) according to the splitting type of the bundles it parametrizes on the exceptional line \( E \subset F_1 \)

\[
\hat{M}(-1, n) = Z_0(-1, n) \supset Z_1(-1, n) \supset Z_2(-1, n) \supset \ldots
\]

defined as follows: if \( Z_0^0(-1, n) = Z_k(-1, n) \setminus Z_{k+1}(-1, n) \) then

\[
Z_0^0(-1, n) = \{ E \in \hat{M}(-1, n) \mid \mathcal{E}|_E \simeq \mathcal{O}_E(-k) \oplus \mathcal{O}_E(k + 1) \}.
\]

**Proposition 5.1.** There is a map

\[
F_1: \hat{M}(-1, n) \to \prod_{k=0}^n M(n-k)
\]
which restricted to the subset $Z^0_k(-1, n)$ yields a morphism

$$Z^0_k(-1, n) \to M(n - k)$$

whose fibre is an open set in $\text{Hom}(\sigma^* E_1, O_E(k))/\mathbb{C}^* \simeq \mathbb{P}^{2k+1}$, made by $k$-linear forms that have no common zeroes on the exceptional line.

**Proof.** We start by considering $Z^0_0(-1, n)$. The morphism $Z^0_0(-1, n) \to M(n)$ is given by $E_1 \mapsto E = (\pi_* E)^{**}$. The fibre of this morphism includes a $\mathbb{P}^1$. To show that this is indeed a $\mathbb{P}^1$-fibration we need to check that $E_1$ has no other deformations than those coming from the choice of a point in $M(n)$ and a point in this $\mathbb{P}^1$. This follows from the equalities

$$\dim \text{Ext}^1(E_1, E_1(-E)) = \dim \text{Ext}^1(E, E(-\ell_\infty)) + 1$$

$$\text{Ext}^2(E_1, E_1(-E)) = 0$$

Note that this result is compatible with the isomorphism $\mathfrak{M}^1(r, k, N) \simeq G(k, r)$ mentioned in Section 5.1.

In general, if $E_1 \in Z^0_k(-1, n)$ with $k \geq 1$, so that $E_1|E \simeq O_E(k+1) \oplus O_E(-k)$, the direct image $\pi_*(E_1(kE))$ is locally free. This defines the morphism $Z^0_k(-1, n) \to M(n - k)$.

We consider now the case $k = 0$. One has $Z^0_0(0, n) \simeq M(n)$. We study the other strata by reducing to the odd case. If $E_1 \in Z^0_k(0, n)$, there is a unique surjection $\alpha : E_1 \to O_E(-k)$; let $F$ be the kernel. Restricting $0 \to F \to E_1 \to O_E(-k) \to 0$ we get an exact sequence

$$0 \to O_E(1-k) \to F|E \to O_E(k) \to 0$$

so that

$$F|E \simeq O_E(a+1) \oplus O_E(-a) \quad \text{with} \quad -k \leq a \leq k - 1.$$ 

A detailed analysis shows that $a = k - 1$. As a result we have:

**Proposition 5.2.** For all $k \geq 1$ there is a morphism

$$Z^0_k(0, n) \to M(n - 2k + 1)$$

whose fibres have dimension $2k - 1$.

**References**

[1] D. Ben-Zvi and T. Nevins, *Flows of Calogero-Moser systems*, Int. Math. Res. Not. IMRN (2007), no. 23, Art. ID rnm105, 38 pp.

[2] U. Bruzzo, F. Fucito, J. F. Morales, and A. Tanzini, *Multi-instanton calculus and equivariant cohomology*, J. High Energy Phys. (2003), no. 5, 054, 24 pp. (electronic).

[3] U. Bruzzo, R. Poghossian, and A. Tanzini, *Poincaré polynomial of moduli spaces of framed sheaves on (stacky) Hirzebruch surfaces*, arXiv:0909.1458 [math.AG]. Comm. Math. Phys. (2011), to appear.
[4] N. P. Buchdahl, *Blowups and gauge fields*, Pacific J. Math. 196 (2000), pp. 69–111.
[5] S. K. Donaldson, *Instantons and geometric invariant theory*, Comm. Math. Phys. 93 (1984), pp. 453–460.
[6] E. Gasparim and C.-C. M. Liu, *The Nekrasov conjecture for toric surfaces*, Comm. Math. Phys. 293 (2010), pp. 661–700.
[7] A. A. Henri, *Monads for torsion-free sheaves on multi-blowups of the projective plane*, arXiv:0903.3190 [math.AG].
[8] D. Huybrechts and M. Lehn, *Stable pairs on curves and surfaces*, J. Alg. Geom. 4 (1995), pp. 67–104.
[9] ———, *Framed modules and their moduli*, Internat. J. Math. 6 (1995), pp. 297–324.
[10] ———, *The geometry of moduli spaces of sheaves*, 2nd edition, Cambridge University Press, Cambridge, 2010.
[11] L. Illusie, *Complexe cotangent et déformations. I*, vol. 239 of Lecture Notes in Mathematics, Springer-Verlag, Berlin-New York, 1971.
[12] A. Kapustin, A. Kuznetsov, and D. Orlov, *Noncommutative instantons and twistor transform*, Comm. Math. Phys. 221 (2001), pp. 385–432.
[13] A. King, *Instantons and holomorphic bundles on the blown-up plane*. Thesis, Oxford University, 1989.
[14] M. Lehn, *Modulrümme gerahmter Vektorbündel*. Ph.D. thesis, Bonn, 1992.
[15] M. Lübke, *The analytic moduli space of framed vector bundles*, J. Reine Angew. Math. 441 (1993), pp. 45–59.
[16] S. Mukai, *Symplectic structure of the moduli space of sheaves on an abelian or K3 surface*, Invent. Math. 77 (1984), pp. 101–116.
[17] H. Nakajima, *Lectures on Hilbert schemes of points on surfaces*, vol. 18 of University Lecture Series, American Mathematical Society, Providence, RI, 1999.
[18] H. Nakajima and K. Yoshioka, *Lectures on instanton counting*, in Algebraic structures and moduli spaces, vol. 38 of CRM Proc. Lecture Notes, Amer. Math. Soc., Providence, RI, 2004, pp. 31–101.
[19] ———, *Instanton counting on blowup. I. 4-dimensional pure gauge theory*, Invent. Math. 162 (2005), pp. 313–355.
[20] N. A. Nekrasov, *Seiberg-Witten prepotential from instanton counting*, Adv. Theor. Math. Phys. 7 (2003), pp. 831–864.
[21] T. A. Nevins, *Moduli spaces of framed sheaves on certain ruled surfaces over elliptic curves*, Int. J. Math. 13 (2002), pp. 1117–1151.
[22] ———, *Representability for some moduli stacks of framed sheaves*, Manuscr. Math. 109 (2002), pp. 85–91.
[23] C. Okonek, M. Schneider, and H. Spindler, *Vector bundles on complex projective spaces*, vol. 3 of Progress in Mathematics, Birkhäuser Boston, Mass., 1980.
[24] C. L. S. Rava, Ph.D. thesis, SISSA (Trieste), 2011.
[25] ———, *ADHM data for framed sheaves on Hirzebruch surfaces*. In preparation, 2011.
[26] F. Sala and P. Tortella, Representations of the Heisenberg algebra and moduli spaces of framed sheaves. arXiv:1004.2814 [math.AG].

[27] M. E. Sorokina, Birational properties of moduli spaces of rank 2 semistable sheaves on the projective plane. Ph.D. thesis (in Russian), Yaroslavl, 2006.

[28] A. S. Tikhomirov, On birational transformations of Hilbert schemes of an algebraic surface, Mat. Zametki 73 (2003), pp. 281–294.
Abstract. Based on previous works, we compare over a perfect field \( k \) the category of homotopy invariant sheaves with transfers introduced by V. Voevodsky and the category of cycle modules introduced by M. Rost: the former is a full subcategory of the latter. Using the recent construction by D.C. Cisinski and the author of a non effective version \( DM(k) \) of the category of motivic complexes, we show that cycle modules form the heart of a natural t-structure on \( DM(k) \), generalizing the homotopy t-structure on motivic complexes.

2010 Mathematics Subject Classification: 14F42, (14C15, 14C35)
Keywords and Phrases: motifs mixtes, complexes motiviques, modules de cycles, filtration par coniveau

Table des matières

Introduction ................................................. 412
Notations ................................................. 416

Partie I. Modules homotopiques et modules de cycles ................................................. 417
1. Modules homotopiques ........................................ 417
   1.1. Rappels sur les faisceaux avec transferts .............. 417
   1.2. Définition ........................................... 420
   1.3. Réalisation des motifs géométriques ................. 422
2. Modules de cycles ........................................ 424
   2.1. Rappels ........................................... 424
   2.2. Fonctorialité ........................................ 425
   2.3. Suite exacte de localisation ......................... 426
   2.4. Module homotopique associé ......................... 429

(1)L’auteur est partiellement financé par l’ANR, projet no. ANR-07-BLAN-0142 “Méthodes à la Voevodsky, motifs mixtes et Géométrie d’Arakelov”.

Documenta Mathematica 16 (2011) 411–455
INTRODUCTION.

Théorie de Voevodsky. — Dans sa théorie des complexes motiviques sur un corps parfait $k$, V. Voevodsky introduit le concept central de faisceau Nisnevich invariant par homotopie avec transferts, que nous appellerons simplement faisceau homotopique. Rappelons qu’un faisceau homotopique $F$ est un préfaisceau de groupes abéliens sur la catégorie des $k$-schémas algébriques lisses, fonctoriel par rapport aux correspondances finies à homotopie près, qui est un faisceau pour la topologie de Nisnevich. Un exemple central d’un tel faisceau est donné par le préfaisceau $\mathbb{G}_m$ qui à un schéma lisse $X$ associe le groupe des sections globales inversibles sur $X$. La catégorie des faisceaux homotopiques, notée ici $\mathcal{H}I(k)$, a de bonnes propriétés que l’on peut résumer essentiellement en disant que c’est une catégorie abélienne de Grothendieck, monoïdale symétrique fermée.

Un des points centraux de la théorie est la démonstration par Voevodsky que tout faisceau homotopique $F$ admet une résolution de Gersten\(^{(2)}\). Un cas
particulier de ce résultat est le fait que pour tout schéma lisse $X$, le groupe abélien $F(X)$ admet une résolution par un complexe de la forme:

$$\begin{align*}
(G) \quad C^*(X, \hat{F}_*) : \bigoplus_{x \in X^{(n)}} \hat{F}(\kappa(x)) & \rightarrow \ldots \bigoplus_{x' \in X^{(n)}} \hat{F}_{-n}(\kappa(x')) & \rightarrow \ldots
\end{align*}$$

Suivant Voevodsky, $F_{-n} = \text{Hom}_{\mathcal{H}_{I}(k)}(\mathbb{G}_{m}^{\otimes n}, F)$. On a noté $X^{(n)}$ l’ensemble des points de codimension $n$ de $X$. Pour un entier $r \geq 0$ et un point $x$ de $X$, $\hat{F}_r(\kappa(x))$ désigne la fibre du faisceau homotopique $F_r$ au point Nisnevich qui correspond au corps résiduel $\kappa(x)$, vu comme un corps de fonctions.

Un corollaire de cette résolution de Gersten est que les faisceaux homotopiques sont essentiellement déterminés par leurs fibres en un corps de fonctions. La question centrale de cet article est de savoir jusqu’à quel point ils le sont.

**Théorie de Rost.** — Pour définir un complexe de Gersten, du type $(G)$, on remarque qu’il faut essentiellement se donner un groupe abélien pour chaque corps résiduel d’un point de $X$. M. Rost axiomatise cette situation en introduisant les modules de cycles. Un module de cycles est un foncteur $\phi$ de la catégorie des corps de fonctions au-dessus de $k$ vers les groupes abéliens gradués, muni d’une fonctorialité étendue qui permet de définir un complexe $C^*(X, \phi)$ du type $(G)$. Pour avoir une idée de cette fonctorialité, le lecteur peut se référer aux propriétés de la K-théorie de Milnor – mais aussi à la théorie des modules galoisiens. Rost note l’analogie entre ce complexe et le groupe des cycles de $X$ – comme l’avaient fait Bloch et Quillen avant lui – et utilise le traitement de la théorie de l’intersection par Fulton pour montrer que la co-homologie du complexe, notée $A^*(X, \phi)$, est naturelle en $X$ par rapport aux morphismes de schémas lisses.

Une comparaison. — Répondant à la question finale du premier paragraphe, nous comparons la théorie de Rost et celle de Voevodsky. D’une manière vague, notre résultat principal affirme que l’association $F \mapsto \hat{F}_*$ définit un foncteur pleinement fidèle des faisceaux homotopiques dans les modules de cycles, avec pour quasi-inverse à gauche le foncteur $\phi \mapsto A^0(\ldots, \phi)$.

Pour être plus précis dans la formulation de ce résultat, on est conduit à élargir la catégorie des faisceaux homotopiques. On définit un module homotopique $F_*$ comme un faisceau homotopique $\mathbb{Z}$-gradué muni d’isomorphismes $\epsilon_n : F_n \rightarrow (F_{n+1})_{-1}$. La catégorie obtenue, notée $\mathcal{H}_{I}(k)$, est encore abélienne de Grothendieck, symétrique monoïdale fermée. De plus, elle contient comme sous-catégorie pleine la catégorie $HI(k)$ – si $F$ est un faisceau homotopique, le module homotopique associé a pour valeur $\mathbb{G}_{m}^{\otimes n} \otimes F$ (resp. $F_{-n}$) en degré $n \geq 0$ (resp. $n < 0$).

Dès lors, on peut montrer que le système $\hat{F}_*$ des fibres d’un module homotopique $F_*$ en un corps de fonctions définit un module de cycles. De plus, pour

traitée dans le cas de la K-théorie par H. Gillet (voir [GIL85]) puis étendue dans le cas des modules de cycles par M. Rost.

Documenta Mathematica 16 (2011) 411–455
tout module de cycles φ, le groupe $A^0(X, φ)$, dépendant fonctoriellement d’un schéma lisse $X$, définit un module homotopique.

**Théorème (cf. 3.7).** — Les deux associations décrites ci-dessus définissent des foncteurs quasi-inverses l’un de l’autre.

La résolution de Gersten obtenue par Voevodsky est maintenant équivalente au résultat suivant:

**Corollaire (cf. 3.12).** — Si $F_*$ est un module homotopique et $X$ un schéma lisse, $H^n(X, F_*) = A^n(X, \hat{F}_*)$.[3]

Notons que ce corollaire est étendu au cas singulier à la fin de l'article (Proposition 6.10). Cette extension nécessite d’interpréter le théorème 3.7 en termes motiviques.

**L’interprétation motivique.** — Rappelons qu’un complexe motivique suivant Voevodsky est un complexe[4] de faisceaux Nisnevich avec transferts dont les faisceaux de cohomologie sont des faisceaux homotopiques. La catégorie des complexes motiviques $DM^{eff}(k)$ porte ainsi naturellement une t-structure au sens de Beilinson, Bernstein et Deligne dont le cœur est la catégorie $HI(k)$. La catégorie $DM^{eff}(k)$ est triangulée monoïdale symétrique fermée. Elle contient comme sous catégorie pleine la catégorie des motifs purs modulo équivalence rationnelle définie par Grothendieck. C’est ainsi une catégorie “effective”, dans le sens où le motif de Tate $\mathbb{I}(1)$ n’a pas de $\otimes$-inverses. Suivant l’approche initiale de Grothendieck, on est conduit à introduire une version non effective des complexes motiviques ; c’est ce qui est fait par D.C. Cisinski et l’auteur dans [CD09b]. Il est naturel dans le contexte des complexes motiviques de remplacer la construction habituelle pour inverser $\mathbb{I}(1)$ par l’approche des topologues pour définir la catégorie homotopique stable. La catégorie $DM(k)$, dont les objets seront appelés les spectres motiviques, est ainsi construite à partir du formalisme des spectres et des catégories de modèles. C’est la catégorie monoïdale homotopique[5] universelle munie d’un foncteur dérivé monoïdal

$$\Sigma^\infty : DM^{eff}(k) \to DM(k)$$

admettant un adjoint à droite $\Omega^\infty$ et telle que l’objet $\Sigma^\infty \mathbb{I}(1)$ est $\otimes$-inversible. Notons que dans le cadre des complexes motiviques, le foncteur $\Sigma^\infty$ est pleinement fidèle d’après le théorème de simplification de Voevodsky [Voe02].

Dans cet article, nous montrons que l’on peut étendre la définition de la t-structure homotopique à la catégorie $DM(k)$, de telle manière que le foncteur $\Omega^\infty$ est t-exact. Le cœur de la t-structure homotopique sur $DM(k)$ est

---

[3] L’identification obtenue ici est naturelle, non seulement par rapport au pullback (lemme 3.3), mais aussi par rapport aux correspondances finies (proposition 3.10) et par rapport au pushout par un morphisme projectif (proposition 3.16).

[4] Originellement, ces complexes sont supposés bornés supérieurement. Nous abandonnons cette hypothèse dans tout l’article suivant [CD09b].

[5] C’est-à-dire la catégorie homotopique associée à une catégorie de modèles monoïdale.
la catégorie $H_L(k)$ des modules homotopiques, qui est donc canoniquement identifiée à la catégorie des modules de cycles d’après le théorème 3.7 déjà cité. Ceci nous permet de donner une interprétation frappante du module de cycles $\tilde{F}_*$ associé à un module homotopique $F_*$, à travers la notion de motifs génériques de [DÉG08b].(6) Le motif générique associé à un corps de fonctions $E$ est le promoteur défini par tous les modèles lisses de $E$. On considère la catégorie $DM_{gm}^{(0)}(k)$ formée par tous les twists de motifs génériques par $\mathbb{P}(n)[n] = \mathbb{P}\{n\}$ pour $n \in \mathbb{Z}$. Alors, $\tilde{F}_*$ est simplement la restriction du foncteur représenté par $F_*$ dans $DM(k)$ à la catégorie $DM_{gm}^{(0)}(k)$. La catégorie $DM_{gm}^{(0)}(k)$ est une catégorie de “points” pour les spectres motiviques, et la fonctorialité des modules de cycles est interprétée en termes de morphismes de spécialisations entre ces points.

De ce point de vue, les modules homotopiques correspondent à des systèmes locaux où le groupoïde fondamental est remplacé par la catégorie $DM_{gm}^{(0)}(k)$. L’interprétation motivique nous sert finalement à introduire une condition de finitude (définition 6.6) sur les modules de cycles qui implique que leur graduation naturelle est bornée inférieurement (Corollaire 6.8) – comme c’est le cas de la plupart des modules de cycles définis par des moyens géométriques.

**Plan du travail.** — L’article est divisé en deux parties, l’une consacrée au théorème principal 3.7 et l’autre à sa signification en termes de la théorie motivique de Voevodsky.

La première partie est faite de trois sections. Dans la section 1, on rappelle les propriétés principales des faisceaux homotopiques, et on introduit la catégorie des modules homotopiques. Dans la section 2, on rappelle brièvement la théorie des modules de cycle de M. Rost et on établit quelques résultats supplémentaires utiles dans cet article. La section 3 est consacrée à la preuve du théorème central 3.7 cité précédemment. De plus, on établit plusieurs propriétés concernant la fonctorialité de l’identification 3.12 citée ci-dessus.

La deuxième partie est aussi constituée de trois sections. La section 4 contient des rappels concernant la théorie des complexes motiviques de Voevodsky ainsi que la version stable qu’on a introduite avec Cisinski dans [CD09b]. La section 5 est consacrée à la définition de la t-structure homotopique et à l’identification de son coeur avec les modules homotopiques. La section 6 est consacrée aux applications du point de vue motivique: construction de modules de cycles (section 6.1), borne inférieure (section 6.2) et extension du corollaire 3.12 au cas singulier (section 6.3).

**Mise en perspective.** — Ce travail a été utilisé récemment par B. Kahn dans [KAH10] pour étendre un théorème de Merkurjev. Kahn démontre par exemple que le théorème de Merkurjev est conséquence de notre théorème 3.7 (voir remarque 6.2).

Nous avons aussi utilisé les résultats de cet article dans deux travaux indépendants:

(6) Cette notion a aussi été introduite par A. Beilinson dans [BE102].

---

Documenta Mathematica 16 (2011) 411–455
F. Dégilde

- F. Morel a introduit une t-structure homotopique sur la catégorie homotopique stable des schémas, analogue à celle qu’on définit sur $\mathcal{D}M(k)$. Il a conjecturé une relation très précise entre le cœur de sa t-structure, noté $\Pi_*(k)$, et les modules homotopiques (avec transferts) considérés ici: la catégorie $\mathcal{H}L_*(k)$ est une sous-catégorie pleine de $\Pi_*(k)$, formée des objets sur lesquels l’application de Hopf agit trivialement. On démontre cette conjecture à partir des résultats de cet article dans [Dég10].

- On approfondit aussi la relation entre modules homotopiques et résolution de Gersten en montrant que la suite spectrale du coniveau associée à la cohomologie représentée par un spectre motivique $E$ s’identifie canoniquement à la suite spectrale d’hyper-cohomologie à coefficients dans $E$ associée à la t-structure homotopique (voir [Dég09, sec. 6]). Ce théorème prolonge un résultat de Bloch-Ogus (cf. [BO74, 6.4]).

Remerciements. — Mes remerciements vont en premier lieu à F. Morel qui a dirigé ma thèse, dans laquelle le résultat central de cet article a été établi.

Notations. — On fixe un corps parfait $k$. Tous les schémas considérés sont des $k$-schémas séparés. Nous dirons qu’un schéma $X$ est lisse si il est lisse de type fini sur $k$. La catégorie des schémas lisses est notée $\mathcal{L}_k$. Nous disons qu’un schéma $X$ est essentiellement de type fini s’il est localement isomorphe au spectre d’une $k$-algèbre qui est une localisation d’une $k$-algèbre de type fini.

On appelle corps de fonctions toute extension de corps $E/k$ de degré de transcendance fini. Un corps de fonctions valué est un couple $(E, v)$ où $E$ est un corps de fonctions et $v$ est une valuation sur $E$ dont l’anneau des entiers est essentiellement de type fini sur $k$. Un modèle de $E/k$ est un $k$-schéma lisse connexe $X$ muni d’un $k$-isomorphismisme entre son corps des fonctions et $E$. On définit le pro-schéma des modèles de $E$ :

\[
(E) = \lim_{\leftarrow \mathcal{A} \subset E} \text{Spec}(A)
\]

où $\mathcal{A}$ parcourt l’ensemble ordonné filtrant des sous-$k$-algèbres de type fini de $E$ dont le corps des fractions est $E$.

Voici une liste des catégories principales utilisées dans ce texte:

- $\mathcal{D}M_{gm}^{eff}(k)$ (resp. $\mathcal{D}M_{gm}(k)$) désigne la catégorie des motifs géométriques effectifs (resp. non nécessairement effectifs).
- $\mathcal{D}M^{eff}(k)$ désigne la catégorie des complexes motiviques (que l’on ne suppose pas nécessairement bornés inférieurement).
Modules Homotopiques

DM(k) désigne la catégorie des spectres motiviques, version non effective de $DM^{eff}(k)$.

HI(k) (resp. $HI_*(k)$) désigne la catégorie des faisceaux (resp. modules) homotopiques. C’est le coeur de la t-structure homotopique sur $DM^{eff}(k)$ (resp. $DM(k)$).

$\mathcal{M}Cycl(k)$ désigne la catégorie des modules de cycles.

PARTIE I
MODULES HOMOTOPIQUES ET MODULES DE CYCLES

1. Modules homotopiques

1.1. Rappels sur les faisceaux avec transferts. — Dans cette partie préliminaire, on rappelle la théorie de Voevodsky des faisceaux avec transferts et des faisceaux homotopiques. Nous nous référons à [Dég07] pour les détails.(7)

1.1. — Soient $X$ et $Y$ des schémas lisses. Rappelons qu’une correspondance finie de $X$ vers $Y$ est un cycle de $X \times Y$ dont le support est fini équidimensionel sur $X$. La formule habituelle permet de définir un produit de composition pour les correspondances finies qui donne lieu à une catégorie additive $L_k^{cor}$ (cf. [Dég07, 4.1.19]). On obtient un foncteur $\gamma : L_k \rightarrow L_k^{cor}$, égal à l’identité sur les objets, en associant à tout morphisme le cycle associé à son graphe. La catégorie $L_k^{cor}$ est enfin monoïdale symétrique. Le produit tensoriel sur les objets est donné par le produit cartésien des schémas lisses; sur les morphismes, il est induit par le produit extérieur des cycles (cf. [Dég07, 4.1.23]).

1.2. — Un faisceau avec transferts est un foncteur $F : (L_k^{cor})^{op} \rightarrow \mathsf{Ab}$additif contravariant tel que $F \circ \gamma$ est un faisceau Nisnevich. On note $Sh^{tr}(k)$ la catégorie des faisceaux avec transferts munis des transformations naturelles. Cette catégorie est abélienne de Grothendieck (cf. [Dég07, 4.2.8]). Une famille génératrice est donnée par les faisceaux représentables par un schéma lisse $X$ :

$$Z^{tr}(X) : Y \mapsto c(Y, X).$$

Il existe un unique produit tensoriel symétrique $\otimes^{tr}$ sur $Sh^{tr}(k)$ telle que le foncteur $Z^{tr}$ est monoïdal symétrique. La catégorie $Sh^{tr}(k)$ est de plus monoïdale symétrique fermée (cf. [Dég07, 4.2.14]).

Définition 1.3. — Un faisceau homotopique est un faisceau avec transferts $F$ invariant par homotopie : pour tout schéma lisse $X$, le morphisme induit par la projection canonique $F(X) \rightarrow F(A^1_X)$ est un isomorphisme.

(7) Cette référence contient une relecture des preuves originales de Voevodsky ainsi que quelques compléments qui nous seront utiles.
On note $HI(k)$ la sous-catégorie pleine de $Sh^{tr}(k)$ formée des faisceaux homotopiques. Le foncteur d’oubli évident $O : HI(k) \to Sh^{tr}(k)$ admet un adjoint à gauche $h_0 : Sh^{tr}(k) \to HI(k)$, $h_0(F)$ étant défini comme le faisceau associé au préfaisceau

$$\text{coKer} \left( F(\mathbb{A}^1_k) \xrightarrow{s_0 - s_1} F(X) \right)$$

avec $s_0$ (resp. $s_1$) la section nulle (resp. unité) de $\mathbb{A}^1_k/X$ (cf. [DÉG07, 4.4.4, 4.4.15]). D’après loc. cit., le foncteur $O$ est exact. La catégorie $HI(k)$ est donc une sous-catégorie épaisse de $Sh^{tr}(k)$. En particulier, c’est une catégorie abélienne de Grothendieck dont une famille génératrice est donnée par les faisceaux de la forme $h_0(X) := h_0(\mathcal{Z}^r(X))$. On vérifie aisément que le foncteur $O$ commute de plus à toutes les limites projectives ce qui implique que $HI(k)$ admet des limites projectives.

1.4. — Pour un corps de fonctions $E$, on définit la fibre de $F$ en $E$ comme la limite inductive de l’application de $F$ au pro-schéma $(E)$ :

$$\hat{F}(E) = \lim_{A \subset E} F(\text{Spec}(A))$$

Les foncteurs $F \mapsto \hat{F}(E)$ forment une famille conservative de foncteurs fibres de $HI(k)$ (cf. [DÉG07, 4.4.7]).

Remarque 1.5. — Ce dernier résultat repose sur la propriété très intéressante des faisceaux homotopiques suivante:

Proposition 1.6. — Pour toute immersion ouverte dense $j : U \to X$ dans un schéma lisse, le morphisme induit

$$j_* : h_0(U) \to h_0(X)$$

est un épimorphisme dans $HI(k)$.

Cette proposition est une conséquence du corollaire 4.3.22 de [DÉG07]: il existe un recouvrement ouvert $W \xrightarrow{\alpha} X$ et une correspondance finie $\alpha : W \to U$ telle que le diagramme suivant est commutatif à homotopie près

$$
\begin{array}{ccc}
& W \\
\alpha \downarrow & \searrow & \downarrow \pi \\
U & \to & X
\end{array}
$$

On peut la reformuler en disant que pour tout faisceau homotopique $F$, le morphisme $F(X) \to F(U)$ induit par $j$ est un monomorphisme. On déduit de ce dernier résultat que pour tout schéma lisse connexe $X$ de corps des fonctions $E$, le morphisme canonique $F(X) \to \hat{F}(E)$ est un monomorphisme.

1.7. — Dans une catégorie abélienne de Grothendieck $\mathcal{A}$, une classe de flèches $W$ est dite localisante si :

(i) $W$ est stable par limite inductive.

(8) i.e. exacts commutant aux limites inductives.
(ii) Soit $f$ et $g$ des flèches composables de $\mathcal{A}$. Si deux des constituants de $(f, g, gf)$ appartiennent à $W$, le troisième appartient à $W$.

Si $\mathcal{S}$ est un classe de flèches essentiellement petite, on peut parler de la classe de flèches localisante engendrée par $\mathcal{S}$.

**Lemme 1.8.** — Il existe un unique produit tensoriel symétrique $\otimes^{Htr}$ sur $HI(k)$ tel que le foncteur $h_0$ est monoïdal symétrique.

**Démonstration.** — D’après ce qui précède, $HI(k)$ s’identifie à la localisation de la catégorie $Sh^{tr}(k)$ par rapport à la classe de flèches localisante engendrée par les morphismes $\mathbb{Z}^{tr}(\mathbb{A}^1_k) \to \mathbb{Z}^{tr}(X)$ pour un schéma lisse $X$ arbitraire. Ainsi, pour tout schéma lisse $X$, $W \otimes^{tr} \mathbb{Z}^{tr}(X) \subset W$. Donc le produit tensoriel $\otimes^{tr}$ satisfait la propriété de localisation par rapport à $W$ ce qui démontre le lemme.

La catégorie $HI(k)$ munie du produit tensoriel $\otimes^{Htr}$ obtenu dans le lemme précédent est monoïdale symétrique fermée. Ce produit tensoriel est caractérisé par la relation $h_0(X) \otimes^{Htr} h_0(Y) = h_0(X \times Y)$ déduite du lemme précédent.

**Définition 1.9.** — Soit $s : \{1\} \to \mathbb{G}_m$ l’immersion du point unité. On appelle sphère de Tate le conoyau de $h_0(s)$ dans la catégorie $HI(k)$. On la note $S^1_t$.

Pour tout entier $n \geq 0$, on note $S^n_t$ la puissance tensorielle $n$-ième de $S^1_t$ dans la catégorie monoïdale $HI(k)$.

D’après l’invariance par homotopie, on obtient encore une suite exacte courte scindée dans $HI(k)$:

$$0 \to S^1_t \to h_0(\mathbb{G}_m) \xrightarrow{j} h_0(\mathbb{A}^1_k) \to 0,$$

où $j$ est l’immersion ouverte évidente.

**1.10.** — Soit $n \geq 0$ un entier et $E/k$ un corps de fonction. Pour un groupe abélien $M$, on note $T_n(M)$ la puissance tensorielle $n$-ième de $M$ pour $\otimes \mathbb{Z}$.

En utilisant le morphisme canonique $\mathbb{G}_m \to h_0(\mathbb{G}_m)$ et la définition du produit tensoriel $\otimes^{Htr}$, on obtient un morphisme canonique:

$$\lambda^n_E : T_n(E^\times) \to \widehat{S}^n_t(E).$$

Notons encore

$$\pi_n : T_n(E^\times) \to K_n^M(E)$$

l’épimorphisme canonique à valeur dans le $n$-ème groupe de K-théorie de Milnor de $E$. On utilisera de manière centrale le résultat suivant dû à Suslin et Voevodsky (voir [SV00, th. 3.4]):

**Théorème 1.11** (Suslin-Voevodsky). — Avec les notations qui précèdent, le morphisme $\lambda^n_E$ se factorise de manière unique par $\pi_n$ et induit un isomorphisme:

$$K_n^M(E) \to \widehat{S}^n_t(E).$$

On déduit de ce théorème le lemme suivant:
Lemme 1.12. — L’automorphisme $\epsilon$ de permutation des facteurs sur $S^2_t = S^1_s \otimes \text{Htr} S^1_s$ est égal à $-1$.

Démonstration. — Compte tenu de la proposition 1.6, il suffit de montrer que pour tout corps de fonctions $E/k$, $\epsilon$ agit par $-1$ sur la fibre $\hat{S}^2_t(E)$.

D’après le théorème précédent, la flèche canonique :

$$\lambda^2_t : E^\times \otimes_Z E^\times \rightarrow \hat{S}^2_t(E)$$

est un épimorphisme. De plus, pour tout couple $(a, b)$ d’unités de $E$, la relation suivante $\lambda^2_t(b, a) = -\lambda^2_t(a, b)$ est vérifiée, d’après la relation analogue bien connue dans $K^2_t(E)$. On conclut du fait que $\epsilon \lambda^2_t(a, b) = \lambda^2_t(b, a)$. 

1.13. — Pour un entier $n \geq 0$ et un faisceau homotopique $F$, on pose $F_{-n} = \text{Hom}_{H1(k)}(S^n_t, F)$. Par définition, pour tout schéma lisse $X$,

$$F_{-1}(X) = F(\mathbb{G}_m \times X)/F(X).$$

Le foncteur $?_{-n}$ est le $n$-ième itéré du foncteur $?_{-1}$. Ainsi la proposition 3.4.3 de [Dég08b] entraîne :

Lemme 1.14. — L’endofoncteur $H1(k) \rightarrow H1(k)$, $F \mapsto F_{-n}$ est exact.

Le résultat suivant est un corollaire du théorème de simplification de Voevodsky [Voe02].

Proposition 1.15. — L’endofoncteur $H1(k) \rightarrow H1(k)$, $F \mapsto S^n_t \otimes \text{Htr} F$ est pleinement fidèle.

Démonstration. — Il suffit de considérer le cas $n = 1$. La preuve anticipe la suite de l’exposé puisqu’elle utilise la catégorie $DM^{eff}_{-2}(k)$ des complexes motiviques de Voevodsky définie dans [Voe02]. Le théorème central de loc. cit. affirme que le twist de Tate est pleinement fidèle dans $DM^{eff}_{-2}(k)$. Il en résulte que le morphisme canonique $F \rightarrow \text{Hom}_{DM^{eff}_{-2}(k)}(\mathbb{Z}^1(1)[1], F(1)[1])$ est un isomorphisme. D’après [Dég08b, 3.4.4], le membre de droite est égal à $H^0(F(1)[1])_{-1}$. Or par définition, $H^0(F(1)[1]) = S^1_t \otimes \text{Htr} F$ et la transformation naturelle correspondante $F \rightarrow (S^1_t \otimes \text{Htr} F)_{-1}$ est l’application d’adjonction. 

1.2. Définition. —

1.16. — On note $\mathbb{Z} - H1(k)$ la catégorie des faisceaux homotopiques $\mathbb{Z}$-gradués. Pour un tel faisceau $F$, et un entier $n \in \mathbb{Z}$, on note $F, \{n\}$ le faisceau gradué dont la composante en degré $i$ est $F_{i+n}$. Si $F$ est un faisceau homotopique, on note encore $F, \{n\}$ le faisceau gradué concentré en degré $-n$ égal à $\mathbb{Z}$. La catégorie $\mathbb{Z} - H1(k)$ est abélienne de Grothendieck avec pour générateurs la famille $(h_0(X)\{i\})$ indexée par les schémas lisses $X$ et les entiers $i \in \mathbb{Z}$.

Cette catégorie est monoïdale symétrique :

$$\left( F \otimes \text{Htr} G \right)_n = \oplus_{p+q=n} F_p \otimes \text{Htr} G_q.$$
Pour la symétrie, on adopte la convention donnée par la règle de Koszul :

$$\oplus_{p+q=n} F_p \otimes^{Htr} G_q \sum (-1)^{pq} \epsilon_{pq} \oplus_{p+q=n} G_q \otimes^{Htr} F_p$$

où $\epsilon_{pq}$ désigne l’isomorphisme de symétrie pour la structure monoïdale des faisceaux homotopiques.

On note $S^*_t$ le monoïde libre dans $\mathbb{Z} - HI(k)$ engendré par le faisceau $S^*_t$ placé
en degré 1. Il est égal en degré $n$ à $S^n_t$. Compte tenu de la règle de Koszul ci-dessus et du lemme 1.12, c’est un monoïde commutatif dans $\mathbb{Z} - HI(k)$. On note $S^*_t -mod$ la catégorie des modules sur $S^*_t$. C’est une catégorie abélienne
monoïdale de Grothendieck avec pour générateurs $(S^*_t \otimes^{Htr} h_0(X) \{i\})$ pour $X$
un schéma lisse et $i \in \mathbb{Z}$. Comme $S^*_t$ est un monoïde libre, se donner un
$S^*_t$-module

$$\tau : S^*_t \otimes^{Htr} F_* \to F_*$$

revient à se donner une suite de morphismes

$$S^1_t \otimes^{Htr} F_n \tau_n \to F_{n+1}$$

appelés morphismes de suspension.

Définition 1.17. — Un module homotopique est un $S^*_t$-module $(F_*, \tau)$ tel que
le morphisme adjoint à $\tau_n$

$$\epsilon_n : F_n \to \text{Hom}_{HI(k)}(S^1_t, F_{n+1}) = (F_{n+1})_{-1}$$

est un isomorphisme. On note $HI_*(k)$ la sous-catégorie de $S^*_t -mod$ formée des
modules homotopiques.

Il revient au même de se donner la suite de morphismes $(\tau_n)_{n \in \mathbb{N}}$ ou la suite
de d’isomorphismes $(\epsilon_n)_{n \in \mathbb{N}}$ pour définir une structure de module homotopique
sur un faisceau homotopique gradué $F_*$. Par la suite, la notation $(F_*, \epsilon_*)$ pour
un module homotopique fera toujours référence aux isomorphismes $\epsilon_n$.

1.18. — Compte tenu du lemme 1.14, le foncteur d’oubli $HI_*(k) \to S^*_t -mod$
est exact et conservatif. Il admet de plus un adjoint à gauche $L$ défini pour
tout faisceau homotopique $F$ et tout entier $i \in \mathbb{Z}$ par la formule

$$L(S^*_t \otimes^{Htr} F\{i\})_n = \begin{cases} S^{n+i}_t \otimes^{Htr} F & \text{si } n + i \geq 0 \\ F_{n+i} & \text{si } n + i \leq 0 \end{cases}$$

en adoptant la notation de 1.13. Le fait que $L$ prenne ses valeurs dans les
faisceaux homotopiques résulte de 1.15. On pose plus simplement $\sigma^\infty F\{i\} = L(S^*_t \otimes^{Htr} F\{i\})$. La catégorie $HI_*(k)$ est donc une sous-catégorie abélienne
de $S^*_t -mod$, avec pour générateurs la famille

(1.18.a) $h_{0,*}(X) = \sigma^\infty h_0(X)\{i\}$

pour un schéma lisse $X$ et un entier $i \in \mathbb{Z} -$ le symbole $*$ correspond à la
graduation naturelle de module homotopique.
Si $(F_*, e_*)$ est un module homotopique, on pose $\omega^\infty F_* = F_0$. On obtient ainsi un couple de foncteurs adjoints

(1.18.b) \[ \sigma^\infty : HI(k) \cong HI_*(k) : \omega^\infty \]

tels que $\sigma^\infty$ est pleinement fidèle (prop. 1.15) et $\omega^\infty$ est exact (lemme 1.14). Ainsi, pour tout schéma lisse $X$, tout module homotopique $F_*$ et tout $(n, i) \in \mathbb{Z}^2$,

(1.18.c) \[ \text{Hom}_{HI_*(k)}(h_{0,*}(X), F_*\{i\}[n]) = H^0_{\text{Nis}}(X; F_i). \]

**Lemme 1.19.** — Il existe sur $HI_*(k)$ une unique structure monoïdale symétrique telle que le foncteur $L$ est monoïdal symétrique.

**Démonstration.** — Compte tenu de ce qui précède, le foncteur $L$ est un foncteur de localisation: pour tout schéma lisse $X$, tout module homotopique $F_*$ et tout mutuple $(n, i) \in \mathbb{Z}^2$.

(1.18.e) \[ \text{Hom}_{HI_*(k)}(h_{0,*}(X), F_*\{i\}[n]) = H^0_{\text{Nis}}(X; F_i). \]

Enfin, l’objet $\sigma^\infty S^1_*$ est inverse pour le produit tensoriel avec pour inverse $\sigma^\infty Z^{tr} \{-1\}$.

**Remarque 1.20.** — La catégorie $HI_*(k)$ est la catégorie monoïdale abélienne de Grothendieck universelle pour les propriétés qui viennent d’être énoncées. La construction donnée ici est parfaitement analogue à la construction de la catégorie des spectres en topologie algébrique, comme le suggère nos notations et en particulier pour le faisceau $S^1_*$ qui joue le rôle de la sphère topologique. La construction ici est facilitée parce que nous sommes dans un cadre abélien et que la sphère $S^1_*$ est anti-commutative. Le théorème de simplification 1.15 rend la construction du foncteur $L$ plus facile mais n’est pas indispensable.

1.3. Réalisation des motifs géométriques. — Rappelons que la catégorie des motifs géométriques effectifs $DM^eff_{gm}(k)$ définie par Voevodsky est l’enveloppe pseudo-abélienne de la localisation de la catégorie $K^b(\mathcal{L}_{\text{cor}}^\text{cor})$ des complexes de $\mathcal{L}_{\text{cor}}^\text{cor}$ à équivalence d’homotopie près par la sous-catégorie triangulée épaisse engendrée par les complexes suivants :

1. $\ldots 0 \to U \cap V \to U \oplus V \to X \to 0 \ldots$
   pour un recouvrement ouvert $U \cup V$ d’un schéma lisse $X$.

2. $\ldots 0 \to H^1_{\text{cor}} \to X \to 0 \ldots$
   induit par la projection canonique pour un schéma lisse $X$.
Rappelons que cette catégorie est triangulée monoïdale symétrique. Pour un schéma lisse $X$, on note simplement $M(X)$ le complexe concentré en degré 0 égal à $X$ vu dans $DM_{gm}^{eff}(k)$.

Pour tout complexe borné $C$ de $L^c_{\text{cor}}$, on note $Z^r(C)$ le complexe de faisceau avec transferts évident. Pour un faisceau homotopique $F$, posons $\varphi_F(C) = \text{Hom}_{D(Sh^{\text{eff}}(k))}(Z^r(C), F)$. Rappelons que pour un schéma lisse $X$,

\[
\text{Hom}_{D(Sh^{\text{eff}}(k))}(Z^r(X)[-n], F) = H^n_{\text{Nis}}(X; F) \quad (\text{cf. } [\text{Voe00b}, 3.1.9]);
\]

la cohomologie Nisnevich de $F$ est de plus invariante par homotopie (cf. [Voe00a, 5.6]). On en déduit que le foncteur $\varphi_F$ ainsi défini se factorise et induit un foncteur cohomologique encore noté $\varphi_F : DM_{gm}^{eff}(k)^{op} \to \mathcal{A}$. On définit le motif de Tate suspendu\(^{(9)}\) $Z[1]$, comme le complexe

\[
\ldots \to \text{Spec}(k) \to \mathbb{G}_m \to 0 \ldots
\]

où $\mathbb{G}_m$ est placé en degré 0, vu dans $DM_{gm}^{eff}(k)$. Avec une convention légèrement différente de celle de Voevodsky, adaptée à nos besoins, on définit la catégorie des motifs géométriques $DM_{gm}(k)$ comme la catégorie monoïdale symétrique universelle obtenue en inversant $Z[1]$ pour le produit tensoriel. Un objet de $DM_{gm}(k)$ est un couple $(C, n)$ où $C$ est un complexe de $L^c_{\text{cor}}$ et $n$ un entier, noté suggestivement $C[n]$. Les morphismes sont définis par la formule

\[
\text{Hom}_{DM_{gm}(k)}(C[n], D[m]) = \lim_{r \geq -n, -m} \text{Hom}_{DM_{gm}^{eff}(k)}(C[r + n], D[r + n]).
\]

Cette catégorie est de manière évidente équivalente à la catégorie définie dans [Voe00b] obtenue en inversant le motif de Tate $Z(1) = Z[1][-1]$. Elle est donc triangulée monoïdale symétrique.

Considérons maintenant un module homotopique $(F_\ast, \epsilon_\ast)$. Pour tout motif géométrique $C[n]$, on pose

\[
\varphi(C[n]) = \lim_{r \geq -n} \text{Hom}_{D(Sh^{\text{eff}}(k))}(Z^r(C)[r + n], F_r),
\]

où les morphismes de transitions sont

\[
\text{Hom}(Z^r(C)[r + n], F_r) \xrightarrow{\epsilon_{r+1}} \text{Hom}(Z^r(C)[r + n], (F_{r+1})_{-1})
\]

\[
= \text{Hom}(Z^r(C)[r + n + 1], F_{r+1}),
\]

les morphismes étant considérés dans la catégorie $D(Sh^{\text{eff}}(k))$. Comme dans le cas des motifs effectifs, ceci induit un foncteur de réalisation cohomologique associé à $(F_\ast, \epsilon_\ast) :$ \[
\varphi : DM_{gm}(k)^{op} \to \mathcal{A}.
\]

Notons que ce foncteur est naturellement gradué $\varphi_n(Z^r(C)[r]) = \varphi(Z^r(C)[r - n])$ de sorte que, d’après le théorème de simplification 1.15, pour tout schéma lisse $X$, \(\varphi_n(Z^r(X)) = F_n(X)\) et

\[
\text{Remarque 1.21. — On déduit du théorème de simplification 1.15 la relation suivante: } \varphi(M(X)[n]) = F_{-n}(X).
\]

\(^{(9)}\)En effet, $Z[1] = Z[1][1]$. 

Documenta Mathematica 16 (2011) 411–455
2. Modules de cycles

Dans cette section, on rappelle la théorie de Rost des modules de cycles exposée dans [Ros96] ainsi que les compléments que nous lui avons apportés dans [Dég08a]. L'étude de la fonctorialité de la suite exacte longue de localisation à l'aide d'un morphisme de Gysin raffiné est nouvelle (cf. proposition 2.6).

2.1. Rappels. — Un pré-module de cycles $\phi$ (cf. [Ros96, (1.1)]) est la donnée pour tout corps de fonctions $E$ d'un groupe abélien $\mathbb{Z}$-gradué $\phi(E)$ satisfaisant à la fonctorialité suivante :

(D1) Pour toute extension de corps $f : E \to L$, on se donne un morphisme appelé restriction $f_* : \phi(E) \to \phi(L)$ de degré 0.

(D2) Pour toute extension finie de corps $f : E \to L$, on se donne un morphisme appelé norme $f^* : \phi(L) \to \phi(E)$ de degré 0.

(D3) Pour tout élément $\sigma \in K^M_r(E)$ du $r$-ième groupe de K-théorie de Milnor de $E$, on se donne un morphisme $\gamma_\sigma : \phi(E) \to \phi(E)$ de degré $r$.

(D4) Pour tout corps de fonctions valué $(E, v)$, on se donne un morphisme appelé résidu $\partial_v : \phi(E) \to \phi(\kappa(v))$ de degré $-1$.

Considérant ces données, on introduit fréquemment un cinquième type de morphisme, associé à un corps de fonctions valué $(E, v)$ et à une uniformisante $\pi$ de $v$, de degré 0, $s^\pi_v = \partial_v \circ \gamma_\pi$, appelé spécialisation.

Ces données sont soumises à un ensemble de relations (cf. [Ros96, (1.1)]). On peut se faire une idée de ces relations en considérant le foncteur de K-théorie de Milnor qui est l’exemple le plus simple de pré-module de cycles.

Considérons un schéma $X$ essentiellement de type fini sur $k$. Soit $x, y$ deux points de $X$. Soit $Z$ l’adhérence réduite de $x$ dans $X$, $\tilde{Z}$ sa normalisation et $f : \tilde{Z} \to Z$ le morphisme canonique. Supposons que $y$ est un point de codimension 1 dans $Z$ et notons $\tilde{Z}_y^{(0)}$ l’ensemble des points génériques de $f^{-1}(y)$. Tout point $z \in \tilde{Z}_y^{(0)}$ correspond alors à une valuation $v_z$ sur $\kappa(x)$ de corps résiduel $\kappa(z)$. On note encore $\varphi_z : \kappa(y) \to \kappa(z)$ le morphisme induit par $f$. On définit un morphisme $\partial_y^\varphi : \phi(\kappa(x)) \to \phi(\kappa(y))$ par la formule suivante :

$$\partial_y^\varphi = \begin{cases} \sum_{z \in \tilde{Z}_y^{(0)}} \varphi_z^* \circ \partial_{v_z} & \text{si } y \in Z^{(1)}, \\ 0 & \text{sinon.} \end{cases}$$

Considérons ensuite le groupe abélien :

$$C^p(X; \phi) = \bigoplus_{x \in X^{(p)}} \phi(\kappa(x)).$$

On dit que le pré-module de cycles $\phi$ est un module de cycles (cf. [Ros96, (2.1)]) si pour tout schéma essentiellement de type fini $X$,

(FD) Le morphisme

$$d_{X, \phi}^p : \sum_{x \in X^{(p)}, y \in X^{(p+1)}} \partial_y^\varphi : C^p(X; \phi) \to C^{p+1}(X; \phi)$$

est bien défini.
(C) La suite
\[ \ldots \to C^p(X;\phi) \xrightarrow{d^p_{X;\phi}} C^{p+1}(X;\phi) \to \ldots \]
est un complexe.

Les modules de cycles forment de manière évidente une catégorie que l'on note
\[ MCycl(k) \].

On introduit une graduation sur le complexe de la propriété (C) :
\[ C^p(X;\phi) = \bigoplus_{x \in X^{(p)}} M_{n-p}(\kappa(x)) \],

On note \[ A^p(X;\phi) \] le \( p \)-ième groupe de cohomologie de ce complexe, appelé parfois groupe de Chow à coefficients dans \( \phi \).

Pour un schéma lisse \( X \) de corps des fonctions \( E \), le groupe \( A^0(X;\phi) \) est donc le noyau de l'application bien définie
\[ \phi_n(E) \xrightarrow{\sum_{x \in X^{(1)}} \partial_x} \phi_{n-1}(\kappa(x)) \]

où \( \partial_x \) désigne le morphisme résidu associé à la valuation sur \( E \) correspondant au point \( x \).

2.2. Fonctorialité. —

2.1. — Le complexe gradué \( C^\ast(X;\phi) \) est contravariant en \( X \) par rapport aux morphismes plats (cf. [Ros96, (3.4)]). Il est covariant par rapport aux morphismes propres équidimensionnels (cf. [Ros96, (3.5)]).

2.2. — Dans [Dég06, 3.18], nous avons prolongé le travail original de Rost et nous avons associé à tout morphisme \( f : Y \to X \) localement d'intersection complète ([Dég06, 3.12]) tel que \( Y \) est lissifiable ([Dég06, 3.13]) un morphisme de Gysin
\[ f^\ast : C^\ast(X;\phi) \to C^\ast(Y;\phi) \]
qui est un composé d’un morphisme de complexes et de l’inverse formel d’un morphisme de complexe qui est un quasi-isomorphisme (plus précisément, il s’agit de l’inverse formel d’un morphisme \( p^\ast \) pour \( p \) la projection d’un fibré vectoriel). Pour désigner une telle flèche formelle, on utilise la notation abrégée
\[ f^\ast : X \bullet \to Y \].

(10) Ce morphisme de Gysin \( f^\ast \) satisfait les propriétés suivantes :
1. Lorsque \( f \) est de plus plat, \( f^\ast \) coïncide avec le pullback plat évoqué plus haut.
2. Si \( g : Z \to Y \) est un morphisme localement d’intersection complète avec \( Z \) lissifiable, \( (fg)^\ast = g^\ast f^\ast \).

(10) Les flèches de ce type sont bien définies dans la catégorie dérivée des groupes abéliens et induisent en particulier un morphisme sur les groupes de cohomologie.
Dans le cas où $f$ est une immersion fermée régulière, l’hypothèse que $Y$ est lissifiable est inutile ; le morphisme $f^*$ est défini en utilisant la déformation au cône normal, suivant l’idée originale de Rost (cf. [Dég06, 3.3]). On utilisera par ailleurs le résultat suivant dû à Rost ([Ros96, (12.4)]) qui décrit partiellement ce morphisme de Gysin :

**Proposition 2.3.** — Soit $X$ un schéma intègre de corps des fonctions $E$, et $i : Z \to X$ l’immersion fermée d’un diviseur principal régulier irréductible paramétré par $\pi \in \mathcal{O}_X(X)$. Soit $v$ la valuation de $E$ correspondant au diviseur $Z$. Alors, le morphisme $i^* : A^0(X;\phi) \to A^0(Z;\phi)$ est la restriction de $s_v^\pi : \phi(E) \to \phi(\kappa(v))$.

2.4. — A tout carré cartésien

$\begin{array}{ccc}
Y' & \xrightarrow{j} & X' \\
\downarrow \Delta & & \downarrow f \\
Y & \xrightarrow{i} & X
\end{array}$

tel que $i$ est une immersion fermée régulière, on associe un morphisme de Gysin raffiné $\Delta^* : X' \to Y'$. Ce morphisme $\Delta^*$ vérifie les propriétés suivantes :

1. Si $j$ est régulier et le morphisme des cônes normaux $N_{Y'}(X') \to g^{-1}N_Y(X)$ est un isomorphisme, $\Delta^* = j^*$.
2. Si $f$ est propre, $i^* f^* = g_* \Delta^*$.

De plus, si l’immersion canonique $C_{Y'}(X') \to g^{-1}N_Y(X)$ du cône de $j$ dans le fibré normal de $i$ est de codimension pure égale à $e$, le morphisme $\Delta^*$ est de degré cohomologique $e$.

2.5. — Pour tout couple de schémas lisses $(X,Y)$ et pour toute correspondance finie $\alpha \in c(X,Y)$, on définit un morphisme $\alpha^* : Y' \to X$ (cf. [Dég06, 6.9]). On peut décrire ce dernier comme suit. Supposons que $\alpha$ est la classe d’un sous-schéma fermé irréductible $Z$ de $X \times Y$. Considérons les morphismes:

$X \xrightarrow{\ell} Z \xrightarrow{\alpha} Z \times X \times Y \xrightarrow{q} Y$

où $p$ et $q$ désignent les projections canoniques et $i$ le graphe de l’immersion fermée $Z \to X \times Y$. Alors,

$\alpha^* = p_* i^* q^*$

(2.5.a)

où $i^*$ désigne le morphisme de Gysin de l’immersion fermée régulière $i$, $q^*$ le pullback plat et $p_*$ le pushout fini.

La propriété $(\beta\alpha)^* = \alpha^* \beta^*$ est démontrée dans [Dég06, 6.5].

2.3. Suite exacte de localisation. — La suite exacte de localisation n’est pas étudiée (ni rappelée) dans [Dég06]. Nous la rappelons maintenant suivant [Ros96] et démontrons un résultat supplémentaire concernant sa fonctorialité. Pour une immersion fermée $i : Z \to X$ purement de codimension $c$, d’immersion
ouverte complémentaire $j : U \to X$, on obtient en utilisant la fonctorialité rappelée ci-dessus une suite exacte courte scindée de complexes

$$(2.5.b) \quad 0 \to C^{p-c}(Z; \phi)_{n-c} \xrightarrow{i_*} C^p(X; \phi)_n \xrightarrow{j^*} C^p(U; \phi)_n \to 0.$$ 

On en déduit une suite exacte longue de localisation

$$(2.5.c) \quad \cdots \to A^{p-c}(Z; \phi)_{n-c} \xrightarrow{i_*} A^p(X; \phi)_n \xrightarrow{j^*} A^p(U; \phi)_n \xrightarrow{\partial^Z_*} A^{p-c+1}(Z; \phi)_{n-c} \to \cdots$$

où le morphisme $\partial^Z_*$ est défini au niveau des complexes par la formule $\sum_{x \in (T; \phi), z \in (Z-c+1)} \partial^x_* Z$.

Cette suite est naturelle par rapport au pushout propre et au pullback plat.

La proposition suivante est nouvelle :

**Proposition 2.6.** — Considérons un carré cartésien

\[
\begin{array}{ccc}
T & \xrightarrow{i} & Z \\
\downarrow k & & \downarrow j \\
Y & \xrightarrow{\iota} & X
\end{array}
\]

tel que $i$ est une immersion fermée régulière. Supposons que $i$ (resp. $k$) est une immersion fermée d’immersion ouverte complémentaire $j : U \to X$ (resp. $l : V \to X$). Notons $h : V \to U$ le morphisme induit par $i$. Supposons enfin que $i$ (resp. $k$) est de codimension pure égale à $c$ (resp. $d$). Alors, le diagramme suivant est commutatif :

\[
\begin{array}{ccc}
\cdots & \xrightarrow{\Delta} & A^{p-c}(Z; \phi)_{n-c} \xrightarrow{i_*} A^p(X; \phi)_n \xrightarrow{j^*} A^p(U; \phi)_n \xrightarrow{\partial^Z_*} A^{p-c+1}(Z; \phi)_{n-c} & \cdots \\
\cdots & \xrightarrow{\Delta} & A^{p-d}(T; \phi)_{n-d} \xrightarrow{k_*} A^p(Y; \phi)_n \xrightarrow{l^*} A^p(V; \phi)_n \xrightarrow{\partial^Y_*} A^{p-d+1}(T; \phi)_{n-d} & \cdots
\end{array}
\]

**Remarque 2.7.** — 1. On peut généraliser la proposition précédente au cas des morphismes de Gysin raffinés comme dans la proposition 4.5 de [Dég06]. Nous laissons au lecteur le soin de formuler cette généralisation.

2. Alors que l’hypothèse sur la codimension pure de $i$ est naturelle, celle sur $k$ ne l’est pas, en particulier dans un cas non transverse. Elle ne nous sert qu’à exprimer les degrés cohomologiques de tous les morphismes et peut aisément être supprimée si on accepte des morphismes non homogènes par rapport au degré cohomologique.

**Démonstration.** — Il suffit de reprendre la preuve de la proposition 4.5 de loc. cit. dans le cas du diagramme commutatif :

\[
\begin{array}{ccc}
T & \xrightarrow{k} & Z \\
\downarrow \iota & & \downarrow j \\
Y & \xrightarrow{\iota} & X
\end{array}
\]
On obtient ainsi un diagramme commutatif\(^{(11)}\), avec les notations analogues de loc. cit.

\[
\begin{array}{cccccc}
Z & \rightarrow & C_T Z & \rightarrow & k^* N_X Y & \rightarrow & T \\
\downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
X & \rightarrow & N_X Y & \rightarrow & N_Y X & \rightarrow & Y \\
\downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
U & \rightarrow & N_Y V & \rightarrow & N_X V & \rightarrow & U \\
\end{array}
\]

Les carrés (1), (2), (3) sont commutatifs d’après loc. cit. et les carrés (1’), (2’), (3’) le sont pour des raisons triviales. Les flèches \(\bullet \rightarrow\) qui apparaissent dans ce diagramme sont bien des morphismes de complexes et induisent donc des morphismes de suite exacte longue de localisation. Il suffit alors d’appliquer le fait que les morphismes \(p^*, p_T^*\) et \(p_U^*\) sont des quasi-isomorphismes pour conclure.

\[\square\]

**Corollaire 2.8.** — Considérons un carré cartésien

\[
\begin{array}{ccc}
T & \rightarrow & Z \\
\downarrow & \downarrow & \downarrow \\
Y & \rightarrow & X
\end{array}
\]

de schémas lisses tels que \(i\) (resp. \(k\)) est une immersion fermée de codimension pure égale à \(c\), d’immersion ouverte complémentaire \(j\) : \(U \rightarrow X\) (resp. \(l\) : \(V \rightarrow X\)). Notons \(h : V \rightarrow U\) le morphisme induit par \(f\). Alors, le diagramme suivant est commutatif :

\[
\begin{array}{cccccc}
\cdots & \rightarrow & A^{p-c}(Z; \phi)_{n-c} & \rightarrow & A^p(X; \phi)_{n} & \rightarrow & A^p(U; \phi)_{n} & \rightarrow & A^{p-c+1}(Z; \phi)_{n-c} \\
\downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
\cdots & \rightarrow & A^{p-c}(T; \phi)_{n-c} & \rightarrow & A^p(Y; \phi)_{n} & \rightarrow & A^p(V; \phi)_{n} & \rightarrow & A^{p-c+1}(T; \phi)_{n-c} \\
\end{array}
\]

**Remarque 2.9.** — Dans l’article [Dég08b], une **paire fermée** est un couple \((X, Z)\) tel que \(X\) est un schéma lisse et \(Z\) un sous-schéma fermé. On dit que \((X, Z)\) est lisse (resp. de codimension \(n\)) si \(Z\) est lisse (resp. purement de codimension \(n\) dans \(X\)).

Si \(i : Z \rightarrow X\) est l’immersion fermée associée, un morphisme de paires fermées \((f, g)\) est un carré commutatif

\[
\begin{array}{ccc}
T & \rightarrow & Z \\
\downarrow & \downarrow & \downarrow \\
Y & \rightarrow & X
\end{array}
\]

qui est topologiquement cartésien. On dit que \((f, g)\) est **cartésien** (resp. **transverse**) quand le carré est cartésien (resp. et le morphisme induit sur les cônes

\(^{(11)}\)Il y a une faute de frappe dans le diagramme commutatif de loc. cit. Il faut lire \(t^* N_X Y\) au lieu de \(N_Y X\).
normaux $C_TY \to g^{-1}C_TX$ est un isomorphisme).\(^{(12)}\)
Le corollaire précédent montre que la suite de localisation associée à un module de cycles $\phi$ et une paire fermée $(X, Z)$ est naturelle par rapport aux morphismes transverses.

2.4. Module homotopique associé. —

2.10. — Considérons un module de cycles $\phi$. D'après 2.5, $A^0(\cdot; \phi)$, définit un préfaisceau gradué avec transferts. D'après [Dég06, 6.9], c'est un faisceau homotopique gradué. On le note $F^\phi_*$ et on lui définit une structure de module homotopique comme suit:
Soit $X$ un schéma lisse. On considère le début de la suite exacte longue de localisation (2.5.c) associée à la section nulle $X \to \mathbb{A}^1_X$:

$$0 \to F^\phi_n(\mathbb{A}^1_X) \xrightarrow{j_{X}} F^\phi_n(\mathbb{G}_m \times X) \xrightarrow{\partial^X_0} F^\phi_{n-1}(X) \to \ldots$$

On peut décrire le morphisme $\partial^X_0$ si $X$ est connexe de corps des fonctions $E$ comme étant induit par le morphisme

$$\partial^E_0 : \phi_n(E(t)) \to \phi_{n-1}(E)$$

associé à la valuation standard de $E(t)$.

Soit $s_1 : X \to \mathbb{G}_m \times X$ la section unité. Rappelons que $(F^\phi_n)_1(X) = \text{Ker}(s_1^*)$. Or par invariance par homotopie de $F^\phi_n$, le morphisme canonique $\text{Ker}(s_1^*) \to \text{coKer}(j^*)$ est un isomorphisme. Ainsi, le morphisme $\partial^X_0$ induit un morphisme

$$\epsilon_{n, X} : (F^\phi_n)_1(X) \to F^\phi_{n-1}(X).$$

On vérifie que la suite de localisation précédente est compatible aux transferts en $X$, comme cela résulte de la description des transferts rappelée en 2.5 et du corollaire 2.8. Ainsi, $\epsilon_n$ définit un morphisme de faisceaux homotopiques. Pour tout corps de fonctions $E$, $A^1(\mathbb{A}^1_E; \phi) = 0$ (cf. [Ros96, (2.2)(H)]). Donc la fibre de $\epsilon_n$ en $E$ est un isomorphisme ce qui implique que c’est un isomorphisme de faisceaux homotopiques d’après 1.4.
Ainsi, $(F^\phi_*, \epsilon_0^{-1})$ définit un module homotopique qui dépend fonctoriellement de $\phi$.

3. Equivalence de catégories

3.1. Transformée générique. — Considérons un couple $(E, n)$ formé d’un corps de fonctions $E$ et d’un entier relatif $n$. Rappelons que l’on a associé dans [Dég08B, 3.3.1] au couple $(E, n)$ un motif générique

$$M(E)\{n\} = \varprojlim_{\mathbb{A}^1_E} M(\text{Spec}(A))\{n\}$$

\(^{(12)}\)Lorsque $(X, Z)$ est lisse de codimension $n$ le fait que le morphisme $(f, g)$ est transverse entraîne que $(Y, T)$ est lisse de codimension $n$ ($k$ est régulier).
dans la catégorie des pro-objets de $DM_{gm}(k)$. On note $DM_{gm}^{(0)}(k)$ la catégorie des motifs génériques.

3.1. — Considérons un module homotopique $(F_\ast, \epsilon_\ast)$ ainsi que le foncteur de réalisation $\varphi : DM_{gm}(k)^{op} \to \mathcal{L}b$ qui lui est associé dans la section 1.3. On note $\hat{\varphi}$ le prolongement évident de $\varphi$ à la catégorie des pro-objets. Il résulte de [DÉG08B, 6.2.1] que la restriction de $\hat{\varphi}$ à la catégorie $DM_{gm}^{(0)}(k)$ est un module de cycles, que l’on note $\hat{F}_\ast$ et que l’on appelle la transformée générique de $F_\ast$. Rappelons brièvement certaines parties de la construction de [DÉG08B].

De même que pour tout motif générique $M(E)\{n\}$, $\hat{\varphi}(M(E)\{n\}) = \hat{F}_{-n}(E)$ n’est autre que le cycle de $f_\ast$ en $E$ (cf. 1.4). La transformée $\hat{F}_\ast$ s’interprète donc comme le système des cycles de $F_\ast$. Ce sont les morphismes de spécialisation entre ces fibres qui donnent la structure de pré-module de cycles :

(D1) Fonctorialité évidente de $F_\ast$.  
(D2) ([DÉG08B, 5.2]) Pour une extension finie $L/E$, on trouve des modèles respectifs $X$ et $Y$ de $E$ et $L$ ainsi qu’un morphisme fini surjectif $f : Y \to X$ dont l’extension induite des corps de fonctions est isomorphe à $L/E$. Le graphe de $f$ vu comme cycle de $X \times Y$ définit une correspondance finie de $X$ vers $Y$ notée $f^1$ la transposée de $f$. On en déduit un morphisme $(f^1)^* : F_\ast(X) \to F_\ast(Y)$. On montre que ce morphisme est compatible à la restriction à un ouvert de $X$ et il induit donc la fonctorialité attendue.

(D3) ([DÉG08B, 5.3]) Soit $E$ un corps de fonctions et $x \in E^\times$ une unité. Considérons un modèle $X$ de $E$ munit d’une section inversible $x : X \to \mathbb{G}_m$ qui correspond à $x$. Considérons l’immersion fermée $s_x : X \to \mathbb{G}_m \times X$ induite par cette section. On en déduit un morphisme

$$\gamma_x : F_{n-1}(X) \xrightarrow{\epsilon_{n-1}} (F_n)_{-1}(X) \xrightarrow{\nu} F_n(\mathbb{G}_m \times X) \xrightarrow{\kappa} F_n(X)$$

où $\nu$ est l’inclusion canonique. Ce morphisme est compatible à la restriction suivant un ouvert de $X$ et induit la donnée D3 pour $F_\ast$.  

(D4) ([DÉG08B, 5.4]) Soit $(E,v)$ un corps de fonctions valué. On peut trouver un schéma lisse $X$ munit d’un point $x$ de codimension 1 tel que l’adhérence réduite $Z$ de $x$ dans $X$ est lisse et l’anneau local $O_{X,x}$ est isomorphe à l’anneau des entiers de $v$. On pose $U = X - Z$, $j : U \to X$ l’immersion ouverte évidente. Rappelons que le motif $M_Z(X)$ de la paire $(X,Z)$ est défini comme l’objet de $DM_{gm}^{f}(k)$ représenté par le complexe concentré en degré 0 et $-1$ avec pour seule différentielle non nulle le morphisme $j$. Ce motif s’inscrit naturellement dans le triangle distingué

$$M_Z(X)[-1] \xrightarrow{\partial_X^*} M(U) \xrightarrow{j^*} M(X) \xrightarrow{+1}$$

On a défini dans [DÉG08B, sec. 2.2.5] un isomorphisme de pureté

$$\mathfrak{p}_{X,Z} : M_Z(X) \to M(Z)(1)[2].$$

Documenta Mathematica 16 (2011) 411–455
On en déduit un morphisme

$$\partial_{U,Z} : F_n(U) = \varphi_n(M(U)) \xrightarrow{\varphi_n(p^{-1}_{U,Z})} \varphi_n(M_Z(X)[-1])$$

ayant posé $\varphi_n(M) = \varphi(M{-n})$ pour un motif $M$. Le morphisme résidu du module de cycles $F_\ast$ est donné par la limite inductive des morphismes $\partial_{U,Z \cap U}$ suivant les voisinages ouverts $U$ de $x$ dans $X$.

### 3.2. Résolution de Gersten: fonctorialité I

3.2. — Considérons un module de cycles $\phi$ et $F^\phi$ le module homotopique qui lui est associé dans le paragraphe 2.10 – jusqu’au paragraphe 3.5, on n’indique pas la graduation pour alléger les notations. D’après [Ros96, 6.5], on dispose pour tout schéma lisse $X$ et tout entier $p \in \mathbb{Z}$ d’un isomorphisme canonique $A^p(X; \phi) = H^p_{Zar}(X; F^\phi)$.

On rappelle la construction de cet isomorphisme tout en le généralisant au cas de la topologie Nisnevich. Soit $X$ un schéma lisse et $X_{Nis}$ le petit site Nisnevich associé. Les morphismes de $X_{Nis}$ étant étales, on obtient, en utilisant la fonctorialité rappelée dans 2.1, un préfaisceau de complexes de groupes abéliens sur $X_{Nis}$:

$$C^*_X(\phi) : V/X \mapsto C^*(V; \phi).$$

On vérifie que c’est un faisceau Nisnevich (voir [Dég08b], preuve de 6.10). On pose de plus:

$$F^\phi_X = H^0(C^*_X(\phi)).$$

Ainsi, $F^\phi_X$ est la restriction du faisceau $F^\phi$, défini sur le site Nisnevich $\mathcal{L}_k$, au petit site $X_{Nis}$. D’après [Ros96, 6.1], le morphisme évident

(3.2.a) $$F^\phi_X \rightarrow C^*_X(\phi)$$

est un quasi-isomorphisme. Il induit donc un isomorphisme

$$H^p_{Nis}(X; F^\phi_X) \rightarrow H^p_{Nis}(X; C^*_X(\phi)).$$

Notons par ailleurs que le complexe $C^*_X(\phi)$ vérifie la propriété de Brown-Gersten au sens de [CD09A, 1.1.9] (voir à nouveau [Dég08b], preuve de 6.10).

D’après la démonstration de [CD09A, 1.1.10], on en déduit que le morphisme canonique

$$H^p(C^*(X; \phi)) \rightarrow H^p_{Nis}(X; C^*_X(\phi))$$

est un isomorphisme. Ces deux isomorphismes définissent comme annoncé :

(3.2.b) $$\rho_X : A^p(X; \phi) = H^p(C^*(X; \phi)) \xrightarrow{\sim} H^p_{Nis}(X; F^\phi_X) \cong H^p_{Nis}(X; F^\phi).$$

**Lemme 3.3.** — L’isomorphisme $\rho_X$ construit ci-dessus est naturel en $X$ par rapport aux morphismes de schémas.

---

(13) Le complexe de faisceaux $C^*_X(\phi)$ est la *résolution de Gersten* du faisceau $F^\phi_X$. C’est en fait la version Nisnevich de la résolution de Cousin au sens de [HAR66].
Démonstration. — Notons que, du fait que $F^\phi_X$ est la restriction d’un faisceau $F^\phi$ sur $\mathcal{L}_k$, pour tout morphisme $f : Y \to X$ de schémas lisses, on obtient une transformation naturelle canonique:

$$F^\phi_X \to f_! F^\phi_Y$$

qui induit dans la catégorie dérivée:

$$\tau_f : F^\phi_X \to Rf_! F^\phi_Y.$$

La preuve consiste à relever cette transformation naturelle au niveau de la résolution $C^*_X(\phi)$. On considère d’abord le cas où $f$ est plat. Suivant le paragraphe 2.1, on dispose d’un morphisme de complexes

$$f^* : C^*(X; \phi) \to C^*(Y; \phi)$$

qui est naturel en $X$ par rapport aux morphismes étalés. La transformation naturelle sur $X_{Nis}$ correspondante définit un morphisme dans la catégorie dérivée des faisceaux abéliens sur $X_{Nis}$:

$$(3.3.a) \quad \eta_f : C^*_X(\phi) \to f_! C^*_Y(\phi) = Rf_! C^*_Y(\phi).$$

(La dernière identification résulte du fait que $C^*_Y(\phi)$ vérifie la propriété de Brown-Gersten.) Par définition de la structure de faisceau sur $F^\phi$, le diagramme suivant est commutatif:

$$\begin{array}{ccc}
F^\phi_X & \xrightarrow{\tau_f} & Rf_! F^\phi_Y \\
\downarrow & & \downarrow \\
C^*_X(\phi) & \xrightarrow{\eta_f} & Rf_! C^*_Y(\phi).
\end{array}$$

On en déduit la naturalité de $\rho$ par rapport aux morphismes plats. Remarquons au passage que si $f$ est la projection d’un fibré vectoriel, $\eta_f$ est un quasi-isoïsmorphisme.

Il reste à considérer le cas d’une immersion fermée $f = i : Z \to X$ entre schémas lisses. Notons $N$ le fibré normal associé à $i$. La spécialisation au fibré normal définie par Rost (cf. [Ros96, sec. 11]) est un morphisme de complexes

$$\sigma_Z X : C^*(X; \phi) \to C^*(N; \phi)$$

qui est de plus naturel en $X$ par rapport aux morphismes étalés (cf. [Dég06, 2.2]). Notons $\nu$ le morphisme composé

$$N \xrightarrow{\nu} Z \xrightarrow{i} X.$$  

On en déduit dans la catégorie dérivée un morphisme canonique

$$\sigma_i : C^*_X(\phi) \to R\nu_* C^*_N(\phi).$$

Puisque le morphisme $\eta_0$ est un quasi-isoïsmorphisme, on obtient alors un morphisme canonique dans la catégorie dérivée

$$(3.3.b) \quad \eta_i : C^*_X(\phi) \to R\iota_* C^*_Z(\phi).$$
Rappelons enfin que, par définition du pullback sur $F^\phi$, le diagramme suivant est commutatif:

$$
\begin{array}{ccc}
F^\phi(X) & \xrightarrow{i^*} & F^\phi(Z) \\
\downarrow & & \downarrow \\
C^*(X;\phi) & \xrightarrow{\sigma\phi(X)} & C^*(N;\phi) \xrightarrow{\rho^*} C^*(Z;\phi).
\end{array}
$$

On en déduit que le diagramme suivant est commutatif:

$$
\begin{array}{ccc}
F^\phi_X & \xrightarrow{\tau_i} & R_i F^\phi_Z \\
\downarrow & & \downarrow \\
C^*_X(\phi) & \xrightarrow{\eta_i} & R_i C^*_Z(\phi)
\end{array}
$$

ce qui conclut.

\[ \square \]

\textbf{Remarque 3.4.} — 1. On généralisera ce lemme au cas des correspondances finies dans la proposition 3.10.

2. Les constructions (3.3.a) et (3.3.b) de la preuve précédente permettent d’associer à tout morphisme de schémas $f : Y \to X$ un diagramme commutatif dans la catégorie dérivée des faisceaux sur $X_{nis}$:

$$
\begin{array}{ccc}
F^\phi_X & \xrightarrow{\tau_f} & R_f F^\phi_Y \\
\downarrow & & \downarrow \\
C^*_X(\phi) & \xrightarrow{\eta_f} & R_f C^*_Y(\phi),
\end{array}
$$

en considérant la factorisation de $f$ par son morphisme graphe qui est une immersion régulière. On peut montrer par ailleurs que $\eta_f$ est compatible à la composition des morphismes.

3.5. — On reprend les notations du paragraphe 3.2 Considérons par ailleurs le foncteur de réalisation

$$
\varphi : DM_{gm}(k)^{op} \to \ab
$$

associé au module homotopique $F^\phi$ suivant la section 1.3. L’isomorphisme $\rho_X$ correspond par définition à un isomorphisme:

$$
A^p(X,\phi)_n \to \varphi_n(M(X)[-p]).
$$

Considérons de plus une immersion fermée $i : Z \to X$ entre schémas lisses et $j : U \to X$ l’immersion ouverte du complémentaire. Supposons que $i$ est de codimension pure égale à $c$. On déduit de la suite exacte de localisation (2.5.b) une unique flèche pointillée qui fait commuter le diagramme de complexes suivant (on utilise à nouveau le fait que $C^*_X(\phi)$ vérifie la propriété de Brown-Gersten):

$$
\begin{array}{c}
0 \to C^*(Z,\phi)_n-c[-c] \xrightarrow{i^*} C^*(X,\phi)_n \xrightarrow{j^*} C^*(U,\phi)_n \to 0 \\
\downarrow \quad \downarrow \quad \downarrow \\
0 \xrightarrow{(1)} \Gamma(Z,\phi)_n \xrightarrow{R\Gamma(X,\phi)_n} \Gamma(U,\phi)_n \xrightarrow{R\Gamma(U,\phi)_n} 0.
\end{array}
$$
La flèche (1) est un quasi-isomorphisme, puisqu’il en est de même des deux autres flèches verticales. Considérons le motif $M_Z(X)$ associé à la paire fermée $(X, Z)$ – cf. 3.1, (D4). En utilisant l’isomorphisme (1) et l’identification canonique $H^0_\mathbb{Z}(X; F\phi)_n = \varphi_n(M_Z(X)[-p])$, on obtient un diagramme commutatif:

\[
\begin{array}{cccccc}
A^{p-1}(U, \phi)_n & \xrightarrow{\partial^U_Z} & A^{p-c}(Z, \phi)_{n-c} & \xrightarrow{i_*} & A^p(X, \phi)_n \\
\downarrow{\rho_V} & & \downarrow{\rho_{X,Z}} & & \downarrow{\rho_X} \\
\varphi_n(M(U)[-p]) & \xrightarrow{\varphi_n(M_Z(X)[-p])} & \varphi_n(M(X)[-p])
\end{array}
\]

dans lequel les flèches verticales sont des isomorphismes. Le morphisme $\rho'_{X,Z}$ est de plus naturel en $(X, Z)$ par rapport aux morphismes transverses (définis en 2.9). Cela résulte en effet du corollaire 2.8, ou plus précisément du diagramme commutatif apparaissant dans la démonstration de 2.6, en utilisant d’une part l’unicité de la flèche pointillée (1) et d’autre part la description de la fonctorialité dérivée de $C^*_X(\phi)$ établie ci-dessus – i.e. les transformations naturelles $\tau_f$ et $\tau_i$.

Comme conséquence de cette construction, on obtient le lemme clé suivant:

**Lemme 3.6.** — Reprenons les notations qui précèdent. Considérons le triangle de Gysin (cf. [Voe00b, 3.5.4]) associé à $(X, Z)$:

\[
M(U) \to M(X) \xrightarrow{\iota_*} M(Z)(c)[2c] \xrightarrow{\partial_{X,Z}} M(U)[1].
\]

Alors, le diagramme suivant est commutatif:

\[
\begin{array}{cccccc}
A^{p-1}(U, \phi)_n & \xrightarrow{\partial^U_Z} & A^{p-c}(Z, \phi)_{n-c} & \xrightarrow{i_*} & A^p(X, \phi)_n \\
\downarrow{\rho_V} & & \downarrow{\rho_{X,Z}} & & \downarrow{\rho_X} \\
\varphi_n(M(U)[-p]) & \xrightarrow{\varphi_n(M_Z(X)(c - p))} & \varphi_n(M(Z)(c)[2c - p]) & \xrightarrow{\varphi_n(\iota'_*)} & \varphi_n(M(X)[-p]).
\end{array}
\]

**Démonstration.** — On utilise la construction du triangle de Gysin effectuée dans [Dég08b]. Considérons l’isomorphisme de pureté défini dans [Dég08b, sec. 2.2.5]

\[
\varphi_{X,Z} : M_Z(X) \to M(Z)(c)[2c].
\]

D’après ce qui précède, l’isomorphisme composé

\[
\rho_{X,Z} : A^{p-c}(Z, \phi)_{n-c} \xrightarrow{\rho'_{X,Z}} \varphi_n(M_Z(X)[-p]) \xrightarrow{\varphi_{(p_{X,Z})}} \varphi_n(M(Z)(c)[2c - p]) = \varphi_{n-c}(M(Z)(c - p))
\]
s’inscrit dans le diagramme commutatif:

\[
\begin{array}{cccccc}
A^{p-1}(U, \phi)_{n} & \xrightarrow{\partial Z_{c}^{u}} & A^{p-c}(Z, \phi)_{n-c} & \xrightarrow{i_{*}} & A^{p}(X, \phi)_{n} \\
\rho_{U} & & \phi_{n-c}(M(Z)[-c-1]) & & \rho_{X, Z} \\
\varphi_{n}(M(U)[-p]) & & \varphi_{n}(M(Z)(-c)) & & \varphi_{n}(M(X)[-p]).
\end{array}
\]

Il s’agit de voir que \( \rho_{X, Z} = \rho_{Z} \). Notons que d’après ce qui précède, le morphisme \( \rho_{X, Z} - \rho_{Z} \) est naturel en \((X, Z)\) par rapport aux morphisme transverses (définis en 2.9). Soit \( P_{Z}X \) la complétion projective du fibré normal de \( Z \) dans \( X \). Considérons l’éclatement \( B_{Z}(A_{1}^{k}) \) de \( Z \times \{0\} \) dans \( X \), ainsi que le diagramme de déformation classique qui lui est associé

\[
(X, Z) \xrightarrow{(d, i_1)} (B_{Z}(A_{1}^{k}), A_{1}^{k}) \xleftarrow{(d', i_0)} (P_{Z}X, Z).
\]

Les carrés correspondants à \((d, i_1)\) et \((d', i_0)\) sont transverses. On est donc réduit au cas où \((X, Z) = (P_{Z}X, Z)\). Dans ce cas, l’immersion fermée \( i \) admet une rétraction et le morphisme \( \rho_{X, Z} \) (resp. \( \rho_{Z} \)) est déterminé de manière unique par \( \rho_{X} \).

\[\Box\]

3.3. Théorème et démonstration. —

Théorème 3.7. — Les foncteurs

\[
\begin{align*}
HI_{*}(k) & \Rightarrow \mathcal{M}Cycl(k) \\
F_{*} & \Rightarrow \hat{F}_{*} \\
F_{*}^{\phi} & \leftarrow \phi
\end{align*}
\]

sont des équivalences de catégories quasi-inverses l’une de l’autre.

Démonstration. — Il s’agit de construire les deux isomorphismes naturels qui réalisent l’équivalence. Premier isomorphisme : Considérons un module de cycles \( \phi \), \( F_{*}^{\phi} \) le module homotopique associé. Par définition, pour tout corps de fonctions \( E \), il existe une flèche canonique

\[
a_{E} : \hat{F}_{*}^{\phi}(E) = \lim_{A \subset E} A^{0}(\text{Spec}(A); \phi)_{n} \rightarrow \phi_{n}(E).
\]

C’est trivialement un isomorphisme et il reste à montrer que \( a \) définit un morphisme de modules de cycles. La compatibilité à (D1) est évidente. La compatibilité à (D2) résulte du fait que pour un morphisme fini surjectif \( f : Y \rightarrow X \), le morphisme \( A^{0}(f, \phi) \) est le pushout \( f_{*} \) propre (cf. [Dég08b, 6.6]).

Compatibilité à (D3) : On reprend les notations du point (D3) de 3.1 pour le module homotopique \( F_{*}^{\phi} \) et pour une unité \( x \in E^{\times} \). On considère la flèche canonique

\[
a_{E}^{0} : \hat{F}_{n}^{\phi}(G_{m} \times (E)) = \lim_{A \subset E} A^{0}(\text{Spec}(A[t, t^{-1}]; \phi)_{n} \rightarrow \phi_{n}(E(t)).
\]

Documenta Mathematica 16 (2011) 411–455
Pour tout $E$-point $y$ de $\text{Spec}(E[t])$, on note $v_y$ la valuation de $E(t)$ correspondante, d’uniformisante $t - y$. D’après la proposition 2.3, le diagramme suivant est commutatif :

$$
\begin{array}{ccc}
\hat{F}_n^\phi(G_m \times (E)) & \xrightarrow{s'_\nu} & \hat{F}_n^\phi(E) \\
\downarrow^{a_E} & & \downarrow^{a_E} \\
\phi_n(E(t)) & \xrightarrow{s'_n} & \phi_n(E).
\end{array}
$$

Par définition du morphisme structural $\epsilon_*$ de $F_n^\phi$ (cf. 2.10), le morphisme $\nu' : \hat{F}_n^\phi(E(t)) \xrightarrow{(s_n - 1)} (\hat{F}_n^\phi)_{-1}(E) \xrightarrow{\partial} \hat{F}_n^\phi(G_m \times (E))$ est la section de la suite exacte courte

$$0 \rightarrow \hat{F}_n^\phi(E) \xrightarrow{\partial} \hat{F}_n^\phi(G_m \times (E)) \xrightarrow{a_E} \hat{F}_{n-1}^\phi(E) \rightarrow 0$$

qui correspond à la rétraction $s^*_1$ de $p^*$, pour $s_1 : (E) \rightarrow G_m \times (E)$ la section unité de la projection $p : G_m \times (E) \rightarrow (E)$. En particulier, $\nu'$ est caractérisé par les propriétés $\partial \nu' = 1$ et $s^*_1 \nu' = 0$.

Notons $\varphi : E \rightarrow E(t)$ l’inclusion canonique. On peut vérifier en utilisant les relations des pré-modules de cycles les formules suivantes :

1. $\forall \rho \in \phi_n(E), \partial_{\nu}(\{t\}.\varphi_*(\rho)) = \rho.$
2. $\forall y \in E^\times, \forall \rho \in \phi_n(E), \partial_{\nu}(\{t\}.\varphi_*(\rho)) = 0.$
3. $\forall y \in E^\times, \forall \rho \in \phi_n(E), s_{\nu}(\{t - y\}.\varphi_*(\rho)) = \{y\}.\rho.$

D’après (2), l’application $\phi_n(E) \rightarrow \phi_n(E(t)), \rho \mapsto \{t\}.\varphi_*(\rho)$ induit une unique flèche pointillée rendant le diagramme suivant commutatif :

$$
\begin{array}{ccc}
\hat{F}_n^\phi(E) & \xrightarrow{a_E} & \hat{F}_n^\phi(G_m \times (E)) \\
\downarrow^{a_E} & & \downarrow^{a_E} \\
\phi_n(E(t)) & \xrightarrow{(t).\varphi_*} & \phi_n(E).
\end{array}
$$

D’après la relation (1) et la relation (3) avec $y = 1$, cette flèche pointillée satisfait les deux propriétés caractérisant $\nu'$. On déduit donc de la relation (3) avec $y = x$ que $\nu' \circ s^*_1(\rho) = \{x\}.\rho$ ce qui prouve la relation attendue.

**Compatibilité à (D4)** : Considérons les notations du point (D4) dans 3.1. La compatibilité au résidu est alors une conséquence directe du lemme 3.6 appliqué, pour tout voisinage ouvert $U$ de $x$ dans $X$, à l’immersion fermée $i : Z \cap U \rightarrow U$ dans le cas $c = 1$, $p = 1$.

Deuxième isomorphisme : Considérons un module homotopique $(F_*, \epsilon_*)$. Pour tout schéma lisse $X$, en considérant la limite inductive des morphismes de restriction $F(X) \rightarrow F(U)$ pour les ouverts $U$ de $X$, on obtient une flèche $F_*(X) \rightarrow \mathcal{C}(X; \hat{F}_*)$ qui induit par définition des différentielles un morphisme $b_X : F_*(X) \rightarrow A^0(X; \hat{F}_*)$ homogène de degré 0.

Le point clé est de montrer que cette flèche est naturelle par rapport aux correspondances finies. Soit $\alpha \in c(X, Y)$ une correspondance finie entre schémas lisses, que l’on peut supposer connexes. Rappelons que pour tout ouvert dense $j : U \rightarrow X$, le morphisme $j^* : A^0(X; \hat{F}_*) \rightarrow A^0(U; \hat{F}_*)$ est injectif d’après la
suit exacte de localisation (2.5.c). Ainsi, on peut remplacer \( \alpha \) par \( \alpha \circ j \) et \( X \) par \( U \). Par additivité, on se ramène encore au cas où \( \alpha \) est la classe d’un sous-schéma fermé intègre \( Z \) de \( X \times Y \), fini et dominant sur \( X \). Dès lors, \( \alpha \circ j = [Z \times_X U] \). Donc puisque \( k \) est parfait, quitte à réduire \( X \), on peut supposer que \( Z \) est lisse sur \( k \). Rappelons que d’après 2.5, \( \alpha^* = p_* t^* q^* \) pour les morphismes évidents suivants

\[
X \xrightarrow{p} Z \xrightarrow{t} Z \times X \xrightarrow{q} Y.
\]

On est donc ramené à vérifier la naturalité dans les trois cas suivants :

Premier cas : Si \( \alpha = q \) est un morphisme plat, la compatibilité résulte alors de la définition du pullback plat sur \( A^0(\cdot ; F_* \cdot) \) est de la définition de D1.

Deuxième cas : Si \( \alpha = t p \), \( p : Z \to X \) morphisme fini surjectif entre schémas lisses. Ce cas résulte de la définition du pushout propre sur \( A^0 \) et de la définition de D2.

Troisième cas : Supposons \( \alpha = i \), pour \( i : Z \to X \) immersion fermée régulière entre schémas lisses. Comme on l’a déjà vu, l’assertion est locale en \( X \). On se réduit donc en factorisant \( i \) au cas de codimension 1. On peut aussi supposer que \( Z \) est un diviseur principal paramétré par \( \pi \in O_X(U) \), pour \( U = X - Z \). D’après la proposition 2.3, on est ramené à montrer que le diagramme suivant est commutatif :

\[
\begin{array}{ccc}
F_* (X) & \xrightarrow{i^*} & F_* (Z) \\
\parallel & & \parallel \\
\hat{F}_* (\kappa (X)) & \xrightarrow{s^*} & \hat{F}_* (\kappa (Z)).
\end{array}
\]

Tenant compte de la naturalité du morphisme structural \( \epsilon_* \) du module homotopique \( F_* \), on se ramène à la commutativité du diagramme :

\[
\begin{array}{ccc}
\varphi (M (X) \{1\}) & \xrightarrow{i^*} & \varphi (M (Z) \{1\}) \\
\varphi (M (U) \{1\}) & \xrightarrow{\nu} & \varphi (M (G_m \times U)) & \xrightarrow{\gamma_\pi} & \varphi (M (U)) & \xrightarrow{\partial_{X,Z}} & \varphi (M (Z) \{1\}) \\
\end{array}
\]

où \( \nu \) est l’inclusion canonique, \( \gamma_\pi \) est induit par \( \pi : U \to G_m \) et \( \partial_{X,Z} = \partial_X \circ p^{-1} \circ p_1 \) avec les notations de 3.1(D4) est le morphisme résidu au niveau des motifs. Or la commutativité de ce diagramme résulte exactement de [Dég08b, 2.6.5].

Le morphisme \( b : F_* \to A^0(\cdot ; F_* \cdot) \) est donc un morphisme de faisceaux avec transferts. Or, il est évident que le morphisme induit sur les fibres en un corps de fonctions quelconque est un isomorphisme. Il en résulte (cf. 1.4) que \( b \) est un isomorphisme. Enfin, on établit facilement la compatibilité de \( b \) avec les morphismes structuraux des modules homotopiques \( F_* \) et \( A^0(\cdot ; F_* \cdot) \) compte tenu de la construction 2.10 – on utilise simplement la fonctorialité de \( b \) par rapport à \( j_X : G_m \times X \to A^1_X \) et \( s_1 : X \to G_m \times X \).

\[\square\]

3.8. — Le théorème précédent montre que la catégorie des modules de cycles est monoidale symétrique avec pour élément neutre le foncteur de K-théorie de
Milnor. Le produit tensoriel est de plus compatible au foncteur de décalage de
la graduation des modules de cycles – i.e. le foncteur noté \{±\} dans \( HI_*(k) \).
A tout schéma lisse \( X \), on associe un module de cycles
\[
\hat{h}_0,*(X) = (h_0,*(X))^\wedge.
\]
D’après le théorème précédent, la famille de modules de cycles \( \hat{h}_0,*(X)\{n\} \)
pour un schéma lisse \( X \) et un entier \( n \in \mathbb{Z} \) est génératrice dans la catégorie
abélienne \( \mathcal{MCycl}(k) \).
Notons que ces générateurs caractérisent le produit tensoriel des modules de
cycles:
\[
\hat{h}_0,0(X,E) = \hat{h}_0,*(X) \otimes \hat{h}_0,*(Y) = \hat{h}_0,*(X \times Y)\{n + m\}.
\]
On peut enfin donner une formule explicite pour calculer ces modules de cycles.
Considérons pour tous schémas lisses \( X \) et \( Y \)
\[
\pi(Y, X) = \ker \left( c(A^1_{Y,X}) \rightarrow c(Y, X) \right).
\]
Notons que ce groupe s’étend de manière évidente aux schémas réguliers essen-
tiellement de type fini sur \( k \) et que l’on dispose d’un théorème de commutation
aux limites projectives de schémas pour ces groupes étendus (cf. \[Dég07, 4.1.24\]). Par ailleurs, si \( E \) est un corps de fonctions, et \( X \) un schéma projectif
lisse, \( \pi(\text{Spec}(E), X) = CH_0(X_E) \), groupe de Chow des 0-cycles de \( X \) étendu à
\( E \).
On déduit de tout cela les calculs suivants: pour tout corps de fonctions \( E \) et
tout schéma projectif lisse \( X \),
\[
\hat{h}_0,0(X,E) = CH_0(X_E).
\]
De plus, pour tout entier \( n > 0 \),
\[
\hat{h}_0,n(X,E) = \ker \left( \oplus_{i=0}^n CH_0(G_{m,E}^{n-1} \times X_E) \rightarrow CH_0(G_{m,E}^n \times X_E) \right)
\]
où les flèches sont induites par les injections évidentes \( G_{m,E}^i \times \{1\} \times G_{m,E}^{n-1-i} \rightarrow G_{m,E}^n \).

3.4. Résolution de Gersten: fonctorialité II. —

3.9. — Dans ce paragraphe, on complète les résultats du paragraphe 3.2. On
fixe un module de cycles \( \phi \) et on note \( F^2 \) le module homotopique qui lui est
associé. On peut étendre la construction de \textit{loc. cit.} au cas d’un \( k \)-schéma de
type fini \( X \): on associe à ce schéma un complexe de faisceaux sur \( X_{\text{Nis}} \):
\[
C^r_X(\phi) : V/X \mapsto C^r(V; \phi)
\]
et un faisceau \( F^2_X = H^0(C^\infty_X) \). Le complexe \( C^\infty_X(\phi) \) vérifie encore la propriété
de Brown-Gersten mais par contre, le morphisme canonique:
\[
F^2_X \rightarrow C^\infty_X(\phi)
\]
n’est plus nécessairement un isomorphisme. Cette construction nous sert à
montrer le résultat suivant:
Proposition 3.10. — Considérons les notations ci-dessus. Alors, l’isomorphisme $\rho_X : A^p(X,\phi) \to H^p(X; F^\phi)$ pour un schéma lisse $X$ (cf. (3.2.b)) est naturel par rapport aux correspondances finies.

Démonstration. — La preuve reprend le principe de la preuve du lemme 3.3. Soit $f : Y \to X$ un morphisme plat ou une immersion fermée régulière entre schémas de type fini. Il est clair que les constructions de loc. cit. se généralisent et permettent de définir un morphisme canonique $\eta_f$ qui s’insère dans un diagramme commutatif de la catégorie dérivée des faisceaux abéliens sur $X_{Nis}$:

$\xymatrix{ F_X^\phi \ar[r]^\tau_f \ar[d] & Rf_Y F_Y^\phi \ar[d] \\
C_X^\phi(\phi) \ar[r]^\eta_f & Rf_Y C_Y^\phi(\phi).}$

Par ailleurs, si $p : Z \to X$ est un morphisme fini, il induit (cf. 2.1) un morphisme de complexes

$p_* : C^*(Z;\phi) \to C^*(X;\phi)$

qui est naturel en $X$ par rapport aux morphismes étalés (cf. [Ros96, (4.1)]). On en déduit un morphisme canonique $tr_p : p_* C_Z^\phi(\phi) \to C_X^\phi(\phi)$ qui induit un diagramme commutatif dans la catégorie dérivée ($p_*$ est exact):

$\xymatrix{ Rp_* F_Z^\phi \ar[r]^{tr_p^0} \ar[d] & F_X^\phi \ar[d] \\
Rp_* C_Z^\phi(\phi) \ar[r]^{tr_p} & C_X^\phi(\phi).}$

Revenons à la preuve de la proposition. Il suffit de montrer la naturalité de $\rho_X$ pour une correspondance finie $\alpha \in c(X,Y)$ telle que $\alpha$ est la classe d’un sous-schéma fermé intègre $Z$ de $X \times Y$. Suivant le paragraphe 2.4, on considère les morphismes:

$X \xleftarrow{p} Z \xrightarrow{q} ZXY \xrightarrow{i} Y.$

D’après loc. cit., $\alpha^* = p_* i^* q^*$. Appliquant les constructions qui précèdent, on obtient un diagramme commutatif dans la catégorie dérivée des groupes abéliens:

$\xymatrix{ H^p(Y; F_Y^\phi) \ar[r]^{\tau_q^*} \ar[d] & H^p(ZXY; F_Y^\phi) \ar[r]^{\tau_i^*} \ar[d] & H^p(Z; F_Z^\phi) \ar[r]^{\tau_p^*} \ar[d] & H^p(X; F_X^\phi) \\
H^p(Y; C_Y^\phi(\phi)) \ar[r]^{\eta_q^*} & H^p(ZXY; C_Y^\phi(\phi)) \ar[r]^{\eta_i^*} & H^p(Z; C_Z^\phi(\phi)) \ar[r]^{\eta_p^*} & H^p(X; C_X^\phi(\phi))}$

On vérifie que la composée des flèches de la première ligne coïncide avec $\alpha^*$ et cela permet de conclure. 

\[\Box\]
3.11. — Soit $F_*$ un module homotopique, $\phi = \hat{F}_*$ sa transformée générique. Considérons l’isomorphisme $b : F_* \to F_*^p$ qui lui est associé d’après le théorème 3.7. Compte tenu de l’isomorphisme (3.2.b), on en déduit un isomorphisme

$$(3.11.a) \quad \epsilon_X : H^p_{\text{Nis}}(X; F_*) \xrightarrow{b} H^p_{\text{Nis}}(X; F_*^p) \xrightarrow{\delta^*} A^p(X; \hat{F}_*).$$

La proposition précédente a comme corollaire immédiat:

**Corollaire 3.12.** — Avec les notations ci-dessus, $\epsilon_X$ est un isomorphisme naturel en $X$ par rapport aux correspondances finies.

3.13. — Considérons le module homotopique $S_\mathfrak{t}$. D’après le théorème de Suslin-Voevodsky rappelé en 1.11, pour tout corps de fonctions $E$, $S_\mathfrak{t}(E) \simeq K^M_*(E)$. Cet isomorphisme est de plus compatible aux structures de module de cycles. Pour la norme, cela résulte de [SV00, 3.4.1]. Pour le résidu associé à un corps de fonctions valué $(E, v)$, on se réduit à montrer que $\delta_0(\pi) = 1$ pour le module de cycle $S_\mathfrak{t}$, ce qui résulte de [Dégl07, 2.6.5].

On en déduit l’isomorphisme de Bloch(14) pour tout schéma lisse $X$ :

$$\epsilon^B_X : H^p_{\text{Nis}}(X; S^\mathfrak{t}_*) \to A^n(X; K^M_*)_n = CH^n(X).$$

On a obtenu ci-dessus que cet isomorphisme est compatible aux transferts. Rappelons que pour tout module de cycles $\phi$, il existe un accouplement de modules de cycles $K^M_\times \phi = \phi$ au sens de [Ros96, 1.2]. Il induit d’après [Ros96, par. 14] un accouplement

$$CH^n(X) \otimes A^m(X; \phi)_r \to A^{m+n}(X; \phi)_{r+n}.$$ 

Considérant un module homotopique $F_*$, on dispose d’un (iso)morphisme de modules homotopiques $S_\mathfrak{t} \otimes F_* \to F_*$. Pour un schéma lisse $X$, de diagonale $\delta : X \to X \times X$, on en déduit un accouplement

$$(H^n(X; S^\mathfrak{t}_*)_n \otimes H^m(X; F_*)_r) \to H^{m+n}(X; F_*^{r+n}).$$

défini en associant à deux morphismes $a : h_{0,*}(X) \to S^\mathfrak{t}_* \{n\}[n]$ et $b : h_{0,*}(X) \to F_* \{r\}[m]$ la composée

$h_{0,*}(X) \xrightarrow{b} h_{0,*}(X) \otimes h_{0,*}(X) \xrightarrow{\Delta \otimes b} S^\mathfrak{t}_* \otimes F_* \{n+r\}[n+m] \xrightarrow{\tilde{\sim}} F_* \{n+r\}[n+m].$

Nous laissons au lecteur le soin de vérifier la compatibilité suivante :

**Lemme 3.14.** — Avec les notations introduites ci-dessus, le diagramme suivant est commutatif :

$$\begin{array}{c}
H^n(X; S^\mathfrak{t}_*)_n \otimes H^m(X; F_*)_r, \xrightarrow{\epsilon^B_X \otimes \epsilon_X} H^{n+m}(X; F_*^{r+n}), \xrightarrow{\epsilon_X} CH^n(X) \otimes A^m(X; \hat{F}_*)_r, \xrightarrow{\epsilon_X} A^{m+n}(X; \hat{F}_*)^{r+n}.
\end{array}$$

(14) En effet, d’après l’isomorphisme que l’on vient d’expliquer, le faisceau gradué $S^\mathfrak{t}_*$ est le faisceau de K-théorie de Milnor non ramifié.

**Documenta Mathematica** 16 (2011) 411–455
Ainsi, l’isomorphisme $\epsilon_X$ est compatible au produit, et l’isomorphisme $\epsilon^b_X$ est compatible aux structures de module décrites ci-dessus.

3.15. — Notons $\varphi : DM_{gm}(k)^{op} \to \mathcal{S}h$ le foncteur de réalisation associé à $F_\ast$ (cf. section 1.3). D’après la proposition précédente, le foncteur $\varphi$ prolonge le foncteur $A^\ast(\cdot ; \hat{F}_\ast)$. Ainsi, on a étendu canoniquement la cohomologie à coefficients dans un module de cycles quelconque en un foncteur de réalisation triangulé de $DM_{gm}(k)$. Nous notons encore

$$\epsilon^b_X : \varphi(M(X) \{-r\}[-n]) \to A^n(X; \hat{F}_\ast)_r$$

l’isomorphisme qui se déduit par construction de l’isomorphisme (3.11.a).

Soit $f : Y \to X$ un morphisme projectif entre schémas lisses, de dimension relative constante $d$. Dans [Dég08a, 2.7], on a construit $f^* : M(X)(d)[2d] \to M(Y)$, morphisme de Gysin associé à $f$ dans $DM_{gm}(k)$.

**Proposition 3.16.** — Considérons les notations introduites ci-dessus. Alors, le carré suivant est commutatif :

$$\begin{array}{ccc}
\varphi(M(X)\{d-r\}[d-n]) & \xrightarrow{\varphi(f^*)} & \varphi(M(Y)\{-r\}[-n]) \\
\epsilon_X \downarrow & & \downarrow \epsilon_Y \\
A^{n-d}(X; \hat{F}_\ast)_{r-d} & \xrightarrow{f^*} & A^n(Y; \hat{F}_\ast)_r
\end{array}$$

**Démonstration.** — Dans cette preuve, on utilisera particulièremment le lemme suivant :

**Lemme 3.17.** — Soit $X$ un schéma lisse et $E/X$ un fibré vectoriel de rang $n$. Soit $p : P \to X$ le fibré projectif associé, et $\lambda$ le fibré inversible canonique sur $P$ tel que $\lambda \subset p^{-1}(E)$. On note $c = c_1(\lambda) \in CH^1(X)$ la première classe de Chern de $\lambda$.

Alors, le morphisme suivant est un isomorphisme :

$$\bigoplus_{i=0}^n A^\ast(X; \hat{F}_\ast) \to A^\ast(P; \hat{F}_\ast),$$

$$x_i \mapsto p^\ast(x_i).c^\prime.$$

en utilisant la structure de $CH^\ast(X)$-module (ici notée à droite) de $A^\ast(X; \hat{F}_\ast)$ rappelée en 3.13.

Pour obtenir ce lemme, il suffit d’appliquer le théorème du fibré projectif dans $DM_{gm}(k)$ (cf. [Voe00b, 2.5.1]) et de regarder son image par $\varphi$ compte tenu du lemme 3.14.

Soit $E$ un fibré vectoriel sur $X$ et $P$ sa complétion projective. On déduit de ce lemme le cas où $f = p$. En effet, d’après la formule de projection

$$p_\ast(p^\ast(x_i).c^\prime) = x_i.p_\ast(c')$$

pour les groupes de Chow à coefficients (cf. [Dég06, 5.9]), on déduit que $p_\ast$ est la projection évidente à travers le théorème du fibré projectif. L’analogue de ce calcul pour $\varphi(p^\ast)$ résulte des définitions de [Dég08a].
Compte tenu de la définition du morphisme de Gysin et du cas que l'on vient de traiter, nous sommes ramenés au cas où \( f = i : Z \to X \) est une immersion fermée, que l'on peut supposer de codimension pure égale à \( c \). Ce cas est alors une conséquence directe du lemme 3.6.

PARTIE II

MOTIFS MIXTES TRIANGULÉS

4. Rappels

Dans cette section, on rappelle la théorie de Voevodsky des complexes motiviques et l'extension qu'on lui a donnée avec D.C. Cisinski suivant les lignes de Morel et Voevodsky.

4.1. Catégorie effective. —

4.1. — La catégorie \( Sh^r(k) \) (cf. paragraphe 1.2) est abélienne de Grothendieck. On note \( \mathcal{F}_A \), la sous-catégorie triangulée localisante de la catégorie dérivée \( D(Sh^r(k)) \) engendrée par les complexes de la forme:
\[
\cdots \to \mathbb{Z}^r(A^1_X) \to \mathbb{Z}^r(X) \to 0 \cdots
\]

Définition 4.2 (Voevodsky). — On définit la catégorie des motifs effectifs comme le quotient de Verdier:

\[
DM^{eff}(k) = D(Sh^r(k))/\mathcal{F}_A.
\]

Suivant les idées de la théorie de l'homotopie des schémas de Morel et Voevodsky, on peut décrire cette catégorie grâce à la notion de localisation de Bousfield. Le concept central dans cette théorie est le suivant:

Définition 4.3. —

1. Soit \( C \) un complexe de faisceaux avec transferts.
   On dit que \( C \) est \( \mathbb{A}^1 \)-local si pour tout schéma lisse \( X \) et tout entier \( n \in \mathbb{Z} \), le morphisme suivant, induit par la projection canonique, est un isomorphisme:
   \[
   H^n_{\text{Nis}}(X, C) \to H^n_{\text{Nis}}(\mathbb{A}^1_X, C).
   \]
   On dit que \( C \) est \( \text{Nis-fibrant} \) si pour tout schéma lisse \( X \) et tout entier \( n \in \mathbb{Z} \), le morphisme canonique suivant est un isomorphisme:
   \[
   H^n(C(X)) \to H^n_{\text{Nis}}(X, C).
   \]
   On dit que \( C \) est \( \mathbb{A}^1 \)-fibrant si il est \( \mathbb{A}^1 \)-local et \( \text{Nis-fibrant} \).

2. Soit \( f : C \to D \) un morphisme de \( C(Sh^r(k)) \).
   On dit que \( f \) est une \( \mathbb{A}^1 \)-équivalence si pour tout complexe \( \mathbb{A}^1 \)-local \( L \), le morphisme induit
   \[
   \text{Hom}_{D(Sh^r(k))}(D, L) \to \text{Hom}_{D(Sh^r(k))}(C, L)
   \]
est un isomorphisme.

On dit que $f$ est une $A_1$-fibration si c'est un épimorphisme dans $C(Sh^{tr}(k))$ et son noyau est $A_1$-fibrant.

La proposition suivante (voir [CD09B, ex. 3.3, 4.12]) donne une bonne structure homotopique à la catégorie $DM^{eff}(k)$.

**Proposition 4.4.** — 1. La catégorie $C(Sh^{tr}(k))$ avec pour équivalences faibles les quasi-isomorphismes et pour fibrations les épimorphismes dont le noyau est Nis-fibrant est une catégorie de modèles symétrique monoïdale dont la catégorie homotopique associée est $D(Sh^{tr}(k))$.

2. La localisation de Bousfield de la catégorie de modèles précédente par rapport à la classe d'objets $\mathcal{S}_A$ est encore une catégorie de modèles symétrique monoïdale avec pour équivalences faibles les $A_1$-équivalences et pour fibrations les $A_1$-fibrations.

On déduit de cette proposition que le morphisme de projection canonique

$$L_A : D(Sh^{tr}(k)) \rightarrow DM^{eff}(k)$$

est monoïdal et admet un adjoint à droite $O : DM^{eff}(k) \rightarrow D(Sh^{tr}(k))$ pleinement fidèle. L'image essentielle de ce dernier est formée des complexes $A_1$-locaux. Par la suite, on identifie $DM^{eff}(k)$ à cette image essentielle. En particulier, pour tout complexe $C$, $L_A(C)$ est $A_1$-local.

On notera simplement $\otimes$ le produit tensoriel sur $DM^{eff}(k)$ et l'unité pour $\otimes$.

**Remarque 4.5.** — Cette description des objets de $DM^{eff}(k)$ comme complexes $A_1$-locaux n'a rien d'original. Elle est essentiellement due à Voevodsky suivant son article fondateur [Voe00a] (voir aussi le théorème 5.1 ci-dessous).

Rappelons au passage que d'après [Voe00a, 3.2.6], le foncteur canonique

$$(4.5.a) \quad DM^{eff}_{gm}(k) \rightarrow DM^{eff}(k)$$

est pleinement fidèle. Notons de plus que son image essentielle est égale à la sous-catégorie pleine formée des objets compacts de $DM^{eff}(k)$ (cf. [CD09B, ex. 5.5]).

**4.2. Catégorie non effective.** —

4.6. — Soit $T$ le conoyau du morphisme $\mathbb{Z}^{tr}(k) \xrightarrow{\Delta} \mathbb{Z}^{tr}(\mathbb{G}_m)$ induit par la section unité. Utilisant la notation de la section 1.3, on définit le complexe motivique de Tate suspendu dans $DM^{eff}(k)$ par la formule: $\mathbb{1}\{1\} := L_A(T)$.\(^{(15)}\)

Dans l'énoncé qui suit, nous dirons qu'une catégorie est monoïdale homotopique si c'est un quotient de Verdier d'une catégorie dérivée munie d'un produit tensoriel dérivé. La proposition qui suit est bien connue (cf. [Hov01] pour\(^{(15)}\) Rappelons à nouveau qu'avec la notation habituelle $1\{1\} = 1(1)[1]$. Le twist $1\{1\}$ apparaît plus naturel pour notre propos c'est pourquoi on adopte cette notation ici.)
l’aspect catégorie de modèles pure et [CD09b, 7.15] pour l’aspect catégorie dérivée):

**Proposition 4.7.** — Il existe une unique catégorie monoidal homotopique $DM(k)$ munie d’une adjonction de catégories homotopiques

\[(4.7.a) \quad \Sigma^\infty : DM^{eff}(k) \rightleftharpoons DM(k) : \Omega^\infty\]

de vérifiant

1. $\Sigma^\infty$ est monoidal,
2. $\Sigma^\infty L_{A^1}(T)$ est inversible pour le produit tensoriel sur $DM(k)$,

et qui soit universelle (initiale) pour ces propriétés.

On note encore $\otimes$ le produit tensoriel sur $DM(k)$ et $\mathbb{I}$ l’objet unité. Pour tout entier $n \in \mathbb{Z}$, on note $\mathbb{I}\{n\}$ la puissance tensorielle $n$-ième de l’objet inversible $\mathbb{I}\{1\}$ dans $DM(k)$ ; pour tout spectre motivique $K$, on pose $K\{n\} := K \otimes \mathbb{I}\{n\}$.

**4.8.** — Rappelons que la construction de $DM(k)$ reprend celle de la catégorie d’homotopie stable de la topologie algébrique. On utilise ici la version classique (i.e. non symétrique) des spectres qui sera plus commode pour notre propos.\(^{(16)}\)

On définit un spectre motivique comme une famille $(E_n, \sigma_n)_{n \in \mathbb{N}}$ telle que $E_n$ est un complexe de faisceaux avec transferts et $\sigma_n : T \otimes^r E_n \to E_{n+1}$ un morphisme de faisceaux avec transferts dit de suspension. On notera simplement $E$ pour le spectre $(E_n, \sigma_n)_{n \in \mathbb{N}}$.

On dit que $E$ est un $\Omega$-spectre si pour tout $n \in \mathbb{N}$, le complexe de faisceaux avec transferts $E_n$ est $A^1$-fibration et le morphisme adjoint à $\sigma_n$:

\[(4.8.a) \quad \tau_n : E_n \to \text{Hom}(T, E_{n+1})\]

est un quasi-isomorphisme (voir [Hov01, 3.1]).

Un morphisme $f$ de spectres motiviques est un morphisme de complexes gradués compatible avec les morphismes de suspensions. On dit que $f$ est une équivalence stable (resp. quasi-isomorphisme) si pour tout $\Omega$-spectre $E$, $\text{Hom}(f, E)$ est un isomorphisme (resp. $f$ est un quasi-isomorphisme degré par degré). La catégorie $DM(k)$ est la localisation de la catégorie des spectres motiviques par rapport aux équivalences stables (voir [Hov01, 3.4]).

**Exemple 4.9.** — Soit $X$ un schéma lisse. On peut donner la description suivante du spectre associé au complexe $A^1$-local $L_{A^1}(\mathbb{Z}^r(X))$ par le foncteur $\Sigma^\infty$:

\[(\Sigma^\infty L_{A^1}(\mathbb{Z}^r(X)))_n := L_{A^1}(T^{\otimes^r n} \otimes^r \mathbb{Z}^r(X)),\]

les morphismes de suspensions étant donnés par les applications évidentes.

On déduit de plus du théorème de simplification de Voevodsky [Voe00A], que $\Sigma^\infty L_{A^1}(\mathbb{Z}^r(X))$ est un $\Omega$-spectre. Par la suite, on le notera simplement $M(X)$.

\(^{(16)}\)Dans [CD09b], on utilise les spectres symétriques pour définir $DM(k)$ et sa structure monoidale symétrique. L’équivalence de cette définition avec celle présentée ici résulte de [Hov01, 9.1, 9.3].
**Remarque 4.10.** — Nous utiliserons par la suite les points techniques suivants concernant les spectres motiviques\(^{(17)}\):

1. Si $E$ et $E'$ sont des $\Omega$-spectres, un morphisme $f : E \to E'$ est une équivalence faible si et seulement si c'est un quasi-isomorphisme. De plus, la catégorie $DM(k)$ s'identifie à la localisation de la catégorie des $\Omega$-spectres par rapport aux quasi-isomorphismes.

2. Un triangle entre $\Omega$-spectres $E' \to E \to E'' \to E[1]$ est distingué dans $DM(k)$ si et seulement si pour tout entier $n \geq 0$, le triangle correspondant $E'_n \to E_n \to E''_n \to E_n[1]$ est distingué dans $DM^{eff}(k)$. Ce dernier triangle est en particulier distingué dans $D(Sh^{tr}(k))$.

3. Si $E$ est un $\Omega$-spectre et $n \in \mathbb{Z}$ un entier, $\Omega^\infty (E\{n\}) = E_n$.

**Remarque 4.11.** — D'après la propriété universelle de $DM(k)$, le foncteur (4.5.a) s'étend de manière unique en un foncteur:

\[
DM_{gm}(k) \to DM(k).
\]

On démontre dans [CD07] – à la suite de la définition 10.1.4 – que ce foncteur est pleinement fidèle avec pour image essentielle la sous-catégorie formée des objets compacts.

---

5. **t-structure homotopique**

Notre référence pour les t-structures est [BBD82, sec. 1.3].

5.1. **Cas effectif.** — Le théorème suivant est une reformulation du résultat central de la théorie des complexes motiviques (cf. [VOE00b, 3.1.12]):

**Théorème 5.1** (Voevodsky). — Soit $C$ un complexe de faisceaux avec transferts. Les conditions suivantes sont équivalentes :

(i) $C$ est $\mathbb{A}^1$-local.

(ii) Pour tout entier $n \in \mathbb{Z}$, $\mathbb{H}^n(C)$ est $\mathbb{A}^1$-local.

(iii) Pour tout entier $n \in \mathbb{Z}$, $\mathbb{H}^n(C)$ est invariant par homotopie.

**Démonstration.** — L'équivalence de (i) et (ii) résulte de la suite spectrale d'hypercohomologie Nisnevich. L'implication (ii) $\Rightarrow$ (iii) est évidente et sa réciproque résulte du théorème de Voevodsky loc. cit.

\(^{(17)}\)Ces assertions résultent de [HOV01, 3.4] ; plus précisément du fait que les $\Omega$-spectres sont les objets fibrants pour la *structure de catégorie de modèles stable* sur la catégorie des spectres motiviques.
Rappelons que la catégorie $D(Sh^{tr}(k))$ porte naturellement une t-structure (cf. [BBD82, 1.3.2]): un complexe $C$ est positif (resp. négatif) si pour tout $i < 0$ (resp. $i > 0$) le faisceau de cohomologie $H^i(C)$ est nul.

Corollaire 5.2. — Il existe une unique t-structure sur $DM^eff(k)$ telle que le foncteur $O : DM^eff(k) \to D(Sh^{tr}(k))$ est t-exact\(^{(18)}\).

Le foncteur canonique:

$$DM^eff(k) \to HI(k), C \mapsto H^0(C)$$

induit une équivalence de catégories entre $HI(k)$ et le cœur de $DM^eff(k)$ pour cette t-structure.

Suivant Voevodsky, on appelle cette t-structure sur $DM^eff(k)$ la t-structure homotopique.

Considérons les notations qui suivent la définition 1.3. Si $F$ est un faisceau avec transferts, on obtient l’identification: $H^0(L_{A^1}F) = h_0(F)$ (\(^{(19)}\)) En particulier, pour tout schéma lisse $X$:

\[(5.2.a) \quad H^0(L_{A^1}Z^{tr}(X)) = h_0(X).\]

Rappelons au passage le calcul du motif de Tate (voir [SV00, 3.2]):

\[(5.2.b) \quad \mathbb{1}\{1\} := L_{A^1}(T) = S^1_1.\]

5.2. Cas non effectif. — Notons que le Hom interne des complexes de faisceaux avec transferts se dérive à droite pour la structure de catégorie de modèles du point 2 de la proposition 4.4. On le note $R_{A^1}\text{Hom}$. Le théorème suivant nous sera essentiel. C’est un corollaire de la théorie de Voevodsky: il résulte de [Voe00a, 4.34].

Théorème 5.3. — L’endofoncteur de $DM^eff(k)$ défini par la formule

$$K \mapsto R_{A^1}\text{Hom}(S^1_1, K)$$

est t-exact et sa restriction au cœur de $DM^eff(k)$, identifié avec $HI(k)$, coïncide avec le foncteur $F \mapsto F_{-1}$ défini au paragraphe 1.13.

5.4. — Considérons un $\Omega$-spectre $E = (E_n, \sigma_n)$ au sens du paragraphe 4.8. Fixons un entier $n \geq 0$. On associe à $E$ un faisceau homotopique par la formule:

$$H^0_n(E) = \mathbb{H}^0(E_n).$$

Considérons l’isomorphisme (4.8.a). Notons que, puisque $T$ est cofibrant et $E_{n+1}$ est $A^1$-fibrant pour la structure de catégorie de modèles du point (2) de 4.4, on obtient:

$$\text{Hom}(T, E_{n+1}) = R_{A^1}\text{Hom}(T, E_{n+1}) = R_{A^1}\text{Hom}(S^1_1, E_{n+1})$$

\(^{(18)}\) i.e. il respecte les objets positifs ainsi que les objets négatifs.

\(^{(19)}\) En effet, $h_0(F)$ est $A^1$-local d’après le théorème précédent et le morphisme canonique $F \to h_0(F)$ est une $A^1$-équivalence faible (voir aussi [VOE00B, 3.2.3]).
D’après le théorème précédent, le morphisme $H^0(\tau_n)$ induit donc un isomorphisme:

$$\epsilon_n : H^0_n(E) \to (H^0_{n+1}(E))_{-1}$$

Si l’on pose $H^0_{-n}(E) := (H^0_n(E_0))_{-n}$, on a défini ainsi un module homotopique (cf. définition 1.17) que l’on note $H^0_n(E)$. On a construit ainsi un foncteur $H^0_n$ sur les $\Omega$-spectres. Puisqu’il respecte manifestement les quasi-isomorphismes, il induit, d’après le premier point de la remarque 4.10, un unique foncteur:

$$H^0 : DM(k) \to HI_*(k).$$

Pour tout entier $m \in \mathbb{Z}$, on pose: $H^m_0(E) = H^0(E[m])$.

**Lemme 5.5.** — Considérons les notations introduites ci-dessus.

1. Le foncteur $H^0 : DM(k) \to HI_*(k)$ est un foncteur cohomologique qui commute aux sommes quelconques.
2. La famille de foncteurs $(H^m_0)_{m \in \mathbb{Z}}$ est conservative.

**Démonstration.** — Le point 1 résulte des propriétés analogues du foncteur $H^0 : DM^{eff}(k) \to HI(k)$, du deuxième point de la remarque 4.10, et du lemme 1.14. Concernant le point 2, d’après la remarque 4.10(1), on se ramène à montrer que pour un morphisme $f : E \to E'$ entre $\Omega$-spectres, les conditions suivantes sont équivalentes:

(i) $f$ est un quasi-isomorphisme (au sens du paragraphe 4.8).

(ii) pour tout $m \in \mathbb{Z}$, $H^m_0(f)$ est un isomorphisme de modules homotopiques.

Par définition du foncteur $H^m_0$, cette équivalence résulte du corollaire 5.2 et du lemme 1.14.

On dit qu’un spectre motivique est positif (resp. négatif) si pour tout $n < 0$ (resp. $n > 0$), $H^m_n(E) = 0$. Soit $\tau_{\leq 0}$ le foncteur de troncature négative pour la $t$-structure homotopique sur $DM^{eff}(k)$. On vérifie en utilisant à nouveau le théorème 5.3 que l’application de $\tau_{\leq 0}$ degré par degré à un $\Omega$-spectre $E$ définit un $\Omega$-spectre négatif $\tau_{\leq 0}E$ et un morphisme canonique:

$$\tau_{\leq 0}E \to E.$$
1. Le diagramme suivant est commutatif:

\[
\begin{array}{ccc}
DM(k) & \xrightarrow{\omega^\infty} & H_\infty(k) \\
\downarrow & & \downarrow \\
DM^{eff}(k) & \xrightarrow{\omega} & HI(k)
\end{array}
\]

avec la notation de (1.18.b) pour \( \omega^\infty \). En particulier, \( \Omega^\infty \) est t-exact.

2. Pour tout objet \( E \) de \( DM(k) \) et tout entier \( n \in \mathbb{Z} \),

\[
H^0(E[n]) = H^0(E)\{n\}
\]

en utilisant la notation du paragraphe 1.16 pour le membre de droite. Ainsi, le produit tensoriel par \( \mathbb{I}\{1\} \) est t-exact.

3. Pour tout schéma lisse \( X \),

\[
H^0(M(X)) = h_0, (X)
\]

avec la notation de l’exemple 4.9 (resp. (1.18.a)) à gauche (resp. à droite). La première assertion résulte du troisième point de la remarque 4.10. La deuxième assertion se déduit du cas \( n = -1 \) qui résulte lui-même de la définition. La troisième résulte de la remarque 4.9 et de (5.2.a).

Notons finalement qu’un objet \( E \) de \( DM(k) \) est positif si et seulement si pour tout schéma lisse \( X \) et tout couple d’entiers \( (n, i) \in \mathbb{Z} \times \mathbb{N}^* \),

\[
\text{Hom}_{DM(k)}(M(X)\{n\}[i], E) = 0.
\]

On en déduit aisément la proposition suivante:

**Proposition 5.8.** — Le produit tensoriel sur \( DM(k) \) est t-exact à droite (i.e. préserve les objets négatifs).

### 5.3. Coeur homotopique

— Notons \( H^0(DM(k)) \) le coeur de \( DM(k) \) pour la t-structure homotopique de la proposition 5.6. Notons le corollaire suivant de la proposition 5.8:

**Corollaire 5.9.** — Le produit tensoriel sur \( DM(k) \) induit une structure monoïdale symétrique sur \( H^0(DM(k)) \) dont l’objet unité est \( H^0(\mathbb{I}) \) et le produit tensoriel est défini par la formule:

\[
E \otimes E' := H^0(E \otimes E').
\]

Ce corollaire résulte plus précisément du lemme suivant laissé au lecteur\(^{(20)}\):

**Lemme 5.10.** — Soit \( \mathcal{F} \) une catégorie triangulée monoïdale munie d’une t-structure telle que le produit tensoriel est t-exact à droite. Alors, pour tous objets \( K \) et \( L \) négatifs de \( \mathcal{F} \),

\[
H^0(K \otimes L) \simeq H^0(H^0(K) \otimes L).
\]

\(^{(20)}\)On fera attention toutefois que ce lemme est faux sans bornes sur \( K \) et \( L \).
L’identification du cœur de la t-structure homotopique sur $DM(k)$ est maintenant aisée:

\[ \text{Théorème 5.11. — Le foncteur } H^0 \text{ induit une équivalence de catégories abéliennes monoïdales:} \]

\[ H^0(DM(k)) \rightarrow HI_*(k) \simeq \mathcal{M}\text{Cycl}(k), \]

l’équivalence de droite étant celle du théorème 3.7.

\[ \text{Démonstration. — Tout module homotopique définit évidemment un } \Omega\text{-spectre au sens du paragraphe 4.8. On vérifie facilement que cela définit un quasi-inverse au foncteur de l’énoncé.} \]

Concernant les structures monoïdales, on note tout d’abord que $H^0(\Omega_1) = K^M_*$ compte tenu du théorème 1.11 et du point 2 du paragraphe 5.7. D’après la formule (5.9.a), et les points 2 et 3 du paragraphe loc. cit., pour tous schémas lisses $X$, $Y$ et tout couple d’entiers $(n, m) \in \mathbb{Z}^2$, on obtient:

\[ h^0_*(X) \{n\} \otimes h^0_*(Y) \{m\} = H^0(\Sigma^\infty M(X) \{n\}) \otimes H^0(\Sigma^\infty M(X) \{n\}) \]

\[ = H^0(\Sigma^\infty M(X \times Y) \{n + m\}) = h^0_*(X \times Y) \{n + m\}. \]

Cela conclut d’après le lemme 1.19. \qed

\[ \text{Remarque 5.12. — Reprenons les notations de la section 3.1. Utilisant le foncteur pleinement fidèle (4.11.a), on définit un foncteur:} \]

\[ DM_{gen}(k) \rightarrow \text{pro-}DM(k) \xrightarrow{H^0} \text{pro-}HI_*(k). \]

On vérifie que ce foncteur est pleinement fidèle.\(^{(21)}\)

Il en résulte que la catégorie des motifs génériques est la sous-catégorie pleine de la catégorie abélienne pro-$HI_*(k)$, formée des pro-objets de la forme $H^0(M(L) \{n\})$ pour un corps de fonctions $L/k$ et un entier $n \in \mathbb{Z}$. Ces pro-objets correspondent à des foncteurs fibres de la catégorie $HI_*(k)$ (i.e. exacts, commutant aux sommes infinies). Cette interprétation des motifs génériques est donc très proche de la notion de points d’un topos. La transformée générique d’un module homotopique $F_*$ est finalement donnée par la restriction de $F_*$ à cette catégorie de points.\(^{(22)}\)

6. Applications et compléments

6.1. Construction de modules de cycles. —
6.1. — Pour un objet $M$ de $DM(k)$, on note $\hat{H}^0(M)$ la transformée générique (par. 3.1) du module homotopique $H^0(M)$ (par. 5.4). Cette construction nous permet de construire des modules de cycles intéressants. Ainsi, pour tout schéma algébrique $X$, on peut définir suivant [Voe00b] un complexe motivique $C^\infty_{\text{ring}} Z^{tr}(X)$ – qui coïncide avec le complexe motivique $\hat{M}(X)$ lorsque $X$ est lisse. Pour tout entier $i \geq 0$, on pose donc:

$$(6.1.a) \quad \hat{h}_i(X)_L := H^\text{sing}_i(\Sigma^\infty C^\infty_{\text{ring}} Z^{tr}(X)).$$

Pour tout corps de fonctions $L$, le gradué de degré 0 de ce module de cycles est donné par l'homologie de Suslin de $X$:

$$\hat{h}_{0,0}(X)_L = H^\text{sing}_s(\mathcal{G}^m_X \times X_L/L)$$

avec les notations de [SV96].

Si $X$ est projectif lisse de dimension $d$, le motif $\hat{M}(X) = \Sigma^\infty C^\infty_{\text{ring}} Z^{tr}(X)$ dans $DM(k)$ est fortement dualisable avec pour dual fort $\hat{M}(X)(-d)[-2d]$.

Il en résulte que pour tout corps de fonctions $L$,

$$(6.1.b) \quad \hat{h}_{i,n}(X)_L = H^s_d(X_L)_L,$$

où $H^s_d(X_L)_L$ désigne la cohomologie motivique de $X$ étendue à $L$ en degré $s$ et twist $t$.

Remarque 6.2. — Supposons $k$ de caractéristique 0. Comme remarqué par B. Kahn dans [KAH10], on obtient un théorème de Merkurjev (cf. [MER08, Th. 2.10]) comme corollaire du théorème 3.7. En effet, pour $X$ propre et lisse, on peut identifier le module de cycles $\hat{h}_{i,*}(X)$ introduit ici avec le module de cycles $K_X$ construit par Merkurjev. On renvoie le lecteur à [KAH10] pour d'autres utilisations des modules de cycles $\hat{h}_{i,*}(X)$.

6.2. Modules de cycles constructibles. — On introduit l'hypothèse suivante sur le corps $k$:

($\mathcal{M}_k$) Pour tout corps de fonctions $E/k$, il existe un $k$-schéma projectif lisse dont le corps des fonctions est $k$-isomorphe à $E$.

Cette hypothèse est évidemment une conséquence de la résolution des singularités au sens classique pour $k$.

(23) Si $X$ est lisse ou si l'on admet la résolution des singularités pour $k$, on peut montrer que pour tout couple $(i, n) \in \mathbb{Z} \times \mathbb{N}$, le groupe $\hat{h}_{i,n}(X)_L$ est le conoyau de la flèche

$$\bigoplus_{i=1}^n H^{\text{sing}}_i(G^{n-1}_m \times X_L/L) \to H^{\text{sing}}_i(G^n_m \times X_L/L)$$

induite par la somme des inclusions $G^{n-1}_m \to G^n_m$ de la forme $Id \times s_1 \times Id$ où $s_1$ est la section unité.

(24) C'est une conséquence de [Voe00B, chap. 5, 2.1.4] et de la dualité dans les motifs de Chow. On obtient une démonstration plus directe à l'aide du morphisme de Gysin suivant [DÉG08A, 2.16].

450  F. DÉGLISE

Documenta Mathematica 16 (2011) 411–455
Le résultat suivant est bien connu\(^{25}\):\n
**Proposition 6.3.** — Soit \(d\) un entier et \(\mathcal{P}_{\leq d}\) la sous-catégorie triangulée de \(\text{DM}(k)\) engendrée par les motifs de schémas projectifs lisses de dimension inférieure à \(d\).

Soit \(X\) un schéma lisse de dimension inférieure à \(d\).

(i) Si \(\mathcal{M}_k\) est vérifiée, \(M(X)\) appartient à \(\mathcal{P}_{\leq d}\).

(ii) Le motif rationel \(M(X) \otimes \mathbb{Q}\) appartient à \(\mathcal{P}_{\leq d} \otimes \mathbb{Q}\).

On en déduit le résultat suivant:

**Proposition 6.4.** — Soit \(X\) un schéma lisse de dimension \(d\) et \((n,i) \in \mathbb{Z}^2\) un couple d’entiers.

(i) Si \(X\) est projectif lisse, \(\hat{h}_{i-n}(X) = 0\) si \(n > d\).

(ii) Si \(\mathcal{M}_k\) est vérifiée, \(\hat{h}_{i-n}(X) = 0\) si \(n > d\).

(iii) Dans tous les cas, \(\hat{h}_{i-n}(X) \otimes \mathbb{Q} = 0\) si \(n > d\).

**Démonstration.** — Le point (i) est un corollaire de la formule (6.1.b) et du théorème de simplification de Voevodsky car ce dernier affirme qu’il n’y a pas de cohomologie motivique en poids strictement négatif.

Si \(\mathcal{C}_{\leq d}\) la sous-catégorie pleine de \(\text{DM}(k)\) formée des motifs \(\mathcal{M}\) tel que pour tout corps de fonctions \(E\) et tout couple \((n,i) \in \mathbb{Z}^2\), \(n > d\),

\[
\text{Hom}_{\text{DM}(k)}(M(E), \mathcal{M}_{\{-n\}}[-i]) = 0.
\]

Cette catégorie est une sous-catégorie triangulée. D’après (i), elle contient les motifs \(\mathcal{M}(P)\) pour \(P\) projectif lisse de dimension inférieure à \(d\). La proposition précédente permet donc de conclure.

**Remarque 6.5.** — Considérons un schéma algébrique \(X\) de dimension \(d\). Sous l’hypothèse de résolution des singularités, on peut trouver un hyper-recouvrement \(p : X \to X\) pour la topologie cdh tel que pour tout entier \(n \geq 0\), \(X_n\) est projectif lisse de dimension inférieure à \(d\). Utilisant les techniques de \([\text{Voe00a}]\), on peut montrer que le morphisme induit \(\mathbb{Z}^{tr}(X) \to \mathbb{Z}^{tr}(X)\) est un isomorphisme dans \(\text{DM}^{eff}(k)\). Le point (ii) de la proposition ci-dessus est dès lors valable sans hypothèse de lissité sur \(X\).

Notons que d’après le théorème de De Jong, on peut toujours trouver un hyper-recouvrement \(p\) comme ci-dessus pour la \(h\)-topologie. D’après \([\text{CD07}, 10.4.4, 15.1.2]\), le morphisme \(p_* : M(X) \to M(X)\) est un isomorphisme dans \(\text{DM}(k)\otimes \mathbb{Q}\). Le point (iii) est donc valable sans hypothèse de lissité.

**Définition 6.6.** — Nous dirons qu’un module homotopique (resp. module de cycles) est **constructible** s’il appartient à la sous-catégorie épaisse\(^{26}\) de \(\text{HI}^*_\ast(k)\) (resp. \(\mathcal{M}_\text{Cycl}(k)\)) engendrée par les objets \(\sigma^\infty \hat{h}_i(X)\{n\}\) (resp. \(\hat{h}_i(X)\{n\}\)) pour un schéma lisse \(X\) et un couple d’entiers \((n,i) \in \mathbb{Z}^2\).

\(^{25}\)On obtient une preuve très élégante en utilisant un argument dû à J. Riou facilement adapté de la preuve de [RI005, th. 1.4].

\(^{26}\)i.e. stable par noyau, conoyau, extension, sous-objet et quotient.
Remarque 6.7. — 1. Grâce à la t-structure homotopique, on peut considérer une autre condition de finitude sur les modules homotopiques. Un module homotopique $F_*$ est dit fortement constructible s’il est de la forme $H^0(E)$ pour un motif géométrique $E$. (27) Dans ce cas, $F_*$ est constructible dans le sens précédent mais la réciproque n’est pas claire.

2. Les modules homotopiques constructibles ne jouissent pas des propriétés de finitude de leur analogue $l$-adique. Ainsi, il y a lieu de considérer parallèlement la notion plus forte de module homotopique de type fini (28): $F_*$ est de type fini s’il existe un épimorphisme $\sigma^\infty h_0(X) \to F_*$. Ces subtilités interviennent car le foncteur $H^0$ ne préserve pas la propriété d’être géométrique (i.e. compact) – contrairement à son analogue $l$-adique, le foncteur cohomologique associé à la t-structure canonique, qui préserve la constructibilité.

3. Dans le prolongement de la remarque précédente, notons qu’il est probable que la plupart des modules homotopiques constructibles ne soient pas fortement dualisables. La seule exception que l’on connaisse à cette règle est le cas d’un $k$-schéma étale $X$ et du module homotopique $\sigma^\infty h_0(X)$. Ce dernier est fortement dualisable dans $HI(k)$ (ou même dans $HI(k)$) et il est son propre dual fort.

Corollaire 6.8. — La graduation d’un module de cycles constructible $M$ est bornée inférieurement dès que l’une des deux propriétés suivantes est réalisée:

- La propriété $(M_k)$ est satisfaite.
- $M$ est sans torsion.

6.3. HOMOLOGIE DE BOREL-MOORE. —

6.9. — Pour la proposition qui suit, on suppose l’existence pour tout schéma algébrique $X$ d’un motif à support compact $M^c(X)$ dans $DM(k)$ satisfaisant les propriétés suivantes:

(C1) $M^c(X)$ est covariant par rapport aux immersion fermées et contrariant par rapport aux immersions ouvertes.

(C2) Si $i : Z \to X$ est une immersion fermée et $j : U \to X$ l’immersion ouverte complémentaire, il existe un triangle distingué canonique:

$$M^c(Z) \xrightarrow{i_*} M^c(X) \xrightarrow{j^*} M^c(U) \xrightarrow{+1}$$

(C3) Si $X$ est lisse de dimension pure $d$, $M(X)$ est fortement dualisable avec pour dual fort $M^c(X)[-d][-d]$. De plus, la contravariance de $M^c(X)$ par rapport aux immersions ouvertes correspond par dualité à la covariance naturelle de $M(X)$.

(27) De même, un module de cycles est fortement constructible si le module homotopique associé l’est.

(28) Cette notion, introduite dans la thèse de l’auteur [Dég02], a été étudiée indépendamment par J. Ayoub dans l’appendice de [HK06].
Si $k$ vérifie la résolution des singularités, Voevodsky a obtenu cette construction dans [VOE00b, par. 4].
Soit $\phi$ un module de cycles. Rappelons la numérotation homologique du complexe des cycles à coefficients dans $\phi$ (cf. [Ros96, §5]): pour $(n,s) \in \mathbb{Z}^2$, et un schéma algébrique $X$,

$$C_n(X;\phi)_s = \bigoplus_{x \in X(n)} \phi_{s+n}(\kappa(x))$$

où $X(n)$ désigne l’ensemble des points de dimension $n$ de $X$.

**Proposition 6.10.** — Soit $\phi$ un module de cycles et $F_\ast$ le module homotopique associé (théorème 3.7).
Alors, utilisant l’hypothèse et les notations qui précèdent, pour tout schéma lissifiable $(29)$ $X$ et tout couple $(n,s) \in \mathbb{Z}^2$,

$$A_i(X,\phi)_s \simeq \text{Hom}_{DM(k)}(\mathbb{I}[i], M^c(X) \otimes F_\ast \{s\}).$$

**Démonstration.** — La catégorie $DM(k)$ est naturellement munie d’un Hom enrichi en complexes (en tant que localisation d’une catégorie dérivée). On introduit les complexes suivants, gradués par rapport à $s \in \mathbb{Z}$:

$$C_\ast(X)_s = C_\ast(X;\phi)_s, \quad D_\ast(X)_s = R \text{Hom}(\mathbb{I}, M^c(X) \otimes F_\ast \{s\}).$$

Supposons tout d’abord que $X$ est lisse de dimension pure $d$. On obtient alors un quasi-isomorphisme canonique:

$$\epsilon_X : D_\ast(X)_s = R \text{Hom}(\mathbb{I}, M^c(X) \otimes F_\ast \{s\})$$

$$\simeq^{(1)} R \text{Hom}(M(X), F_\ast \{s + d\}[d]) \simeq \Gamma(X, F_{s+d})[d]$$

$$\simeq^{(2)} C^\ast(X,\phi)_{s+d}[d] = C_\ast(X,\phi)_s.$$

L’isomorphisme $(1)$ résulte de la propriété $(C3)$ ci-dessus et l’isomorphisme $(2)$ est induit par $(3.2.a)$. Cet isomorphisme est naturel en $X$ par rapport aux immersions ouvertes.

Dans le cas général, on peut supposer que $X$ est connexe. Puisque il est lissifiable, il existe un schéma lisse $Y$, qu’on peut supposer connexe, et une immersion fermée $i : X \to Y$. Soit $j : U \to Y$ l’immersion ouverte complémentaire. Utilisant la propriété $(C2)$ et la suite $(2.5.c)$, on obtient une flèche pointillée dans $D(\mathbb{I}/b)$ formant un morphisme de triangles distingués:

$$\begin{array}{c}
D_\ast(X) \longrightarrow D_\ast(Y) \longrightarrow D_\ast(U) \longrightarrow D_\ast(U) +1 \\
\epsilon_X \downarrow \quad \epsilon_Y \downarrow \quad \epsilon_U \\
C_\ast(X) \longrightarrow C_\ast(Y) \longrightarrow C_\ast(U) \longrightarrow C_\ast(U) +1
\end{array}$$

$(29)$ i.e. il existe une immersion fermée de $X$ dans un schéma lisse.
Alors, $\epsilon_X$ est un quasi-isomorphisme compatible à la graduation. Il induit l’isomorphisme attendu.

Remarque 6.11. — Rappelons que suivant [Voe00b, par. 4], $M^\alpha(X)$ est covariant par rapport aux morphismes propres et contravariant par rapport aux morphismes plats équidimensionnels. On peut montrer que l’isomorphisme de la proposition précédente est canonique, contravariant par rapport aux morphismes plats équidimensionnels et covariant par rapport aux morphismes propres, en utilisant les techniques des sections 3.2 et 3.4.

En caractéristique 0, on pourrait aussi utiliser la méthode de descente par hyper-enveloppes de [GS96] pour obtenir la proposition précédente, remplaçant le choix d’une lissification par celui d’un hyper-recouvrement cdh – on exploite la fonctorialité covariante de $C_\ast(X)$ et $D_\ast(X)$. Ceci permet de se débarrasser de l’hypothèse: $X$ lissifiable.

RÉFÉRENCES

[BBD82] A. A. BEILINSON, J. BERNSTEIN & P. DELIGNE – “Faisceaux pervers”, Analysis and topology on singular spaces, I (Luminy, 1981), Astérisque, vol. 100, Soc. Math. France, Paris, 1982, p. 5–171.

[Bei02] A. BEILINSON – “Remarks on $n$-motives and correspondences at the generic point”, Motives, polylogarithms and Hodge theory, Part I (Irvine, CA, 1998), Int. Press Lect. Ser., vol. 3, Int. Press, Somerville, MA, 2002, p. 35–46.

[BO74] S. BLOCH & A. OGUS – “Gersten’s conjecture and the homology of schemes”, Ann. Sci. École Norm. Sup. (4) 7 (1974), p. 181–201 (1975).

[CD07] D.-C. CISINSKI & F. DÉGLISE – “Triangulated categories of motives”, 2007.

[CD09a] , “Mixed weil cohomologies”, arXiv:0712.3291v2, 2007, revised in 2009.

[CD09b] , “Local and stable homological algebra in Grothendieck abelian categories”, HHA 11 (2009), no. 1, p. 219–260.

[Dég02] F. DÉGLISE – “Modules homotopiques avec transferts et motifs génériques”, Thèse, Univ. Paris VII, 2002.

[Dég06] , “Transferts sur les groupes de chow à coefficients”, Mathematische Zeitschrift 252 (2006), no. 2, p. 315–343.

[Dég07] , “Correspondences and transfers”, Algebraic cycles and motives (J. Nagel & P. Chris, éds.), Cambridge university press, April 2007.

[Dég08a] , “Around the gysin triangle I”, arXiv:0904.2415, avril 2008.

[Dég08b] , “Motifs génériques”, Rend. Semin. Mat. Univ. Padova 119 (2008), p. 173–244.

[Dég09] , “Modules homotopiques”, arXiv:0904.4747v1, 2009.

[Dég10] , “Orientable homotopy modules”, arXiv:1005.4187, 2010.

[Gil85] H. GILLET – “Some new Gysin homomorphisms for the Chow homology of varieties”, Proc. London Math. Soc. (3) 50 (1985), no. 1, p. 57–68.

[GS96] H. GILLET & C. SOULÉ – “Descent, motives and $K$-theory”, J. Reine Angew. Math. 478 (1996), p. 127–176.
[Har66] R. Hartshorne – Residues and duality. Lecture notes of a seminar on the work of A. Grothendieck, given at Harvard 1963/64. With an appendix by P. Deligne. Lecture Notes in Mathematics, No. 20, Springer-Verlag, Berlin, 1966.

[HK06] A. Huber & B. Kahn – “The slice filtration and mixed Tate motives”, Compos. Math. 142 (2006), no. 4, p. 907–936.

[Hov01] M. Hovey – “Spectra and symmetric spectra in general model categories”, J. Pure Appl. Alg. 165 (2001), p. 63–127.

[Kah10] B. Kahn – “Relatively unramified elements in cycle modules”, http://www.math.jussieu.fr/~kahn/preprints/modcyclnr.pdf, 2010.

[Mer08] A. Merkurjev – “Unramified elements in cycle modules”, J. Lond. Math. Soc. (2) 78 (2008), no. 1, p. 51–64.

[Rio05] J. Riou – “Dualité de Spanier-Whitehead en géométrie algébrique”, C. R. Math. Acad. Sci. Paris 340 (2005), no. 6, p. 431–436.

[Ros96] M. Rost – “Chow groups with coefficients”, Doc. Math. J. (1996), p. 319–393.

[SV96] A. Suslin & V. Voevodsky – “Singular homology of abstract algebraic varieties”, Invent. Math. 123 (1996), no. 1, p. 61–94.

[SV00] ———, “Bloch-Kato conjecture and motivic cohomology with finite coefficients”, The arithmetic and geometry of algebraic cycles (Banff, AB, 1998), NATO Sci. Ser. C Math. Phys. Sci., vol. 548, Kluwer Acad. Publ., Dordrecht, 2000, p. 117–189.

[Voe00a] V. Voevodsky – “Cohomological theory of presheaves with transfers”, Cycles, transfers, and motivic homology theories, Ann. of Math. Stud., vol. 143, Princeton Univ. Press, Princeton, NJ, 2000, p. 87–137.

[Voe00b] ———, “Triangulated categories of motives over a field”, Cycles, transfers, and motivic homology theories, Ann. of Math. Stud., vol. 143, Princeton Univ. Press, Princeton, NJ, 2000, p. 188–238.

[Voe02] ———, “Cancellation theorem”, preprint, 2002.
THE ADDITIVITY THEOREM IN ALGEBRAIC $K$-THEORY

DANIEL R. GRAYSON

Received: April 20, 2011
Communicated by Max Karoubi

Abstract. The additivity theorem in algebraic $K$-theory, due to Quillen and Waldhausen, is a basic tool. In this paper we present a new proof, which proceeds by constructing an explicit homotopy combinatorially.

2010 Mathematics Subject Classification: 19D99
Keywords and Phrases: additivity, algebraic $K$-theory

Introduction

In this paper, we present a new proof of the additivity theorem of Quillen [7, §3, Theorem 2] and Waldhausen [8, 1.3.2(4)]. See also [6] and [5]. Previous proofs used Theorem A or Theorem B of Quillen [7], but this one proceeds by constructing an explicit combinatorial homotopy, which is made possible by suitably subdividing one of the spaces involved.

This research was partially supported by the National Science Foundation under grant NSF DMS 08-10948. I thank Mona Merling for valuable remarks that helped me improve the exposition substantially.

1. The additivity theorem

Let $Ord$ denote the category of finite nonempty ordered sets. We regard a simplicial object in a category $C$ as a functor $Ord^P \to C$. For $A \in Ord$ let $\Delta^A$ denote the simplicial set it represents. For each $n \in \mathbb{N}$ let $[n]$ denote the ordered set $\{0 < 1 < \cdots < n\} \in Ord$, and let $\Delta^n$ denote the simplicial set it represents. Let $\Delta^A_{top}$ denote the corresponding topological simplex, consisting of the functions $p : A \to [0, 1]$ that sum to 1; for $A = [n]$ we may also write $p = (p_0, \ldots, p_n)$.

If $X$ is a simplicial set, we let $[A, x, p]$ denote the point of the geometric realization $|X|$ corresponding to $A \in Ord$, $x \in X(A)$, and $p \in \Delta^A_{top}$.

For objects $A$ and $B$ in $Ord$, let $A \ast B \in Ord$ denote their concatenation; it is the disjoint union, with the ordering extended so the elements of $A$ are smaller.
than the elements of $B$. We make that precise by setting $A \ast B := \{(0) \times A\} \cup \{(1) \times B\}$, so $(0,a)$ and $(1,b)$ denote typical elements, and the ordering is lexicographic. We do the analogous thing with multiple concatenation, e.g., $A \ast B \ast C := \{(0) \times A\} \cup \{(1) \times B\} \cup \{(2) \times C\}$. Given functions $p : A \to \mathbb{R}$ and $q : B \to \mathbb{R}$, we let $p \ast q : A \ast B \to \mathbb{R}$ be the function defined by $(0,a) \mapsto p(a)$ and $(1,b) \mapsto q(b)$. An embedding $\Delta^A_{\text{top}} \times \Delta^B_{\text{top}} \to \Delta^{A \ast B}_{\text{top}}$ is defined by $(p,q) \mapsto (p/2) \ast (q/2)$.

The reason for using $\text{Ord}$ in this paper, instead of its full subcategory whose objects are the ordered sets $[n]$, is that it is closed under the concatenation change $(A,B) \mapsto A \ast B$ and under various other constructions used later in the paper. Since the two categories are equivalent, nothing essential is changed. Since $\text{Ord}$ is not a small category, to make the definition of geometric realization of a simplicial set work, one should either replace $\text{Ord}$ by a small subcategory containing each $[n]$ and closed under the constructions used in this paper, or one should interpret the point $[A,x,p]$ introduced above as the point $[[n],[n],\theta^* x,p\theta]$ where $\theta : [n] \to A$ is the unique isomorphism of its form.

For a simplicial object $X$, its two-fold edge-wise subdivision $\text{sub}_2 X$ (see [3, §4], [2], and [1]) is the simplicial object defined by $A \mapsto X(A \ast A)$. For a simplicial set $X$, there is a natural homeomorphism $\Psi : |\text{sub}_2 X| \cong |X|$ (defined in [3, §4]). It can be defined on each simplex as the affine map that sends each vertex of $|\text{sub}_2 X|$ to the midpoint of the corresponding (possibly degenerate) edge of $|X|$. More precisely, it sends a point $[A,x,p] \in |\text{sub}_2 X|$ to $[A \ast A, x, (p/2) \ast (p/2)] \in |X|$.

The edges of $|\text{sub}_2 X|$ that map onto the two parts of each edge of $|X|$ are oriented in the same direction. There is another edge-wise subdivision where the edges are oriented in opposite directions, defined by $A \mapsto X(A \ast A^{op})$. Subdivision into more parts can be accomplished by adding additional factors of $A$ or $A^{op}$. Our use of $\text{sub}_2 X$ in this paper, rather than one of the other available subdivisions, was based on rough sketches in low dimension of the homotopy $\Theta$ produced in Lemma 7 below.

Let $C$ be a category. Let $\text{Ar} C$ denote the category of arrows in $C$. If $f$ is an arrow of $C$, let $[f]$ denote the corresponding object of $\text{Ar} C$.

As defined in [8, 1.1 and 1.2] a category with cofibrations and weak equivalences consists of a category $\mathcal{N}$ equipped with a subcategory $\text{co} \mathcal{N}$ of cofibrations and a subcategory $\text{w} \mathcal{N}$ of weak equivalences satisfying five axioms, not repeated here. Its $K$-theory space is denoted by $\Omega \mathcal{N}$ or $\Omega \text{w} \mathcal{N}$, and is defined as the loop space $\Omega \text{w} \mathcal{N}$, where $\text{w} \mathcal{N}$ is defined in [8, (1.3)] as follows. Given $A \in \text{Ord}$, we regard it as a category in the usual way, and we let $\text{Exact}(A, \mathcal{N})$ denote the category of functors $N : A \to \mathcal{N}$ that are exact in the sense that (1) $N[a \to a] = *$ for all $a \in A$, and (2) the sequence

$$N[a \to b] \to N[a \to c] \to N[b \to c]$$

is a cofibration sequence, for all $a \leq b \leq c$ in $A$. (In the presence of condition (1), condition (2) is equivalent to...
A sequence $M \to N$ is a cofibration sequence if: (1) for all $M \in \mathcal{M}$ the sequence $F(M) \to G(M) \to H(M)$ is a cofibration sequence of $\mathcal{M}$; and (2) for any cofibration $M' \to M$ in $\mathcal{M}$ the map $G(M') \cup_{F(M')} F(M) \to G(M)$ is a cofibration in $\mathcal{N}$.

Given a cofibration sequence $F \to G \to H$ as in the definition above, the additivity theorem (Theorem 8 below) states that $F \vee H$ and $G$ yield homotopic maps $wS.M \to wS.N$. We will prove it by showing first that $G$ and $\Phi_{H,F}$ yield homotopic maps, and then composing two such homotopies. To construct this homotopy we need a new triangulation of the cylinder $[0, 1] \times [wS.M]$ that agrees with that of $[wS.M]$ at one end and with that of $[\text{sub}_2 wS.M]$ at the other end. Geometrically, it’s sort of clear that such a thing should exist, for another description of the triangulation on $[\text{sub}_2 X]$ for a simplicial set $X$, or rather of its bisimplicial variant, is that it comes by intersecting the simplices of $|\Delta^1 \times X| \cong |\Delta^1| \times |X|$ with $\{p\} \times |X|$, where $p$ denotes the midpoint of $|\Delta^1|$, the new triangulation (called $I_X$ in Definition 4 below), or rather a bisimplicial variant of it, arises by intersecting the simplices of $|\Delta^2 \times X| \cong |\Delta^2| \times |X|$ with

\[
\begin{array}{ccc}
N[a \to b] & \rightarrow & N[b \to b] \\
\downarrow & & \downarrow \\
N[a \to c] & \rightarrow & N[b \to c]
\end{array}
\]

being a pushout square.) Then $S.N$ is the simplicial category that is defined on objects by sending $A \in \text{Ord}$ to $\text{Exact}(A \mathcal{A}, N)$, and is defined on arrows in the natural way. Since $\mathcal{N}$ is equipped with a category of weak equivalences $w\mathcal{N}$, so is the exact category $\text{Exact}(A \mathcal{A}, N)$, as Waldhausen proves, yielding a simplicial category denoted $wS.N$.

Now suppose $F$ and $G$ are exact functors $\mathcal{M} \to \mathcal{N}$ between categories with cofibrations and weak equivalences. Choose a coproduct operation on $\mathcal{N}$ satisfying the identities $N \vee * = N$ and $* \vee N = N$. We define a map $\Phi = \Phi_{F,G} : \text{sub}_2 S.M \to S.N$ by $(\Phi M)[a \to b] := FM((0, a) \to (0, b)] \vee GM((1, a) \to (1, b)]$; here we have $A \in \text{Ord}$, an exact functor $M : Ar(A \ast A) \to \mathcal{M}$ regarded as an element of $(\text{sub}_2 S.M)(A)$, and an arrow $a \to b$ in $A$. One extends the definition of $\Phi M$ from objects to arrows by naturality and checks that it is exact (using the identity $(\Phi M)[a \to a] = * \vee * = *$ and exactness of the coproduct of two cofibration sequences), so $\Phi$ is well defined. The idea is that each edge of $S.M$ gets subdivided into two parts, and we apply $F$ to the first part and $G$ to the second. (The same thing works for two homomorphisms between abelian groups, with $S$. replaced by the nerve of the group.) Let $\text{sub}_2 wS.M$ denote the simplicial category obtained by applying edge-wise subdivision in the simplicial direction. The functor $\Phi$ preserves weak equivalences, because $F$, $G$, and sum do, yielding a map $\Phi : \text{sub}_2 wS.M \to wS.N$ of simplicial categories.

The following definition comes from the text above [8, Proposition 1.3.2].
Given objects $A$ and $B$ of $\text{Ord}$, define $A \times B \in \text{Ord}$ to be $A \times B$ equipped with the lexicographic ordering, where $(a,b) \leq (a',b')$ if and only if (1) $a < a'$, or (2) $a = a'$ and $b \leq b'$. (The notation is chosen to suggest that the projection $A \times B \to A$ is an order preserving map, but the projection $A \times B \to B$ is, in general, not.)

**Definition 3.** Given maps $A \xleftarrow{c} C \xrightarrow{\epsilon} B$ in $\text{Ord}$, define $\varphi^{-1}(\sigma) \in \text{Ord}$ to be the ordered subset $\{(a,b) \mid \sigma a = \varphi b\} \subseteq A \times B$. (The notation is chosen as a reminder that when $\sigma$ is injective, then projection to the second factor gives an isomorphism $\varphi^{-1}(\sigma) \cong \varphi^{-1}(\sigma(A)) \subseteq B$. On the other hand, if $\sigma$ is the map $[n] \to [0]$, then $\varphi^{-1}(\sigma) = B * \ldots * B$, the concatenation of $n + 1$ copies of $B$.)

**Definition 4.** Let $s : [2] \to [1]$ be the map in $\text{Ord}$ defined by $s(0) = 0$, $s(1) = 1$, and $s(2) = 1$. For a simplicial set $X$ we define a simplicial set $IX$ on objects by setting $IX(A) := \{ (\varphi, x) \mid \varphi : A \to [1], x \in X(\varphi^{-1}(s))\}$ for $A \in \text{Ord}$; its definition on arrows arises from naturality. We point out that $\varphi^{-1}(s) = \varphi^{-1}(\emptyset) * \varphi^{-1}(1) * \varphi^{-1}(1)$, so $\varphi^{-1}(s) \cong A$ if $\varphi = 0$, and $\varphi^{-1}(s) \cong A * A$ if $\varphi = 1$. Consequently, the simplicial subset of $IX$ defined by the equation $\varphi = 0$ is isomorphic to $X$, and the simplicial subset of $IX$ defined by the equation $\varphi = 1$ is isomorphic to $\text{sub}_2 X$. We regard those isomorphisms as identifications.

**Definition 5.** We define a map $\Psi : IX \to |\Delta^1| \times |X|$ as follows. The first component $|IX| \to |\Delta^1|$ arises from the simplicial map $IX \to \Delta^1$ defined by $(\varphi, x) \mapsto \varphi$, and thus it sends a point $[A, (\varphi, x), p]$ to the point $[A, \varphi, p]$. The second component $|IX| \to |X|$ is the unique map, affine on each simplex, whose behavior on vertices is the map $|A| \to |X|$ from the simplicial set $IA$ to the simplicial set $IX$. More precisely, the map sends a point $[A, (\varphi, x), p] \in |IX|$ to $[\varphi^{-1}(s), x, \varphi \circ p] \in |X|$, where $\varphi \circ p \in \Delta_{\text{top}}^{-1}(s)$ is defined by $(0, a) \mapsto p(a)$ for $a \in \varphi^{-1}(0)$, and by $(1, a) \mapsto p(a)/2$ and $(2, a) \mapsto p(a)/2$ for $a \in \varphi^{-1}(1)$. (Writing $p'$ for the restriction of $p$ to $\varphi^{-1}(0)$ and $p''$ for the restriction of $p$ to $\varphi^{-1}(1)$, we see that $\varphi \circ p = p' * (p''/2) * (p''/2)$.)

**Lemma 6.** For a simplicial set $X$, the map $\Psi : |IX| \to |\Delta^1| \times |X|$ is a homeomorphism.

**Proof.** By commutativity with colimits, we may assume $X = \Delta^n$. The simplicial set $IX$ has only a finite number of nondegenerate simplices, so the source and target of $\Psi$ are compact Hausdorff spaces, and thus it is enough to show that $\Psi$ is a bijection.

To show surjectivity, consider a point $([1], \beta, q, [t], x, r)$ in $|\Delta^1| \times |X|$, with $r$ in the interior of $\Delta^n_{\text{top}}$. Let $k = q(0)$. We may assume that the partial sums $s_j :=$
\[ \sum_{i=0}^{j-1} r_i, \text{ for } j = 0, \ldots, t+1, \text{ include } k, \text{ for if not, then picking } j \text{ so that } s_j < k < s_{j+1}, \text{ we may construct } r' = (r_0, \ldots, r_{j-1}, k - s_j, s_{j+1} - k, r_{j+1}, \ldots, r_t) \in \Delta_{\text{top}}^{t+1}; \]
its partial sums are those of \( r \), together with \( k \), and there is a surjective map \( f : [t+1] \to [t] \) that collapses \( r' \) to \( r \). Letting \( x' = f^*(x) = x \circ f \) be the corresponding degeneracy of \( x \), we have \( [[t], x, r] = [[t+1], x', r'] \). Similarly, we may assume that each number \( w \) with \( k \leq w \leq k + (1 - k)/2 \) is a partial sum of \( r \) if and only if \( w + (1 - k)/2 \) is.
Pick \( b \) with \( s_b = k \) and \( c \) with \( s_{b+c} = k + (1 - k)/2 \). Then, due to the symmetry of the partial sums, \( r_{b+i} = r_{b+c+i} \) if \( 0 \leq i < c \), and \( b + 2c = t + 1 \). In more detail, one deduces the equality as follows: one has \( r_{b+i} = s_{b+i+1} - s_{b+i} \), in which \( s_{b+i+1} \) and \( s_{b+i} \) are adjacent partial sums between \( k \) and \( k + (1 - k)/2 \), so by symmetry of the partial sums, \( s_{b+i+1} + (1 - k)/2 = s_{b+c+i+1} \) and \( s_{b+i} + (1 - k)/2 = s_{b+c+i} \), hence \( r_{b+c+i} = s_{b+c+i+1} - s_{b+c+i} = s_{b+i+1} - s_{b+i} = r_{b+i} \). Now let \( p \in \Delta_{\text{top}}^{b+c+1} \) be defined by \( p = (r_0, \ldots, r_{b-1}, 2r_b, \ldots, 2r_{b+c-1}) \), and let \( \varphi : [b + c - 1] \to [1] \) be defined by \( \varphi(i) = 0 \) for \( 0 \leq i < b \) and \( \varphi(i) = 1 \) for \( b \leq i < b + c \). Then \((|[1], [\beta], [[x], r]) = \Psi([b+c-1], (\varphi, x', p)) \), where \( x' \in \Xi((\varphi^{-1}(s))) \) corresponds to \( x \in X([i]) \) via the unique isomorphism \( \varphi^{-1}(s) \cong [t] \).
To show injectivity, consider a point \([A, (\varphi, x), p]) \in |JX| \) where \((\varphi, x)\) is non-degenerate and \( p \) is an interior point of \( \Delta_{\text{top}}^b \). Observe that \( x \) is a function \( \varphi^{-1}(s) \to [n] \), and that \( \varphi \circ p \) is an interior point of its simplex. The deterministic procedure described in the previous paragraph recovers \( A, \varphi, x, \) and \( p \), up to isomorphism, from the unique nondegenerate interior representatives of the two components of \( \Psi([A, (\varphi, x), p]) \), showing injectivity.

\[ \text{Lemma 7.} \] Let \( F \Rightarrow G \Rightarrow H : \mathcal{M} \to \mathcal{N} \) be a cofibration sequence of exact functors between categories with cofibrations and weak equivalences. There is a map \( \Theta : \text{IwS} \mathcal{M} \to \text{wS} \mathcal{N} \) such that \( \Theta \) agrees with \( G \) on the simplicial subset of \( \text{IwS} \mathcal{M} \) where \( \varphi = 0 \) and with \( \Phi_{H,F} \) on the simplicial subset of \( \text{IwS} \mathcal{M} \) where \( \varphi = 1 \).

\[ \text{Proof.} \] The construction will be natural in the direction of the nerve of the weak equivalences, so we don’t explicitly mention the weak equivalences in the rest of the proof. For each object \([M' \xrightarrow{f} M]\) of \( \text{Ar} \mathcal{M} \) we choose a value in \( \mathcal{N} \) for

\[ P[f] := \text{colim} \left( \begin{array}{cc} F(M') & F(M) \\ \downarrow & \downarrow \\ G(M') & \end{array} \right) \]
a cofibration \([A \rightarrow B] \rightarrow [A' \rightarrow B']\) is an arrow having the property that both
\([A \rightarrow A']\) and \([A \cup A' \rightarrow B] \rightarrow [B']\) are cofibrations; the latter part of the condition
ensures that cofibrations are stable under pushout. It follows that \(P\) sends each (horizontal) cofibration sequence

\[
\begin{array}{ccc}
L & \rightarrow & M' \\
\downarrow f & & \downarrow g \\
L' & \rightarrow & M
\end{array}
\]

of (vertical) maps (in which the rows are cofibration sequences of \(M\)) to a cofibration sequence \(P[f] \rightarrow P[g] \rightarrow P[h]\) of \(N\). The point is that, according
to definition 1, the left vertical map in the pushout diagram

\[
\begin{array}{ccc}
[FL'] & \rightarrow & [FM'] \\
\downarrow & & \downarrow \\
[GL'] & \rightarrow & [P[f] \rightarrow P[g]]
\end{array}
\]

is a cofibration in \(F_1N\), that the upper horizontal map is an arrow in \(F_1N\), and
thus that the pushout \([P[f] \rightarrow P[g]]\) lies in \(F_1N\) and is therefore a cofibration.
One also sees, using the gluing lemma \([8, 1.2: \text{Weq } 2]\), that \(P\) sends each (horizontal) weak equivalence

\[
\begin{array}{ccc}
L' & \sim & M' \\
\downarrow f & & \downarrow g \\
L & \sim & M
\end{array}
\]

of (vertical) maps (in which the horizontal maps are weak equivalences of \(M\)) to a weak equivalence \(P[f] \sim P[g]\) in \(wN\).
We say that \(P\) is an exact functor, in the sense that it preserves cofibration sequences and weak equivalences, as proved above.
We point out two special cases.

(A) if \(f = 1\) is an identity map (or an isomorphism), then there is a natural
isomorphism \(P[f] \cong G(M')\)

(B) if \(f = 0\) is a map that factors through \(*\), then there is a natural
isomorphism \(P[f] \cong F(M) \lor H(M')\)
Thus, in a precise sense, \(P\) includes \(G\) and \(F \lor H\) as special cases, allowing it
to play the lead role in the construction of \(\Theta\), which somehow deforms \(f = 1\)
to \(f = 0\) continuously. (This basic idea was also used in \([4, (10.3)\) and \((10.4)\)]
to prove a different sort of additivity theorem.)
We define \(\Theta : IwS.M \rightarrow wS.N\) as follows. Given \(A \in \text{Ord}\) and \((\varphi, M) \in (IwS.M)(A)\), we define \(\Theta(\varphi, M) \in (wS.N)(A)\) as follows. Recall from definition 4 that \(\varphi\) is a map \(A \rightarrow [1]\), that \(s\) is a certain map \(s : [2] \rightarrow [1]\), and
that \(M \in (wS.M)(\varphi^{-1}(s))\). Introduce maps \(d \leq e : [1] \rightarrow [2]\) defined by
\(d(0) = e(0) = 0, d(1) = 1, \text{ and } e(1) = 2;\) they are the sections of \(s\), and thus,
for any \( a \in A \), we have \( (d\varphi a, a) \in \varphi^{-1}(s) \) and \( (e\varphi a, a) \in \varphi^{-1}(s) \). Our task is to define an exact functor \( \Theta(\varphi, M) : Ar A \to N \), so given an object \( [a \to b] \) in \( Ar A \), we define an object of \( N \) as follows, introducing the label \( f \) for future reference.

\[
(\Theta(\varphi, M))[a \to b] := P[M[(d\varphi a, a) \to (d\varphi b, b)] \xrightarrow{f} M[(e\varphi a, a) \to (e\varphi b, b)]]
\]

We extend the definition of \( \Theta(\varphi, M) \) to arrows by naturality and by pointing out that the construction preserves weak equivalences. Exactness of \( \Theta(\varphi, M) \) follows from exactness of \( M \) and of \( P \), completing the definition of \( \Theta \).

The rest of the statement follows from the following two special cases, which result from the previous ones.

(A) if \( \varphi a = \varphi b = 0 \) then \( f = 1 \) is an identity map, and thus there is a natural isomorphism

\[
(\Theta(\varphi, M))[a \to b] \cong GM[(0, a) \to (0, b)]
\]

(B) if \( \varphi a = \varphi b = 1 \), then \( (d\varphi b, b) = (1, b) \prec (2, a) = (e\varphi a, a) \), which implies that \( f = 0 \) (because it factors through the object \( M[(1, b) \to (1, b)] = \ast \)), and thus that there is a natural isomorphism

\[
(\Theta(\varphi, M))[a \to b] \cong HM[(1, a) \to (1, b)] \vee FM[(2, a) \to (2, b)]
\]

\[\Box\]

**Theorem 8** (Additivity, [8, 1.3.2(4)]). Let \( F \Rightarrow G \Rightarrow H \) be a cofibration sequence of exact functors \( M \to N \) between categories with cofibrations and weak equivalences. Then \( F \vee H \) and \( G \) induce homotopic maps \( KM \to KN \).

**Proof.** Combining lemma 7 and lemma 6 we see that \( G \) and \( \Phi_{H,F} \) induce homotopic maps \( |wS,M| \to |wS,N| \). There is a cofibration sequence \( F \Rightarrow F \vee H \Rightarrow H \), so \( F \vee H \) and \( \Phi_{H,F} \) also induce homotopic maps. Composing the two homotopies (after reversing one of them) yields the result. \[\Box\]

**Remark 9.** Waldhausen’s Additivity Theorem provides four equivalent formulations of the result, so it is sufficient to prove only the fourth of them, as we do here. Quillen’s version [7, §3, Theorem 2] of the additivity theorem was stated for the \( Q \)-construction as a homotopy equivalence \( (s,q) : QE \to QM \times QM \), where \( M \) is an exact category, and \( E \) is the exact category of short exact sequences \( E = (0 \to sE \to tE \to qE \to 0) \) in \( M \). Here \( s,q : E \to M \) are the exact functors that extract \( sE \) and \( qE \) from the exact sequence \( E \). Quillen’s formulation is analogous to Waldhausen’s first formulation [8, 1.3.2(1)] and is implied by it.
References

[1] Herbert Edelsbrunner and Daniel R. Grayson, *Edgewise subdivision of a simplex*, Proceedings of the Fifteenth Annual Symposium on Computational Geometry (Miami Beach, FL, 1999) (New York), ACM, 1999, pp. 24–30 (electronic), links: cv art doi pdf.

[2] ———, *Edgewise subdivision of a simplex*, Discrete Comput. Geom. 24 (2000), no. 4, 707–719, ACM Symposium on Computational Geometry (Miami, FL, 1999), links: cv doi pdf.

[3] Daniel R. Grayson, *Exterior power operations on higher K-theory*, K-Theory 3 (1989), no. 3, 247–260, links: cv art doi pdf.

[4] ———, *Weight filtrations via commuting automorphisms*, K-Theory 9 (1995), no. 2, 139–172, links: cv art doi pre pdf.

[5] Kevin Charles Jones, Youngsoo Kim, Andrea H. Mhoon, Rekha Santhanam, Barry J. Walker, and Daniel R. Grayson, *The additivity theorem in K-theory*, K-Theory 32 (2004), no. 2, 181–191, links: cv art doi pre pdf.

[6] Randy McCarthy, *On fundamental theorems of algebraic K-theory*, Topology 32 (1993), no. 2, 325–328.

[7] Daniel Quillen, *Higher algebraic K-theory. I*, Algebraic K-theory, I: Higher K-theories (Proc. Conf., Battelle Memorial Inst., Seattle, Wash., 1972), Springer, Berlin, 1973, pp. 85–147. Lecture Notes in Math., Vol. 341.

[8] Friedhelm Waldhausen, *Algebraic K-theory of spaces*, Algebraic and geometric topology (New Brunswick, N.J., 1983), Lecture Notes in Math., vol. 1126, Springer, Berlin, 1985, pp. 318–419, links: homepage pdf djvu pdf-amended djvu-amended.

Daniel R. Grayson
2409 S. Vine St.
Urbana, IL 61801
USA
dan@math.uiuc.edu
WITT GROUPS OF COMPLEX CELLULAR VARIETIES

MARCUS ZIBROWIUS

Received: December 30, 2010
Revised: May 10, 2011
Communicated by Max Karoubi

Abstract. We show that the Grothendieck-Witt and Witt groups of smooth complex cellular varieties can be identified with their topological KO-groups. As an application, we deduce the values of the Witt groups of all irreducible hermitian symmetric spaces, including smooth complex quadrics, spinor varieties and symplectic Grassmannians.

2010 Mathematics Subject Classification: Primary 11E81; Secondary 19G99, 19L99, 32M15
Keywords and Phrases: Witt groups, KO-theory, projective homogeneous varieties, hermitian symmetric spaces

Contents

Introduction 466

1 PRELIMINARIES 467
1.1 Witt groups and hermitian K-theory 467
1.2 KO-theory 470
1.3 Comparison 474
1.4 \text{A}^1\text{-homotopy theory} 475
1.5 Representing algebraic and hermitian K-theory 476

2 THE COMPARISON MAPS 479
2.1 Twisting by line bundles 481
2.2 The comparison for cellular varieties 483

3 THE ATIYAH-HIRZEBRUCH SPECTRAL SEQUENCE 488
3.1 The Atiyah-Hirzebruch spectral sequence for cellular varieties 489
3.2 The Atiyah-Hirzebruch spectral sequence for Thom spaces 491

Documenta Mathematica 16 (2011) 465–511
INTRODUCTION

The purpose of this paper is to demonstrate that the Grothendieck-Witt and Witt groups of complex projective homogeneous varieties can be computed in a purely topological way. That is, we show in Theorem 2.5 how to identify them with the topological KO-groups of these varieties, and we illustrate this with a series of known and new examples.

Our theorem holds more generally for any smooth complex cellular variety. By this we mean a smooth complex variety $X$ with a filtration by closed subvarieties $\emptyset = Z_0 \subset Z_1 \subset Z_2 \cdots \subset Z_N = X$ such that the complement of $Z_k$ in $Z_{k+1}$ is an open “cell” isomorphic to $\mathbb{A}^{n_k}$ for some $n_k$. Let us put our result into perspective. It is well-known that for such cellular $X$ we have an isomorphism

$$K_0(X) \cong K^0(X(\mathbb{C}))$$

between the algebraic K-group of $X$ and the complex K-group of the underlying topological space $X(\mathbb{C})$. In fact, both sides are easy to compute: they decompose as direct sums of the K-groups of the cells, each of which is isomorphic to $\mathbb{Z}$. Such decompositions are characteristic of oriented cohomology theories. Witt groups, however, are strictly non-oriented, and this makes computations much harder. It is true that the Witt groups of complex varieties decompose into copies of $\mathbb{Z}/2$, the Witt group of $\mathbb{C}$, but even in the cellular case there is no general understanding of how many copies to expect.

Nonetheless, we can prove our theorem by an induction over the number of cells of $X$. The main issue is to define the map from Witt groups to the relevant KO-groups in such a way that it respects various exact sequences. The basic idea is clear: the Witt group $W^0(X)$ classifies vector bundles equipped with non-degenerate symmetric forms, and in topology symmetric complex vector bundles are in one-to-one correspondence with real vector bundles, classified by $KO^0(X)$. More precisely, we have two natural maps:

$$GW^0(X) \to KO^0(X(\mathbb{C}))$$
$$W^0(X) \to \frac{KO^0(X(\mathbb{C}))}{K^0(X(\mathbb{C}))}$$

Here, $GW^0(X)$ is the Grothendieck-Witt group of $X$, and in the second line $K^0(X)$ is mapped to $KO^0(X)$ by sending a complex vector bundle to the underlying real bundle. It is possible to extend these maps to shifted groups and

---

**4 Examples**

- **4.1 Notation**
- **4.2 Projective spaces**
- **4.3 Grassmannians**
- **4.4 Maximal symplectic Grassmannians**
- **4.5 Quadrics**
- **4.6 Spinor varieties**
- **4.7 Exceptional hermitian symmetric spaces**
groups with support in a concrete and “elementary” way, as was done in [Zib09]. The method advocated here is to rely instead on a result in $\mathbb{A}^1$-homotopy theory: the representability of hermitian K-theory by a spectrum whose complex realization is the usual topological KO-spectrum. Currently, our only reference is a draft paper of Morel [Mor06], but the result is well-known to the experts and a full published account will undoubtedly become available in due course. In the unstable homotopy category at least, the statement is immediate from Schlichting and Tripathi’s recent description of a geometric representing space for hermitian K-theory (see Section 1.5).

The structure of the paper is as follows: In the first section we assemble the basic definitions, reviewing some representability results along the way before finally stating in 1.9 the results in $\mathbb{A}^1$-homotopy theory that we ultimately take as our starting point. Our main result, Theorem 2.5, is stated and proved in the second section. Section 3 reviews mostly well-known facts about the Atiyah-Hirzebruch spectral sequence, on which the computations of examples in the final section rely.

1 Preliminaries

1.1 Witt groups and hermitian K-theory

From a modern point of view, the theory of Witt groups represents a K-theoretic approach to the study of quadratic forms. We briefly run through some of the basic definitions.

Recall that the algebraic K-group $K_0(X)$ of a scheme $X$ can be defined as the free abelian group on isomorphism classes of vector bundles over $X$ modulo the following relation: for any short exact sequence of vector bundles $0 \rightarrow E \rightarrow F \rightarrow G \rightarrow 0$ over $X$ we have $[F] = [E] + [G]$ in $K_0(X)$. In particular, as far as $K_0(X)$ is concerned, we may pretend that all exact sequences of vector bundles over $X$ split.

Now let $(E, \epsilon)$ be a symmetric vector bundle, by which we mean a vector bundle $E$ equipped with a non-degenerate symmetric bilinear form $\epsilon$. We may view $\epsilon$ as an isomorphism from $E$ to its dual bundle $E^\vee$, in which case its symmetry may be expressed by saying that $\epsilon$ and $\epsilon^\vee$ agree under the canonical identification of the double-dual $(E^\vee)^\vee$ with $E$. Two symmetric vector bundles $(E, \epsilon)$ and $(F, \phi)$ are isometric if there is an isomorphism of vector bundles $i: E \rightarrow F$ compatible with the symmetries, i.e. such that $i^\vee \phi i = \epsilon$. The orthogonal sum of two symmetric bundles has the obvious definition $(E, \epsilon) \perp (F, \phi) := (E \oplus F, \epsilon \oplus \phi)$. Any vector bundle $E$ gives rise to a symmetric bundle $H(E) := (E \oplus E^\vee, (0 1 \ 1 0))$, the hyperbolic bundle associated with $E$. These hyperbolic bundles are the simplest members of a wider class of so-called metabolic bundles: symmetric bundles $(M, \mu)$ which contain a subbundle $j: \mathcal{L} \rightarrow \mathcal{M}$ of half their own rank.
on which \( \mu \) vanishes. In other words, \((\mathcal{M}, \mu)\) is metabolic if it fits into a short exact sequence of the form

\[
0 \to \mathcal{L} \to \mathcal{M} \to \mathcal{L}^\vee \to 0
\]

The subbundle \( \mathcal{L} \) is then called a Lagrangian of \( \mathcal{M} \). If the sequence splits, \((\mathcal{M}, \mu)\) is isometric to \(H(\mathcal{L})\), at least in any characteristic other than two. This motivates the definition of the Grothendieck-Witt group.

**Definition 1.1 ([Wal03a, Sch10a]).** The Grothendieck-Witt group \(GW^0(X)\) of a scheme \(X\) is the free abelian group on isometry classes of symmetric vector bundles over \(X\) modulo the following two relations:

- \([([\mathcal{E}, \epsilon] \perp [\mathcal{G}, \gamma]) = [([\mathcal{E}, \epsilon]) + ([\mathcal{G}, \gamma])]
- \([([\mathcal{M}, \mu])] = [H(\mathcal{L})]\) for any metabolic bundle \((\mathcal{M}, \mu)\) with Lagrangian \(\mathcal{L}\)

The Witt group \(W^0(X)\) is defined similarly, except that the second relation reads \([([\mathcal{M}, \mu])] = 0\). Equivalently, we may define \(W^0(X)\) by the exact sequence

\[
K_0(X) \overset{H}{\to} GW^0(X) \to W^0(X) \to 0
\]

**Shifted Witt Groups.** The groups above can be defined more generally in the context of exact or triangulated categories with dualities. The previous definitions are then recovered by considering the category of vector bundles over \(X\) or its bounded derived category. However, the abstract point of view allows for greater flexibility. In particular, a number of useful variants of Witt groups can be introduced by passing to related categories or dualities. For example, if we take a line bundle \(\mathcal{L}\) over \(X\) and replace the usual duality \(\mathcal{E}^\vee := \mathcal{H}om(\mathcal{E}, \mathcal{O}_X)\) on vector bundles by \(\mathcal{H}om(-, \mathcal{L})\) we obtain “twisted” Witt groups \(W^0(X; \mathcal{L})\). On the bounded derived category, we can consider dualities that involve shifting complexes, leading to the definition of “shifted” Witt groups \(W^i(X)\). This approach, pioneered by Paul Balmer in [Bal00, Bal01a], elevates the theory of Witt groups into the realm of cohomology theories. We illustrate the meaning and significance of these remarks with a few of the key properties of the theory, concentrating on the case when \(X\) is a smooth scheme over a field of characteristic not equal to two. The interested but unacquainted reader may prefer to consult [Bal01b] or [Bal05].

- For any line bundle \(\mathcal{L}\) over \(X\) and any integer \(i\), we have a Witt group

\[
W^i(X; \mathcal{L})
\]

This is the \(i\)th Witt group of \(X\) “with coefficients in \(\mathcal{L}\),” or “twisted by \(\mathcal{L}\”). When \(\mathcal{L}\) is trivial it is frequently dropped from the notation.
The Witt groups are four-periodic in $i$ and “two-periodic in $\mathcal{L}$” in the sense that, for any $i$ and any line bundles $\mathcal{L}$ and $\mathcal{M}$ over $X$, we have canonical isomorphisms
\[ W^i(X; \mathcal{L}) \cong W^{i+4}(X; \mathcal{L}) \]
\[ W^i(X; \mathcal{L}) \cong W^i(X; \mathcal{L} \otimes \mathcal{M} \otimes \mathcal{M}^2) \]

More generally, for any closed subset $Z$ of $X$ we have Witt groups “with support on $Z$”, written $W^i_{\mathcal{L}}(X)$ for $Z = X$ these agree with $W^i(X; \mathcal{L})$.

We have long exact “localization sequences” relating the Witt groups of $X$ and $X - Z$, which can be arranged as 12-term exact loops by periodicity. Balmer’s approach already works on the level of Grothendieck-Witt groups, as shown in [Wal03a]. In this context, the localization sequences take the form
\[ \text{GW}^i_Z(X) \to \text{GW}^i(X) \to \text{GW}^i(X - Z) \to \text{GW}^{i+1}_Z(X) \to \text{GW}^{i+1}(X - Z) \to \text{GW}^{i+2}_Z(X) \to \cdots \]
continuing to the right with shifted Witt groups of $X$, and similarly for arbitrary twists $\mathcal{L}$ [Wal03a, Theorem 2.4]. However, if one wishes to continue the sequences to the left, one has to revert to the methods of higher algebraic K-theory.

**HERMITIAN K-THEORY.** Recall that the higher algebraic K-groups of a scheme $X$ can be defined as the homotopy groups of a topological space $K(X)$ associated with $X$. If one replaces $K(X)$ by an appropriate spectrum one can similarly define groups $K_n(X)$ in all degrees $n \in \mathbb{Z}$. On a smooth scheme $X$, however, the groups in negative degrees vanish.

An analogous construction of hermitian K-theory is developed in [Sch10b], Given a scheme $X$ and a line bundle $\mathcal{L}$ over $X$, Schlichting constructs a family of spectra $GW^i(X; \mathcal{L})$ from which hermitian K-groups can be defined as
\[ GW^i_n(X; \mathcal{L}) := \pi_n(GW^i(X; \mathcal{L})) \]
In degree $n = 0$, one recovers Balmer and Walter’s Grothendieck-Witt groups, and the Witt groups appear as hermitian K-groups in negative degrees. To be precise, for any smooth scheme $X$ over a field of characteristic not equal to two one has the following natural identifications:
\[ GW^i_0(X; \mathcal{L}) \cong GW^i(X; \mathcal{L}) \]
\[ GW^i_n(X; \mathcal{L}) \cong W^{i-n}(X; \mathcal{L}) \text{ for } n < 0 \]
For affine varieties, the identifications of Witt groups may be found in [Hor05]: see Proposition A.4 and Corollary A.5. For a general smooth scheme $X$, we can pass to a vector bundle torsor $T$ over $X$ such that $T$ is affine [Jou73, Lemma 1.5;
Hor05, Lemma 2.1]. Both Balmer’s Witt groups and Schlichting’s hermitian K-groups are homotopy invariant in the sense that the groups of $T$ may naturally be identified with those of $X$. This is proved for Witt groups in [Gil03, Corollary 4.2] and may be deduced for hermitian K-theory from the Mayer-Vietoris sequences established in [Sch10b, Theorem 1]. The identifications also hold more generally for hermitian K-groups with support $GW^i_{n,Z}(X)$ [Sch]. They will be used implicitly throughout.

For completeness, we mention how the 4-periodic notation used here translates into the traditional notation in terms of KO- and U-theory, as used for example in [Hor05]. Namely, we have

\[
GW^i_n(X) = \begin{cases} 
KO_n(X) & \text{for } i \equiv 0 \mod 4 \\
U_n(X) & \text{for } i \equiv -1 \\
-KO_n(X) & \text{for } i \equiv -2 \\
-U_n(X) & \text{for } i \equiv -3
\end{cases}
\]

(This notation will not be used elsewhere in this paper.)

1.2 KO-theory

We now turn to the corresponding theories in topology. To ensure that the definitions given here are consistent with the literature, we restrict our attention to finite-dimensional CW complexes. Since we are ultimately only interested in topological spaces that arise as complex varieties, this is not a problem.

The definitions of $K_0$ and $GW^0$ given above applied to complex vector bundles over a finite-dimensional CW complex $X$ yield its complex and real topological K-groups $K^0(X)$ and $KO^0(X)$. Since short exact sequences of vector bundles over $X$ always split, the definitions may even be simplified:

**Definition 1.2.** For a finite-dimensional CW complex $X$, the complex K-group $K^0(X)$ is the free abelian group on isomorphism classes of complex vector bundles over $X$ modulo the relation $[E \oplus G] = [E] + [G]$. Likewise, the KO-group $KO^0(X)$ is the free abelian group on isometry classes of symmetric complex vector bundles over $X$ modulo the relation $[(E, \epsilon) \perp (G, \gamma)] = [(E, \epsilon)] + [(G, \gamma)]$.

There is a more common description of $KO^0(X)$ as the K-group of real vector bundles. The equivalence with the definition given here can be traced back to the fact that the orthogonal group $O(n)$ is a maximal compact subgroup of both $GL_n(\mathbb{R})$ and $O_n(\mathbb{C})$, but also seen very concretely along the following lines. We say that a complex bilinear form $\epsilon$ on a real vector bundle $F$ is real if $\epsilon: F \otimes F \to \mathbb{C}$ factors through $\mathbb{R}$.

---

1This step is known as Jouanolou’s trick.
2The key property we need is that any vector bundle over a finite-dimensional CW complex has a stable inverse. See the proof of Theorem 1.5.
Lemma 1.3. Let \((E, \varepsilon)\) be a symmetric complex vector bundle. There exists a unique real subbundle \(\mathbb{R}(E, \varepsilon) \subset E\) such that \(\mathbb{R}(E, \varepsilon) \otimes \mathbb{C} = E\) and such that the restriction of \(\varepsilon\) to \(\mathbb{R}(E, \varepsilon)\) is real and positive definite. Concretely, a fibre of \(\mathbb{R}(E, \varepsilon)\) is given by the real span of any orthonormal basis of the corresponding fibre of \(E\).

Corollary 1.4. For any CW complex \(X\), the monoid of isomorphism classes of real vector bundles over \(X\) is isomorphic to the monoid of isometry classes of symmetric complex vector bundles over \(X\) (with respect to the operations \(\oplus\) and \(\perp\), respectively).

Proof of Lemma 1.3. In the case of a vector bundle over a point we may assume without loss of generality that 
\[
(E, \varepsilon) = \left(\mathbb{C}^r, \begin{pmatrix} 1 \\ 0 & \cdots & 0 \\ 0 & 1 \end{pmatrix}\right)
\]
Clearly, the subspace \(\mathbb{R}^r \subset \mathbb{C}^r\) has the required properties. Uniqueness follows from elementary linear algebra. If \((E, \varepsilon)\) is an arbitrary symmetric complex vector bundle over a space \(X\), then any point of \(X\) has some neighbourhood over which \((E, \varepsilon)\) can be trivialized in the form above. We know how to define \(\mathbb{R}(E, \varepsilon)\) over each such neighbourhood, and by uniqueness these local bundles can be glued together.

Proof of Corollary 1.4. A map in one direction is given by sending a symmetric complex vector bundle \((E, \varepsilon)\) to \(\mathbb{R}(E, \varepsilon)\). Conversely, with any real vector bundle \(E\) over \(X\) we may associate a symmetric complex vector bundle \((E \otimes \mathbb{C}, \sigma)\), where \(\sigma\) is the \(\mathbb{C}\)-linear extension of some inner product \(\sigma\) on \(E\). Since \(\sigma\) is defined uniquely up to isometry, so is \((E \otimes \mathbb{C}, \sigma)\). See [MH73, Chapter V, § 2] for a proof that avoids the uniqueness part of the preceding lemma.

Representing topological K-groups. A standard construction of the cohomology theories associated with \(K^0\) and \(KO^0\) is based on the fact that these functors are representable in the homotopy category \(\mathcal{H}\) of topological spaces. The starting point is the homotopy classification of vector bundles: Let us write \(Gr_{r,n}\) for the Grassmanian \(Gr(r, \mathbb{C}^{r+n})\) of complex \(r\)-bundles in \(\mathbb{C}^{r+n}\), and let \(Gr_r\) be the union of \(Gr_{r,n} \subset Gr_{r,n+1} \subset \cdots\) under the obvious inclusions. Denote by \(U_{r,n}\) and \(U_r\) the universal \(r\)-bundles over these spaces. For any connected paracompact Hausdorff space \(X\) we have a one-to-one correspondence between the set \(\text{Vec}_r(X)\) of isomorphism classes of rank \(r\) complex vector bundles over \(X\) and homotopy classes of maps from \(X\) to \(Gr_r\): a homotopy class \([f]\) in \(\mathcal{H}(X, Gr_r)\) corresponds to the pullback of \(U_r\) along \(f\) [Hus94, Chapter 3, Theorem 7.2].

To describe \(K^0(X)\), we need to pass to \(Gr_r\), the union of the \(Gr_r\) under the embeddings \(Gr_r \hookrightarrow Gr_{r+1}\) that send a complex \(r\)-plane \(W\) to \(\mathbb{C} \oplus W\).
THEOREM 1.5. For finite-dimensional CW complexes $X$ we have natural isomorphisms
\[ \text{K}^0(X) \cong \mathcal{H}(X, \mathbb{Z} \times \text{Gr}) \] (4)
such that, for $X = \text{Gr}_{r,n}$, the class $[\mathcal{U}_{r,n}] + (d-r)[\mathcal{C}]$ in $\text{K}^0(\text{Gr}_{r,n})$ corresponds to the inclusion $\text{Gr}_{r,n} \hookrightarrow \{d\} \times \text{Gr}_{r,n} \hookrightarrow \mathbb{Z} \times \text{Gr}$.

Proof. The theorem is of course well-known, see for example [Ada95, page 204]. To deduce it from the homotopy classification of vector bundles, we note first that any CW complex is paracompact and Hausdorff [Hat09, Proposition 1.20]. Moreover, we may assume that $X$ is connected. The product $\mathbb{Z} \times \text{Gr}$ can be viewed as the colimit of the inductive system
\[ \coprod_{d \geq 0} \{d\} \times \text{Gr}_d \hookrightarrow \coprod_{d \geq -1} \{d\} \times \text{Gr}_{d+1} \hookrightarrow \coprod_{d \geq -2} \{d\} \times \text{Gr}_{d+2} \hookrightarrow \cdots \subset \mathbb{Z} \times \text{Gr} \]

Any continuous map from $X$ to $\mathbb{Z} \times \text{Gr}$ factors through one of the components $\text{colim}_n \{d\} \times \text{Gr}_n$. By cellular approximation, it is in fact homotopic to a map that factors through $\{d\} \times \text{Gr}_n$ for some $n$. Thus,
\[ \mathcal{H}(X, \mathbb{Z} \times \text{Gr}) \cong \coprod_{d \in \mathbb{Z}} \text{colim}_n \text{Vect}_n(X) \]
where the colimit is taken over the maps $\text{Vect}_n(X) \to \text{Vect}_{n+1}(X)$ sending a vector bundle $\mathcal{E}$ to $\mathbb{C} \oplus \mathcal{E}$. We define a map from the coproduct to $\text{K}^0(X)$ by sending a vector bundle $\mathcal{E}$ in the $d$th component to the class $[\mathcal{E}] + (d - \text{rank} \mathcal{E})[\mathcal{C}]$ in $\text{K}^0(X)$. To see that this is a bijection, we use the fact that every vector bundle $\mathcal{E}$ over a finite-dimensional CW complex has a stable inverse: a vector bundle $\mathcal{E}^\perp$ over $X$ such that $\mathcal{E} \oplus \mathcal{E}^\perp$ is a trivial bundle [Hus94, Chapter 3, Proposition 5.8].

If we replace the complex Grassmannians by real Grassmannians $\mathbb{R} \text{Gr}_{r,n}$, we obtain the analogous statement that $\text{KO}^0$ can be represented by $\mathbb{Z} \times \mathbb{R} \text{Gr}$. Equivalently, but more in the spirit of Definition 1.2, we could work with the following spaces:

DEFINITION 1.6. Let $(V, \nu)$ be a symmetric complex vector space, and let $\text{Gr}(r, V)$ be the Grassmannian of complex $k$-planes in $V$. The “non-degenerate Grassmannian”
\[ \text{Gr}^{\text{nd}}(r, (V, \nu)) \]
is the open subspace of $\text{Gr}(r, V)$ given by $r$-planes $T$ for which the restriction $\nu|_T$ is non-degenerate.

Complexification induces an inclusion of $\mathbb{R} \text{Gr}(k, \mathbb{R}(V, \nu))$ into $\text{Gr}^{\text{nd}}(r, (V, \nu))$, which, by Lemma 1.7 below, is a homotopy equivalence. So let $\text{Gr}^{\text{nd}}_{r,n}$ abbreviate $\text{Gr}^{\text{nd}}(r, \mathbb{H}^{r+n})$, where $\mathbb{H}$ is the hyperbolic plane $(\mathbb{C}^2, (\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}))$, and let $\mathcal{U}^{\text{nd}}_{r,n}$ denote the restriction of the universal bundle over $\text{Gr}(r, \mathbb{C}^{2r+2n})$ to $\text{Gr}^{\text{nd}}_{r,n}$. Then
colimits $\text{Gr}^r_r$ and $\text{Gr}^{nd}$ can be defined in the same way as for the usual Grassmannians, and, for finite-dimensional CW complexes $X$, we obtain natural isomorphisms

$$KO^0(X) \cong \mathcal{H}(X, \mathbb{Z} \times \text{Gr}^{nd})$$

(5)

Here, for even $(d-r)$, the inclusion $\text{Gr}^{nd}_{r,n} \hookrightarrow \{d\} \times \text{Gr}^{nd}_{r,n} \hookrightarrow \mathbb{Z} \times \text{Gr}^{nd}$ corresponds to the class of $[\mathbb{H}^{nd}_{r,n}] + \frac{d-r}{2}$ in $GW^0(\text{Gr}^{nd}_{r,n})$.

**Lemma 1.7.** For any symmetric complex vector space $(V, \nu)$, the following inclusion is a homotopy equivalence:

$$\text{RGr}(k, \mathbb{R}(V, \nu)) \xrightarrow{j} \text{Gr}^{nd}(k, (V, \nu))$$

$$U \mapsto U \otimes_{\mathbb{R}} \mathbb{C}$$

**Proof.** Consider the projection $\pi: V = \mathbb{R}(V, \nu) \oplus i\mathbb{R}(V, \nu) \twoheadrightarrow \mathbb{R}(V, \nu)$. We define a retract $r$ of $j$ by sending a complex $k$-plane $T \in \text{Gr}^{nd}(k, (V, \nu))$ to $\pi(\mathbb{R}(T, \nu|_T)) \subset \mathbb{R}(V, \nu)$. This is indeed a linear subspace of real dimension $k$: since $\nu$ is positive definite on $\mathbb{R}(T, \nu|_T)$ but negative definite on $i\mathbb{R}(V, \nu)$, the intersection $\mathbb{R}(T, \nu|_T) \cap i\mathbb{R}(V, \nu)$ is trivial.

More generally, we can define a family of endomorphisms of $V$ parametrized by $t \in [0, 1]$ by

$$\pi_t: \mathbb{R}(V, \nu) \oplus i\mathbb{R}(V, \nu) \twoheadrightarrow \mathbb{R}(V, \nu) \oplus i\mathbb{R}(V, \nu)$$

$$(x, y) \mapsto (x, ty)$$

This family interpolates between the identity $\pi_1$ and the projection $\pi_0$, which we can identify with $\pi$. We claim that

$$\pi_t(\mathbb{R}(T, \nu|_T)) \subset V$$

is a real linear subspace of dimension $k$ on which $\nu$ is real and positive definite.

The claim concerning the dimension has already been verified in the case $t = 0$ and follows for non-zero $t$ from the fact that $\pi_t$ is an isomorphism. Now take a non-zero vector $v \in \pi_t(\mathbb{R}(T, \nu|_T))$ and write it as $v = x + ty$, where $x, y \in \mathbb{R}(V, \nu)$ and $x + iy \in \mathbb{R}(T, \nu|_T)$. Since $\nu(x, x), \nu(y, y)$ and $\nu(x + iy, x + iy)$ are all real we deduce that $\nu(x, y) = 0$; it follows that $\nu(v, v)$ is real as well. Moreover, since $\nu(x + iy, x + iy)$ is positive we have $\nu(x, x) > \nu(y, y)$, so that $\nu(v, v) > (1 - t^2)\nu(y, y)$. In particular, $\nu(v, v) > 0$ for all $t \in [0, 1]$, as claimed. It follows that $T \mapsto \pi_t(\mathbb{R}(T, \nu|_T)) \otimes_{\mathbb{R}} \mathbb{C}$ defines a homotopy from $j \circ r$ to the identity on $\text{Gr}^{nd}(k, (V, \nu))$. \qed

**K-spectra and cohomology theories.** The infinite Grassmannian $\text{Gr}$ can be identified with the classifying space $BU$ of the infinite unitary group. Consequently, $K^0$ can be represented by $\mathbb{Z} \times BU$, which by Bott periodicity is equivalent to its own two-fold loopspace $\Omega^2(\mathbb{Z} \times BU)$. This can be used to
construct a 2-periodic \( \Omega \)-spectrum \( K_{\text{top}} \) in the stable homotopy category \( \mathcal{SH} \) whose even terms are all given by \( \mathbb{Z} \times \text{BU}. \) Similarly, \( R \text{Gr} \) is equivalent to the classifying space \( BO \) of the infinite orthogonal group, and Bott periodicity in this case says that \( \mathbb{Z} \times BO \) is equivalent to \( \Omega^8(\mathbb{Z} \times BO) \). Thus, one obtains a spectrum \( KO_{\text{top}} \) in \( \mathcal{SH} \) which is 8-periodic. The associated cohomology theories are given by

\[
\begin{align*}
K^i(X) &:= SH(\Sigma\infty(X_+), S^i \wedge K_{\text{top}}) \\
KO^i(X) &:= SH(\Sigma\infty(X_+), S^i \wedge KO_{\text{top}})
\end{align*}
\]

where \( X_+ \) denotes the union of \( X \) and a disjoint base point, and \( \Sigma\infty \) is the functor assigning to a pointed space its suspension spectrum. We refer the reader to [Ada95, III.2] for background and details.

For convenience and later reference, we include here the values of the theories on a point. Since we are in fact dealing with multiplicative theories, these can be summarized in the form of coefficient rings:

\[
\begin{align*}
K^*(\text{point}) &= \mathbb{Z}[g, g^{-1}] \\
KO^*(\text{point}) &= \mathbb{Z}[\eta, \alpha, \lambda, \lambda^{-1}]/(2\eta, \eta^3, \eta\alpha, \alpha^2 - 4\lambda)
\end{align*}
\]

where \( g \) is of degree \(-2\) and \( \eta, \alpha \) and \( \lambda \) have degrees \(-1, -4 \) and \(-8\), respectively [Bot69, pages 66–74]\\]

1.3 Comparison

Now suppose \( X \) is a smooth complex variety. We write \( X(\mathbb{C}) \) for the set of complex points of \( X \) equipped with the analytic topology. If \( E \) is a vector bundle over \( X \) then \( E(\mathbb{C}) \) has the structure of a complex vector bundle over \( X(\mathbb{C}) \), so that we obtain natural maps

\[
\begin{align*}
K_0(X) &\to K^0(X(\mathbb{C})) \\
GW^0(X) &\to KO^0(X(\mathbb{C}))
\end{align*}
\]

and an induced map

\[
W^0(X) \to \frac{KO^0(X(\mathbb{C}))}{K^0(X(\mathbb{C}))}
\]

We now wish to extend these maps to be defined on \( GW^i(X) \) and \( W^i(X) \) for arbitrary \( i \), and also on groups with support and twisted groups. Let us comment on some “elementary” constructions that are possible before outlining the approach that we will ultimately follow here.

---

3The multiplicative relations among the generators are given on page 74, but unfortunately the relation \( \eta\alpha = 0 \) is missing. This omission seems to have pervaded much of the literature, and I am indebted to Ian Grojnowski for pointing out the same mistake in an earlier version of this paper. Of course, the relation follows from the fact that \( KO^{-5}(\text{point}) = 0 \).
Firstly, one way to extend the maps to the groups $GW^i(X)$ and $W^i(X)$ is to use the multiplicative structure of the theories together with Walter’s results on projective bundles [Wal03b]. Namely, for any variety $X$ one has isomorphisms

$$GW^i(X \times \mathbb{P}^1) \cong GW^i(X) \oplus GW^{i-1}(X)$$
$$KO^{2i}(X(\mathbb{C}) \times S^2) \cong KO^{2i}(X(\mathbb{C})) \oplus KO^{2i-2}(X(\mathbb{C}))$$

This allows an inductive definition of comparison maps, at least for all negative $i$. Basic properties of these maps, for example compatibility with the periodicities of Grothendieck-Witt and KO-groups, can be checked by direct calculations.

It is less clear how to obtain maps on Witt groups with restricted supports. One possibility, pursued in [Zib09], is to work on the level of complexes of vector bundles and adapt a construction of classes in relative K-groups described in [Seg68] to the case of KO-theory. However, it remains unclear to the author how to see in this approach that the resulting maps are compatible with the boundary morphisms in localization sequences.

Theorem 2.5 below could in fact be proved without knowing that the comparison maps respect the boundary morphisms in localization sequences in general. However, $A^1$-homotopy theory provides an alternative construction of a comparison map for which this property immediately follows from the construction, and which in any case is so compellingly elegant that it would be difficult to argue in favour of any other approach.

### 1.4 $A^1$-HOMOTOPY THEORY

Theorem 1.5 describing $K^0$ in terms of homotopy classes of maps to Grassmanians has an analogue in algebraic geometry, in the context of $A^1$-homotopy theory. Developed mainly by Morel and Voevodsky, the theory provides a general framework for a homotopy theory of schemes emulating the situation for topological spaces. The authoritative reference is [MV99]; closely related texts by the same authors are [Voe98], [Mor99] and [Mor04]. See [DLØ07] for a textbook introduction and [Dug01] for an enlightening perspective on one of the main ideas.

We summarize the main points relevant for us in just a few sentences. The category $\text{Sm}_k$ of smooth schemes over a field $k$ can be embedded into some larger category of “spaces” $\text{Spc}_k$ which is closed under small limits and colimits, and which can be equipped with a model structure. The $A^1$-homotopy category $\mathcal{H}(k)$ over $k$ is the homotopy category associated with this model category.

In fact, there are several possible choices for $\text{Spc}_k$ and many possible model structures yielding the same homotopy category $\mathcal{H}(k)$. One possibility is to consider the category of simplicial presheaves over $\text{Sm}_k$, or the category of simplicial sheaves with respect to the Nisnevich topology. Both categories contain $\text{Sm}_k$ as full subcategories via the Yoneda embedding, and they also contain simplicial sets viewed as constant (pre)sheaves. One may thus apply
a general recipe for equipping the category of simplicial (pre)sheaves over a site with a model structure (see [Jar87]). In a crucial last step, one forces the affine line \( \mathbb{A}^1 \) to become contractible by localizing with respect to the set of all projections \( \mathbb{A}^1 \times X \rightarrow X \).

As in topology, we also have a pointed version \( \mathcal{H}_\bullet(k) \) of \( \mathcal{H}(k) \). Remarkably, these categories contain several distinct “circles”: the simplicial circle \( S^1 \), the “Tate circle” \( G_m = \mathbb{A}^1 - 0 \) (pointed at 1) and the projective line \( \mathbb{P}^1 \) (pointed at \( \infty \)). These are related by the intriguing formula \( \mathbb{P}^1 = S^1 \wedge G_m \).

A common notational convention which we will follow is to define

\[
S^{p,q} := (S^1)^{p-q} \wedge G_m^q
\]

for any \( p \geq q \). In particular, we then have \( S^1 = S^{1,0} \), \( G_m = S^{1,1} \) and \( \mathbb{P}^1 = S^{2,1} \).

One can take the theory one step further by passing to the stable homotopy category \( \mathcal{SH}(k) \), a triangulated category in which the suspension functors \( S^{p,q} \wedge - \) become invertible. This category is usually constructed using \( \mathbb{P}^1 \)-spectra. The triangulated shift functor is given by suspension with the simplicial sphere \( S^{1,0} \).

Finally and crucially, the analogy with topology can be made precise: when we take our ground field \( k \) to be the complex numbers, or more general any subfield of \( \mathbb{C} \), we have a complex realization functor

\[
\mathcal{H}(k) \to \mathcal{H}
\]

that sends a smooth scheme \( X \) to its set of complex points \( X(\mathbb{C}) \) equipped with the analytic topology. There is also a pointed realization functor and, moreover, a triangulated functor of the stable homotopy categories

\[
\mathcal{SH}(k) \to \mathcal{SH}
\]

which takes \( \Sigma^\infty(X_+) \) to \( \Sigma^\infty(X(\mathbb{C})_+) \) for any smooth scheme \( X \) [Rio06, Théorème I.123; Rio07a, Théorème 5.26].

### 1.5 Representing algebraic and hermitian K-theory

Grassmannians of \( r \)-planes in \( k^{n+r} \) can be constructed as smooth projective varieties over any field \( k \). Viewing them as objects in \( \mathbf{Spc}_k \), we can form their colimits \( \text{Gr}_r \) and \( \text{Gr} \) in the same way as in topology. The following analogue of Theorem 1.5 is established in [MV99, § 4]; see Théorème III.3 and Assertion III.4 in [Rio06].

**Theorem 1.8.** For smooth schemes \( X \) over \( k \) we have natural isomorphisms

\[
K_0(X) \cong \mathcal{H}(k)(X, \mathbb{Z} \times \text{Gr})
\]

such that the inclusion of \( \text{Gr}_{r,n} \hookrightarrow \{d\} \times \text{Gr}_{r,n} \hookrightarrow \mathbb{Z} \times \text{Gr} \) corresponds to the class \( [\mathcal{U}_{r,n}] + (d-r)[\mathcal{O}] \) in \( K_0(\text{Gr}_{r,n}) \).

*Documenta Mathematica 16 (2011) 465–511*
An analogous result for hermitian K-theory has recently been obtained by Schlichting and Tripathi:\footnote{Talk “Geometric representation of hermitian K-theory in $\mathbb{A}^1$-homotopy theory” at the Workshop “Geometric Aspects of Motivic Homotopy Theory”, 6.–10. September 2010 at the Hausdorff Center for Mathematics, Bonn} Let $\Gr_{r,n}^{nd}$ denote the “non-degenerate Grassmanians” defined as open subvarieties of $\Gr_{r,r+2n}$ as above, and let $\Gr_{r}^{nd}$ and $\Gr^{nd}$ be the respective colimits. Then for smooth schemes over $k$ we have natural isomorphisms

$$GW^0(X) \cong \mathcal{H}(k)(X, \mathbb{Z} \times \Gr^{nd}) \quad (14)$$

It follows from the construction that, when $(d - r)$ is even, the inclusion of $\Gr_{r,n}^{nd} \hookrightarrow \{d\} \times \Gr_{r,n}^{nd} \hookrightarrow \mathbb{Z} \times \Gr^{nd}$ corresponds to the class of $[U_{r,n}] + \frac{d-r}{2}[\mathbb{H}]$ in $GW^0(\Gr_{r,n}^{nd})$, where $U_{r,n}^{nd}$ is the universal symmetric bundle over $\Gr_{r,n}^{nd}$.

The fact that hermitian K-theory is representable in $\mathcal{H}(k)$ has been known for longer, see [Hor05]. One of the advantages of having a geometric description of a representing space, however, is that one can easily see what its complex realization is. In particular, this gives us an alternative way to define the comparison maps. For any smooth complex scheme $X$ we have the following commutative squares, in which the left vertical arrows are the comparison maps (8) and (9), the right vertical arrows are induced by the complex realization functor (11).

$$\begin{align*}
K_0(X) & \cong \mathcal{H}(\mathbb{C})(X, \mathbb{Z} \times \Gr) \\
K^0(\mathbb{C})(X) & \cong \mathcal{H}(\mathbb{C})(X, \mathbb{Z} \times \Gr) \\
KO_0(X) & \cong \mathcal{H}(\mathbb{C})(X, \mathbb{Z} \times \Gr)
\end{align*}$$

Some of the results quoted here are in fact known in a much greater generality. Firstly, higher algebraic and hermitian K-groups of $X$ are obtained by passing to suspensions of $X$ in (13) and (14). Even better, algebraic and hermitian K-theory are representable in the stable $\mathbb{A}^1$-homotopy category $\mathcal{S}\mathcal{H}(k)$. Let us make the statement a little more precise by fixing some notation. Given a spectrum $E$ in $\mathcal{S}\mathcal{H}(k)$, we obtain a bigraded reduced cohomology theory $\tilde{E}^{p,q}$ on $\mathcal{H}(k)$ and a corresponding unreduced theory $E^{p,q}$ on $\mathcal{H}(k)$ by setting

$$\begin{align*}
\tilde{E}^{p,q}(X) & := \mathcal{S}\mathcal{H}(k)(\Sigma^n X, S^{p,q} \wedge \mathbb{E}) & \text{for } X \in \mathcal{H}(k) \\
E^{p,q}(X) & := \tilde{E}^{p,q}(X_p) & \text{for } X \in \mathcal{H}(k)
\end{align*}$$

A spectrum $\mathcal{K}$ representing algebraic K-theory was first constructed in [Voe98, § 6.2]; see [Ri06] or [Ri07b] for some further discussion. It is $(2,1)$-periodic, meaning that in $\mathcal{S}\mathcal{H}(k)$ we have an isomorphism

$$S^{2,1} \wedge \mathcal{K} \cong \mathcal{K}$$
Thus, the bigrading of the corresponding cohomology theory $K^{p,q}$ is slightly artificial. The identification with the usual notation for algebraic K-theory is given by

$$K^{p,q}(X) = K_{2q-p}(X)$$

For hermitian K-theory we have an $(8,4)$-periodic spectrum $KO$, and the corresponding cohomology groups $KO^{p,q}$ are honestly bigraded. The translation into the notation used for hermitian K-groups in Section 1.1 is given by

$$KO^{p,q}(X) = GW_{2q-p}^{q}(X)$$

We will refer to the number $2q - p$ as the degree of the group $KO^{p,q}(X)$. The relation with Balmer’s Witt groups obtained by combining (16) and (2) is illustrated by the following table:

| $KO^{p,q}$ | $p = 0$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|------------|---------|---|---|---|---|---|---|---|
| $q = 0$    | $GW^0$  | $W^1$ | $W^2$ | $W^3$ | $W^0$ | $W^1$ | $W^2$ | $W^3$ |
| $q = 1$    | $GW^1_2$ | $GW^1_4$ | $GW^1_6$ | $GW^1_8$ | $W^2$ | $W^3$ | $W^0$ | $W^1$ |
| $q = 2$    | $GW^2_4$ | $GW^2_8$ | $GW^2_8$ | $GW^2_8$ | $GW^2_8$ | $W^3$ | $W^0$ | $W^1$ |
| $q = 3$    | $GW^3_8$ | $GW^3_8$ | $GW^3_8$ | $GW^3_8$ | $GW^3_8$ | $GW^3_8$ | $W^0$ | $W^1$ |

As for the representing spaces in the unstable homotopy category, it is known that the complex realizations of $KO$ and $K$ represent real and complex topological K-theory. This is well-documented in the latter case, see for example [Rio06, Proposition VI.12]. For $KO$, our references are slightly thin. Since the emphasis in this article is on showing how such a result in $A^1$-homotopy theory can be used for some concrete computations, we will at this point succumb to an “axiomatic” approach — the key statements we will be using are as follows:

**Standing assumptions 1.9.** There exist spectra $K$ and $KO$ in $SH(\mathbb{C})$ representing algebraic K-theory and hermitian K-theory in the sense described above, such that:

(a) The complex realization functor (12) takes $K$ to $K^{top}$ and $KO$ to $KO^{top}$.

(b) We have an exact triangle in $SH(\mathbb{C})$ of the form

$$KO \wedge S^{1,1} \to KO \to K \to S^{1,0} \wedge \ldots$$

which corresponds to the usual triangle in $SH$.

These results are announced in [Mor06]. Independent constructions of spectra representing hermitian K-theory can be found in [Hor05] and in a recent preprint of Panin and Walter [PW10].
2 The comparison maps

It follows immediately from 1.9 that complex realization induces comparison maps

\[ \tilde{k}^{p,q} : \tilde{\mathcal{K}}^{p,q}(X) \to \tilde{\mathcal{K}}^{p}(X(\mathbb{C})) \]

\[ \tilde{k}^{p,q}_h : \tilde{\mathcal{K}}^{p,q}(X) \to \widetilde{\mathcal{K}}^{p}(X(\mathbb{C})) \]

and hence comparison maps \( k^{p,q} \) and \( k^{p,q}_h \) on \( K \)- and hermitian \( K \)-groups. In particular, in degrees 0 and \(-1\) we have maps

\[ k^{0,0}_h : K_0(X) \to K_0(X(\mathbb{C})) \]

\[ gw^q := k^{2q,q}_h : GW^q(X) \to KO^{2q}(X(\mathbb{C})) \]

\[ w^q := k^{2q-1,q-1}_h : W^q(X) \to KO^{2q-1}(X(\mathbb{C})) \]

for any smooth complex scheme \( X \). Some good properties of these maps follow directly from the construction:

- They commute with pullbacks along morphisms of smooth schemes.
- They are compatible with suspension isomorphisms.
- They are compatible with the periodicity isomorphisms, so we can identify \( k^{p,q}_h \) with \( k^{p+8,q+4}_h \) (and hence \( w^q \) with \( w^{q+4} \) and \( gw^q \) with \( gw^{q+4} \)).

It is also clear that they are compatible with long exact sequences arising from exact triangles in \( SH(\mathbb{C}) \). This will be particularly useful in the following two cases.

Localization sequences. Given a smooth closed subscheme \( Z \) of a smooth scheme \( X \), we have an exact triangle

\[ \Sigma^\infty(X - Z)_+ \to \Sigma^\infty X_+ \to \Sigma^\infty \left( \frac{X}{X - Z} \right) \to S^{1,0} \wedge \ldots \]

in \( SH(\mathbb{C}) \). It induces long exact “localization sequences” for cohomology theories. For example, for hermitian \( K \)-theory we obtain sequences of the form

\[ \cdots \to \tilde{\mathcal{K}}O^{p,q} \left( \frac{X}{X - Z} \right) \to KO^{p,q}(X) \to KO^{p,q}(X - Z) \to \tilde{\mathcal{K}}O^{p+1,q} \left( \frac{X}{X - Z} \right) \to \ldots \]  

(18)

The comparison maps commute with all maps appearing in this sequence and the corresponding sequence of topological \( KO \)-groups.

The space \( X/(X - Z) \) depends only on the normal bundle \( N \) of \( Z \) in \( X \). To make this precise, we introduce the Thom space of a vector bundle \( \mathcal{E} \) over an arbitrary smooth scheme \( Z \), defined as the homotopy quotient of \( \mathcal{E} \) by the complement of the zero section:

\[ \text{Thom}_Z(\mathcal{E}) := \mathcal{E}/(\mathcal{E} - Z) \]
Using a geometric construction known as deformation to the normal bundle, Morel and Voevodsky show in Theorem 2.23 of [MV99, Chapter 3] that \(X/(X - Z)\) is canonically isomorphic to \(\text{Thom}_Z(N)\) in the unstable pointed \(\mathbb{A}^1\)-homotopy category. Thus, sequence (18) can be rewritten in the following form:

\[
\ldots \to \overline{KO}^{p,q}(\text{Thom}_Z N) \to KO^{p,q}(X) \to KO^{p,q}(X - Z) \to \overline{KO}^{p+1,q}(\text{Thom}_Z N) \to KO^{p+1,q}(X) \to KO^{p+1,q}(X - Z) \to \ldots
\]

Karoubi/Bott sequences. The KO- and K-groups of a topological space \(X\) fit into a long exact sequence known as the Bott sequence [Bot69, pages 75 and 112; BG10, 4.I.B]. It has the form

\[
\ldots \to KO^{2i-1}X \to KO^{2i-2}X \to K^0X \to KO^{2i}X \to KO^{2i-1}X \to K^1X \\
\to KO^{2i+1}X \to KO^{2i}X \to K^0X \to KO^{2i+2}X \to KO^{2i+1}X \to \ldots
\]

(19)

The maps from KO- to K-groups are essentially given by complexification (or, depending on our choice of description of KO-groups, by forgetting the symmetric structure of a complex symmetric bundle), and the maps from K- to KO-groups are given by sending a complex vector bundle to its underlying real bundle (or to the associated hyperbolic bundle). The maps between KO-groups are given by multiplication with the generator \(\eta\) of \(KO^{-1}(\text{point})\) (see (7)).

This long exact sequence is induced by the triangle described in 1.9. The sequence arising from the corresponding triangle (17) in the stable \(\mathbb{A}^1\)-homotopy category is known as the Karoubi sequence. The comparison maps induce a commutative ladder diagram that allows us to compare the two. Near degree zero, this takes the following form:

\[
\ldots \to KO^{2i-1}X \to GW^{i-1}X \to K_0X \to GW^{i}X \to W^iX \to 0 \to \ldots \\
\ldots \to KO^{2i-1}X \to KO^{2i-2}X \to K^0X \to KO^{2i}X \to KO^{2i-1}X \to K^1X \to \ldots
\]

(20)

As a consequence, the comparison maps \(w^i\) factor as

\[
W^i(X) \to \frac{KO^{2i}(X)}{K^0(X)} \to KO^{2i-1}(X)
\]

For cellular varieties, or more generally for spaces for which the odd topological K-groups vanish, the second map in this factorization is an isomorphism.

\[\text{Unfortunately, there are misprints on both pages. In particular, the central group in the diagram on page 112 should be } K^0.\]
Groups with restricted support. Comparing the localization sequences (1) and (18), we see that the groups $\widetilde{KO}^p,q(X)$ play the role of hermitian K-groups of $X$ supported on $Z$. This should be viewed as part of any representability statement, see for example [PW10, Theorem 6.5]. Alternatively, a formal identification of the groups in degrees zero and below using only the minimal assumptions we have stated could be achieved as follows:

**Lemma 2.1.** Let $Z$ be a smooth closed subvariety of a smooth quasi-projective variety $X$. We have the following isomorphisms:

$$\widetilde{KO}^{2q,q}(X) \cong GW^q_Z(X)$$
$$\widetilde{KO}^{p,q}(X) \cong W^{p-q}_Z(X) \text{ for } 2q - p < 0$$

**Proof.** Consider $Z = Z \times \{0\}$ as a subvariety of $X \times \mathbb{A}^1$. Its open complement $(X \times \mathbb{A}^1) - Z$ contains $X = X \times \{1\}$ as a retract. Thus, the projection from $X \times \mathbb{A}^1$ onto $X$ induces a splitting of the localization sequences associated with $(X \times \mathbb{A}^1 - Z) \hookrightarrow X \times \mathbb{A}^1$, and we have

$$GW^{i+1}_Z(X \times \mathbb{A}^1) \cong \text{coker} \left( GW^{i+1}_1(X \times \mathbb{A}^1) \hookrightarrow GW^{i+1}_1(X \times \mathbb{A}^1 - Z) \right)$$
$$\widetilde{KO}^{n+2,i+1}(X \times \mathbb{A}^1) \cong \text{coker} \left( KO^{n+1,i+1}(X \times \mathbb{A}^1) \hookrightarrow KO^{n+1,i+1}(X \times \mathbb{A}^1 - Z) \right)$$

By (16), we can identify the groups appearing on the right, so we obtain an induced isomorphism of the cokernels. The quotient $X \times \mathbb{A}^1/(X \times \mathbb{A}^1 - Z)$ can be identified with the suspension of $X/(X - Z)$ by $S^{2,1}$, so we have an isomorphism

$$\widetilde{KO}^{n+2,i+1}(X \times \mathbb{A}^1) \cong \widetilde{KO}^{2i+1}(X)$$

On the other hand, we have analogous isomorphisms

$$GW^{i+1}_Z(X \times \mathbb{A}^1) \cong GW^i_Z(X)$$

for Grothendieck-Witt and Witt groups. For Witt groups, this is a special case of Theorem 2.5 in [Nen07], the case when $Z = X$ being Theorem 8.2 in [BG05]. The corresponding isomorphisms of Grothendieck-Witt groups may be deduced via Karoubi induction. The proof in lower degrees is analogous.

### 2.1 Twisting by line bundles

As described in Section 1.1, there is a natural notion of Witt groups twisted by line bundles. In the homotopy theoretic approach, such a twist can be encoded by passing to the Thom space of the bundle.

**Definition 2.2.** For a vector bundle $\mathcal{E}$ of constant rank $r$ over a smooth scheme $X$, we define the hermitian K-groups of $X$ with coefficients in $\mathcal{E}$ by

$$KO^{p,q}(X; \mathcal{E}) := \widetilde{KO}^{p+2r,q+r}(\text{Thom} \mathcal{E})$$
Likewise, for any complex vector bundle of rank $r$ over a topological space $X$, we define

$$\text{KO}^p(X; \mathcal{E}) := \overline{\text{KO}}^{p+2r}(\text{Thom } \mathcal{E})$$

When $\mathcal{E}$ is a trivial bundle, its Thom space is just a suspension of $X$, so that $\text{KO}^{p,q}(X; \mathcal{E})$ and $\text{KO}^{p,q}(X)$ agree.

**Lemma 2.3.** For any vector bundle $\mathcal{E}$ over a smooth quasi-projective variety $X$, we have isomorphisms

$$\text{KO}^{2q,q}(X; \mathcal{E}) \cong \text{GW}^q(X; \det \mathcal{E})$$

$$\text{KO}^{p,q}(X; \mathcal{E}) \cong \text{W}^{p-q}(X; \det \mathcal{E}) \text{ for } 2q - p < 0$$

**Proof.** This follows from Lemma 2.1 and Nenashev’s Thom isomorphisms for Witt groups: for any vector bundle $\mathcal{E}$ of rank $r$ there is a canonical Thom class in $\text{W}^r_X(\mathcal{E})$ which induces an isomorphism $\text{W}^i(X; \det \mathcal{E}) \cong \text{W}^{i+r}_X(\mathcal{E})$ by multiplication [Nen07, Theorem 2.5]. This Thom class actually comes from a class in $\text{GW}^r_X(\mathcal{E})$, and, as in the proof of Lemma 2.1, we can deduce that it induces an analogous isomorphism on Grothendieck-Witt groups via Karoubi induction.

**Remark.** The isomorphisms of Lemmas 2.1 and 2.3 are constructed here in a rather ad hoc fashion, and we have taken little care in recording their precise form. Whenever we give an argument concerning the comparison maps on “twisted groups” in the following, we do all constructions on the level of representable groups of Thom spaces. The identifications with the usual twisted groups are only needed to identify the final output of concrete calculations as in Section 4.

It follows similarly from Thom isomorphisms in topology that the groups $\text{KO}(X; \mathcal{E})$ only depend on the determinant line bundle of $\mathcal{E}$:

**Lemma 2.4.** For complex vector bundles $\mathcal{E}$ and $\mathcal{F}$ on a topological space $X$ with identical first Chern class modulo 2, we have

$$\text{KO}^p(X; \mathcal{E}) \cong \text{KO}^p(X; \mathcal{F})$$

**Proof.** A complex vector bundle $\mathcal{E}$ whose first Chern class vanishes modulo 2 has a spin structure and is therefore oriented with respect to KO-theory [ABS64, § 12]. That is, we have a Thom isomorphism

$$\text{KO}^p X \xrightarrow{\cong} \overline{\text{KO}}^{p+2r}(\text{Thom } \mathcal{E})$$

Now suppose $c_1(\mathcal{E}) \equiv c_1(\mathcal{F}) \mod 2$. We may view $\mathcal{E} \oplus \mathcal{E} \oplus \mathcal{F}$ both as a vector bundle over $\mathcal{E}$ and as a vector bundle over $\mathcal{F}$, and by assumption it is oriented with respect to KO-theory in both cases. Thus, both groups in the lemma can be identified with $\text{KO}^p(X; \mathcal{E} \oplus \mathcal{E} \oplus \mathcal{F})$. \qed
Remark 2.1. In general the identifications of Lemma 2.4 are non-canonical. Given a spin structure on a real vector bundle, the constructions in [ABS64] do yield a canonical Thom class, but there may be several different spin structures on the same bundle. Still, canonical identifications exist in many cases. For example, there is a canonical spin structure on the square of any complex line bundle, yielding canonical identifications
\[ KO^p(X; \mathcal{L}) \cong KO^p(X; \mathcal{L} \otimes \mathcal{M}^{\otimes 2}) \]
for any two complex line bundles \( \mathcal{L} \) and \( \mathcal{M} \) over \( X \). Moreover, as different spin structures on a spin bundle over \( X \) are classified by the singular cohomology group \( H^1(X; \mathbb{Z}/2) \), all spin structures arising in the context of complex cellular varieties below will be unique.

2.2 The comparison for cellular varieties

Theorem 2.5. For a smooth cellular complex variety \( X \), the following comparison maps are isomorphisms:
\[ K_0(X) \xrightarrow{\cong} K^0(X(\mathbb{C})) \]
\[ gw^q : GW^q(X) \xrightarrow{\cong} KO^{2q}(X(\mathbb{C})) \]
\[ w^q : W^q(X) \xrightarrow{\cong} KO^{2q-1}(X(\mathbb{C})) \]

This remains true for twisted groups (see Section 2.1).

As indicated in the introduction, the first isomorphism is well-known and almost self-evident, given that both \( K_0(X) \) and \( K^0(X(\mathbb{C})) \) are free abelian of rank equal to the number of cells of \( X \). In particular, both the algebraic group \( K_0(\mathbb{C}) \) and the topological K-group \( K^0(\text{point}) \) are isomorphic to the integers, generated by the trivial line bundle, and the comparison map is the obvious isomorphism.

Let us begin the proof of the theorem by also considering the other two maps first in the case when \( X \) is just a point \( \text{Spec}(\mathbb{C}) \). We can easily see that the corresponding groups are isomorphic by direct comparison:

| \( KO^{p,q}(\mathbb{C}) \) | \( p \) | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|---|
| \( q = 0 \) | \( \mathbb{Z} \) | 0 | 0 | 0 | \( \mathbb{Z}/2 \) | 0 | 0 | 0 | \( \mathbb{Z}/2 \) |
| \( q = 1 \) | ... | ... | 0 | 0 | 0 | \( \mathbb{Z}/2 \) | 0 | 0 | \( \mathbb{Z}/2 \) |
| \( q = 2 \) | ... | ... | ... | ... | \( \mathbb{Z} \) | 0 | \( \mathbb{Z}/2 \) | 0 | \( \mathbb{Z}/2 \) |
| \( q = 3 \) | ... | ... | ... | ... | ... | ... | \( \mathbb{Z}/2 \) | \( \mathbb{Z}/2 \) | \( \mathbb{Z}/2 \) |

Table 1: (Grothendieck-)Witt and KO-groups of a point.
To see that the isomorphisms are given by our comparison maps, we can use the comparison of the Karoubi and Bott sequences. First, setting $i = 0$ in Diagram (20), we see that $gw$ and $w$ are isomorphisms on a point. As $W(\mathbb{C})$ is the only non-trivial Witt group of a point, it follows that $w$ is an isomorphism on a point in general, so that we have

$$
\begin{array}{cccccccc}
\cdots & \rightarrow & GW^{q-1} & \rightarrow & K_0 & \rightarrow & GW^q & \rightarrow & W^q & \rightarrow & 0 & \rightarrow & \cdots \\
& \downarrow{gw^{q-1}} & \cong & & \downarrow{gw^q} & \cong & & & & & & & \\
\cdots & \rightarrow & KO^{2q-2} & \rightarrow & K^0 & \rightarrow & KO^{2q} & \rightarrow & KO^{2q-1} & \rightarrow & 0 & \rightarrow & \cdots 
\end{array}
$$

Given the periodicity of the Grothendieck-Witt groups, repeated applications of the Five Lemma now show that $gw$ is an isomorphism on a point for all values of $q$. (This strategy of proof is known as “Karoubi induction”.)

We now treat the hermitian case in general. The case of algebraic/complex $K$-theory could be dealt with similarly, or deduced from the hermitian case using triangle (17). It will be helpful to consider not only the maps $gw = k^{2q,q}$ and $w^{q+1} = k^{2q+1,q}$ in degrees 0 and $-1$, respectively, but also the maps $k^{2q-1,q}$ in degree 1 and the maps $k^{2q+2,q}$ in degree $-2$. We will prove the following extended statement:

**Theorem 2.6.** For a smooth cellular variety $X$, the hermitian comparison maps in degrees 1, 0, $-1$ and $-2$ have the properties indicated:

$$
\begin{align*}
KO^{2q-1,q}(X) & \rightarrow KO^{2q-1}(X(\mathbb{C})) \\
KO^{2q,q}(X) & \cong KO^{2q}(X(\mathbb{C})) \\
KO^{2q+1,q}(X) & \cong KO^{2q+1}(X(\mathbb{C})) \\
KO^{2q+2,q}(X) & \rightarrow KO^{2q+2}(X(\mathbb{C}))
\end{align*}
$$

The analogous statements for twisted groups are also true.

**Remark 2.2.** The map in degree 1 is not an isomorphism even when $X$ is a point. For example, it is known that $KO^{-1,0}(\mathbb{C}) = \mathbb{Z}/2$ (see [Kar05, Example 18]), from which we may deduce via the Karoubi sequence that $KO^{1,1}(\mathbb{C}) \cong \mathbb{C}^*$. In particular, $KO^{1,1}(\mathbb{C})$ cannot be isomorphic to $KO^1(\text{point}) = 0$.

The map in degree $-2$ can be identified with the inclusion of the 2-torsion subgroup of $KO^{2q+2}(X(\mathbb{C}))$ into $KO^{2q+2}(X(\mathbb{C}))$ for any cellular variety $X$. This follows from the theorem and the description of the KO-groups of cellular varieties given in Lemma 3.1.

In degrees less than $-2$, the comparison map is necessarily zero. The problem is that while $\eta: W^{p-q}(X) \rightarrow W^{p-q}(X)$ is an isomorphism in all negative degrees, the topological $\eta$ is nilpotent ($\eta^3 = 0$).

The proof of Theorem 2.6 will proceed by induction over the number of cells of $X$ and occupy the remainder of this section. To begin the induction, we need
to consider the case of only one cell, which immediately reduces to the case of a point by homotopy invariance. In this case, degrees 0 and −1 have already been dealt with above. In degrees 1 and −2, on the other hand, most of the statements are trivial, and we only need to look at a few particular cases, which we postpone to the end of the proof.

Spheres. Assuming the theorem to be true for a point, the compatibility of the comparison maps with suspensions immediately shows that the theorem is also true for the reduced cohomology of the spheres \( (\mathbb{P}^1)^{\times d} = S^{2d,d} \). To be precise, the following maps in degrees 1, 0, −1 and −2 have the properties indicated:

\[
\begin{align*}
\tilde{KO}^{2q-1,q}(S^{2d,d}) & \rightarrow \tilde{KO}^{2q-1}(S^{2d}) \\
\tilde{KO}^{2q,q}(S^{2d,d}) & \cong \tilde{KO}^{2q}(S^{2d}) \\
\tilde{KO}^{2q+1,q}(S^{2d,d}) & \cong \tilde{KO}^{2q+1}(S^{2d}) \\
\tilde{KO}^{2q+2,q}(S^{2d,d}) & \rightarrow \tilde{KO}^{2q+2}(S^{2d})
\end{align*}
\]

Cellular varieties. Now let \( X \) be a smooth cellular variety. By definition, \( X \) has a filtration by closed subvarieties \( \emptyset = Z_0 \subset Z_1 \subset Z_2 \cdots \subset Z_N = X \) such that the open complement of \( Z_k \) in \( Z_{k+1} \) is isomorphic to \( \mathbb{A}^{n_k} \) for some \( n_k \). In general, the subvarieties \( Z_k \) will not be smooth. Their complements \( U_k := X - Z_k \) in \( X \), however, are always smooth as they are open in \( X \). So we obtain an alternative filtration \( X = U_0 \supset U_1 \supset U_2 \cdots \supset U_N = \emptyset \) of \( X \) by smooth open subvarieties \( U_k \). Each \( U_k \) contains a closed cell \( C_k \cong \mathbb{A}^{n_k} \) with open complement \( U_{k+1} \).

Our inductive hypothesis is that we have already proved the theorem for \( U_{k+1} \), and we now want to prove it for \( U_k \). We can use the following exact triangle in \( \mathcal{SH}(\mathbb{C}) \):

\[
\Sigma^\infty(U_{k+1})_+ \rightarrow \Sigma^\infty(U_k)_+ \rightarrow \Sigma^\infty\text{Thom}(N_{C_k \setminus U_k}) \rightarrow S^{1,0} \wedge \ldots
\]

As \( C_k \) is a cell, the Quillen-Suslin theorem tells us that the normal bundle \( N_{C_k \setminus U_k} \) of \( C_k \) in \( U_k \) has to be trivial. Thus, \( \text{Thom}(N_{C_k \setminus U_k}) \) is \( \mathbb{A}^1 \)-weakly equivalent to \( S^{2d,d} \), where \( d \) is the codimension of \( C_k \) in \( U_k \). Figure 1 displays the comparison between the long exact cohomology sequences induced by this triangle. The inductive step is completed by applying the Five Lemma to each dotted map in the diagram.

The twisted case. To obtain the theorem in the case of coefficients in a vector bundle \( \mathcal{E} \) over \( X \), we replace the exact triangle above by the triangle

\[
\Sigma^\infty\text{Thom}(\mathcal{E}|_{U_{k+1}}) \rightarrow \Sigma^\infty\text{Thom}(\mathcal{E}|_{U_k}) \rightarrow \Sigma^\infty\text{Thom}(\mathcal{E}|_{C_k} \oplus N_{C_k \setminus U_k}) \rightarrow S^{1,0} \wedge \ldots
\]

The existence of this exact triangle is shown in the next lemma. The Thom space on the right is again just a sphere, so we can proceed as in the untwisted case.
Figure 1: The inductive step.
Lemma 2.7. Given a smooth subvariety $Z$ of a smooth variety $X$ with complement $U$, and given any vector bundle $E$ over $X$, we have an exact triangle

$$\Sigma^\infty \Theta \text{hom}(E_U) \to \Sigma^\infty \Theta \text{hom} E \to \Sigma^\infty \Theta \text{hom}(E|_Z \oplus N_{Z\setminus X}) \to S^{1,0} \wedge \ldots$$

Proof. From the Thom isomorphism theorem we know that the Thom space of a vector bundle over a smooth base is $h^1$-weakly equivalent to the quotient of the vector bundle by the complement of the zero section. Consider the closed embeddings $U \hookrightarrow (E - Z), X \hookrightarrow E$ and $Z \hookrightarrow E$. Computing the normal bundles, we obtain

$$\frac{(E - Z)}{(E - X)} \cong \Theta \text{hom}_U(E_U)$$

$$\frac{E}{(E - X)} \cong \Theta \text{hom}_X E$$

$$\frac{E}{(E - Z)} \cong \Theta \text{hom}_Z(E|_Z \oplus N_{Z\setminus X})$$

The claim follows by passing to the stable homotopy category and applying the octahedral axiom to the composition of the embeddings $(E - X) \subseteq (E - Z) \subseteq E$.

Remaining details concerning a point. To finish the proof of Theorem 2.6, we now return to the maps of degrees 1 and $-2$ in the case of a point, which we skipped above. First, let us deal with degree 1. The odd KO-groups of a point are all trivial except for $KO^{-1}$, so $k_{-1}^{2q-1,0}$ is trivially a surjection unless $q \equiv 0 \mod 4$. In that case, surjectivity of $k_{-1}^{1,0}$ is clear from the following diagram:

$$\ldots \to KO^{-1,0} \to GW^{-1} \to K_0 \to \ldots$$

$$\ldots \to KO^{-1} \to KO^{-2} \to K^0 \to \ldots$$

Lastly, we consider what happens in degree $-2$. Again, three out of four cases are trivial as $KO^{2q+2,q} = W^{q+2}$ is zero unless $q \equiv 2 \mod 4$. For the non-trivial case, consider the map $\eta$ appearing in triangle (17). As the negative algebraic K-groups of $\mathbb{C}$ are zero, $\eta$ yields automorphisms of $W^{p-q}$ in negative degrees. In topology, the corresponding maps are given by multiplication with a generator.
\[ \eta \text{ of } KO^{-1}, \text{ and } \eta^2 \text{ generates } KO^{-2}. \] So the commutative square

\[
\begin{array}{ccc}
W^0 \xrightarrow{\cong} W^0 \\
\downarrow \cong \\
KO^{-1} \xrightarrow{\eta \cong} KO^{-2}
\end{array}
\]

shows that \( k_{h}^{0,-2} \) is an injection (in fact, an isomorphism), as claimed. This completes the proof of Theorem 2.6.

**Remark 2.3.** We indicate briefly how Theorem 2.5 can alternatively be obtained by working only with the maps in degrees 0 and \(-1\) that can be defined by more elementary means. The basic strategy — comparing the localization sequences arising from the inclusion of a closed cell \( C_k \) into the union of “higher” cells \( U_k \) — still works. But we cannot deduce that the comparison maps are isomorphisms on \( U_k \) from the fact that they are isomorphisms on \( U_{k+1} \) because the parts of the sequences that we can actually compare are now too short.

We can, however, still deduce that the maps in degree 0 with domains the Grothendieck-Witt groups of \( U_k \) are surjective, and that the maps in degree \(-1\) with domains the Witt groups of \( U_k \) are injective. The inductive step can then be completed with the help of the Bott/Karoubi sequences. This argument works even without assuming that the comparison maps are compatible with the boundary maps in localization sequences in general: in the relevant cases the cohomology groups involved are so simple that this property can be checked by hand.

### 3 The Atiyah-Hirzebruch spectral sequence

We now aim to prepare the ground for the discussion of the KO-theory of some examples in the next section. The main computational tool will be the Atiyah-Hirzebruch spectral sequence, which in topology exists for any generalized cohomology theory and any finite-dimensional CW complex \( X \) [Ada95, III.7; Koc96, Theorem 4.2.7]. For KO-theory, it has the form

\[ E_2^{p,q} = HP(X; KO^q(point)) \Rightarrow KO^{p+q}(X) \]

with differential \( d_r \) of bidegree \((r, -r + 1)\). The \( E_2 \)-page is thus concentrated in the half-plane \( p \geq 0 \) and 8-periodic in \( q \): we have the integral cohomology of \( X \) in rows \( q \equiv 0 \) and \( q \equiv -4 \) mod 8, its cohomology with \( \mathbb{Z}/2 \)-coefficients in rows \( q \equiv -1 \) and \( q \equiv -2 \), and all other rows are zero. The differential \( d_2 \) is given by \( Sq^2 \circ \pi_2 \) and \( Sq^2 \) on rows \( q \equiv 0 \) and \( q \equiv -1 \), respectively, where

\[ Sq^2 : H^*(X; \mathbb{Z}/2) \rightarrow H^{*+2}(X; \mathbb{Z}/2) \]

is the second Steenrod square and \( \pi_2 \) is mod-2 reduction [Fuj67, 1.3].
The spectral sequence is multiplicative [Koc96, Proposition 4.2.9]. That is, the multiplication on the $E_2$-page induced by the cup product on singular cohomology and the ring structure of $\text{KO}^*(\text{point})$ (see (7)) descends to a multiplication on all subsequent pages, such that the multiplication on the $E_\infty$-page is compatible with the multiplication on $\text{KO}^*(X)$. In particular, each page is a module over $\text{KO}^*(\text{point})$. The differentials of the spectral sequence are derivations, i.e. they satisfy a Leibniz rule.

3.1 The Atiyah-Hirzebruch spectral sequence for cellular varieties

For cellular varieties, or more generally for CW complexes with only even-dimensional cells, the spectral sequence becomes simple enough to make some general deductions. We summarize some lemmas of Hoggar and Kono and Hara.

**Lemma 3.1.** [Hog69, 2.1 and 2.2] Let $X$ be a CW complex with only even-dimensional cells. Then:

- The ranks of the free parts of $\text{KO}^0X$ and $\text{KO}^4X$ are equal to the number $t_0$ of cells of $X$ of dimension a multiple of 4.
- The ranks of the free parts of $\text{KO}^2X$ and $\text{KO}^6X$ are equal to the number $t_1$ of cells of $X$ of dimension 2 modulo 4.
- The groups of odd degrees are two-torsion, i.e. $\text{KO}^{2i-1}X = (\mathbb{Z}/2)^{s_i}$ for some $s_i$.
- $\text{KO}^{2i}X$ is isomorphic to the direct sum of its free part and $\text{KO}^{2i+1}X$.

Table 3 in Section 4.1 summarizes these statements.

**Proof.** The cohomology of $X$ is free on generators given by the cells and concentrated in even degrees. The first two statements thus follow easily from the Atiyah-Hirzebruch spectral sequence for KO-theory (e.g. after tensoring with $\mathbb{Q}$). On the other hand, we see from the Atiyah-Hirzebruch spectral sequence for complex K-theory that $K^0(X)$ is a free abelian group on the cells while $K^1(X)$ is zero. The last two statements thus become consequences of the Bott sequence (19). □

The free part of $\text{KO}^*$ is thus very simple. In good cases, the spectral sequence also provides a nice description of the 2-torsion. To see this, note that $\text{Sq}^2 \text{Sq}^2 = \text{Sq}^4 \text{Sq}^1$ must vanish when the cohomology of $X$ with $\mathbb{Z}/2$-coefficients is concentrated in even degrees. So we can view $(H^*(X;\mathbb{Z}/2), \text{Sq}^2)$ as a differential graded algebra over $\mathbb{Z}/2$. To lighten the notation, we will write

$$H^*(X, \text{Sq}^2) := H^*(H^*(X;\mathbb{Z}/2), \text{Sq}^2)$$
for the cohomology of this algebra. We keep the same grading as before, so that it is concentrated in even degrees. The row \( q \equiv -1 \) on the \( E_3 \)-page is given by \( H^*(X, \text{Sq}^2) \cdot \eta \), where \( \eta \) is the generator of \( \text{KO}^{-1}(\text{point}) \). Since it is the only row that contributes to \( \text{KO}^* \) in odd degrees, we arrive at the following lemma, which will be central to our computations.

**Lemma 3.2.** Let \( X \) be as above. If the Atiyah-Hirzebruch spectral sequence of \( \text{KO}^*(X) \) degenerates on the \( E_3 \)-page, then

\[
\text{KO}^{2i-1}(X) \cong \bigoplus_k H^{2i+8k}(X, \text{Sq}^2)
\]

In all the examples we consider below, the spectral sequence does indeed degenerate at this stage. However, showing that it does can be tricky. One step in the right direction is the following observation of Kono and Hara [KH91, Proposition 1].

**Lemma 3.3.** Let \( X \) be as above. If the differentials \( d_3, d_4, \ldots, d_r \) are trivial and \( d_r \) is non-trivial, then \( r \equiv 2 \mod 8 \). In other words, the first non-trivial differential after \( d_2 \) can only appear on a page \( E_r \) with page number \( r \equiv 2 \mod 8 \).

Such a differential is non-zero only on rows \( q \equiv 0 \) and \( q \equiv -1 \mod 8 \). If it is non-zero on some \( x \) in row \( q \equiv 0 \), then it is also non-zero on \( \eta x \) in row \( q \equiv -1 \). Conversely, if it is non-zero on some \( y \) in row \( q \equiv -1 \), there exists some \( x \) in row \( q \equiv 0 \) such that \( y = x\eta \) and \( d_r \) is non-zero on \( x \).

**Proof.** We see from the spectral sequence of a point that \( d_r \eta = 0 \) for all differentials. Thus, multiplication by \( \eta \) gives a map of bidegree \((0, -1)\) on the spectral sequence that commutes with the differentials. On the \( E_2 \)-page this map is mod-2 reduction from row \( q \equiv 0 \) to row \( q \equiv -1 \) and the identity between rows \( q \equiv -1 \) and \( q \equiv -2 \). It follows that on the \( E_3 \)-page multiplication by \( \eta \) induces a surjection from row \( q \equiv 0 \) to row \( q \equiv -1 \) and an injection of row \( q \equiv -1 \) into row \( q \equiv -2 \). This implies all statements above.

We derive a corollary that we will use to deduce that the spectral sequence collapses for certain Thom spaces:

**Corollary 3.4.** Suppose we have a continuous map \( p: X \rightarrow T \) of CW complexes with only even-dimensional cells. Suppose further that the Atiyah-Hirzebruch spectral sequence for \( \text{KO}^*(X) \) collapses on the \( E_3 \)-page, and that \( p^* \) induces an injection in row \( q \equiv -1 \):

\[
p^*: H^*(T, \text{Sq}^2) \hookrightarrow H^*(X, \text{Sq}^2)
\]

Then the spectral sequence for \( \text{KO}^*(T) \) also collapses at this stage.

**Proof.** Write \( d_r \) for the first non-trivial higher differential, so \( r \equiv 2 \mod 8 \). Then, for any element \( x \) in row \( q \equiv 0 \), we have \( p^*(d_r x) = d_r p^*(x) = 0 \) since the spectral sequence for \( X \) collapses. From our assumption on \( p^* \) we can deduce that \( d_r x = 0 \). By the preceding lemma, this is all we need to show.
3.2 The Atiyah-Hirzebruch spectral sequence for Thom spaces

In order to compute twisted KO-groups, we need to apply the Atiyah-Hirzebruch spectral sequence of KO-theory to Thom spaces. So let $X$ be a finite-dimensional CW complex, and let $\pi: E \to X$ be a vector bundle of constant rank over $X$. Though we will be mainly interested in the case when $E$ is complex, we may more generally assume here that $E$ is any real vector bundle which is oriented. Then the Thom isomorphism for singular cohomology tells us that the reduced cohomology of the Thom space $\text{Thom} \ E$ is additively isomorphic to the cohomology of $X$ itself, apart from a shift in degrees by $r := \text{rank}_R E$. The isomorphism is given by multiplication with a Thom class $\theta$ in $\widetilde{H}_r(\text{Thom} \ E; \mathbb{Z})$:

$$H^*(X; \mathbb{Z}) \xrightarrow{\cong} \widetilde{H}^{*+r}(\text{Thom} \ E; \mathbb{Z})$$

$$x \mapsto \pi^*(x) \cdot \theta$$

Similarly, the reduction of $\theta$ modulo two induces an isomorphism of the respective singular cohomology groups with $\mathbb{Z}/2$-coefficients. Thus, apart from a shift of columns, the entries on the $E_2$-page of the spectral sequence for $\widetilde{KO}^*(\text{Thom} \ E)$ are identical to those on the $E_2$-page for $KO^*(X)$. However, the differentials may differ.

**Lemma 3.5.** Let $E \xrightarrow{\pi} X$ be a complex vector bundle of constant rank over a topological space $X$, with Thom class $\theta$ as above. The second Steenrod square on $H^*(\text{Thom} \ E; \mathbb{Z}/2)$ is given by “$\text{Sq}^2 + c_1(E)$”, where $c_1(E)$ is the first Chern class of $E$ modulo two. That is,

$$\text{Sq}^2(\pi^* x \cdot \theta) = \pi^* (\text{Sq}^2(x) + c_1(E)x) \cdot \theta$$

for any $x \in H^*(X; \mathbb{Z}/2)$. More generally, if $E$ is a real oriented vector bundle, the second Steenrod square on the cohomology of its Thom space is given by “$\text{Sq}^2 + w_2(E)$”, where $w_2$ is the second Stiefel-Whitney class of $E$.

**Proof.** This is a special case of an identity of Thom, which he in fact used to define Stiefel-Whitney classes:

$$\text{Sq}^i(\pi^* x \cdot \theta) = \pi^* (\text{Sq}^i(x) + w_i(E)x) \cdot \theta$$

See [MS74, page 91].

When $X$ is a CW complex with cells only in even dimensions, the operation $\text{Sq}^2 + c_1$ can be viewed as a differential on $H^*(X; \mathbb{Z}/2)$ for any $c_1$ in $H^2(X; \mathbb{Z}/2)$. Extending our previous notation, we denote the cohomology with respect to this differential by

$$H^*(X, \text{Sq}^2 + c_1) := H^*(H^*(X; \mathbb{Z}/2), \text{Sq}^2 + c_1)$$

(21)
Corollary 3.6 (of Lemmas 3.2 and 3.5). If the Atiyah-Hirzebruch spectral sequence of $\widetilde{KO}^* (\text{Thom } \mathcal{E})$ degenerates on the $E_3$-page, then

$$KO^{2i-1}(X; \mathcal{E}) \cong \bigoplus_k H^{2i+8k}(X, Sq^2 + c_1 \mathcal{E})$$

It is true more generally that the differentials in the spectral sequence for $\widetilde{KO}^* (\text{Thom } \mathcal{E})$ depend only on the second Stiefel-Whitney class of $E$. This follows from the observation that the Atiyah-Hirzebruch spectral sequence is compatible with Thom isomorphisms, as is made more precise by the next lemma:

Fix a vector bundle $E$ of constant rank $r$ over a finite-dimensional CW complex $X$. Suppose $E$ is oriented with respect to ordinary cohomology and let $\theta \in \widetilde{E}^r_0 (\text{Thom } \mathcal{E})$ be a Thom class.

Lemma 3.7. If $E$ is oriented with respect to $KO^*$, then $\theta$ survives to the $E_\infty$-page of the Atiyah-Hirzebruch spectral sequence computing $\widetilde{KO}^* (\text{Thom } \mathcal{E})$, and the Thom isomorphism for $H^*$ extends to an isomorphism of spectral sequences. That is, for each page right multiplication with the class of $\theta$ in $\widetilde{E}_s^r (\text{Thom } \mathcal{E})$ gives an isomorphism of $E_*^s (X)$-modules

$$E_*^s (X) \mathbin{\overset{\theta}{\rightarrow}} \widetilde{E}_s^{r+s} (\text{Thom } \mathcal{E})$$

Moreover, any lift of $\theta \in \widetilde{E}_\infty^{r,0} (\text{Thom } \mathcal{E})$ to $\widetilde{KO}^r (\text{Thom } \mathcal{E})$ defines a Thom class of $E$ with respect to $KO^*$. The isomorphism of the $E_\infty$-pages of the spectral sequences is induced by the Thom isomorphism given by multiplication with any such class.

Proof. We may assume without loss of generality that $X$ is connected. Fix a point $x$ on $X$. The inclusion of the fibre over $x$ into $\mathcal{E}$ induces a map $i_x : S^r \hookrightarrow \text{Thom } \mathcal{E}$. By assumption, the pullback $i_x^*$ on ordinary cohomology maps $\theta$ to a generator of $\widetilde{H}^r (S^r)$, and the pullback on $\widetilde{KO}^*$ gives a surjection

$$\widetilde{KO}^* (\text{Thom } \mathcal{E}) \mathbin{\overset{i_x^*}{\twoheadrightarrow}} \widetilde{KO}^r (S^r)$$

Consider the pullback along $i_x$ on the $E_\infty$-pages of the spectral sequences for $S^r$ and $\text{Thom } \mathcal{E}$. Since we can identify $\widetilde{E}_\infty^{r,0} (\text{Thom } \mathcal{E})$ with a quotient of $\widetilde{KO}^r (\text{Thom } \mathcal{E})$ and $\widetilde{E}_\infty^{r,0} (S^r)$ with $\widetilde{KO}^r (S^r)$, we must have a surjection

$$i_x^* : \widetilde{E}_\infty^{r,0} (\text{Thom } \mathcal{E}) \twoheadrightarrow \widetilde{E}_\infty^{r,0} (S^r)$$

On the other hand, the behaviour of $i_x^*$ on $\widetilde{E}_\infty^{r,0}$ is determined by its behaviour on $\widetilde{H}^r$, whence we can only have such a surjection if $\theta$ survives to the $E_\infty$-page of $\text{Thom } \mathcal{E}$. Thus, all differentials vanish on $\theta$, and if multiplication by $\theta$ induces an isomorphism from $E_*^{r,s} (X)$ to $E_*^{r+s,*}$ on page $s$, it also induces an
isomorphism on the next page. Lastly, consider any lift of $\theta$ to an element $\Theta$ of $\widetilde{KO}^*(\text{Thom} \mathcal{E})$. It is clear by construction that right multiplication with $\Theta$ gives an isomorphism from $E_\infty(X)$ to $\tilde{E}_\infty(\text{Thom} E)$, and thus it also gives an isomorphism from $KO^*(X)$ to $\widetilde{KO}^*(\text{Thom} \mathcal{E})$. Thus, $\Theta$ is a Thom class for $\mathcal{E}$ with respect to $KO^*$.

Lemma 3.7 allows the following strengthening of Lemma 2.4:

**Corollary 3.8.** For complex vector bundles $\mathcal{E}$ and $\mathcal{F}$ over $X$ with identical first Chern class modulo 2, the spectral sequences computing $\widetilde{KO}^*(\text{Thom} \mathcal{E})$ and $\widetilde{KO}^*(\text{Thom} \mathcal{F})$ can be identified up to a possible shift of columns when $\mathcal{E}$ and $\mathcal{F}$ have different ranks.

4 Examples

We now turn to the study of projective homogeneous varieties, that is, varieties of the form $G/P$ for some complex simple linear algebraic group $G$ with a parabolic subgroup $P$. Any such variety has a cell decomposition [BGG73, Proposition 5.1], so that our comparison theorem applies. As far as we are only interested in the topology of $G/P$, we may alternatively view it as a homogeneous space for the compact real Lie group $G^c$ corresponding to $G$:

**Proposition 4.1.** Let $P$ be a parabolic subgroup of a simple complex algebraic group $G$. Then we have a diffeomorphism

$$G/P \cong G^c/K$$

where $K$ is a compact subgroup of maximal rank in a maximal compact subgroup $G^c$ of $G$. More precisely, $K$ is a maximal compact subgroup of a Levi subgroup of $P$.

**Proof.** The Iwasawa decomposition for $G$ viewed as a real Lie group implies that we have a diffeomorphism $G \cong G^c \cdot P$ [GOV94, Ch. 6, Prop. 1.7], inducing a diffeomorphism of quotients as claimed for $K = G^c \cap P$. Since $G^c \hookrightarrow G$ is a homotopy equivalence, so is the inclusion $G^c \cap P \hookrightarrow P$. On the other hand, if $L$ is a Levi subgroup of $P$ then $P = U \rtimes L$, where $U$ is unipotent and hence contractible. So the inclusion $L \hookrightarrow P$ is also a homotopy equivalence. It follows that any maximal compact subgroup $L^c$ of $L$ is also maximal compact in $P$, and conversely that any maximal compact subgroup of $P$ will be contained as a maximal compact subgroup in some Levi subgroup of $P$. We may therefore assume that $K \subset L^c \subset L \subset P$ and conclude that $K \hookrightarrow L^c$ is a homotopy equivalence. Since both groups are compact, it follows that in fact $K \cong L^c$.

The KO-theory of homogeneous varieties has been studied intensively. In particular, the papers [KH91] and [KH92] of Kono and Hara provide complete computations of the (untwisted) KO-theory of all compact irreducible hermitian symmetric spaces, which we list in Table 2. For the convenience of the
reader, we indicate how each of these arises as a quotient of a simple complex algebraic group $G$ by a parabolic subgroup $P$, describing the latter in terms of marked nodes on the Dynkin diagram of $G$ as in [FH91, § 23.3]. The last column gives an alternative description of each space as a quotient of a compact real Lie group.

On the following pages, we will run through this list of examples and, in each case, extend Kono and Hara’s computations to include KO-groups twisted by a line bundle. Since each of these spaces is a “Grassmannian” in the sense that the parabolic subgroup $P$ in $G$ is maximal, its Picard group is free abelian on a single generator. Thus, there is exactly one non-trivial twist that we need to consider. In most cases, we — reassuringly — recover results for Witt groups that are already known. In a few other cases, we consider our results new.

The untwisted KO-theory of complete flag varieties is also known in all three classical cases thanks to Kishimoto, Kono and Ohsita. We do not reproduce

| $G/p$ | $G$ | Diagram of $P$ | $G^c/K$ |
|-------|-----|----------------|---------|
| Grassmannians (AIII) $\text{Gr}_{m,n}$ | $\text{SL}_{m+n}$ | | $\text{U}(m+n)/\text{U}(m) \times \text{U}(n)$ |
| Maximal symplectic Grassmannians (CI) $X_n$ | $\text{Sp}_{2n}$ | | $\text{Sp}(n)/\text{U}(n)$ |
| Projective quadrics of dimension $n \geq 3$ (BDI) $Q^n$ | $\text{SO}_{n+2}$ | $(n$ odd) | $\text{SO}(n+2)/\text{SO}(n) \times \text{SO}(2)$ |
| | | $(n$ even) | |
| Spinor varieties (DIII) $S_n$ | $\text{SO}_{2n}$ | | $\text{SO}(2n)/\text{U}(n)$ |
| Exceptional hermitian symmetric spaces: EIII | $E_6$ | | $\text{Spin}(10) \cdot S^7$ |
| | | $(\text{Spin}(10) \cap S^7 = \text{Z}/4)$ | |
| EVII | $E_7$ | | $\text{E}_7 \cdot S^1$ |
| | | $(\text{E}_7 \cap S^3 = \text{Z}/3)$ | |

Table 2: List of irreducible compact hermitian symmetric spaces. The symbols AIII, CI, . . . refer to E. Cartan’s classification. In the description of $\text{Gr}_{m,n}$ we use $\text{U}(m+n)$ instead of $G^c = \text{SU}(m+n)$.
their result here but instead refer the reader directly to [KKO04]. By a recent result of Calmèes and Fasel, all Witt groups with non-trivial twists vanish for these varieties [CF11].

4.1 Notation

Topologically, a cellular variety is a CW complex with cells only in even (real) dimensions. For such a CW complex $X$ the KO-groups can be written in the form displayed in Table 3 below. This was shown in Section 3.1 in the case when the twist $\mathcal{L}$ is trivial, and the general case follows: if $X$ is a CW complex with only even-dimensional cells, so is the Thom space of any complex vector bundle over $X$ [MS74, Lemma 18.1].

In the following examples, results on $\text{KO}^*$ will be displayed by listing the values of the $t_i$ and $s_i$. Since the $t_i$ are just given by counting cells, and since the numbers of odd- and even-dimensional cells of a Thom space $\text{Thom}_X \mathcal{E}$ only depend on $X$ and the rank of $\mathcal{E}$, the $t_i$ are in fact independent of $\mathcal{L}$. The $s_i$, on the other hand, certainly will depend on the twist, and we will sometimes acknowledge this by writing $s_i(\mathcal{L})$.

$$
\begin{align*}
\text{KO}^6(X; \mathcal{L}) &= \mathbb{Z}^{t_1} \oplus (\mathbb{Z}/2)^{s_0} = \text{GW}^3(X; \mathcal{L}) \\
\text{KO}^7(X; \mathcal{L}) &= (\mathbb{Z}/2)^{s_0} = \text{W}^0(X; \mathcal{L}) \\
\text{KO}^0(X; \mathcal{L}) &= \mathbb{Z}^{t_0} \oplus (\mathbb{Z}/2)^{s_1} = \text{GW}^0(X; \mathcal{L}) \\
\text{KO}^1(X; \mathcal{L}) &= (\mathbb{Z}/2)^{s_1} = \text{W}^1(X; \mathcal{L}) \\
\text{KO}^2(X; \mathcal{L}) &= \mathbb{Z}^{t_1} \oplus (\mathbb{Z}/2)^{s_2} = \text{GW}^1(X; \mathcal{L}) \\
\text{KO}^3(X; \mathcal{L}) &= (\mathbb{Z}/2)^{s_2} = \text{W}^2(X; \mathcal{L}) \\
\text{KO}^4(X; \mathcal{L}) &= \mathbb{Z}^{t_0} \oplus (\mathbb{Z}/2)^{s_3} = \text{GW}^2(X; \mathcal{L}) \\
\text{KO}^5(X; \mathcal{L}) &= (\mathbb{Z}/2)^{s_3} = \text{W}^3(X; \mathcal{L})
\end{align*}
$$

Table 3: Notational conventions in the examples. Only the $s_i$ depend on $\mathcal{L}$.

4.2 Projective spaces

Complex projective spaces are perhaps the simplest examples for which Theorem 2.5 asserts something non-trivial, so we describe the results here separately before turning to complex Grassmannians in general. The computations of the Witt groups of projective spaces were certainly landmark events in the history of the theory. In 1980, Arason was able to show that the Witt group $\text{W}^0(\mathbb{P}^n)$ of $\mathbb{P}^n$ over a field $k$ agrees with the Witt group of $k$ [Ara80]. The shifted Witt groups of projective spaces, and more generally of arbitrary projective bundles, were first computed by Walter in [Wal03b]. Quite recently, Nenashev deduced the same results via different methods [Nen09].
In the topological world, complete computations of $\text{KO}^i(\mathbb{CP}^n)$ were first published in a 1967 paper by Fujii [Fuj67]. It is not difficult to deduce the values of the twisted groups $\text{KO}^i(\mathbb{CP}^n; \mathcal{O}(1))$ from these: the Thom space $\text{Thom}(\mathcal{O}_{\mathbb{C}P^n}(1))$ can be identified with $\mathbb{CP}^{n+1}$, so

$$\text{KO}^i(\mathbb{CP}^n; \mathcal{O}(1)) = \tilde{\text{KO}}^{i+2}(\text{Thom}(\mathcal{O}(1))) = \tilde{\text{KO}}^{i+2}(\mathbb{CP}^{n+1})$$

Alternatively, we could do all required computations directly following the methods outlined in Section 3. The result, in any case, is displayed in Table 4, coinciding with the known results for the (Grothendieck-)Witt groups.

| $\text{KO}^*(\mathbb{CP}^n; \mathcal{L})$ | $t_0$ | $t_1$ | $\mathcal{L} \equiv \mathcal{O}$ | $\mathcal{L} \equiv \mathcal{O}(1)$ |
|---|---|---|---|---|
| $n \equiv 0 \pmod{4}$ | $(n/2) + 1$ | $n/2$ | 1 0 0 0 | 1 0 0 0 |
| $n \equiv 1$ | $(n+1)/2$ | $(n+1)/2$ | 1 1 0 0 | 0 0 0 0 |
| $n \equiv 2$ | $(n/2) + 1$ | $n/2$ | 1 0 0 0 | 0 0 1 0 |
| $n \equiv 3$ | $(n+1)/2$ | $(n+1)/2$ | 1 0 0 1 | 0 0 0 0 |

Table 4: KO-groups of projective spaces

### 4.3 Grassmannians

We now consider the Grassmannians $\text{Gr}_{m,n}$ of complex $m$-planes in $\mathbb{C}^{m+n}$. Again both the Witt groups and the untwisted KO-groups are already known: the latter by Kono and Hara [KH91], the former by the work of Balmer and Calmès [BC08]. A detailed comparison of the two sets of results in the untwisted case has been carried out by Yagita [Yag09]. We provide here an alternative topological computation of the twisted groups.

Balmer and Calmès state their result by describing an additive basis of the total Witt group of $\text{Gr}_{m,n}$ in terms of certain “even Young diagrams”. This is probably the most elegant approach, but needs some space to explain. We will stick instead to the tabular exposition used in the other examples. Let $\mathcal{O}(1)$ be a generator of $\text{Pic}(\text{Gr}_{m,n})$, say the dual of the determinant line bundle of the universal $m$-bundle over $\text{Gr}_{m,n}$. The result is displayed in Table 5.

Our computation is based on the following geometric observation. Let $\mathcal{U}_{m,n}$ and $\mathcal{U}_{m,n}^\perp$ be the universal $m$-bundle and the orthogonal $n$-bundle on $\text{Gr}_{m,n}$, so that $\mathcal{U} \oplus \mathcal{U}^\perp = \mathcal{O}^{\oplus(m+n)}$. We have various natural inclusions between the Grassmannians of different dimensions, of which we fix two:

$\text{Gr}_{m,n-1} \hookrightarrow \text{Gr}_{m,n}$ via the inclusion of the first $m + n - 1$ coordinates into $\mathbb{C}^{m+n}$

$\text{Gr}_{m-1,n} \hookrightarrow \text{Gr}_{m,n}$ by sending an $(m-1)$-plane $\Lambda$ to the $m$-plane $\Lambda \oplus (e_{m+n})$, where $e_1, e_2, \ldots, e_{m+n}$ are the canonical basis vectors of $\mathbb{C}^{m+n}$.
KO*(Gr_{m,n}; \mathcal{L}) & t_0 & t_1 & \mathcal{L} \equiv \mathcal{O} & s_0 & s_1 & s_2 & s_3 & \mathcal{L} \equiv \mathcal{O}(1) & s_0 & s_1 & s_2 & s_3 \\

m \text{ and } n \text{ odd s.t. } & a & a & b & b & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\

m \equiv n & a & a & b & b & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\

m \text{ and } n \text{ odd s.t. } & a & a & b & b & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\

m \not\equiv n & a & a & b & b & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\

\begin{align*}
& \begin{cases} m \equiv n \equiv 0 \\
& \begin{cases} m \equiv 0 \text{ and } n \text{ odd} \\
& \begin{cases} n \equiv 0 \text{ and } m \text{ odd} \\
& \begin{cases} m \equiv 0 \text{ and } n \equiv 2 \\
& \begin{cases} m \equiv 2 \text{ and } n \text{ odd} \\
& \begin{cases} a \equiv 2 \text{ and } m \text{ odd} \\
\end{cases} \\
\end{cases} \\
\end{cases} \\
\end{cases} \\
\end{cases} \\
\end{align*}

\begin{align*}
& \begin{cases} m \equiv n \equiv 0 \\
& \begin{cases} m \equiv 0 \text{ and } n \text{ odd} \\
& \begin{cases} n \equiv 0 \text{ and } m \text{ odd} \\
& \begin{cases} m \equiv 0 \text{ and } n \equiv 2 \\
& \begin{cases} m \equiv 2 \text{ and } n \text{ odd} \\
& \begin{cases} a \equiv 2 \text{ and } m \text{ odd} \\
\end{cases} \\
\end{cases} \\
\end{cases} \\
\end{cases} \\
\end{cases} \\
\end{align*}

\begin{align*}
& \begin{cases} m \equiv n \equiv 0 \\
& \begin{cases} m \equiv 0 \text{ and } n \equiv 2 \\
& \begin{cases} m \equiv 2 \text{ and } n \equiv 0 \\
\end{cases} \\
\end{cases} \\
\end{cases} \\
\end{align*}

All equivalences (\equiv) are modulo 4. For the values of \(a = b_1 + b_2\), put 
\(k := \lfloor \frac{m}{2} \rfloor\) and \(l := \lfloor \frac{n}{2} \rfloor\). Then 
\[ a := \binom{m+n}{m} \quad b := \binom{k+l}{k} \quad b_1 := \binom{k+l-1}{k} \quad b_2 := \binom{k+l-1}{k-1} \]

Table 5: KO-groups of Grassmannians

**Lemma 4.2.** The normal bundle of \(Gr_{m,n-1}\) in \(Gr_{m,n}\) is the dual \(U_{m,n-1}^\vee\) of the universal \(m\)-bundle. Similarly, the normal bundle of \(Gr_{n-1,m}\) in \(Gr_{m,n}\) is given by \(U_{m,n-1}^\perp\). In both cases, the embeddings of the subspaces extend to embeddings of their normal bundles, such that one subspace is the closed complement of the normal bundle of the other.

This gives us two cofibration sequences of pointed spaces:

\[
Gr_{m-1,n,+} \overset{i}{\hookrightarrow} Gr_{m,n,+} \overset{p}{\twoheadrightarrow} Thom(U_{m,n-1}^\vee) \quad (22)
\]

\[
Gr_{m,n-1,+} \overset{i}{\hookrightarrow} Gr_{m,n,+} \overset{p}{\twoheadrightarrow} Thom(U_{m-1,n}^\perp) \quad (23)
\]

These sequences are the key to relating the untwisted KO-groups to the twisted ones. Following the notation in [KH91], we write \(A_{m,n}\) for the cohomology of \(Gr_{m,n}\) with \(\mathbb{Z}/2\)-coefficients, denoting by \(a_i\) and \(b_i\) the Chern classes of \(U\) and \(U^\perp\), respectively, and by \(a\) and \(b\) the total Chern classes \(1 + a_1 + \cdots + a_m\) and
1 + b_1 + \cdots + b_n: 

\[ A_{m,n} = \frac{\mathbb{Z}/2[a_1, a_2, \ldots, a_m, b_1, b_2, \ldots, b_n]}{a \cdot b = 1} \]

We write \( d \) for the differential given by the second Steenrod square \( Sq^2 \), and \( d' \) for \( Sq^2 + a_1 \). To describe the cohomology of \( A_{m,n} \) with respect to these differentials, it is convenient to introduce the algebra

\[ B_{k,l} = \frac{\mathbb{Z}/2[a_2^2, a_4^2, \ldots, a_{2k}^2, b_2^2, b_4^2, \ldots, b_{2l}^2]}{(1 + a_2^2 + \cdots + a_{2k}^2)(1 + b_2^2 + \cdots + b_{2l}^2)} \]

Note that this subquotient of \( A_{2k,2l} \) is isomorphic to \( A_{k,l} \) up to a “dilatation” in grading. Proposition 2 in [KH91] tells us that

\[ H^*(A_{m,n}, d) = \begin{cases} 
B_{k,l} & \text{if } (m, n) = (2k, 2l), (2k + 1, 2l) \text{ or } (2k, 2l + 1) \\
B_{k,l} \oplus B_{k,l} \cdot a_m b_{n-1} & \text{if } (m, n) = (2k + 1, 2l + 1)
\end{cases} \]

Here, the algebra structure in the case where both \( m \) and \( n \) are odd is determined by \((a_m b_{n-1})^2 = 0\).

**Lemma 4.3.** The cohomology of \( A_{m,n} \) with respect to the twisted differential \( d' \) is as follows:

\[ H^*(A_{m,n}, d') = \begin{cases} 
B_{k,l-1} \cdot a_m \oplus B_{k-1,l} \cdot b_n & \text{if } (m, n) = (2k, 2l) \\
B_{k,l} \cdot a_m & \text{if } (m, n) = (2k, 2l + 1) \\
B_{k,l} \cdot b_n & \text{if } (m, n) = (2k + 1, 2l) \\
0 & \text{if } (m, n) = (2k + 1, 2l + 1)
\end{cases} \]

**Proof.** Let us shift the dimensions in the cofibration sequences (22) and (23) in such a way that we have the Thom spaces of \( U^\vee_{m,n} \) and \( U^\perp_{m,n} \) on the right. Since the cohomologies of the spaces involved are concentrated in even degrees, the associated long exact sequence of cohomology groups falls apart into short exact sequences. Reassembling these, we obtain two short exact sequences of differential \((A_{m,n+1}, d)\)- and \((A_{m+1,n}, d)\)-modules, respectively:

\[ 0 \to (A_{m,n}, d') \cdot \theta^\vee \xrightarrow{p^\ast} (A_{m,n+1}, d) \xrightarrow{i^\ast} (A_{m-1,n+1}, d) \to 0 \quad (24) \]

\[ 0 \to (A_{m,n}, d') \cdot \theta^\perp \xrightarrow{p^\ast} (A_{m+1,n}, d) \xrightarrow{i^\ast} (A_{m+1,n-1}, d) \to 0 \quad (25) \]

Here, \( \theta^\vee \) and \( \theta^\perp \) are the respective Thom classes of \( U^\vee_{m,n} \) and \( U^\perp_{m,n} \). The map \( i^\ast \) in the first row is the obvious quotient map annihilating \( a_m \). Its kernel, the image of \( A_{m,n} \) under multiplication by \( a_m \), is generated as an \( A_{m,n+1} \)-module by its unique element in degree \( 2m \), and thus we must have \( p^\ast(\theta^\vee) = a_m \). Likewise, in the second row we have \( p^\ast(\theta^\perp) = b_n \).
The Atiyah-Hirzebruch spectral sequence for
By Proposition 4 of [KH91] we know that the spectral sequence for
the Chern roots of
that
The other cases are simpler.
define a splitting of
one may check explicitly that the boundary map
summand
count shows that together they encompass all of
we work with sequence (25), we see that
B
map is
in the (2
For the Steenrod square Sq
B
ko
The lemma can be deduced from here case by case. For example, when both
m and n are even, \( i^* \) maps \( H^*(A_{m,n+1}, d) = B_{k,l} \) to the first summand of
H
that
\( \partial = 0 \).
Thus, we obtain a short exact sequence
\[
0 \to B_{k-1,l} \cdot a_{m-1}b_n \xrightarrow{\partial} H^*(A_{m,n}, d') \cdot \theta^\tau \xrightarrow{p^*} B_{k,l-1} \cdot a_m^2 \to 0 \tag{26}
\]
For the Steenrod square Sq
, of the top Chern class \( a_m \) of \( U \), we have Sq\( ^2(a_m) = a_1 a_m \). This can be checked, for example, by expressing \( a_m \) as the product of the Chern roots of \( U \). Consequently, \( d'(a_m) = 0 \). Together with the fact that
H
that we can define a splitting of \( p^* \) by sending \( a_m^2 \) to \( a_m \theta^\tau \). Thus,
H
B
summand
B
\( \partial \) above sends \( a_{m-1}b_n \) to \( b_n \theta \).
The other cases are simpler.

**Lemma 4.4.** The Atiyah-Hirzebruch spectral sequence for \( \tilde{KO}^*(\text{Thom} U_{m,n}^0) \) collapses at the \( E_3 \)-page.

**Proof.** By Proposition 4 of [KH91] we know that the spectral sequence for
KO
(\( \text{Gr}_{m,n} \)) collapses as this stage, for any \( m \) and \( n \). Now, if both \( m \) and \( n \) are even, we have
\[
(B_{k,l-1} \cdot a_m \oplus B_{k-1,l} \cdot b_n) \cdot \theta
\]
in the \((-1)^{k+l}\) row of the \( E_3 \)-pages of the spectral sequences for Thom
\( U^\vee \) and
\( \text{Thom} U^\perp \), where \( \theta = \theta^\vee \) or \( \theta^\perp \), respectively. In the case of \( U^\vee \) we see from (26) that \( p^* \) maps the second summand injectively to the \( E_3 \)-page of the spectral sequence for \( KO^*(\text{Gr}_{m,n+1}) \). Similarly, in the case of \( U^\perp \), the first summand is mapped injectively to the \( E_3 \)-page of \( KO^*(\text{Gr}_{m+1,n}) \). Since the spectral sequences for Thom
\( U^\vee \) and
\( \text{Thom} U^\perp \) can be identified via Corollary 3.8, we can argue as in Corollary 3.4 to see that they must collapse at this stage. Again, the cases when at least one of \( m, n \) is odd are similar but simpler.

We may now apply Corollary 3.6. The entries of Table 5 that do not appear in [KH91], i.e. those of the last four columns, follow from Lemma 4.3 by noting that \( B_{k,l} \) is concentrated in degrees \( 8i \) and of dimension \( \dim B_{k,l} = \dim A_{k,l} = (k+l) \).

**4.4 Maximal Symplectic Grassmannians**

The Grassmannian of isotropic \( n \)-planes in \( \mathbb{C}^{2n} \) with respect to a non-degenerate skew-symmetric bilinear form is given by \( X_n = \text{Sp}(n)/U(n) \). The
universal bundle $U$ on the usual Grassmannian $\text{Gr}(n, 2n)$ restricts to the universal bundle on $X_n$, and so does the orthogonal complement bundle $U^\perp$. We will continue to denote these restrictions by the same letters. Thus, $U \oplus U^\perp \cong \mathbb{C}^{2n}$ on $X_n$, and the fibres of $U$ are orthogonal to those of $U^\perp$ with respect to the standard hermitian metric on $\mathbb{C}^{2n}$. The determinant line bundles of $U$ and $U^\perp$ give dual generators $\mathcal{O}(1)$ and $\mathcal{O}(-1)$ of the Picard group of $X_n$.

**Theorem 4.5.** The additive structure of $\text{KO}^*(X_n; \mathcal{L})$ is as follows:

|   | $t_0$ | $t_1$ | $s_i(\mathcal{O})$ | $s_i(\mathcal{O}(1))$ |
|---|---|---|---|---|
| $n$ even | $2^{n-1}$ | $2^{n-1}$ | $\rho(\frac{n}{2}, i)$ | $\rho(\frac{n}{2}, i - n)$ |
| $n$ odd | $2^{n-1}$ | $2^{n-1}$ | $\rho(\frac{n+1}{2}, i)$ | 0 |

Here, for any $i \in \mathbb{Z}/4$ we write $\rho(n, i)$ for the dimension of the $i$-graded piece of a $\mathbb{Z}/4$-graded exterior algebra $\Lambda_{\mathbb{Z}/2}(g_1, g_2, \ldots, g_n)$ on $n$ homogeneous generators $g_1, g_2, \ldots, g_n$ of degree 1, i.e.

$$\rho(n, i) = \sum_{d \equiv i \mod 4} \binom{n}{d}$$

A table of the values of $\rho(n, i)$ can be found in [KH92, Proposition 4.1].

It turns out to be convenient to work with the vector bundle $U^\perp \oplus \mathcal{O}$ for the computation of the twisted groups $\text{KO}^*(X_n; \mathcal{O}(1))$. Namely, we have the following analogue of Lemma 4.2.

**Lemma 4.6.** There is an open embedding of the bundle $U^\perp \oplus \mathcal{O}$ over the symplectic Grassmannian $X_n$ into the symplectic Grassmannian $X_{n+1}$ whose closed complement is again isomorphic to $X_n$.

**Proof.** To fix notation, let $e_1, e_2$ be the first two canonical basis vectors of $\mathbb{C}^{2n+2}$, and embed $\mathbb{C}^{2n}$ into $\mathbb{C}^{2n+2}$ via the remaining coordinates. Assuming $X_n$ is defined in terms of a skew-symmetric form $Q_{2n}$, define $X_{n+1}$ with respect to the form

$$Q_{2n+2} := \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & Q_{2n} \end{pmatrix}$$

Then we have embeddings $i_1$ and $i_2$ of $X_n$ into $X_{n+1}$ sending an $n$-plane $\Lambda \subset \mathbb{C}^{2n}$ to $e_1 \oplus \Lambda$ or $e_2 \oplus \Lambda$ in $\mathbb{C}^{2n+2}$, respectively.

We extend $i_1$ to an embedding of $U^\perp \oplus \mathcal{O}$ by sending an $n$-plane $\Lambda \subset X_n$ together with a vector $v$ in $\Lambda^\perp \subset \mathbb{C}^{2n}$ and a complex scalar $z$ to the graph $\Gamma_{\Lambda, v, z} \subset \mathbb{C}^{2n+2}$ of the linear map

$$\begin{pmatrix} z & Q_{2n}(-, v) \\ 0 & 0 \end{pmatrix} : (e_1) \oplus \Lambda \rightarrow (e_2) \oplus \Lambda^\perp$$
To avoid confusion, we emphasize that \( v \) is orthogonal to \( \Lambda \) with respect to a hermitian metric on \( \mathbb{C}^{2n} \). The value of \( Q_{2n}(-,v) \), on the other hand, may well be non-zero on \( \Lambda \). Consider the above embedding of \( \mathcal{U}^\perp \oplus \mathcal{O} \) together with the embedding \( i_2 \):

\[
\mathcal{U}^\perp \oplus \mathcal{O} \quad \rightarrow \quad X_{n+1} \quad \rightarrow \quad X_n
\]

\[
(\Lambda, v, z) \quad \rightarrow \quad \Gamma_{\Lambda,v,z}
\]

\[
\langle e_2 \rangle \oplus \Lambda \quad \rightarrow \quad \Lambda
\]

To see that the two embeddings are complementary, take an arbitrary \((n+1)\)-plane \( W \) in \( X_{n+1} \). If \( e_2 \in W \) then we can consider a basis

\[
e_{2*} = \left(\begin{array}{c}
am_1 \\
0 \end{array}\right), \ldots, \left(\begin{array}{c}
am_n \\
0 \end{array}\right)
\]

of \( W \), and the fact that \( Q_{2n+2} \) vanishes on \( W \) implies that all \( a_i \) are zero. Thus \( W \) can be identified with \( i_2((v_1, \ldots, v_n)) \).

If, on the other hand, \( e_2 \) is not contained in \( W \) then we must have a vector of the form \( i^*(1, z', v') \) in \( W \), for some \( z' \in \mathbb{C} \) and \( v' \in \mathbb{C}^{2n} \). Extend this vector to a basis of \( W \) of the form

\[
\left(\begin{array}{c}
1 \\
z'
\end{array}\right), \left(\begin{array}{c}
0 \\
v_1
\end{array}\right), \ldots, \left(\begin{array}{c}
0 \\
v_n
\end{array}\right)
\]

and let \( \Lambda := \langle v_1, \ldots, v_n \rangle \). The condition that \( Q_{2n+2} \) vanishes on \( W \) implies that \( Q \) vanishes on \( \Lambda \) and that \( b_i = Q_{2n}(v_i, v') \) for each \( i \). In particular, \( \Lambda \) is \( n \)-dimensional. Moreover, we can replace the first vector of our basis by a vector \( i^*(1, z, v) \) with \( v \in \Lambda^\perp \), by subtracting appropriate multiples of the remaining basis vectors. Since \( Q \) vanishes on \( \Lambda \) we have \( Q_{2n}(v_i, v') = Q_{2n}(v_i, v) \) and our new basis has the form

\[
\left(\begin{array}{c}
1 \\
0 \\
Q(v_1,v)
\end{array}\right), \left(\begin{array}{c}
0 \\
v_1
\end{array}\right), \ldots, \left(\begin{array}{c}
0 \\
v_n
\end{array}\right)
\]

This shows that \( W = \Gamma_{\Lambda,v,z} \).

**Corollary 4.7.** We have a cofibration sequence

\[
X_{n+1} \rightarrow X_{n+1} \rightarrow \text{Thom}_{\mathcal{X}_n}(\mathcal{U}^\perp \oplus \mathcal{O})
\]

The associated long exact cohomology sequence splits into a short exact sequence of \( H^*(X_{n+1}) \)-modules since all cohomology here is concentrated in even degrees:

\[
0 \rightarrow \tilde{H}^*(\text{Thom}_{\mathcal{X}_n}(\mathcal{U}^\perp \oplus \mathcal{O})) \rightarrow H^*(X_{n+1}) \rightarrow H^*(X_n) \rightarrow 0
\]

**Lemma 4.8.** Let \( c_i \) denote the \( i \)th Chern classes of \( \mathcal{U} \) over \( X_n \). We have

\[
H^*(X_n, \mathbb{Q}) = \begin{cases} \Lambda(a_1, a_5, a_9, \ldots, a_{4m-3}) & \text{if } n = 2m \\ \Lambda(a_1, a_5, a_9, \ldots, a_{4m-3}, a_{4m+1}) & \text{if } n = 2m + 1 \end{cases}
\]

\[
H^*(X_n, \mathbb{Q}^2 + c_1) = \begin{cases} \Lambda(a_1, a_5, \ldots, a_{4m-3}) \cdot e_{2m} & \text{if } n = 2m \\ 0 & \text{if } n \text{ is odd} \end{cases}
\]
for certain generators $a_i$ of degree $2i$.

Proof. Consider the short exact sequence (27). The mod-2 cohomology of $X_n$ is an exterior algebra on the Chern classes $c_i$ of $U$,

$$H^*(X_n; \mathbb{Z}/2) = \Lambda(c_1, c_2, \ldots, c_n)$$

and $i^*$ is given by sending $c_{n+1}$ to zero. Thus, $p^*$ is the unique morphism of $H^*(X_{n+1}; \mathbb{Z}/2)$-modules that sends the Thom class $\theta$ of $U^+ \oplus O$ to $c_{n+1}$.

This short exact sequence induces a long exact sequence of cohomology groups with respect to the Steenrod square $Sq^2$. The algebra $H^*(X_n, Sq^2)$ was computed in [KH92, 2–2], with the result displayed above, so we already know two thirds of this sequence. Explicitly, we have $a_{4i+1} = c_{2i} c_{2i+1}$, so $i^*$ is the obvious surjection sending $a_i$ to $a_i$ (or to zero). Thus, the long exact sequence once again splits.

If $n = 2m$ we obtain a short exact sequence

$$0 \to H^*(X_{2m}, Sq^2 + c_1) \cdot \theta \overset{i^*}{\to} \Lambda(a_1, \ldots, a_{4m-3}, a_{4m+1}) \overset{p^*}{\to} \Lambda(a_1, \ldots, a_{4m-3}) \to 0$$

We see that $H^*(X_{2m}, Sq^2 + c_1) \cdot \theta$ is isomorphic to $\Lambda(a_1, \ldots, a_{4m-3}) \cdot a_{4m+1}$ as a $\Lambda(a_1, \ldots, a_{4m+1})$-module. It is thus generated by a single element, which is the unique element of degree $8m + 2$. Since $p^*(c_{2m} \theta) = a_{4m+1}$, the class of $c_{2m} \theta$ is the element we are looking for, and the result displayed above follows. If, on the other hand, $n$ is odd, then $i^*$ is an isomorphism and $H^*(X_n, Sq^2 + c_1)$ must be trivial. \hfill \square

We see from the proof that $p^*$ induces an injection of $H^*(X_n, Sq^2 + c_1) \cdot \theta$ into $H^*(X_n, Sq^2)$. Since we already know from [KH92, Theorem 2.1] that the Atiyah-Hirzebruch spectral sequence for $KO^*(X_n)$ collapses, we can apply Corollary 3.4 to deduce that the spectral sequence for $\tilde{KO}^*(\text{Thom}_{X_n}(U^+ \oplus O))$ collapses at the $E_3$-page as well. This completes the proof of Theorem 4.5.

4.5 Quadrics

We next consider smooth complex quadrics $Q^n$ in $\mathbb{P}^{n+1}$. As far as we are aware, the first complete results on (shifted) Witt groups of split quadrics were due to Walter: they are mentioned together with the results for projective bundles in [Wal03a] as the main applications of that paper. Unfortunately, they seem to have remained unpublished. Partial results are also included in Yagita’s preprint [Yag04], see Corollary 8.3. More recently, Nenashev obtained almost complete results by considering the localization sequences arising from the inclusion of a linear subspace of maximal dimension [Nen09]. Calmès informs me that the geometric description of the boundary map given in [BC09] can be used to show that these localization sequences split in general, yielding a complete computation. The calculation described here is completely independent of these results.

\footnote{In [KH92] the generators are written as $c_{2i} c'_{2i+1}$ with $c'_{2i+1} = c_{2i+1} + c_1 c_{2i}$.}
KO∗(Qn; Ł) & t0 & t1 & Ł ≡ O & Ł ≡ O(1) \\ \hline n ≡ 0 \ mod 8 & (n/2) + 2 & n/2 & 2 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ n ≡ 1 & (n + 1)/2 & (n + 1)/2 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ n ≡ 2 & (n/2) + 1 & (n/2) + 1 & 1 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ n ≡ 3 & (n + 1)/2 & (n + 1)/2 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ n ≡ 4 & (n/2) + 2 & n/2 & 2 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ n ≡ 5 & (n + 1)/2 & (n + 1)/2 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ n ≡ 6 & (n/2) + 1 & (n/2) + 1 & 1 & 0 & 1 & 2 & 0 & 0 & 0 & 0 \\ n ≡ 7 & (n + 1)/2 & (n + 1)/2 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ \hline

Table 6: KO-groups of projective quadrics \((n \geq 3)\)

For \(n \geq 3\) the Picard group of \(Q^n\) is free abelian on a single generator given by the restriction of the universal line bundle \(O(1)\) over \(\mathbb{P}^{n+1}\). We will use the same notation \(O(1)\) for this restriction.

**Theorem 4.9.** The KO-theory of a smooth complex quadric \(Q^n\) of dimension \(n \geq 3\) is as described in Table 6.

**Untwisted KO-groups.** Before turning to \(\text{KO}^∗(Q^n; O(1))\) we review the initial steps in the computation of the untwisted KO-groups. The integral cohomology of \(Q^n\) is well-known:

If \(n\) is even, write \(n = 2m\). We have a class \(x\) in \(H^2(Q^n)\) given by a hyperplane section, and two classes \(a\) and \(b\) in \(H^n(Q^n)\) represented by linear subspaces of \(Q\) of maximal dimension. These three classes generate the cohomology multiplicatively, modulo the relations

\[
\begin{align*}
x^m &= a + b \\
ab &\begin{cases} 0 & \text{if } n \equiv 0 \\
ax^m & \text{if } n \equiv 2\end{cases} \\
a^2 &= b^2 \begin{cases} ax^m & \text{if } n \equiv 0 \mod 4 \\
0 & \text{if } n \equiv 2 \mod 4\end{cases}
\end{align*}
\]

Additive generators can thus be given as follows:

\[
\begin{array}{c|cccccccc}
    d & 0 & 2 & 4 & \ldots & n-2 & n & n+2 & n+4 & \ldots & 2n \\
    \hline
    H^d(Q^n) & 1 & x & x^2 & \ldots & x^{m-1} & a, b & ax & ax^2 & \ldots & ax^m
\end{array}
\]

If \(n\) is odd, write \(n = 2m + 1\). Then similarly multiplicative generators are given by the class of a hyperplane section \(x\) in \(H^2(Q^n)\) and the class of a linear subspace \(a\) in \(H^{n+1}(Q^n)\) modulo the relations \(x^{m+1} = 2a\) and \(a^2 = 0\).

\[
\begin{array}{c|cccccccc}
    d & 0 & 2 & 4 & \ldots & n-1 & n+1 & n+3 & n+5 & \ldots & 2n \\
    \hline
    H^d(Q^n) & 1 & x & x^2 & \ldots & x^m & a & ax & ax^2 & \ldots & ax^m
\end{array}
\]
The action of the Steenrod square on $H^*(Q^n;\mathbb{Z}/2)$ is also well-known; see for example [Ish92, Theorem 1.4 and Corollary 1.5] or [EKM08, §78]:

\[
\begin{align*}
\text{Sq}^2(x) &= x^2 \\
\text{Sq}^2(a) &= \begin{cases} 
ax & \text{if } n \equiv 0 \text{ or } 3 \mod 4 \\
0 & \text{if } n \equiv 1 \text{ or } 2 
\end{cases} \\
\text{Sq}^2(b) &= \text{Sq}^2(a) \quad (\text{for even } n)
\end{align*}
\]

As before, we write $H^*(Q^n,\text{Sq}^2)$ for the cohomology of $H^*(Q^n;\mathbb{Z}/2)$ with respect to the differential $\text{Sq}^2$.

**Lemma 4.10.** Write $n = 2m$ or $n = 2m+1$ as above. The following table gives a complete list of the additive generators of $H^*(Q^n,\text{Sq}^2)$.

| $d$ | 0 | \ldots | $n-1$ | $n$ | $n+1$ | \ldots | $2n$ |
|-----|---|--------|-------|----|-------|--------|------|
| $H^d(Q^n,\text{Sq}^2)$ | 1 | $ax^m$ if $n \equiv 0 \mod 4$ | 1 | $a$ if $n \equiv 1$ | 1 | $a,b$ ab if $n \equiv 2$ | 1 | $x^m$ if $n \equiv 3$ |

The results of Kono and Hara on KO$^*(Q)$ follow from here provided there are no non-trivial higher differentials in the Atiyah-Hirzebruch spectral sequence. This is fairly clear in all cases except for the case $n \equiv 2 \mod 4$. In that case, the class $a+b = x^m$ can be pulled back from $Q^{n+1}$, and therefore all higher differentials must vanish on $a+b$. But one has to work harder to see that all higher differentials vanish on $a$ (or $b$). Kono and Hara proceed by relating the KO-theory of $Q^n$ to that of the spinor variety $S_{2+1}$ discussed in Section 4.6.

**Twisted KO-groups.** We now compute KO$^*(Q^n;\mathcal{O}(1))$.

Let $\theta \in H^2(\text{Thom}_{Q^n}\mathcal{O}(1))$ be the Thom class of $\mathcal{O}(1)$, so that multiplication by $\theta$ maps the cohomology of $Q^n$ isomorphically to the reduced cohomology of $\text{Thom}_{Q^n}\mathcal{O}(1)$. The Steenrod square on $H^*(\text{Thom}_{Q^n}\mathcal{O}(1);\mathbb{Z}/2)$ is determined by Lemma 3.5: for any $y \in H^*(Q^n;\mathbb{Z}/2)$ we have $\text{Sq}^2(y \cdot \theta) = (\text{Sq}^2 y + xy) \cdot \theta$.

We thus arrive at

**Lemma 4.11.** The following table gives a complete list of the additive generators of $H^*(\text{Thom}_{Q^n}\mathcal{O}(1),\text{Sq}^2)$.

| $d$ | \ldots | $n+1$ | $n+2$ | $n+3$ | \ldots | $2n+2$ |
|-----|--------|-------|-------|-------|--------|------|
| $\tilde{H}^d(\ldots)$ | $a\theta, b\theta$ | $ax^m\theta$ if $n \equiv 0 \mod 4$ | $x^m\theta$ | $ax^m\theta$ if $n \equiv 1$ | $a\theta$ | $ax^m\theta$ if $n \equiv 3$ |
We claim that all higher differentials in the Atiyah-Hirzebruch spectral sequence for $K^*(\text{Thom}_{Q^n}\mathcal{O}(1))$ vanish. For even $n$ this is clear. But for $n = 8k + 1$ the differential $d_{8k+2}$ might a priori take $x^m\theta$ to $ax^m\theta$, and for $n = 8k + 3$ the differential $d_{8k+3}$ might take $a\theta$ to $ax^m\theta$. 

We therefore need some geometric considerations. Namely, the Thom space $\text{Thom}_{Q^n}\mathcal{O}(1)$ can be identified with the projective cone over $Q^n$ embedded in $\mathbb{P}^{n+2}$. This projective cone can be realized as the intersection of a smooth quadric $Q^{n+2} \subset \mathbb{P}^{n+3}$ with its projective tangent space at the vertex of the cone [Har92, p. 283]. Thus, we can consider the following inclusions:

$$Q^n \xrightarrow{i} \text{Thom}_{Q^n}\mathcal{O}(1) \xrightarrow{j} Q^{n+2}$$

The composition is the inclusion of the intersection of $Q^{n+2}$ with two transversal hyperplanes.

**Lemma 4.12.** All higher differentials ($d_k$ with $k > 2$) in the Atiyah-Hirzebruch spectral sequence for $K^*(\text{Thom}_{Q^n}\mathcal{O}(1))$ vanish.

**Proof.** We need only consider the cases when $n$ is odd. Write $n = 2m + 1$. When $n \equiv 1 \pmod{4}$ we claim that $i^*$ maps $x^{m+1}$ in $H^{n+1}(\text{Thom}_{Q^n}\mathcal{O}(1), Sq^2)$ to $x^m\theta$ in $H^{n+1}(\text{Thom}_{Q^n}\mathcal{O}(1), Sq^2)$. Indeed, $j^*i^*$ maps the class of the hyperplane section $x$ in $H^2(Q^{n+2})$ to the class of the hyperplane section $x$ in $H^2(Q^n)$. So $i^*x$ in $H^2(\text{Thom}_{Q^n}\mathcal{O}(1))$ must be non-zero, hence equal to $\theta$ modulo 2. It follows that $i^*(x^{m+1}) = x^m\theta$. Since $\theta^2 = Sq^2(\theta) = x\theta$, we have $\theta^{m+1} = x^m\theta$, proving the claim. As we already know that all higher differentials vanish on $H^*(Q^{n+2}, Sq^2)$, we may now deduce that they also vanish on $H^*(\text{Thom}_{Q^n}\mathcal{O}(1), Sq^2)$.

When $n \equiv 3 \pmod{4}$ we claim that $i^*$ maps $a$ in $H^{n+3}(Q^{n+2}, Sq^2)$ to $a\theta$ in $H^{n+3}(\text{Thom}_{Q^n}\mathcal{O}(1), Sq^2)$. Indeed, $a$ represents a linear subspace of codimension $m + 2$ in $Q^{n+2}$ and is thus mapped to the class of a linear subspace of the same codimension in $Q^n$: $j^*i^*(a) = ax$ in $H^{n+3}(Q^n)$. Thus, $i^*(a)$ is non-zero in $H^{n+3}(\text{Thom}_{Q^n}\mathcal{O}(1))$, equal to $a\theta$ modulo 2. Again, this implies that all higher differentials vanish on $H^*(\text{Thom}_{Q^n}\mathcal{O}(1), Sq^2)$ since they vanish on $H^*(Q^{n+2}, Sq^2)$. 

The additive structure of $K^*(Q^n; \mathcal{O}(1))$ thus follows directly from the result for $H^2(Q^n, Sq^2 + x) = \tilde{H}^{d+2}(\text{Thom}_{Q^n}\mathcal{O}(1))$ displayed in Lemma 4.11 via Corollary 3.6.

### 4.6 Spinor varieties

Let $\text{Gr}_{SO}(n, N)$ be the Grassmanian of $n$-planes in $\mathbb{C}^N$ isotropic with respect to a fixed non-degenerate symmetric bilinear form, or, equivalently, the Fano variety of projective $(n-1)$-planes contained in the quadric $Q^{N-2}$. For each $N > 2n$, this is an irreducible homogeneous variety. In particular, for $N = 2n + 1$ we obtain the spinor variety $S_{n+1} = \text{Gr}_{SO}(n, 2n+1)$. The variety $\text{Gr}_{SO}(n, 2n)$
falls apart into two connected components, both of which are isomorphic to $S_n$. This is reflected by the fact that we can equivalently identify $S_n$ with $SO(2n-1)/U(n-1)$ or $SO(2n)/U(n)$.

As for all Grassmannians, the Picard group of $S_n$ is isomorphic to $\mathbb{Z}$; we fix a line bundle $S$ which generates it. The KO-theory twisted by $S$ vanishes:

**Theorem 4.13.** For all $n \geq 2$ the additive structure of $KO^*(S_n; \mathcal{L})$ is as follows:

| $n \equiv 2 \mod 4$ | $t_0$ | $t_1$ | $s_i(O)$ | $s_i(S)$ |
|---------------------|-------|-------|-----------|-----------|
| otherwise           | $2^{n-2}$ | $2^{n-2}$ | $\rho \left( \frac{n}{2}, 1 - i \right)$ | $0$ |
|                     | $2^{n-2}$ | $2^{n-2}$ | $\rho \left( \frac{n}{2}, -i \right)$ | $0$ |

Here, the values $\rho(n, i)$ are defined as in Theorem 4.5.

**Proof.** The cohomology of $S_n$ with $\mathbb{Z}/2$-coefficients has simple generators $e_2, e_4, \ldots, e_{2n-2}$, i.e. it is additively generated by products of distinct elements of this list. Its multiplicative structure is determined by the rule $e_2^{2i} = e_{4i}$, and the second Steenrod square is given by $Sq^2(e_{2i}) = i e_{2i+2}$ [Ish92, Proposition 1.1]. In both formulae it is of course understood that $e_j = 0$ for $j \geq n$. What we need to show is that for all $n \geq 2$ we have

$$H^*(S_n, Sq^2 + e_2) = 0$$

Let us abbreviate $H^*(S_n, Sq^2 + e_2)$ to $(H_n, d')$. We claim that we have the following short exact sequence of differential $\mathbb{Z}/2$-modules:

$$0 \rightarrow (H_n, d') \xleftarrow{c_{2n}} (H_{n+1}, d') \rightarrow (H_n, d') \rightarrow 0 \quad (28)$$

This can be checked by a direct calculation. Alternatively, it can be deduced from the geometric considerations below. Namely, it follows from the cofibration sequence of Corollary 4.15 that we have such an exact sequence of $\mathbb{Z}/2$-modules with maps respecting the differentials given by $Sq^2$ on all three modules. Since they also commute with multiplication by $e_2$, they likewise respect the differential $d' = Sq^2 + e_2$.

The long exact cohomology sequence associated with (28) allows us to argue by induction: if $H^*(H_n, d') = 0$ then also $H^*(H_{n+1}, d') = 0$. Since we can see by hand that $H^*(H_2, d') = 0$, this completes the proof.

We close with a geometric interpretation of the exact sequence (28), via an analogue of Lemmas 4.2 and 4.6. Let us write $\mathcal{U}$ for the universal bundle over $S_n$, i.e. for the restriction of the universal bundle over $Gr(n-1,2n-1)$ to $S_n$, and $\mathcal{U}^\perp$ for the restriction of the orthogonal complement bundle, so that $\mathcal{U} \oplus \mathcal{U}^\perp$ is the trivial $(2n-1)$-bundle over $S_n$. As in Section 4.4, we emphasize that under these conventions the fibres of $\mathcal{U}$ and $\mathcal{U}^\perp$ are perpendicular with respect to a hermitian metric on $\mathbb{C}^{2n-1}$ — they are not orthogonal with respect to the chosen symmetric form.
Lemma 4.14. The spinor variety $S_n$ embeds into the spinor variety $S_{n+1}$ with normal bundle $U^\perp$ such that the embedding extends to an embedding of this bundle. The closed complement of $U^\perp$ in $S_{n+1}$ is again isomorphic to $S_n$.

Corollary 4.15. We have a cofibration sequence

$$S_{n+1} \overset{i}{\rightarrow} S_{n+1}^+ \overset{p}{\rightarrow} \text{Thom}_{S_n} U^\perp$$

Note however that, unlike in the symplectic case, the first Chern classes of $U$ and $U^\perp$ pull back to twice a generator of the Picard group of $S_n$. For example, the embedding of $S_2$ into $\text{Gr}(1,3)$ can be identified with the embedding of the one-dimensional smooth quadric into the projective plane, of degree 2, and the higher dimensional cases can be reduced to this example. Thus, $c_1(U)$ and $c_1(U^\perp)$ are trivial in Pic$(S_n)/2$.

Proof of Lemma 4.14. The proof is similar to the proof of Lemma 4.6. Let $e_1, e_2$ be the first two canonical basis vectors of $\mathbb{C}^{2n+1}$, and let $\mathbb{C}^{2n-1}$ be embedded into $\mathbb{C}^{2n+1}$ via the remaining coordinates. Let $S_n$ be defined in terms of a symmetric form $Q$ on $\mathbb{C}^{2n-1}$, and define $S_{n+1}$ in terms of $Q_{2n+1} := \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & Q \end{pmatrix}$.

Let $i_1$ and $i_2$ be the embeddings of $S_n$ into $S_{n+1}$ sending an $(n-1)$-plane $\Lambda \subset \mathbb{C}^{2n-1}$ to $e_1 \oplus \Lambda$ or $e_2 \oplus \Lambda$ in $\mathbb{C}^{2n+1}$, respectively. Given an $(n-1)$-plane $\Lambda \in S_n$ together with a vector $v$ in $\Lambda^\perp \subset \mathbb{C}^{2n-1}$, consider the linear map

$$\begin{pmatrix} -\frac{1}{2}Q(v,v) & -Q(-,v) \\ v & 0 \end{pmatrix} : \langle e_1 \rangle \oplus \Lambda \rightarrow \langle e_2 \rangle \oplus \Lambda^\perp$$

Sending $(\Lambda, v)$ to the graph of this function defines an open embedding of $U^\perp$ whose closed complement is the image of $i_2$.

4.7 Exceptional hermitian symmetric spaces

Lastly, we turn to the exceptional hermitian symmetric spaces EIII and EVII. We write $O(1)$ for a generator of the Picard group in both cases.

Theorem 4.16. The KO-groups of the exceptional hermitian symmetric spaces EIII and EVII are as follows:

| | $t_0$ | $t_1$ | $L \equiv O$ | $L \equiv O(1)$ |
|---|---|---|---|---|
| $s_0$ | 15 | 12 | 3 | 0 |
| $s_1$ | 3 | 0 | 0 | 0 |
| $s_2$ | 0 | 0 | 0 | 0 |
| $s_3$ | 0 | 0 | 0 | 0 |

KO*(EIII; $L$) | 15 | 12 | 3 | 0 | 0 | 0 | 0 | 0 |
KO*(EVII; $L$) | 28 | 28 | 1 | 3 | 3 | 1 | 0 | 0 | 0 | 0 |
Proof. The untwisted KO-groups have been computed in [KH92], the main difficulty as always being to prove that the Atiyah-Hirzebruch spectral sequence collapses. For the twisted groups, however, there are no problems. We quote from § 3 of said paper that the cohomologies of the spaces in question can be written as

\[ H^*(\text{EIII}; \mathbb{Z}/2) = \mathbb{Z}/2[t, u]/(u^2 t, u^3 + t^{12}) \]

\[ H^*(\text{EVII}; \mathbb{Z}/2) = \mathbb{Z}/2[t, v, w]/(t^{14}, v^2, w^2) \]

with \( t \) of degree 2 in both cases, and \( u, v \) and \( w \) of degrees 8, 10 and 18, respectively. The Steenrod squares are determined by \( \text{Sq}^2 u = ut \) and \( \text{Sq}^2 v = \text{Sq}^2 w = 0 \). Thus, we find

\[ H^*(\text{EIII}, \text{Sq}^2 + t) = \mathbb{Z}/2 \cdot u \oplus \mathbb{Z}/2 \cdot u^2 \oplus \mathbb{Z}/2 \cdot u^3 \]

\[ H^*(\text{EVII}, \text{Sq}^2 + t) = 0 \]

By Lemma 3.3 the Atiyah-Hirzebruch spectral sequence for \( \text{EIII} \) must collapse. This gives the result displayed above.

Acknowledgements

The present paper is based on research carried out in the context of the author’s PhD studies, which have been generously funded by the Engineering and Physical Sciences Research Council (EPSRC). I wish to thank my supervisor Burt Totaro for his commitment, time and patience, for directing me towards the problems and ideas discussed here, and for his continuing support. Special thanks are moreover due to Baptiste Calmès and Marco Schlichting for highly helpful conversations and explanations of their respective work. Finally, I am grateful to Jens Hornbostel for constructive comments on an earlier version of this paper, and to the anonymous referee for careful proofreading.

References

[Ada95] Frank Adams, *Stable homotopy and generalised homology*, Chicago Lectures in Mathematics, University of Chicago Press, Chicago, IL, 1995. Reprint of the 1974 original.

[Ara80] Jón Kristinn Arason, *Der Wittring projektiver Räume*, Math. Ann. 253 (1980), no. 3, 205–212.

[ABS64] Michael Atiyah, Raoul Bott, and Arnold Shapiro, *Clifford modules*, Topology 3 (1964), no. suppl. 1, 3–38.

[Bal00] Paul Balmer, *Triangular Witt groups. I. The 12-term localization exact sequence*, K-Theory 19 (2000), no. 4, 311–363.

[Bal01a] ———, *Triangular Witt groups. II. From usual to derived*, Math. Z. 236 (2001), no. 2, 351–382.

[Bal01b] ———, *Witt cohomology, Mayer-Vietoris, homotopy invariance and the Gersten conjecture*, K-Theory 23 (2001), no. 1, 15–30.
Witt Groups of Complex Cellular Varieties

[Bal05] Paul Balmer, Witt groups, Handbook of K-theory. Vol. 1, 2, Springer, Berlin, 2005, pp. 539–576.

[BC09] Paul Balmer and Baptiste Calmès, Geometric description of the connecting homomorphism for Witt groups, Doc. Math. 14 (2009), 525–550.

[BC08] Paul Balmer and Stefan Gille, Koszul complexes and symmetric forms over the punctured affine space, Proc. London Math. Soc. (3) 91 (2005), no. 2, 273–299.

[BGG73] Joseph Bernstein, Israel Gelfand, and Sergei Gelfand, Schubert cells, and the cohomology of the spaces G/P, Uspehi Mat. Nauk 28 (1973), no. 3(171), 3–26 (Russian); English transl., Russian Math. Surveys 28 (1973), no. 3, 1–26.

[Bot69] Raoul Bott, Lectures on K(X), Mathematics Lecture Note Series, W. A. Benjamin, Inc., New York-Amsterdam, 1969.

[BG10] Robert Bruner and John Greenlees, Connective real K-theory of finite groups, Mathematical Surveys and Monographs, vol. 169, American Mathematical Society, Providence, RI, 2010.

[CF11] Baptiste Calmès and Jean Fasel, Trivial Witt groups of flag varieties (2011), available at arxiv.org/abs/1103.4412. Preprint.

[DH91] William Fulton and Joe Harris, Representation theory, Graduate Texts in Mathematics, vol. 129, Springer-Verlag, New York, 1991.

[Dug91] Daniel Dugger, Universal homotopy theories, Adv. Math. 164 (2001), no. 1, 144–176.

[FH91] William Fulton and Joe Harris, Algebraic geometry, Graduate Texts in Mathematics, vol. 133, Springer-Verlag, New York, 1992. A first course.

[Hat92] Allen Hatcher, Algebraic geometry, Graduate Texts in Mathematics, vol. 133, Springer-Verlag, New York, 1992. A first course.

[Hog99] Stuart Hoggard, On KO theory of Grassmannians, Quart. J. Math. Oxford Ser. (2) 20 (1969), 447–463.

[Hus94] Dale Husemoller, Fibre bundles, 3rd ed., Graduate Texts in Mathematics, vol. 20, Springer-Verlag, New York, 1994.

[Ish87] John Frederick Jardine, Simplicial presheaves, J. Pure Appl. Algebra 47 (1987), no. 1, 35–87.

[Jou73] Jean-Pierre Jouanolou, Une suite exacte de Mayer-Vietoris en K-théorie algébrique, Algebraic K-theory, I: Higher K-theories (Proc. Conf., Battelle Memorial Inst., Seattle, Wash., 1972), Springer, Berlin, 1973, pp. 293–316. Lecture Notes in Math., Vol. 341.
Max Karoubi, *Bott periodicity in topological, algebraic and Hermitian $K$-theory*, Handbook of $K$-theory. Vol. 1, 2, Springer, Berlin, 2005, pp. 111–137.

Daisuke Kishimoto, Akira Kono, and Akihiro Ohsita, *KO-theory of flag manifolds*, J. Math. Kyoto Univ. 44 (2004), no. 1, 217–227.

Stanley Kochman, *Bordism, stable homotopy and Adams spectral sequences*, Fields Institute Monographs, vol. 7, American Mathematical Society, Providence, RI, 1996.

Max Karoubi, *KO-theory of complex Grassmannians*, J. Math. Kyoto Univ. 31 (1991), no. 3, 827–833.

Daisuke Kishimoto, Akira Kono, and Akihiro Ohsita, *KO-theory of flag manifolds*, J. Math. Kyoto Univ. 44 (2004), no. 1, 217–227.

Stanley Kochman, *Bordism, stable homotopy and Adams spectral sequences*, Fields Institute Monographs, vol. 7, American Mathematical Society, Providence, RI, 1996.

Daisuke Kishimoto, Akira Kono, and Akihiro Ohsita, *KO-theory of flag manifolds*, J. Math. Kyoto Univ. 44 (2004), no. 1, 217–227.

Stanley Kochman, *Bordism, stable homotopy and Adams spectral sequences*, Fields Institute Monographs, vol. 7, American Mathematical Society, Providence, RI, 1996.

Fabien Morel, *Théorie homotopique des schémas*, Astérisque 256 (1999).

Fabien Morel and Vladimir Voevodsky, *A¹-homotopy theory of schemes*, Inst. Hautes Études Sci. Publ. Math. 90 (1999), 45–143 (2001).

Fabien Morel, *An introduction to A¹-homotopy theory*, Contemporary developments in algebraic K-theory, ICTP Lect. Notes, XV, Abdus Salam Int. Cent. Theoret. Phys., Trieste, 2004, pp. 357–441.

Fabien Morel, *Rational stable splitting of Grassmannians and rational motivic sphere spectrum* (2006), available at www.mathematik.uni-muenchen.de/~morel/preprint.html. Draft.

Vladimir Gorbatsevich, Arkadi˘ı Onishchik, and ‘Ernest Vinberg, *Lie Groups and Lie Algebras III: Structure of Lie Groups and Lie Algebras*, Encyclopaedia of Mathematical Sciences, vol. 41, Springer-Verlag, Berlin, 1994.

Ivan Panin and Charles Walter, *On the motivic commutative ring spectrum BO* (2010), available at www.math.uiuc.edu/K-theory/0978/. Preprint.

Joël Riou, *Opérations sur la K-théorie algébrique et régulateurs via la théorie homotopique des schémas*, Ph. D. Thesis, l’Université Paris 7 – Denis Diderot, 2006, www.math.u-psud.fr/~riou/these/these.pdf.

Joël Riou, *Catégorie homotopique stable d’un site suspendu avec intervalle*, Bull. Soc. Math. France 135 (2007), no. 4, 495–547.

Joël Riou, *Opérations sur la K-théorie algébrique et régulateurs via la théorie homotopique des schémas*, C. R. Math. Acad. Sci. Paris 344 (2007), no. 1, 27–32.

Marco Schlichting, *Hermitian K-theory of exact categories*, J. K-Theory 5 (2010), no. 1, 105–165.

Marco Schlichting, *The Mayer-Vietoris principle for Grothendieck-Witt groups of schemes*, Invent. Math. 179 (2010), no. 2, 349–433.

Marco Schlichting, *Hermitian K-theory, derived equivalences and Karoubi’s fundamental theorem*. In preparation.

Graeme Segal, *Equivariant K-theory*, Inst. Hautes Études Sci. Publ. Math. 34 (1968), 129–151.
Vladimir Voevodsky, $A^1$-homotopy theory, Proceedings of the International Congress of Mathematicians, Vol. I (Berlin, 1998), 1998, pp. 579–604.

Charles Walter, Grothendieck-Witt groups of triangulated categories (2003), available at www.math.uiuc.edu/K-theory/0643/. Preprint.

Charles Walter, Grothendieck-Witt groups of projective bundles (2003), available at www.math.uiuc.edu/K-theory/0644/. Preprint.

Nobuaki Yagita, Coniveau spectral sequence and motivic cohomology of quadrics and classifying spaces (2004), available at www.math.uiuc.edu/K-theory/0709/. Preprint.

Nobuaki Yagita, Note on Witt group and KO-theory of complex Grassmannians (2009), available at www.math.uni-bielefeld.de/LAG/man/325.html. Preprint.

Marcus Zibrowius, Comparing Grothendieck-Witt groups of a complex variety to its real topological $K$-groups (2009), available at www.dspace.cam.ac.uk/handle/1810/236624. Smith-Knight/Rayleigh-Knight Essay.

Marcus Zibrowius
Department of Pure Mathematics and Mathematical Statistics
University of Cambridge
Wilberforce Road
Cambridge
CB3 0WB
United Kingdom
THE HILBERT-CHOW MORPHISM
AND THE INCIDENCE DIVISOR

JOSEPH ROSS

Abstract. For a smooth projective variety $P$ of dimension $n$, we construct a Cartier divisor supported on the incidence locus in the product of Chow varieties $\mathcal{C}_a(P) \times \mathcal{C}_{n-a-1}(P)$. There is a natural definition of the corresponding line bundle on a product of Hilbert schemes, and we show this bundle descends to the Chow varieties. This answers a question posed by Barry Mazur.

2010 Mathematics Subject Classification: Primary 14C05
Keywords and Phrases: Chow variety, Hilbert scheme

Section 1. Introduction

Let $(P, \mathcal{O}_P(1))$ be a smooth projective variety of dimension $n$ over an algebraically closed field $k$. The Chow variety $\mathcal{C}_{d,d'}(P)$ parameterizes algebraic cycles on $P$. In particular, there is a bijection between $\mathcal{C}_{d,d'}(P)(k)$ and the set of effective algebraic cycles on $P$ of dimension $d$ and degree $d'$. (We suppress the degree since it plays no role.) Let $a$ and $b$ be nonnegative integers such that $a + b + 1 = n$. The product $\mathcal{C}_a(P) \times \mathcal{C}_b(P)$ parameterizes pairs $(A, B)$ of an $a$-dimensional cycle $A$ and an $b$-dimensional cycle $B$. Since $a + b$ is less than $n$, one expects a generic $A$ and $B$ to be disjoint. But one expects that the incidence locus $\mathcal{I}$ parameterizing incident pairs is a codimension 1 closed subvariety in $\mathcal{C}_a(P) \times \mathcal{C}_b(P)$.

For $k$ the field of complex numbers, Mazur [14] constructed a Weil divisor supported on the incidence locus as follows. Consider the diagram of schemes:

\[
P \times \mathcal{C}_a(P) \times \mathcal{C}_b(P) \xrightarrow{\Delta} P \times P \times \mathcal{C}_a(P) \times \mathcal{C}_b(P)
\]

\[
\mathcal{C}_a(P) \times \mathcal{C}_b(P)
\]
Let \( U_a, U_b \) denote the universal cycles on \( P \times \mathcal{C}_a(P) \), \( P \times \mathcal{C}_b(P) \) respectively (these exist in characteristic zero). Since \( \Delta \) is a local complete intersection morphism, there is a refined Gysin homomorphism \( \Delta' \), as constructed in [7, 6.2]. Using standard operations in intersection theory, one has \( pr_{23} \Delta'(U_a \boxtimes U_b) \), a cycle of codimension 1 on \( \mathcal{C}_a(P) \times \mathcal{C}_b(P) \). The main question of [14] is whether \( pr_{23} \Delta'(U_a \boxtimes U_b) \) is a Cartier divisor. The main result of this paper is a positive answer to Mazur’s question.

**Theorem 1.1.** Let \( U \subset \mathcal{C}_a(P) \times \mathcal{C}_b(P) \) denote the locus of disjoint cycles, i.e., the complement of the incidence locus \( \mathcal{I} \). Let \( U' \subset U \) denote the union of products \( C_a \times C_b \) of irreducible components \( C_a \subset \mathcal{C}_a(P), C_b \subset \mathcal{C}_b(P) \) over which the universal cycles intersect properly.

- There is a Cartier divisor \( D \) on \( U \) which is supported on \( U - U' \).
- The restriction of \( D \) to \( U' \) is an effective Cartier divisor.
- Let \( T \) be the spectrum of a discrete valuation ring \( R \supset k \), with generic point \( \eta \). Let \( g : T \to U \subset \mathcal{C}_a(P) \times \mathcal{C}_b(P) \) be a morphism corresponding to cycles \( Z, W \) on \( P \times T \), and such that \( g(\eta) \in U \). Let \( s_D \) denote the canonical section of the line bundle \( O_U(D) \). Then we have
  \[
  \text{ord } g^*(s_D) = \deg(Z \cdot W) \in \mathbb{Z},
  \]
  where \( Z \cdot W \in A_0(P \times T) \) is the class constructed in [7, 20.2].

**Remark 1.2.** Our methods will suggest there is a line bundle on the whole of the product \( \mathcal{C}_a(P) \times \mathcal{C}_b(P) \), but it does not seem reasonable to expect a Cartier divisor beyond the locus \( U' \). On the locus \( U' \), the operation \( \Delta' \) is defined on the cycle level, and all of the coefficients appearing in \( pr_{23} \Delta'(U_a \boxtimes U_b) \) are positive. On the locus \( U - U' \), negative coefficients may appear.

**Techniques.** Our approach to Mazur’s question, initiated in [20], is to define the incidence line bundle \( \mathcal{L} \) on a product of Hilbert schemes mapping to the corresponding Chow varieties, and then show \( \mathcal{L} \) is the pullback of a line bundle \( \mathcal{M} \) on the Chow varieties. Our \( \mathcal{L} \) will be equipped with a canonical nonvanishing rational section on the locus of disjoint subschemes, and we will show this section is induced by a trivialization of \( \mathcal{M} \) on \( U \). Briefly, \( \mathcal{L} \) is the determinant of a perfect complex formed from the universal flat families. Then we form a proper hypercovering of \( U \) along the Hilbert-Chow morphism, and a descent datum for \( \mathcal{L} \) on this hypercovering. This amounts to an identification \( \phi \) between two pullbacks of \( \mathcal{L} \), satisfying a cocycle condition.

At first we define the descent datum \( \phi \) over a normal base provided the incidence has the expected dimension (5.18); this boils down to the Serre Tor-formula for intersection multiplicities and basic properties of the determinant functor (additivity on short exact sequences). To extend the descent datum over families with more complicated incidence structure, we establish some moving lemmas to produce local trivializations (5.13, 5.14), then apply Grothendieck-Riemann-Roch to show these local sections glue (5.17, 5.32). A useful tool is the following
result, which characterizes functions on a seminormal scheme (4.3; see also Definition 4.2): a Noetherian ring $A$ is seminormal if and only if every pointwise function on $\text{Spec} \, A$ which varies algebraically along (complete) DVRs is induced by an element of $A$.

As for the effectiveness of $(\mathcal{L}, \phi)$, i.e., that $\mathcal{L}$ is induced by a line bundle on the Chow varieties, an outgrowth of 4.3 is a criterion for effective descent (4.6) applicable to our Hilbert-Chow hypercovering: the bundle $\mathcal{L}$ descends to $\mathcal{M} \in \text{Pic}(\mathcal{U})$ if it can be trivialized locally on $\mathcal{U}$, compatibly with the descent datum $\phi$. The compatible local trivializations are built into the definition of the descent datum.

Motivation. In the classical construction, the Chow variety $C_{d,d'}(\mathbb{P}^n)$ is realized as a closed subvariety of the scheme of Cartier divisors of the Grassmannian $\mathcal{G}$ of $(n-d-1)$-planes: to a $d$-dimensional cycle $Z$ on $\mathbb{P}^n$ we associate the codimension one set of $(n-d-1)$-planes in $\mathbb{P}^n$ which intersect $Z$. Thus the natural ample line bundle on $\mathcal{G} \times \text{CDiv}(\mathcal{G})$ simultaneously shows the projectivity of the Chow variety, and endows the incidence locus (in $C_{d,d'}(\mathbb{P}^n) \times \mathcal{G}$, a special case of the $\mathcal{F}$ considered above) with the structure of an effective Cartier divisor. This generalizes to the case $C_{d}(\mathbb{P}^n) \times C_{n-d-1}(\mathbb{P}^n)$ using the ruled join; see [17].

This direct geometric construction does not extend to general smooth projective $P$. However, the Hilbert scheme (the moduli space for closed subschemes of $P$) and the Hilbert-Chow morphism $\mathcal{H} \to \mathcal{C}$ suggest another approach. The pullback of the line bundle associated to the incidence divisor via $\mathcal{H}(\mathbb{P}^n) \times \mathcal{G} \to \mathcal{C}(\mathbb{P}^n) \times \mathcal{G}$ is the determinant of a perfect complex formed from the universal flat families (see the end of Section 3), and the determinant construction can be defined for any smooth projective $P$. Thus one is naturally led to wonder, for a general pair of Hilbert schemes parameterizing subschemes of dimension $a,b$ as above, whether the determinant line bundle descends to the corresponding product of Chow varieties. The direct geometric construction for $P = \mathbb{P}^n$ and the determinant formula are in fact compatible; see the end of Section 3.

Further motivation comes from the case of zero-cycles and divisors, where the Hilbert-Chow morphism admits a reasonably explicit description. The equality of the families of zero-cycles associated to two families of zero-dimensional subschemes has a natural expression in terms of determinants; and similarly two families of codimension one subschemes determine the same family of cycles if the determinants of their structure sheaves agree. For a detailed study of the determinant bundle in the case of zero-cycles and divisors, in particular the descent to the Chow varieties, see [20].

Contents. In Section 2 we recall background material on determinant functors and $K$-theory. In Section 3 we discuss the relevant properties of the Chow variety and the Hilbert-Chow morphism, and define the incidence line bundle and the Hilbert-Chow hypercovering along which the incidence bundle descends. In Section 4 we explain the role of seminormality both in defining the descent datum and demonstrating its effectiveness. In Section 5 we construct the descent datum and show it is effective.
OTHER WORK. In [24] Wang proved that a certain multiple of the incidence divisor, namely $(n-1)!pr_{23}^!(\Delta(U_a\boxtimes U_b))$, is Cartier by using the Archimedean height pairing on algebraic cycles (over $\mathbb{C}$). (See the references in [24] for history of the height pairing.) Given disjoint cycles $A, B$ on $P$ as above, one has the pairing $\langle A, B \rangle := \int_A [G_B]$ defined by integrating a normalized Green’s current for $B$ over $A$. Wang views $\langle A, B \rangle$ as a function on the open set $U$ in $\mathcal{C}_a(P) \times \mathcal{C}_b(P)$ consisting of disjoint cycles, and by studying the behavior of the function as the cycles collide, obtains [24, Thm. 1.1.2] a metrized line bundle $L$ on $U$ and a rational section $s$ that is regular and nowhere zero on $U$, such that $\log \|s(A, B)\|^2 = (\dim(P) - 1)!\langle A, B \rangle$. Using relative fundamental classes in Deligne cohomology (and again over $\mathbb{C}$), Barlet and Kaddar [4] constructed an incidence Cartier divisor in the analytic setting under the assumption that the incidence is generically finite over the parameter space. It would be interesting to “go back” from the Chow varieties to the height pairing.

CONVENTIONS. We use the definition of the Chow variety from [13]. In characteristic 0, there is a functor of effective algebraic cycles (of dimension $d$ and degree $d'$) defined on the category of seminormal $k$-schemes; and this functor is represented by a seminormal, projective $k$-scheme $\mathcal{C}_{d, d'}(P)$ [13, I.3.21]. In characteristic $p > 0$, there are several plausible notions of a family of effective algebraic cycles, stemming from the ambiguity of the field of definition of a cycle [13, I.4.11]. In this case we work with the seminormal, projective $k$-scheme $\mathcal{C}_{d, d'}(P)$ constructed in [13, I.4.13]. This coarsely represents at least two reasonable functors of effective algebraic cycles. Since we rely on the method of “seminormal descent,” our methods do not apply to other definitions of Chow varieties (e.g., those of Barlet [3], Angéniol [1], and Rydh [21]) when those constructions yield Chow varieties/schemes which are not seminormal. All schemes considered in this paper are locally Noetherian. A variety over a field $k$ is an integral separated scheme of finite type over $k$.

ACKNOWLEDGMENTS. This paper is derived from and improves upon the author’s PhD thesis. The author thanks his thesis advisor Aise Johan de Jong. This work was completed while the author was a wissenschaftlicher Mitarbeiter at the Universität Duisburg-Essen.

SECTION 2. DETERMINANT FUNCTORS AND $K$-GROUPS

A determinant functor assigns an invertible sheaf to each perfect complex. We discuss this notion following [12]. Then we quickly review some background material on $K$-groups.

NOTATION 2.1. Let $X$ be a scheme. Let $D(X)$ denote the derived category of the abelian category $\text{Mod}(X)$ of $\mathcal{O}_X$-modules. Denote by $D^+(X)$, respectively $D^-(X)$, the full subcategory of $D(X)$ whose objects are complexes of $\mathcal{O}_X$-modules which are bounded below, respectively bounded above. Denote by $D^b(X)$ the full subcategory whose objects are complexes which are both bounded below and bounded above.
We denote by $D_{(q)coh}(X)$ the full triangulated subcategory of $D(X)$ consisting of pseudo-(quasi)coherent complexes, and by $D^*_{(q)coh}(X)$ the corresponding bounded category for $* = +, -, b [10, 1.4 \text{ p.38}, 2.1 \text{ p.85}].$ We denote by $\text{Parf}(X) \subset D^b(X)$ the full triangulated subcategory consisting of perfect complexes $[9, \text{ Exp. I Def. 4.7 p.44}].$

Let $\text{Parf-is}(X)$ denote the category whose objects are perfect complexes on $X$, with morphisms isomorphisms in $D(X)$. Let $\text{Pic}(X)$ denote the category whose objects are invertible sheaves on $X$, and whose morphisms are isomorphisms. This is a Picard category in the sense of $[2, \text{ Exp. XVIII Def. 1.4.2}].$

**Determinants.** The main result of $[12, \text{ Thm. 2}]$ is that there exists up to canonical isomorphism a unique determinant functor $\det_X : \text{Parf-is}(X) \to \text{Pic}(X)$ extending the usual determinant (top exterior power) of a locally free sheaf. Indeed the idea is to locally replace a perfect complex by a bounded complex of locally free sheaves, and take the signed tensor product of the usual determinants of the locally free terms. Then one shows this patches to give a global invertible sheaf. For every true triangle of complexes $0 \to F_1 \xrightarrow{\alpha} F_2 \xrightarrow{\beta} F_3 \to 0$ in $\text{Parf-is}(X)$, we require an isomorphism $i_X(\alpha, \beta) : \det(F_1) \otimes \det(F_3) \cong \det(F_2)$, and the isomorphisms $i$ (extending the obvious $i$ for short exact sequences of locally free sheaves) are required to be compatible with isomorphisms of triangles, and more generally triangles of triangles. Associated to morphisms of schemes we have base change isomorphisms interchanging the determinant with pullback, and these are required to be compatible with composition of morphisms of schemes.

**Remark 2.2.** When $X$ is reduced, $i$ extends to the class of distinguished triangles, is functorial over isomorphisms (in $D(X)$) of distinguished triangles, and is compatible with distinguished triangles of distinguished triangles $[12, \text{ Prop. 7}].$

**Associated Cartier divisors.** If $F \in \text{Parf}(X)$ is acyclic at every $x \in X$ of depth $0$, then $[12, \text{ Ch.II}]$ constructs a Cartier divisor $\text{Div}(F)$ on $X$ and a canonical isomorphism $\mathcal{O}_X(\text{Div}(F)) \cong \det_X(F)$ extending the trivialization $\mathcal{O}_{X,x} \cong \det_x F$ at $x \in X$ of depth $0$. The formation of this divisor and isomorphism is additive on short exact sequences $[12, \text{ Thm.3(ii)}]$, and is compatible with base change $f : X' \to X$ such that $LF^*(F)$ is acyclic at every $x' \in X'$ of depth $0$ $[12, \text{ Thm.3(v)}]$. Furthermore, in case $X$ is normal and $x \in X$ is a point of depth $1$, the coefficient of $\{x\}$ in the Weil divisor associated to $\text{Div}(F)$ is the alternating sum of the lengths (at $x$) of the cohomology sheaves, i.e., $\sum (-1)^i \ell_x(\mathcal{H}^i(F)).$ This construction is also studied in $[6, \text{ Sect. 3}]$ and $[15, \text{ Ch. 5 Sect. 3}].$

We mention two further properties implicit in $[12, \text{ Thm.3}].$ If the complex $F$ is acyclic, then $\text{Div}(F) = 0$ and the canonical isomorphism $\mathcal{O}_X(\text{Div}(F)) = \mathcal{O}_X \cong \det_X(F)$ is the trivialization of the determinant of an acyclic complex $[12, \text{ Lemma 2}].$ Finally, the construction is determined by the quasi-isomorphism class of the perfect complex $F$: since all filtration levels and subquotients.
appearing in the canonical filtration ("good truncation") of \( F \) are generically acyclic so long as \( F \) is, the additivity implies Div(\( F \)) = \( \sum (-1)^{i} \text{Div}(H^i(\mathcal{F})) \).

**K-groups.** Let \( X \) be a variety. Then \( K_0(X)\) is the Grothendieck group of \( X \), generated by coherent sheaves on \( X \) with relations for short exact sequences of sheaves; \( K^0(X) \) is the Grothendieck group of vector bundles. When \( X \) is regular, we have an isomorphism \( K_0(X) \cong K^0(X) \). We have also the Chern character \( \text{ch} : K_0(X) \to A_*(X) \), where \( A_*(X) \) is the Chow group of cycles on \( X \), graded by dimension. We note \( F \in \text{Perf}(X) \) determines a class in \( K_0(X) \) since for any abelian category \( \mathcal{A} \) (e.g., \( \text{Coh}(X) \)), \( K_0(\mathcal{A}) \cong K_0(D^b(\mathcal{A})) \); the latter group is generated by objects of the triangulated category \( D^b(\mathcal{A}) \) with relations for distinguished triangles.

The group \( K_0(X) \) has a topological filtration: the subgroup \( F \in \text{Coh}(X) \) is generated by those \( F \in \text{Coh}(X) \) such that \( \dim(\text{Supp}(F)) \leq k \). For a proper morphism of schemes \( f : X \to T \) we obtain a homomorphism \( f_* : K_0(X) \to K_0(T) \) sending (the class of) a coherent sheaf to the alternating sum of (the classes of) its higher direct image sheaves. This preserves the topological filtration. If \( T \) is a point and \( F \in \text{Coh}(X) \), then \( \chi(F) = \text{ch}(f_*(F)) \).

**Section 3. The Hilbert-Chow morphism and the incidence divisor**

In this section we define the Chow variety, the Hilbert-Chow morphism, and construct our proper hypercovering. Then we define the incidence line bundle on the product of Hilbert schemes.

We recall an application of the characterization of seminormal schemes [19, 5.1], where it is shown that properties (1)-(5) below characterize the Chow variety. For properties (6) and (7) we refer to [13].

**Definition-Theorem 3.1 (Existence of the Chow variety).** Let \( P \) be a smooth projective variety over a field \( k \). The Chow variety \( \mathcal{C}_{d,d'} \) of \( P \) is a \( k \)-scheme with the following properties:

1. It is projective over \( k \).
2. It is seminormal.
3. For every point \( w \in \mathcal{C}_{d,d'} \) there exist purely inseparable field extensions \( \kappa(w) \subset L_i \) and cycles \( Z_i \) on \( P_{L_i} \) such that:
   a. \( Z_i \) and \( Z_j \) are essentially equivalent [13, I.3.8]: they agree as cycles over the perfection \( \kappa(w)^{\text{perf}} \) of \( \kappa(w) \);
   b. the intersection of the fields \( L_i \) is \( \kappa(w) \), which is the Chow field (field of definition of the Chow form in any projective embedding of \( P \)) of any of the \( Z_i \) [13, I.3.24.1]; and
   c. for any cycle \( Z \) on \( P_{M} \) defined over a subfield \( k \subset M \subset \kappa(w)^{\text{perf}} \) which agrees with the \( Z_i \) over \( \kappa(w)^{\text{perf}} \) (equivalently, agrees with one \( Z_i \)), we have \( \kappa(w) \subset M \) (the Chow field is the intersection of all fields of definition of the cycle).
4. Points \( w \) of \( \mathcal{C}_{d,d'} \) are in bijective correspondence with systems \( (k \subset \kappa(w), \{Z_i \}_{i \in I}) \) up to an obvious equivalence relation.
(5) For any DVR $R ⊃ k$ and any cycle $Z$ on $P_R$ of relative dimension $d$ and degree $d'$ in the generic fiber, we obtain a morphism $g : \text{Spec } R → \mathcal{C}_{d,d'}$ such that the generic fiber $Z_0$ and the special fiber $Z_s$ agree with the systems of cycles of the previous property at $g(η)$ and $g(s)$.

(6) For any numerical polynomial $q$ of degree $d$ and with leading coefficient $d'/d!$, we obtain a morphism (the Hilbert-Chow morphism)

$$FC : (\mathcal{C}_{q})_{\text{red}}^{\text{sn}} → \mathcal{C}_{d,d'}$$

by taking the fundamental cycle of the components of maximal relative dimension (= $d$) [13, I.6.3.1]. A finite number of $(\mathcal{C}_{q})_{\text{red}}^{\text{sn}}$’s surject onto $\mathcal{C}_{d,d'}$.

(7) Let $η ∈ \mathcal{C}_{d,d'}$ be a generic point. Then either $\dim(η) = 0$ or there exists a cycle $Z_0$ on $P_0$ defined over $κ(η)$. In particular, if $k$ is perfect then there exists a $Z_η$ for every generic point $η$ of $\mathcal{C}_{d,d'}$ [13, I.4.14].

Construction 3.2. Let $P$ be a smooth projective variety, and $r ∈ \mathbb{Z}_{≥ 0}$. Let $\mathcal{H}_r$ denote the Hilbert scheme of $r$-dimensional subschemes of $P$. Let $\mathcal{H}_r'$ denote the seminormalization of the (closed) subscheme of $\mathcal{H}_r'$ consisting of subschemes $Z$ such that $Z$ has pure $r$-dimensional support (this is different from the notion of a pure sheaf: $Z$ may have embedded components of smaller dimension so long as they are set-theoretically contained in the top-dimensional components). We have the product of the Hilbert-Chow morphisms (3.1) $π : Y_0 = \mathcal{H}_0 × \mathcal{H}_r → \mathcal{C}_0 × \mathcal{C}_r =: C$. Because seminormalization is a functor, we may form a proper hypercovering $π_∗ : Y_∗ → C$ augmented towards $C$ whose $i$-th term $Y_i$ is the seminormalization of $Y_0 × C \ldots × C Y_0 (i + 1 \text{ factors})$, with the (seminormalizations of the) canonical morphisms.

Remarks 3.3. (3.3.1) We explain property (7) in more detail. For any positive-dimensional component of $\mathcal{C}_{d,d'}$, its generic point corresponds to a cycle all of whose coefficients are 1, i.e., a subscheme [13, I.4.14]. Hence we can find a (component of some) $(\mathcal{C}_{q})_{\text{red}}^{\text{sn}}$ admitting a birational morphism to that component.

(3.3.2) Over a field of characteristic zero, the seminormality of the Chow variety and [19, 4.1] imply $\mathcal{O}_η = \pi_∗(\mathcal{O}_{X_∗})$ for a proper hypercovering $X_*$ augmented towards $C$. In characteristic $p > 0$, we have the characterization of the residue fields on the Chow variety as the intersection of all fields of definition [13, I.4.5]. So by [19, 4.1 Corrigendum], we have $\mathcal{O}_η = \pi_∗(\mathcal{O}_{X_∗})$ for a proper hypercovering such that $X_0 = \mathcal{H}_{r,\text{sn}}$.

In more detail and in the language of [19, 4.1], we explain how to construct (locally) a pointwise function on $C$ from a pointwise function on $\mathcal{H}_{r,\text{sn}}$ which belongs to $π_∗(\mathcal{O}_{X_∗})$. So suppose $z ∈ C(P_k)$ corresponds via a morphism $\text{Spec } κ(z) → C(P_k)$ to the cycle $Z$ on $P_κ(z)$. Consider an algebraically closed field $K$ containing $κ(z)$ and the cycle associated to the base change $Z_{\overline{K}}$. Then by [13, I.4.5], the residue field $κ(z)$ is characterized as the intersection in $K$ of all fields of definition of $Z$, i.e., the intersection of all $E_i$ such that $k ⊂ E_i ⊂ K$ and there exists a subscheme $Y_i ⊂ P_{E_i}$ whose associated cycle agrees with $Z_{\overline{K}}$ upon base change. Consider fields $E_0, E_1$ satisfying these conditions. Then

Documenta Mathematica 16 (2011) 513–543
we have morphisms $\text{Spec } E_i \to \mathcal{H}$ with the property that the compositions $\text{Spec } K \to \text{Spec } E_i \to \mathcal{H} \to \mathcal{C}$ are the same. Thus we have a commutative diagram:

$$
\begin{array}{c}
\text{Spec } K \\
\downarrow \\
\text{Spec } E_0, \text{Spec } E_1 \\
\downarrow \\
\mathcal{H}
\end{array}
\xrightarrow{p_0} \xrightarrow{p_1} \xrightarrow{(\mathcal{H} \times \mathcal{H})_{\text{red}}^{\text{sn}}} \text{Spec } E_i
$$

(The morphism from $\text{Spec } K$ factors through the seminormalization.) Considering $a \in \pi_{\bullet}^*(\mathcal{O}_{X_\bullet})$ as a pointwise function, we obtain elements $a_i \in E_i$. The preceding diagram shows $a_0 = a_1$ in $K$, therefore $a_0 \in E_0 \cap E_1$. By the same argument we find $a_0 \in E_0 \cap E_i$ for all $i$, therefore $a_0 \in \kappa(z)$. Thus we made an element in the residue field $\kappa(z)$. It varies algebraically along DVRs by [19, 4.1].

(3.3.3) If $X = Y \cup Z$ is a reducible scheme with irreducible components $Y, Z$, then the field of definition of $Y$ is contained in the field of definition of $X$. Also, a scheme and its seminormalization have the same residue fields. Hence to cut out the residue fields on $\mathcal{C}$, it is enough to consider the subscheme $\mathcal{H}_r \hookrightarrow \mathcal{H}'$ defined in 3.2. So we have $\mathcal{O}_\mathcal{C} = \pi_{\bullet}^*(\mathcal{O}_{X_\bullet})$ for a proper hypercovering such that $X_0 = \mathcal{H}_r$.

We record the (presumably known) fact that the Hilbert-Chow morphism is compatible with products.

**Lemma 3.4.** If $P, P'$ are smooth projective varieties over a field $k$, then the following diagram commutes. (We suppose $p$ has leading coefficient $d'/(d!)$ and $\deg(p) = d$; and $q$ has leading coefficient $e'/(e!)$ and $\deg(q) = e$.)

$$
\begin{array}{c}
(\mathcal{H}^p(P))^{\text{sn}}_{\text{red}} \times (\mathcal{H}^q(P'))^{\text{sn}}_{\text{red}} \\
\downarrow_{\mathcal{C} \times \mathcal{C}} \\
(\mathcal{H}^{pq}(P \times P'))^{\text{sn}}_{\text{red}}
\end{array}
\xrightarrow{\mathcal{C} \times \mathcal{C}}
\begin{array}{c}
\mathcal{C}_{d,d'}(P) \times \mathcal{C}_{e,e'}(P') \\
\downarrow_{\mathcal{C}} \\
\mathcal{C}_{d+e,d'+e'}(P \times P')
\end{array}
$$

**Proof.** We describe the map in the top row: if $Z \hookrightarrow P \times T, Z' \hookrightarrow P' \times T$ constitute a $T$-point of $(\mathcal{H}^{pq}(P))^{\text{sn}}_{\text{red}} \times (\mathcal{H}^{pq}(P'))^{\text{sn}}_{\text{red}}$, then the scheme theoretic intersection $pr_{13}^*Z \cap pr_{13}^*Z'$ in $P \times P' \times T$ is a $T$-point of $(\mathcal{H}^{pq}(P \times P'))^{\text{sn}}_{\text{red}}$. A top-dimensional component in the product scheme is induced by a pair of top-dimensional components; and length multiplies, so the coefficients in the product cycle are the products of the coefficients of the factors. □

The main goal of this paper is to construct a Cartier divisor supported on the incidence locus. Now we define an invertible sheaf (the “incidence bundle”) on a product of Hilbert schemes, and show the incidence bundle is pulled back from the product of Chow varieties in the case $P = \mathbb{P}^n$. 

---

**Documenta Mathematica 16 (2011) 513–543**
Let $B$ be a base scheme, and $P$ a smooth projective $B$-scheme. Let $\mathcal{H}^1, \mathcal{H}^2$ denote the Hilbert schemes corresponding to numerical polynomials $q_1, q_2$, and set $\mathcal{H}^{1,2} := \mathcal{H}^1 \times_B \mathcal{H}^2$.

Over each $\mathcal{H}^i$ we have a universal flat family, a closed subscheme of $P \times_B \mathcal{H}^i$. Standard facts about the behavior of perfect complexes under certain operations (stability under tensor product; pullback; and pushfoward via a proper morphism of finite Tor-dimension; and the perfectness of a coherent sheaf on the source of a smooth morphism which is flat over the target) imply $\text{Rpr}_{23}^*(\mathcal{O}_{\mathcal{H}_1} \otimes \mathcal{O}_{\mathcal{H}_2})$ is a perfect complex on $\mathcal{H}^{1,2}$.

For details on the necessary facts about perfect complexes, see section 2 of \[20\].

The incidence bundle $L$ is defined to be its determinant:

$$L := \det_{\mathcal{H}^{1,2}} \text{Rpr}_{23}^*(\mathcal{O}_{\mathcal{H}_1} \otimes \mathcal{O}_{\mathcal{H}_2}).$$

In fact we will be interested in this construction only on the locus $Y_0$ defined earlier in this section. Furthermore, since the complex $\text{Rpr}_{23}^*(\mathcal{O}_{\mathcal{H}_1} \otimes \mathcal{O}_{\mathcal{H}_2})$ is acyclic over the locus $U_0$ of disjoint pairs of subschemes, this construction determines a Cartier divisor on the closure of the locus of disjoint pairs of subschemes (the Hilbert scheme analogue of the locus $U$ defined in 1.1). See \[20, 2.5\].

As motivation for pursuing the determinant formula (mentioned in the Introduction), we make contact with the classical construction of the Chow variety of $P = \mathbb{P}^n$. As explained in the Introduction, the construction of the Chow variety endows the incidence locus $\mathcal{I} \hookrightarrow \mathcal{C}(\mathbb{P}^n) \times \mathcal{G}$ with the structure of a Cartier divisor. Let $FC_{\mathbb{P}^n} : \mathcal{H}(\mathbb{P}^n) \to \mathcal{C}(\mathbb{P}^n)$ denote the Hilbert-Chow morphism (and its product with $\mathcal{G}$). In the special case of Construction 3.5 with $P = \mathbb{P}^n$ and $\mathcal{H}^2 = \mathcal{G}$, it follows from [5, Thms. 1.2, 1.4] that there is a canonical isomorphism $L \cong FC_{\mathbb{P}^n}^*(\mathcal{O}(\mathcal{I}))$ of invertible sheaves on $\mathcal{H} \times \mathcal{G}$. This isomorphism is canonical in the following sense. Over the locus $U_0$ of disjoint subschemes, there is a canonical trivialization $L|_{U_0} \cong \mathcal{O}_{U_0}$. This rational section is the pullback via $FC_{\mathbb{P}^n}$ of the canonical trivialization of $\mathcal{O}(\mathcal{I})$ on the complement of $\mathcal{I}$.

Section 4. Seminormal schemes and descent criteria

In this section we explain the role of seminormality both in defining the descent datum and demonstrating its effectiveness.

Definition 4.1 ([8]). A ring $A$ is a Mori ring if it is reduced and its integral closure $A^\nu$ (in its total quotient ring $Q$) is finite over it; if $A$ is a Mori ring, $A^{\text{sn}}$ denotes its seminormalization, the largest subring $A \subset A^{\text{sn}} \subset A^\nu$ such that $\text{Spec } A^{\text{sn}} \to \text{Spec } A$ is bijective and all maps on residue fields are isomorphisms. The seminormalization is described elementwise in [23, 1.1]. We say $A$ is seminormal if $A = A^{\text{sn}}$ (so we only define seminormality for Mori rings). A locally Noetherian scheme $X$ is Mori if and only if it has an affine cover by Noetherian Mori rings [8, Def. 3.1].
Definition 4.2. Let $A$ be a ring, and let $S = \{ f_y \in \kappa(y) | y \in \text{Spec} A \}$ be a collection of elements, one in each residue field. Then we say $S$ is a pointwise function on $\text{Spec} A$. We say the pointwise function $S$ varies algebraically along (complete) DVRs if it has the following property: for every specialization $p_1 \subset p_2$ in $A$ and every (complete) discrete valuation ring $R$ covering that specialization via a ring homomorphism $g : A \to R$, there exists a (necessarily) unique $f_R \in R$ such that $g_{p_1}(f_{p_1}) = f_R$ (in $K$) and $g_{p_2}(f_{p_2}) = f_R$ (in $k_0$).

The main result of [19, 2.2, 2.6] is the following.

Theorem 4.3. Let $A$ be a seminormal (in particular, Mori) ring which is Noetherian. Let $\{ f_y \in \kappa(y) | y \in \text{Spec} A \}$ be a pointwise function on $\text{Spec} A$ which varies algebraically along (complete) DVRs. Then there exists a unique $f \in A$ whose image in $\kappa(y)$ is $f_y$ for all $y \in \text{Spec} A$.

This simplifies greatly the problem of defining a descent datum for a line bundle on a seminormal scheme, as seen in the following corollary.

Corollary 4.4. Let $X$ be a seminormal locally Noetherian (in particular, Mori) scheme, and let $L, M \in \text{Pic}(X)$. Then an isomorphism $L \cong M$ is equivalent to an “identification of fibers varying algebraically along DVRs,” that is:

for any field or (complete) DVR $R$, any Spec $R \to X$, an identification $\beta_f : f^* L \cong f^* M$ compatible with restriction to the closed and generic points: if $s : \text{Spec} R \to \text{Spec} R$, $\eta : \text{Spec} R \to \text{Spec} R$ denote the inclusions, then $\beta_{f_1} = s^* \beta_f$ and $\beta_{f_2} = \eta^* \beta_f$.

Proof. Fix an open cover $X = \cup_i \text{Spec} S_i$ with $S_i$ a seminormal (Noetherian and Mori) ring which trivializes both $L$ and $M$, and fix trivializations $\varphi_i : L_i := L|_{\text{Spec} S_i} \cong O_{\text{Spec} S_i}$, $\psi_i : M_i := M|_{\text{Spec} S_i} \cong O_{\text{Spec} S_i}$. Then defining $L \cong M$ is equivalent to identifying $\Gamma(\text{Spec} S_i, L_i) \cong \Gamma(\text{Spec} S_i, M_i)$ as $S_i$-modules (for all $i$), compatibly with restrictions. Then considering the diagram:

$$
\begin{array}{ccc}
\Gamma(L_i) & \cong & \Gamma(M_i) \\
\varphi_i & & \psi_i \\
S_i & \to & S_i \\
\end{array}
$$

and its pullbacks to spectra of fields and DVRs, we see that relative to the fixed $\varphi_i, \psi_i$, a family $\beta_f$ as in the statement is equivalent to an invertible pointwise function on each $S_i$ varying algebraically along DVRs. By Theorem 4.3 this is equivalent to a family of elements $f_i \in S_i^\times = \text{Isom}_{S_i}(S_i, S_i)$. The $f_i$ thus obtained agree on overlaps by the uniqueness statement in Theorem 4.3. Then using the above diagram again we see that relative to the fixed trivializations, the family $f_i$ is equivalent to a family of isomorphisms $\Gamma(L_i) \cong \Gamma(M_i)$ compatible with restrictions. \qed

We will need the following general fact later.
**Lemma 4.5.** Let $X$ be a seminormal $k$-scheme and $Y \hookrightarrow X$ a closed subscheme. Suppose every $y \in Y$ admits a generization to a point in $X - Y$, i.e., that $Y$ does not contain any generic points of $X$. Let $S$ be a pointwise function on $X$ which varies algebraically along DVRs covering specializations within $X - Y$, and along those from $X - Y$ to $Y$. Then $S$ varies algebraically along DVRs.

**Proof.** This follows readily from the techniques used in [19]. We may assume $X$ is affine. Let $\nu : X^\nu \to X$ denote the normalization. Then our pointwise function $S$ determines a pointwise function on $X^\nu$ which is constant along the fibers of $\nu$. The normalization is birational, so identifies generic points of $X^\nu$ with those of $X$. Hence as a pointwise function on $X^\nu$, $S$ varies algebraically along specializations of the form $\eta \leadsto x$, with $\eta$ generic. This is enough to conclude $S$ is induced by an element of $\Gamma(X^\nu, \mathcal{O}_{X^\nu})$ (see [19, 2.4]). But then because $S$ is constant along the fibers of $\nu$, this element comes from $\Gamma(X, \mathcal{O}_X)$ and a fortiori varies algebraically along all DVRs. $\square$

As for the effectiveness of a descent datum, we recall the following result from [20]. Let $\pi : X_\bullet \to X$ be a proper hypercovering augmented towards a scheme $X$. We denote by $(\mathcal{L}, \phi)$ an element of $\text{Pic}(X_\bullet)$, i.e., $\mathcal{L}$ is an element of $\text{Pic}(X_0)$ and $\phi : p_0^* \mathcal{L} \sim \to p_1^* \mathcal{L}$ is an isomorphism on $X_1$ satisfying the cocycle condition on $X_2$. As in [20, 3.3], we say $(\mathcal{L}, \phi) \in \text{Pic}(X_\bullet)$ is Zariski locally effective if for every $x \in X$, there exists an open $U \subset X$ containing $x$ and a trivialization $T_x : \mathcal{L}|_{\pi_0^{-1}(U)} \sim \to \mathcal{O}_{\pi_0^{-1}(U)}$ compatible with $\phi$ in the sense that the diagram

$$
p_0^*(\mathcal{L}|_{\pi_0^{-1}(U)}) \xrightarrow{p_0^*T} \mathcal{O}_{(\pi_0)^{-1}(\pi_0^{-1}(U))} \downarrow \phi = \downarrow \phi \xrightarrow{p_1^*T} \mathcal{O}_{(\pi_1)^{-1}(\pi_0^{-1}(U))}
$$

commutes.

**Proposition 4.6.** [20, 3.4] Let $X$ be a scheme, and let $\pi : X_\bullet \to X$ be a proper hypercovering augmented towards $X$ which satisfies $\mathcal{O}_X = \pi_\bullet*(\mathcal{O}_{X_\bullet})$. Then:

- $\pi^\bullet : \text{Pic}(X) \to \text{Pic}(X_\bullet)$ is injective; and
- the image of $\pi^\bullet$ consists of those $(\mathcal{L}, \phi)$ that are Zariski locally effective.

** Remark 4.7.** The Proposition applies when $X$ is seminormal and $X_\bullet$ satisfies any of the conditions in [19, 4.1 Corrigendum], for example the proper hypercovering $\pi : Y_\bullet \to C$ defined in 3.2.

**Section 5. Definition of the descent datum**

In this section we prove the main result, in the following form. Having established this result, we will consider the refinements and further properties stated in 1.1.
Theorem 5.1 (5.36, 5.37). With the notation as in 2.2, let \( U \hookrightarrow P \times \mathcal{H}_r \) denote the (pullback of the) universal flat family. Using 3.5 we may form the determinant line bundle \( L \) on \( Y_0 \). Now base change everything to \( \overline{U} \subset C \), the closure of the locus of disjoint cycles. Then the following hold.

- The sheaf \( L \) lifts to an invertible sheaf on \( Y_* \), i.e., there is an isomorphism \( \phi : p_0^* \mathcal{H} \cong p_1^* \mathcal{L} \) on \( Y_1 \) satisfying the cocycle condition on \( Y_2 \).
- The descent datum \( \phi \) is effective: there is a unique \( \mathcal{M} \in \text{Pic}(U) \) such that \((\pi^* \mathcal{M}, \text{can}) \cong (\mathcal{L}, \phi)\).

Subsection 5.1. Notation and preliminary reductions.

Definition 5.2. Let \( P \) be a smooth projective \( k \)-variety of dimension \( n \). A Hilbert datum for \( P \) over \( T \) consists of the following:

1. a seminormal \( k \)-scheme \( T \);
2. \( Z \mapsto P_T := P \times_k T \) a \( T \)-flat closed subscheme of relative dimension \( a \), such that the support of \( Z \) has pure dimension \( a \) in every fiber; and
3. \( W \mapsto P_T \) a \( T \)-flat closed subscheme of relative dimension \( b \), such that the support of \( W \) has pure dimension \( b \) in every fiber;

such that \( a + b + 1 \leq n \) and every point \( t \in T \) admits a generalization to the locus of disjoint subschemes. Thus a Hilbert datum \((Z, W)\) is simply a morphism \( T \to \mathcal{H}_a \times \mathcal{H}_b \) such that the image of every generic point of \( T \) lies in a component of \( \mathcal{H}_a \times \mathcal{H}_b \) with at least one pair of disjoint subschemes. Typically we will make some construction from \((Z, W)\) and then show the construction only depends on \([Z], [W]\), the cycles underlying \( Z \) and \( W \). Therefore we make the following definition. A Hilbert-Chow datum for \( P \) over \( T \) is a pair of Hilbert data \((Z, W), (Z', W')\) for \( P \) over \( T \) such that \([Z] = [Z']\) and \([W] = [W']\). Since the supports of \( Z, W \) are assumed pure-dimensional, we have also \( \text{Supp}(Z) = \text{Supp}(Z') \) and \( \text{Supp}(W) = \text{Supp}(W') \). Thus a Hilbert-Chow datum \((Z, Z', W, W')\) for \( P \) over \( T \) is nothing more than a morphism \( T \to (\mathcal{H}_a \times \mathcal{H}_b \times \mathcal{E}_a \times \mathcal{E}_b, \mathcal{H}_a \times \mathcal{H}_b)^{\alpha} \) such that (after projecting to either \( \mathcal{H}_a \times \mathcal{H}_b \) factor) every generic point of \( T \) lands in a pair of irreducible components with at least one pair of disjoint subschemes. Because we work on the subscheme \( \mathcal{H}_r \) of the Hilbert scheme, disjointness of subschemes on \( \mathcal{H}_a \times \mathcal{H}_b \) corresponds exactly to disjointness of their associated cycles on \( \mathcal{E}_a \times \mathcal{E}_b \). In general, two subschemes could have disjoint associated cycles but lower-dimensional components which coincide. So in the notation above we have \( Z \cap W = \emptyset \) if and only if \( Z' \cap W = \emptyset \).

Note that given a morphism \( S \to T \) of seminormal \( k \)-schemes and a Hilbert-Chow datum for \( P \) over \( T \), by pullback we obtain a Hilbert-Chow datum for \( P \) over \( S \).

Notation 5.3. The structure morphism \( P_T \to T \) will be called \( \pi \).

For \( F, G \in \text{Parf}(P_T) \), we set \( f_T(F, G) := \det_T R\pi_*(F \otimes^L G) \in \text{Pic}(T) \).

If \( \alpha \) is a \( b \)-dimensional cycle on \( P_T \) with \( \alpha = \sum a_i W_i \), we put \( f_T(\mathcal{O}_Z, \alpha) := \otimes_i (f_T(\mathcal{O}_Z, \mathcal{O}_{W_i})^{\otimes a_i}) \). In general we use the notation \([\cdot] \) to denote the cycle.
associated to a subscheme or coherent sheaf: this means the top-dimensional components and their geometric multiplicities, even if, for example, $b < n - a - 1$. (In fact we used this in the preceding definition.)

If $T$ is affine and equal to $\text{Spec} R$, we may write $f_R$ for $f_T$.

We will use the subscripts $(-)_0$ and $(-)_\eta$ to denote the base change of some object to closed and generic fibers, respectively.

By the incidence $Z \cap W$, we mean the underlying reduced algebraic subset $\text{Supp}(Z) \cap \text{Supp}(W)$. Stated properties of $Z \cap W$ will depend only on the underlying supports $\text{Supp}(Z)$, $\text{Supp}(W)$.

**Goal.** For every Hilbert-Chow datum, we aim to construct an isomorphism $\phi_{Z,Z',W,W'}: f_T(\mathcal{O}_Z, \mathcal{O}_W) \cong f_T(\mathcal{O}_{Z'}, \mathcal{O}_{W'})$ varying functorially in $T$, in such a way that the resulting descent datum $\{\phi_T\}$ is Zariski locally effective. The essential case is $b = n - a - 1$.

**Proposition 5.4 (reduction to fields and DVRs).** To define an isomorphism $\phi_T: f_T(\mathcal{O}_Z, \mathcal{O}_W) \cong f_T(\mathcal{O}_{Z'}, \mathcal{O}_{W'})$ for each Hilbert-Chow datum, for all smooth projective $P$, so that for each $P$, the collection $\{\phi_T\}$

1. is compatible with base change $S \to T$; and
2. satisfies the cocycle condition;

it is sufficient to define an isomorphism $\phi_T$ for each Hilbert-Chow datum with $T$ the spectrum of a field or complete DVR, compatible with base change among fields and complete DVRs, and satisfying the cocycle condition on fields.

**Proof.** This is a consequence of 4.4. □

**Proposition 5.5 (reduction to the diagonal).** To define an isomorphism $\phi_T: f_T(\mathcal{O}_Z, \mathcal{O}_W) \cong f_T(\mathcal{O}_{Z'}, \mathcal{O}_{W'})$ for each Hilbert-Chow datum, for all smooth projective $P$, so that for each $P$, the collection $\{\phi_T\}$:

1. is compatible with base change $S \to T$; and
2. satisfies the cocycle condition; and
3. is Zariski locally effective;

it is sufficient to define an isomorphism $\phi_T$ for each Hilbert-Chow datum with $W = W'$, for all $P$, so that for each $P$, the collection $\{\phi_T\}$ has the stated properties.

**Proof.** On $P \times P \times T$, let $\mathcal{O}_\Delta$ denote the structure sheaf of the diagonal $(\times T)$, i.e., the image of the closed immersion $P \times T \xrightarrow{\Delta \times 1_T} P \times P \times T$. Given $T$-flat closed subschemes $Z, W \hookrightarrow P \times T$, we let $Z \times W \hookrightarrow P \times P \times T$ denote the scheme-theoretic intersection $\text{pr}_{13}^*Z \cap \text{pr}_{23}^*W$. Then there is a canonical isomorphism of line bundles on $T$: $f_T(\mathcal{O}_Z, \mathcal{O}_W; P) \cong f_T(\mathcal{O}_{Z \times W}, \mathcal{O}_\Delta; P \times P)$. Then the proposition follows from the fact that the Hilbert-Chow morphism is compatible with products (3.4). □

**Remark 5.6.** We may even assume $W$ is constant, i.e., there is a $k$-subscheme $W_k$ such that $W = W_k \times_k T$; and we may assume $W_k$ is integral (even smooth).
Notation 5.7. When $W$ and $W'$ are omitted from the notations, this means $W = W'$.

We start with some easy cases of our goal.

Lemma 5.8. Let $(Z, W)$ be a Hilbert datum over any base $T$ such that $Z \cap W = \emptyset$. Then there is a canonical trivialization $\varphi_T^Z : f_T(O_Z, O_W) \cong O_T$ which is compatible with base change.

Proof. The hypothesis implies $O_Z \otimes L O_W$, hence also $R\pi_* (O_Z \otimes L O_W)$, is acyclic. The pullback via $S \to T$ is also acyclic, and the trivialization of the determinant of an acyclic complex is compatible with base change. □

Remark 5.9. The canonical isomorphism $\varphi$ of the lemma has an additivity property in each variable. For example, if $F_1 \to F_2 \to F_3 \to +1$ is a distinguished triangle in $\text{Parf}(P_T)$ such that $\text{Supp}(F_i) \cap W = \emptyset$ for all $i$, then the isomorphism $f_T(F_1, O_W) \otimes f_T(F_3, O_W) \cong f_T(F_2, O_W)$ induced by the triangle corresponds, via the identifications $\varphi_T^Z$, to multiplication $O_T \otimes O_T \to O_T$.

Corollary 5.10. Among Hilbert-Chow data satisfying $Z \cap W = Z' \cap W = \emptyset$, there exists a collection of isomorphisms $\phi_T^{Z, Z'} : f_T(O_Z, O_W) \cong f_T(O_{Z'}, O_W)$ which is compatible with base change and satisfies the cocycle condition.

Proof. We define $\phi_T^{Z, Z'; W} := (\varphi_T^{Z'})^{-1} \circ \varphi_T^Z : f_T(O_Z, O_W) \cong f_T(O_{Z'}, O_W)$ to be the composition of the canonical trivializations. This is compatible with base change because each $\varphi_T^Z$ is. We check the cocycle condition:

$$\phi_T^{Z', Z''; W} \circ \phi_T^{Z, Z'; W} = ((\varphi_T^{Z''})^{-1} \circ \varphi_T^{Z'}) \circ (\varphi_T^Z) = ((\varphi_T^{Z''})^{-1} \circ \varphi_T^{Z'}) \circ \varphi_T^{Z'} = \phi_T^{Z, Z''; W}.$$

□

From now on we keep the collection $\{ \phi_T \}$ whose existence is asserted in 5.10. The idea is to gradually extend it to a collection over Hilbert-Chow data with increasingly complicated incidence structure, until we have covered the whole moduli space. Note that an isomorphism of line bundles on a reduced (e.g., seminormal) scheme is determined by its restriction to generic points (i.e., points of depth 0). Since our base $T$ is always reduced, when we have defined an isomorphism $\phi$ for a more general class of Hilbert-Chow data, to check agreement with previously defined isomorphisms it suffices to check agreement on generic points.

Lemma 5.11. Let $(Z, W)$ be a Hilbert datum over $T$, and suppose that $(Z \cap W)_\eta = \emptyset$ for all generic points $\eta \in T$. (This holds, for example, whenever $Z$ and $W$ intersect properly on $P_T$.) Then there exists a Cartier divisor $D_{Z, W}$ on $T$ and a canonical isomorphism $\varphi_T^Z : f_T(O_Z, O_W) \cong O_T(D_{Z, W})$ characterized by agreeing with the trivialization $\varphi_T^Z_\eta$ for every generic point $\eta \in T$. When $Z \cap W = \emptyset$, $D_{Z, W} = 0$ and $\varphi_T^Z$ is the canonical trivialization. The formation of the divisor $D_{Z, W}$ and the isomorphism $\varphi_T^Z$ are compatible with base change $S \to T$ preserving the generic disjointness.
Furthermore, the formation of $D_{Z,W}$ is additive in each variable: if $\mathcal{F}_1 \to \mathcal{F}_2 \to \mathcal{F}_3 \to \mathcal{F}_4 + 1$ is a distinguished triangle in $\text{Parf}(T)$ such that $(\text{Supp}(\mathcal{F}_i) \cap W)_\eta = \emptyset$ for all generic points $\eta \in T$, all $i$; then $D_{\mathcal{F}_1,W} + D_{\mathcal{F}_2,W} = D_{\mathcal{F}_3,W}$; and the triangle induces, upon application of $\varphi_T^*$, the canonical isomorphism $\mathcal{O}_T(D_{\mathcal{F}_1,W}) \otimes \mathcal{O}_T(D_{\mathcal{F}_2,W}) \cong \mathcal{O}_T(D_{\mathcal{F}_3,W})$. Similarly we have an additivity property in the variable $W$.

Proof. To see the generic disjointness is satisfied when $Z$ and $W$ intersect properly, note that $Z$ (resp. $W$) has codimension $\geq b+1$ (resp. $\geq a+1$) in $P_T$, hence $Z \cap W$ has codimension $\geq a + b + 2$ in $P_T$. Therefore $\dim(Z \cap W) < \dim(T)$, so the support of $\mathcal{O}_Z \otimes \mathcal{O}_W$ cannot dominate any component of $T$. The hypothesis on the incidence means the construction of $[\mathcal{I}_Z \otimes \mathcal{O}_W]$ is additive in each variable.

Remark 5.12. Our essential task is to show that given a Hilbert-Chow datum $(Z, Z', W)$, we have $D_{Z,W} = D_{Z',W}$.

Subsection 5.2. Moving Lemmas.

Proposition 5.13. Let $(Z,W)$ be a Hilbert datum over $T$ the spectrum of a local ring, with $W = W_k \times_k T$ for a $b$-dimensional $k$-scheme $W_k$, and suppose $(Z \cap W)_\eta = \emptyset$ for every generic point $\eta \in T$. Then there exist subvarieties $B_1, \ldots, B_n \subset P$ of dimension $b + 1$, $M_i \in \text{Pic}(B_i)$, and short exact sequences:

$$0 \to M_i \xrightarrow{\alpha_i} \mathcal{O}_{B_i} \to Q_i^0 \to 0$$

$$0 \to M_i \xrightarrow{\beta_i} \mathcal{O}_{B_i} \to Q_i^\infty \to 0$$

such that:

1. $(Z \cap \text{Supp}(Q_i^k))_\eta = \emptyset$ for all $i$, all generic $\eta$; and
2. the $b$-dimensional cycle $[W] + \sum_i([Q_i^0] - [Q_i^\infty])$ is disjoint from $Z$.

Proof. We let $Z_0$ denote the cycle over the closed fiber of $T$. By Chow’s moving lemma [16, Thm.], we can find a cycle $\alpha$ rationally equivalent to $[W_k]$ and satisfying $\alpha \cap Z_0 = \emptyset$; hence also $\alpha \cap Z = \emptyset$ on $P_T$. This shows we can achieve the second property; the issue is to show we can move $W$ in such a way that the first property is satisfied.

Suppose we have a closed immersion $P \hookrightarrow \mathbb{P}^{2n+1}$. Then every step in moving a cycle involves essentially two choices: a linear space $L \cong \mathbb{P}^n \hookrightarrow \mathbb{P}^{2n+1}$, disjoint from $P$, from which projection induces a finite morphism $\pi_L : P \to \mathbb{P}^n$, and an element $g \in \text{PGL}(n+1)$. The excess intersection $e(Z_0, \pi_L^* \pi_L[W]) - [W]$ is smaller than $e(Z_0, [W])$ for generic $L$; and $\pi_L^* (g \cdot \pi_L[W])$ is disjoint from $Z_0$ for generic $g$.

If $(Z \cap W)_\eta = \emptyset$ for all generic points $\eta \in T$, then for generic choices of $L, g$, we have $(Z \cap \pi_L^* (g \cdot \pi_L[W]))_\eta = (Z \cap \pi_L^* (\pi_L[W]))_\eta = \emptyset$. The $Q_i$s are supported in subsets of the form $\pi_L^* (g \cdot \pi_L[W])$ and $\pi_L^* (\pi_L[W])$, hence the result. □
We need a slight variation for subvarieties $W$ as in 5.13 of dimension strictly smaller than $n - a - 1$; eventually we need that such subvarieties do not contribute to $D_{Z,W}$.

**Proposition 5.14.** Let $(Z,W)$ be a Hilbert datum over $T$ a base of dimension $\leq 1$, with $W = W_k \times_k T$ for a $k$-scheme $W_k$, and suppose $(Z \cap W)_\eta = \emptyset$ for every generic point $\eta \in T$.

Suppose further that $\dim(W_k) = b \leq n - a - 2$. Then there exist subvarieties $B_1, \ldots, B_n \subset P$ of dimension $b + 1$, $M_i \in \text{Pic}(B_i)$, and short exact sequences:

$$0 \to M_i \xrightarrow{i^*_i} \mathcal{O}_{B_i} \to Q_i^b \to 0$$

such that:

1. $(Z \cap B_i)_\eta = \emptyset$ for all $i$, all generic $\eta$; and
2. the $b$-dimensional cycle $[W] + \sum_i ([Q_i^b] - [Q_i^\infty])$ is disjoint from $Z$.

**Remark 5.15.** The first condition in the conclusion implies $(Z \cap \text{Supp}(Q_i^b))_\eta = \emptyset$ for all $i$, all generic $\eta$.

**Proof.** Again we are intersecting a finite number of open conditions. Without loss of generality we may assume $W_k$ is an integral subscheme of dimension $n - a - 2$. Let $pr_1(Z) \xleftarrow{\sim} P$ denote the “sweep” of the family $Z$ (with the reduced structure); this is a subscheme of dimension $\leq a + 1$.

Now $pr_1(Z)$ and $W_k$ are not expected to meet, and we have an open dense $U \subset T$ such that $pr_1(Z_U) \cap W_k = \emptyset$. For a generic finite morphism $\pi : P \to \mathbb{P}^n$ (as in the proof of 5.13) we have, possibly after shrinking $U$, that $\pi(pr_1(Z_U)) \cap \pi(W_k) = \emptyset$; and that the pair $(Z, \pi^*\pi_*(W_k) - W_k)$ has smaller excess intersection than does $(Z, W_k)$. Now we move $\pi_*(W_k)$ along a general smooth (affine) rational curve $C \leftarrow \text{PGL}(n + 1)$. Let $\mathcal{Y} \leftarrow \mathbb{P}^n \times C$ note the total space of the resulting family,

Write $\mathcal{Y} = \sum m_i Y_i$, and let $Y_i^{fl} \hookrightarrow \mathbb{P}^n \times \mathbb{P}^1$ denote the flat limit of the family $Y_i \hookrightarrow \mathbb{P}^n \times C$. Then $\mathcal{Y}^{fl} := \sum m_i Y_i^{fl}$ is the unique way to complete $\mathcal{Y}$ to a family of cycles over $\mathbb{P}^1$. Let $pr_1(\mathcal{Y}^{fl}) \leftarrow \mathbb{P}^n$ be the sweep; this is a subscheme of dimension $n - a - 1$. Choose some $t \in T$ such that $Z_t \cap W_k = \emptyset$. For a general choice of $C$, since $\dim(Z_t) + \dim(\mathcal{Y}^{fl}) = a + (n - a - 1) < n$, we will have $\pi(pr_1(Z_t)) \cap \mathcal{Y}^{fl} = \emptyset$. Hence the disjointness holds on an open dense of $T$. This process can be iterated until we have a cycle $\alpha \sim W_k$ such that $pr_1(Z) \cap \alpha = \emptyset$.

The subvarieties $B_i \hookrightarrow P$ lie in subsets of the form $\pi^{-1}(\mathcal{Y}^{fl})$. (This follows from the proof that flat pullback preserves rational equivalence [7, 1.7].) Since a general $\mathcal{Y}^{fl}$ used in one step of the moving process is disjoint from a general member of the family $Z$, this holds after pullback by $\pi$ as well. \hfill $\square$

**Subsection 5.3. Grothendieck-Riemann-Roch.**

**Lemma 5.16.** Let $T$ be the spectrum of a field, and suppose $\mathcal{F}, \mathcal{G} \in \text{Parf}(P_T)$ satisfy $\dim(\text{Supp}(\mathcal{F})) + \dim(\text{Supp}(\mathcal{G})) < n = \dim(P)$. Then $\chi(\mathcal{F} \otimes^L \mathcal{G}) = 0$. 

---

**Documenta Mathematica 16 (2011) 513–543**
Since $F_a(K_0(P))$ is generated by $[O_V]$, $V \subset P$ a subvariety of dimension $\leq a$ [7, Ex. 15.1.5], we may assume $\mathcal{F}, \mathcal{G}$ are structure sheaves of subvarieties of dimensions $a, b$ respectively, with $a + b < n$.

Now since $P$ is smooth, any coherent sheaf has a finite length resolution by finite rank locally free sheaves, so we may apply [7, 18.3.1 (c)] to the closed immersion $i : V \to P$ with $\beta = O_V$. This gives, in the Chow group $A_*(P)$:

$$i_*(\text{ch}(O_V) \cap \text{Td}(V)) = \text{ch}(i_*O_V) \cap \text{Td}(P).$$

As $\text{ch}(O_V) = 1$ and $\text{Td}(V) = [V] + r_V$ with $r_V \in A_{<a}(V)$, the left hand side lies in $A_{\leq a}(P)$.

Since $P$ is smooth, $A^p(P) \cap A_q(P) \subset A_{q-p}(P)$ by [7, 8.3 (b)]. As $\text{Td}(P) = [P] + r_p$ with $r_p \in A_{<n}(P)$, by equating terms in each degree, we find $\text{ch}(i_*O_V) \in A^{\geq a}(P)$.

By Grothendieck-Riemann-Roch (for the smooth $P$, as in [7, 15.2.1]) and the action of $\text{ch}$ on $\otimes$, $\chi(\mathcal{F} \otimes \mathcal{G}^*) = \int_P \text{ch}(\mathcal{F}) \cdot \text{ch}(\mathcal{G}) \cdot \text{Td}_P$. Here $\cdot$ means intersection product of cycle classes. The first possible nonzero term in $\text{ch}(\mathcal{F}) \cdot \text{ch}(\mathcal{G})$ would come from $\text{ch}_{n-a}(\mathcal{F}) \cdot \text{ch}_{n-b}(\mathcal{G})$, but this term is zero for degree reasons. □

PROPOSITION 5.17. Let $T$ be the spectrum of a field, and $Z \hookrightarrow P_T$ an $a$-dimensional sub scheme. Let $M, N \in \text{Coh}(P_T)$ be invertible sheaves on some subvariety of $P_T$ of dimension $\leq n - a$, and suppose we have exact sequences:

$$0 \to M \xrightarrow{\alpha} N \to Q_s \to 0$$

$$0 \to M \xrightarrow{\beta} N \to Q_t \to 0$$

such that $Z \cap \text{Supp}(Q_s) = Z \cap \text{Supp}(Q_t) = \emptyset$.

Then the unique $a_Z \in \Gamma(T, O_T)$ making the following diagram commute:

$$\begin{array}{cccc}
\varphi_Z & \alpha & \varphi_Z & \\
\downarrow \text{via } s & & & \downarrow \text{via } t \\
(f_T(O_Z, M))^{-1} \otimes f_T(O_Z, N) & & & f_T(O_Z, Q_t) \\
\downarrow \varphi_Z & & & \downarrow \varphi_Z \\
O_T & & & O_T
\end{array}$$

depends only on $[Z]$, i.e., $a_Z = a_{Z'}$ if $[Z] = [Z']$.

Proof. To prove the claim it is equivalent to show that the difference between $f(1 \otimes s), f(1 \otimes t) : f_T(O_Z, M) \cong f_T(O_Z, N)$ depends only on $[Z]$.

STEP 1. By taking a filtration of $O_Z$ such that the graded pieces are isomorphic to (twists of) structure sheaves of subvarieties, and using the additivity of the determinant on filtrations, we are reduced to showing that if $F \in \text{Coh}(P_T)$ with $\text{Supp}(F) \subset \text{Supp}(Z)$ and $\dim(\text{Supp}(F)) \leq a - 1$, the induced isomorphisms $f(1 \otimes s), f(1 \otimes t) : f(F, M) \cong f(F, N)$ are equal. (Such a filtration exists by
The subquotients of the filtration of $\mathcal{O}_Z$ depend on the filtration chosen, but the top-dimensional components always appear with their correct multiplicities, i.e., the cycle $[Z]$ can be extracted from the filtration.

**Step 2.** Let $Q_U$ denote the cokernel of the universal $\mathcal{O}_p$-homomorphism $M \to N$; note $Q_U$ is flat over $\text{Hom}_{\mathcal{O}_p}(M, N) \setminus 0$ since a morphism between invertible sheaves is either injective or zero, hence the Euler characteristic of every cokernel is $\chi(M) - \chi(N)$. We consider the line bundle $\det R\pi_s(p_1^*(\mathcal{F}) \otimes L_{Q_U})$ on $\text{Hom}_{\mathcal{O}_p}(M, N) \setminus 0$. Its fiber over $s \in \text{Hom}_{\mathcal{O}_p}(M, N)$ is precisely $f(\mathcal{F}, Q_s)$. Since $Q_s = Q_{\lambda s}$ for $\lambda \in \Gamma(T, \mathcal{O}_T^1)$, we consider $\det R\pi_s(p_1^*(\mathcal{F}) \otimes L_{Q_U})$ as a line bundle on the projective space $\mathbb{P}(\text{Hom}_{\mathcal{O}_p}(M, N) \setminus 0)$. We claim this line bundle is trivial. To prove this it suffices to show it is trivial along a line $\mathbb{P}^1 \cong L \to \mathbb{P}(\text{Hom}_{\mathcal{O}_p}(M, N) \setminus 0)$.

For this purpose Grothendieck-Riemann-Roch (i.e., ignoring torsion) is adequate. More precisely, we consider the GRR diagram:

\[
\begin{array}{ccc}
K_0(P \times L) & \xrightarrow{\text{ch}(-) \cdot \text{Td}(P) \cdot \text{Td}(L)} & A_*(P \times L)_{\mathbb{Q}} \\
R\pi_* & & \pi_* \\
K_0(L) & \xrightarrow{\text{det}} & A_*(L)_{\mathbb{Q}} \\
\text{Pic}(L) & \xrightarrow{c_1} &
\end{array}
\]

We have $\dim(\text{Supp}(p_1^*(\mathcal{F}))) \leq a, \dim(\text{Supp}(Q_U)) \leq n - a$, and $\dim(P \times L) = n + 1$. Hence $\text{ch}(p_1^*(\mathcal{F})) \cdot \text{ch}(Q_U) = 0$ in $A_*(P \times L)_{\mathbb{Q}}$ for degree reasons (as in the proof of 5.16), so $p_1^*(\mathcal{F}) \otimes L_{Q_U} \in K_0(P \times L)$ maps to 0 in the top row. Hence $c_1(\det R\pi_s(p_1^*(\mathcal{F}) \otimes L_{Q_U}))$ is a torsion class, and therefore $\det R\pi_s(p_1^*(\mathcal{F}) \otimes L_{Q_U})$ is trivial.

**Step 3.** For $s \in \text{Hom}_{\mathcal{O}_p}(M, N) \setminus 0$, consider the induced identification $f(s) : f(\mathcal{F}, Q_s) \otimes f(\mathcal{F}, M) \cong f(\mathcal{F}, N)$. Since $f(\mathcal{F}, Q_s)$ is canonically trivial, i.e., the trivialization induced by $\text{Supp}(\mathcal{F}) \cap \text{Supp}(Q_s) = \emptyset$ extends over all $\text{Hom}_{\mathcal{O}_p}(M, N) \setminus 0$, we consider $f(s)$ as an isomorphism $f(\mathcal{F}, M) \cong f(\mathcal{F}, N)$. Since $\chi(\mathcal{F} \otimes L_{M}) = \chi(\mathcal{F} \otimes \mathcal{O}_N) = 0$ by 5.16, we have $f(s) = f(\lambda s)$ for $\lambda \in \Gamma(T, \mathcal{O}_T^1)$. Therefore we have a commutative diagram:

\[
\begin{array}{ccc}
\text{Hom}(M, N) \setminus 0 & \xrightarrow{s \mapsto f(s)} & \text{Isom}(f(\mathcal{F}, M), f(\mathcal{F}, N)) \\
\mathbb{P}(\text{Hom}(M, N) \setminus 0) & \xrightarrow{f} &
\end{array}
\]

But there are no nonconstant functions on $\mathbb{P}(\text{Hom}(M, N) \setminus 0)$, hence $f(s) = f(t) : f(\mathcal{F}, M) \cong f(\mathcal{F}, N)$.
Subsection 5.4. Further Properties of $D_{Z,W}$. Assuming the base and incidence are optimal, we can already prove the following important property of $D_{Z,W}$.

**Proposition 5.18.** Let $(Z, Z', W)$ be a Hilbert-Chow datum over a normal base $T$, and suppose the incidence $Z \cap W$ satisfies:

1. $(Z \cap W)_q = \emptyset$ for all generic points $q \in T$; and
2. $Z \cap W$ is finite over all points of depth 1 in $T$.

Then the Cartier divisors $D_{Z,W}, D_{Z',W}$ are equal.

**Proof.** Since the smooth locus of $T$ contains all points of depth 1, and the formation of $D_{Z,W}$ is compatible with the inclusion $T^{sm} \subset T$, we may assume $T$ is smooth. For $T$ regular there is a canonical isomorphism [12, Prop. 8]:

$$f_T(O_Z, O_W) \cong \otimes_{p,q} (\det_T R^q \pi_* (H^p(O_Z \otimes^L O_W)))^{(-1)^{p+a}}.$$  

To calculate the coefficient of a depth 1 point $t \in T$ in $D_{Z,W}$, we may replace $T$ with the spectrum of the DVR $O_{T,t}$. Then the support of $H^p(O_Z \otimes^L O_W)$ is finite over $T$ (indeed, over $t$), so in the displayed expression only the terms with $q = 0$ can contribute.

By [12, Thm. 3(vi)], the multiplicity of a depth one point is determined by the sum $\sum_{i} (-1)^{i} \ell_t (H^i(R\pi_* (O_Z \otimes^L O_W)))$. This last sum is equal to

$$\sum_{i} (-1)^{i} \ell_t (\pi_* H^i(O_Z \otimes^L O_W)) = (\deg \pi)(\sum_{p,p' \neq t} (-1)^{p} \ell_{p'} (H^p(O_Z \otimes^L O_W))),$$

hence it suffices to show $\gamma(O_Z) := \sum_{p,p' \neq t} (-1)^{p} \ell_{p'} (H^p(O_Z \otimes^L O_W))$ depends only the underlying cycle $[Z]$. We remark that if $Z$ and $W$ are integral and $b = n - a - 1$, then the contribution of a point $t'$ (lying over $t$) to $\gamma(O_Z)$ is exactly Serre’s Tor-formula for the intersection index of $Z$ and $W$ at $t'$ [22, V.C.Thm. 1(b)].

Without loss of generality we assume $W$ is integral and $\dim(W) = b = n - a - 1$; we will see in the proof all sums are 0 if $b < n - a - 1$. Given an exact sequence $0 \to F_1 \to F_2 \to F_3 \to 0$ of coherent sheaves on $P_T$ with support of relative dimension $\leq a$ and satisfying the incidence hypothesis with respect to $W$, by the long exact cohomology sequence we obtain $\gamma(F_1) + \gamma(F_3) = \gamma(F_2)$. It then follows $\gamma$ is additive on filtrations.

Write $[Z] = \sum a_i Z_i$. Again by [11, I.7.4], the sheaf $O_Z$ admits a filtration whose subquotients are invertible sheaves $L_i$ on subvarieties contained in $Z$; and each top-dimensional component $Z_i$ appears exactly $a_i$ times. Since every $L_i$ is some tensor power of a very ample class on $P_T$, we may assume there is either an injective map $O_{Z_i} \to L_i$ or an injective map $L_i \to O_{Z_i}$. Therefore $\gamma(O_Z) \equiv \sum_i \gamma(L_i) = \sum_i a_i \gamma(O_{Z_i})$ modulo summands of the form $\gamma(F)$ where $F$ is a sheaf on $P_T$ whose support over the generic point of $T$ has dimension $\leq a - 1$. So it suffices to show $\gamma$ vanishes on sheaves of this type. Again we may assume $F$ is isomorphic to the structure sheaf of a subvariety $Y \hookrightarrow Z$ in $P_T$. Note that $\dim Y \leq a$, else, being contained in $Z$, $Y$ would dominate $T$ and...
have all fibers of dimension \( \geq a \); and then \( Y \) would contribute to the cycle \([Z]\) of \( Z \).

There are two cases to consider: if \( Y \cap W = \emptyset \), then \( \gamma(F) = 0 \) since \( F_{\leq 1}O_W \) is acyclic. If \( Y \cap W \neq \emptyset \), then the intersection \( Y \cap W \) is improper, and the Tor-formula vanishes at components of improper intersection (due to Serre in the equal characteristic case \cite[22, V.C.Thm.1(a)]{}).

\begin{proof}

Proof.

\end{proof}

**Remarks 5.19.** (5.19.1) The proof shows the relation between the incidence line bundle and Serre's Tor-formula for intersection multiplicities; and also that our construction agrees with Mazur's at least on the normalization of the locus \( U' \). Thus the essential tasks are to extend the divisor through the locus where the expected incidence condition (the hypothesis in 5.18) fails, and to remove the assumption of normality.

(5.19.2) The assumption (C2) of [4] is that the incidence is generically finite over, and nowhere dense in, its image in the base \( S \). These assumptions imply the map \( S \rightarrow \mathcal{C}_n(P) \times \mathcal{C}_{n-a-1}(P) \) factors through \( U' \).

(5.19.3) Considering the GRR diagram as in the proof of 5.17 with \( L \) replaced by a general regular base \( T \), one sees that the first Chern class of the incidence line bundle \( f_T(O_Z,O_W) \) modulo torsion depends only on the underlying cycles \([Z],[W]\), independent of any assumption of properness of intersection. By contrast in 5.18 we have the result integrally.

(5.19.4) We may write \( D_{[Z],W} \) in the case we have a Hilbert datum as in 5.18, e.g., with proper intersection over a regular base \( T \).

**Corollary 5.20.** Among Hilbert-Chow data:

1. over normal bases \( T \); and
2. such that \( Z,W \) (and hence \( Z',W \)) are generically disjoint and have finite incidence over points of depth 1 in \( T \);

there exists a collection of isomorphisms \( \phi_T^{Z',Z} : f_T(O_Z,O_W) \cong f_T(O_{Z'},O_{W}) \) which:

1. is compatible with base change preserving the incidence condition (2);
2. satisfies the cocycle condition; and
3. agrees with the collection on disjoint families defined in 5.10.

**Proof.** We define \( \phi_T^{Z',Z} := (\varphi_T^{Z'})^{-1} \circ \varphi_T^Z : f_T(O_Z,O_W) \cong O_T(D_{[Z],W}) \cong f_T(O_{Z'},O_{W}) \).

**Construction-Notation 5.21.** We continue with the Cartier divisor \( D_{Z,W} \hookrightarrow T \) associated to a Hilbert datum (5.11). For \( Z \hookrightarrow P_T \) a \( T \)-flat family of \( a \)-dimensional subschemes of \( P \) and \( s \in \mathbb{Z}_{\geq 0} \), let \( \text{Coh}_{\leq s,Z_T}(P) \) denote the abelian category of coherent sheaves \( \mathcal{G} \) on \( P \) such that \( \dim(\text{Supp}(\mathcal{G})) \leq s \) and \( (Z \cap \text{Supp}(\mathcal{G}))_\eta = \emptyset \) for all generic points \( \eta \in T \). For \( \mathcal{G} \in \text{Coh}_{\leq n-a-1,Z_T}(P) \), we obtain a Cartier divisor \( D_{Z,\mathcal{G}} \hookrightarrow T \) and a canonical isomorphism \( f_T(O_Z,\mathcal{G}) \cong O_T(D_{Z,\mathcal{G}}) \).

We denote by \( K_0^{*}(P) \) the \( K_0 \)-group of the abelian category \( \text{Coh}_{\leq 1;Z_T}(P) \): we take the free abelian group on sheaves in \( \text{Coh}_{\leq 1;Z_T}(P) \), then impose relations...
from short exact sequences whose terms all lie in \( \text{Coh}_{\leq a, \mathbb{Z}_T} (P) \). Let \( \text{CDiv}(T) \) denote the group of Cartier divisors on \( T \).

We summarize some elementary properties of this construction in the following proposition.

**Proposition 5.22.**

1. The map \( D_{Z,-} : \text{Coh}_{\leq n-a-1, \mathbb{Z}_T} (P) \to \text{CDiv}(T) \)
   
   defined by \( G \mapsto D_{Z,G} \) descends to a homomorphism \( K_0^{n-a-1, \mathbb{Z}} (P) \to \text{CDiv}(T) \) (which we also denote by \( D_{Z,-} \)).

2. If \( f : S \to T \) is a morphism and \( G \in \text{Coh}_{\leq n, \mathbb{Z}_T} (P) \), then \( f^*(D_{Z,G}) = D_{Z,f^*G} \).

3. If \( G \in \text{Coh}_{\leq n-a-1, \mathbb{Z}_T} (P) \) satisfies \( Z \cap G = \emptyset \), then \( D_{Z,G} = 0 \).

4. If \( G \in \text{Coh}_{\leq n, \mathbb{Z}_T} (P) \) and the divisorial part of the family \( Z \hookrightarrow P_T \) is trivial (i.e., \( a \leq \dim(P) - 2 \)), then \( D_{Z,G} = 0 \).

**Proof.** The first three properties follow immediately from the additivity, compatibility with base change, and compatibility with the trivialization of an acyclic complex, of the associated divisor construction discussed in Section 2.

To prove the last property, by the first property we may assume \( G \) is the structure sheaf of a zero-dimensional subvariety \( W \hookrightarrow P \). (If \( k = \mathbb{K} \), this is just a single closed point, but we give an argument here valid for any \( T \)-flat family \( W \hookrightarrow P_T \) of zero-dimensional subschemes, such that \( W \) is integral.)

By [20, 5.3] there is a canonical isomorphism:

\[
\det_T R\pi_* (\mathcal{O}_Z \otimes^L G) \cong (\det_T \pi_* G)^{rk(O_Z)-1} \otimes \det_T (\pi_*(\det_{P_T} (\mathcal{O}_Z)|_W)).
\]

Since the divisorial part of \( \mathcal{O}_Z \) was assumed to be empty, the line bundle \( \det_{P_T} (\mathcal{O}_Z) \) is canonically trivial; and \( \text{rk}(\mathcal{O}_Z) - 1 = -1 \). Therefore the right hand side is canonically trivial, so \( D_{Z,G} = 0 \).

**Remark 5.23.** In case \( a = \dim(P) - 1 \), we refer to [20]. The reduction to the diagonal shows we may assume \( \dim(P) = 1 \) or \( a \leq \dim(P) - 2 \); but to the extent we rely on (4) in 5.22, we must understand the case of zero-cycles and divisors.

**Proposition 5.24.** Suppose \( T \) has dimension \( \leq 1 \) and \( G \in \text{Coh}_{\leq n-a-2, \mathbb{Z}_T} (P) \). Then \( D_{Z,G} = 0 \).

**Remark 5.25.** The condition \( D_{Z,G} = 0 \) is equivalent to the canonical trivialization (induced by the generic acyclicity of \( R\pi_* (\mathcal{O}_Z \otimes^L G) \)) extending over all of \( T \).

**Proof.** We may also assume \( G \) is the structure sheaf of a subvariety \( W \hookrightarrow P \) of dimension \( b \leq n-a-2 \), as these sheaves generate the \( K_0 \)-group.

Since \( D_{Z,G} = 0 \) for \( \dim(\text{Supp}(G)) \leq 0 \) (again by 5.22), it suffices to prove the following claim: if \( D_{Z,G} = 0 \) for all \( G \in \text{Coh}_{\leq b-1, \mathbb{Z}_T} (P) \), and \( 1 \leq b \leq n-a-2 \), then \( D_{Z,G} = 0 \) for all \( G \in \text{Coh}_{\leq b, \mathbb{Z}_T} (P) \).

We prove this claim: by 5.14 there are short exact sequences:

\[
0 \to M_i \xrightarrow{\delta_i} \mathcal{O}_{B_i} \to Q_i^0 \to 0
\]
Suppose we For

The formation of the divisor $D$ is disjoint from $Z$. Then $D = \sum \nu (Q_i) = \sum D_{Z,Q_i}$ in $K_0^{\cdot-a-1}(P)$. Furthermore the difference $Q^*_g - [Q^*_f]$ lies in $\mathcal{F}$. But then $D_{Z,W} = D_{Z,W} + \sum_i (D_{Z,Q_i} - D_{Z,Q_i}^*)$ from the SES $= D_{Z,W} + \sum_i (D_{Z,Q_i} - D_{Z,\{Q_i\}}^*)$ since we assumed $D_{Z,\cdot}$ vanishes on $\mathcal{F}$. Set $\mathcal{G} = 0$.

But then

**Corollary 5.26.** Suppose $T$ is normal and $\mathcal{G} \in \text{Coh}_{\leq n-a-2, Z}$ then $D_{Z,\mathcal{G}} = 0$.

**Proof.** For $t \in T$ of depth 1, the formation of $D_{Z,\mathcal{G}}$ is compatible with the morphism $g_t : \text{Spec} \mathcal{O}_{T,t} \rightarrow T$. By 5.24, we have $g_t^*(D_{Z,\mathcal{G}}) = 0$ for all such $t$. Since $D_{Z,\mathcal{G}}$ is determined its restriction to points of depth 1, the result follows.

**Proposition 5.27.** Suppose $T$ is seminormal and $\mathcal{G} \in \text{Coh}_{\leq n-a-2, Z}$ then $D_{Z,\mathcal{G}} = 0$.

**Proof.** The formation of the divisor $D_{Z,\mathcal{G}}$ is compatible with the (finite, birational) normalization morphism $\nu : T^{\nu} \rightarrow T$. By the previous result we know local equations for $\nu^*(D_{Z,\mathcal{G}})$ are units. We need to show these units are constant along the fibers of $\nu$. Suppose $t \in T$ has branches $b_1, \ldots, b_r$ in $T^{\nu}$. For each $b_t$ there exists a DVR $R_t$ and a morphism $g_t : \text{Spec} R_t \rightarrow T$ such that $R_t$ has residue field $k(t)$, and $g_t$ covers a generalization of $t$ to the locus of subschemes disjoint from $\text{Supp}(\mathcal{G})$. Denote by $S$ the union of the Spec $R_t$s glued along $\text{Spec} k(t)$. Then we have a morphism $g : S \rightarrow T$, and by 5.24 we conclude $g^*(D_{Z,\mathcal{G}}) = 0$. The corresponding trivialization $f_S(\mathcal{O}_Z, \mathcal{G}) \cong \mathcal{O}_S$ is our candidate for extending the trivialization through $t$.

Let $\text{Exc}(\nu) \rightarrow T$ denote the locus over which $\nu$ is not an isomorphism, and let $I_T \rightarrow T$ denote the locus of subschemes $Z$ such that $Z \cap \text{Supp}(\mathcal{G}) \neq \emptyset$. Set $U := T - (\text{Exc}(\nu) \cap I_T)$. Then we have a trivialization $\varphi_U : f_U(\mathcal{O}_Z, \mathcal{G}) \cong \mathcal{O}_U$, and this extends to $\varphi_T : f_T(\mathcal{O}_Z, \mathcal{G}) \cong \mathcal{O}_T$. For $t \notin U$, we constructed in the previous paragraph the isomorphism $t_S : f_S(\mathcal{O}_Z, \mathcal{G}) \cong \mathcal{O}_S$, which we may restrict to $t$. Together these define a pointwise trivialization $f_T(\mathcal{O}_Z, \mathcal{G}) \cong \mathcal{O}_T$, i.e., a nonzero element in every fiber of the line bundle $f_T(\mathcal{O}_Z, \mathcal{G})$. We form the cartesian diagram:

\[
\begin{array}{ccc}
S & \xrightarrow{g} & T \\
\downarrow{\nu} & & \downarrow{\nu} \\
S & \xrightarrow{g} & T \\
\end{array}
\]
Since the formation of $D_{Z,G}$ is compatible with all of the morphisms appearing in this diagram, we obtain $(\nu|S)^*(ts) = (g^\nu)^* (\varphi_T)$. It then follows $\varphi_T$ is constant along the fibers of $\nu$, hence it descends to a trivialization $\varphi_T : f_T(\mathcal{O}_Z, \mathcal{G}) \cong \mathcal{O}_T$. □

**Corollary 5.28.** Suppose $T$ is seminormal and $\mathcal{G} \in \operatorname{Coh}_{\leq n-2;Z}(P)$. Then $D_{Z,G} = D_{Z,[\mathcal{G}]}$.

**Remark 5.29.** For $T$ smooth and $\mathcal{G} \in \operatorname{Coh}_{\leq n-2;Z}(P)$, one can deduce the line bundle $f_T(\mathcal{O}_Z, \mathcal{G})$ is trivial as follows. The filtration of the $K_0(X)$-group by dimension of support with multiplication, if $X$ is a smooth quasi-projective scheme over a field [9, Exp.0 Ch.2 Sect.4 Thm.2.12 Cor.1]. From $\mathcal{O}_Z \in F_{n+\dim T}(K_0(P_T))$ and $\mathcal{G} \in F_{n-2+\dim T}(K_0(P_T))$ it follows that $\mathcal{O}_Z \otimes L_G \in F_{\dim T-2}(K_0(T))$. Therefore $R\pi_*(\mathcal{O}_Z \otimes L_G) \in F_{\dim T-2}(K_0(T))$, and hence $f_T(\mathcal{O}_Z, \mathcal{G}) \cong \mathcal{O}_T$.

For a general base $T$, this reasoning is valid rationally, hence we can conclude $f_T(\mathcal{O}_Z, \mathcal{G})$ is a torsion line bundle.

Since the dimension filtration’s compatibility with multiplication can be viewed as a consequence of the moving lemma, in some sense we have given this proof. Because we need to keep track of the trivialization and not just the abstract invertible sheaf, we work with Cartier divisors rather than line bundles.

**Corollary 5.30.** Let $(Z,W)$ be a Hilbert datum over any base $T$ such that $(Z \cap W)_\eta = \emptyset$ for every generic point $\eta \in T$. Suppose further that $W = W_k \times_k T$ for a $k$-scheme $W_k$ with $[W_k] = \sum a_i W_i$. Then there is a canonical isomorphism $f_T(\mathcal{O}_Z, \mathcal{O}_W) \cong \otimes (f_T(\mathcal{O}_Z, \mathcal{O}_{W_i}))^{\alpha_i}$.

**Subsection 5.5. Application to the incidence line bundle.**

**Construction 5.31.** The facts 5.13 and 5.30 produce a construction; in the description we suppress the base $T$ and we use $f(\cdot) := f(\mathcal{O}_Z, \cdot)$ as before, since the first factor is constant. We use the identification of 5.30: $f(\mathcal{O}_W) \cong \otimes (f(\mathcal{O}_{W_i}))^{\alpha_i} := f([\mathcal{O}_W])$. (In our situation $W$ is a Cartier divisor on a $(b + 1)$-dimensional subvariety $B \hookrightarrow P$ with $[W] = \sum a_i W_i$.)

Now given $(Z,W); B_i, M_i, s_{0,\infty}$ as in 5.13, let $f(s_*): f(M) \otimes f(Q_\chi) \cong f(\mathcal{O}_B)$ denote the isomorphism induced by the short exact sequence. We set

$$\alpha := [W] + \sum_i ([Q_0^i] - [Q_{\infty}^i])$$

to be the moved $b$-dimensional cycle. Then we have a canonical isomorphism $f([\mathcal{O}_W]) \otimes (\otimes_i f([Q_0^i])) = f(\alpha) \otimes (\otimes_i f([Q_{\infty}^i]))$.

Let $\beta_{W,\alpha}^Z : f([\mathcal{O}_W]) \cong f(\alpha)$ be the unique isomorphism making the following diagram commute:

$$
\begin{array}{ccc}
f([\mathcal{O}_W]) \otimes (\otimes_i f([Q_0^i])) & \longrightarrow & f(\alpha) \otimes (\otimes_i f([Q_{\infty}^i]) \otimes f(M_i)) \\
\downarrow^1 \otimes (\otimes_i f(s_{0,\infty}^i)) & & \downarrow^1 \otimes (\otimes_i f(s_{0,\infty}^i)) \\
\beta_{W,\alpha}^Z \otimes 1 & & f(\alpha) \otimes (\otimes_i f(\mathcal{O}_{B_i}))
\end{array}
$$

**Documenta Mathematica 16 (2011) 513–543**


Let \( \phi_T \) have been chosen. Then the isomorphism:

\[
\phi_T: (Z', W') \rightarrow (Z, W)
\]

is independent of the choice of \( \phi_T \). Suppose subvarieties \( B_1, \ldots, B_n \subset P \) and short exact sequences as in 5.13 have been chosen. Then the isomorphism:

\[
(\varphi^{Z', \alpha} \circ \beta^{Z'}_{W, \alpha})^{-1} \circ (\varphi^{Z, \alpha} \circ \beta^{Z}_{W, \alpha}) : f_T(O_Z, O_W) \cong f_T(O_{Z'}, O_{W'}) \cong f_T(O_Z, \alpha) \cong f_T(O_{Z'}, \alpha) \cong f_T(O_Z, O_W)
\]

is independent of the choice of \( \phi_T \). Suppose subvarieties \( B_1, \ldots, B_n \subset P \) and short exact sequences as in 5.13 have been chosen. Then the isomorphism:

\[
(\varphi^{Z', \alpha} \circ \beta^{Z'}_{W, \alpha})^{-1} \circ (\varphi^{Z, \alpha} \circ \beta^{Z}_{W, \alpha}) = (\varphi^{Z', \hat{\alpha}} \circ \beta^{Z'}_{W, \hat{\alpha}})^{-1} \circ (\varphi^{Z, \hat{\alpha}} \circ \beta^{Z}_{W, \hat{\alpha}}).
\]

Furthermore \( (\varphi^{Z', \alpha} \circ \beta^{Z'}_{W, \alpha})^{-1} \circ (\varphi^{Z, \alpha} \circ \beta^{Z}_{W, \alpha}) \) agrees with the canonical identifications at every generic \( \eta \in T \). If in addition \( T \) is the spectrum of a regular local ring and the incidence \( Z \cap W \) satisfies the hypotheses of 5.20, then \( \phi^T_Z \) is defined by a constant depending only on \( [Z] \), and this is exactly canceled by the map back to \( f_T(O_{Z'}, O_W) \). In other words, since \( a_Z = a_{Z'} \), the following diagram commutes.

\[
\begin{array}{c}
\phi^{Z', \alpha}_{\eta} \circ (\beta^{Z'}_{W, \alpha})_{\eta} \circ \varphi^{Z, \alpha}_{\eta} = (\beta^Z_{W, \alpha})_{\eta} \circ \varphi^{Z', \alpha}_{\eta} \circ \phi^T_{\eta}
\end{array}
\]

From this our claim follows: two choices of moving data produce isomorphisms which agree at the generic points of \( T \), hence they agree; and the isomorphism \( \phi^T_Z \), for \( T \) regular is characterized by agreeing with the composition of the canonical trivializations over the generic points of \( T \).

**Corollary 5.33.** Among Hilbert-Chow data over bases \( T \) satisfying:

1. \( T \) is either regular, or local; and
2. all generic points of \( T \) correspond to disjoint subschemes;

**Proof.** We claim any choice of moving data produces the canonical isomorphism over the generic points of \( T \). Now 5.17 shows the choice of short exact sequences affects the map \( \varphi^{Z, \alpha}_{\eta} \circ (\beta^{Z}_{W, \alpha})_{\eta} : f_{\eta}(O_Z, O_W) \cong f_{\eta}(O_Z, \alpha) \cong O_{\eta} \) by a constant depending only on \( [Z] \), and this is exactly canceled by the map back to \( f_{\eta}(O_{Z'}, O_W) \). In other words, since \( a_Z = a_{Z'} \), the following diagram commutes.

\[
\begin{array}{c}
\phi^{Z', \alpha}_{\eta} \circ (\beta^{Z'}_{W, \alpha})_{\eta} \circ \varphi^{Z, \alpha}_{\eta} = (\beta^Z_{W, \alpha})_{\eta} \circ \varphi^{Z', \alpha}_{\eta} \circ \phi^T_{\eta}
\end{array}
\]

From this our claim follows: two choices of moving data produce isomorphisms which agree at the generic points of \( T \), hence they agree; and the isomorphism \( \phi^T_Z \), for \( T \) regular is characterized by agreeing with the composition of the canonical trivializations over the generic points of \( T \). \( \square \)
there exists a collection of isomorphisms \{\phi_T\} which:

(1) is compatible with base change preserving generic disjointness;
(2) satisfies the cocycle condition; and
(3) extends the collection of 5.20.

Proof. To see our construction commutes with base change preserving generic disjointness, note that moves valid over \(T\) (i.e., producing \(\beta_{W,a}\)) pullback via \(S \rightarrow T\) to suitable moves on \(S\). The cocycle condition (an equality of two isomorphisms of line bundles on a reduced scheme) holds at the generic points, hence it holds everywhere. \(\square\)

The previous corollary provides us an isomorphism \(f_s(\mathcal{O}_Z, \mathcal{O}_W) \cong f_s(\mathcal{O}_{Z'}, \mathcal{O}_W)\) for a Hilbert-Chow datum over \(s\) the spectrum of a field corresponding to a point in the incidence locus, namely the restriction of the isomorphism over the local ring, possibly followed by a field extension. In the next proposition we observe this is compatible with specializations from the locus of disjoint subschemes into the incidence locus.

**Proposition 5.34.** Let \((Z, Z', W)\) be a Hilbert-Chow datum over \(s\) the spectrum of a field \(\kappa(s)\) corresponding to a point of incidence, i.e., \(Z \cap W, Z' \cap W \neq \emptyset\). Let \((Z_T, Z'_T, W)\) be a Hilbert-Chow datum over \(T\) the spectrum of a DVR covering a generization from \(s\) to the locus of disjoint subschemes, and with \(T_0 = s\). Then the isomorphism

\[
\begin{align*}
    f_s(\mathcal{O}_Z, \mathcal{O}_W) & \xrightarrow{\text{can}} f_T(\mathcal{O}_{Z_T}, \mathcal{O}_W) \\
    \phi_{T,s} & \xrightarrow{\text{can}} f_T(\mathcal{O}_{Z^*_T}, \mathcal{O}_W) \\
    f_s(\mathcal{O}_Z', \mathcal{O}_W) & \xleftarrow{\text{can}} f_T(\mathcal{O}_{Z'^*_T}, \mathcal{O}_W)
\end{align*}
\]

is equal to the isomorphism induced by \(\phi_R := \phi_{\text{Spec} R}\), where \((R, \mathfrak{m}, K = R/\mathfrak{m})\) is the (seminormal) local ring of the image of \(s\) on \((\mathcal{M}_a \times \mathcal{M}_a)^{an}\). In other words, the previously displayed isomorphism is equal to \((- \otimes_K \kappa(s))\) of the following isomorphism:

\[
\begin{align*}
    f_K(\mathcal{O}_Z, \mathcal{O}_W) & \xrightarrow{\text{can}} f_R(\mathcal{O}_{Z^*_R}, \mathcal{O}_W) \\
    \phi_{R,s} & \xrightarrow{\text{can}} f_R(\mathcal{O}_{Z'^*_R}, \mathcal{O}_W) \\
    f_K(\mathcal{O}_Z', \mathcal{O}_W) & \xleftarrow{\text{can}} f_R(\mathcal{O}_{Z'^*_R}, \mathcal{O}_W)
\end{align*}
\]

In particular, if we generize to the locus of disjoint subschemes and then restrict, the resulting isomorphism at \(s\) is independent of the choice of generization.

Proof. Since the collection \{\phi_T\} is compatible with base change preserving generic disjointness, we may replace \(R\) by \(R/I\) for some ideal \(I \subset R\) such that \(R/I\) is a local domain whose generic point \(\text{Spec} L\) is the image of the generic point of \(T\). Therefore we have commutative squares:
To show $\phi_R \times_R T = \phi_T$, it suffices to show they agree at $\eta$, i.e., that $(\phi_R \times_R T) \times_T \eta = \phi_T \times_T \eta$. But this is equivalent to $(\phi_L) \times_L \eta = \phi_\eta$, which is a consequence of the compatibility with base change on pairs of disjoint subschemes (5.10).

**Corollary 5.35.** Among Hilbert-Chow data satisfying at least one of the following conditions:

1. the conditions of 5.33;
2. the base $T$ is a field;

there exists a collection of isomorphisms $\{\phi_T\}$ which:

1. satisfies the cocycle condition;
2. extends the collection of 5.33 (so is compatible with base change preserving generic disjointness); and
3. is compatible with specialization from the locus of disjoint subschemes to the incidence locus.

**Proof.** The compatibility with specialization is built into the construction. The new feature to check is the cocycle condition on field points mapping to the incidence locus. But if the cocycle condition holds after generization and the covering isomorphism is compatible with base change, then the collection must also satisfy the cocycle condition at new (field) points.

We augment 5.35 to include specializations fully within the incidence locus, hence we have the first part of 5.1.

**Theorem 5.36.** Among Hilbert-Chow data satisfying at least one of the following conditions:

1. the conditions of 5.33;
2. the base $T$ is a field;
3. the base $T$ is a DVR;

there exists a collection of isomorphisms $\{\phi_T\}$ which:

1. satisfies the cocycle condition;
2. extends the collection of 5.35 (so is compatible with base change preserving generic disjointness); and
3. is compatible with arbitrary specialization.

Therefore, in the notation of 5.1, the incidence bundle $\mathcal{L} \in \text{Pic}(Y_0)$ lifts to an element $(\mathcal{L}, \phi) \in \text{Pic}(Y_\bullet)$.
Proof. This follows from the general lemma 4.5, the preceding result 5.35, and
the hypothesis that $T$ is seminormal. □

Now we conclude the proof of 5.1.

Theorem 5.37 (Zariski local effectiveness). The element $(\mathcal{L}, \phi = \{\phi_T\})$ of 5.36
is Zariski locally effective: for any cycle $(z, W) \in \mathcal{U} \subset \mathcal{C}_a \times \mathcal{C}_b$ there exists an
open subscheme $V \subset \mathcal{U}$ containing $(z, W)$ and an isomorphism $t : \mathcal{L}|_{\pi^{-1}_0(V)} \cong
\mathcal{O}_{\mathcal{H}}$ which is compatible with $\phi$. Therefore, the incidence bundle $\mathcal{L}$ descends
to $\mathcal{U}$.

Proof. This structure is built into the definition of the descent datum $\phi = \{\phi_T\}$. By 5.5 we may assume $W$ is fixed. Suppose $z$ and $W$ are disjoint. Then $W$ is
disjoint from all cycles in a neighborhood $V$ of $z$, and also from all subschemes
in $V_0 := \pi_0^{-1}(V) \subset \mathcal{H}$. We let $Z_{V_0} \hookrightarrow P \times V_0$ denote the corresponding
family. On $V_0$ we use the canonical trivialization $\varphi_{V_0}^{Z_{V_0}} : f_{V_0}(O_{Z_{V_0}}, O_W) :=
det_{V_0}(R\pi_* (O_{Z_{V_0}} \otimes L O_W)) \cong O_{V_0}$ induced by the acyclicity of $O_{Z_{V_0}} \otimes L O_W$.
Then by our definition of $\phi$ on disjoint subschemes, the following diagram commutes:

$$
\begin{array}{ccc}
\mathcal{L}|_{\pi^{-1}_0(V_0)} & \xrightarrow{\varphi_{V_0}^{Z_{V_0}}} & \mathcal{O}_{V_0} \\
\uparrow{\phi_{V_0}^Z} & & \uparrow{\varphi_{V_0}^Z} \\
\mathcal{L}|_{\pi^{-1}_0(V_0)} & \xrightarrow{\varphi_{V_0}^{W}} & \mathcal{O}_{V_0}
\end{array}
$$

For a pair $(z, W)$ in the incidence locus, choose a collection of short exact
sequences as in 5.13 moving $W$ to a rationally equivalent $\alpha$ such that $z \cap \alpha = \emptyset$.
Then also $z' \cap \alpha = \emptyset$ for $z'$ in a neighborhood $\mathcal{V} \ni z$, and $\alpha$ is disjoint from all
subschemas parameterized by $V_0 := \pi_0^{-1}(V)$. Then we define $t$ to be the
trivialization induced by the move, then the acyclicity of $O_{Z_{V_0}} \otimes L O_{W_i}$ (for all
$W_i \in \text{Supp}(\alpha)$):

$$
t : f_{V_0}(O_{Z_{V_0}}, O_W) \cong f_{V_0}(O_{Z_{V_0}}, \alpha) \xrightarrow{\varphi_{V_0}^{Z_{V_0}}} \mathcal{O}_{V_0}.
$$

This is compatible with $\phi$ by 5.32. Zariski local effectiveness implies effective-
ness by 3.1 and 4.6. □

Remark 5.38. In [18] we proved 1.1 for pairs of 1-dimensional cycles on a
threefold $P$ (the case $n = 3, a = b = 1$) by constructing isomorphisms $\varphi_T^{Z, W} : f_T(O_Z, O_W) \cong f_T(Z, [W])$ for all Hilbert data $(Z, W)$ for $P$ over spectra of
fields and complete DVRs $T$. (We use the evident extension of the “determinant
of a cycle” notation from 5.3 when both variables are cycles.) Thus we obtained
for all Hilbert-Chow data $(Z, W), (Z', W')$ over spectra of fields and DVRs a
system of isomorphisms $\phi_T^{Z,Z';W,W'} := (\varphi_T^{Z',W'})^{-1} \circ \varphi_T^{Z,W} : f_T(\mathcal{O}_Z, \mathcal{O}_W) \cong f_T(\mathcal{O}_{Z'}, \mathcal{O}_{W'})$.

We explain the strategy used in [18] to construct the $\varphi$ isomorphisms. Consider first the case where $T$ is the spectrum of a field, and write $[Z] = \sum n_i Z_i$. Then $\mathcal{O}_Z, \mathcal{O}_{Z_i} \in F_1(K_0(P_T))$ and the difference $\mathcal{O}_Z - \sum n_i \mathcal{O}_{Z_i}$ lies in $F_0(K_0(P_T))$.

Since the determinant is additive on filtrations, to define $\varphi_T^{Z,W}$ it is enough to trivialize in a sufficiently canonical way the determinant $f_T(\mathcal{O}_Z, \mathcal{O}_W)$ when at least one of $Z, W$ is zero-dimensional. To achieve this we used the explicit form of the isomorphism in [20, 5.3]. For $T$ the spectrum of a DVR, one has also to trivialize the determinant $f_T(\mathcal{O}_Z, \mathcal{O}_W)$ when at least one of $Z, W$ is supported in the closed fiber. For this we used an exact sequence given by a uniformizer. To summarize, as far as the incidence bundle is concerned, [20, 5.3] trivializes canonically the difference between the Hilbert scheme and the Chow variety. Thus what was lacking in [18] was a generalization of [20, 5.3] to higher dimensions.

To obtain the functoriality of the resulting collection $\{\phi_T\}$ in [18], we deduced from the construction of the $\varphi$ isomorphisms a natural list of properties (essentially: being additive on triangles, and agreeing with prescribed normalizations on structure sheaves of subvarieties) sufficient to characterize them, then checked the properties were stable under base change. This method worked on all of $\mathcal{E} \times \mathcal{E}$, not just the locus $\mathcal{U}$. By contrast, our approach here is to restrict to the locus $\mathcal{U}$ and compare with the canonical trivialization on the locus $U$ of disjoint cycles. As for the effectiveness of the descent datum, both in [18] and in the present work the verification of the Zariski local effectiveness of the collection $\{\phi_T\}$ employs Grothendieck-Riemann-Roch to analyze the effect on $\varphi$ of a choice of “moving collection” of short exact sequences.

As a final point of contrast, in [18] we made use of the product structure of the Hilbert-Chow proper hypercovering and showed, using [20, 3.12], that the incidence bundle could be descended to the product of a Hilbert scheme and a Chow variety. Then we showed the descended bundle inherited a Zariski locally effective descent datum of its own. In other words, we wrote $\pi = (FC \times \text{Id}) \circ (\text{Id} \times FC) : \mathcal{H} \times \mathcal{H} \rightarrow \mathcal{E} \times \mathcal{E}$ and descended along $(FC \times \text{Id})$ and $(\text{Id} \times FC)$ separately.

Subsection 5.6. Conclusion of the proof of 1.1. Finally we verify the properties stated in 1.1.

Descent of rational section. To see we have actually constructed a Cartier divisor in $\mathcal{U} \subset \mathcal{E}_a \times \mathcal{E}_b$ supported on the incidence locus, consider the diagram whose vertical arrows are the restriction maps:

\[
\begin{array}{ccc}
\text{Pic}(\mathcal{U}) & \rightarrow & \text{Pic}(\mathcal{U}_*) \\
| & & | \\
\text{Pic}(U) & \rightarrow & \text{Pic}((\pi^{-1}(U))_*)
\end{array}
\]
The arrow in the bottom row is injective by 4.6. On the locus of disjoint subschemes, the isomorphism constructed in 5.8, \( \varphi : \mathcal{L}|_{U_0} \cong \mathcal{O}_{U_0} \), is an isomorphism of pairs \( \mathcal{L}|_{(\pi^{-1}(U))} \) (corresponding to subschemes \( U \subset \mathcal{U} \)). Hence by the injectivity of the bottom row, the trivialization \( \varphi \) descends to \( U \subset \mathcal{U} \).

**Restriction to** \( U' \)** is effective. To check that the restriction \( D|_{U'} \) is effective, we may replace \( U' \) with its normalization \( U'' \). Then we may replace \( U'' \) with the local ring of some depth 1 point \( t \) on \( U'' \). By the assumption that the universal cycles intersect properly, over a given component \( C \subset U'' \), the incidence has dimension \( \dim(C) - 1 \). This is preserved by the finite base change \( U'' \to U' \).

Suppose first the incidence is generically finite onto its image. Then 5.18 applies, and the coefficient of \( t \) in \( D|_{U''} \) is a sum of intersection multiplicities of properly intersecting components (weighted with positive coefficients). If the incidence dominates \( t \), this coefficient is positive by [22, V.C. Thm.1(b)]]; in any case the coefficient is nonnegative.

If the incidence has generic positive dimension over its image, then its image must have dimension \( \leq \dim(C) - 2 \). Hence in this case the associated coefficient is 0.

**Intersection multiplicity.** On the Chow varieties we have the incidence bundle \( \mathcal{M} \) and its rational section over the locus of disjoint cycles, giving the Cartier divisor \( D \to \mathcal{U} \). This pulls back via the Hilbert-Chow morphism \( \pi : Y_0 \to \mathcal{U} \) to the determinant line bundle \( \mathcal{L} \) and its rational section over the locus of disjoint subschemes. Our goal is to relate the order of vanishing of a local defining equation of \( D \), to intersection numbers. So let \( s_D \) be the canonical (rational) section of the line bundle \( \mathcal{O}_D(D) \).

If \( g : T \to \mathcal{U} \) is a morphism from the spectrum of a discrete valuation ring \( R \supset k \) (corresponding to cycles \( Z, W \)), there exists a discrete valuation ring \( R' \) which is finite over \( R \), and such that the composition \( g' : T' := \text{Spec} R' \to T \to \mathcal{U} \) factors through \( Y_0 \). (Note that if we start with a specialization from a generic point of \( \mathcal{U} \), we can find a component of the Hilbert scheme so that no generic extension is necessary.) If \( \text{ord}(g'(s_D)) = \text{deg}(Z_{T'} \cdot W_{T'}) \), it follows that \( \text{ord}(g'(s_D)) = \text{deg}(Z \cdot W) \). Thus we may assume our specialization factors through the Hilbert scheme, corresponding to subschemes \( Z, W \) such that \( [\tilde{Z}] = Z \) and \( [\tilde{W}] = W \). Now we assume disjointness over the generic point \( \eta \in T \), i.e., \( g(\eta) \in U \). Let \( t \in T \) denote the closed point.

First we have:

\[
\text{ord}_T(s_D) = \text{ord}_T(s_{\pi \cdot D}) = \sum_p (-1)^p t_p(\mathcal{H}^p(R\pi_*(\mathcal{O}_Z \otimes^L \mathcal{O}_W)))
\]

since each \( \mathcal{H}^p(R\pi_*(\mathcal{O}_Z \otimes^L \mathcal{O}_W)) \) is a torsion \( T \)-module, and by [12, Thm.3(v)]. Since the scheme \( P_T \) is smooth, the filtration of the \( K_0 \)-groups by dimension is compatible with multiplication, thus \( \mathcal{O}_Z \otimes^L \mathcal{O}_W \) and \( R\pi_*(\mathcal{O}_Z \otimes^L \mathcal{O}_W) \) are classes of dimension zero. Then \( \sum_p (-1)^p t_p(\mathcal{H}^p(R\pi_*(\mathcal{O}_Z \otimes^L \mathcal{O}_W))) \) is equal to

\[
\text{ord}_T(s_D) = \text{ord}_T(s_{\pi \cdot D}) = \sum_p (-1)^p t_p(\mathcal{H}^p(R\pi_*(\mathcal{O}_Z \otimes^L \mathcal{O}_W)))
\]
the degree of the $K_0$-classes $\mathcal{O}_Z \otimes L \mathcal{O}_{\bar{W}}$ and $R\pi_*(\mathcal{O}_Z \otimes L \mathcal{O}_{\bar{W}})$. Note also the refined class $Z \cdot W$ is of the expected dimension, i.e., $Z \cdot W \in A_0(P_T)$. We have $\mathcal{O}_Z \otimes L \mathcal{O}_{\bar{W}} = \sum_i (-1)^i [\text{Tor}^1_i (\mathcal{O}_Z, \mathcal{O}_{\bar{W}})] \in K_0(P_T)$. The degree of this class is computed by [7, 20.4]: it is simply the degree of the refined class $Z \cdot W \in A_0(P_T)$, since the terms of dimension $< 0$ necessarily vanish.

References

[1] B. Angéniol. Familles de cycles algébriques - Schéma de Chow, volume 896. Springer-Verlag, 1981.
[2] M. Artin, A. Grothendieck, and Verdier J.-L. Théorie des topos et cohomologie étale des schémas: séminaire de géométrie algébrique du Bois Marie, 1963-64, SGA 4, volume 305. Springer-Verlag, 1972.
[3] D. Barlet. Espace analytique réduit des cycles analytiques complexes compacts d’un espace analytique complexe de dimension finie, volume 482. Springer-Verlag, 1975.
[4] D. Barlet and M. Kaddar. Incidence divisor. Internat. J. Math., 14(4):339–359, 2003.
[5] David Eisenbud, Frank-Olaf Schreyer, and Jerzy Weyman. Resultants and Chow forms via exterior syzygies. J. Amer. Math. Soc., 16(3):537–579 (electronic), 2003.
[6] John Fogarty. Truncated Hilbert functors. J. Reine Angew. Math., 234:65–88, 1969.
[7] William Fulton. Intersection theory, volume 2 of Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]. Springer-Verlag, Berlin, second edition, 1998.
[8] S. Greco and C. Traverso. On seminormal schemes. Compositio Math., 40(3):325–365, 1980.
[9] A. Grothendieck, P. Berthelot, and L. Illusie. SGA 6. Lecture Notes in Mathematics, 225.
[10] Robin Hartshorne. Residues and duality. Lecture notes of a seminar on the work of A. Grothendieck, given at Harvard 1963/64. With an appendix by P. Deligne. Lecture Notes in Mathematics, No. 20. Springer-Verlag, Berlin, 1966.
[11] Robin Hartshorne. Algebraic geometry. Springer-Verlag, New York, 1977. Graduate Texts in Mathematics, No. 52.
[12] Finn Faye Knudsen and David Mumford. The projectivity of the moduli space of stable curves. I. Preliminaries on “det” and “Div”. Math. Scand., 39(1):19–55, 1976.
[13] János Kollár. Rational curves on algebraic varieties, volume 32 of Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern
The Hilbert-Chow Morphism and the Incidence Divisor

Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]. Springer-Verlag, Berlin, 1996.

[14] Barry Mazur. Private letter to Bruno Harris, Mar 10, 1993.
[15] David Mumford and John Fogarty. Geometric invariant theory, volume 34 of Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas]. Springer-Verlag, Berlin, second edition, 1982.
[16] Joel Roberts. Chow’s moving lemma. In Algebraic geometry, Oslo 1970 (Proc. Fifth Nordic Summer School in Math.), pages 89–96. Wolters-Noordhoff, Groningen, 1972. Appendix 2 to: “Motives” (ıt Algebraic geometry, Oslo 1970 (Proc. Fifth Nordic Summer School in Math.), pp. 53–82, Wolters-Noordhoff, Groningen, 1972) by Steven L. Kleiman.
[17] Joseph Ross. The incidence divisor of cycles in projective space. To appear in Communications in Algebra.
[18] Joseph Ross. The Hilbert-Chow morphism and the incidence divisor. PhD thesis, Columbia University, 2009.
[19] Joseph Ross. A characterization of seminormal schemes. J. Algebra, 323(5):1494–1502, 2010. Corrigendum 323(11):3155-3156, 2010.
[20] Joseph Ross. The Hilbert-Chow morphism and the incidence divisor: Zero-cycles and divisors. Journal of Pure and Applied Algebra, 215(11):2586 – 2599, 2011.
[21] D. Rydh. Families of cycles and the Chow scheme. PhD thesis, Kungliga Tekniska högskolan, 2008.
[22] Jean-Pierre Serre. Local algebra. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2000. Translated from the French by CheeWhye Chin and revised by the author.
[23] Carlo Traverso. Seminormality and Picard group. Ann. Scuola Norm. Sup. Pisa (3), 24:585–595, 1970.
[24] Bin Wang. Mazur’s incidence structure for projective varieties. I. Compositio Math., 115(3):303–327, 1999.
PRESENTATION OF AN Iwasawa Algebra:
THE CASE OF $\Gamma_1 SL(2, \mathbb{Z}_p)$

LAURENT CLOZEL

Received: July 1, 2010
Revised: February 17, 2011
Communicated by Takeshi Saito

Abstract. We give an explicit presentation of the $p$-adic Iwasawa algebra of the subgroup of level one of $SL(2, \mathbb{Z}_p)$ for $p \neq 2$.

2010 Mathematics Subject Classification: 11F85, 11S31, 22E50
Keywords and Phrases: Non-commutative Iwasawa agebras, p-adic groups

Assume $G$ is a semi-simple Chevalley group, so $G(\mathbb{Z}_p) \subset G(\mathbb{Q}_p)$ is a maximal compact subgroup. Both the $p$-adic representation theory of $G(\mathbb{Q}_p)$ and non-commutative Iwasawa theory involve the Iwasawa algebra of $G(\mathbb{Z}_p)$ or suitable congruence subgroups. It seems to have been assumed that explicit descriptions, by generators and relations, of these algebras were inaccessible. However, it is a general principle that natural objects coming from semi-simple (split) groups have explicit presentations. Famous examples are Serre’s presentation of the semi-simple algebras and Steinberg’s presentation of the Chevalley groups [7, 8]. In this paper we will give a presentation for the Iwasawa algebra of the subgroup of level 1 in $SL(2, \mathbb{Z}_p)$ ($p \neq 2$).

I thank M. Duflo and, once more, J.-P. Serre for useful discussions. Thanks are also due to P. Schneider for reading a first draft and suggesting several corrections, and to the referee for his careful work which eliminated an embarrassing mistake.

1Membre de l’Institut Universitaire de France
Let $G = SL(2)$ and let $G$ be the subgroup of level 1 in $G(Z_p)$:

$$G = \{ g \in SL(2, \mathbb{Z}_p) : g \equiv 1[p] \}.$$ 

We assume $p > 2$, so $G$ has no $p$–torsion. It has a triangular decomposition

$$G = N^- T N^+$$

where $N^- = \begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix}$, $N^+ = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$ (entries * in $p\mathbb{Z}_p$) and $T = \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}$ (entries in $1 + p\mathbb{Z}_p$). We identify $N^-$, $N^+$ with $\mathbb{Z}_p$ by $* = px (x \in \mathbb{Z}_p)$. Similarly $T \sim \mathbb{Z}_p$ by $x \mapsto -x (1 + p) x (1 + p) - x (x \in \mathbb{Z}_p)$.

We consider the Iwasawa algebra $\Lambda_G$ of $\mathbb{Z}_p$–valued measures (or distributions, in the sense of [9]), on $G$, which we will denote by $D(G, \mathbb{Z}_p)$. The triangular decomposition of $G$, as an analytic manifold, yields a decomposition of $D(G, \mathbb{Z}_p)$ as a topological $\mathbb{Z}_p$–module:

$$D(G, \mathbb{Z}_p) = D(N^-, \mathbb{Z}_p) \hat{\otimes} D(T, \mathbb{Z}_p) \hat{\otimes} D(N^+, \mathbb{Z}_p),$$

the factors of (1.1) being the spaces of distributions on the factors of $G$. If $f$ is a function on $G$ and $U$, $V$, $W$ distributions on $N^-$, $T$, $N^+$,

$$< U \otimes V \otimes W, f > := < U \otimes V \otimes W, f(ush) >$$

where $u \in N^-$, $h \in T$, $n \in N^+$ and $f$ is therefore seen as a function on $N^- \times T \times N^+$. The natural definition of the completed tensor product is equivalent to the explicit description of $D(G, \mathbb{Z}_p)$ reviewed below.

The algebra $\Lambda_{\mathbb{Z}_p} = D(\mathbb{Z}_p, \mathbb{Z}_p)$ is identified with the ring of power series $\mathbb{Z}_p[[T]]$ by Iwasawa’s theorem. For $\mu \in \Lambda_{\mathbb{Z}_p}$, the associated series is given by the Fourier–Amice transform

$$\hat{\mu}(t) = \int_{\mathbb{Z}_p} (1 + t)^x d\mu(x) \quad (t \in \mathbb{Z}_p, |t| < 1).$$

In particular, $\delta(x)$ being the Dirac measure at $x$:

$$\hat{\delta}(1) = 1 + T,$$

so

$$T = \hat{\delta}(1) - \hat{\delta}(0).$$

\textbf{Documenta Mathematica 16 (2011) 545–559}
In each factor of the decomposition (1.1), we therefore have the Dirac measures:

\[
\begin{align*}
\mu_- &= \delta\left(\frac{1}{p} \ 1\right), \\
\hat{\mu}_- &= 1 + Y \in \mathcal{D}(N^-, \mathbb{Z}_p) \cong \\
&\cong \mathbb{Z}_p[[Y]] \\
\mu_+ &= \delta\left(1 \ \frac{p}{1}\right), \\
\hat{\mu}_+ &= 1 + X \in \mathcal{D}(N^+, \mathbb{Z}_p) \cong \\
&\cong \mathbb{Z}_p[[X]] \\
\mu_0 &= \delta\left(\frac{(1+p)}{(1+p)^{-1}}\right), \\
\hat{\mu}_0 &= 1 + H \in \mathcal{D}(T, \mathbb{Z}_p) \cong \\
&\cong \mathbb{Z}_p[[H]]
\end{align*}
\]

For each factor, \( U = N^-, T \) or \( N^+ \) of \( G \), \( \mathcal{D}(U, \mathbb{Z}_p) \) is naturally sent to \( \mathcal{D}(G, \mathbb{Z}_p) \), by integrating a function \( f \in C(G, \mathbb{Z}_p) \) against \( \mu \in \mathcal{D}(U, \mathbb{Z}_p) \) on the \( U \)-factor. This map is compatible with the convolution product. We therefore write, unambiguously, \( Y^n, X^n, H^n \) (\( n \geq 0 \)) in \( \mathcal{D}(G, \mathbb{Z}_p) \). A distribution \( \lambda \) in this space can then be written uniquely

\[
(1.3) \quad \lambda = \sum_n \lambda_n \ Y^n \ H^n \ X^n \quad (n \in \mathbb{N}^3)
\]

with \( \lambda_n \in \mathbb{Z}_p \). This is the meaning of the completed tensor product (1.1). The expansion is convergent in \( \mathcal{D}(G, \mathbb{Z}_p) \). Of course the product \( Y^n H^n X^n := Y^n \otimes H^n \otimes X^n \) is defined as above. This easily follows from Mahler’s theorem in several variables (cf. Lazard [4, Théorème 1.2.4]).

It immediately follows from formula (1.2) that the distributions \( Y, H, X \in \mathcal{D}(G, \mathbb{Z}_p) \) multiply in the obvious fashion when the variables are taken in the “natural order”, i.e.

\[
\begin{align*}
Y \otimes H &= Y \ast H \\
Y \otimes X &= Y \ast X \\
H \otimes X &= H \ast X 
\end{align*}
\]

the convolution product being taken on \( G \). We will simply write, consistent with previous notation:

\[
(1.4) \quad YH = Y \ast H , \ YX = Y \ast X , \HX = H \ast X.
\]

To determine the product structure in \( \mathcal{D}(G, \mathbb{Z}_p) \) is to understand first the product of monomials in a different order.

Consider first the product \( HY \). It suffices to compute, in \( G \), the product \( \mu_0 \mu_- = \delta(h_0)\delta(u_0) \), say. We compute \( h_0u_0h_0^{-1} \).

Since

\[
\left(\begin{array}{cc}
t & \ t^{-1} \\
\ t^{-1} & \ x \\
\end{array}\right) \left(\begin{array}{cc}
1 & \ 1 \\
\ x^-1 & \ t \\
\end{array}\right) = \left(\begin{array}{cc}
1 & \ t^{-2} \\
\ t^{-2} & \ 1 \\
\end{array}\right),
\]

Documenta Mathematica 16 (2011) 545–559
we have \( h_0 u_0 h_0^{-1} = u_0^{(1+p)^{-2}} \) if we write the group \( N^+ \) multiplicatively. The equation

\[
\mu_0 \mu_- = \delta(h_0 u_0 h_0^{-1}) \delta(h_0),
\]
and the fact that \( D(N^-, \mathbb{Z}_p) \cong \mathbb{Z}_p[[Y]] \) is a homomorphism, show that

\[
(1.5) \quad (1 + H)(1 + Y) = (1 + Y)^q (1 + H)
\]
where we have set

\[
(1.6) \quad q = (1 + p)^{-2} \equiv 1 \pmod{p}.
\]

Similarly consider \( XH \). Let \( n_0 \) be the generator of \( N^+ \). Now \( \delta(n_0) \delta(h_0) \) reduces to \( h_0^{-1} n_0 h_0 \). Again

\[
\begin{pmatrix} t^{-1} \\ t \end{pmatrix} \begin{pmatrix} 1 & x \\ 1 & t^{-1} \end{pmatrix} \begin{pmatrix} 1 & t^{-2}x \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ t \end{pmatrix},
\]
so \( h_0^{-1} n_0 h_0 = n_0^{(1+p)^{-2}} = n_0^q \), whence

\[
(1.7) \quad (1 + X)(1 + H) = (1 + H)(1 + X)^q.
\]

Finally, to express \( XY \) we have to decompose

\[
n_0 u_0 = \begin{pmatrix} 1 \\ p \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ p \\ 1 \end{pmatrix} = \begin{pmatrix} 1 + p^2 & p \\ p & 1 \end{pmatrix}.
\]

Since

\[
\begin{pmatrix} 1 \\ a \\ 1 \end{pmatrix} \begin{pmatrix} t \\ t^{-1} \end{pmatrix} \begin{pmatrix} 1 \\ b \\ 1 \end{pmatrix} = \begin{pmatrix} t \\ ta \\ abt + t^{-1} \end{pmatrix},
\]
we see that

\[
p = ta = tb
\]
with

\[
1 + p^2 = t = (1 + p)^P, \quad P \in \mathbb{Z}_p.
\]

This yields, since \( t_0 = 1 + p \) is the parameter of \( h_0 \):

\[
(1.8) \quad (1 + X)(1 + Y) = (1 + Y)^Q (1 + H)^P (1 + X)^Q
\]
with

\[
(1.9) \quad Q = (1 + p^2)^{-1} \equiv 1[p^2], \quad P = \frac{\log(1 + p^2)}{\log(1 + p)}.
\]

For \( p > 2 \), we have

\[
\log(1 + p) = p - \frac{p^2}{2} + \frac{p^3}{3} \cdots = p(1 + O(p))
\]
\[
\log(1 + p^2) = p^2(1 + O(p^2))
\]
whence

\[(1.10) \quad P = p(1 + O(p)).\]

Note that we have simply written \(HY\) for \(H \ast Y\), etc... This will cause no confusion if we remember that a product such as \(HY\), for variables not in the natural order, is not given by the ostensible product of monomials in the expression (1.3).

To summarize, we have:

**Proposition 1.1.** Set \(Q = (1 + P^2)^{-1}\), \(q = (1 + p)^{-2}\), \(P = \log(1 + P^2)\), \(log(1 + P)\). Then the elements \(X, Y, H\) of \(D(G, \mathbb{Z}_p)\) verify the relations

\[(a) \quad (1 + H)(1 + Y) = (1 + Y)^q(1 + H)\]

\[(b) \quad (1 + X)(1 + H) = (1 + H)(1 + X)^q\]

\[(c) \quad (1 + X)(1 + Y) = (1 + Y)^Q(1 + H)^P(1 + X)^Q.\]

Consider now the universal, non–commutative \(p\)–adic algebra in the variables \(Y, H, X\): thus 

\[A = \mathbb{Z}_p\{Y, H, X\}\]

is composed of all the non–commutative series

\[(1.11) \quad f = \sum_{n \geq 0} \sum_i a_i x^i\]

where the coefficients \(a_i \in \mathbb{Z}_p\) and, for all \(n \geq 0\), \(i\) runs over all maps \(\{1, 2, \ldots n\} \rightarrow \{1, 2, 3\}\); we set \(x_1 = Y, x_2 = H, x_3 = X\) and \(x^i = x_{i(1)} \cdots x_{i(n)}\). The topology on \(A\) is the product topology on \(\prod_{n} \mathbb{Z}_p^{I(n)}\) where \(I(n)\) is the set of maps (\(\equiv\) of non–commutative monomials of degree \(n\)). The algebra \(A\) has a maximal ideal \(\mathcal{M}_A\) generated by \((p, x_1, x_2, x_3)\) and a prime ideal \(\mathcal{P}_A\) generated by \((x_1, x_2, x_3)\). Its topology is given by the powers of \(\mathcal{M}_A\).

The non–commutative polynomial algebra

\[A = \mathbb{Z}_p\{Y, H, X\}\]

is a dense subalgebra of \(A\).

Let \(R\) be the closed two–sided ideal generated in \(A\) by the relations \((a, b, c)\).

Our main result is

**Theorem 1.2.** The Iwasawa algebra \(\Lambda_G\) is naturally isomorphic to \(A/R\).

The proof will in fact rely on the corresponding result with coefficients in \(\mathbb{F}_p\). So let \(\Omega_G = \Lambda_G \otimes_{\mathbb{Z}_p} \mathbb{F}_p\) be the Iwasawa algebra with finite coefficients, 

\[\overline{A} = \mathbb{F}_p\{Y, H, X\}\]

the algebra of non–commutative series with coefficients in \(\mathbb{F}_p\), with its natural linearly compact topology, given by its maximal ideal \(\mathcal{M}_{\overline{A}}\).

Let \(\overline{R}\) be the image of \(R\) in \(\overline{A}\).
Lemma 1.3. \( \overline{R} \) is the closed two-sided ideal generated in \( \overline{A} \) by the image of the relations \((a,b,c)\).

Proof. Denote by \( \mathcal{I} \subset A \) the ideal generated by the relations; let \( \mathcal{J} \subset \overline{A} \) be the similar ideal. Then \( \mathcal{J} \) is obviously the image of \( \mathcal{I} \) in \( \overline{A} \); we denote it by \( \overline{\mathcal{I}} \).

Let \( \overline{R} \) be the reduction of \( \overline{\mathcal{I}} \), and consider the closure \( cl(\overline{\mathcal{I}}) \) of \( \overline{\mathcal{I}} \) in \( \overline{A} \); we denote it by \( \overline{\mathcal{I}} \).

Let \( R \) be the reduction of \( R \), and consider the closure \( cl(I) \) of \( I \) in \( A \). If \( f \in R \), we have \( f = \lim f_n \) for the topology given by \( (M \subseteq N) \). This implies that \( f = \lim f_n \) for the topology given by \( M \subseteq N \) on \( A \), thus \( f \in cl(I) \).

Conversely assume \( f_n \in I \subset R \). Then \( f_n \) is the reduction of a series \( f_n \in I \subset R \). Since \( R \) is closed and \( A \) compact, we may assume that \( f_n \) converges to \( g \in R \). Then, by definition of the topologies, \( f = \lim f_n = g \). Thus \( cl(I) = R \), which finishes the proof.

Theorem 1.4. The Iwasawa algebra \( \mod p, \Omega_G \), is naturally isomorphic to \( A/\overline{R} \).

The proof of these results will occupy \( \S \) 2, 3.

2

We consider the natural map

\[ A \rightarrow \Lambda_G \]

given by the universal property of \( A \). Note that the topology of \( \Lambda_G \), as a distribution algebra, coincides with its topology when it is seen as the algebra of distributions on the commutative group \( \mathbb{Z}_p^3 \). In particular a basis of neighbourhoods of 0 is given by the family of \( \mathbb{Z}_p \)-modules \( M^N_{\Lambda} \), where

\[ (2.1) \quad M^N_{\Lambda} = \{ \lambda \in \Lambda_G, \lambda = \sum n \lambda_n Y^{n_1} H^{n_2} X^{n_3}, v(\lambda_n) + |n| \geq N \} \]

with the usual notation \( |n| = n_1 + n_2 + n_3 \). For a linear monomial \( x = Y, H \) or \( X \), we have \( w(x) = 1 \), \( w \) being the function on \( \Lambda \) given by

\[ (2.2) \quad w(\lambda) = \inf_n (v(\lambda_n) + |n|) . \]

We will use the following deep result of Lazard:

Proposition 2.1 (Lazard). The valuation \( w \) is additive : \( w(\lambda * \mu) = w(\lambda) + w(\mu) \) \( (\lambda, \mu \in \Lambda_G) \).

Cf. [4, III 2.3.3]. Lazard proves, in fact, that the associated graded ring is an enveloping algebra, thus an integral domain, and this implies the additivity. I am indebted to the paper of Schneider and Teitelbaum [6] for a lucid exposition of Lazard’s results.

In fact, it follows from Lazard’s results that \( M^N_{\Lambda} \) is indeed the \( N \)-th power of the maximal ideal \( M_{\Lambda} \) of \( \Lambda_G \). Indeed, let \( J_N \) be defined by \( w(\lambda) \geq N \). It is
easy to check that $J_1 = \mathcal{M}_A$. The additivity implies that $\mathcal{M}_A^N$ is contained in $J_N$. Since every linear monomial belongs to the maximal ideal, the expression (2.1) implies the converse inclusion since $\mathcal{M}_A^N$ is closed. Consider now the filtration of $A$ by the powers of its maximal ideal. It is defined by a valuation $w_A$ given by a formula similar to (2.2) : if

$$f = \sum_i a_i x^i,$$

$$w_A(f) = \inf_i (v(a_i) + |i|)$$

where $|i| = n$ is the degree of $i$ (cf. after (1.11)). We now have the following (“ideal” means two–sided ideal unless otherwise indicated).

**Proposition 2.2.** The natural map $\varphi : A \to \Lambda_G$ extends continuously to a surjective homomorphism $A \to \Lambda_G$. In fact,

$$\varphi(\mathcal{M}_A^N) \subseteq \mathcal{M}_\Lambda^N \quad (N \geq 0).$$

**Proof :** The continuity is implied by the stronger property

$$w(\varphi(x^i)) = n = |i|$$

where $n$, as after (1.11), is the degree of the monomial. By induction on $n$, this follows from Proposition 2.1. If $f \in \mathcal{M}_A^N$, we have $w_A(f) \geq N$ and the continuity follows from (2.3) by $\mathbb{Z}_p$–linearity. The surjectivity follows from the fact that $\varphi$ is already surjective if $A$ is replaced by the set of linear combinations of well-ordered monomials ($i$ increasing).

**Corollary 2.3.** There is a natural, continuous surjection

$$\mathcal{B} = A/R \to \Lambda_G.$$

**Corollary 2.4.** There is a continuous surjection

$$\overline{\varphi} : \overline{\mathcal{B}} = \overline{A}/\overline{R} \to \Omega_G.$$

This follows from Lemma 1.3.

It follows from Abelian distribution theory that $\Omega_G$ is, as a space, isomorphic to

$$\mathbb{F}_p[[Y, H, X]]$$

with the compact topology. An obvious computation shows that

$$\mathcal{M}_\Omega^N = \{ \lambda \in \Omega_G : v_\Omega(\lambda) \geq N \},$$

$v_\Omega$ being the usual valuation on power series, is the image of $\mathcal{M}_A^N$. In particular it is an ideal ; for $N = 1$, $\mathcal{M}_\Omega$ is the maximal (two–sided) ideal, and $(\mathcal{M}_\Omega)^N \subseteq \mathcal{M}_\Omega$. (Reduce mod $p$ the corresponding property for $\Lambda.$)
Similarly in $\mathcal{A}$, we find that the reduction mod $p$ (image in $\mathfrak{A}$) of $\mathcal{M}_N^\mathfrak{A}$ is the ideal of series
$$\mathcal{F} = \sum_i \alpha_i \cdot x^i \quad (\alpha_i \in \mathbb{F}_p)$$
such that $|i| \geq N$. For $N = 1$ we obtain the maximal ideal in $\mathfrak{A}$. Furthermore in this case too $(\mathcal{M}_N^\mathfrak{A})^N = \mathcal{M}_N^\mathfrak{A}$.

3

In this paragraph we will directly study the quotient algebra $\overline{\mathcal{B}} = \overline{\mathcal{A}} / \overline{\mathcal{R}}$, using the properties of the relations $(a,b,c)$.

Consider the natural filtration of $\mathcal{A}$ by the powers of $M^\mathcal{A}$, which we denote by $F_n^\mathcal{A}$. We have $F_n^\mathcal{A} / F_{n+1}^\mathcal{A} = \mathfrak{A}^I(n)$ where $I(n)$ is the set of maps $\{1, \ldots, n\} \to \{1, 2, 3\}$ (§1). The filtration $F_n^\mathcal{B}$ induces a filtration on $\overline{\mathcal{B}} = \overline{\mathcal{A}} / \overline{\mathcal{R}}$:
$$F_n^\overline{\mathcal{B}} = F_n^\mathcal{A} + \overline{\mathcal{R}}$$
whence a graduation
$$gr^{n} \overline{\mathcal{B}} = F_n^\mathcal{A} / F_{n+1}^\mathcal{A} + \overline{\mathcal{R}}$$

Let $S_n = S_n(X,Y,Z)$ be the space of commutative polynomials over $\mathbb{F}_p$ of degree $n$ ; thus $\dim S_n = \frac{(n+1)(n+2)}{2}$. Let $\Sigma_n$ be the space of homogeneous non-commutative polynomials of degree $n$ ; thus $\Sigma_n \to F_n^\mathcal{A} / F_{n+1}^\mathcal{A}$, and therefore $\Sigma_n \to gr^n \overline{\mathcal{B}}$, is surjective.

**Proposition 3.1.** $\dim gr^n \overline{\mathcal{B}} \leq \dim S_n$.

In order to prove this we consider the relations defining $\mathcal{R}$ (or rather $\mathcal{R}$). Consider first the relation (a) :
$$(1 + H)(1 + Y) = (1 + Y)^q(1 + H)$$
with $q \equiv 1 \ [p]$. Expanding the power series gives
$$1 + H + Y + HY = (1 + qY + \left(\frac{q}{2}\right)Y^2 + \cdots)(1 + H).$$
We note that $\left(\frac{q}{2}\right) = \frac{q(q-1)}{2} \equiv 0 \ [p]$. Thus in $\overline{\mathcal{A}} / \overline{\mathcal{R}}$ :
$$1 + H + Y + HY = (1 + qY)(1 + H) + R(Y)(1 + H),$$
the term $R(Y)$ being of degree $\geq 3$, so
$$HY = (q - 1)Y + qYH + R_1(Y,H)$$
$$= YH + R_1(Y,H)$$
since \( q \equiv 1 \), \( R_1(Y, H) \) of degree \( \geq 3 \). This shows that in \( \mathcal{B} = \mathcal{A}/\mathcal{K} \):

\[
(3.1) \quad HY = YH \mod F^3\mathcal{B} \quad \text{i.e.} \quad HY = YH \mod gr^2\mathcal{B}.
\]

The computation for relation (b) is obviously similar, yielding in \( \mathcal{B} \)

\[
(3.2) \quad XH = HX \mod F^3\mathcal{B}.
\]

Consider now the identity (c):

\[
(1 + X)(1 + Y) = (1 + Y)^Q(1 + H)^P(1 + X)^Q.
\]

We have \( Q \equiv 1 \mod p^2 \), \( P \equiv p \mod p^2 \). Again the coefficients \( \frac{Q(Q - 1)}{2} \) of \( Y^2, X^2 \) in the power series vanish \( \mod p \). Modulo \( \mathcal{A} \mathcal{M} \), whose image is in \( F^3\mathcal{B} \), we then have

\[
(1 + X)(1 + Y) \equiv (1 + QY)(1 + H)^P(1 + QX).
\]

Since \( P \equiv p \mod p^2 \) and since \( 2 \) is invertible, \( (1 + H)^P \equiv 1 \mod (p, H^3) \). Thus

\[
1 + X + Y + XY \equiv 1 + QX + QY + Q^2YX \quad \mod F^3\mathcal{B},
\]

and since \( Q \equiv 1 \):

\[
(3.3) \quad XY \equiv YX \quad \mod F^3\mathcal{B}.
\]

Since \( gr^2\mathcal{B} \) is generated by these three monomials and the squares \( Y^2, H^2, X^2 \), the identities (3.1)–(3.3) show that \( \dim gr^2\mathcal{B} \leq 6 \), whence the result for \( n = 2 \). The proposition for general \( n \) is deduced from this case. Consider an arbitrary monomial of degree \( n \),

\[ x^i = x_{i_1} \ldots x_{i_n}. \]

The following lemma is obvious:

**Lemma 3.2.** We can change \( x^i \) into a well-ordered monomial \( x^{i'} \) (\( i' \) increasing) by a sequence of transpositions \( x_{i_n}x_{i_{n+1}} \mapsto x_{i_{n+1}}x_{i_n} \).

(Consider the set of inversions \( \{ \alpha < \beta : i_\alpha > i_\beta \} \). Assume \( i_\gamma > i_{\gamma+1} \), and replace in \( x^i \) the term \( x_{i_\gamma}x_{i_{\gamma+1}} \) by \( x_{i_{\gamma+1}}x_{i_\gamma} \). It is easy to check that the set of inversions decreases by one element.)

We now write \( x^i = x^jx_{i_n}x_{i_{n+1}}x^{j'} \). We will prove by induction

**Lemma 3.3.** In \( \mathcal{B} \), \( x^i \equiv x^{i'} \mod F^{n+1}\mathcal{B} \), where \( i' \) is well-ordered.

But this is now equally obvious. Let \( r, s \) be the degrees of \( x^j, x^{j'} \), so \( n = r + s + 2 \). Then \( x^i \equiv x^j[F^{r+1}\mathcal{B}], x^{j'} \equiv x^{j'}[F^{s+1}\mathcal{B}] \) and \( x_{i_n}x_{i_{n+1}} \equiv x_{i_{n+1}}x_{i_n}[F^3\mathcal{B}] \); we are of course assuming \( i_\alpha > i_{\alpha+1} \). Factoring the congruences gives
Laurent Clozel

$x^i \equiv x'^i x_{i,n+1} x_{i,n} x'^i [F^{n+1}B]$ since the filtration $F^n$, image of $F^n$ on $\mathcal{A}$, verifies $F^n F^m \subset F^{n+m}$. Using induction if necessary, we obtain the Lemma, whence Proposition 3.1.

**Proof of Theorem 1.4.**– The natural map $\varphi : \mathcal{A} \to \Lambda_G$ sends $M^\bullet$ to $M^\bullet_G$. Since $F^\bullet$ is on $\mathcal{B}$ the filtration inherited from the natural filtration on $\mathcal{A}$, we see that $\varphi$ sends $F^n B$ to $M^n_G$. We then have a natural map $gr \varphi : gr^\bullet B \to gr^\bullet \Omega_G$, surjective since $\varphi$ is so. It is an isomorphism since $\dim \ gr^n B \leq \ dim \ S_n = \ dim \ gr^n \Omega_G$. (The last equality follows from the considerations after Cor. 2.4; cf also [3, Theorem 7.24]). Therefore $\varphi$ is isomorphic since the filtration on $\mathcal{B}$ is complete. The last point follows from the fact that $\mathcal{B} = \mathcal{A}/R$ where $\mathcal{R}$ is closed and therefore complete for the filtration induced from that of $\mathcal{A}$: see e.g. [5, Thm 4 (5) p. 31].

**Proof of Theorem 1.2.**– The reduction of $\varphi : \mathcal{A}/R \to \Lambda_G$ is $\varphi$. Recall that $\mathcal{R}$ is the image of $\mathcal{R}$ in $\mathcal{A}$. Assume $f \in \mathcal{A}$ satisfies $\varphi(f) = 0$. We then have $f \in \mathcal{R}$ by Theorem 1.3, so $f = r_1 + pf_1$, $r_1 \in \mathcal{R}$, $f_1 \in \mathcal{A}$. Then $\varphi(f_1) = 0$. Inductively, we obtain an expression $f = r_n + p^n f_n$ of the same type. Since $p^n f_n \to 0$ in $\mathcal{A}$ and $\mathcal{R}$ is closed, we see that $f \in \mathcal{R}$, QED.

In this section, we show that the description of $\Lambda_G$ given in § 1 allows one to give different proofs of some results of Ardakov and to understand them in terms of the growth of coefficients in the Iwasawa expansion.

Ardakov’s main result in [1] is that the centre of the Iwasawa algebra reduces to the Iwasawa algebra of the centre of $G$, trivial in our case. We will see that the fact of being central is incompatible with the boundedness of the Iwasawa coefficients.

It will be instructive to compare this behaviour with what happens for the centre of the enveloping algebra. Recall that instead of the Iwasawa distributions, or measures, we can consider the analytic distributions (or hyperfunctions), dual to the locally analytic functions on $G$ (cf. Schneider–Teitelbaum [6]). They admit an expansion (1.3), but with now

\[
|\lambda_n| r^n \to 0 \quad \forall r < 1, \ |n| = n_1 + n_2 + n_3.
\]

Among these we have the Casimir operator (seen as a distribution with support at 1)

\[
\omega = h^2 + 2(xy + yx) = h^2 - 2h + 4xy
\]

(cf. e.g. Borel [2, p. 19]) where $h, x, y$ are the infinitesimal generators of the groups $T, N^+, N^-$. It suffices to compute $\omega$ on a function $f$ given by

\[
f(utn) = (1 + Y)^{x_1} (1 + H)^{x_2} (1 + X)^{x_3}
\]
where \( u, t, n \) have parameters \( x_1, x_2, x_3 \in \mathbb{Z}_p \) and \( Y, H, X \) belong to the disc \(|w| < 1\) in \( \mathbb{C}_p \) or even \( \mathbb{Q}_p \) (such functions are dense). Now

\[
(xyf)(1) = \frac{d}{dt} \left| \frac{d}{ds} f(e^{xy}e^{tx}) \right| = \frac{d}{dt} \left| \frac{d}{ds} \left( \begin{pmatrix} 1 & t \\ s & 1 \end{pmatrix} \right) \right| \\
= \frac{d^2}{dsdt} \left| (1 + Y)^{s/p} (1 + X)^{t/p} \right| \\
= \frac{1}{p^2} \log(1 + Y) \log(1 + X),
\]

\[
hf(1) = \frac{d}{dt} \left| \frac{d}{ds} f(e^{t} e^{-t}) \right| \\
= \frac{1}{\log(1 + p)} \frac{d}{dt} \left| \frac{d}{ds} \left( (1 + p)^t \right)^{-} \right| \\
= \frac{1}{\log(1 + p)} \log(1 + H),
\]

\[
h^2f(1) = \frac{1}{\log^2(1 + p)} \frac{d^2}{dp^2} \left| \frac{d}{ds} f((1 + p)^t) \right| \\
= \frac{1}{\log^2(1 + p)} \log(1 + H),
\]

Thus the Amice transform of \( \omega \) is

\[
F(Y, H, X) = \\
= \frac{1}{\log^2(1 + p)} \log^2(1 + H) - \frac{2}{\log(1 + p)} \log(1 + H) + \frac{4}{p^2} \log(1 + Y) \log(1 + X).
\]

This obviously has an expansion (4.1) – and is an element of the ring of convergent series on \( D(1)^4, D(1) \subset \mathbb{Q}_p \) being the open unit disc – but it is not an element of \( \Lambda_G \).

We will see that the invariance under \( T \) suffices to impose such a logarithmic behaviour. This leads to:

**Theorem 4.1.** The space of elements on \( \Lambda_G \) invariant by conjugation under \( T \) is equal to the Iwasawa algebra \( \Lambda_T \subset \Lambda_G \).

Assume indeed \( \lambda \in \Lambda_G \) is \( T \)-invariant, with Amice transform

\[
F(Y, H, X).
\]

We have \( Y = u_0 - 1 \), with \( h_0 u_0 h_0^{-1} = u_0(1 + p)^{-2} \); thus the automorphism \( Ad(h_0) \) of \( G \) sends \( 1 + Y \) to \( (1 + Y)^{(1 + p)^{-2}} \). Similarly, \( h_0 u_0 h_0^{-1} = u_0(1 + p)^2 \), so \( 1 + X \)
Laurent Clozel

is sent to $(1 + X)^{(1+p)^2}$. Of course $H$ is left invariant. If $\lambda$ is $T$–invariant we therefore have

$$F(Y, H, X) = F(Y', H, X')$$  \hspace{5cm} (4.2)

where $1 + Y' = (1 + Y)^{(1+p)^{-2}}$, $1 + X' = (1 + X)^{(1+p)^2}$. Since $p \not= 2$, $(1 + p)^2$ is a topological generator of $1 + \mathbb{Z}_p$. Therefore (4.2) remains true if

$$1 + Y' = (1 + Y)^u, \ 1 + X' = (1 + X)^{u^{-1}}, \ u \in 1 + p\mathbb{Z}_p.$$  \hspace{5cm} (4.3)

In the following computations consider $F$ as an element of the Lazard ring in three variables. If we fix a value of $H$ in $\mathbb{C}_p$ such that $|H| < 1$, say $H_0$, $F(Y, H_0, X) := F_1(Y, X)$ becomes an Iwasawa series in the two variables, still invariant under (4.3). Now set

$$U = \log(1 + Y), \ V = \log(1 + X),$$

two series convergent in $D(1)$. We have

$$F_1(Y, X) = G_1(U, V)$$

where $G_1$ converges absolutely in the domain of convergence of the exponential, i.e., for $|U|, |V| < r_0 = p^{-1}$. Moreover $G_1$ is invariant by $U \mapsto uU, \ V \mapsto u^{-1}V, \ |u - 1| < p^{-1}$. This implies that

$$G_1(U, V) = G_2(UV)$$

with $G_2(z)$ convergent for $|z| < r_0^2$.

Let

$$G_2(z) = \sum_{0}^{\infty} b_n z^n,$$

$$F_1(Y, X) = \sum_{m,n} a_{mn} Y^m X^n \ (|a_{mn}| \leq 1).$$

Then

$$F_1(Y, X) = G_2(\log(1 + Y) \log(1 + X)),$$

$$\log(1 + Y) = Y \sum_{0}^{\infty} \frac{(-1)^k}{k + 1} Y^k := Y L_1(Y)$$

$$\log(1 + X) = X \sum_{0}^{\infty} \frac{(-1)^\ell}{\ell + 1} X^\ell := X L_1(X)$$

Thus $(\log(1 + Y) \log(1 + X))^q$ contains only terms the degree of which in $Y$ and $X$ is at least $q$. We have of course $b_0 = a_0$, and the previous remark implies that

$$\sum_{n \geq 0} a_{1n} Y^n + \sum_{m \geq 0} a_{m1} Y^m X^n \hspace{5cm} \text{Documenta Mathematica 16 (2011) 545–559}$$
is identical with the sum of terms of these degrees in
\[ b_1 YX L_1(Y) L_1(X), \]
i.e. with
\[ b_1 YX (L_1(Y) + L_1(X) - 1). \]
Since the \( a_{mn} \) are integral, this implies that \( b_1 = 0 \) as the denominators in the
log–series are not bounded.
By induction assume that \( b_1 = \cdots b_{N-1} = 0 \), so
\[ G_2 = \sum_{N} b_q z^q. \]
We then find that
\[ F_1(Y, X) = b_N Y^N X^N L_1(Y)^N L_1(X)^N \]
+ terms of degree \( > N \) in \( X \) and \( Y \).
Now
\[ L_1(Y) = 1 + Y M_1(Y), \quad \text{say}, \]
\[ L_1(X) = 1 + X M_1(X) \]
so (4.4) implies that
\[ F_1(Y, X) = b_N Y^N X^N (1 + NY M_1(Y) + XM_1(X)) \]
+ terms of degree \( > N \) in \( X \) and \( Y \).
Since \( M_1 \) does not have bounded denominators, we deduce that \( b_N = 0 \).
Finally we have proved that \( F_1 = b_0 \), i.e. \( F(Y, H, X) \equiv b_0(H) \) for any \( H \in \mathbb{C}_p, |H| < 1 \). This implies that \( F(Y, H, X) = F(H) \) has no terms involving \( X \) or \( Y \),
whence the result.

**Corollary 4.2.** The centre of \( \Lambda_G \) is composed of the multiples of the Dirac
measure at 1.

For assume that \( \lambda \in \Lambda_G \) is central, so invariant by all conjugates of \( T \) in \( G \).
By Thm. 4.1 its support is contained in the intersection of the tori \( g T g^{-1} \)
\( (g \in G) \). This intersection is reduced to \{1\}.
We note that Theorem 4.1 itself follows from Ardakov’s results [1, Proposition 2.2]: a simple computation shows that the only finite orbits of \( T \) in \( G \) are the
elements of \( T \) (use the triangular decomposition).

5

This section is devoted to conjectural remarks on a formal extension of the
main result.
Consider the formulas of Proposition 1, for example
(a) \((1 + H)(1 + Y) = (1 + Y)^{(1+p)^{-2}}(1 + H)\)

(c) \((1 + X)(1 + Y) = (1 + Y)^{(1+p^2)^{-1}}(1 + H)^{\frac{\log(1+p^2)}{1+p^2}}(1 + X)^{(1+p^2)^{-1}}\).

In the \(p\)-adic computation the series for, say, \((1 + X)^x (x \in \mathbb{Z}_p)\) converges as an Iwasawa expansion because of the integrality of the binomial function \(\binom{x}{n}\).

However, replace now \(\Lambda_G\) by \(k[[Y,H,X]]\) where \(k\) is a field of characteristic zero. Set \(p = \varepsilon\), another formal variable, which should however be considered as a small parameter. The binomial coefficients, namely

\[
\binom{(1 + \varepsilon)^{-2}}{n} = \frac{(1 + \varepsilon)^{-2}((1 + \varepsilon)^{-2} - 1) \cdots ((1 + \varepsilon)^{-2} - n + 1)}{n!}
\]

and similarly

\[
\frac{\log(1 + \varepsilon^2)}{\log(1 + \varepsilon)}
\]

are well-defined series in \(k[[\varepsilon]]\). Formulas \((a,b,c)\) therefore define the products \(HY, XH\) and \(XY\) in \(k[[\varepsilon]][[Y,H,X]]\). The \(p\)-adic results do not seem to imply that this extends to an associative product in this ring of power series. Note that if it were so, equations \((a,b,c)\) at \(\varepsilon = 0\) would simply yield \(HY = YH\), \(XH = HX\) and \(XY = YX\). Such an extension would therefore define, quite naturally, a formal deformation of the algebra of power series \(k[[Y,H,X]]\) associated to the group \(SL(2)\). It would be interesting to understand this deformation in group-theoretic terms (or in terms of the Lie algebra) — assuming, of course, it exists. In this respect one should note that formulas \((a,b)\) allow one to define inductively the products \(H^nY^m\) and \(X^nH^m\). However I do not see how to define \(X^nY^m\), even granting \((c)\).

References

[1] K. Ardakov, *The centre of completed algebras of pro-\(p\) groups*, Documenta Math. 9 (2004), 599–606.
[2] A. Borel, *Automorphic forms on \(SL(2,\mathbb{R})\)*, Cambridge Univ. Press, 1997.
[3] J.D. Dixon, M.P.F. Du Sautoy, A. Mann, D. Segal, *Analytic pro-\(p\) groups*, 2nd edition, Cambridge University Press (1999).
[4] M. Lazard, *Groupes analytiques \(p\)-adiques*, Publ. IHES 26 (1965), 389–603.
[5] Li Huishi, F. van Oystaeyen, *Zariskian filtrations*, Kluwer, 1997.
[6] P. Schneider, J. Teitelbaum, *Algebras of \(p\)-adic distributions and admissible representations*, Inv. Math. 153 (2003), 145–196.
[7] J.-P. Serre, *Algèbres de Lie semi-simples complexes*, Benjamin, 1966.

[8] R. Steinberg, *Lectures on Chevalley groups*, mimeographed notes, Yale University, 1967.

[9] L. Washington, *Introduction to cyclotomic fields*, Springer, 1982.

Laurent Clozel
Mathématique
Université Paris-Sud 11
Bât. 425
91405 Orsay CEDEX
France
ON THE EXISTENCE OF STATIONARY SOLUTIONS FOR SOME NON-FREDHOLM INTEGRO-DIFFERENTIAL EQUATIONS

VITALI VOUGALTER, VITALY VOLPERT

Received: February 4, 2011
Communicated by Heinz Siedentop

ABSTRACT. We show the existence of stationary solutions for some reaction-diffusion type equations in the appropriate $H^2$ spaces using the fixed point technique when the elliptic problem contains second order differential operators with and without Fredholm property.

2010 Mathematics Subject Classification: 35J10, 35P30, 35K57
Keywords and Phrases: solvability conditions, non Fredholm operators, integro-differential equations, stationary solutions

1 INTRODUCTION

Let us recall that a linear operator $L$ acting from a Banach space $E$ into another Banach space $F$ satisfies the Fredholm property if its image is closed, the dimension of its kernel and the codimension of its image are finite. As a consequence, the equation $Lu = f$ is solvable if and only if $\phi_i(f) = 0$ for a finite number of functionals $\phi_i$ from the dual space $F^*$. These properties of Fredholm operators are widely used in many methods of linear and nonlinear analysis.

Elliptic problems in bounded domains with a sufficiently smooth boundary satisfy the Fredholm property if the ellipticity condition, proper ellipticity and Lopatinskii conditions are satisfied (see e.g. [1], [9], [10]). This is the main result of the theory of linear elliptic problems. In the case of unbounded domains, these conditions may not be sufficient and the Fredholm property may not be satisfied. For example, Laplace operator, $Lu = \Delta u$, in $\mathbb{R}^d$ does not satisfy the Fredholm property when considered in Hölder spaces, $L : C^{2+\alpha}(\mathbb{R}^d) \to C^{\alpha}(\mathbb{R}^d)$, or in Sobolev spaces, $L : H^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$.

Linear elliptic problems in unbounded domains satisfy the Fredholm property if and only if, in addition to the conditions cited above, limiting operators are
invertible (see [11]). In some simple cases, limiting operators can be explicitly constructed. For example, if

\[ Lu = a(x)u'' + b(x)u' + c(x)u, \quad x \in \mathbb{R}, \]

where the coefficients of the operator have limits at infinity,

\[ a_\pm = \lim_{x \to \pm \infty} a(x), \quad b_\pm = \lim_{x \to \pm \infty} b(x), \quad c_\pm = \lim_{x \to \pm \infty} c(x), \]

the limiting operators are:

\[ L_\pm u = a_\pm u'' + b_\pm u' + c_\pm u. \]

Since the coefficients are constant, the essential spectrum of the operator, that is the set of complex numbers \( \lambda \) for which the operator \( L - \lambda \) does not satisfy the Fredholm property, can be explicitly found by means of the Fourier transform:

\[ \lambda_\pm(\xi) = -a_\pm \xi^2 + b_\pm i\xi + c_\pm, \quad \xi \in \mathbb{R}. \]

Invertibility of limiting operators is equivalent to the condition that the essential spectrum does not contain the origin.

In the case of general elliptic problems, the same assertions hold true. The Fredholm property is satisfied if the essential spectrum does not contain the origin or if the limiting operators are invertible. However, these conditions may not be explicitly written.

In the case of non-Fredholm operators the usual solvability conditions may not be applicable and solvability conditions are, in general, not known. There are some classes of operators for which solvability conditions are obtained. Let us illustrate them with the following example. Consider the equation

\[ Lu \equiv \Delta u + au = f \tag{1.1} \]

in \( \mathbb{R}^d \), where \( a \) is a positive constant. The operator \( L \) coincides with its limiting operators. The homogeneous equation has a nonzero bounded solution. Hence the Fredholm property is not satisfied. However, since the operator has constant coefficients, we can apply the Fourier transform and find the solution explicitly. Solvability conditions can be formulated as follows. If \( f \in L^2(\mathbb{R}^d) \) and \( xf \in L^1(\mathbb{R}^d) \), then there exist a solution of this equation in \( H^2(\mathbb{R}^d) \) if and only if

\[ \left( f(x), \frac{e^{ipx}}{(2\pi)^\frac{d}{2}} \right)_{L^2(\mathbb{R}^d)} = 0, \quad p \in S^d_{\sqrt{a}} \text{ a.e.} \]

(see [19]). Here and further down \( S^d_r \) denotes the sphere in \( \mathbb{R}^d \) of radius \( r \) centered at the origin. Thus, though the operator does not satisfy the Fredholm property, solvability conditions are formulated in a similar way. However, this similarity is only formal since the range of the operator is not closed.

In the case of the operator with a potential,
Lu ≡ ∆u + a(x)u = f,

Fourier transform is not directly applicable. Nevertheless, solvability conditions in \( \mathbb{R}^3 \) can be obtained by a rather sophisticated application of the theory of self-adjoint operators (see [13]). As before, solvability conditions are formulated in terms of orthogonality to solutions of the homogeneous adjoint equation. There are several other examples of linear elliptic operators without Fredholm property for which solvability conditions can be obtained (see [11]-[19]).

Solvability conditions play an important role in the analysis of nonlinear elliptic problems. In the case of non-Fredholm operators, in spite of some progress in understanding of linear problems, there exist only few examples where nonlinear non-Fredholm operators are analyzed (see [4]-[6]). In the present article we consider another class of nonlinear equations, for which the Fredholm property may not be satisfied:

\[
\frac{\partial u}{\partial t} = \Delta u + au + \int_{\Omega} G(x-y)F(u(y), y)dy = 0, \quad a \geq 0.
\]

Here \( \Omega \) is a domain in \( \mathbb{R}^d, \quad d = 1, 2, 3 \), the more physically interesting dimensions. In population dynamics the integro-differential equations describe models with intra-specific competition and nonlocal consumption of resources (see e.g. [2], [3], [7]). The linear part of the corresponding operator is the same as in equation (1.1) above. We will use the explicit form of solvability conditions and will study the existence of stationary solutions of the nonlinear equation.

2 Formulation of the results

The nonlinear part of equation (1.2) will satisfy the following regularity conditions.

**Assumption 1.** Function \( F(u, x) : \mathbb{R} \times \Omega \to \mathbb{R} \) is such that

\[ |F(u, x)| \leq k|u| + h(x) \quad \text{for} \quad u \in \mathbb{R}, \quad x \in \Omega \]  

with a constant \( k > 0 \) and \( h(x) : \Omega \to \mathbb{R}^+, \quad h(x) \in L^2(\Omega) \). Moreover, it is a Lipschitz continuous function, such that

\[ |F(u_1, x) - F(u_2, x)| \leq l|u_1 - u_2| \quad \text{for any} \quad u_{1,2} \in \mathbb{R}, \quad x \in \Omega \]

with a constant \( l > 0 \).

Clearly, the stationary solutions of (1.2), if they exist, will satisfy the nonlocal elliptic equation

\[ \Delta u + \int_{\Omega} G(x-y)F(u(y), y)dy + au = 0, \quad a \geq 0. \]
Let us introduce the auxiliary problem

\[-\Delta u - au = \int_{\Omega} G(x - y)F(v(y), y)dy.\] (2.3)

We denote \((f_1(x), f_2(x))_{L^2(\Omega)} := \int_{\Omega} f_1(x)f_2(x)dx\), with a slight abuse of notations when these functions are not square integrable, like for instance those used in the one dimensional Lemma A1 of the Appendix. In the first part of the article we study the case of \(\Omega = \mathbb{R}^d\), such that the appropriate Sobolev space is equipped with the norm

\[\|u\|_{H^2(\mathbb{R}^d)}^2 := \|u\|_{L^2(\mathbb{R}^d)}^2 + \|\Delta u\|_{L^2(\mathbb{R}^d)}^2.\]

The main issue for the problem above is that the operator \(-\Delta - a : H^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d), a \geq 0\) does not satisfy the Fredholm property, which is the obstacle to solve equation (2.3). The similar situations but in linear problems, both self-adjoint and non self-adjoint involving non Fredholm second or fourth order differential operators or even systems of equations with non Fredholm operators have been studied extensively in recent years (see [13]-[18]). However, we manage to show that equation (2.3) in this case defines a map \(T_\alpha : H^2(\mathbb{R}^d) \to H^2(\mathbb{R}^d), a \geq 0\), which is a strict contraction under certain technical conditions.

**Theorem 1.** Let \(\Omega = \mathbb{R}^d, G(x) : \mathbb{R}^d \to \mathbb{R}, G(x) \in L^1(\mathbb{R}^d)\) and Assumption 1 holds.

I) When \(a > 0\) we assume that \(xG(x) \in L^1(\mathbb{R}^d)\), orthogonality relations (6.4) hold if \(d = 1\) and (6.9) when \(d = 2, 3\) and \(\sqrt{2}(2\pi)^{\frac{d}{2}}N_{\alpha}, d l < 1\). Then the map \(T_\alpha v = u\) on \(H^2(\mathbb{R}^d)\) defined by equation (2.3) has a unique fixed point \(v_\alpha\), which is the only stationary solution of problem (1.2) in \(H^2(\mathbb{R}^d)\).

II) When \(a = 0\) we assume that \(x^2G(x) \in L^1(\mathbb{R}^d)\), orthogonality relations (6.10) hold, \(d = 1, 2, 3\) and \(\sqrt{2}(2\pi)^{\frac{d}{2}}N_0, d l < 1\). Then the map \(T_0 v = u\) on \(H^2(\mathbb{R}^d)\) defined by equation (2.3) admits a unique fixed point \(v_0\), which is the only stationary solution of problem (1.2) with \(a = 0\) in \(H^2(\mathbb{R}^d)\).

In both cases I) and II) the fixed point \(v_\alpha, a \geq 0\) is nontrivial provided the intersection of supports of the Fourier transforms of functions \(\text{supp}\hat{F}(0, x) \cap \text{supp}\hat{G}\) is a set of nonzero Lebesgue measure in \(\mathbb{R}^d\).

In the second part of the work we study the analogous problem on the finite interval with periodic boundary conditions, i.e. \(\Omega = I := [0, 2\pi]\) and the appropriate functional space is

\[H^2(I) = \{u(x) : I \to \mathbb{R} \mid u(x), u''(x) \in L^2(I), u(0) = u(2\pi), u'(0) = u'(2\pi)\}.

Let us introduce the following auxiliary constrained subspaces

\[H_0^2(I) := \{u \in H^2(I) \mid \left(u(x), \frac{e^{\pm in_0 x}}{\sqrt{2\pi}}\right)_{L^2(I)} = 0\}, n_0 \in \mathbb{N}\] (2.4)
and

\[ H^2_{0,0}(I) = \{ u \in H^2(I) \mid (u(x), 1)_{L^2(I)} = 0 \} , \tag{2.5} \]

which are Hilbert spaces as well (see e.g. Chapter 2.1 of [8]). We prove that equation (2.3) in this situation defines a map \( \tau_n \), \( a \geq 0 \) on the above mentioned spaces which will be a strict contraction under our assumptions.

**Theorem 2.** Let \( \Omega = I, \ G(x) : I \to \mathbb{R}, \ G(x) \in L^1(I), \ G(0) = G(2\pi), \ F(u,0) = F(u,2\pi) \) for \( u \in \mathbb{R} \) and Assumption 1 holds.

I) When \( a > 0 \) and \( a \neq n^2, \ n \in \mathbb{Z} \) we assume that \( 2\sqrt{\pi N_1}l < 1 \). Then the map \( \tau_l u = u \) on \( H^2(I) \) defined by equation (2.3) has a unique fixed point \( v_a \), the only stationary solution of problem (1.2) in \( H^2(I) \).

II) When \( a = n_0^2, \ n_0 \in \mathbb{N} \) assume that orthogonality relations (6.17) hold and \( 2\sqrt{\pi N_0}l < 1 \). Then the map \( \tau_{n_0} u = u \) on \( H^2_{0,0}(I) \) defined by equation (2.3) has a unique fixed point \( v_{n_0} \), the only stationary solution of problem (1.2) in \( H^2_{0,0}(I) \).

III) When \( a = 0 \) assume that orthogonality relation (6.18) holds and \( 2\sqrt{\pi N_0}l < 1 \). Then the map \( \tau_{n_0} u = u \) on \( H^2_{0,0}(I) \) defined by equation (2.3) has a unique fixed point \( v_0 \), the only stationary solution of problem (1.2) in \( H^2_{0,0}(I) \).

In all cases I), II) and III) the fixed point \( v_a, \ a \geq 0 \) is nontrivial provided the Fourier coefficients \( G_n F(0,x)_n \neq 0 \) for some \( n \in \mathbb{Z} \).

**Remark.** We use the constrained subspaces \( H^2_{0,0}(I) \) and \( H^2_{0,0}(I) \) in cases II) and III) respectively, such that the operators \(-\frac{d^2}{dx^2} - n_0^2 : H^2_{0,0}(I) \to L^2(I) \) and \(-\frac{d^2}{dx^2} : H^2_{0,0}(I) \to L^2(I) \), which possess the Fredholm property, have empty kernels.

We conclude the article with the studies of our problem on the product of spaces, where one is the finite interval with periodic boundary conditions as before and another is the whole space of dimension not exceeding two, such that in our notations \( \Omega = I \times \mathbb{R}^d = [0,2\pi] \times \mathbb{R}^d, \ d = 1, 2 \) and \( x = (x_1, x_\perp) \) with \( x_1 \in I \) and \( x_\perp \in \mathbb{R}^d \). The appropriate Sobolev space for the problem is \( H^2(\Omega) \) defined as

\[ \{ u(x) : \Omega \to \mathbb{R} \mid u(x), \Delta u(x) \in L^2(\Omega), \]

\[ u(0, x_\perp) = u(2\pi, x_\perp), \ u_{x_1}(0, x_\perp) = u_{x_1}(2\pi, x_\perp) \}, \]

where \( x_\perp \in \mathbb{R}^d \) a.e. and \( u_{x_1} \) stands for the derivative of \( u(x) \) with respect to the first variable \( x_1 \). As in the whole space case covered in Theorem 1, the operator

\[ -\Delta - a \ : \ H^2(\Omega) \to L^2(\Omega), \ a \geq 0 \]

does not possess the Fredholm property. Let us show that problem (2.3) in this context defines a map \( t_a : H^2(\Omega) \to H^2(\Omega), \ a \geq 0 \), a strict contraction under appropriate technical conditions.
Theorem 3. Let $\Omega = I \times \mathbb{R}^d$, $d = 1, 2$, $G(x) : \Omega \to \mathbb{R}$, $G(x) \in L^1(\Omega)$, $G(0, x_{\perp}) = G(2\pi, x_{\perp})$, $F(u, 0, x_{\perp}) = F(u, 2\pi, x_{\perp})$ for $x_{\perp} \in \mathbb{R}^d$ a.e. and $u \in \mathbb{R}$ and Assumption 1 holds.

I) When $\nu^2 < a < (\nu_0 + 1)^2$, $\nu \in \mathbb{N} = \mathbb{N} \cup \{0\}$ let $x_{\perp} G(x) \in L^1(\Omega)$, condition (6.29) holds if dimension $d = 1$ and (6.30) if $d = 2$ and $\sqrt{2}(2\pi)^{\frac{d+1}{2}} M_d l < 1$. Then the map $t_n v = u$ on $H^2(\Omega)$ defined by equation (2.3) has a unique fixed point $v_n$, the only stationary solution of problem (1.2) in $H^2(\Omega)$.

II) When $a = \nu_0^2$, $\nu_0 \in \mathbb{N}$ let $x_{\perp}^2 G(x) \in L^1(\Omega)$, conditions (6.25), (6.27) hold if dimension $d = 1$ and conditions (6.26), (6.27) hold if $d = 2$ and $\sqrt{2}(2\pi)^{\frac{d+1}{2}} M_d l < 1$. Then the map $t_0 v = u$ on $H^2(\Omega)$ defined by equation (2.3) has a unique fixed point $v_0$, the only stationary solution of problem (1.2) in $H^2(\Omega)$.

III) When $a = 0$ let $x_{\perp}^2 G(x) \in L^1(\Omega)$, conditions (6.23) hold and $\sqrt{2}(2\pi)^{\frac{d+1}{2}} M_d l < 1$. Then the map $t_0 v = u$ on $H^2(\Omega)$ defined by equation (2.3) has a unique fixed point $v_0$, the only stationary solution of problem (1.2) in $H^2(\Omega)$.

In all cases I), II) and III) the fixed point $v_n$, $a \geq 0$ is nontrivial provided that for some $n \in \mathbb{Z}$ the intersection of supports of the Fourier images of functions $\text{supp} F(0, x) \cap \text{supp} G_n$ is a set of nonzero Lebesgue measure in $\mathbb{R}^d$.

Remark. Note that the maps discussed above act on real valued functions due to the assumptions on $F(u, x)$ and $G(x)$ involved in the nonlocal term of (2.3).

3 The Whole Space Case

Proof of Theorem 1. We present the proof of the theorem in case I) and when $a = 0$ the argument will be similar. Let us first suppose that in the case of $\Omega = \mathbb{R}^d$ for some $v \in H^2(\mathbb{R}^d)$ there exist two solutions $u_{1, 2} \in H^2(\mathbb{R}^d)$ of problem (2.3). Then their difference $w := u_1 - u_2 \in H^2(\mathbb{R}^d)$ will satisfy the homogeneous problem $-\Delta w = w$. Since the Laplacian operator acting in the whole space does not have any nontrivial square integrable eigenfunctions, $w(x)$ vanishes a.e. in $\mathbb{R}^d$. Let $v(x) \in H^2(\mathbb{R}^d)$ be arbitrary. We apply the standard Fourier transform to both sides of (2.3) and arrive at

$$\hat{w}(p) = (2\pi)^d \frac{\hat{G}(p) \hat{f}(p)}{p^2 - a}$$

(3.1)

with $\hat{f}(p)$ denoting the Fourier image of $F(v(x), x)$. Clearly, we have the upper bounds

$$|\hat{w}(p)| \leq (2\pi)^d N_{a, d} |\hat{f}(p)| \quad \text{and} \quad |p^2 \hat{w}(p)| \leq (2\pi)^d N_{a, d} |\hat{f}(p)|$$

with $N_{a, d} < \infty$ by means of Lemma A1 of the Appendix in one dimension and via Lemma A2 for $d = 2, 3$ under orthogonality relations (6.4) and (6.9).
Fourier image given by (3.1) and the map $u|\mathcal{H}$ defined. This enables us to choose arbitrarily where $\hat{\mathcal{H}}$ constrained subspaces (2.4) and (2.5) respectively instead of which is finite by means of (2.1) of Assumption 1. Therefore, for any $H$ only stationary solution of equation (1.2) in $\supseteq I$ and when $a \neq 0$. But $a \neq n^2$, $n \in \mathbb{Z}$ and therefore, it is not an eigenvalue of the operator $\frac{d^2}{dx^2}$ on $L^2(I)$ with periodic boundary conditions. Therefore, $w(x)$ vanishes a.e. in $I$. Suppose $v(x) \in H^2(I)$ is arbitrary. Let us apply the Fourier transform to problem (2.3) considered on the interval $I$ which yields

$$u_n = \sqrt{2\pi} \frac{G_n f_n}{n^2 - a}, \quad n \in \mathbb{Z}$$

Proof of Theorem 2. Let us demonstrate the proof of the theorem in case I) and when $a = n_0^2$, $n_0 \in \mathbb{N}$ or $a = 0$ the ideas will be similar, using the constrained subspaces (2.4) and (2.5) respectively instead of $H^2(I)$. First we suppose that for $v \in H^2(I)$ there are two solutions $u_{1,2} \in H^2(I)$ of problem (2.3) with $\Omega = I$. Then function $w := u_1 - u_2 \in H^2(I)$ will be a solution to the problem $-w'' = aw$. But $a \neq n^2$, $n \in \mathbb{Z}$ and therefore, it is not an eigenvalue of the operator $\frac{d^2}{dx^2}$ on $L^2(I)$ with periodic boundary conditions. Therefore, $w(x)$ vanishes a.e. in $I$. Suppose $v(x) \in H^2(I)$ is arbitrary. Let us apply the Fourier transform to problem (2.3) considered on the interval $I$ which yields

$$u_n = \sqrt{2\pi} \frac{G_n f_n}{n^2 - a}, \quad n \in \mathbb{Z}$$
with $f_n := F(v(x), x)_n$. Clearly for the transform of the second derivative we have
\[ (-w'')_n = \sqrt{2\pi} \frac{n^2 G_n f_n}{n^2 - a}, \quad n \in \mathbb{Z}, \]
which enables us to estimate
\[ \|w\|_{H^2(I)}^2 = \sum_{n=-\infty}^{\infty} |u_n|^2 + \sum_{n=-\infty}^{\infty} n^2 |u_n|^2 \leq 4\pi N_n^2 \|F(v(x), x)\|_{L^2(I)}^2 < \infty \]
due to (2.1) of Assumption 1 and Lemma A3 of the Appendix. Hence, for an arbitrary $v(x) \in H^2(I)$ there is a unique $u(x) \in H^2(I)$ solving equation (2.3) with its Fourier image given by (4.1) and the map $\tau_n : H^2(I) \to H^2(I)$ in case I is well defined. Let us consider any $v_{1,2} \in H^2(I)$ with their images under the map mentioned above $u_{1,2} = \tau_n v_{1,2} \in H^2(I)$ and arrive easily at the upper bound
\[ \|u_1 - u_2\|_{H^2(I)}^2 = \sum_{n=-\infty}^{\infty} |u_{1n} - u_{2n}|^2 + \sum_{n=-\infty}^{\infty} n^2 (u_{1n} - u_{2n})^2 \leq 4\pi N_n^2 \|F(v_1(x), x) - F(v_2(x), x)\|_{L^2(I)}^2. \]
Obviously $v_{1,2}(x) \in H^2(I) \subset L^\infty(I)$ due to the Sobolev embedding. By means of (2.2) we easily obtain
\[ \|\tau_n v_1 - \tau_n v_2\|_{H^2(I)} \leq 2\sqrt{\pi N_n} \|v_1 - v_2\|_{H^2(I)}, \]
such that the constant in the right side of this upper bound is less than one as assumed. Thus, the Fixed Point Theorem implies the existence and uniqueness of a function $v_n \in H^2(I)$ satisfying $\tau_n v_n = v_n$, which is the only stationary solution of problem (1.2) in $H^2(I)$. Suppose $v_n(x) = 0$ a.e. in $I$. Then we obtain the contradiction to the assumption that $G_n F(0, x)_n \neq 0$ for some $n \in \mathbb{Z}$. Note that in the case of $a \neq n^2$, $n \in \mathbb{Z}$ the argument does not require any orthogonality conditions.

5 The Problem on the Product of Spaces

Proof of Theorem 3. We present the proof of the theorem for case II) since when the parameter $a$ vanishes or is located on the open interval between squares of two nonnegative integers the ideas are similar. Suppose there exists $v(x) \in H^2(\Omega)$ which generates $u_{1,2}(x) \in H^2(\Omega)$ solving equation (2.3). Then the difference $w := u_1 - u_2 \in H^2(\Omega)$ will satisfy $-\Delta w = n_0^2 w$ in our domain $\Omega$. By applying the partial Fourier transform to this equation we easily arrive at $-\Delta \omega_n(x) = (n_0^2 - n^2) \omega_n(x)$. Clearly $\|w\|_{L^2(\Omega)}^2 = \sum_{n=-\infty}^{\infty} \|w_n\|_{L^2(\mathbb{R}^d)}^2$ such that $w_n(x) \in L^2(\mathbb{R}^d)$, $n \in \mathbb{Z}$. Since the transversal Laplacian operator $-\Delta \omega$ on $L^2(\mathbb{R}^d)$ does not have any nontrivial square integrable eigenfunctions.
where $\tilde{u}_n(p)$ stands for the Fourier image of $F(v(x), x)$. Obviously,
\[
|\tilde{u}_n(p)| \leq (2\pi)^{d+1} M_{n_0}^2 |\tilde{f}_n(p)| \quad \text{and} \quad |(p^2 + n^2)\tilde{u}_n(p)| \leq (2\pi)^{d+1} M_{n_0}^2 |\tilde{f}_n(p)|,
\]
where $M_{n_0} < \infty$ by means of Lemma A5 of the Appendix under the appropriate orthogonality conditions stated in it. Thus
\[
\|u\|_{H^2(\Omega)}^2 = \sum_{n=-\infty}^\infty \int_{\mathbb{R}^d} |\tilde{u}_n(p)|^2 dp + \sum_{n=-\infty}^\infty \int_{\mathbb{R}^d} |(p^2 + n^2)\tilde{u}_n(p)|^2 dp \leq (2\pi)^{d+1} M_{n_0}^2 \|F(v(x), x)\|_{L^2(\Omega)}^2 < \infty
\]
by means of (2.1) of Assumption 1, such that for any $v(x) \in H^2(\Omega)$ there exists a unique $u(x) \in H^2(\Omega)$ solving equation (2.3) with its Fourier image given by (5.1) and the map $t_g : H^2(\Omega) \rightarrow H^2(\Omega)$ in case II) of the Theorem is well defined. Then we consider arbitrary $v_{1,2} \in H^2(\Omega)$ such that the images under the map are $u_{1,2} = t_g v_{1,2} \in H^2(\Omega)$ and obtain
\[
\|u_{1} - u_{2}\|_{H^2(\Omega)}^2 = \sum_{n=-\infty}^\infty \int_{\mathbb{R}^d} |\tilde{u}_{1n}(p) - \tilde{u}_{2n}(p)|^2 dp + \sum_{n=-\infty}^\infty \int_{\mathbb{R}^d} |(p^2 + n^2)(\tilde{u}_{1n}(p) - \tilde{u}_{2n}(p))|^2 dp \leq (2\pi)^{d+1} M_{n_0}^2 \|F(v_1(x), x) - F(v_2(x), x)\|_{L^2(\Omega)}^2.
\]
Clearly $v_{1,2} \in H^2(\Omega) \subset L^\infty(\Omega)$ via the Sobolev embedding theorem. Using (2.2) we easily arrive at the estimate
\[
\|t_{n_0^2} v_1 - t_{n_0^2} v_2\|_{H^2(\Omega)} \leq \sqrt{2}(2\pi)^{d+1} M_{n_0^2} \|v_1 - v_2\|_{H^2(\Omega)}
\]
with the constant in the right side of it less than one by assumption. Therefore, the Fixed Point Theorem yields the existence and uniqueness of a function $v_{n_0^2} \in H^2(\Omega)$ which satisfies $t_{n_0^2} v_{n_0^2} = v_{n_0^2}$ and is the only stationary solution of problem (1.2) in $H^2(\Omega)$ in case II) of the theorem. Suppose $v_{n_0^2}(x) = 0$ a.e. in $\Omega$. This yields the contradiction to the assumption that there exists $n \in \mathbb{Z}$ for which $\text{supp} G_n \cap \text{supp} F(0, x)_n$ is a set of nonzero Lebesgue measure in $\mathbb{R}^d$. ■

**6 Appendix**

Let $G(x)$ be a function, $G(x) : \mathbb{R}^d \rightarrow \mathbb{R}$, $d \leq 3$ for which we denote its standard Fourier transform using the hat symbol as
\[
\hat{G}(p) := \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} G(x) e^{-ipx} dx, \quad p \in \mathbb{R}^d
\]
such that
\[ \| \hat{G}(p) \|_{L^\infty(\mathbb{R}^d)} \leq \frac{1}{(2\pi)^{\frac{d}{2}}} \| G \|_{L^1(\mathbb{R}^d)} \] (6.1)
and
\[ G(x) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \hat{G}(q)e^{iqx}dq, \ x \in \mathbb{R}^d. \]

Let us define the auxiliary quantities
\[ N_{a, d} := \max \left\{ \| \hat{\chi} \|_{L^\infty(\mathbb{R}^d)} \right\} \] (6.2)
and
\[ N_{0, d} := \max \left\{ \| \hat{\chi} \|_{L^\infty(\mathbb{R}^d)} \right\} \] (6.3)
when \( a = 0 \).

**Lemma A1.** Let \( G(x) \in L^1(\mathbb{R}) \).

a) If \( a > 0 \) and \( xG(x) \in L^1(\mathbb{R}) \) then \( N_{a, 1} < \infty \) if and only if
\[ \left( \frac{G(x)}{a} \right)_{L^2(\mathbb{R})} = 0. \] (6.4)

b) If \( a = 0 \) and \( xG(x) \in L^1(\mathbb{R}) \) then \( N_{0, 1} < \infty \) if and only if
\[ (G(x), 1)_{L^2(\mathbb{R})} = 0 \quad \text{and} \quad (G(x), x)_{L^2(\mathbb{R})} = 0. \] (6.5)

**Proof.** In order to prove part a) of the lemma we write the function
\[ \frac{G(p)}{p^2 - a} = \frac{G(p)}{p^2 - a} \chi_{I_+} + \frac{G(p)}{p^2 - a} \chi_{I^-}, \] (6.6)
where \( \chi_A \) here and further down stands for the characteristic function of a set \( A \). \( A^c \) for its complement, the set \( I_+ = I^+ \cup I^- \) with \( I^+ = \{ p \in \mathbb{R} \mid \sqrt{a} - \delta < p < \sqrt{a} + \delta \} \), \( I^- = \{ p \in \mathbb{R} \mid -\sqrt{a} - \delta < p < -\sqrt{a} + \delta \} \) and \( 0 < \delta < \sqrt{a} \).

The second term in the right side of (6.6) can be easily estimated in absolute value from above using (6.1) as
\[ \frac{1}{\sqrt{2\pi b^d}} \| G \|_{L^1(\mathbb{R})} < \infty \]
and the remaining term in the right side of (6.6) can be written as
\[ \frac{G(p)}{p^2 - a} \chi_{I_+} + \frac{G(p)}{p^2 - a} \chi_{I^-}. \]

We will use the expansions near the points of singularity given by
\[ \hat{G}(p) = \hat{G}(\sqrt{a}) + \int_{\sqrt{a}}^p \frac{d\hat{G}(s)}{ds}ds, \quad \hat{G}(p) = \hat{G}(-\sqrt{a}) + \int_{-\sqrt{a}}^p \frac{d\hat{G}(s)}{ds}ds. \]
with \( \left\| \frac{d\hat{G}(p)}{dp} \right\|_{L^\infty(\mathbb{R})} \leq \frac{1}{\sqrt{2\pi}} \|xG\|_{L^1(\mathbb{R})} < \infty \) by the assumption of the lemma. This enables us to obtain the bound

\[
\left| \int_0^p \frac{d\hat{G}(a)}{da} \, da \right| \leq \frac{C}{2\sqrt{\alpha - \delta}} < \infty, \quad \left| \int_{-\sqrt{\alpha - \delta}}^p \frac{d\hat{G}(a)}{da} \, da \right| \leq \frac{C}{2\sqrt{\alpha - \delta}} < \infty.
\]

Therefore it remains to estimate

\[
\frac{\hat{G}(\sqrt{\alpha})}{p^2 - a} \chi_{1^+} \leq \frac{\hat{G}(\sqrt{\alpha})}{p^2 - a} \chi_{1^+} + \frac{\hat{G}(\sqrt{a})}{p^2 - a} \chi_{1^+}.
\]

which belongs to \( L^\infty(\mathbb{R}) \) if and only if \( \hat{G}(\pm \sqrt{\alpha}) = 0 \), which is equivalent to the orthogonality relations \( (6.4) \). To estimate the second term in the right side of \( (6.2) \) under these orthogonality relations we consider the two situations. The first one is when \( |p| \leq \sqrt{\alpha + \delta} \) and we have the bound

\[
\left| \frac{p^2 \hat{G}(p)}{p^2 - a} \right| \leq (\sqrt{\alpha + \delta})^2 \left\| \frac{\hat{G}(p)}{p^2 - a} \right\|_{L^\infty(\mathbb{R})} < \infty.
\]

In the second one \( |p| > \sqrt{\alpha + \delta} \) which yields \( \frac{p^2}{p^2 - a} \in L^\infty(\mathbb{R}) \) and \( \hat{G}(p) \) is bounded via \( (6.1) \), which completes the proof of part a) of the lemma. Then we turn our attention to the situation of \( a = 0 \), such that

\[
\frac{\hat{G}(p)}{p^2} = \frac{\hat{G}(p)}{p^2} \chi_{\{|p| \leq 1\}} + \frac{\hat{G}(p)}{p^2} \chi_{\{|p| > 1\}}.
\]

The second term in the right side of the identity above can be easily estimated as

\[
\left| \frac{\hat{G}(p)}{p^2} \right| \leq \|\hat{G}(p)\|_{L^\infty(\mathbb{R})} < \infty
\]

due to \( (6.1) \). We will make use of the representation

\[
\hat{G}(p) = \hat{G}(0) + \frac{d\hat{G}(0)}{dp} + \int_0^p \left( \int_0^s \frac{d^2\hat{G}(q)}{dq^2} \, dq \right) \, ds.
\]

Obviously

\[
\left| \frac{d^2\hat{G}(p)}{dp^2} \right| \leq \frac{1}{\sqrt{2\pi}} \|x^2G(x)\|_{L^1(\mathbb{R})} < \infty
\]

by the assumption of the lemma. Hence

\[
\left| \int_0^p \left( \int_0^s \frac{d^2\hat{G}(q)}{dq^2} \, dq \right) \, ds \right| \chi_{\{|p| \leq 1\}} \leq \frac{C}{2} < \infty
\]

and the only expression which remains to estimate is given by

\[
\left[ \frac{\hat{G}(0)}{p^2} + \frac{d\hat{G}(0)}{dp} \right] \chi_{\{|p| \leq 1\}}, \text{ which is contained in } L^\infty(\mathbb{R}) \text{ if and only if } \hat{G}(0) \text{ and}
\]

\[
\hat{G}(\sqrt{\alpha}) \text{ is contained in } L^\infty(\mathbb{R}).
\]
\[ \frac{d\hat{G}}{dp}(0) \] vanish. This is equivalent to the orthogonality relations (6.5). Note that \( \|\hat{G}(p)\|_{L^\infty(\mathbb{R})} < \infty \) by means of (6.1).

The proposition above can be generalized to higher dimensions in the following statement.

**Lemma A2.** Let \( G(x) \in L^1(\mathbb{R}^d), \ d = 2, 3. \)

a) If \( a > 0 \) and \( xG(x) \in L^1(\mathbb{R}^d) \) then \( N_{a, d} < \infty \) if and only if

\[
\left( G(x), \frac{e^{ipx}}{(2\pi)^{\frac{d}{2}}} \right)_{L^2(\mathbb{R}^d)} = 0 \quad \text{for} \quad p \in S^d_{\sqrt{a}} \ \text{a.e.} \tag{6.9}
\]

b) If \( a = 0 \) and \( x^2G(x) \in L^1(\mathbb{R}^d) \) then \( N_{0, d} < \infty \) if and only if

\[
(G(x), 1)_{L^2(\mathbb{R}^d)} = 0 \quad \text{and} \quad (G(x), x_k)_{L^2(\mathbb{R}^d)} = 0, \quad 1 \leq k \leq d. \tag{6.10}
\]

**Proof.** To prove part a) of the lemma we introduce the auxiliary spherical layer in the space of \( d = 2, 3 \) dimensions

\[
A_\delta := \{ p \in \mathbb{R}^d \mid \sqrt{a} - \delta < |p| < \sqrt{a} + \delta \}, \ 0 < \delta < \sqrt{a}
\]

and write

\[
\hat{G}(p) = \hat{G}(p) \left( \frac{p^2 - a}{p^2 - \sqrt{a}} \right) \chi_{A_\delta} + \frac{\hat{G}(p)}{p^2 - a} \chi_{A_\delta}. \tag{6.11}
\]

For the second term in the right side of (6.11) we have the upper bound in the absolute value as \( \|\hat{G}(p)\|_{L^\infty(\mathbb{R}^d)} < \infty \) due to (6.1). Let us expand

\[
\hat{G}(p) = \int_{|s|}^{\sqrt{a}} \frac{\partial \hat{G}(|s|, \sigma)}{\partial |s|} |s| \ d|s| + \hat{G}(\sqrt{a}, \sigma),
\]

where \( \sigma \) stands for the angle variables on the sphere. Using the elementary inequality

\[
\left| \frac{\partial \hat{G}(p)}{\partial |s|} \right| \leq \frac{1}{(2\pi)^{\frac{d}{2}}} \|xG(x)\|_{L^1(\mathbb{R}^d)} \]

with its right side finite by the assumption of the lemma we estimate

\[
\left| \int_{|s|}^{\sqrt{a}} \frac{\partial \hat{G}(|s|, \sigma)}{\partial |s|} |s| \ d|s| \right| \leq \frac{C}{\sqrt{a}} < \infty.
\]

The only remaining term \( \hat{G}(\sqrt{a}, \sigma) \chi_{A_\delta} \in L^\infty(\mathbb{R}^d), \ d = 2, 3 \) if and only if \( \hat{G}(\sqrt{a}, \sigma) \) vanishes a.e. on the sphere \( S^d_{\sqrt{a}} \), which is equivalent to orthogonality relations (6.9). The proof of the fact that the second norm in the right
side of (6.2) under conditions (6.9) is finite is analogous to the one presented in Lemma A1 in one dimension. For the proof of part b) of the lemma we apply the two and three dimensional analog of formula (6.7), such that for the second term in its right side there is a bound analogous to (6.8). Let us use the representation formula

\[ \hat{G}(p) = \hat{G}(0) + |p| \frac{\partial \hat{G}}{\partial |p|}(0, \sigma) + \int_0^{(|p|} \left( \int_0^s \frac{\partial^2 \hat{G}(|q|, \sigma)}{\partial |q|^2} d|q| \right) ds. \]

Apparently

\[ \frac{\partial \hat{G}}{\partial |p|}(0, \sigma) = -\frac{i}{2\pi} \int_{\mathbb{R}^d} G(x)|x| \cos \theta dx, \]  

(6.12)

where \( \theta \) is the angle between vectors \( p \) and \( x \) in \( \mathbb{R}^d \) and for the second derivative

\[ \left| \frac{\partial^2 \hat{G}(p)}{\partial |p|^2} \right| \leq \frac{1}{(2\pi)^\frac{d}{2}} \|x^2 G(x)\|_{L^1(\mathbb{R}^d)} < \infty \]

by the assumption of the lemma. This yields

\[ \left| \int_0^{(|p|} \left( \int_0^s \frac{\partial^2 \hat{G}(|q|, \sigma)}{\partial |q|^2} d|q| \right) ds \right| \leq \frac{C_2}{2} < \infty, \]

such that the only expression remaining to estimate is given by

\[ \left[ \frac{\hat{G}(0)}{|p|^2} + \frac{\frac{\partial \hat{G}}{\partial |p|}(0, \sigma)}{|p|} \right] \chi_{\{|p| \leq 1\}} \]  

(6.13)

with the first derivative (6.12) containing the angular dependence. We consider first the case of \( d = 2 \) such that \( p = (|p|, \theta_p), \; x = (|x|, \theta_x) \in \mathbb{R}^2 \) and the angle between them \( \theta = \theta_p - \theta_x \). A straightforward computation yields that the right side of (6.12) is given by

\[ -\frac{i}{2\pi} \sqrt{Q_1^2 + Q_2^2} \cos (\theta_x - \alpha) \]  

where

\[ Q_1 := \int_{\mathbb{R}^2} G(x)x_1 dx, \quad Q_2 := \int_{\mathbb{R}^2} G(x)x_2 dx, \quad \tan \alpha := \frac{Q_2}{Q_1} \]  

(6.14)

and \( x = (x_1, x_2) \in \mathbb{R}^2 \) such that (6.13) is equal to

\[ \left[ \frac{\hat{G}(0)}{|p|^2} - \frac{i}{2\pi} \sqrt{Q_1^2 + Q_2^2} \cos (\theta_x - \alpha) \right] \chi_{\{|p| \leq 1\}}. \]

Note that the situation of \( Q_1 = 0 \) and \( Q_2 \neq 0 \) corresponds to the cases of \( \alpha \) equal to \( \frac{\pi}{2} \) or \( -\frac{\pi}{2} \). Obviously, the expression above is contained in \( L^\infty(\mathbb{R}^2) \).
if and only if the quantities \( \hat{G}(0) \), \( Q_1 \) and \( Q_2 \) vanish, which is equivalent to orthogonality relations (6.10) in two dimensions. In the case of \( d = 3 \) the argument is quite similar. The coordinates of vectors

\[
x = (x_1, x_2, x_3) = (|x| \sin \theta x \cos \varphi, |x| \sin \theta x \sin \varphi, |x| \cos \theta) \in \mathbb{R}^3
\]

and

\[
p = (|p| \sin \theta p \cos \varphi, |p| \sin \theta p \sin \varphi, |p| \cos \theta) \in \mathbb{R}^3
\]

are being used to compute \( \cos \theta = \frac{(p, x)}{|p| |x|} \), involved in the right side of (6.12). Here \( (p, x) \) stands for the scalar product of the vectors in three dimensions.

An easy calculation shows that when \( d = 3 \) the right side of (6.12) can be written as

\[
-\frac{i}{(2\pi)^{\frac{d}{2}}} \left\{ \sqrt{Q_1^2 + Q_2^2} \sin \theta_p \cos (\varphi_p - \alpha) + Q_3 \cos \theta_p \right\}
\]

with \( \alpha \) given by (6.14) and here \( Q_k = \int_{\mathbb{R}^3} G(x)x_k dx, \ k = 1, 2, 3 \), which are the three dimensional generalizations of the correspondent expressions given by (6.14) and term (6.13) will be equal to

\[
\left[ \hat{G}(0) \right] \left( \frac{p}{|p|^2} - \frac{i}{(2\pi)^{\frac{d}{2}}} \right) \left\{ \sqrt{Q_1^2 + Q_2^2} \sin \theta_p \cos (\varphi_p - \alpha) + Q_3 \cos \theta_p \right\} \chi_{(|p| \leq 1)}
\]

and will belong to \( L^\infty(\mathbb{R}^3) \) if and only if \( \hat{G}(0) \) along with \( Q_k, \ k = 1, 2, 3 \) vanish, which is equivalent to orthogonality conditions (6.10) in three dimensions. The second norm in the right side of (6.3) is finite under relations (6.1).

Let the function \( G(x) : I \to \mathbb{R}, \ G(0) = G(2\pi) \) and its Fourier transform on the finite interval is given by

\[
G_n := \int_0^{2\pi} G(x) e^{-inx} \sqrt{2\pi} dx, \ n \in \mathbb{Z}
\]

and \( G(x) = \sum_{n=-\infty}^{\infty} G_n e^{inx} \sqrt{2\pi} \). Similarly to the whole space case we define

\[
N_a := \max \left\{ \left\| \frac{G_n}{n^2 - a} \right\|_{l^\infty}, \left\| \frac{n^2 G_n}{n^2 - a} \right\|_{l^\infty} \right\}
\]

(6.15)

for \( a > 0 \). In the situation of \( a = 0 \)

\[
N_0 := \max \left\{ \left\| \frac{G_n}{n^2} \right\|_{l^\infty}, \left\| G_n \right\|_{l^\infty} \right\}
\]

(6.16)
We have the following elementary statement.

**Lemma A3.** Let \( G(x) \in L^1(I) \) and \( G(0) = G(2\pi) \).

a) If \( a > 0 \) and \( a \neq n^2 \), \( n \in \mathbb{Z} \) then \( N_a < \infty \).

b) If \( a = n_0^2 \), \( n_0 \in \mathbb{N} \) then \( N_a < \infty \) if and only if
\[
\left( G(x), \frac{e^{\pm in_0 x}}{\sqrt{2\pi}} \right)_{L^2(I)} = 0. \tag{6.17}
\]

c) If \( a = 0 \) then \( N_0 < \infty \) if and only if
\[
\left( G(x), 1 \right)_{L^2(I)} = 0. \tag{6.18}
\]

**Proof.** Clearly we have the bound
\[
\|G_n\|_{L^\infty} \leq \frac{1}{\sqrt{2\pi}} \|G\|_{L^1(I)} < \infty. \tag{6.19}
\]
Thus in case a) when \( a \neq n^2 \), \( n \in \mathbb{Z} \) the expressions under the norms in the right side of (6.15) do not contain any singularities and the result of the lemma is obvious. When \( a = n_0^2 \) for some \( n_0 \in \mathbb{N} \) or \( a = 0 \) conditions (6.17) and (6.18) respectively are necessary and sufficient for eliminating the existing singularities by making the corresponding Fourier coefficients equal to zero: \( G_{\pm n_0} \) in case b) and \( G_0 \) in case c).

Let \( G(x) \) be a function on the product of spaces studied in Theorem 3, \( G(x) : \Omega = I \times \mathbb{R}^d \rightarrow \mathbb{R}, \ d = 1, 2, \ G(0, x_\perp) = G(2\pi, x_\perp) \) for \( x_\perp \in \mathbb{R}^d \) a.e. and its Fourier transform on the product of spaces equals to
\[
\hat{G}_n(p) := \frac{1}{(2\pi)^{d+1}} \int_{\mathbb{R}^d} dx_\perp e^{-ipx_\perp} \int_0^{2\pi} G(x_1, x_\perp) e^{-inx_1} dx_1, \quad p \in \mathbb{R}^d, \ n \in \mathbb{Z}
\]
such that
\[
|\hat{G}_n(p)|_{L^\infty_{\mathbb{R}^d}} := \sup_{p \in \mathbb{R}^d, \ n \in \mathbb{Z}} |\hat{G}_n(p)| \leq \frac{1}{(2\pi)^{d+1}} \|G\|_{L^1(\Omega)} \tag{6.20}
\]
and \( G(x) = \sum_{n=-\infty}^{\infty} \int_{\mathbb{R}^d} \hat{G}_n(p) e^{ipx_\perp} e^{inx_1} dp \). It is also useful to consider the Fourier transform only in the first variable, such that
\[
G_n(x_\perp) := \int_0^{2\pi} G(x_1, x_\perp) e^{-inx_1} dx_1, \quad n \in \mathbb{Z}.
\]

Let us introduce \( \xi_n^a(p) := \frac{\hat{G}_n(p)}{p^2 + n^2 - a} \) and define
\[
M_a := \max\{\|\xi_n^a(p)\|_{L^\infty_{\mathbb{R}^d}}, \|p^2 + n^2\xi_n^a(p)\|_{L^\infty_{\mathbb{R}^d}}\} \tag{6.21}
\]
for $a > 0$ and

$$M_0 := \max \left\{ \left\| \frac{\hat{G}_n(p)}{p^2 + n^2} \right\|_{L^\infty_p}, \left\| \hat{G}_n(p) \right\|_{L^\infty_p} \right\}$$  \hspace{1cm} (6.22)$$

when $a = 0$. Here the momentum vector $p \in \mathbb{R}^d$.

**Lemma A4.** Let $G(x) \in L^1(\Omega)$, $x_2^2 G(x) \in L^1(\Omega)$ and $G(0, x_\perp) = G(2\pi, x_\perp)$ for $x_\perp \in \mathbb{R}^d$ a.e., $d = 1, 2$. Then $M_0 < \infty$ if and only if

$$(G(x), 1)_{L^2(\Omega)} = 0, \ (G(x), x_\perp, k)_{L^2(\Omega)} = 0, \ 1 \leq k \leq d, \ d = 1, 2. \hspace{1cm} (6.23)$$

**Proof.** Let us expand

$$\frac{\hat{G}_n(p)}{p^2 + n^2} = \frac{\hat{G}_0(p)}{p^2} \chi_{p \in \mathbb{R}^d, n=0} + \frac{\hat{G}_n(p)}{p^2 + n^2} \chi_{p \in \mathbb{R}^d, n \neq 0}. \hspace{1cm}$$

The second term in the right side of this identity can be estimated above in the absolute value by means of (6.20) by

$$\frac{1}{(2\pi)^{d+1}} \left\| G \right\|_{L^1(\Omega)} < \infty. \hspace{1cm}$$

Clearly we have the bound on the norm

$$\left\| x_\perp^2 G_n(x_\perp) \right\|_{L^1(\mathbb{R}^d)} \leq \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} dx_1 \int_{\mathbb{R}^d} dx_\perp x_\perp^2 |G(x)| < \infty, \ n \in \mathbb{Z} \hspace{1cm} (6.24)$$

by the assumption of the lemma. Thus the term $\frac{\hat{G}_0(p)}{p^2} \in L^\infty(\mathbb{R}^d)$ if and only if the orthogonality conditions (6.23) hold, which is guaranteed for $d = 1$ by Lemma A1 and when dimension $d = 2$ by Lemma A2. Note that the last term in the right side of (6.22) is bounded via (6.20).

Next we turn our attention to the situation when the parameter $a$ is nontrivial.

**Lemma A5.** Let $G(x) \in L^1(\Omega)$, $x_2^2 G(x) \in L^1(\Omega)$ and $G(0, x_\perp) = G(2\pi, x_\perp)$ for $x_\perp \in \mathbb{R}^d$ a.e., $d = 1, 2$ and $a = n_0^2$, $n_0 \in \mathbb{N}$. Then $M_a < \infty$ if and only if

$$\left( G(x_1, x_\perp), \frac{e^{inx_1} e^{i\pm\sqrt{n_0^2-n_1^2}x_\perp}}{\sqrt{2\pi} \sqrt{2\pi}} \right)_{L^2(\Omega)} = 0, \ |n| \leq n_0 - 1, \ d = 1, \hspace{1cm} (6.25)$$

$$\left( G(x_1, x_\perp), \frac{e^{inx_1} e^{ipx_\perp}}{\sqrt{2\pi} \sqrt{2\pi}} \right)_{L^2(\Omega)} = 0, \ p \in S^2, \ n \in \mathbb{N}, \ \sqrt{n_0^2-n^2} \ a.e., \ |n| \leq n_0 - 1, \ d = 2, \hspace{1cm} (6.26)$$

and

$$\left( G(x_1, x_\perp), \frac{e^{i\pm n_0 x_1}}{\sqrt{2\pi}} \right)_{L^2(\Omega)} = 0, \ G(x_1, x_\perp) \frac{e^{i\pm n_0 x_1}}{\sqrt{2\pi} \sqrt{2\pi}} x_\perp k \right)_{L^2(\Omega)} = 0, \ 1 \leq k \leq d. \hspace{1cm} (6.27)$$
Lemma A6. Between the squares of two consecutive nonnegative integers.

x(\sum_{\Omega} x)^{(n)} \in R^d, |n| \leq n_0, n_0 \in Z.

Obviously \|G_n(p)\|_{L^2(\Omega)} < \infty by means of (6.20). We have trivial estimates on the norms for \( n \in Z \)

\[ \|G_n(x_\perp)\|_{L^1(\Omega^2)} \leq \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \int_{\Omega^2} dx_\perp |G(x_1, x_\perp)| < \infty \]

and

\[ \|x_\perp G_n(x_\perp)\|_{L^1(\Omega^2)} \leq \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \int_{\Omega^2} dx_\perp |x_\perp G(x_1, x_\perp)| < \infty. \]

Note that \( G(x) \in L^1(\Omega) \) and \( x_\perp^2 G(x) \in L^1(\Omega) \) by the assumptions of the lemma, which yields \( x_\perp G(x) \in L^1(\Omega) \). Thus when dimension \( d = 1 \) by means of Lemma A1 \( \xi_n^a(p)(p)\chi_{\{p \in R^d, |n| \leq n_0\}} \leq \|G_n(p)\|_{L^\infty(\bar{\Omega})} < \infty \) if and only if orthogonality relations (6.25) hold. For \( d = 2 \) the necessary and sufficient conditions for the boundedness of the second term in (6.28) via Lemma A2 are given by (6.26). Lemmas A1 and A2 yield that the third term in (6.28) belongs to \( L^\infty(\bar{\Omega}) \) if and only if conditions (6.27) with the positive sign under the exponents are satisfied. Clearly \( x_\perp^2 G_n(x_\perp) \in L^1(\Omega^2) \) due to the assumption of the lemma and estimate (6.24). Similarly, we obtain that the necessary and sufficient conditions for the the last term in (6.28) to be contained in \( L^\infty(\bar{\Omega}) \) are given by (6.27) with the negative sign under the exponents. Then we represent \( (p^2 + n^2)\xi_n^a(p) \) as the sum

\[ (p^2 + n^2)\xi_n^a(p)(p)\chi_{\{p \in R^d, n \in Z, p^2 + n^2 \leq n_0^2 + 1\}} + (p^2 + n^2)\xi_n^a(p)\chi_{\{p \in R^d, n \in Z, p^2 + n^2 > n_0^2 + 1\}} \]

in which the absolute value of the first term has the upper bound \( (n_0^2 + 1)\|\xi_n^a(p)\|_{L^\infty(\bar{\Omega})} < \infty \) under the orthogonality conditions of the lemma and of the second one \( (1 + n_0)\|G_n(p)\|_{L^\infty(\bar{\Omega})} < \infty \) via (6.20).

Finally, we study the case when the parameter \( a \) is located on an open interval between the squares of two consecutive nonnegative integers.

Lemma A6. Let \( G(x) \in L^1(\Omega) \), \( x_\perp G(x) \in L^1(\Omega) \) and \( G(0, x_\perp) = G(2\pi, x_\perp) \) for \( x_\perp \in R^d \) a.e., \( d = 1, 2 \) and \( n_0^2 < a < (n_0 + 1)^2 \), \( n_0 \in Z^+ = N \cup \{0\} \). Then \( M_n < \infty \) if and only if

\[ \left( G(x_1, x_\perp), \frac{e^{-in_1 x_\perp}}{\sqrt{2\pi}} \right)_{L^2(\Omega)} = 0, \quad |n| \leq n_0, \quad d = 1. \]
\[
\left( G(x_1, x_\perp), \frac{e^{inx_1} e^{ipx_\perp}}{\sqrt{2\pi} 2\pi} \right) _{L^2(\Omega)} = 0, \quad p \in S^2_{\sqrt{\alpha - n^2}} \text{ a.e., } |n| \leq n_0, \quad d = 2.
\]

(6.30)

**Proof.** Let us expand \( \xi_n^a(p) \) as the sum of two terms

\[
\xi_n^a(p) \chi_{\{p \in \mathbb{R}^d, |n| \geq n_0 + 1\}} + \xi_n^a(p) \chi_{\{p \in \mathbb{R}^d, |n| \leq n_0\}},
\]

such that the absolute value of the first one is bounded above by

\[
\left\| \hat{G}_n(p) \right\|_{L^\infty_{n,p}} < \infty \quad \text{and the second one belongs to } L^\infty_{n,p} \text{ if and only if orthogonality relations (6.29) are satisfied in one dimension by means of Lemma A1 and if and only if conditions (6.30) are fulfilled in two dimensions via Lemma A2.}
\]

We write \((p^2 + n^2)\xi_n^a(p)\) as the sum

\[
(p^2 + n^2)\xi_n^a(p) \chi_{\{p \in \mathbb{R}^d, |n| \geq n_0 + 1\}} + (p^2 + n^2)\xi_n^a(p) \chi_{\{p \in \mathbb{R}^d, |n| \leq n_0\}},
\]

in which the first and the second terms can be easily bounded above in their absolute values by the quantities finite under the conditions of the lemma, namely

\[
\left(1 + \frac{a}{(n_0 + 1)^2 - a}\right)\|\hat{G}_n(p)\|_{L^\infty_{n,p}} \quad \text{and} \quad (n_0 + 1)^2\|\xi_n^a(p)\|_{L^\infty_{n,p}}
\]

respectively.

**Acknowledgement.** The first author thanks James Colliander for stimulating discussions.

**References**

[1] M.S. Agranovich. *Elliptic boundary problems*, Encyclopaedia Math. Sci., vol. 79, Partial Differential Equations, IX, Springer, Berlin, 1997, pp. 1–144.

[2] N. Apreutesei, N. Bessonov, V. Volpert, V. Vougalter: *Spatial Structures and Generalized Travelling Waves for an Integro-Differential Equation*, Discrete Contin. Dyn. Syst. Ser. B 13, 3 (2010), 537–557

[3] H. Berestycki, G. Nadin, B. Perthame, L. Ryzhik: *The non-local Fisher-KPP equation: traveling waves and steady states*, Nonlinearity 22, 12 (2009), 2813–2844

[4] A. Ducrot, M. Marion and V. Volpert: *Systemes de réaction-diffusion sans propriété de Fredholm*, CRAS, 340, 9 (2005), 659–664

*Documenta Mathematica* 16 (2011) 561–580
[5] A. Ducrot, M. Marion and V. Volpert: Reaction-diffusion problems with non Fredholm operators, Advances Diff. Equations, 13, 11-12 (2008), 1151–1192

[6] A. Ducrot, M. Marion and V. Volpert: Reaction-diffusion waves (with the Lewis number different from 1). Publibook, Paris, 2009.

[7] S. Genieys, V. Volpert, P. Auger: Pattern and waves for a model in population dynamics with nonlocal consumption of resources, Math. Model. Nat. Phenom. 1, 1 (2006), 63–80

[8] P.D. Hislop, I.M. Sigal: Introduction to spectral theory with applications to Schrödinger operators. Springer, 1996

[9] J.L. Lions, E. Magenes. Problèmes aux limites non homogènes et applications. Volume I. Dunod, Paris, 1968.

[10] L.R. Volevich. Solvability of boundary problems for general elliptic systems, Mat. Sb., 68 (1965), 373–416; English translation: Amer. Math. Soc. Transl., 67 (1968), Ser. 2, 182–225.

[11] V. Volpert. Elliptic partial differential equations. Volume I. Fredholm theory of elliptic problems in unbounded domains. Birkhäuser, 2011.

[12] V. Volpert, B. Kazmierczak, M. Massot, Z. Peradzynski: Solvability conditions for elliptic problems with non-Fredholm operators, Appl. Math., 29, 2 (2002), 219–238

[13] V. Vougalter, V. Volpert: Solvability conditions for some non Fredholm operators, Proc. Edinb. Math. Soc. (2), 54, 1 (2011), 249–271

[14] V. Vougalter, V. Volpert: On the solvability conditions for some non Fredholm operators, Int. J. Pure Appl. Math., 60, 2 (2010), 169–191

[15] V. Vougalter, V. Volpert: On the solvability conditions for the diffusion equation with convection terms. To appear in Commun. Pure Appl. Anal.

[16] V. Vougalter, V. Volpert: Solvability relations for some non Fredholm operators, Int. Electron. J. Pure Appl.Math., 2, 1 (2010), 75–83

[17] V. Volpert, V. Vougalter: On the solvability conditions for a linearized Cahn-Hilliard equation. To appear in Rend. Istit. Mat. Univ. Trieste.

[18] V. Vougalter, V. Volpert: Solvability conditions for some systems with non Fredholm operators, Int. Electron. J. Pure Appl.Math., 2, 3 (2010), 183–187

[19] V. Volpert, V. Vougalter: Solvability conditions for some linear and nonlinear non-Fredholm elliptic problems. Preprint 2011.
FUNDAMENTAL GROUP OF SCHURIAN CATEGORIES
AND THE HUREWICZ ISOMORPHISM

CLAUDE CIBILS, MARÍA JULIA REDONDO AND ANDREA SOLOTAR

Abstract. Let $k$ be a field. We attach a CW-complex to any Schurian $k$-category and we prove that the fundamental group of this CW-complex is isomorphic to the intrinsic fundamental group of the $k$-category. This extends previous results by J.C. Bustamante in [8]. We also prove that the Hurewicz morphism from the vector space of abelian characters of the fundamental group to the first Hochschild-Mitchell cohomology vector space of the category is an isomorphism.

2010 Mathematics Subject Classification: 55Q05 18D20 16E40 16W25 16W50
Keywords and Phrases: Fundamental group, linear category, Schurian category, Hurewicz, Hochschild

1 Introduction

In this paper we consider Schurian categories, that is, small linear categories over a field $k$ such that each vector space of morphisms is either of dimension one or zero.

Recall that there is no homotopy theory available for a $k$-algebra or, more generally, for a $k$-linear category. More precisely there is neither homotopy equivalence nor a definition of loops as in algebraic topology taking into account the $k$-linear structure. As an alternative we consider an intrinsic fundamental
Claude Cibils, María Julia Redondo, and Andrea Solotar

`a la Grothendieck, that we have defined in [13] and [14] using connected gradings. In [14] we have computed this group for matrix algebras $M_p(k)$ where $p$ is a prime and $k$ is an algebraically closed field of characteristic zero, obtaining that $\pi_1(M_p(k)) = F_{p-1} \times C_p$ where $F_{p-1}$ is the free group with $p-1$ generators and $C_p$ is the cyclic group of order $p$, using classifications of gradings provided in [2, 4, 23].

The intrinsic fundamental group is defined in terms of Galois coverings provided by connected gradings. It is the automorphism group of the fibre functor over a fixed object. In case a universal covering $U$ exists for a $k$-linear category $C$, its fundamental group $\pi_1(C)$ is isomorphic to the automorphism group of the covering $U$.

It is intrinsic in the sense that it does not depend on the presentation of the $k$-category by generators and relations. If a universal covering exists, then we obtain that the fundamental groups constructed by R. Martínez-Villa and J.A. de la Peña (see [27], and [1, 6, 20]) depending on the choice of a presentation of the category by a quiver and relations are in fact quotients of the intrinsic $\pi_1$ that we introduce. Note that those groups can vary according to different presentations of the same $k$-category (see for instance [1, 10, 25]).

The definition of $\pi_1(C)$ is inspired in the topological case considered for instance in R. Douady and A. Douady’s book [16]. They are closely related to the way in which the fundamental group is viewed in algebraic geometry after A. Grothendieck and C. Chevalley.

Note that the existence of a universal covering for a $k$-linear category is equivalent to the existence of a universal grading, namely a connected grading such that any other connected grading is a quotient of it.

In this paper we will prove that a Schurian category $C$ admits a universal covering. In fact a universal grading is obtained through the topological fundamental group of a CW-complex $CW(C)$ that we attach to $C$. We infer that $\pi_1(C) = \pi_1(CW(C))$. The CW-complex we define is very close to a simplicial complex. It is a simplicial complex when $C$ is such that if $y_Cx \neq 0$ then $x_Cy = 0$ for $x \neq y$ (where $y_Cx$ is the vector space of morphisms from $x$ to $y$).

J.C. Bustamante considers in [8] $k$-categories with a finite number of objects subject to the above conditions which he calls "Schurian almost triangular" in order to prove a similar result through the fundamental group of a presentation as defined in [6, 21, 27]. He uses the simplicial complex from [5, 7, 28] whose 2-skeleton coincides with ours in the Schurian almost triangular context. We do not require that the category has a finite number of objects, neither that it admits an admissible presentation. Moreover we provide an example of a Schurian category which has no admissible presentation and we compute its fundamental group. Note also that we generalize the result by M. Bardzell and E. Marcos in [3], where they prove that the fundamental group of a Schurian basic algebra does not depend on the admissible presentation.

We thank G. Minian for interesting discussions, and in particular for pointing out that cellular approximation enables to provide homotopy arguments from algebraic topology using the 1-skeleton. In [8], J.C. Bustamante uses the edge
path group instead, which requires a finite number of objects. In [9] a CW-complex attached to a presentation of a finite number of objects category is considered in order to compute the fundamental group of the presentation.

In case of a complete Schurian category $\mathcal{C}$, that is all the vector spaces of morphisms are one dimensional and composition of non-zero morphisms is non-zero, the CW-complex attached to $\mathcal{C}$ enables to retrieve that $\pi_1(\mathcal{C}) = 1$, see [14, Corollary 4.6].

Finally we consider the Hurewicz morphism (see [1, 12, 31]) for a Schurian category $\mathcal{C}$. We show that this morphism from the vector space of abelian characters of $\pi_1(\mathcal{C})$ to the first Hochschild-Mitchell cohomology vector space of $\mathcal{C}$ is an isomorphism.

Even though the best understood class of coverings is that of Galois coverings, general non-Galois coverings have also been considered. For instance, in [17, 30], almost Galois coverings and balanced coverings, respectively, have been defined. The approach is different since the focus is to get results in the representation theory of algebras, but they also use gradings, and some notions and results may have a connection with some parts of this paper.

The authors want to thank the referee for the comments and suggestions he/she made, which contributed to improvements in the presentation of the results.

2 Fundamental group

Recall that, given a field $k$, a $k$-category is a small category $\mathcal{C}$ such that each morphism set $\mathcal{C}_x$ from an object $x \in \mathcal{C}_0$ to an object $y \in \mathcal{C}_0$ is a $k$-vector space, the composition of morphisms is $k$-bilinear and the $k$-multiples of the identity at each object are central in its endomorphism ring.

A grading $X$ of a $k$-category $\mathcal{C}$ by a group $\Gamma_X$ is a direct sum decomposition

$$\mathcal{C}_x = \bigoplus_{s \in \Gamma_X} X^s \mathcal{C}_x$$

for each $x, y \in \mathcal{C}_0$, where $X^s \mathcal{C}_x$ is called the homogeneous component of degree $s$ from $x$ to $y$, such that for $s, t \in \Gamma_X$

$$X^t \mathcal{C}_y \cdot X^s \mathcal{C}_x \subset X^{ts} \mathcal{C}_x.$$  

In case $f \in X^s \mathcal{C}_x$ and $f \neq 0$ we write $\deg_X f = s$ and we say that $f$ is homogeneous of degree $s$.

In order to define a connected grading, we consider virtual morphisms. More precisely, each non-zero morphism $f$ from its source $s(f) = x$ to its target $t(f) = y$ gives rise to a virtual morphism $(f, -1)$ from $y$ to $x$, and we put $s(f, -1) = y$ and $t(f, -1) = x$. We consider neither zero virtual morphisms nor composition of virtual morphisms. A non-zero usual morphism $f$ is identified with the virtual morphism $(f, 1)$ with the same source and target objects as $f$.

A walk $w$ in $\mathcal{C}$ is a sequence of virtual morphisms

$$(f_n, \epsilon_n), \ldots, (f_1, \epsilon_1)$$

Documenta Mathematica 16 (2011) 581–595
where $\epsilon_i \in \{+1, -1\}$, such that the target of $(f_i, \epsilon_i)$ is the source of $(f_{i+1}, \epsilon_{i+1})$.

We put $s(w) = s(f_i, \epsilon_i)$ and $t(w) = t(f_n, \epsilon_n)$.

The category $\mathcal{C}$ is connected if for any pair of objects $(x, y)$ there exists a walk $w$ from $x$ to $y$.

A homogeneous virtual morphism is a virtual morphism $(f, \epsilon)$ with $f$ homogeneous. We put $\deg_X(f, 1) = \deg_X(f)$ and $\deg_X(f, -1) = \deg_X(f)^{-1}$.

A homogeneous walk $w$ is a walk made of homogeneous virtual morphisms, and its degree is the ordered product of the degrees of the virtual morphisms.

By definition the grading $X$ is connected if for any pair of objects $(x, y)$ and any group element $s \in \Gamma_X$ there exists a homogeneous walk $w$ from $x$ to $y$ such that $\deg_X w = s$. Hence if a connected grading exists the category is necessarily connected. In case the category $\mathcal{C}$ is already connected, a grading is connected if for a fixed pair of objects $(x_0, y_0)$ there exists a homogeneous walk from $x_0$ to $y_0$ of degree $s$ for any $s \in \Gamma_X$, see [13].

In general a linear category does not admit a universal covering. However, in case a universal covering $\mathcal{U}$ exists, according to the theory developed in [13, 14], we have that the intrinsic fundamental group $\pi_1(\mathcal{C})$ is isomorphic to the automorphism group of the universal covering. In this paper we will not provide the general definition of the fundamental group since we will only consider $k$-categories with a universal covering.

3 CW-complex

Let $\mathcal{C}$ be a connected Schurian $k$-category, that is a small linear category over a field $k$ such that each vector space of morphisms is either of dimension one or zero. We choose a non-zero morphism $y e_x$ in each one-dimensional space of morphisms $y \mathcal{C}_x$, where $x e_x = x 1_x$ is the unit element of the endomorphism algebra of $x$.

Observe that $x e_y y e_x \neq 0$ is equivalent to $y e_x x e_y \neq 0$, since if $x e_y y e_x = \lambda (x 1_x)$ with $\lambda \in k^*$, then $y e_x x e_y \neq 0$ since otherwise $y e_x x e_y y e_x$ is simultaneously zero and a non-zero multiple of $y e_x$.

**Definition 3.1** The associated CW-complex $CW(\mathcal{C})$ is defined as follows

- The 0-cells are given by the set of objects $\mathcal{C}_0$.
- Each morphism $y e_x$ with $x \neq y$ gives rise to a 1-cell still denoted $y e_x$ attached to $x$ and $y$.
- If $x, y$ and $z$ are pairwise distinct objects such that $y \mathcal{C}_x, z \mathcal{C}_y$ and $z \mathcal{C}_x$ are non-zero, and $z e_y y e_x \neq 0$, we add a 2-cell attached to the 1-cells $y e_x, z e_y$ and $z e_x$.
- If $x$ and $y$ are distinct objects such that $y \mathcal{C}_x$ and $z \mathcal{C}_y$ are non-zero, and $x e_y y e_x \neq 0$ (equivalently $y e_x x e_y \neq 0$, as mentioned above), we add exactly one 2-cell attached to the 1-cells $y e_x$ and $z e_y$. 
Remark 3.2 Note that in case $x$ and $y$ are distinct objects such that $yC_x \neq 0 \neq xC_y$, two 1-cells are attached to $x$ and $y$.

Observe that in case $x, y$ and $z$ are distinct objects such that $zC_y yC_x = 0$, there is no 2-cell attached, even in case $zC_x \neq 0$.

The associated CW-complex we have just defined has no $n$-cells for $n \geq 3$, it coincides with its 2-skeleton. We do not need to go further since the fundamental group of any CW-complex coincides with the fundamental group of its 2-skeleton, see for instance [22, Chapter 2].

Example 3.3 (see [14, Corollary 4.6]) Let $D^n$ be a complete Schurian category with $n$-objects $1, \ldots, n$: for each pair of objects $(x, y)$, the morphism space $D^n_{xy}$ is one dimensional with a basis element $yC_x$, where $yC_x = 1_{xy}$.

Composition is defined by $zC_y yC_x = zC_x$ for any triple of objects. Note that the direct sum algebra of morphisms for $D^n$ is the matrix algebra $M_n(k)$.

We assert that $CW(D^n)$ is contractible, that is, it has the homotopy type of a point. Note that $CW(D^2)$ is a disk. For $n \geq 3$ consider the CW-subcomplex $L^n$ consisting of all 0-cells of $D^n$ and a chosen 1-cell attached to $i$ and $i + 1$ for $i = 1, \ldots, n - 1$ (there are no 2-cells in $L^n$). This CW-subcomplex is closed and contractible.

Consequently the quotient $CW(D^n)/L^n$ has the same homotopy type than $CW(D^n)$, see for instance [22, p.11]. Moreover $CW(D^n)/L^n$ has only one 0-cell. We assert that each 1-cell not in $L^n$ is the border of at least one disk in $CW(D^n)/L^n$. Indeed, in case of the 1-cell not in $L^n$ between $j$ and $j + 1$, for $j = 1, \ldots, n - 1$, the 2-cell attached to the two 1-cells between $j$ and $j + 1$ becomes the required disk in the quotient. In case the 1-cell is between $j$ and $j + k$, for $j = 1, \ldots, n - 2$ with $k = 2, \ldots, n - j$, the 2-cells given by the triples $(j, j + 1, j + 2), (j, j + 2, j + 3), \ldots, (j, j + k - 1, j + k)$ provide a disk in the quotient having the original 1-cell as border. Finally there are two 1-cells attached to $n$ and $1$, both are not in $L^n$ and can be identified since a 2-cell is attached to them; they are the border of the disk obtained with the 2-cells $(1, 2, 3), (1, 3, 4), \ldots, (1, n - 1, n)$.

Let $w = (f_n, \epsilon_n), \ldots, (f_1, \epsilon_1)$ be a walk in $C$ from $x$ to $y$. The inverse walk $w^{-1}$ is the walk $(f_1, -\epsilon_1), \ldots, (f_n, -\epsilon_n)$ from $y$ to $x$. Note that in case $w$ is a homogeneous walk for a grading $X$, then

$$\deg_X w^{-1} = (\deg_X w)^{-1}.$$
subdivided in $n$ intervals $I_i = \left[ \frac{i-1}{n}, \frac{i}{n} \right)$, where $I_i$ corresponds to the 1-cell defined by the non-zero morphism $f_i$ corresponding to the virtual one $(f_i, e_i)$ and where $w(\frac{i-1}{n}) = s(f_i, e_i)$ and $w(\frac{i}{n}) = t(f_i, e_i)$.

**Proposition 3.4** Let $\mathcal{C}$ be a connected Schurian $k$-category, let $X$ be a connected grading of $\mathcal{C}$ and let $u$ be a set of connector walks for $X$ for an object $c_0$. There exists a connected grading $Z_{X,u}$ of $\mathcal{C}$ by the group $\pi_1(CW(\mathcal{C}), c_0)$, where $c_0$ is considered as a base point of the CW-complex.

**Proof.** Let $u$ be a set of connector walks for $X$ and let $ye_x$ be a non-zero morphism of $yC_x$. We define its $Z_{X,u}$-degree as the homotopy class of the loop described by the walk $yu_{c_0}^{-1}, ye_x, xu_{c_0}$ in $CW(\mathcal{C})$, that is,

$\deg_{Z_{X,u}} ye_x = [yu_{c_0}^{-1}, ye_x, xu_{c_0}]$.

In order to prove that this defines a grading, let $x, y, z$ be objects in $\mathcal{C}$. In case $ze_y ye_x = 0$ there is nothing to prove. In case $ze_y ye_x \neq 0$ we have that

$ze_y ye_x = z\lambda y ye_x$

with $\lambda y$ a non-zero element in $k$. We have to show that the following equality holds:

$(\deg_{Z_{X,u}} ze_y)(\deg_{Z_{X,u}} ye_x) = \deg_{Z_{X,u}} ze_x$.

The left hand side is the following homotopy class

$[ze_y ye_x yu_{c_0}] [yu_{c_0}^{-1}, ye_x, xu_{c_0}] = [zu_{c_0}^{-1}, ze_y ye_x yu_{c_0}^{-1}, ye_x, xu_{c_0}] = [zu_{c_0}^{-1}, ze_y ye_x, xu_{c_0}]$.

Observe that since $ze_y ye_x$ is a non-zero morphism in $\mathcal{C}$, the CW-complex has a 2-cell attached, which means that the path described by the walk $ze_y ye_x$ is homotopic to $ze_x$. This observation provides the required result.

**Lemma 3.5** Let $\mathcal{C}$ be a connected Schurian category with a given base object $c_0$, let $X$ be a connected grading of $\mathcal{C}$ and let $Z_{X,u}$ be the grading considered above by the group $\pi_1(CW(\mathcal{C}), c_0)$. Let $w$ be a closed walk at $c_0$ in $\mathcal{C}$. Then

$\deg_{Z_{X,u}} w = [w] \in \pi_1(CW(\mathcal{C}), c_0)$

where $[w]$ is the homotopy class of the loop described by $w$ in $CW(\mathcal{C})$.

**Proof.** Observe first that the degree of a pure virtual morphism $(ye_x, -1)$ is the homotopy class $[yu_{c_0}^{-1}, ye_x, xu_{c_0}]^{-1} = [xu_{c_0}^{-1}, ye_x^{-1}, yu_{c_0}]$. Hence the connector walks $yu_{c_0}$ annihilate successively in $\pi_1(CW(\mathcal{C}), c_0)$, enabling us to obtain the result (recall that $c_0 u_{c_0} = c_0 I_{c_0}$).

Documenta Mathematica 16 (2011) 581–595
Proposition 3.6 Let $C$ be a connected Schurian $k$-category and let $X$ be a connected grading. Then the grading $Z_{X,u}$ obtained in Proposition 3.4 is connected.

Proof. Since $C$ is connected, it is enough to prove that for any element $[l] \in \pi_1(CW(C), c_0)$ there exists a closed walk $w$ at $c_0$ in $C$ such that $\deg_{Z_{X,u}} w = [l]$. Recall that $[l]$ is a homotopy class, more precisely $l$ is a continuous map $[0,1] \rightarrow CW(C)$ such that $l(0) = l(1) = c_0$. We use cellular approximation (see for instance [22, Theorem 4.8]) in order to obtain a homotopic cellular loop $l'$ such that the image of $l'$ is contained in the 1-skeleton. Its image is compact. A compact set in a CW-complex meets only finitely many cells (see for instance [22, Proposition A.1, page 520]). We infer that $l$ is homotopic to a loop $l'$ such that its image is a closed walk $w$ at $c_0$ in $C$. The previous Lemma asserts that the $Z_{X,u}$-degree of $w$ is precisely $[l'] = [l]$. $\diamond$

Definition 3.7 Let $X$ and $Z$ be gradings of a $k$-category $C$. We say that $X$ is a quotient of $Z$ if there exists a surjective group map $\varphi : \Gamma_Z \rightarrow \Gamma_X$ such that for any pair of objects $(x,y)$ we have that $X^y_C x = \bigoplus_{\varphi(r)=s} Z^r_y C_x$.

Theorem 3.8 Let $C$ be a connected Schurian $k$-category and let $X$ be a connected grading of $C$. Let $Z_{X,u}$ be the connected grading of $C$ by $\pi_1(CW(C), c_0)$ defined in the Proof of Proposition 3.4. Then $X$ is a quotient of $Z_{X,u}$ through a unique group map $\varphi$.

Proof. Let $[l]$ be a homotopy class in $\pi_1(CW(C), c_0)$. As in the previous proof, using cellular approximation we can assume that the image of $l$ is a closed walk $w$ at $c_0$ in $C$. In order to define a group morphism $\varphi : \pi_1(CW(C), c_0) \rightarrow \Gamma_X$ we put $\varphi([l]) = \deg_X w$. In order to check that $\varphi$ is well defined, we have to prove that $\deg_X w = \deg_X w'$ whenever $w$ and $w'$ are closed walks at $c_0$ providing homotopic loops in $CW(C)$. Assume first that $w$ and $w'$ only differ by a 2-cell, that is, $z e_y, y e_x$ is part of $w$, $z e_y, y e_x \not= 0$ and $w'$ coincide with $w$ except that $z e_y, y e_x$ is replaced by $z e_x$ through the corresponding 2-cell in $CW(C)$. Since $C$ is Schurian we have that $z e_y, y e_x$ is a non-zero multiple of $z e_x$. Now since $X$ is a grading $\deg_X (z e_y, y e_x) = \deg_X z e_x$. 

Documenta Mathematica 16 (2011) 581–595
and \( \deg_X w = \deg_X w' \).

For the general case, let \( h \) be a homotopy from \( w \) to \( w' \). Using again the result in [22, Proposition A.1, page 520], we can assume that the compact image of \( h \) meets a finite number of 2-cells. Consequently \( w \) and \( w' \) only differ by a finite number of 2-cells. By induction we obtain \( \deg_X w = \deg_X w' \).

The map is clearly a group morphism. In order to prove that \( \varphi \) is surjective, let \( s \in \Gamma_X \). Since \( X \) is connected, there exists a closed homogeneous walk \( w \) at \( c_0 \) of \( X \)-degree \( s \). Clearly there is a loop \( l \) with image \( w \), hence \( \varphi([l]) = s \).

It remains to prove that the homogeneous component of a given \( X \)-degree \( s \) is the direct sum of the corresponding \( Z_{X,u} \)-homogeneous components. Observe that since \( \mathcal{C} \) is Schurian, the direct sum decomposition is reduced to only one component. Let \( y e_x \) be a morphism which has \( X \)-degree \( s \). By definition, its \( Z_{X,u} \)-degree is \([y u_{c_0}^{-1} y e_x x u_{c_0}]\) and we have to prove that \( \varphi[y u_{c_0}^{-1} y e_x x u_{c_0}] = s \), that is, \( \deg_X(y u_{c_0}^{-1} y e_x x u_{c_0}) = s \). The result follows since the connectors \( x u_{c_0} \) have trivial \( X \)-degree.

Concerning uniqueness, let \( \varphi' : \pi_1(CW(\mathcal{C})), c_0 \to \Gamma_X \) be a surjective group map such that for each morphism \( y e_x \) we have \( \varphi'(\deg_{Z_{X,u}} y e_x) = \deg_X y e_x \), that is,

\[
\varphi' \left( [y u_{c_0}^{-1} y e_x x u_{c_0}] \right) = \varphi \left( [y u_{c_0}^{-1} y e_x x u_{c_0}] \right).
\]

This shows that \( \varphi \) and \( \varphi' \) coincide on loops of this form. Let now \( l \) be an arbitrary loop. In order to prove that \( \varphi'([l]) = \varphi([l]) \), we first replace \( l \) by a cellular approximation in such a way that \( l \) describes a walk in \( \mathcal{C} \). Clearly any loop at \( c_0 \) in \( CW(\mathcal{C}) \) is homotopic to a product of loops as above and their inverses. We infer that \( \varphi \) and \( \varphi' \) are equal on any loop.

\( \diamond \)

We will prove next that \( Z_{X,u} \) depends neither on the choice of the set \( u \) nor on the connected grading \( X \). We will consider a slightly more general situation in order to prove these facts.

First recall that a set of connector walks depends on a given grading. In case there is no grading, a set of connector walks means a set of connector walks for the trivial grading by the trivial group. In other words a set of connector walks for a linear category without a given grading is just a choice of a set of walks from a given object \( c_0 \) to each object \( x \), where the walk from \( c_0 \) to itself is \( c_0 \).

Let \( \mathcal{C} \) be a connected Schurian \( k \)-category with a base object \( c_0 \) and let \( u \) be a set of connector walks. By definition the grading \( Z_u \) of \( \mathcal{C} \) with group \( \pi_1(CW(\mathcal{C}), c_0) \) is given by \( \deg Z_u y e_x = [y u_{c_0}^{-1} y e_x x u_{c_0}] \). Next we will prove that given two sets of connector walks \( u, v \), the corresponding gradings \( Z_u \) and \( Z_v \) differ in a simple way that we will call conjugation.

**Definition 3.9** Let \( X \) be a grading of a connected \( k \)-category \( \mathcal{C} \). Let \( a = (a_x)_{x \in \mathcal{C}_0} \) be a set of group elements of \( \Gamma_X \). The conjugated grading \( a X \) has the same homogeneous components than \( X \) but the degree is changed as follows:

\[
(a X)^s y C x = X^{a_x a_x^{-1}} y C x
\]

Documenta Mathematica 16 (2011) 581–595
In order to consider morphisms between gradings, they must be understood in the setting of Galois coverings, see [14]. More precisely any grading gives rise to a Galois covering through a smash product construction, see [11]. The Galois coverings obtained by smash products form a full subcategory of the category of Galois coverings. Moreover, both categories are equivalent. Consequently morphisms between gradings are morphisms between the corresponding smash product Galois coverings.

Now, to each grading $X$ of a $k$-category we associate a new $k$-category $C\#X$ and a functor $F_X : C\#X \to C$ as follows.

\[
(C\#X)_0 = C_0 \times \Gamma_X
\]

\[
(y,t)(C\#X)_{(x,s)} = X^{t^{-1}sy}C_x
\]

\[
F_X(x,s) = x
\]

\[
F_X : (y,t)(C\#X)_{(x,s)} \hookrightarrow yC_x
\]

In particular $F_X$ is a Galois covering and any Galois covering is isomorphic to one of this type. Note that $C\#X$ is a connected category if and only if the grading $X$ is connected.

**Proposition 3.10** Let $C$ be a connected $k$-category and $X$ be a connected grading of $C$. Let $a = (a_x)_{x \in C_0}$ be a set of group elements of $\Gamma_X$ and $^\pi X$ the conjugated grading. The Galois coverings $C\#X$ and $C\#^\pi X$ are isomorphic, more precisely there exists a functor $H : C\#^\pi X \to C\#X$ such that $F_XH = F_X^a$.

**Proof.** Recall that $X^{(a_X)y}C_x = X^{a_xsy}C_x$. Consequently

\[
(y,t)(C\#^\pi X)_{(x,s)} = (X^{a_XY})^{t^{-1}sy}C_x = X^{a_xt^{-1}sa_x^{-1}y}C_x = (y,t)(C\#^a X)_{(x,s)a_x^{-1}}.
\]

This computation shows that defining $H$ on objects by $H(x,s) = (x, sa_x^{-1})$ and by the identity on morphisms provides the required isomorphism.

**Proposition 3.11** Let $C$ be a connected Schurian $k$-category, $c_0$ a base object and $X, Y$ two connected gradings of $C$. Let $Z_{X,u}$ and $Z_{Y,v}$ be the connected gradings by the group $\pi_1(CW(C),c_0)$, associated to the sets $u$ and $v$ of homogeneous connector walks for $X$ and $Y$ respectively, given by the choices $xu_{u_0}$ and $xv_{v_0}$ for any $x \in C_0$. Then $Z_{Y,v}$ and $Z_{X,u}$ are conjugated through the set of group elements $a_x = xu_{u_0}^{-1}xv_{v_0}$.

**Proof.** Recall that $\deg_{Z_{X,u}}yex = [yux_{u_0}^{-1}yex xu_{u_0}]$, then by definition

\[
\deg_{Z_{X,u}}yex = a_x^{-1}(\deg_{Z_{X,u}}yex) a_x
\]

\[
= [yv_{v_0}^{-1}yu_{u_0}yux_{u_0}^{-1}yex xu_{u_0} xu_{u_0}^{-1}xv_{v_0}]
\]

\[
= \deg_{Z_{Y,v}}yex.
\]
Remark 3.12 Since all the gradings \( Z_{X,u} \) are isomorphic, we can choose the trivial grading by the trivial group. However we still need to choose connector walks. Moreover we have shown that each connected grading is a unique quotient of the grading by the group \( \pi_1(CW(C), c_0) \).

Corollary 3.13 Let \( C \) be a connected Schurian \( k \)-category, and let \( c_0 \) be a base object. Then
\[
\pi_1(C, c_0) = \pi_1(CW(C), c_0).
\]

Proof. From [13], we know that in case a universal covering exists, the fundamental group of the category is its group of automorphisms. The results we have proven show that the grading by the fundamental group of \( CW(C) \) is a universal grading, consequently the smash product Galois covering is a universal Galois covering with automorphism group \( \pi_1(CW(C), c_0) \).

Next we compute the intrinsic fundamental group of a \( k \)-category with an infinite number of objects and without admissible presentation.

Example 3.14 Let \( C \) be the \( k \)-category given by the quiver:

\[
\begin{array}{c}
\vdots \\
\downarrow \\
a_1 \\
\downarrow \\
\beta_1 \\
\alpha_1 \\
b_1 \\
\uparrow \\
\vdots \\
a_0 \\
\downarrow \\
\beta_0 \\
\alpha_0 \\
b_0 \\
\uparrow \\
\vdots \\
a_{-1} \\
\downarrow \\
\beta_{-1} \\
\alpha_{-1} \\
b_{-1} \\
\uparrow \\
\vdots
\end{array}
\]

with the relations \( \gamma_\alpha \alpha_i \beta_{i+1} = \alpha_{i+1} \) for all \( i \neq 0 \) and \( \gamma_0 \alpha_0 \beta_1 = 0 \).

In \( CW(C) \) there is a 2-cell attached to each square except the 0-one. Consequently \( \pi_1(C) = \mathbb{Z} \).

4 Hurewicz isomorphism

Let \( C \) be a \( k \)-category. A \( k \)-derivation \( d \) with coefficients in \( C \) is a set of linear morphisms \( y_d x : y_c^C x \rightarrow y_c^C x \) for each pair \((x, y)\) of objects, verifying
\[
z_d_x(gf) = z_d_y(g)f + g_y d_x(f)
\]
for any \( f \in \gamma C_x \) and \( g \in \gamma C_y \). Let \( a = (a_x)_{x \in C_0} \) be a family of endomorphisms of each object \( x \in C_0 \), namely \( a_x \in \gamma C_x \). The inner derivation \( d_a \) associated to \( a \) is defined by

\[
y(d_a)_x(f) = a_y f - f a_x.
\]

The first Hochschild-Mitchell cohomology \( HH^1(C) \) is the quotient of the vector space of derivations by the subspace of inner ones (see [29] for the general definition).

**Remark 4.1** In fact \( HH^1(C) \) has a Lie algebra structure, where the bracket of derivations is given by

\[
y[d, d']_x = y d_x y d'_x - y d'_x y d_x.
\]

**Definition 4.2** Let \( X \) be a grading of a \( k \)-category \( C \). The Hurewicz morphism

\[
h : \text{Hom}(\Gamma_X, k^+) \rightarrow HH^1(C)
\]

is defined as follows. Let \( \chi : \Gamma_X \rightarrow k^+ \) be an abelian character and let \( f \) be a homogeneous morphism in \( \gamma C_x \). Then

\[
yh(\chi)_x(f) = \chi(\deg_X f) f.
\]

An arbitrary morphism is decomposed as a sum of its homogeneous components in order to extend linearly the definition of \( yh(\chi)_x \).

**Remark 4.3** The set \( h(\chi) \) is a derivation. This can be verified in a simple way relying on the fact that \( X \) is a grading. Derivations of this type are called “Eulerian derivations”, see for instance [18, 19].

The following result is immediate.

**Lemma 4.4** The image of the Hurewicz morphism is an abelian Lie subalgebra of \( HH^1(C) \).

We recall that, under some assumptions, the Hurewicz morphism is injective.

**Proposition 4.5** Let \( C \) be a \( k \)-category and assume the endomorphism algebra \( \gamma C_x \) of each object \( x \) in \( C_0 \) is equal to \( k \). Let \( X \) be a connected grading of \( C \). Then the Hurewicz morphism is injective.

**Proof.** If \( h(\chi) \) is an inner derivation,

\[
\epsilon(f) h(\chi)(f) = \chi(\deg_X f) f = a(f) f - f a(f)
\]

for any homogeneous non-zero morphism \( f \), where \( (a_x)_{x \in C_0} \) is a set of endomorphisms which are elements of \( k \) by hypothesis. Then \( \chi(\deg_X f) = a(f) - a_x(f) \).
Now we assert that the same equality holds for any homogeneous walk $w$, that is, 
\[ \chi(\deg_X w) = a_t(w) - a_s(w). \]

For instance let $w = (g, -1), (f, 1)$ be a homogeneous walk where $f \in yC_x$ and $g \in yC_z$. Then

\[
\chi(\deg_X w) = \chi((\deg_X g)^{-1}(\deg_X f)) = -\chi(\deg_X g) + \chi(\deg_X f) = a_s(g) + a_t(f) - a_s(f) = a_z - a_y + a_y - a_x = a_z - a_x = a_t(w) - a_s(w).
\]

Let $c_0$ be any fixed object of $C$. Since $X$ is a connected grading, for any group element $s \in \Gamma_X$ there exists a homogeneous walk $w$, closed at $c_0$, such that $\deg_X w = s$. Consequently

\[
\chi(s)w = (a_{c_0} - a_{c_0})w = 0
\]

hence $\chi(s) = 0$ for any $s \in \Gamma_X$.

**Theorem 4.6** Let $C$ be a connected Schurian $k$-category and let $U$ be its universal grading by the fundamental group $\pi_1(CW(C), c_0)$. The corresponding Hurewicz morphism is an isomorphism.

**Proof.** The previous result insures that $h$ is injective. In order to prove that $h$ is surjective, let $d$ be a derivation. We choose a non-zero morphism $y e_x$ in each 1-dimensional space of morphisms $yC_x$, with $x e_x = x I_x$. Let $c_0$ be a fixed object in $C$. To describe the universal grading, recall that we choose a set of connector walks, hence

\[
\deg_{\pi_1(CW(C))} y e_x = [y u_c^{-1} y e_x x u_c] \in \pi_1(CW(C), c_0).
\]

Since $yC_x$ is one dimensional, $d(y e_x) = y \lambda_x y e_x$ with $y \lambda_x \in k$. In order to define an abelian character $\chi$ such that $h(\chi) = d$, let $l$ be a loop at $c_0$ in $CW(C)$. By cellular approximation we can assume that the image of $l$ is a closed walk $w$ in $C$. In case $w$ is of the form $y u_c^{-1} y e_x x u_c$, we define $\chi[l] = y \lambda_x$. Otherwise the cellular loop $w$ is homotopic to a product of loops of the previous type or of their inverses, and we define $\chi[l]$ to be the corresponding sum of scalars. We have to verify that $\chi$ is well defined. First observe that if $z e_y y e_x \neq 0$, the scalars of the derivation $d$ verify

\[ z \lambda_x = z \lambda_y + y \lambda_x. \]

Indeed, $z e_y y e_x = \mu z e_x$, with $\mu \neq 0$, hence

\[
d(z e_y y e_x) = z e_y d(y e_x) + d(z e_y) y e_x = \mu (z \lambda_y + y \lambda_x) z e_x.
\]
We deduce the result since \( d(\mu z e_x) = \mu x \lambda_x z e_x \). Consider now two cellular loops \( l \) and \( l' \) which are homotopic by a 2-cell, meaning that a walk \( z e_y, y e_x \) is replaced by \( z e_x \). The previous computation shows that \( \chi[l] = \chi[l'] \). We have already verified that any homotopy of cellular loops decomposes as a finite number of homotopies of the previous type, hence we deduce that \( \chi \) is a well defined map. By construction \( \chi : \pi_1(CW(C), c_0) \to k^+ \) is an abelian character and clearly \( h(\chi) = d \).

References

[1] Assem, I.; de la Peña, J. A. The fundamental groups of a triangular algebra. Comm. Algebra 24 (1996), 187–208.

[2] Bahturin, Y.; Zaicev, M. Group gradings on matrix algebras, Canad. Math. Bull. 45 (2002), no. 4, 499–508.

[3] Bardzell, M. J.; Marcos, E. N. \( H^1 \) and presentations of finite dimensional algebras. Representations of algebras (São Paulo, 1999), 31–38, Lecture Notes in Pure and Appl. Math., 224, Dekker, New York, 2002.

[4] Boboc, C.; Dăscălescu, S. Group gradings on \( M_3(k) \), Comm. Algebra 35 (2007), 2654–2670.

[5] Bongartz, K.; A criterion for finite representation type. Math. Ann. 269 (1984), no. 1, 1–12.

[6] Bongartz, K.; Gabriel, P. Covering spaces in representation-theory, Invent. Math. 65 (1981/82) 331–378.

[7] Bretscher, O.; Gabriel, P. The standard form of a representation-finite algebra. Bull. Soc. Math. France 111 (1983), no. 1, 21–40.

[8] Bustamante, J.C. On the fundamental group of a Schurian algebra. Comm. Algebra 30 (2002), no. 11, 5307–5329.

[9] Bustamante, J.C. The classifying space of a bound quiver. J. Algebra 277 (2004), no. 2, 431–455.

[10] Bustamante, J.C.; Castonguay, D. Fundamental groups and presentations of algebras. J. Algebra Appl. 5 (2006), 549–562.

[11] Cibils, C.; Marcos, E. Skew category, Galois covering and smash product of a category over a ring. Proc. Amer. Math. Soc. 134, (2006), no. 1, 39–50.

[12] Cibils, C.; Redondo M.J. Cartan-Leray spectral sequence for Galois coverings of categories. J. Algebra 284 (2005), 310–325.
[13] Cibils, C.; Redondo M.J.; Solotar, A. The intrinsic fundamental group of a linear category. To appear in Algebras and Representation Theory. DOI: 10.1007/s10468-010-9263-1.

[14] Cibils, C.; Redondo M.J.; Solotar, A. Connected gradings and fundamental group. Algebra Number Theory 4 (2010), no. 5, 625–648.

[15] Cibils, C.; Solotar, A. Galois coverings, Morita equivalence and smash extensions of categories over a field. Doc. Math. 11 (2006), 143–159.

[16] Donady, R.; Donady, A. Algèbre et théories galoisiennes. Paris: Cassini. 448 p. (2005).

[17] Dowbor, P; Hajduk, A. On some nice class of non-Galois coverings for algebras, Bol. Soc. Mat. Mex. 39 (in press).

[18] Farkas, D.; Green, E.L.; Marcos, E.N. Diagonalizable derivations of finite-dimensional algebras. II. Pacific J. Math. 196 (2000), no. 2, 341–351.

[19] Farkas, D.R.; Geiss, Ch.; Green, E.L.; Marcos, E.N. Diagonalizable derivations of finite-dimensional algebras. I. Israel J. Math. 117 (2000), 157–181.

[20] Gabriel, P. The universal cover of a representation-finite algebra. Representations of algebras (Puebla, 1980), 68–105, Lecture Notes in Math. 903, Springer, Berlin-New York, 1981.

[21] Green, Edward L. Graphs with relations, coverings and group-graded algebras. Trans. Amer. Math. Soc. 279 (1983), no. 1, 297–310.

[22] Hatcher, A. Algebraic topology. Cambridge University Press, Cambridge, 2002. xii+544 pp.

[23] Khazal, R.; Boboc, C.; Dăscălescu, S. Group gradings of $M_2(K)$, Bull. Austral. Math. Soc. 68 (2003), no. 2, 285–293.

[24] Le Meur, P. The universal cover of an algebra without double bypass, J. Algebra 312 (2007), no. 1, 330–353.

[25] Le Meur, P. The fundamental group of a triangular algebra without double bypasses. C. R. Math. Acad. Sci. Paris 341 (2005), 211–216.

[26] Le Meur, P. Revêtements galoisiens et groupe fondamental d’algèbres de dimension finie. Ph.D. thesis, Université Montpellier 2 (2006). http://tel.archives-ouvertes.fr/tel-00011753

[27] Martínez-Villa, R.; de la Peña, J. A. The universal cover of a quiver with relations. J. Pure Appl. Algebra 30 (1983), 277–292.

[28] Martins, Ma. I. R.; de la Peña, J. A. Comparing the simplicial and the Hochschild cohomologies of a finite-dimensional algebra. J. Pure Appl. Algebra 138 (1999), no. 1, 45–58.

Documenta Mathematica 16 (2011) 581–595
[29] Mitchell, B. Rings with several objects. Adv. Math. 8 (1972), 1–161.

[30] de la Peña, J.A.; Redondo, M.J. Coverings without groups. J. Algebra 321 (2009), 3816–3826.

[31] de la Peña, J.A.; Saorín, M. On the first Hochschild cohomology group of an algebra. Manuscripta Math. 104 (2001), 431–442.

[32] Reynaud, E. Algebraic fundamental group and simplicial complexes. J. Pure Appl. Algebra 177 (2003), no. 2, 203–214.

Claude Cibils
Institut de mathématiques et de modélisation de Montpellier I3M
UMR 5149
Université Montpellier 2
F-34095 Montpellier cedex 5
France
Claude.Cibils@math.univ-montp2.fr

María Julia Redondo
Departamento de Matemática
Universidad Nacional del Sur
Av. Alem 1253
8000 Bahía Blanca
Argentina
mredondo@criba.edu.ar

Andrea Solotar
Departamento de Matemática
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires
Ciudad Universitaria, Pabellón 1
1428, Buenos Aires
Argentina
asolotar@dm.uba.ar
CHARACTERISING WEAK-OPERATOR
CONTINUOUS LINEAR FUNCTIONALS ON $\mathcal{B}(H)$ CONSTRUCTIVELY

DOUGLAS S. BRIDGES

Received: February 7, 2011
Communicated by Anil Nerode

Abstract. Let $\mathcal{B}(H)$ be the space of bounded operators on a not-necessarily-separable Hilbert space $H$. Working within Bishop-style constructive analysis, we prove that certain weak-operator continuous linear functionals on $\mathcal{B}(H)$ are finite sums of functionals of the form $T \rightsquigarrow \langle Tx, y \rangle$. We also prove that the identification of weak- and strong-operator continuous linear functionals on $\mathcal{B}(H)$ cannot be established constructively.

2010 Mathematics Subject Classification: 03F60, 47L50, 46S30
Keywords and Phrases: Constructive, operators, (ultra)weak operator topology, continuous functionals

1 Introduction

Let $H$ be a complex Hilbert space that is nontrivial (that is, contains a unit vector), $\mathcal{B}(H)$ the space of all bounded operators on $H$, and $\mathcal{B}_1(H)$ the unit ball of $\mathcal{B}(H)$. In this paper we carry out, within Bishop-style constructive mathematics (BISH),\(^1\) an investigation of weak-operator continuous linear functionals on $\mathcal{B}(H)$.

Depending on the context, we use, for example, $x$ to represent either the element $(x_1, \ldots, x_N)$ of the finite direct sum $H_N \equiv \bigoplus_{n=1}^{N} H$ of $N$ copies of $H$ or else the element $(x_n)_{n \geq 1}$ of the direct sum $H_{\infty} \equiv \bigoplus_{n \geq 1} H$ of a sequence of copies of $H$. We use $I$ to denote the identity projection on $H$.

The following are the topologies of interest to us here.

\(^{1}\)That is, mathematics that uses only intuitionistic logic and is based on a suitable set- or type-theoretic foundation [1, 2, 12]. For more on BISH see [3, 4, 8].
The weak operator topology: the weakest topology on $B(H)$ with respect to which the mapping $T \mapsto \langle Tx, y \rangle$ is continuous for all $x, y \in H$; sets of the form
\[ \{ T \in B(H) : |\langle Tx, y \rangle| < \varepsilon \}, \]
with $x, y \in H$ and $\varepsilon > 0$, form a sub-base of weak-operator neighbourhoods of 0 in $B(H)$.

The ultraweak operator topology: the weakest topology on $B(H)$ with respect to which the mapping $T \mapsto \sum_{n=1}^{\infty} \langle Tx_n, y_n \rangle$ is continuous for all $x, y \in H_{\infty}$; sets of the form
\[ \left\{ T \in B(H) : \left| \sum_{n=1}^{\infty} \langle Tx_n, y_n \rangle \right| < \varepsilon \right\}, \]
with $x, y \in H_{\infty}$ and $\varepsilon > 0$, form a sub-base of ultraweak-operator neighbourhoods of 0 in $B(H)$.

These topologies are induced, respectively, by the seminorms of the form $T \mapsto |\langle Tx, y \rangle|$ with $x, y \in H$; and those of the form $T \mapsto |\sum_{n=1}^{\infty} \langle Tx_n, y_n \rangle|$ with $x, y \in H_{\infty}$.

An important theorem in classical operator algebra theory states that the weak-operator continuous linear functionals on (any linear subspace of) $B(H)$ all have the form $T \mapsto \sum_{n=1}^{N} \langle Tx_n, y_n \rangle$ with $x, y \in H$ for some $N$; and the ultraweak-operator continuous linear functionals on $B(H)$ have the form $T \mapsto \sum_{n=1}^{\infty} \langle Tx_n, y_n \rangle$, where $x, y \in H_{\infty}$. However, the classical proofs, such as those found in [10, 11, 14], depend on applications of nonconstructive versions of the Hahn-Banach theorem, the Riesz representation theorem, and polar decomposition.

The foregoing characterisation of ultraweak-operator continuous functionals was derived constructively, when $H$ is separable, in [9]. A variant of it was derived in [8] (Proposition 5.4.16) without the requirement of separability, and using not the standard ultraweak operator topology but one that is classically, though not constructively, equivalent to it. Our aim in the present work is to provide a constructive proof of the standard classical characterisation of weak-operator continuous linear functionals (Theorem 10) on $B(H)$, without the requirement of separability but with one hypothesis in addition to the classical ones. In presenting this work, we emphasise that, in contrast to their classical counterparts, our proofs contain extractable, implementable algorithms for the desired representation of weak-operator continuous linear functionals; moreover, the constructive proofs themselves verify that those algorithms meet their specifications.
2 Preliminary Lemmas

The proof of our main theorem depends on a sequence of (at-times-complicated) lemmas. For the first one, we remind the reader of two elementary definitions in constructive analysis: we say that an inhabited set $S$—that is, one in which we can construct an element—is finitely enumerable if there exist a positive integer $N$ and a mapping of $\{1, \ldots, N\}$ onto $S$; if that mapping is one-one, then $S$ is called finite.

**Lemma 1** If $u$ is a weak-operator continuous linear functional on $\mathcal{B}(H)$, then there exist a finitely enumerable subset $F$ of $H \times H$ and a positive number $C$ such that $|u(T)| \leq C \sum_{(x,y)\in F} |\langle Tx, y \rangle|$ for all $T \in \mathcal{B}(H)$.

**Proof.** This is an immediate consequence of Proposition 5.4.1 in [8].

We shall need some information about locally convex spaces. Let $(p_j)_{j \in J}$ be a family of seminorms defining the topology on a locally convex linear space $V$, and let $A$ be a subset of $V$. A subset $S$ of $A$ is said to be located (in $A$) if

$$\inf \left\{ \sum_{j \in F} p_j(x - s) : s \in S \right\}$$

exists for each $x \in A$ and each finitely enumerable subset $F$ of $J$. We say that $A$ is TOTALLY BOUNDED if for each finitely enumerable subset $F$ of $J$ and each $\varepsilon > 0$, there exists a finitely enumerable subset $S$ of $A$—called an $\varepsilon$-APPROXIMATION TO $S$ RELATIVE TO $(p_j)_{j \in F}$—such that for each $x \in A$ there exists $s \in S$ with $\sum_{j \in F} p_j(x - s) < \varepsilon$.

The unit ball $B_1(H)$ is weak-operator totally bounded ([8], Proposition 5.4.15); but, in contrast to the classical situation, it cannot be proved constructively that $B_1(H)$ is weak-operator complete [5].

A mapping $f$ between locally convex spaces $(X, (p_j)_{j \in J})$ and $(Y, (q_k)_{k \in K})$ is UNIFORMLY CONTINUOUS on a subset $S$ of $X$ if for each $\varepsilon > 0$ and each finitely enumerable subset $F$ of $J$ and each finitely enumerable subset $G$ of $K$, there exist $\delta > 0$ and a finitely enumerable subset $F$ of $J$ such that if $x, x' \in S$ and $\sum_{j \in F} p_j(x - x') < \delta$, then $\sum_{k \in G} q_k (f(x) - f(x')) < \varepsilon$.

We recall four facts about total boundedness, locatedness, and uniform continuity in a locally convex space $V$. The proofs are found on pages 129–130 of [8].

- If $f$ is a uniformly continuous mapping of a totally bounded subset $A$ of $V$ into a locally convex space, then $f(A)$ is totally bounded.
- If $f$ is a uniformly continuous, real-valued mapping on a totally bounded subset $A$ of $V$, then $\sup_{x \in A} f(x)$ and $\inf_{x \in A} f(x)$ exist.
- A totally bounded subset of $V$ is located in $V$. 

**Documenta Mathematica** 16 (2011) 597–617
Douglas S. Bridges

> If \( A \subset V \) is totally bounded and \( S \subset A \) is located in \( A \), then \( S \) is totally bounded.

We remind the reader that a bounded linear mapping \( T : X \to Y \) between normed linear spaces is normed if its norm,

\[
\|T\| \equiv \sup \{ \|Tx\| : x \in X, \|x\| \leq 1 \},
\]

exists. If \( X \) is finite-dimensional, then \( \|T\| \) exists; but the statement ‘Every bounded linear functional on an infinite-dimensional Hilbert space is normed’ is essentially nonconstructive.

**Lemma 2** Every weak-operator continuous linear functional on \( B(H) \) is normed.

**Proof.** This follows from observations made above, since, in view of Lemma 1, the linear functional is weak-operator uniformly continuous on the weak-operator totally bounded set \( B_1(H) \).

We note the following stronger form of Lemma 1.

**Lemma 3** Let \( u \) be a weak-operator continuous linear functional on \( B(H) \). Then there exist \( \delta > 0 \), and finitely many nonzero \( \xi_1, \ldots, \xi_N \) and \( \zeta_1, \ldots, \zeta_N \) of \( H \) with \( \sum_{n=1}^N \|\xi_n\|^2 = \sum_{n=1}^N \|\zeta_n\|^2 = 1 \), such that \( |u(T)| \leq \delta \sum_{n=1}^N |\langle T\xi_n, \zeta_n \rangle| \) for each \( T \in B(H) \).

**Proof.** By Lemma 1, there exist a positive integer \( \nu, C > 0 \), and vectors \( x, y \in H_\nu \) such that \( |u(T)| \leq C \sum_{n=1}^\nu |\langle Tx_n, y_n \rangle| \) for all \( T \in B(H) \). For each \( n \leq \nu \), construct nonzero vectors \( x_n', y_n' \) such that \( x_n' \neq x_n \) and \( y_n' \neq y_n \). The desired result follows from the inequality

\[
\sum_{n=1}^\nu |\langle Tx_n, y_n \rangle| \leq \sum_{n=1}^\nu |\langle T(x_n - x_n'), y_n - y_n' \rangle| + \sum_{n=1}^\nu |\langle Tx_n', y_n' \rangle| + \sum_{n=1}^\nu |\langle T(x_n - x_n'), y_n' \rangle| + \sum_{n=1}^\nu |\langle T(x_n', y_n' \rangle|
\]

the fact that each of the vectors \( x_n', x_n - x_n', y_n, y_n' \) is nonzero, and scaling to get the desired norm sums equal to 1 and then the positive \( \delta \).

The next lemma will be used in an application of the separation theorem in the proof of Lemma 6.

---

In fact, a nonzero linear functional on a normed space is normed if and only its kernel is located ([8], Proposition 2.3.6).

A vector in a locally convex space is nonzero if it is mapped to a positive number by at least one seminorm.

At this stage, it is trivial to prove Lemma 3 classically by simply deleting terms \( \langle Tx_n, y_n \rangle \) when either \( x_n \) or \( y_n \) is 0. With intuitionistic logic we need to work a little harder, because we cannot generally decide whether a given vector in \( H \) is, or is not, equal to 0.
Lemma 4 Let $\zeta_1, \ldots, \zeta_N$ be elements of $H$ with $\sum_{n=1}^{N} \|\zeta_n\|^2 = 1$. Let $K$ be a finite-dimensional subspace of $H_N$, and let $\|\| \|$ be the standard norm on the dual space $K^*$ of $K$:

$$\|f\| = \sup \{|f(x)| : x \in K, \|x\| \leq 1\} \quad (f \in K^*).$$

Define a mapping $F$ of $B(H)$ into $(K^*, \|\|)$ by

$$F(T)(x) \equiv \sum_{n=1}^{N} \langle Tx_n, \zeta_n \rangle \quad (x \in K).$$

Then $F$ is weak-operator uniformly continuous on $B_1(H)$.

Proof. Given $\varepsilon > 0$, let $\{x_1, \ldots, x_m\}$ be an $\varepsilon$-approximation to the (compact) unit ball of $K$. Writing $x_i = (x_{i,1}, \ldots, x_{i,N})$, consider $S, T \in B_1(H)$ with

$$\sum_{i=1}^{m} \sum_{n=1}^{N} |\langle (S - T)x_{i,n}, \zeta_n \rangle| < \varepsilon.$$

For each $x$ in the unit ball of $K$, there exists $i$ such that $\|x - x_i\| < \varepsilon$. We compute

$$|F(S)(x) - F(T)(x)| \leq |F(S)(x) - F(S)(x_i)| + |F(S)(x_i) - F(T)(x_i)|$$

$$+ |F(T)(x) - F(T)(x_i)|$$

$$\leq \sum_{n=1}^{N} |\langle S(x_n - x_{i,n}), \zeta_n \rangle| + \sum_{n=1}^{N} |\langle (S - T)x_{i,n}, \zeta_n \rangle|$$

$$+ \sum_{n=1}^{N} |\langle T(x_n - x_{i,n}), \zeta_n \rangle|$$

$$\leq 2 \sum_{n=1}^{N} \|x_n - x_{i,n}\| \|\zeta_n\| + \varepsilon$$

$$\leq 2 \|x - x_i\| \|\zeta\| + \varepsilon < 3\varepsilon.$$

Hence $\|F(S) - F(T)\| \leq 3\varepsilon$. Since $\varepsilon > 0$ is arbitrary, we conclude that $F$ is uniformly continuous on $B_1(H)$.

In order to ensure that the unit kernel $B_1(H) \cap \ker u$ of a weak-operator continuous linear functional $u$ on $B(H)$ is weak-operator totally bounded, and hence weak-operator located, we derive a generalisation of Lemma 5.4.9 of [8].

Lemma 5 Let $\left(V, (p_j)_{j \in J}\right)$ be a locally convex space. Let $V_1$ be a balanced, convex, and totally bounded subset of $V$. Let $u$ be a linear functional on $V$ that, on $V_1$, is both uniformly continuous and nonzero. Then $V_1 \cap \ker u$ is totally bounded.
Proof. Since $u$ is nonzero and uniformly continuous on the totally bounded set $V_1,$
\[
C = \sup\{|u(y)| : y \in V_1\}
\]
exists and is positive. Choose $y_1$ in $V_1$ such that $u(y_1) > C/2.$ Then
\[
y_0 \equiv \frac{C}{2u(y_1)} y_1
\]
belongs to the balanced set $V_1,$ and $u(y_0) = C/2.$ Let $\varepsilon > 0,$ and let $F$ be a finitely enumerable subset of $J.$ Since each $p_j$ is uniformly continuous on $V,$ it maps the totally bounded set $V_1$ onto a totally bounded subset of $\mathbb{R}.$\(^6\) Hence there exists $b > 0$ such that $p_j(x) \leq b$ for each $j \in F$ and each $x \in V_1.$ Using Theorem 5.4.6 of [8], compute $t$ with
\[
0 < t < \frac{C\varepsilon}{C + 4b}
\]
such that
\[
S_t = \{y \in V_1 : |u(y)| \leq t\}
\]
is totally bounded. Pick a $t$-approximation $\{s_1, \ldots, s_n\}$ of $S_t$ relative to $\langle p_j \rangle_{j \in F},$ and set
\[
y_k = \frac{C}{C + 2t}s_k - \frac{2}{C + 2t}u(s_k)y_0 \quad (1 \leq k \leq n).
\]
Then $y_k \in \ker(u).$ Since $|u(s_k)| \leq t$ and $V_1$ is balanced,
\[
\frac{-u(s_k)}{t}y_0 \in V_1.
\]
Thus
\[
y_k = \frac{C}{C + 2t}s_k + \left(1 - \frac{C}{C + 2t}\right)\left(\frac{-u(s_k)}{t}y_0\right) \in V_1.
\]
\(^6\)We use $\mathbb{R}$ and $\mathbb{C}$ for the sets of real and complex numbers, respectively.

Documenta Mathematica 16 (2011) 597–617
Now consider any element $y$ of $V_1 \cap \ker(u)$. Since $y \in S_t$, there exists $k$ such that $\sum_{j \in F} p_j (y - s_k) < t$ and therefore

$$\sum_{j \in F} p_j (y - y_k) \leq \sum_{j \in F} p_j (y - s_k) + \sum_{j \in F} p_j (s_k - y_k)$$

$$< t + \frac{2}{C + 2t} \sum_{j \in F} p_j (ts_k + u(s_k)y_0)$$

$$\leq t + \frac{2}{C + 2t} \sum_{j \in F} (tp_j (s_k) + u(s_k)p_j (y_0))$$

$$\leq t + \frac{2t}{C} \sum_{j \in F} (p_j (s_k) + p_j (y_0))$$

$$\leq t \left( 1 + \frac{4b}{C} \right) < \varepsilon.$$

Thus $\{y_1, \ldots, y_n\}$ is a finitely enumerable $\varepsilon$-approximation to $V_1 \cap \ker(u)$ relative to the family $(p_j)_{j \in F}$ of seminorms.

The next lemma, the most complicated in the paper, extracts much of the sting from the proof of our main theorem by showing how to find finitely many mappings of the form $T \mapsto \langle Tx, \zeta \rangle$ whose sum is small on the unit kernel of $u$.

**Lemma 6** Let $u$ be a nonzero weak-operator continuous linear functional on $B(H)$. Let $\delta$ be a positive number, and $\xi_1, \ldots, \xi_N$ and $\zeta_1, \ldots, \zeta_N$ nonzero elements of $H$, such that

$$\sum_{n=1}^{N} \|\xi_n\|^2 = \sum_{n=1}^{N} \|\zeta_n\|^2 = 1,$$

and

$$|u(T)| \leq \delta \sum_{n=1}^{N} |\langle T\xi_n, \zeta_n \rangle| \quad (T \in B(H)).$$

(1)

Then for each $\varepsilon > 0$, there exists a unit vector $x$ in the subspace

$$K \equiv \mathbb{C}\xi_1 \times \mathbb{C}\xi_2 \times \cdots \times \mathbb{C}\xi_N$$

of $H_N$, such that $x_n \neq 0$ for $1 \leq n \leq N$ and $\sum_{n=1}^{N} (T x_n, \zeta_n) \leq \varepsilon$ for all $T \in B_1(H) \cap \ker u$.

**Proof.** First note that since each $\xi_n$ is nonzero, $K$ is an $N$-dimensional subspace of $H_N$. Now, an application of Lemma 5 tells us that the unit kernel

\footnote{Such $\xi_k, \zeta_k$, and $\delta$ exist, by Lemma 3.}
$B_1(H) \cap \ker u$ of $u$ is weak-operator totally bounded. For each $x \in H_N$, since the mapping $T \mapsto \sum_{n=1}^N (Tx_n, \zeta_n)$ is weak-operator uniformly continuous on the unit kernel, we see that

$$\|x\|_0 = \sup \left\{ \sum_{n=1}^N (Tx_n, \zeta_n) : T \in B_1(H) \cap \ker u \right\}$$

exists. The mapping $x \mapsto \|x\|_0$ is a seminorm on $H_N$ satisfying $\|x\|_0 \leq \|\zeta\| \|x\| = \|x\|$; whence the identity mapping from $(H_N, \|\|)$ to $(H_N, \|\|_0)$ is uniformly continuous. Since the subset

$$\{x \in K : \|x\| = 1\}$$

of the finite-dimensional Banach space $(K, \|\|)$ is totally bounded, it follows that

$$\beta \equiv \inf \{\|x\|_0 : x \in K, \|x\| = 1\},$$

exists. It will suffice to prove that $\beta = 0$. For then, given $\varepsilon$ with $0 < \varepsilon < 1$, we can construct a unit vector $x' \in K$ such that $\left| \sum_{n=1}^N (Tx'_n, \zeta_n) \right| < \varepsilon/2$ for all $T \in B_1(H) \cap \ker u$. Picking nonzero vectors $y_n \in C\xi_n$ such that

$$\left( \sum_{n=1}^N \|x'_n - y_n\|^2 \right)^{1/2} < \varepsilon/8,$$

we have

$$1 - \left( \sum_{n=1}^N \|y_n\|^2 \right)^{1/2} \leq \varepsilon/8,$$

so

$$x = \left( \sum_{n=1}^N \|y_n\|^2 \right)^{-1/2} y$$

is a unit vector in $C\xi_1 \times \cdots \times C\xi_N$ with each $x_n \neq 0$. Moreover,

$$\|x - y\|^2 = \sum_{k=1}^N \left( \sum_{n=1}^N \|y_n\|^2 \right)^{-1/2} - 1 \|y_k\|^2$$

$$\leq \left( \frac{\varepsilon}{8} \right)^2 \sum_{k=1}^N \|y_k\|^2$$

$$\leq \frac{\varepsilon^2}{64} \left( 1 + \frac{\varepsilon}{8} \right)^2 < \frac{\varepsilon^2}{16}.$$
so for each \( T \in B_1(H) \cap \ker u \),
\[
\sum_{n=1}^{N} \langle Tx_n, \zeta_n \rangle \leq \sum_{n=1}^{N} \langle Tx'_n, \zeta_n \rangle + \sum_{n=1}^{N} |\langle (x_n - x'_n), \zeta_n \rangle| \\
\leq \frac{\varepsilon}{2} + \sum_{n=1}^{N} \|x_n - x'_n\| \|\zeta_n\| \\
\leq \frac{\varepsilon}{2} + \|x - x'\| \|\zeta\| \\
\leq \frac{\varepsilon}{2} + \|x - y\| + \|y - x'\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{4} + \frac{\varepsilon}{8} < \varepsilon.
\]

To prove that \( \beta = 0 \), we suppose that \( \beta > 0 \). Then \( \|\|_0 \) is a norm equivalent to the original norm on \( K \), so \((K, \|\|_0)\) is an \( N \)-dimensional Banach space. Define norms \( \|\|_* \) and \( \|\|_{0_*} \) on the dual \( K^* \) of \( K \) by
\[
\|f\|_* \equiv \sup \{ |f(x)| : x \in K, \|x\| \leq 1 \}, \\
\|f\|_{0_*} \equiv \sup \{ |f(x)| : x \in K, \|x\|_0 \leq 1 \}.
\]

For each \( T \in B(H) \) and each \( x \in K \) let
\[
F(T)(x) \equiv \sum_{n=1}^{N} \langle Tx_n, \zeta_n \rangle.
\]

Then, by Lemma 4, \( F \) is weak-operator uniformly continuous as a mapping of \( B_1(H) \) into \((K^*, \|\|_*); \) since the norms \( \|\|_* \) and \( \|\|_{0_*} \) are equivalent on the finite-dimensional dual space \( K^* \), \( F \) is therefore weak-operator uniformly continuous as a mapping of \( B_1(H) \) into \((K^*, \|\|_{0_*}). \) Hence
\[
D = F(B_1(H) \cap \ker u)
\]
is a totally bounded, and therefore located, subset of \((K^*, \|\|_*). \) Moreover, for each \( T \in B_1(H) \cap \ker u \) and each \( x \in K \), \( |F(T)(x)| \leq \|x\|_{0}; \) so \( D \) is a subset of the unit ball \( S_0^* \) of \((K^*, \|\|_{0_*}). \) We shall use the separation theorem from functional analysis to prove that \( D \) is \( \|\|_{0_*}-\)dense in \( S_0^* \). Consider any \( \phi \) in \( S_0^* \), and suppose that
\[
0 < d = \inf \{ \|\phi - F(T)\|_{0_*} : T \in B_1(H) \cap \ker u \}.
\]

Now, \( D \) is bounded, convex, balanced, and located; so, by Corollary 5.2.10 of [8], there exists a linear functional \( v \) on \((K^*, \|\|_{0_*})\) with norm 1 such that
\[
v(\phi) > |v(F(T))| + \frac{d}{2} \quad (T \in B_1(H) \cap \ker u).
\]
It is a simple exercise\(^8\) to show that since \((K^*, \|\|_0)\) is \(N\)-dimensional, there exists \(y \in K\) such that \(\|y\|_0 = 1\) and \(v(f) = f(y)\) for each \(f \in K^*\). Hence

\[
\phi(y) \geq \sup \{|F(T)(y)| : T \in B_1(H) \cap \ker u\} + \frac{d}{2}
\]

\[
> \sup \left\{ \sum_{n=1}^{N} (Ty_n, \zeta_n) : T \in B_1(H) \cap \ker u \right\} = \|y\|_0,
\]

which contradicts the fact that \(\phi \in S_0^*\). We conclude that \(d = 0\) and therefore that \(D\) is \(\|\|_0\)-dense in \(S_0^*\).

Continuing our proof that \(\beta = 0\), pick \(T_0 \in B_1(H)\) with \(u(T_0) > 0\). Replacing \(u\) by \(u(T_0)^{-1} u\) if necessary, we may assume that \(u(T_0) = 1\). Define a linear functional \(\Psi\) on \((K, \|\|_0)\) by setting

\[
\Psi(x) = \beta \sum_{n=1}^{N} (T_0 x_n, \zeta_n) \quad (x \in K).
\]

Note that for \(x \in K\) we have

\[
|\Psi(x)| \leq \beta \sum_{n=1}^{N} \|x_n\| \|\zeta_n\| \leq \beta \|x\| \|\zeta\| \leq \|x\|_0.
\]

Hence \(\Psi \in S_0^*\). By the work of the previous paragraph, we can find \(T \in B_1(H) \cap \ker u\) such that \(\|\Psi - F(T)\|_0 < \beta/2\delta\). In particular, since \(\|\xi\|_0 \leq \|\xi\| = 1\),

\[
\left| \sum_{n=1}^{N} \langle (\beta T_0 - T) \xi_n, \zeta_n \rangle \right| < \frac{\beta}{2\delta}.
\]

In order to apply the defining property of \(\delta\) and thereby obtain a contradiction, we need to estimate not the sum on the left hand side of (2), but \(\sum_{n=1}^{N} |\langle (\beta T_0 - T) \xi_n, \zeta_n \rangle|\). To do so, we write

\[
\{n : 1 \leq n \leq N\} = P \cup Q,
\]

where \(P, Q\) are disjoint sets,

\[
n \in P \Rightarrow |\langle (\beta T_0 - T) \xi_n, \zeta_n \rangle| \neq 0, \text{ and}
\]

\[
n \in Q \Rightarrow |\langle (\beta T_0 - T) \xi_n, \zeta_n \rangle| < \frac{\beta}{2\delta N}.
\]

If \(n \in P\), we set

\[
\lambda_n = \frac{1}{|\langle (\beta T_0 - T) \xi_n, \zeta_n \rangle|}
\]

\(^8\)Alternatively, we can refer to [3] (page 287, Theorem 10) or [8] (Theorem 5.4.14).
and if \( n \in Q \), we set \( \lambda_n = 0 \); in each case, we define \( \gamma_n \equiv \lambda_n \xi_n \). Then \( \gamma \equiv (\gamma_1, \ldots, \gamma_N) \in K \) and

\[
\|\gamma\|_0^2 \leq \|\gamma\|^2 = \sum_{n=1}^{N} |\lambda_n|^2 \|\xi_n\|^2 \leq \|\xi\|^2 = 1.
\]

Hence

\[
\left| \sum_{n=1}^{N} \langle (\beta T_0 - T) \gamma_n, \zeta_n \rangle \right| \leq \|\Psi - F(T)\|_0^* < \frac{\beta}{2\delta}.
\]

Moreover,

\[
\left| \sum_{n=1}^{N} \langle (\beta T_0 - T) \gamma_n, \zeta_n \rangle \right| = \sum_{n \in P} \left| \langle (\beta T_0 - T) \lambda_n \xi_n, \zeta_n \rangle \right| = \sum_{n \in P} \left| \langle (\beta T_0 - T) \xi_n, \zeta_n \rangle \right|,
\]

so

\[
\sum_{n=1}^{N} \left| \langle (\beta T_0 - T) \xi_n, \zeta_n \rangle \right| = \sum_{n \in P} \left| \langle (\beta T_0 - T) \xi_n, \zeta_n \rangle \right| + \sum_{n \in Q} \left| \langle (\beta T_0 - T) \xi_n, \zeta_n \rangle \right| \leq \sum_{n=1}^{N} \left| \langle (\beta T_0 - T) \gamma_n, \zeta_n \rangle \right| + N \left( \frac{\beta}{2\delta N} \right) < \frac{\beta}{\delta}
\]

and therefore \( u(\beta T_0 - T) < \beta \). But \( u(\beta T_0 - T) = \beta u(T_0) - u(T) = \beta \), a contradiction which ensures that \( \beta \) actually equals 0. \( \blacksquare \)

We shall apply Lemma 6 shortly; but its application requires another construction.

**Lemma 7** Let \( N \) be a positive integer, let \( \xi_1, \ldots, \xi_N \) be linearly independent vectors in \( H \), and let \( \zeta_1, \ldots, \zeta_N \) be nonzero elements of \( H \), such that

\[
\sum_{n=1}^{N} \|\xi_n\|^2 = \sum_{n=1}^{N} \|\zeta_n\|^2 = 1.
\]

Then there exists a positive number \( c \) with the following property: for each unit vector \( z \) in the subspace

\[
K \equiv C\xi_1 \times \cdots \times C\xi_N,
\]

there exists \( T \in B_1(H) \) such that \( \sum_{n=1}^{N} \langle Tz_n, \zeta_n \rangle > c \).

**Proof.** Let

\[
m \equiv \inf \left\{ \|\zeta_n\|^2 : 1 \leq n \leq N \right\} > 0.
\]
Define a norm on the $N$-dimensional span $V$ of $\{\xi_1, \ldots, \xi_N\}$ by
\[
\left\| \sum_{n=1}^{N} \alpha_n \xi_n \right\|_1 \equiv \max_{1 \leq n \leq N} |\alpha_n|.
\]
Since $V$ is finite-dimensional, there exists $b > 0$ such that $\|x\|_1 \leq b \|x\|$ for each $x \in V$. Let $z \equiv (\lambda_1 \xi_1, \ldots, \lambda_N \xi_N)$ in $H_N$ satisfy $\|z\|_1 = 1$. If $|\lambda_n| < 1/\sqrt{N}$ for each $n$, then
\[
1 = \sum_{n=1}^{N} |\lambda_n|^2 = \sum_{n=1}^{N} |\alpha_n|^2 \|\xi_n\|^2 < \sum_{n=1}^{N} \left( \frac{1}{\sqrt{N}} \right)^2 = 1,
\]
which is absurd. Hence we can pick $\nu$ such that $|\lambda_\nu| > 1/\sqrt{2N}$. Define a linear mapping $T$ on $H$ such that
\[
T \xi_\nu = \frac{\lambda_\nu^*}{b|\lambda_\nu|} \zeta_\nu, \quad T \xi_n = 0 \ (n \neq \nu),
\]
and $Tx = 0$ whenever $x$ is orthogonal to $V$. Then
\[
\| T \left( \sum_{n=1}^{N} \alpha_n \xi_n \right) \| = \left| \frac{\lambda_\nu^*}{b|\lambda_\nu|} \alpha_\nu \right| \leq \frac{1}{b} \left\| \sum_{n=1}^{N} \alpha_n \xi_n \right\|_1 \leq \left\| \sum_{n=1}^{N} \alpha_n \xi_n \right\|,
\]
so $T \in B_1(H)$. Moreover,
\[
\langle Tz_n, \zeta_n \rangle = \begin{cases} 
0 & \text{if } n \neq \nu \\
\frac{1}{b} |\lambda_\nu| \|\zeta_\nu\|^2 & \text{if } n = \nu,
\end{cases}
\]
so
\[
\sum_{n=1}^{N} \langle Tz_n, \zeta_n \rangle = \frac{1}{b} |\lambda_\nu| \|\zeta_\nu\|^2 > \frac{m}{b \sqrt{2N}}.
\]
It remains to take $c \equiv m/b \sqrt{2N}$.

The next lemma takes the information arising from the preceding two, and shows that when the vectors $\zeta_n$ in (1) are linearly independent, we can approximate $u$ by a finite sum of mappings of the form $T \sim \langle Tx, y \rangle$, not just on its unit kernel but on the entire unit ball of $B(H)$. At the same time, we produce a priori bounds on the sums of squares of the norms of the components of the vectors $x, y$ that appear in the terms $\langle Tx, y \rangle$ whose sum approximates $u(T)$. Those bounds will be needed in the proof of our characterisation theorem.

**Lemma 8** Let $H$ be a Hilbert space, and $u$ a nonzero weak-operator continuous linear functional on $B(H)$. Let $\delta$ be a positive number, $\xi_1, \ldots, \xi_N$ linearly independent vectors in $H$, and $\zeta_1, \ldots, \zeta_N$ nonzero vectors in $H$, such that
\[
\sum_{n=1}^{N} \|\xi_n\|^2 = \sum_{n=1}^{N} \|\zeta_n\|^2 = 1
\]
and (1) holds. Let $c > 0$ be as in Lemma 7. Then for each $\varepsilon > 0$, there exists $x \in C_{\xi_1} \times \cdots \times C_{\xi_N}$ such that $x_n \neq 0$ for each $n$,

$$\|x\| < \frac{2\|u\|}{c},$$

and

$$\left| u(T) - \sum_{n=1}^{N} \langle Tx_n, \zeta_n \rangle \right| < \varepsilon$$

for all $T \in B_1(H)$.

**Proof.** Pick $T_0 \in B_1(H)$ with $u(T_0) > 0$. To begin with, take the case where $u(T_0) = 1$ and therefore $\|u\| \geq 1$. Given $\varepsilon > 0$, set

$$\alpha \equiv \min \{\varepsilon, 1\} \frac{2}{\|u\|(1 + \|u\|)}.$$ 

Applying Lemma 6, we obtain nonzero vectors $z_n \in C_{\xi_n}$ $(1 \leq n \leq N)$ such that

$$\left| \sum_{n=1}^{N} \langle T z_n, \zeta_n \rangle \right| < c \alpha \quad (T \in B_1(H) \cap \ker u).$$

For each $T \in B_1(H)$, since

$$(1 + \|u\|)^{-1} (T - u(T)T_0) \in B_1(H) \cap \ker u,$$

we have

$$\left| \sum_{n=1}^{N} \langle (T - u(T)T_0) z_n, \zeta_n \rangle \right| < (1 + \|u\|) \alpha.$$ 

By Lemma 7, there exists $T_1 \in B_1(H)$ such that $\sum_{n=1}^{N} \langle T_1 z_n, \zeta_n \rangle > c$. We compute

$$c < \sum_{n=1}^{N} \langle T_1 z_n, \zeta_n \rangle$$

$$\leq \sum_{n=1}^{N} \langle (T_1 - u(T_1)T_0) z_n, \zeta_n \rangle + |u(T_1)| \sum_{n=1}^{N} \langle T_0 z_n, \zeta_n \rangle$$

$$\leq (1 + \|u\|) c \alpha + \|u\| \left| \sum_{n=1}^{N} \langle T_0 z_n, \zeta_n \rangle \right|.$$ 

Hence

$$\left| \sum_{n=1}^{N} \langle T_0 z_n, \zeta_n \rangle \right| > \frac{c}{\|u\|} \left( 1 - (1 + \|u\|) \alpha \right)$$

$$\geq \frac{c}{\|u\|} \left( 1 - \frac{1}{2\|u\|} \right) > \frac{c}{2\|u\|},$$

**Documenta Mathematica 16 (2011) 597–617**
since \( \|u\| \geq 1 \). Setting
\[
x \equiv \left( \sum_{n=1}^{N} \langle T_0 z_n, \zeta_n \rangle \right)^{-1} z,
\]
we have \( 0 \neq x_n \in C_\xi_n \) for each \( n \), and
\[
\|x\| = \left| \sum_{n=1}^{N} \langle T_0 z_n, \zeta_n \rangle \right|^{-1} \|z\| < \frac{2 \|u\|}{c}.
\]

Moreover, for each \( T \in B_1(H) \),
\[
\left| u(T) - \sum_{n=1}^{N} \langle Tx_n, \zeta_n \rangle \right| = \left| \sum_{n=1}^{N} \langle T_0 z_n, \zeta_n \rangle \right|^{-1} \left| \sum_{n=1}^{N} \langle (u(T)T_0 - T) z_n, \zeta_n \rangle \right|
\leq \frac{2 \|u\|}{c} (1 + \|u\|) \alpha \leq \varepsilon.
\]

We now remove the restriction that \( u(T_0) = 1 \). Applying the first part of the theorem to \( v \equiv u(T_0)^{-1}u \), we construct \( y \in K \) such that each component \( y_n \neq 0 \), \( \|y\| \leq 2 \|v\| /c \), and
\[
\left| v(T) - \sum_{n=1}^{N} \langle Ty_n, \zeta_n \rangle \right| < u(T_0)^{-1} \varepsilon,
\]
and we obtain the desired conclusion by taking \( x \equiv u(T_0)y \).

**Lemma 9** Under the hypotheses of Lemma 8, but without the assumption that \( u \) is nonzero, for all \( \varepsilon, \varepsilon' > 0 \), there exists \( x \in C_\xi_1 \times \cdots \times C_\xi_N \) such that \( x_n \neq 0 \) for each \( n \),
\[
\|x\| < \frac{2 \left( \|u\| + \varepsilon' \right)}{c},
\]
and
\[
\left| u(T) - \sum_{n=1}^{N} \langle Tx_n, \zeta_n \rangle \right| < \varepsilon
\]
for all \( T \in B_1(H) \).

**Proof.** Either \( \|u\| > 0 \) and we can apply Lemma 8, or else \( \|u\| < \varepsilon/2 \). In the latter event, pick \( x \) in \( C_\xi_1 \times \cdots \times C_\xi_N \) such that \( x_n \neq 0 \) for each \( n \) and
\[
\|x\| < \min \left\{ \frac{\varepsilon}{2}, \frac{2 \left( \|u\| + \varepsilon' \right)}{c} \right\}.
\]
Then for each \( T \in B_1(H) \) we have
\[
\left| \sum_{n=1}^{N} \langle Tx_n, \zeta_n \rangle \right| \leq \sum_{n=1}^{N} \|x_n\| \|\zeta_n\| \leq \|x\| \|\zeta\| < \frac{\varepsilon}{2}.
\]
and therefore
\[ |u(T) - \sum_{n=1}^{N} \langle Tx_n, \zeta_n \rangle| \leq \|u\| + \sum_{n=1}^{N} |\langle Tx_n, \zeta_n \rangle| < \varepsilon. \]

### 3 The Characterisation Theorem

We are finally able to prove our main result, by inductively applying Lemma 9.

**Theorem 10** Let \( H \) be a nontrivial Hilbert space, and \( u \) a nonzero weak-operator continuous linear functional on \( \mathcal{B}(H) \). Let \( \delta \) be a positive number, \( \xi_1, \ldots, \xi_N \) linearly independent vectors in \( H \), and \( \zeta_1, \ldots, \zeta_N \) nonzero vectors in \( H \), such that \( |u(T)| \leq \delta \sum_{n=1}^{N} |\langle T\xi_n, \zeta_n \rangle| \) for all \( T \in \mathcal{B}(H) \). Then there exists \( x \in C\xi_1 \times \cdots \times C\xi_N \) such that
\[ u(T) = \sum_{n=1}^{N} \langle Tx_n, \zeta_n \rangle \]  
for all \( T \in \mathcal{B}(H) \).

**Proof.** Re-scaling if necessary, we may assume that \( \|u\| < 2^{-3} \). In the notation of, and using, Lemma 9, compute \( x^{(1)} \in K \equiv C\xi_1 \times \cdots \times C\xi_N \) such that
\[ \|x^{(1)}\| \leq \frac{2}{c} (\|u\| + 2^{-3}) < \frac{1}{2c} \]  
and
\[ |u(T) - \sum_{n=1}^{N} \langle Tx_n^{(1)}, \zeta_n \rangle| < 2^{-4} \quad (T \in \mathcal{B}_1(H)). \]

Suppose that for some positive integer \( k \) we have constructed vectors \( x^{(i)} \in K \) \((1 \leq i \leq k)\) such that
\[ \|x^{(k)}\| < \frac{1}{2^kc}. \]  

---

9The requirement that the vectors \( \xi_n \) be linearly independent is the one place where we have a stronger hypothesis than is needed in the classical theorem. It is worth noting here that if \( u(T) \) has the desired form \( \sum_{n=1}^{N} \langle T\xi_n, \zeta_n \rangle \), then classically we can find a set \( F \) of indices \( n \leq N \) such that (i) the set \( S \) of those \( \xi_n \) with \( n \in F \) is linearly independent and (ii) if \( \xi_k \notin S \), then \( \xi_k \) is linearly dependent on \( S \).

10In this proof we do not need the fact that, according to Lemma 9, we can arrange for the components of the vector \( x^{(1)} \), and of the subsequently constructed vectors \( x^{(k)} \), to be nonzero.
and

$$\left| u(T) - \sum_{n=1}^{N} \left\langle T \left( x_{n}^{(1)} + \cdots + x_{n}^{(k)} \right), \zeta_{n} \right\rangle \right| < 2^{-k-3} \quad (T \in B_{1}(H)) \quad (5)$$

Consider the weak-operator continuous linear functional

$$v : T \mapsto u(T) - \sum_{n=1}^{N} \left\langle T \left( x_{n}^{(1)} + \cdots + x_{n}^{(k)} \right), \zeta_{n} \right\rangle$$

on $B(H)$. Writing

$$x_{n}^{(1)} + \cdots + x_{n}^{(k)} = \lambda_{n} \xi_{n}$$

and

$$\gamma \equiv \max \{|\lambda_{1}|, \ldots, |\lambda_{n}|\},$$

for each $T \in B(H)$ we have

$$|v(T)| \leq |u(T)| + \sum_{n=1}^{N} \left| \left\langle T \left( x_{n}^{(1)} + \cdots + x_{n}^{(k)} \right), \zeta_{n} \right\rangle \right|$$

$$\leq \delta \sum_{n=1}^{N} |\langle T \xi_{n}, \zeta_{n} \rangle| + \sum_{n=1}^{N} |\lambda_{n}| |\langle T \xi_{n}, \zeta_{n} \rangle|$$

$$\leq (\delta + \gamma) \sum_{n=1}^{N} |\langle T \xi_{n}, \zeta_{n} \rangle|.$$

We can now apply Lemma 9, to obtain

$$x^{(k+1)} = \left( x_{1}^{(k+1)}, \ldots, x_{N}^{(k+1)} \right) \in K$$

such that

$$\left\| x^{(k+1)} \right\| < \frac{2}{c} (\| \nu \| + 2^{-k-3}) < \frac{1}{2^{k+1}c}$$

and

$$\left| u(T) - \sum_{n=1}^{N} \left\langle T \left( x_{n}^{(1)} + \cdots + x_{n}^{(k)} + x_{n}^{(k+1)} \right), \zeta_{n} \right\rangle \right|$$

$$= \left| v(T) - \sum_{n=1}^{N} \left\langle T x_{n}^{(k+1)}, \zeta_{n} \right\rangle \right| < 2^{-k-4}$$

for all $T \in B_{1}(H)$. This completes the inductive construction of a sequence $(x^{(k)})_{k \geq 1}$ in $K$ such that (4) and (5) hold for each $k$. The series $\sum_{k=1}^{\infty} x^{(k)}$ converges to a sum $x$ in the finite-dimensional Banach space $K$, by comparison with $\sum_{k=1}^{\infty} 2^{-k} c^{-1}$. Letting $k \to \infty$ in (5), we obtain (3) for all $T \in B_{1}(H)$ and hence for all $T \in B(H)$.
For nonzero $u$, the proof of our theorem can be simplified at each stage of the induction, since we can use Lemma 8 directly. If $H$ has dimension $> N$, we can then construct the classical representation of $u$ in the general case as follows. Either $\|u\| > 0$ and there is nothing to prove, or else $\|u\| < \delta$ (the same $\delta$ as in the statement of the theorem). In the latter case, we construct a unit vector $\xi_{N+1}$ orthogonal to each of the vectors $\xi_n$ ($1 \leq n \leq N$), set $\zeta_{N+1} = \xi_{N+1}$, and consider the weak-operator continuous linear functional

$$v : T \mapsto u(T) + \delta \langle T\xi_{N+1}, \zeta_{N+1} \rangle.$$ 

We have

$$|v(T)| \leq |u(T)| + \delta |\langle T\xi_{N+1}, \zeta_{N+1} \rangle| \leq \delta \sum_{n=1}^{N+1} |\langle T\xi_{n}, \zeta_{n} \rangle|.$$ 

Moreover,

$$|v(I)| \geq \delta \|\xi_{N+1}\|^2 - |u(I)| \geq \delta - \|u\| > 0,$$

where $I$ is the identity operator on $H$; so $v$ is nonzero. We can therefore apply the nonzero case to $v$, to produce a vector $y \in C\xi_1 \times \cdots \times C\xi_{N+1}$ such that

$$v(T) = \sum_{n=1}^{N+1} \langle Ty_n, \zeta_n \rangle \quad (T \in B(H)).$$ 

Setting $x_n = y_n$ ($1 \leq n \leq N$) and $x_{n+1} = y_{N+1} - \delta\xi_{N+1}$, we obtain

$$u(T) = \sum_{n=1}^{N+1} \langle Tx_n, \zeta_n \rangle$$

for each $T \in B(H)$. Note, however, that this proof gives $x$ in $C\xi_1 \times \cdots \times C\xi_{N} \times C\xi_{N+1}$, not, as in Theorem 10, in $C\xi_1 \times \cdots \times C\xi_{N}$.

As an immediate consequence of Theorem 10, the functional $u$ therein is a linear combination of the functionals $T \mapsto \langle T\xi_{n}, \zeta_{n} \rangle$ associated with the seminorms that describe the boundedness of $u$.

**Corollary 11** Under the hypotheses of Theorem 10, there exist complex numbers $\alpha_1, \ldots, \alpha_N$ such that

$$u(T) = \sum_{n=1}^{N} \alpha_n \langle T\xi_{n}, \zeta_{n} \rangle$$

for each $T \in B(H)$. 

*Documenta Mathematica* 16 (2011) 597–617
4 Strong-operator Continuous Functionals

Next we turn briefly to the strong operator topology on $\mathcal{B}(H)$: the locally convex topology generated by the seminorms $T \mapsto \|Tx\|$ with $x \in H$. (That is, the weakest topology with respect to which the mapping $T \mapsto Tx$ is continuous for each $x \in H$.) Clearly, a weak-operator continuous linear functional on $\mathcal{B}(H)$ is strong-operator continuous. The converse holds classically, but, as we now show by a Brouwerian example, is essentially nonconstructive.

Let $(e_n)_{n \geq 1}$ be an orthonormal basis of unit vectors in an infinite-dimensional Hilbert space, and let $(a_n)_{n \geq 1}$ be a binary sequence with at most one term equal to 1. Then for $k \geq j$ we have

$$
\sum_{n=j}^{k} |a_n \langle Te_1, e_n \rangle| \leq \left( \sum_{n=j}^{k} a_n^2 \right)^{1/2} \left( \sum_{n=j}^{k} |\langle Te_1, e_n \rangle|^2 \right)^{1/2} \\
\leq \left( \sum_{n=j}^{k} |\langle Te_1, e_n \rangle|^2 \right)^{1/2}.
$$

Since $\sum_{n=1}^{\infty} |\langle Te_1, e_n \rangle|^2$ converges to $\|Te_1\|$, we see that $\sum_{n=j}^{k} |a_n \langle Te_1, e_n \rangle| \to 0$ as $j, k \to \infty$. Hence

$$u(T) \equiv \sum_{n=1}^{\infty} a_n \langle Te_1, e_n \rangle$$

defines a linear functional $u$ on $\mathcal{B}(H)$; moreover, $|u(T)| \leq \|Te_1\|$ for each $T$, so (by Proposition 5.4.1 of [8]) $u$ is strong-operator continuous. Suppose it is also weak-operator continuous. Then, by Lemma 2, it is normed. Either $\|u\| < 1$ or $\|u\| > 0$. In the first case, if there exists (a unique) $\nu$ with $a_\nu = 1$, then $u(T) = \langle Te_1, e_\nu \rangle$ for each $T \in \mathcal{B}(H)$. Defining $T$ such that $Te_1 = e_\nu$ and $Te_n = 0$ for all $n \neq \nu$, we see that $T \in \mathcal{B}_1(H)$ and $u(T) = 1$; whence $\|u\| = 1$, a contradiction. Thus in this case, $a_n = 0$ for all $n$. On the other hand, in the case $\|u\| > 0$ we can find $T$ such that $u(T) > 0$; whence there exists $n$ such that $a_n = 1$. It now follows that the statement

If $H$ is an infinite-dimensional Hilbert space, then every strong-operator continuous linear functional on $\mathcal{B}(H)$ is weak-operator continuous.

implies the essentially nonconstructive principle

**LPO:** For each binary sequence $(a_n)_{n \geq 1}$, either $a_n = 0$ for all $n$ or else there exists $n$ such that $a_n = 1$

and so is itself essentially nonconstructive.
5 Concluding Observations

The ideal constructive form of Theorem 10 would have two improvements over the current one. First, the requirement that the vectors \(\xi_n\) be linearly independent would be relaxed to have them only nonzero in Lemma 8, Lemma 9, and Theorem 10. Second, \(B(H)\) would be replaced by a suitable linear subspace \(R\) of itself, and our theorem would apply to linear functionals that are weak-operator continuous on \(R\), where “suitable” probably means “having weak-operator totally bounded unit ball \(R_1 \equiv R \cap B_1(H)\)”. With that notion of suitability and with minor adaptations, Lemma 6 holds and the proof of Lemma 8 goes through as far as the construction of the vector \(z \in K\). In fact, Theorem 10 goes through with \(B(H)\) replaced by any linear subspace \(R\) of \(B(H)\) that has weak-operator totally bounded unit ball and satisfies the following condition (cf. Lemma 7):

\[
(*) \quad \text{Let } N \text{ be a positive integer, let } \xi_1, \ldots, \xi_N \text{ be linearly independent vectors in } H, \text{ and let } \zeta_1, \ldots, \zeta_N \text{ be nonzero elements of } H, \text{ such that } \sum_{n=1}^{N} \|\xi_n\|^2 = \sum_{n=1}^{N} \|\zeta_n\|^2 = 1. \text{ Then there exists a positive number } c \text{ with the following property: for each unit vector } z \text{ in the subspace } \\
K \equiv C\xi_1 \times \cdots \times C\xi_N \\
\text{there exists } T \in R_1 \text{ such that } \sum_{n=1}^{N} \langle Tz_n, \zeta_n \rangle > c.
\]

This condition holds in the special case where \(N = 1\), in which case, if also \(R_1\) is weak-operator totally bounded, we obtain Theorem 1 of [6]. However, there seems to be no means of establishing \((*)\) for \(N > 1\) and a general \(R\). So the ideal form of our theorem remains an ideal and a challenge.

Acknowledgements. This work was supported by a Marie Curie IRSES award from the European Union for the project Construmath, with counterpart funding from the New Zealand Ministry of Research, Science & Technology. It was written while the author enjoyed the hospitality of (i) Professor Helmut Schwichtenberg and his colleagues in the Logic Group at the Mathematisches Institut der Ludwig-Maximilians-Universität München, and (ii) Professor Erik Palmgren in the Matematiska Institution, University of Uppsala. The author also thanks the referee for providing helpful comments that improved the presentation of the paper.

\textsuperscript{11}But the proof of the theorem in [6] is simpler and more direct than the case \(N = 1\) of the proof of our Theorem 10 above.
References

[1] P. Aczel and M. Rathjen, *Notes on Constructive Set Theory*, Report No. 40, Institut Mittag-Leffler, Royal Swedish Academy of Sciences, 2001.

[2] P. Aczel and M. Rathjen, *Constructive Set Theory*, monograph, forthcoming.

[3] E.A. Bishop, *Foundations of Constructive Analysis*, McGraw-Hill, New York, 1967.

[4] E.A. Bishop and D.S. Bridges, *Constructive Analysis*, Grundlehren der Math. Wiss. 279, Springer Verlag, Heidelberg, 1985.

[5] D.S. Bridges, ‘On the weak operator compactness of the unit ball of $L(H)$’, Zeit. math. Logik Grundlagen Math. 24, 493–494, 1978.

[6] D.S. Bridges, ‘Characterising dominated weak-operator continuous functionals constructively’, Ann. Pure Appl. Logic, to appear.

[7] D.S. Bridges and L.S. Dediu (Viță), ‘Weak operator continuity and the existence of adjoints’, Math. Logic Quarterly 45, 203–206, 1999.

[8] D.S. Bridges and L.S. Viță, *Techniques of Constructive Analysis*, Universitext, Springer Verlag, Heidelberg, 2006.

[9] D.S. Bridges and N.F. Dudley Ward, ‘Constructing ultraweakly continuous linear functionals on $B(H)$’, Proc. American Math. Soc. 126(11), 3347–3353, 1998.

[10] J. Dixmier, *Les Algèbres d’Opérateurs dans l’Espace Hilbertien: Algèbres de von Neumann*, Gauthier-Villars, Paris, 1957.

[11] R.V. Kadison and J.R. Ringrose, *Fundamentals of the Theory of Operator Algebras* (Vols. 1,2), Academic Press, New York, 1988.

[12] P. Martin-Löf, *Intuitionistic Type Theory*, Bibliopolis, Naples, 1984.

[13] R.Mines, F. Richman, and W. Ruitenburg, *A Course in Constructive Algebra*, Universitext, Springer Verlag, New York, 1988.

[14] S. Sakai, *$C^*$ Algebras and $W^*$ Algebras*, Springer Verlag, Heidelberg, 1971.

[15] B. Spitters, ‘Constructive results on operator algebras’, J. Univ.Comp. Sci. 11(12), 2096–2113, 2005.

[16] M. Takesaki, *Theory of Operator Algebra 1*, Encyclopaedia of Math. Sciences, Springer Verlag, Berlin, 2002.
Characterising Weak-Operator ...
THE CLASSIFICATION OF REAL
PURELY INFINITE SIMPLE C*-ALGEBRAS

JEFFREY L. BOERSEMA, EFREN RUIZ, P. J. STACEY

Received: February 26, 2011
Communicated by Joachim Cuntz

Abstract. We classify real Kirchberg algebras using united $K$-theory. Precisely, let $A$ and $B$ be real simple separable nuclear purely infinite $C^*$-algebras that satisfy the universal coefficient theorem such that $A_{\mathbb{C}}$ and $B_{\mathbb{C}}$ are also simple. In the stable case, $A$ and $B$ are isomorphic if and only if $K^{\text{CRT}}(A) \cong K^{\text{CRT}}(B)$. In the unital case, $A$ and $B$ are isomorphic if and only if $(K^{\text{CRT}}(A), [1_A]) \cong (K^{\text{CRT}}(B), [1_B])$. We also prove that the complexification of such a real $C^*$-algebra is purely infinite, resolving a question left open from [43]. Thus the real $C^*$-algebras classified here are exactly those real $C^*$-algebras whose complexification falls under the classification result of Kirchberg [26] and Phillips [35]. As an application, we find all real forms of the complex Cuntz algebras $\mathcal{O}_n$ for $2 \leq n \leq \infty$.

2010 Mathematics Subject Classification: 46L35, 46L80, 19K99
Keywords and Phrases: Real $C^*$-algebras, $K$-theory, classification

1. Introduction

One of the highlights of the classification theory of simple amenable $C^*$-algebras is the classification of purely infinite nuclear simple $C^*$-algebras, obtained by Kirchberg and Phillips in [26] and [35]. This classification theorem relies in an essential way on the Universal Coefficient Theorem established by Rosenberg and Schochet in [40], where it was observed that “For reasons pointed out already by Atiyah, there can be no good Künneth Theorem or Universal Coefficient Theorem for the $KKO$ groups of real $C^*$-algebras; this explains why we deal only with complex $C^*$-algebras”. Thus at the time of the Kirchberg and Phillips classification, the lack of a universal coefficient theorem was the primary barrier to extending the classification result to real $C^*$-algebras.

However, in [8], a new invariant called united $K$-theory was introduced for real $C^*$-algebras and in [9] a universal coefficient theorem was proven for real
C*-algebras using united \( K \)-theory. In the present paper, we take advantage of these developments to provide a complete classification of a class of real simple purely infinite C*-algebras in terms of united \( K \)-theory. The real C*-algebras that are classified are exactly those real C*-algebras for which the complexification is covered by the Kirchberg and Phillips theory. As an application of our classification we determine all the real forms of the complex Cuntz algebras \( \mathcal{O}_n \) for \( 1 \leq n \leq \infty \): there are two such forms when \( n \) is odd and one when \( n \) is even or infinite.

The overall framework of the proof will be the same as that in the paper [35] and the underlying theory on which that paper was built. Furthermore, many of the proofs in the development leading to the main theorems of [35] carry over to the real case without significant change. In those cases, we will simply refer to the established proofs in the literature without reproducing them here. However there are many situations where the arguments in the real case require modification and we will then provide full proofs or full discussion of the necessary modifications.

In Section 2, we describe the invariant of united \( K \)-theory and summarize its key properties. In Section 3 we then establish real analogues of some of the fundamental properties of purely infinite algebras, in the course of which we resolve a problem left hanging in [43] and [13] by showing that the complexification of a purely infinite simple real C*-algebra is also purely infinite (using the original definition for simple algebras). Following the complex case, as developed in [38], we then establish (in Theorem 5.2) criteria for two unital homomorphisms from the real Cuntz algebra \( \mathcal{O}_R^R \) (\( n \) even) to be approximately unitarily equivalent. Modifications of the complex arguments are required to establish some of the preliminary results: in Section 4 we modify the required results about exponential rank, noting that the close link between self-adjoint and skew-adjoint elements is absent in a real C*-algebra, and in Section 5 we modify the result from [15] establishing the Rokhlin property of the Bernoulli shift on the CAR-algebra.

Our next step is to establish real analogues of Kirchberg’s tensor product theorems and his embedding theorem. This is achieved in Section 6 by using the relevant complex results and the embedding of \( \mathbb{C} \) into \( M_2(R) \). In Sections 7, 8, 9 and 10, we closely follow [35] indicating how the results achieved for the complex case can be obtained in the real case. In particular, Section 7 contains a key result about uniqueness of homomorphisms from \( \mathcal{O}_R^\infty \) to a real purely infinite C*-algebras. Section 8 contains the theory of asymptotic morphisms in the context of real C*-algebras and Section 9 culminates in a theorem identifying \( KK \)-theory to a group of asymptotic unitary equivalence classes of asymptotic morphisms as in Section 4 of [35]. To accomplish this, we make use of the axiomatic characterization of \( KK \)-theory for real C*-algebras established in [12]. This development culminates in Section 10, which contains the statements and proofs of our classification theorems, and in Section 11, which uses these results to describe the real forms of Cuntz algebras. The notation we use in these sections closely follows that in [35].
The Classification of Real . . .

We will use the notation $H^R$ for a real Hilbert space; and $B(H^R)$ and $\mathcal{K}^R$ for the real C*-algebras of bounded and compact operators $H^R$. For the complex versions of these objects we will use $H$, $B(H)$, and $\mathcal{K}$. For a C*-algebra $A$, we will write $M_n(A)$ for the matrix algebra over $A$, and $M_n(\mathbb{R})$ will stand for $M_n(\mathbb{R})$. Following standard convention, we will use $O_n$ for the complex Cuntz algebras and $O_n^R$ for the real versions. The complexification of a real C*-algebra $A$ will be denoted by $A_C$. We will use $\Phi$ throughout to denote the conjugate linear automorphism of $A_C$ defined by $a + ib \mapsto a - ib$ (for $a, b \in A$). Note that $A$ can be recovered from $\Phi$ as the fixed point set. Finally, a tensor product written as $A \otimes B$ will in most cases be the C*-algebra tensor product over $\mathbb{R}$, but should be understood to be a tensor product over $\mathbb{C}$ if both $A$ and $B$ are known to be complex C*-algebras. Recall that if $A$ and $B$ are real C*-algebras, then $(A \otimes B)_C \cong A_C \otimes B_C$.

2. Preliminaries on United $K$-Theory

United $K$-theory was developed in the commutative context in [14] and subsequently extended to the context of real C*-algebras in [8]. United $K$-theory consists of the three separate $K$-theory modules as well as several natural transformations among them. In this section, we give the definition of united $K$-theory and summarize the features needed in this paper. Details are in [8], [9], [10].

**Definition 2.1.** Let $A$ be a real C*-algebra. The **united $K$-theory** of $A$ is given by

$$K^{\text{CRT}}(A) = \{KO_s(A), KU_s(A), KT_s(A), r,c,\varepsilon, \zeta, \psi, \tau, \gamma, \epsilon, \zeta, \psi, \tau \}.$$  

In this definition, $KO_s(A) = K_s(A)$ is the standard $K$-theory of a real C*-algebra, considered as a graded module over the ring $K_s(\mathbb{R})$. This means in particular that there are operations

$$\eta_0: KO_n(A) \to KO_{n+1}(A)$$  

$$\xi: KO_n(A) \to KO_{n+4}(A)$$  

$$\beta_0: KO_n(A) \to KO_{n+8}(A)$$

corresponding to multiplication by the elements of the same name in $KO_s(\mathbb{R})$. The operation $\beta_0$ is the periodicity isomorphism of real $K$-theory.

The second item $KU_s(A) = K_s(AC)$ is the $K$-theory of the complexification of $A$, having period 2. It is a module over $K_s(\mathbb{C})$, which is to say that there is an isomorphism of period 2 and the two remaining groups are independent with no operations between them.

Finally, $KT_s(A)$ is the period 4 self-conjugate $K$-theory originally defined in the topological setting in [1]. In the non-commutative setting, it is more easily defined as $KT_s(A) = K_s(T \otimes A)$ in terms of the algebra $T = \{ f \in C([0,1], \mathbb{C}) |$
for a real C*-algebra to satisfy the following relations (see Proposition 1.7 of [8]):

\[ \eta_0 : KT_n(A) \rightarrow KT_{n+1}(A) \]
\[ \omega : KT_n(A) \rightarrow KT_{n+1}(A) \]
\[ \beta_r : KT_n(A) \rightarrow KT_{n+4}(A) \]

The important advantage of the full united K-theory is that it yields both a K"unneth formula (Theorem 4.2 of [8]). Self-conjugate \( K\)-theory consists of operations \( \psi, \eta, \xi \) that satisfy the above relations.

\[ rc = 2 \]
\[ cr = 1 + \psi \]
\[ r = r \gamma \]
\[ c = c \zeta \]
\[ (\psi \gamma)^2 = 1 \]
\[ (\psi r)^2 = 1 \]

\[ \psi \eta = \zeta \gamma \]
\[ \zeta \psi = \xi \]

and natural transformations \( \zeta \beta = \beta \zeta \)

\[ \xi \gamma = 0 \]
\[ \xi \tau = 1 \]

United K-theory takes values in the algebraic category of CRT-modules. A CRT-module is acyclic, which means that the sequences

\[ \cdots \rightarrow KO_0(A) \xrightarrow{\eta_0} KO_{n+1}(A) \xrightarrow{\psi} KO_n(A) \xrightarrow{r \beta_{2n-1}^1} KO_{n-1}(A) \rightarrow \cdots \]

\[ \cdots \rightarrow KO_0(A) \xrightarrow{\eta_2} KO_{n+2}(A) \xrightarrow{\psi} KT_{n+1}(A) \xrightarrow{r \beta_{2n-1}^1} KO_{n-1}(A) \rightarrow \cdots \]

\[ \cdots \rightarrow KO_{n+1}(A) \xrightarrow{\gamma} KT_n(A) \xrightarrow{\psi} KO_n(A) \xrightarrow{1- \psi} KO_{n+1}(A) \rightarrow \cdots \]

are exact.

The important advantage of the full united K-theory over ordinary K-theory for a real C*-algebra \( A \) is that it yields both a K"unneth formula (Theorem 4.2 of...
Proposition 2.2. For any real C*-algebra $A$,

1. $K^{CRT}(O^\mathbb{R}_2 \otimes A) = 0$
2. $K^{CRT}(O^\mathbb{R}_\infty \otimes A) \cong K^{CRT}(A)$. 

Proof. By Table IV of [8], we have $K^{CRT}(O^\mathbb{R}_2) = 0$. Then (1) follows by the Künneth formula.

The unital inclusion $\mathbb{R} \to O^\mathbb{R}_\infty$ induces an isomorphism on united $K$-theory. This follows from Theorem 4 of [10] and the fact that the unital inclusion $\mathbb{C} \to O^\mathbb{R}_\infty$ induces an isomorphism on (complex) $K$-theory. Thus, Theorem 3.5 of [8] gives $K^{CRT}(O^\mathbb{R}_\infty) \otimes_{CRT} K^{CRT}(A) \cong K^{CRT}(A)$ and $\text{Tor}(K^{CRT}(O^\mathbb{R}_\infty), K^{CRT}(A)) = 0$. Then the isomorphism of (2) follows by the Main Theorem of [8]. □

Recall from [41] that the bootstrap class $\mathcal{N}$ is the smallest subcategory of complex, separable, nuclear C*-algebras that contains the separable type I C*-algebras; that is closed under the operations of taking inductive limits, stable isomorphisms, and crossed products by $\mathbb{Z}$ and $\mathbb{R}$; and that satisfies the two out of three rule for short exact sequences (i.e. if $0 \to A \to B \to C \to 0$ is exact and two of $A$, $B$, $C$ are in $\mathcal{N}$, then the third is also in $\mathcal{N}$).

Proposition 2.3 (Corollary 4.11 of [9]). Let $A$ and $B$ be real separable C*-algebras such that $A_C$ and $B_C$ are in $\mathcal{N}$. Then $A$ and $B$ are $KK$-equivalent if and only if $K^{CRT}(A) \cong K^{CRT}(B)$.

This last result is the essential preliminary result for our classification of real purely infinite simple C*-algebras. We will also make use of Theorem 1 of [10], which states that every countable acyclic CRT-module can be realized as the united $K$-theory a real separable C*-algebra, indeed the C*-algebra can even be taken to be simple and purely infinite.

We now describe a simpler variation of united $K$-theory that, by results from [23], contains as much information as the full version of united $K$-theory.

Definition 2.4. Let $A$ be a real C*-algebra. Then

$$K^{CR}(A) = \{KO_*(A), KU_*(A), r, c, \psi_U\}$$

For any real C*-algebra, $K^{CR}(A)$ is an acyclic CR-module, which means that the relations

$$rc = 2 \hspace{1cm} \psi_U\beta_U = -\beta_U\psi_U \hspace{1cm} \xi = r\beta_U^2 c$$
$$cr = 1 + \psi_U \hspace{1cm} \psi_U^2 = 1 \hspace{1cm} \psi_U c = c$$

are satisfied and that the sequence

$$\cdots \longrightarrow KO_n(A) \xrightarrow{\delta^0} KO_{n+1}(A) \xrightarrow{\delta} KU_{n+1}(A) \xrightarrow{r\beta_U^{-1}} KO_{n-1}(A) \longrightarrow \cdots$$

is exact.
Let $\Gamma$ be the forgetful functor from the category $CRT$-modules to the category of $CR$-modules. It is immediate from Theorem 4.2.1 of [23] that $\Gamma$ is injective (but not surjective) on the class of acyclic $CRT$-modules. Hence we have the following result.

**Proposition 2.5.** Let $A$ and $B$ be real $C^*$-algebras. Then $K^{CRT}(A) \cong K^{CRT}(B)$ if and only if $K^{CR}(A) \cong K^{CR}(B)$.

Note, however, that the results of [10] do not extend to $CR$-modules. Not every countable acyclic $CR$-module can be realized as $K^{CR}(A)$ for a real $C^*$-algebra $A$.

### 3. Preliminaries on Real Simple Purely Infinite $C^*$-Algebras

In this section, we provide some preliminaries on simple and purely infinite $C^*$-algebras, including a theorem characterizing simple purely infinite real $C^*$-algebras in terms of their complexification. One direction of this characterization was achieved in [43] and [13].

Let $A$ be a real unital $C^*$-algebra, let $\mathcal{U}(A)$ denote the group of unitary elements in $A$, and let $\mathcal{U}_0(A)$ denote the connected component of the identity in $\mathcal{U}(A)$. Note that if $u$ is a unitary in a real $C^*$-algebra, then the spectrum $\sigma(u) \subseteq \mathbb{T}$ satisfies $\sigma(u) = \sigma(u)$ and the real $C^*$-algebra generated by $u$ is isomorphic to the algebra of complex-valued continuous functions $f$ on $\sigma(u)$ that satisfy $f(z) = f(\overline{z})$. (If $a$ is an element of $A$, then by definition the spectrum $\sigma(a)$ is found by passing to $A_\mathbb{C}$.)

We begin by making an explicit mention of a fairly well-known result about real simple $C^*$-algebras.

**Definition 3.1.** A real $C^*$-algebra $A$ is $c$-simple if $A_\mathbb{C}$ is simple.

**Proposition 3.2.** A simple real $C^*$-algebra $A$ is either $c$-simple or is isomorphic to a simple complex $C^*$-algebra.

**Proof.** Let $I$ be a proper ideal in $A_\mathbb{C}$. Then $J = A \cap I \cap \Phi(I) = 0$ and so $I \cap \Phi(I) = 0$. Furthermore, $I + \Phi(I) = A_\mathbb{C}$. It then follows that the map $x \mapsto x + \Phi(x)$ is an isomorphism from the complex $C^*$-algebra $I$ onto $A$.

As the structure of simple complex $C^*$-algebras is comparatively well-understood, our primary interest lies in $c$-simple $C^*$-algebras.

As in the complex case, we will use the tilde $\sim$ to denote the relation of Murray-von Neumann equivalence of projections. A projection is said to be infinite if it is Murray-von Neumann equivalent to a proper subprojection of itself. The following definition of purely infinite is from [43]. Bearing in mind subsequent developments, such as [27] and [28], a different definition should be made in the non-simple case. However the focus in this paper is on simple algebras, for which the definition below is appropriate.

**Definition 3.3.** Let $A$ be a real simple $C^*$-algebra.
(1) A subalgebra $B$ is a regular hereditary subalgebra of $A$ if there is an element $x \in A_+$ such that $B = xA x$.

(2) $A$ is purely infinite if every regular hereditary subalgebra of $A$ contains an infinite projection.

Proposition 3.4. Let $A$ be a separable simple purely infinite real $C^*$-algebra. Then either $A$ is unital or there is a real unital simple purely infinite $C^*$-algebra $A_0$ such that $A \cong K^\mathbb{R} \otimes A_0$.

Proof. As in Section 27.5 of [2]. \hfill \Box

Proposition 3.5. Let $A$ be a simple purely infinite $C^*$-algebra and let $p$ be a projection in $A$. Then $pA p$ and $A$ are stably isomorphic.

Proof. In the complex case, this result follows from Corollary 2.6 of [16]. The proof of that result and the proofs of the preliminary lemmas of Section 2 of [16] work the same in the real case. \hfill \Box

For the rest of this section, $f_\varepsilon$ will denote the real-valued function such that $f_\varepsilon(t) = 0$ for $t \leq \varepsilon/2$, $f_\varepsilon(t) = 1$ for $t \geq \varepsilon$, and $f_\varepsilon(t)$ is linear on $[\varepsilon/2, \varepsilon]$.

**Lemma 3.6.** For any real $C^*$-algebra $A$, the following are equivalent.

1. For any non-zero $a, b \in A$ there exist $x, y \in A$ with $a = xby$.

2. For any non-zero positive $a, b \in A$ there exists $x \in A$ with $a = xbx^*$.

Proof. (2) ⇒ (1). Let $0 \neq a, b \in A$. As in the complex case, described in 1.4.5 of [33], there exists $u \in A$ with $a = u(a^*a)^{1/4}$. Let $x \in A$ with $(a^*a)^{1/4} = xbx^*$ and observe that $a = (ux)b(b^*x^*)$.

(1) ⇒ (2). This uses the argument for the complex case, from Lemma 1.7 and Proposition 1.10 of [18]. If $a, b \in A$ are positive and non-zero and $\varepsilon$ is chosen so that $f_\varepsilon(b) \neq 0$ then $a = (zz^*zk)b(zz^*zk)^*$, where $x, y$ are chosen so that $a^{1/6} = x f_\varepsilon(b)y$, $k \geq 0$ is chosen so that $f_\varepsilon/2(b) = kbb$ and $z = x(f_\varepsilon(b)yy^*f_\varepsilon(b))^{1/2}$. \hfill \Box

**Lemma 3.7.** Let $A$ be a real $C^*$-algebra such that for all non-zero elements $a, b$ there exist $x, y$ with $a = xby$. Suppose that $A$ contains a non-zero projection and let $c$ be a non-zero positive element such that $cA^e \neq A$. Then $cA^e$ contains an infinite projection.

Proof. The argument from (viii) ⇒ (i) of Theorem 2.2 of [31] applies to the real case to show that for any non-trivial projection $p$ and positive element $x$ there is a Murray-von Neumann equivalence between $p$ and a subprojection of $x$. We will repeatedly use this fact.

In the unital case, this shows that the unit 1 is Murray-von Neumann equivalent to a projection of $cA^e$, which is necessarily infinite.

Now suppose that $A$ has no unit but has a non-zero projection $p$. Applying the fact above to a non-zero positive element $d$ in $(1-p)A(1-p)$ gives a projection $q$ such that $p \sim q$ and $p \perp q$. Now apply the fact again using the projection $p + q$ and the positive element $p$ to show that $p + q$ is infinite. Finally, apply the same fact using the projection $p + q$ and the positive element $c$ to show that $p + q$ is Murray-von Neumann equivalent to a projection in $cA^e$. \hfill \Box
Lemma 3.8. Let $A$ be a real simple C*-algebra. Then the following are equivalent:

1. $A$ is purely infinite,
2. $A$ is not isomorphic to $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$ and for each pair of non-zero elements $a, b \in A$ there exist $x, y \in A$ such that $a = xby$,
3. $A$ is not isomorphic to $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$ and for each pair of non-zero positive elements $a, b \in A$ there exists $x \in A$ such that $a = xbx^*$. 

Furthermore, if these conditions are satisfied, then for all $\varepsilon > 0$ the element $x$ in (3) can be chosen to satisfy $\|x\| \leq (\|a\|/\|b\|)^{1/2} + \varepsilon$.

Proof. As the result is well-known in the complex case, we may assume by Theorem 3.1 that $A$ is c-simple. By Lemma 3.6, (2) and (3) are equivalent.

For (1) $\Rightarrow$ (2), let $a, b$ be non-zero elements of $A$, identified with $e_{11}(K^R \otimes A)e_{11}$. We are assuming $A_C$ is simple, so Theorem 2.4 of [17] applied to the unital algebra $pAp$ implies that $K \otimes pA_Cp$ is algebraically simple. Then by Proposition 3.5, $K \otimes A_C$ is algebraically simple, whence $K^R \otimes A$ is. The argument from (ii) $\Rightarrow$ (xi) of Theorem 2.2 of [31], then produces $x, y \in K^R \otimes A$ with $a = xby$, so $a = (e_{11}xe_{11})b(e_{11}ye_{11})$.

For (2) $\Rightarrow$ (1), we use a simplified argument based on the proof of Theorem 1.2 of [31]. Note first that if a nonzero projection can be found in $A$ then Lemma 3.7 gives the result. (In particular, this takes care of the unital case.) Let $a, b$ be non-zero positive elements of $A$ with $da = ad = a$ (for a positive element $x$ with norm 1 take $a = f_{1/2}(x)$ and $d = f_{1/4}(x)$). Then let $s, t \in A$ with $d = sat$ and let $y = (as^*sa)^{1/2}t$. An easy argument shows that $\|y\|\|y^*\| = \|y^*\|$ hence $f_{1/2}(\|y\|) f_{1/4}(\|y^*\|) = f_{1/8}(\|y^*\|)$. Unless $f_{1/4}(\|y\|)$ is a projection, Lemma 4.2 of [7] gives a scaling element $t \in A$. In this case, $p_n = f_n + f_n^{1/2}t f_n^{1/2} + f_n^{1/2}t^* f_n^{1/2}$ (where $f_n = t^n(t^*)^n - t^{n+1}(t^*)^{n+1}$ for $n \geq 2$) is a projection by Theorem 3.1 of [7].

The final condition holds as in Lemma 2.4 of [28].

Theorem 3.9. A real c-simple C*-algebra $A$ is purely infinite if and only if $A_C$ is purely infinite.

Proof. From Theorem 3.3 of [43] we know that $A$ is purely infinite if $A_C$ is.

For the converse, suppose $A$ is purely infinite, let $\omega$ be a free ultrafilter on $\mathbb{N}$ and let $A_\omega$ be the corresponding ultrapower algebra, defined in Definition 6.2.2 of [39]. Note that the proofs of Proposition 6.2.6 of [39] and the preliminary Lemma 6.2.3 carry over directly to the real case (using Lemma 3.8). Therefore $A_\omega$ is simple and purely infinite. Suppose that $D$ is a dimension function, as defined in Definition 1.1.2 of [5], on the complexification $(A_\omega)_C \cong (A_C)_\omega$.

For each positive non-zero $a, b$ in $A_\omega$ there exist $x, y \in A_\omega$ with $b = xax^*$ and $a = yby^*$ so $D(a) = D(b)$. For any infinite projection $p \in A_\omega$, there exists a projection $q < p$ with $D(p) = D(q) + D(p - q) = D(p) + D(p)$, so $D(a) = D(p) = 0$ for each positive $a \in A_\omega$. Then for each positive $a \in (A_\omega)_C$, we have $0 \leq D(a) \leq D(a + \Phi(a)) = 0$. So there is no dimension function on $(A_\omega)_C$ and therefore, by Theorem II.2.2 of [5], no 2-quasitrace. Therefore $A_C$
is weakly purely infinite by Theorem 4.8 of [28]. By Corollary 4.16 of [28] it is therefore purely infinite.

□

**Corollary 3.10.**  
(1) If $A$ and $B$ are stably isomorphic real $C^*$-algebras, and if $A$ is purely infinite and c-simple then so is $B$.

(2) Any inductive limit of real purely infinite c-simple $C^*$-algebras is again purely infinite and c-simple.

(3) If $A$ and $B$ are purely infinite and c-simple, then so is $A \otimes_{\min} B$.

**Proof.** These results follow immediately from Theorem 3.9 and the same results in the complex case (see Proposition 4.1.8 of [39]).

□

We now work toward showing that the $K_0$ and $K_1$ groups of a real purely infinite algebra can be described in a similar way to the complex case. The next two lemmas provide the required modification of Lemma 1.7 of [19].

**Lemma 3.11.** Let $A$ be a real c-simple purely infinite unital $C^*$-algebra and let $u \in U(A)$ and let $\lambda \in \sigma(u)$. For any $\varepsilon > 0$ there exists $v \in U(A)$ such that $\|u - v\| < \varepsilon$ and

1. if $\lambda = \lambda^*$ then $v = v_0 + \lambda p$ where $p$ is a non-zero projection in $A$ and $v_0 \in U(p^+Ap^+)$.

2. if $\lambda \neq \lambda^*$ then $v = v_0 + \lambda p_1 + \lambda^* p_2$ where $p_1$ and $p_2$ are orthogonal non-zero orthogonal projections in $A_C$ satisfying $\Phi(p_1) = p_2$ and $v_0 \in U((p_1 + p_2)^+A(p_1 + p_2)^+)$. 

**Proof.** First assume that $\lambda = \lambda^*$. Let $h$ be a positive function on $\sigma(u)$ such that $\text{supp}(h) \subset N_{\varepsilon_0}(\lambda)$ and $h(z^*) = h(z)$ for all $z \in \sigma(u)$. Then $h(u) \in A$ and let $p$ be a non-zero projection in $h(u)Ah(u)$. As in the proof of Lemma 1.7 of [19], we have $\|u - (p^+Ap^+ + \lambda p)\| \leq 3\varepsilon_0$. Then the polar decomposition of $(p^+Ap^+ + \lambda p)$ yields a unitary $v$ of the required form that, if $\varepsilon_0$ is sufficiently small, will satisfy $\|u - v\| < \varepsilon$.

Now assume $\lambda \neq \lambda^*$. Choose $\varepsilon_0$ small enough so that $N_{\varepsilon_0}(\lambda) \cap N_{\varepsilon_0}(\lambda^*) = \emptyset$. Let $h_1$ be a positive function on $\sigma(u)$ such that $\text{supp}(h_1) \subset N_{\varepsilon_0}(\lambda)$. By Theorem 3.9, $A_C$ is purely infinite so there is a non-zero projection $p_1$ in $B = h_1(u)Ah_1(u)$. Define $p_2 = \Phi(p_1) \in \Phi(B)$ and $p = p_1 + p_2$. Now $\Phi(h_1(u)) = h_2(u)$ where $h_2$ is the continuous function on $\sigma(u)$ defined by $h_2(z) = h_1(z^*)$. Since $\text{supp}(h_2) \subset N_{\varepsilon_0}(\lambda^*)$, we have $h_1(u)h_2(u) = 0$. Thus $p_1$ and $p_2$ are orthogonal projections and $p \in A$.

As in Lemma 1.7 of [19], we have $\|up_1 - \lambda p_1\| \leq \varepsilon_0$ and $\|up_2 - \lambda^* p_2\| \leq \varepsilon_0$ from which it follows that $\|u - (p^+Ap^+ + \lambda p_1 + \lambda^* p_2)\| \leq 8\varepsilon_0$. The required unitary $v$ is obtained by taking the polar decomposition of $p^+Ap^+ + \lambda p_1 + \lambda^* p_2$ in $A$.

□

**Lemma 3.12.** Let $A$ and $u$ be as above. Then there is a projection $p$ in $A$ and a unitary $v$ in $U(p^+Ap^+)$ such that $u \sim v + p$.

**Documenta Mathematica 16 (2011) 619–655**
Proof. If \( I \in \sigma(u) \) then using Lemma 3.11, approximate \( u \) by an element of the form \( v + p \). If the approximation is close enough, then the two unitaries will be in the same path component.

If \( \lambda \in \sigma(u) \) where \( \lambda \neq \lambda^* \), use Lemma 3.11 to approximate \( u \) by \( \lambda p_1 + \lambda^* p_2 \). Then we can easily find a path from \( \lambda p_1 + \lambda^* p_2 \) to \( p_1 + p_2 \) in \( (p_1 + p_2)A(p_1 + p_2) \). The only possibility left is \( u = -1 \). In that case, find two orthogonal projections \( q_1 \) and \( q_2 \) and a partial isometry \( s \) such that \( ss^* = q_1 \) and \( s^* s = q_2 \). Let \( p = q_1 + q_2 \). The projection \( p \) can be rotated to \(-p\) within the \( 2 \times 2 \) matrix algebra generated by \( q_1 \), \( q_2 \) and \( s \). Hence the unitary \(-1 = -(p^2) + p \) can be connected to the unitary \(- (p^2) + p \).

Proposition 3.13. Let \( A \) be a c-simple purely infinite real C*-algebra. Then

\[
\begin{align*}
(1) & \quad K_0(A) = \{ [p] \mid p \text{ is a non-zero projection in } A \} \\
(2) & \quad K_1(A) = U(A)/U_0(A) \quad (\text{for } A \text{ unital}).
\end{align*}
\]

Proof. In the complex case, these results are proven in Section 1 of [19]. The proofs of those results as well as the proofs of the preliminary lemmas carry over to the real case, with two modifications. The first is to the proof of Lemma 1.7 of [19], which we already addressed with the proof of Lemma 3.12 above.

Secondly, in the proof of Lemma 1.1 of [19] the author uses an element of the form

\[
\tilde{w} = w + w^* + (1 - w^* w - w w^*), \quad (\text{where } w^2 = 0)
\]

that is a unitary lying in the finite dimensional C*-algebra generated by \( w \). In the complex case it follows that \( \tilde{w} \in U_0(A) \), whereas in the real case unitary groups of finite dimensional C*-algebras are not connected in general.

However, if instead we take \( \tilde{w} = w - w^* + (1 - w^* w - w w^*) \) then \( \tilde{w} \) is in the connected component of the identity, as it corresponds to a matrix of the form

\[
\begin{pmatrix}
0 & 1 \\
-1 & 0
\end{pmatrix}.
\]

The proof of Lemma 1.1 of [19] can be completed without change using this alternative \( \tilde{w} \).

We note that part (1) of Proposition 3.13 appeared as Proposition 11 of [10].

4. Exponential Rank

Definition 4.1. An element \( a \) in a real C*-algebra \( A \) is skew-adjoint if \( a^* = -a \). The set of skew-adjoint elements is denoted by \( A_{sk} \).

If \( a \) is skew-adjoint, then \( \sigma(a) = -\sigma(a) \subseteq i \mathbb{R} \) and the real unital C*-algebra generated by \( a \) is isomorphic to

\[
\{ f \in C(\sigma(a), \mathbb{C}) \mid f(it) = f(-it) \}.
\]

Furthermore, if \( a \) is a skew-adjoint element in a real unital C*-algebra \( A \), then \( \exp(a) \) is a unitary in \( A \).

Lemma 4.2. Let \( A \) and \( B \) be unital real C*-algebras.

\[
\begin{align*}
(1) & \quad U_0(A) = \{ \prod_{i=1}^n \exp(k_i) \mid k_i \in A_{sk}, n \in \mathbb{N} \} \\
(2) & \quad \text{If } \alpha: A \to B \text{ is unital and surjective, then } \alpha(U_0(A)) = U_0(B).
\end{align*}
\]
Suppose first that $u$ is a unitary element in $A$ with $\|u - 1\| < 2$. Then $-1 \notin \sigma(u)$. We define a continuous function $f: \mathbb{T} \setminus \{0\} \to i(-\pi, \pi)$ by $f(\exp(it)) = it$ for $t \in (-\pi, \pi)$. Then $f(u)$ is in the real $C^*$-algebra generated by $u$, is skew-adjoint, and satisfies $\exp(f(u)) = u$.

More generally, if $u \in \mathcal{U}_0(A)$ then there exists a chain $u = u_0, u_1, u_2, \ldots, u_n = 1$ with $\|u_i - u_{i+1}\| < 2$ for all $i \in \{1, 2, \ldots, n\}$. Then applying the previous paragraph we have $u_{i-1}u_i^* = \exp(k_i)$ for all $i$ with $1 \leq i \leq n$. Then $u = \prod_{i=1}^{n} \exp(k_i)$.

Conversely, if $\{k_i\}_{i=1}^n$ is any collection of skew-adjoint elements, then $u(t) = \prod_{i=1}^{n} \exp(tk_i)$ for $0 \leq t \leq 1$ is a continuous path of unitaries from $1_A$ to $\prod_{i=1}^{n} \exp(k_i)$. This proves (1).

For (2), the inclusion $\alpha(\mathcal{U}_0(A)) \subseteq \mathcal{U}_0(B)$ is immediate. Let $u \in \mathcal{U}_0(B)$. Then $u = \prod_{i=1}^{n} \exp(k_i)$ for some skew-adjoint elements $k_i \in B$. Let $l_i \in A$ be elements such that $\alpha(l_i) = k_i$. We may assume that $l_i$ is skew-adjoint for all $i$, by replacing with $\frac{1}{2}(l_i - l_i^*)$ if necessary. Then $u = \alpha(\prod_{i=1}^{n} \exp(l_i))$.

Let $\mathcal{E} = \{\exp(k) \mid k \in A_{sk}\}$ and let $\mathcal{E}^n$ be the set of all products of at most $n$ elements of $\mathcal{E}$. Thus $\mathcal{U}_0(A) = \bigcup_{n=1}^{\infty} \mathcal{E}^n$. The argument in the proof above also implies that the set $\mathcal{E}^{n+1}$ contains the topological closure of $\mathcal{E}^n$ so that we have the an increasing sequence

$$\mathcal{E} \subseteq \mathcal{E}^2 \subseteq (\mathcal{E}^2)^2 \subseteq \cdots \subseteq (\mathcal{E}^n)^n \subseteq (\mathcal{E}^n)^{n+1} \subseteq \cdots$$

similar to that in [37], motivating the following definition.

**Definition 4.3.**

1. The exponential rank of $A$, written $\text{cer}(A)$, is equal to the integer $n$ if $\mathcal{E}^n$ is the smallest set in this sequence to be equal to $\mathcal{U}_0(A)$ and is equal to the symbol $n + \varepsilon$ if $\mathcal{E}^n$ is the smallest set to be equal to $\mathcal{U}_0(A)$. If $\mathcal{E}^n \neq \mathcal{U}_0(A)$ for all $n$ then $\text{cer}(A) = \infty$.
2. The exponential length of $A$, written $\text{cel}(A)$, is equal to the smallest number $0 < \text{cel}(A) \leq \infty$ such that every unitary $u$ in $\mathcal{U}_0(A)$ can be written in the form

$$u = \exp(k_1) \exp(k_2) \cdots \exp(k_n)$$

where $k_i \in A_{sk}$ and

$$\|k_1\| + \|k_2\| + \cdots + \|k_n\| \leq \text{cel}(A).$$

With these definitions, the proofs of Section 2 of [37] can be applied with minimal modification to prove the following results.

**Lemma 4.4.** Let $A$ be a real unital $C^*$-algebra and let $n$ be a positive integer.

1. If $\text{cel}(A) < n\pi$ then $\text{cer}(A) \leq n$.
2. If $\text{cel}(A) \leq n\pi$ then $\text{cer}(A) \leq n + \varepsilon$.

**Lemma 4.5.** Let $A$ be a real unital $C^*$-algebra. If every unitary $u \in \mathcal{U}_0(A)$ can be connected to the identity by a rectifiable path of length no more than $M$, then $\text{cel}(A) \leq M$. 

*Documenta Mathematica* 16 (2011) 619–655
A real $C^*$-algebra $A$ has real skew rank zero if the elements of $A_{sk}$ with finite spectrum are dense in $A_{sk}$.

In the case of a complex $C^*$-algebra $A$ there is a bicontinuous bijection $A_{sa} \to A_{sk}$ given by multiplication by $i$, showing that $A$ has skew rank zero if and only it has real rank zero. However, in the case of real $C^*$-algebras things are more subtle. For example the condition of being skew-rank zero is not equivalent (in the unital case) to the condition that the invertible elements of $A_{sk}$ are dense. Indeed, all finite dimensional real $C^*$-algebras have real skew rank zero, but the invertibles of $(M_n)_{sk}$ are dense only if $n$ is even.

**Proposition 4.7.** Let $A$ be a real unital $c$-simple purely infinite $C^*$-algebra satisfying $[1] \in 2K_0(A)$. Then the invertibles of $A_{sk}$ are dense in $A_{sk}$ and $A$ has real skew rank zero.

**Proof.** Let $A$ be a real purely infinite $C^*$-algebra such that $[1] \in 2K_0(A)$. Let $a \in A_{sk}$ and let $\varepsilon > 0$ be given. Define functions $g : i\mathbb{R} \to \mathbb{R}$ and $f : i\mathbb{R} \to i\mathbb{R}$ by

$$g(it) = \max\{\varepsilon - |t|, 0\} \quad \text{and} \quad f(it) = \begin{cases} i(t + \varepsilon) & t \leq -\varepsilon \\ 0 & |t| < \varepsilon \\ i(t - \varepsilon) & t \geq \varepsilon . \end{cases}$$

Then $g(a) \in A_+$ and $f(a) \in A_{sk}$.

Since $A$ is purely infinite, there is a projection $p \in g(a)Ag(a)$ with $2[p] = [1] \in K_0(A)$. Then $[1 - p] = [p]$ so there is a partial isometry $s$ such that $s^*s = 1 - p$ and $ss^* = p$. Since $f(a)g(a) = 0$ we have $f(a) = (1 - p)f(a)(1 - p)$.

Let $b = f(a) + \varepsilon(s - s^*)$. In matrix form under the decomposition indicated by the projection sum $1 = (1 - p) + p$ we have

$$b = \begin{pmatrix} f(a) & -\varepsilon \\ \varepsilon & 0 \end{pmatrix}$$

whence $b$ is invertible. This proves the first statement.

For the second statement, again let $a \in A_{sk}$ and let $\varepsilon > 0$ be given. By the first part of the theorem, we may assume that $a$ is invertible, hence $\sigma(a) \subseteq i\mathbb{R} \setminus \{0\}$. Write $a = a_1 + a_2$ where the elements $a_1 \in A_C$ satisfy $\sigma(a_1) \subseteq i(0, \infty)$ and $\sigma(a_2) \subseteq i(-\infty, 0)$. Note also that $\Phi(a_1) = a_2$.

Since $A_C$ is simple and purely infinite it has real rank zero, so there exists $b_1 \in (A_C)_{sk}$ such that $\sigma(b_1)$ is a finite subset of $i\mathbb{R}^+$ and $||a_1 - b_1|| < \varepsilon/2$. Let $b_2 = \Phi(b_1)$ and let $b = b_1 + b_2$. Then $b$ is a skew-adjoint element of $A$ with finite spectrum and $||a - b|| < \varepsilon$.

**Lemma 4.8.** Let $A$ be a real $c$-simple unital $C^*$-algebra such that $[1] \in 2K_0(A)$. Let $u \in \mathcal{U}(A)$ be a unitary such that $\sigma(u) \neq S^1$. Then for every $\varepsilon > 0$ there is a unitary $v$ with finite spectrum such that $||u - v|| < \varepsilon$.

**Proof.** If $-1 \notin \sigma(u)$, then there is a continuous function $f : \sigma(u) \to i[-\pi, \pi]$ that is a right inverse to the function $it \mapsto \exp(it)$ and that satisfies $f(z^*) = f(z)^*$. Then $f(u) \in A_{sk}$ can be approximated within $\delta$ by a skew-adjoint
element $b$ with finite spectrum by Proposition 4.7. For an appropriate choice of $\delta$, this implies that $\exp(b) \in \mathcal{U}(A)$ approximates $u$ within $\epsilon$.

Similarly, if $1 \notin \sigma(u)$, then there is a continuous function $f : \sigma(u) \to [-\pi, \pi]$ that is a right inverse to the function $it \mapsto -\exp(-it)$ and that satisfies $f(z^*) = f(z)^*$. In the general case, suppose that $\lambda \notin \sigma(u)$ for some $\lambda \in S^1$. Let $\sigma_1 = \{w \in \sigma(u) \mid \Re(w) > \Re(\lambda)\}$ and let $\sigma_2 = \{w \in \sigma(u) \mid \Re(w) < \Re(\lambda)\}$. Then $\sigma = \sigma_1 \cup \sigma_2$. Let $u_1 = u_1E_u(\sigma)$, where $E_u(\sigma_i)$ denotes the spectral projection of $u$ associated with the clopen subset $\sigma_i$ of $\sigma$. Then $1 \notin \sigma(u_2)$ and $-1 \notin \sigma(u_1)$. Using the results from the first two paragraphs, let $v_i$ be a unitary that approximates $u_i$ in $E_u(\sigma_i)AE_u(\sigma_i)$ within $\epsilon$. Then since $u = u_1 + u_2$ we have that $v = v_1 + v_2$ is a unitary that approximates $u$ within $\epsilon$. □

**Lemma 4.9.** Let $A$ be a real unital simple purely infinite $C^*$-algebra let $u \in \mathcal{U}(A)$ and let $\{\lambda_1, \ldots, \lambda_n\}$ be a subset of $\sigma(u)$ that is closed under conjugation. For any $\epsilon > 0$ there exist $v \in \mathcal{U}(A)$ and orthogonal projections $p_1, \ldots, p_n \in A_\mathbb{C}$ such that $\|u - v\| < \epsilon$ and $v = v_0 + \lambda_1p_1 + \cdots + \lambda_np_n$ with $v_0 \in \mathcal{U}(\sum p_i)$.

Furthermore, the elements $\sum p_i$ and $\sum \lambda_ip_i$ are both in $A$.

**Proof.** Use the constructions of Lemma 3.11 above as in the proof of Lemma 6 of [34]. □

**Lemma 4.10.** Let $A$ be a real unital $C^*$-algebra and let $u \in \mathcal{U}(A)$. For any $\epsilon > 0$ there exists an $h \in M_2(A)_{sk}$ such that $\|u \oplus u^* - \exp(h)\| < \epsilon$.

**Proof.** As in the proof of Corollary 5 of [34], there exists a continuous path $v(t)$ of unitaries in $M_2(A)$ with $v(0) = 1$ and $v(\pi/2) = u \oplus u^*$ such that $-1 \notin \sigma(v(t))$ for $0 \leq t < \pi/2$. Thus we can find a $t$ close enough to $\pi/2$ such that $\|u \oplus u^* - v(t)\| < \epsilon$ and $v(t) = \exp(h)$ for a skew-adjoint $h$.

**Lemma 4.11.** Let $A$ be a real unital c-simple purely infinite $C^*$-algebra such that $[1] \in 2K_0(A)$. Let $e_1, e_2, e_3, e_4$ be nonzero orthogonal projections in $A$ that sum to 1. Let $a$ be a partial isometry such that $a^*a = e_2$ and $aa^* = e_3$. Let $u \in \mathcal{U}(e_1AE_1)$ and $v \in \mathcal{U}(e_2AE_2)$ be unitaries with $\sigma(u) = S^1$. Then for all $\epsilon > 0$ there is a unitary $z \in \mathcal{U}(A)$ and a unitary $w \in \mathcal{U}(e_4AE_4)$ with finite spectrum such that

$$\|z^*(u + 1 - e_1)z - (u + v + av^*a^* + w)\| < \epsilon .$$

**Proof.** This proof closely follows that of Lemma 7 of [34]. By Lemma 4.10 there is a unitary in $(e_2 + e_3)A(e_2 + e_3)$ that is arbitrarily close to $v + av^*a^*$ and that has the form $\exp h$ for $h \in A_{sk}$. This in turn can be approximated by a unitary that has finite spectrum by Proposition 4.7. The general form of such a unitary is

$$\sum_{k=1}^{n} (\lambda_k g_{k1} + \lambda_k^* g_{k2}) + 1_{q_{01}} + (-1)q_{02}$$

**Documenta Mathematica 16 (2011) 619–655**
where $\lambda_k^* \neq \lambda_k$, the nonzero projections $q_k \in A_\mathbb{C}$ satisfy $\Phi(q_k) = q_{k2}$ for $1 \leq k \leq n$, and the (possibly zero) projections $q_0$ are in $A$. Furthermore, the $q_k$ are orthogonal and sum to $e_2 + e_3$. Without loss of generality, we assume that $v + av^*a^*$ has this form. With an obvious choice of coefficients $\lambda_k$, we can write this as

$$v + av^*a^* = \sum_{k=0}^n \sum_{r=1}^2 \lambda_{ki} q_{ki} = \sum \lambda_{ki} q_{ki}.$$  

(Henceforth in this proof will use an undecorated $\sum$ to represent a double sum indexed as $\sum_{k=0}^n \sum_{r=1}^2$.)

Now we replace $u$ by a nearby element of the form given by Lemma 4.9. Specifically, there are orthogonal projections $p_k \in e_1 A_\mathbb{C} e_1$ and, setting $p = e_1 - \sum p_k \in A$, there is a unitary $u_0 \in pA p$ such that

$$u = u_0 + \sum \lambda_{ki} p_k$$

(where the projection $p_{0i} = 0$ if and only if $q_{0i} = 0$).

For each $k \in \{1, \ldots, n\}$ let $c_k \in A_\mathbb{C}$ be a partial isometry such that $c_k^* c_k = p_k$ and $c_k c_k^* < p_k$. Then $c_k^2 = \Phi(c_k)$ satisfies $c_k^2 c_{k2} = p_k$ and $c_k c_{k2} < p_k$ and $c_k = c_k + c_{k2} \in A$ satisfies $c_k^* c_k = p_k + p_{k2}$ and $c_k c_k^* < p_k + p_{k2}$. For $k = 0$ we obtain partial isometries $c_0 \in A$ such that $c_0 c_0 = p_0$ and $c_0 c_0^* < p_0$. Then $c = p + \sum c_k \in A$ satisfies

$$c^* c = e_1, \quad cc^* = e_1 - \sum (p_k - c_k c_k^*), \quad \text{and} \quad
c^* c = u_0 + \sum \lambda_{ki} c_k c_k^*.$$  

Similarly we can find a collection of partial isometries $d_k$ with domain projection $q_k$ and range projection a subprojection of $p_k - c_k^* c_k$ that also satisfy $\Phi(d_k) = d_{k2}$ for $k \neq 0$ and $\Phi(d_k) = d_k$ for $k = 0$. Then the partial isometry $d = \sum d_k \in A$ satisfies

$$d^* d = e_2 + e_3, \quad dd^* \leq \sum (p_k - c_k c_k^*), \quad \text{and} \quad
d \left( \sum \lambda_{ki} q_{ki} \right) d^* = \sum \lambda_{ki} d_k d_k^*.$$  

Now, choose a partial isometry $b$ such that

$$b^* b < e_4, \quad \text{and} \quad
b^* b = \sum (p_k - c_k c_k^* - d_k d_k^*).$$

and define

$$w_0 = \sum \lambda_{ki} (p_k - c_k c_k^* - d_k d_k^*) b.$$  

Then $z_0 = b + c + d$ is a partial isometry with $z_0^* z_0 = e_1 + e_2 + e_3 + b^* b$ and $z_0 z_0^* = e_1$. So in $K_0(A)$ we have $[c_1] = [e_1 + e_2 + e_3 + b^* b]$, which implies $[1 - e_1] = [e_4 - b^* b]$. By Proposition 11 of [10], there is a partial isometry $z_1 \in A$ such that $z_1^* z_1 = 1 - e_1$ and $z_1^* z_1 = e_4 - b^* b$. Then $w = w_0 + e_4 - b^* b$ is a unitary with finite spectrum in $e_4 A e_4$ and $z = z_0 + z_1$ is a unitary in $A$ that satisfies $z^* (u + 1 - e_1) z = u + \sum \lambda_{ki} q_{ki} + w$. 

\[\square\]
**Theorem 4.12.** Let $A$ be a real unital $c$-simple purely infinite $C^*$-algebra such that $[1] \in 2K_0(A)$. For every $u \in \mathcal{U}_0(A)$ and every $\varepsilon > 0$ there is a unitary $v$ with finite spectrum such that $\|u - v\| < \varepsilon$.

**Proof.** With the lemmas that we have developed, the proof is now the same as that of the unital case of Theorem 1 and Corollary 2 of [34], except that wherever there is an element of the form $\exp(ih)$ where $h$ is self-adjoint, we use $\exp(k)$ where $k$ is skew-adjoint. \(\square\)

As in the complex case, we have the following corollary concerning exponential length.

**Corollary 4.13.** Let $A$ be a real unital $c$-simple purely infinite $C^*$-algebra such that $[1] \in 2K_0(A)$. Then $\operatorname{cel}(A) \leq 4$.

**Proof.** By Theorem 4.12, every unitary $u \in \mathcal{U}_0(A)$ can be approximated within $\varepsilon$ by a unitary $v$ with finite spectrum. For $\varepsilon$ sufficiently small, $\|v^*u - 1\| < \varepsilon$ implies there exists a skew-adjoint $k_2$ such that $v^*u = \exp(k_2)$ with $\|k_2\| \leq 4 - \pi$. As $v$ has finite spectrum, there exists a skew-adjoint $k_1$ such that $v = \exp(k_1)$ and $\|k_1\| \leq \pi$. Then $v = \exp(k_1)\exp(k_2)$ and $\|k_1\| + \|k_2\| \leq 4$. \(\square\)

## 5. Homomorphisms from $\mathcal{O}_n^\mathbb{R}$

The following theorem gives the real version of the Rokhlin property of the Bernoulli shift, established in [15] and summarized in [39]. Let $M_{2\infty} = \lim_{k \to \infty} M_{2k}$ be the real CAR algebra and let $\mathbb{H}$ be the real $C^*$-algebra of quaternions.

**Proposition 5.1.** Let $\sigma$ be the one-sided Bernoulli shift on $M_{2\infty}$. For each $\varepsilon > 0$ and for each $r \in \mathbb{N}$ there exist $k \in \mathbb{N}$ and projections $e_0, e_1,\ldots, e_{2^r} = e_0 \in M_{2k}$ such that $\sum_{j=1}^{2^r} e_j = 1$ and $\|\sigma(e_j) - e_{j+1}\| < \varepsilon$ for all $j = 0, 1, 2,\ldots, 2^r - 1$.

**Proof.** Let $A_k = M_{2k}$ and let $A = M_{2\infty}$. Using the notation of Proposition 5.1.3 of [39], let $S$ denote the unilateral shift on $\ell^2(\mathbb{N}, \mathbb{C})$, let $\omega_k = \exp(2\pi i/2^k)$ for each $k \geq 0$ and, given $\delta > 0$, let

$$f_0 = \frac{1}{\sqrt{\pi_0}}(1, 1,\ldots, 1, 0, 0,\ldots) \in \ell^2(\mathbb{N}, \mathbb{R})$$

be a unit vector with $\|Sf_0 - f_0\| < \delta$ and let

$$f_1 = \frac{1}{\sqrt{\pi_1}}(0, 0,\ldots, 0, 1, -1, 1, -1,\ldots, -1, 0, 0,\ldots)$$

be a unit vector in $\ell^2(\mathbb{N}, \mathbb{R})$, orthogonal to $f_0$, with $\|Sf_1 + f_1\| < \delta$. Then, for $r \in \mathbb{N}$, let $f_2,\ldots, f_r \in \ell^2(\mathbb{N}, \mathbb{C})$ be defined by

$$f_j = \frac{1}{\sqrt{\pi_j}}(0, 0,\ldots, 0, 1, \omega_j, \omega_j^2,\ldots, \omega_j^{n_j-1}, 0, 0,\ldots)$$

where there are sufficiently many initial zeros to make $f_j$ orthogonal to its predecessors and where $n_j$ is chosen so that

$$\langle f_j, \overline{f_j} \rangle = 1 + \omega_j^2 + \ldots \omega_j^{2(n_j-1)} = 0$$

**Documenta Mathematica** 16 (2011) 619–655
and \(\|Sf_j - \omega_j f_j\| < \delta\). If \(f_j = g_j + ih_j\) with \(g_j, h_j \in \ell^2(\mathbb{N}, \mathbb{R})\) then, from the orthogonality of \(f_j\) and \(f_j\), \(\|g_j\| = \|h_j\| = 1/\sqrt{2}\).

Let \(a : \ell^2(\mathbb{N}, \mathbb{C}) \to A_{\mathbb{C}}\) be the map described in [15] and [39] satisfying the canonical anticommutation relations and observe that \(a\) maps \(\ell^2(\mathbb{N}, \mathbb{R})\) into \(A\). Let \(v_1 = w_1 = a(f_1)(a(f_0) + a(f_0)^*)\) and, for each \(2 \leq j \leq r\) let \(v_{2j-2} = a(f_j)(a(f_0) + a(f_0)^*), v_{2j-1} = a(f_j)(a(f_0) + a(f_0)^*), w_{2j-2} = a(\sqrt{2} g_j)(a(f_0) + a(f_0)^*) = (v_{2j-2} + v_{2j-1})/\sqrt{2}\) and \(w_{2j-1} = a(\sqrt{2} h_j)(a(f_0) + a(f_0)^*) = -i(v_{2j-2} - v_{2j-1})/\sqrt{2}\). Note that \(\{w_1, w_2, \ldots, w_{2r-1}\} \subset A_k\) for all sufficiently large \(k\).

It is noted in the proof of Proposition 4.1 of [15] that the elements \(v_i\) for \(1 \leq i \leq 2r - 1\) satisfy the relations \(v_i v_j + v_j v_i = 0\) and \(v_i v_j^* + v_j^* v_i = \delta_{ij} 1\). It follows from this that the elements \(w_i\) for \(1 \leq i \leq 2r - 1\) satisfy the same relations. Therefore, using the matrix units described in the proof of Proposition 4.1 of [15], the real C*-algebra \(B\) generated by \(w_1, \ldots, w_{2r-1}\) is isomorphic to \(M_{2r-1}\).

Slightly varying the proof of Proposition 4.1 of [15], let \(\beta\) be the automorphism of the complexification of \(B\) determined by \(\beta(v_1) = -v_1, \beta(v_{2j}) = \omega_j v_{2j}\) and \(\beta(v_{2j+1}) = \omega_j^* v_{2j+1}\) for each \(1 \leq j \leq r - 1\). Note that \(\beta(w_{2j}) = -\frac{1}{2}(\omega_j + \omega_j^*) w_{2j} + \frac{1}{2}(\omega_j - \omega_j^*) w_{2j+1}\) and \(\beta(w_{2j+1}) = -\frac{1}{2}(\omega_j - \omega_j^*) w_{2j} + \frac{1}{2}(\omega_j + \omega_j^*) w_{2j+1}\), so that \(\beta\) leaves the real algebra \(B\) invariant. Identifying \(B\) with \(M_{2r-1}\), there is an orthogonal matrix \(W\) implementing \(\beta\). By standard linear algebra, described for example in Section 81 of [22], \(W\) is orthogonally conjugate to an orthogonal matrix consisting of diagonal elements \(\pm 1\) and diagonal \(2 \times 2\) rotation matrices, determined by the eigenvalues of \(W\).

As in [39], on the complexification of \(B\), identified with \(M_{22r-1}(\mathbb{C})\), \(\beta\) is implemented by a diagonal unitary with entries \(1, \omega_1, \omega_1^2, \ldots, \omega_1^{2r-1}\), each repeated \(2^{r-1}\) times. (The unitary arises as the tensor product of one diagonal unitary with entries \(1, \omega_1, \omega_1^2, \ldots, \omega_1^{2r-1}\) and another with entries \(1, \omega_1^2, \omega_1^4, \ldots, \omega_1^{2(2r-1)}\).) On \(B \cong M_{2r-1}\), the orthogonal matrix \(W\) implementing \(\beta\) is therefore conjugate to an orthogonal matrix with \(2 \times 2\) diagonal blocks \(\text{diag}(1, -1), R, R^2, \ldots, R^{2^{r-1}-1}\), each repeated \(2^{r-1}\) times, where

\[
R = \begin{pmatrix}
\cos(\pi/2^{r-1}) & -\sin(\pi/2^{r-1}) \\
\sin(\pi/2^{r-1}) & \cos(\pi/2^{r-1})
\end{pmatrix}.
\]

The cyclic shift on \(M_{2r}\) is implemented by the unitary

\[
V = \begin{pmatrix}
0 & 0 & 0 & \ldots & 0 & 1 \\
1 & 0 & 0 & \ldots & 0 & 0 \\
0 & 1 & 0 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & 0 & 1
\end{pmatrix},
\]

which is orthogonally conjugate to \(\text{diag}(1,-1), R, R^2, \ldots, R^{2^{r-1}-1}\). It follows that the orthogonal element \(W\) implementing \(\beta\) on \(B\) is orthogonally conjugate to a direct sum of \(2^{r-1}\) copies of \(V\) and thus that \(\beta\) is conjugate to a direct sum of \(2^{r-1}\) cyclic shifts. It follows that there are \(2^r\) orthogonal
projections $e_0, e_1, \ldots, e_{2^n} = e_0$ in $B$ (each of rank $2^{n-1}$) that are cyclically permuted by $\beta$. As in the proof of Proposition 4.1 of [15], a suitable choice of $\delta$ at the start of the proof ensures that $\|\sigma(e_j) - \beta(e_j)\| < \varepsilon$ for each $j$ and therefore the projections $e_0, e_1, \ldots, e_{2^n} = e_0$ have the required properties. □

**Theorem 5.2.** Let $D$ be a real unital $C^*$-algebra satisfying

(i) the canonical homomorphism $\mathcal{U}(D)/\mathcal{U}_0(D) \to K_1(D)$ is an isomorphism, and

(ii) $\text{cel}(D) < \infty$.

Let $n$ be an even integer, let $\phi, \psi$ be unital homomorphisms from $\mathcal{O}_n^{\mathbb{R}}$ to $D$, let $\lambda$ be the endomorphism of $D$ defined by $\lambda(a) = \sum_{j=1}^n \phi(s_j)a\phi(s_j)^*$ and let $u \in \mathcal{U}(D)$ be defined by $u = \sum_{j=1}^n \psi(s_j)\phi(s_j)^*$, where $s_1, \ldots, s_n$ are the canonical generators of $\mathcal{O}_n^{\mathbb{R}}$. Then the following are equivalent:

1. $u \in \{v\lambda(v)^* \mid v \in \mathcal{U}(D)\}$,
2. $[u] \in (n-1)K_1(D)$,
3. $[\phi] = [\psi] \in K\mathcal{K}(\mathcal{O}_n^{\mathbb{R}}, D)$,
4. $\phi$ and $\psi$ are approximately unitarily equivalent.

In particular, these statements are equivalent if $D$ is a real unital purely infinite $C^*$-algebra.

**Proof of Theorem 5.2.** The proof of the equivalence of the four statements, assuming (i) and (ii), is similar to that of the complex case in Sections 3 and 4 of [38], modified only by the use of unitaries of the form $\exp(h)$ with $h \in A_N$ in the proof of the real version of Lemma 4.6 of [38]. We note that in the proof of the real version of Lemma 3.7 of [38], the required result from [19] holds, as was observed already in the proof of Proposition 3.13 above.

Suppose $D$ is a real unital purely infinite $C^*$-algebra. Then condition (i) holds for $D$ by Proposition 3.13. Since $K_0(\mathcal{O}_n^{\mathbb{R}}) = \mathbb{Z}_{n-1}$ and $n$ is even, we have $[1_{\mathcal{O}_2}] \in 2K_0(\mathcal{O}_n^{\mathbb{R}})$. Using the unital homomorphism $\phi$ (or $\psi$) we obtain $[1_{\mathcal{O}_2}] \in 2K_0(D)$. Then condition (ii) holds by Corollary 4.13. □

**Corollary 5.3.**

1. Let $A$ be a real unital purely infinite c-simple $C^*$-algebra. Any two unital homomorphisms $\phi, \psi : \mathcal{O}_2^{\mathbb{R}} \to A$ are approximately unitarily equivalent.

2. Any inductive limit of the form $\mathcal{O}_2^{\mathbb{R}} \to \mathcal{O}_2^{\mathbb{R}} \to \mathcal{O}_2^{\mathbb{R}} \to \ldots$, with unital connecting homomorphisms, is isomorphic to $\mathcal{O}_2^{\mathbb{R}}$.

3. $\mathcal{O}_2^{\mathbb{R}} \otimes \mathcal{O}_2^{\mathbb{R}} \cong \mathcal{O}_2^{\mathbb{R}}$.

4. $\bigotimes_{n=1}^\infty \mathcal{O}_2^{\mathbb{R}} \cong \mathcal{O}_2^{\mathbb{R}}$.

5. $\mathcal{O}_2^{\mathbb{R}} \otimes M_{2^\infty} \cong \mathcal{O}_2^{\mathbb{R}}$.

6. $\mathcal{O}_2^{\mathbb{R}} \otimes \mathbb{H} \cong \mathcal{O}_2^{\mathbb{R}}$.

**Proof.** We know that $KK^\text{rft}(\mathcal{O}_2^{\mathbb{R}}) = 0$ from Section 5 of [8] so the universal coefficient theorem (Theorem 4.1 of [9]) implies that $KK_0(\mathcal{O}_2^{\mathbb{R}}, D) = 0$. Then part (1) follows immediately from Theorem 5.2.
Parts (2) and (3) can be proven in the same way as in the complex case. See Corollary 5.1.5 and Theorem 5.2.1 in [39]. Then part (4) follows from parts (2) and (3).

There is an isomorphism $O_R^2 \cong M_2(O_R^2)$, established as in the complex case: if $s_1$ and $s_2$ are generators of $O_R^2$ satisfying the canonical relations $s_i^*s_j = \delta_{ij}1_{O_R^2}$ and $\sum_{i=1}^2 s_is_i^* = 1$, then

$$S_1 = \begin{pmatrix} s_1 & s_2 \\ \bar{s}_1 & \bar{s}_2 \end{pmatrix} \quad \text{and} \quad S_2 = \begin{pmatrix} 0 & 0 \\ s_1 & s_2 \end{pmatrix}$$

satisfy the same relations and generate $M_2(O_R^2)$. Using that isomorphism, part (5) follows from part (2).

Finally, part (6) follows from (5) and the formula $M_2 \otimes \mathbb{H} \cong M_2$, which follows from Theorem 10.1 of [21] or from Theorem 4.8 of [42]. \qed

6. Tensor Product Theorems

In this section, we reproduce for real $C^*$-algebras some standard results regarding tensor products with $O^2_R$ and $O^\infty_R$.

Definition 6.1.

1. A real (resp. complex) $C^*$-algebra $A$ is amenable if for all $\varepsilon > 0$ and all finite subsets $F \subset A$, there is a finite dimensional real (resp. complex) $C^*$-algebra $B$ and contractive completely positive linear maps $\phi: A \to B$ and $\psi: B \to A$ such that

$$\|\psi \circ \phi(a) - a\| < \varepsilon \quad \text{for all } a \in F.$$  

2. A real (resp. complex) $C^*$-algebra $A$ is nuclear if for all real (resp. complex) $C^*$-algebras $B$ the algebraic tensor product $A \otimes_R B$ (resp. $A \otimes_C B$) has a unique $C^*$-norm.

3. A real (resp. complex) $C^*$-algebra $A$ is exact if the tensor product functor $B \mapsto A \otimes_{\min} B$ is exact. Here the tensor product is over $\mathbb{R}$ (resp. $\mathbb{C}$) and $B$ can be any real (resp. complex) $C^*$-algebra.

Lemma 6.2. Let $A$ be a real $C^*$-algebra. Then

1. $A$ is amenable if and only if $A_C$ is amenable.

2. $A$ is nuclear if and only if $A_C$ is nuclear.

3. $A$ is exact if and only if $A_C$ is exact.

Consequently, $A$ is amenable if and only if it is nuclear; and in this case it is also exact.

Proof. Part (1) can be found in Proposition 3 of [25] and the preceding text. We claim that there is a one-to-one correspondence between $C^*$-norms on the algebraic tensor product $A \otimes_R B$ and those on $A_C \otimes_C B_C$. Let $\gamma$ be a $C^*$-norm on $A \otimes_R B$, and let $A \otimes \gamma B$ be the real $C^*$-algebra obtained by completion. Then the complexification $(A \otimes \gamma B)_C$ has a unique $C^*$-norm extending that on $A \otimes_R B$. Thus every $C^*$-norm on the algebraic tensor product $A \otimes_R B$ extends uniquely to a $C^*$-norm on $A_C \otimes_C B_C$. Part (2) follows immediately from this claim.
It also follows that the restriction of the minimal $C^*$-norm on $A^C \otimes \mathbb{C} B^C$ gives the minimal $C^*$-norm on $A \otimes \mathbb{R} B$. This fact, plus the fact that the complexification functor $A \mapsto A^C$ is exact, implies (3).

The final statement then follows from the corresponding statement for complex $C^*$-algebras. See Theorem 6.1.3 of [39] and Theorem 6.5.2 of [32].

**Proposition 6.3.** Let $A$ be a real separable $C^*$-algebra $A$. Then $A$ is exact if and only if there is an injective homomorphism $\iota: A \to \mathcal{O}_2^R$. If $A$ is unital then $\iota$ can be chosen to be unital.

**Proof.** Suppose that $A$ is exact. Then $A^C$ is separable and exact. Thus, by Theorem 6.3.11 of [39], there is an injective homomorphism $\iota_C: A^C \to \mathcal{O}_2$ (which is unital if $A^C$ is unital). Then we can take $\iota$ to be the composition

$$A \hookrightarrow A^C \overset{\iota_C}{\longrightarrow} \mathcal{O}_2 \hookrightarrow M_2(\mathcal{O}_2^R) \cong \mathcal{O}_2^R.$$ 

Conversely, if there is an injective homomorphism $\iota: A \to \mathcal{O}_2^R$ then the complexification yields an injective homomorphism from $A^C$ to $\mathcal{O}_2$. By Theorem 6.3.11 of [39] this implies that $A^C$ is exact, hence $A$ is exact. \qed

**Lemma 6.4.** Let $A$ be a real purely infinite c-simple nuclear unital $C^*$-algebra. Then all unital endomorphisms on $A \otimes \mathcal{O}_2^R$ are approximately unitarily equivalent.

**Proof.** In the complex case, this result is found as Theorem 6.3.8 of [39]. We will use that result to prove the real version.

By Corollary 5.3, Part (5) it suffices to show that any unital homomorphism

$$\gamma: A \otimes \mathcal{O}_2^R \otimes M_{2\infty} \to A \otimes \mathcal{O}_2^R \otimes M_{2\infty}$$

is approximately unitarily equivalent to the identity. We write $A' = A \otimes \mathcal{O}_2^R$ and let

$$\alpha_{\ell,k}: A' \otimes M_{2\ell} \hookrightarrow A' \otimes M_{2k} \quad \text{for} \quad k < \ell$$

$$\alpha_k: A' \otimes M_{2k} \hookrightarrow A' \otimes M_{2\infty}$$

be the canonical injections. Then we use the commutative diagram

$$\begin{array}{ccc}
A' \otimes M_{2\ell} & \overset{\alpha_{\ell+1,k}}{\longrightarrow} & A' \otimes M_{2k+1} \\
\downarrow^\epsilon & & \downarrow^\epsilon \\
A' \otimes M_{2\ell+1} & \overset{\alpha_{k+1,k}}{\longrightarrow} & A' \otimes M_{2k+2} \\
\downarrow^\epsilon & & \downarrow^\epsilon \\
A' \otimes M_{2\infty} & \overset{\epsilon}{\longrightarrow} & (A' \otimes M_{2\infty})_C \\
\end{array}$$

By Theorem 6.3.8 of [39], there is a sequence of unitaries $u_n \in (A' \otimes M_{2\infty})_C$ such that

$$\|u_n u_n^* - \gamma(a)\| \to 0 \quad \text{for all} \quad a \in A' \otimes M_{2\infty}. $$

For each $n$ find an integer $k(n)$ and a unitary $v_n \in (A' \otimes M_{2k(n)})_C$ such that $\|\alpha_{k(n)}(v_n) - u_n\| < 1/n$. Let $w_n = r(v_n) \in A' \otimes M_{2k(n)+1}$, where $r$ is induced by the realification map $M_{2k(n)} \otimes \mathbb{C} \to M_{2k(n)+1}$. We may assume that the sequence $\{k(n)\}_{n=1}^\infty$ is increasing. Let $a \in A' \otimes M_{2\infty}$ be given such that $\|a\| = 1$ and let $\varepsilon > 0$. Then find an integer $N$ large enough so that, for all $n \geq N$,
there exist $A'$ such that $A \cong A'$.

As in Section 7.2 of [39], let $A = A' \otimes M_{2k(n)}$ such that

\[
\Vert a - \alpha_k(n)(a)\Vert < \varepsilon \quad \text{and} \quad \Vert \gamma - \alpha_k(n)(b)\Vert < \varepsilon.
\]

Then a calculation shows that, for all $n \geq N$,

\[
\Vert v_n a_n v_n^* - b_n \Vert = \Vert \alpha_k(n)(v_n)\alpha_k(n)(a)\alpha_k(n)(v_n)^* - \alpha_k(n)(b)\Vert < 5\varepsilon.
\]

Now for any element $x \in A' \otimes M_{2k(n)}$, we have

\[
\alpha_{k(n)+1}(x) = \alpha_{k(n)+1,\alpha_{k(n)}(x)} = \alpha_{k(n)}(x).
\]

It follows that

\[
\Vert \alpha_{k(n)+1}(v_n)\alpha_{k(n)+1}(v_n)^* - \gamma(n)\Vert < \Vert \alpha_{k(n)+1}(v_n)\alpha_{k(n)}(a)\alpha_{k(n)+1}(v_n)^* - \alpha_{k(n)}(b)\Vert + 2\varepsilon
\]

\[
= \Vert v_n\alpha_{k(n)+1}(v_n)\alpha_{k(n)+1}(v_n)^* - \alpha_{k(n)}(b)\Vert + 2\varepsilon
\]

\[
= \Vert v_n\alpha_{k(n)+1}(v_n) - c(b)\Vert + 2\varepsilon
\]

\[
= \Vert v_n a_n v_n^* - b_n \Vert + 2\varepsilon < 7\varepsilon.
\]

\[\square\]

**Theorem 6.5.** Let $A$ be a real $C^*$-algebra. Then $A$ is c-simple, separable, unital, and nuclear if and only if $A \otimes O_2^\infty \cong O_2^\infty$.

**Proof.** Suppose that $A$ is c-simple, separable, unital, and nuclear. There is a unital homomorphism $\gamma: O_2^\infty \to A \otimes O_2^\infty$ given by $x \mapsto 1 \otimes x$ and there is a unital homomorphism $\kappa: A \otimes O_2^\infty \to O_2^\infty$ by Lemma 6.2 and Proposition 6.3. Then by Theorem 5.2 we have $\kappa \circ \gamma \approx u_1 \otimes O_2^\infty$ and by Lemma 6.4 we have $\gamma \circ \kappa \approx u_1 \otimes O_2^\infty$. Therefore, by (the real analog of) Corollary 2.3.4 of [39], $A \otimes O_2^\infty \cong O_2^\infty$.

Conversely, if the isomorphism $A \otimes O_2^\infty \cong O_2^\infty$ holds for a real $C^*$-algebra $A$, then we have $A \otimes O_2 \cong O_2$, which implies by Theorem 7.1.2 of [39] that $A_\infty$ is simple, separable, unital, and nuclear. Therefore $A$ is c-simple, separable, unital, and nuclear. \[\square\]

We note that the hypothesis above requiring that $A$ be c-simple cannot be relaxed, as the result does not hold for $A = O_2$ (considered as a real $C^*$-algebra).

**Theorem 6.6.**

1. Any two unital homomorphisms $\phi$, $\psi$ from $O_\infty^\infty$ into a real, unital, purely infinite, nuclear, c-simple $C^*$-algebra $A$ are approximately unitarily equivalent.

2. Let $A$ be a real c-simple, separable, and nuclear $C^*$-algebra. Then $A$ is isomorphic to $A \otimes O_2^\infty$ if and only if $A$ is purely infinite.

3. $O_\infty^\infty \cong \bigotimes_{n=1}^\infty O_\infty^\infty$.

**Proof.** As in Section 7.2 of [39]. \[\square\]
Corollary 6.7. Let $A$ and $B$ be real, $c$-simple, separable, nuclear $C^*$-algebras. If $A$ or $B$ is purely infinite, then $A \otimes B$ is purely infinite.

Proof. From part (2) of Theorem 6.6.

7. Homomorphisms from $O_R^\infty$

The goal of this section is to prove the following theorem, analogous to Proposition 2.2.7 of [35].

Theorem 7.1. Let $D$ be a real unital purely infinite simple $C^*$-algebra, and let $\phi, \psi : O_R^\infty \to D$ be unital homomorphisms. Then $\phi$ is asymptotically unitarily equivalent to $\psi$.

The proof of Theorem 7.1 will be the same as that in [35]. However, there are a couple of background topics that need to be addressed in the context of real $C^*$-algebras.

We begin with a discussion of approximately divisible real $C^*$-algebras, following [6]. It is sufficient to consider only separable unital $C^*$-algebras. Also, we skirt the general topic of completely noncommutative $C^*$-algebras by taking into account Definition 2.6 of [6] and the subsequent comment.

Definition 7.2. A separable unital real $C^*$-algebra $A$ is approximately divisible if for all $x_1, x_2, \ldots, x_n \in A$ and $\varepsilon > 0$, there is a unital subalgebra $B$ isomorphic to $M_2$, $M_3$, or $M_2 \oplus M_3$ such that $\|x_i y - y x_i\| < \varepsilon$ for all $i = 1, 2, \ldots, n$ and all $y$ in the unit ball of $B$.

The following theorem is the real version of Corollary 2.1.6 of [35].

Lemma 7.3. The tensor product $O_R^\infty \otimes D$ is approximately divisible for any real separable unital $C^*$-algebra $D$. In particular, every $c$-simple, separable, nuclear, purely infinite, unital real $C^*$-algebra is approximately divisible.

Proof. Let $A = O_R^\infty \otimes D$. Using the isomorphism $O_R^\infty \cong \bigotimes_{n=1}^\infty O_R^\infty$ of Theorem 6.6 we obtain a sequence of mutually commuting unital homomorphisms $\phi_n : O_R^\infty \to A$ such that $\|\phi_n(a)b - b\phi_n(a)\| \to 0$ for all $a \in O_R^\infty$ and all $b \in A$. Choose a unital map $\gamma : M_2 \oplus M_3 \to O_R^\infty$ and let $\psi_n = \phi_n \circ \gamma$. Then for large enough $n$, the subalgebra $B = \psi_n(M_2 \oplus M_3)$ works.

The second statement follows from part (2) of Theorem 6.6.

Lemma 7.4. Let $p$ and $q$ be full projections in $M_\infty(A)$ where $A$ is a real, separable, unital, approximately divisible $C^*$-algebra. Then $p \sim q$ if and only if $[p] = [q]$ in $K_0(A)$.

Proof. The proof is the same as the proof of (the first part of) Proposition 3.10 in [6] in complex case. That proof relies on a progression of results from Section 2 of [6] which can all be proven in the real case in the same way with one minor caveat. The proof of Proposition 2.1 of [6] (which in that paper was left to the reader) relies on the fact that a complex $C^*$-algebra is spanned by its unitaries. While this fact is not true in general for real $C^*$-algebras, it can

Documenta Mathematica 16 (2011) 619–655
easily be shown to be true for finite dimensional real C*-algebras, which is the relevant case.

The proof of Proposition 3.10 in [6] also relies on Theorem 3.1.4 of [3], which is a ring-theoretic result stated in enough generality to apply to real C*-algebras.

We remark that a more direct proof of Lemma 7.4 can be achieved in the special case (which is sufficient for our purposes) that \( A = \bigotimes_{i=1}^{\infty} \mathcal{O}_R \oplus D \) where \( D \) is separable and unital. In that case, we write \( A = \bigotimes_{i=1}^{\infty} \mathcal{O}_R \oplus D \) and let \( A_n = \bigotimes_{i=1}^{n} \mathcal{O}_R \oplus D \) be the unital subalgebra of \( A \) consisting of the first \( n \) factors in the tensor product. Then for each \( n \) and each \( k \), it is easy to find a unital subalgebra \( B_n \subset A_n \cap A \) that is isomorphic to \( M_{2k} \oplus M_{3k} \). Thus we achieve the result of Corollary 2.10 of [6] without having to recheck all the earlier material of Section 2 of [6] in the real case.

**Lemma 7.5.** Let \( D \) be a unital real C*-algebra and let \( p, q \) be any two full projections in \( K_R \otimes \mathcal{O}_R \oplus D \). Then \( p \) is Murray-von Neumann equivalent to a subprojection of \( q \). Furthermore, \( p \) is homotopic to \( q \) if and only if they represent the same class in \( K_0(K_R \otimes \mathcal{O}_R \oplus D) \approx K_0(D) \).

**Proof.** With our Lemmas 7.3 and 7.4, as well as Theorem 3.6 of [11], the proof is the same as that of Lemma 2.1.8 of [35].

**Proof of Theorem 7.1.** With these preliminary definitions and results, the proof is the same as the proof of Proposition 2.2.7 of [35] including all of the lemmas and intermediate results in Sections 2.1 and 2.2 of [35]. We note that in [35], the proofs of Propositions 2.1.9 and 2.1.10 (having to do with exact stability of the relations defining \( \mathcal{O}_R \) and \( \mathcal{E}_m(\delta) \)) are referred back to the proofs of parts (1) and (2) of Lemma 1.3 of [30]. The proof given there for part (2) produces isometries \( \omega_j \) that live in the real algebra \( \mathcal{E}_n(\delta) \). Therefore the homomorphisms \( \phi_{\delta}^{(m)} \) constructed in the complex case restrict to homomorphisms between the real algebras. The same will be true for the analogous proof of part (1).

We also note that the proofs for the real versions of Lemmas 2.2.1 and 2.2.3 of [35] rely on our Theorem 5.2 which is only established for \( n \) even. Hence for real C*-algebras, we need to take \( m \) to be even in Lemma 2.2.1 and \( n \) to be even in Lemma 2.2.3. This is however, sufficient for all subsequent arguments.

### 8. Asymptotic Morphisms

We appropriate the following definition of an asymptotic morphism from Section 25.1 of [4]. The other definitions in this section and the next are adapted from [35].

**Definition 8.1.** Let \( A \) and \( B \) be real C*-algebras. An asymptotic morphism \( \phi \) from \( A \) to \( B \) is a family \( \{ \phi_t \}_{t \in [0, \infty)} \) of maps \( \phi_t : A \rightarrow B \) such that

1. the map \( t \mapsto \phi_t(a) \) is continuous for each \( a \in A \), and
(2) for all \(a, b \in A\) and all \(\lambda \in \mathbb{R}\), the following functions vanish in norm as \(t \to \infty\):

(a) \(\phi_t(a + b) - \phi_t(a) - \phi_t(b)\),
(b) \(\phi_t(\lambda a) - \lambda \phi_t(a)\),
(c) \(\phi_t(ab) - \phi_t(a)\phi_t(b)\),
(d) \(\phi_t(a^*) - \phi_t(a)^*\).

We say that two asymptotic morphisms \(\phi_t\) and \(\psi_t\) from \(A\) to \(B\) are equivalent if \(\|\phi_t(a) - \psi_t(a)\|\) vanishes as \(t \to \infty\) for all \(a \in A\). We say that \(\phi_t\) and \(\psi_t\) are homotopic if there is an asymptotic morphism \(\Phi_t\) from \(A\) to \(C([0,1],B)\) such that \(\Phi_t(0) = \phi_t\) and \(\Phi_t(1) = \psi_t\) for all \(a \in A\). Equivalent asymptotic morphisms are homotopy equivalent (see Remark 25.1.2 of [4]).

We leave the easy proof of the next lemma to the reader.

**Lemma 8.2.** If \(A\) and \(B\) are real \(C^*\)-algebras and \(\phi\) is an asymptotic morphism from \(A\) to \(B\), then there is an asymptotic morphism \(\phi_C\colon A_\mathbb{C} \to B_\mathbb{C}\) defined by \((\phi_C)_t(a + ib) = \phi_t(a) + i\phi_t(b)\).

It can be proven, then, from the same result in the complex case, that for any asymptotic morphism \(\phi\) we have \(\limsup_{t \to \infty} \|\phi_t(a)\| \leq \|a\|\) for all \(a \in A\) (see Proposition 25.1.3 of [4]). Thus, an asymptotic morphism \(\{\phi_t\}\) gives rise to a unique homomorphism

\[\phi\colon A \to C_b([0, \infty), B)/C_0([0, \infty), B)\]

defined in the natural way; and every such homomorphism represents an asymptotic morphism, unique up to equivalence.

**Lemma 8.3.** Let \(A\) be separable and nuclear. Every asymptotic morphism from \(A\) to \(B\) is equivalent to one that is completely positive and contractive. Furthermore, if \(\phi\) and \(\psi\) are homopic completely positive and contractive asymptotic morphisms from \(A\) to \(B\), then in fact there is a homotopy from \(\phi\) to \(\psi\) consisting of completely positive and contractive asymptotic morphisms.

**Proof.** Let \(\phi\) be an asymptotic morphism from \(A\) to \(B\). Then by Proposition 1.1.5 of [35], the complexification \(\phi_C\) is equivalent to an asymptotic morphism \(\psi\) that is completely positive and contractive. The map \(\alpha\colon B_\mathbb{C} \to B\) defined by \(\alpha(a + ib) = a\) is completely positive and contractive. Then the restriction of \(\alpha \phi \psi\) to \(A\) is a completely positive, contractive asymptotic morphism from \(A\) to \(B\) and is equivalent to \(\phi\).

The same construction can be applied to a homotopy to prove the second statement. \(\square\)

**Definition 8.4.** Let \(\phi\) and \(\psi\) be asymptotic morphisms from \(A\) to \(K^\mathbb{R} \otimes D\). We define an asymptotic morphism \(\phi \oplus \psi\), also from \(A\) to \(K^\mathbb{R} \otimes D\), as follows. Choose an isomorphism \(\delta\colon M_2(K^\mathbb{R}) \to K^\mathbb{R}\) and define

\[(\phi \oplus \psi)_t(a) = (\delta \otimes 1_D) \begin{pmatrix} \phi_t(a) & 0 \\ 0 & \psi_t(a) \end{pmatrix}.\]
Lemma 8.5. The asymptotic morphism $\phi \oplus \psi$ is well defined up to unitary equivalence, as well as up to homotopy.

Proof. As in the complex case every automorphism of $K^R$ is implemented by a unitary in $U(B(H^R))$ (the proof in, for example, Lemma V.6.1 of [20] works in the real case). Furthermore, by [36], $U(B(H^R))$ is path connected. (In fact, by Theorem 3 of [29], it is contractible.) □

Definition 8.6. Let $\phi: A \to B$ be an asymptotic morphism of real C*-algebras and let $p \in A$ be a projection. A tail projection for $\phi(p)$ is a continuous path $p_t$ of projections for $t \in [0, \infty)$ such that $\lim_{t \to \infty} \|\phi_t(p) - p_t\| = 0$. We say that $\phi$ is full if there is a full projection $p \in A$ such that $\phi(p)$ has a full tail projection.

Definition 8.7. Let $A$ and $B$ be real C*-algebras. Two asymptotic morphisms $\phi$ and $\psi$ from $A$ and $B$ are asymptotically unitarily equivalent if there is a continuous family of unitary elements $u_t \in B$ such that $\lim_{t \to \infty} \|\phi_t(a)u_t^* - \psi_t(a)\| = 0$ for all $a \in A$.

With these definitions, all the results of Sections 1.2 and 1.3 of [35] hold for real C*-algebras.

Definition 8.8. Let $A$ and $D$ be real C*-algebras. An asymptotic morphism $\phi: A \to D$ has a standard factorization through $O^R_\infty \otimes A$ if there is an asymptotic morphism $\psi: O^R_\infty \otimes A \to D$ such that the asymptotic morphisms $\phi(a)$ and $\psi(1 \otimes a)$ (both from $A$ to $D$) are asymptotically unitarily equivalent.

Similarly, $\phi$ is asymptotically trivially factorizable if there is an asymptotic morphism $\psi: O^R_2 \otimes A \to D$ such that $\phi(a)$ and $\psi(1 \otimes a)$ are asymptotically unitarily equivalent.

Theorem 8.9 (Theorem 2.3.7 of [35]). Let $A$ be a separable, nuclear, unital, and $\epsilon$-simple. Let $D_0$ be a unital C*-algebra, and let $D = O^R_\infty \otimes D_0$. Then two full asymptotic morphisms from $A$ to $K^R \otimes D$ are asymptotically unitarily equivalent if and only they are homotopic.

Proof. The proof of Theorem 2.3.7 in [35] as well as the proofs of all of the preceding lemmas in Section 2.3 of [35] can be proven in the real case with the same proofs, with some extra attention paid to the issue of connectedness of unitary groups.

In a few places Phillips uses the fact that the unitary group of $O_2$ is connected. It is also true that $O^R_2$ is connected since $K_1(O^R_2) \cong 0$. However, on page 85 of [35], Phillips also uses the fact that the unitary group of a corner algebra of $O_\infty$ is connected. The corresponding statement in the real case is not true since $K_1(O^R_\infty) \cong Z_2$. We will show how to adjust the proof so that it works in the real case.

At this point in the proof we are (using Phillips’ notation) trying to find a path of partial isometries from $w_n + f_{n+2}$ to $\nu_{n+1} + w_{n+1}$ (these are partial isometries from $f_{n+1} + f_{n+2}$ to $f_{n+2} + c$). If the unitaries $(w_n + f_{n+2})^*(\nu_{n+1} + w_{n+1})$ and $f_{n+1} + f_{n+2}$ are not in the same connected component of $(f_{n+1} + c$
\[ f_{n+1}O^\infty f_{n+2}, \text{ then this can be changed by by multiplying } w_{n+1} \text{ on the right by a suitable unitary in } f_{n+2}O^\infty f_{n+2}. \]  
Thus by re-choosing the \( w_n \)'s inductively, we can be sure that there is an appropriate path of partial isometries at each step. \( \square \)

9. Groups of Asymptotic Morphisms

**Definition 9.1.** Let \( A \) be a real, separable, nuclear, unital, c-simple C*-algebra and let \( D \) be unital. We define \( E_A(D) \) to be the the set of homotopy classes of full asymptotic morphisms from \( A \) to \( K^R \otimes O^\infty \otimes D \). That is,

\[ E_A(D) = [[A, K^R \otimes O^\infty \otimes D]]_+. \]

More generally, for \( D \) unital or not, we define 

\[ \tilde{E}_A(D) = \ker (E_A(D^+) \to E_A(\mathbb{R})). \]

**Proposition 9.2.** Let \( A \) be real, separable, nuclear, unital, and c-simple. Then \( \tilde{E}_A(-) \) is a functor from the category of separable real C*-algebras with homotopy classes of asymptotic morphisms to abelian groups, that is homotopy invariant, stable, half exact, and split exact.

**Proof.** In the complex case, these results are proven in Section 3.1 of [35]. In the real case, they are proven the same way. Note that split exactness follows from homotopy invariance and half exactness by Corollary 3.5 of [12]. \( \square \)

**Lemma 9.3.** Let \( A \) and \( B \) be C*-algebras (real or complex). Let \( \phi: A \to B \) be an asymptotic morphism. If \( p, q \) are projections in \( A \) with \( p \leq q \), then there are tail projections \( p_t \) (for \( \phi(p) \)) and \( q_t \) (for \( \phi(q) \)) in \( B \) with \( p_t \leq q_t \) for all \( t \).

**Proof.** Let \( \tilde{p}_t \) and \( q_t \) be arbitrary tail projections corresponding to \( \phi(p) \) and \( \phi(q) \), respectively (these exist as in Remark 1.2.2 of [35]). One can easily show that

\[ \lim_{t \to \infty} \| \tilde{p}_t - q_t \tilde{p}_t q_t \| = 0. \]

For each \( t \), the element \( q_t \tilde{p}_t q_t \) is a self adjoint and asymptotically idempotent element of \( q_t B_{q_t} \). Therefore, there is a continuous path of projections \( p_t \in q_t B_{q_t} \) such that

\[ \lim_{t \to \infty} \| q_t \tilde{p}_t q_t - p_t \| = 0. \]

The tail projections \( p_t \) and \( q_t \) have the desired properties. \( \square \)

We note that if \( A \) and \( D \) are complex C*-algebras there are two groups one might consider: we let \( E^c_A(D) \) denote the functor of [35] that is based on complex asymptotic morphisms. On the other hand, according to the notation established in Definition 9.1, the asymptotic morphisms comprising \( E_A(D) \) are only required to be asymptotically linear over \( \mathbb{R} \) (thus the complex structures of \( A \) and \( D \) are forgotten). The following theorem relates the two groups.
Proposition 9.4. If $A$ is a real C*-algebra satisfying the hypotheses of Definition 9.1 and $D$ is a complex unital C*-algebra, then there is an isomorphism
\[ \tilde{E}_A(D) \cong \tilde{E}_{A_C}(D) \]
which is natural with respect to complex homomorphisms.

Proof. We show that for a real unital C*-algebra $A$ and a complex C*-algebra $B$, there is a bijection
\[ [[A,B]]_+ \cong [[A_C,B]]_+^C \]
of equivalence classes of full asymptotic morphisms.

Given a complex asymptotic morphism $\phi$ from $A_C$ to $B$, then we let $\Gamma(\phi)$ be the restriction of $\phi$ to $A$. If $\phi$ is full, then we claim that $\Gamma(\phi)$ is full. Since $\phi$ is full, there is a full projection $p \in A_C$ and a full tail projection $r_t \in B$ such that $\|\phi_t(p) - r_t\| \to 0$. Applying Lemma 9.3 to $p \leq 1$ we obtain tail projections $p_t$ and $q_t$ for $p$ and $1$, respectively, such that $p_t \leq q_t$ for all $t$. Since the tail projections $p_t$ and $r_t$ are asymptotically equal, it must be that $p_t$ are full projections. It follows that $q_t$ are also full projections; and since they are tail projections for the full projection $1_A$ in $A$, it follows that $\Gamma(\phi)$ is full.

Given a real asymptotic morphism $\psi$ from $A$ to $B$, then $\Delta(\psi)$ defines a complex asymptotic morphism from $A_C$ to $B$. Suppose that $\psi$ is full. Let $p$ be a full projection in $A$ and let $q_t \in B$ be a full tail projection for $\psi(p)$. Then clearly $p$ is full in $A_C$ and $q_t$ is a full tail projection for $\Delta(\psi(p))$. Hence $\Delta(\psi)$ is full.

It is immediate that $\Delta$ is a two-sided inverse for $\Gamma$. Furthermore, in the case that $B$ is stable, it is easy to see that $\Gamma$ preserves the semigroup operation of Definition 8.4. Therefore, under the hypotheses of the theorem, there is a group isomorphism $\tilde{E}_A(D) \cong \tilde{E}_{A_C}(D)$. \hfill \Box

Proposition 9.5. Let $A$ be a separable, nuclear, $c$-simple unital, real C*-algebra. Let $B$ be a separable real C*-algebra. Then there is a natural isomorphism $KK(A,B) \cong \tilde{E}_A(B)$.

The proof in the complex case takes place in Section 3.2 of [35]. Rather than reconstructing all of the arguments in the real case, we give a proof that uses results from [12] to reduce the real case to the complex case.

Proof of Proposition 9.5. Fix $A$ satisfying the hypotheses above. Let $e$ be a rank one projection in $K^R$ and let $\iota_A : A \to K^R \otimes O^R_C \otimes A$ be the homomorphism defined by $\iota_A(a) = e \otimes 1 \otimes a$. Let $[[\iota_A]]$ be the induced element of $\tilde{E}_A(A)$. Let $[1_A] \in KK(A,A)$ be the class of the identity. By Corollary 3.3 of [12], there is a unique natural transformation $\alpha$ from $KK(A, -)$ to $\tilde{E}_A(\cdot)$ such that $\alpha([1_A]) = [[\iota_A]]$. We will show that
\[ \alpha : KK(A,B) \to \tilde{E}_A(B) \]
is an isomorphism for all separable real C*-algebras $B$. By Theorem 3.9 of [12] it suffices to show that $\alpha$ is an isomorphism when $B$ is complex. In the complex case we have the element $[1_{A_C}] \in KK^C(A_C, A_C)$ and the homomorphism $(i_A)_C: A_C \rightarrow K^R \otimes O_\infty \otimes A_C \cong K \otimes O_\infty \otimes A_C$.

By Theorem 3.7 of [24] there is a unique natural transformation $\alpha_C$ from $KK^C(A_C, -)$ to $\tilde{E}_{A_C}^C(-)$ such that $\alpha_C([1_{A_C}]) = \llbracket \iota_{A_C} \rrbracket$. A special case of Theorem 3.2.6 of [35] shows that $\alpha_C$ is an isomorphism for all separable complex C*-algebras $B$.

Consider the following diagram for a complex C*-algebra $B$,

\[
\begin{array}{ccc}
KK^C(A_C, B) & \xrightarrow{\alpha_C} & \tilde{E}_{A_C}(B) \\
\downarrow{\nu} & & \downarrow{\mu} \\
KK(A, B) & \xrightarrow{\alpha} & \tilde{E}_{A}(B)
\end{array}
\]

where $\mu$ is the isomorphism of Proposition 9.4 above and $\nu$ is the isomorphism of Lemma 4.3 of [9]. To complete the proof, we only need to show that the diagram commutes. Since the homomorphism $\alpha_C$ is characterized by the value of $\alpha_C([1_{A_C}])$ it suffices to consider the case $B = A_C$ as in the diagram

\[
\begin{array}{ccc}
KK^C(A_C, A_C) & \xrightarrow{\alpha_C} & \tilde{E}_{A_C}(A_C) \\
\downarrow{\nu} & & \downarrow{\mu} \\
KK(A, A_C) & \xrightarrow{\alpha} & \tilde{E}_{A}(A_C)
\end{array}
\]

and to show that $\alpha_C([1_{A_C}]) = (\mu^{-1} \circ \alpha \circ \nu)([1_{A_C}])$ or, equivalently, $(\mu \circ \alpha_C)([1_{A_C}]) = (\alpha \circ \nu)([1_{A_C}])$.

From the construction of $\nu$ in the proof of Lemma 4.3 of [9] it is apparent that $\nu([1_{A_C}]) = [c_A] = (c_A)_*([1_A])$ where $c_A: A \rightarrow A_C$ is the real C*-algebra homomorphism induced by the unital inclusion $c: \mathbb{R} \hookrightarrow \mathbb{C}$. Thus

\[
(\alpha \circ \nu)([1_{A_C}]) = \alpha((c_A)_*([1_A])) = (c_A)_*(\alpha([1_A])) = (c_A)_*([c_A]) = [c_A].
\]

On the other hand, it is apparent from the construction of $\mu$ in the proof of Proposition 9.4 above that $\mu([c_A]) = [c_A]$. Thus

\[
(\mu \circ \alpha_C)([1_{A_C}]) = \mu(\alpha_C([1_{A_C}])) = \mu([c_A]) = [c_A].
\]

□

The following is the real version of Theorems 4.1.1 and 4.1.3 of [35].

**Theorem 9.6.** Let $A$ be a real separable unital nuclear c-simple C*-algebra and let $D$ be a separable unital C*-algebra. Then the following groups are naturally isomorphic, via the obvious maps.

1. $KK(A, D)$
(2) The set of asymptotic unitary equivalence classes of full homomorphisms from $A$ to $K^\mathbb{R} \otimes O^\mathbb{R}_\infty \otimes D$.

(3) The set of homotopy classes of full homomorphisms from $A$ to $K^\mathbb{R} \otimes O^\mathbb{R}_\infty \otimes D$.

(4) The set of asymptotic unitary equivalence classes of full homomorphisms from $K^\mathbb{R} \otimes O^\mathbb{R}_\infty \otimes A$ to $K^\mathbb{R} \otimes O^\mathbb{R}_\infty \otimes D$.

(5) The set of homotopy classes of full homomorphisms from $K^\mathbb{R} \otimes O^\mathbb{R}_\infty \otimes A$ to $K^\mathbb{R} \otimes O^\mathbb{R}_\infty \otimes D$.

**Proof.** The proof of the isomorphism of (1), (2), and (3) is the same as the proof of Theorem 4.1.1 in [35]. The proof of the isomorphism of (1), (4), and (5) relies on Lemma 4.1.2 of [35]. Once that lemma is established, the proof of the isomorphism of (1), (4), and (5) is the same as the proof of Theorem 4.1.3 of [35].

**Lemma 9.7.** Let $A$ be separable, nuclear, unital, and c-simple; let $D_0$ be separable and unital; and let $D = O^\mathbb{R}_\infty \otimes D_0$. Let $t \mapsto \phi_t$, for $t \in [0, \infty)$, be a continuous path of full homomorphisms from $K^\mathbb{R} \otimes A$ to $K^\mathbb{R} \otimes D$, and let $\psi: K^\mathbb{R} \otimes A \to K^\mathbb{R} \otimes D$ be a full homomorphism. Assume that $[\phi_0] = [\psi]$ in $KK_0(A, D)$. Then there is an asymptotic unitary equivalence from $\phi$ to $\psi$ that consists of unitaries in $U_0((K^\mathbb{R} \otimes D)^+)$. The proof will be essentially the same as the proof of Lemma 4.1.2 of [35]. However, that proof has an error in the third paragraph. The element $w_t$ introduced there does not seem to be a unitary as purported. Also, the order of the product in the definition of $z_t$ seems wrong. Fortunately, there is an easy fix and most of the proof can be left as it is. For clarity and completeness we present the entire proof, but the only significant difference is the unitary $w$ in the third paragraph and following. In places where the proof does not change (such as the entire first and second paragraphs, and most of the final paragraph), we use exactly the same language as in [35], except for the references to previous results in the present paper.

**Proof of Lemma 9.7.** Let $\{e_{ij}\}$ be a system of matrix units for $K^\mathbb{R}$. Identify $A$ with the subalgebra $e_{11} \otimes A$ of $K^\mathbb{R} \otimes A$. Define $\psi^{(0)}_t$ and $\psi^{(0)}$ to be the restrictions of $\phi_t$ and $\psi$ to $A$. Then $[\phi^{(0)}_0] = [\psi^{(0)}]$ in $KK_0(A, D)$. It follows from (the equivalence of (1) and (3) of) Theorem 9.6 that $\phi^{(0)}_0$ is homotopic to $\psi^{(0)}$. Therefore $\phi^{(0)}_0$ and $\psi^{(0)}$ are homotopic as asymptotic morphisms, and Theorem 8.9 provides an asymptotic unitary equivalence $t \mapsto u_t$ in $U(((K^\mathbb{R} \otimes D)^+)\text{f}\mbox{rom }\phi^{(0)}$ to $\psi^{(0)}$. Let $c \in U((K^\mathbb{R} \otimes D)^+)$ be a unitary with $c\psi^{(0)}(1) = \psi^{(0)}(1)c = \psi^{(0)}(1)$ and such that $c$ is homotopic to $u_0^{-1}$. Then $c$ commutes with every $\psi^{(0)}(a)$. Replacing $u_t$ by $cu_t$, we obtain an asymptotic unitary equivalence, which we again call $t \mapsto u_t$, from $\phi^{(0)}$ to $\psi^{(0)}$ which is in $U_0((K^\mathbb{R} \otimes D)^+)$. Define $\tau_{ij} = e_{ij} \otimes 1$. Then in particular $u_t\phi_t(\tau_{11})u_t^* \to \psi(\tau_{11})$ as $t \to \infty$. Therefore there is a continuous path $t \to z^{(1)}_t \in U_0((K^\mathbb{R} \otimes D)^+)$ such that $z^{(1)}_t \to$
1 and $z_t^{(1)} u_t \phi_t(\mathbf{1}_{11}) u_t^*(z_t^{(1)})^* = \psi(\mathbf{1}_{11})$ for all $t$. We still have $z_t^{(1)} u_t \phi_t(e_{11} \otimes a) u_t^*(z_t^{(1)})^* \rightarrow \psi(e_{11} \otimes a)$ for $a \in A$.

For convenience, set $f_{ij}^{(1)} = z_t^{(1)} u_t \phi_t(\mathbf{1}_{ij}) u_t^*(z_t^{(1)})^*$ and set $g_{ij} = \psi(\mathbf{1}_{ij})$. For each fixed $t$, the $f_{ij}^{(1)}$ are matrix units for $K_\mathbb{R}$ as are the $g_{ij}$. Also, we have $f_{11}^{(1)} = g_{11}$.

The projections $f_{11}^{(1)} + f_{22}^{(1)}$ and $g_{11} + g_{22}$ represent the same element of $K_\mathbb{R}(D)$ so (using Lemma 7.4) there is a continuous path of partial isometries $x_t^{(1)}$ in $K_\mathbb{R}(D)$ that $x_t^{(1)}(x_t^{(1)})^* = 1 - g_{11} - g_{22}$ and $(x_t^{(1)})^* x_t^{(1)} = 1 - f_{11}^{(1)} - f_{22}^{(1)}$. Set $w_t^{(1)} = g_{11} + g_{22} f_{11}^{(1)} + f_{22}^{(1)} \in \mathcal{U}((K_\mathbb{R}(D))^+)$. Then one checks that $w_t^{(1)} f_{ij}^{(1)}(w_t^{(1)})^* = g_{ij}$ for all $t$ and for $1 \leq i, j \leq 2$. Choose $c^{(1)} \in \mathcal{U}((K_\mathbb{R}(D))^+)$ with

$$c^{(1)}(g_{11} + g_{22}) = (g_{11} + g_{22}) c^{(1)} = g_{11} + g_{22}$$

and $c^{(1)} w_t^{(1)} \in \mathcal{U}(K_\mathbb{R}(D)^+)$.

Set $z^{(2)} = c^{(1)} w_t^{(1)}$ for $t \geq 1$ and extend $z^{(2)}$ continuously over $[0, 1]$ through unitaries so that $z_0^{(2)} = 1$, retaining the property that $z_t^{(2)}$, $g_{11} = g_{11} z_t^{(2)} = g_{11}$. This gives $z_t^{(2)} = 1$ for $t = 0$, $z_1^{(2)} g_{11} = g_{11} z_2^{(2)} = g_{11}$ for all $t$, and

$$z_t^{(2)}(z_t^{(1)})^* (z_t^{(2)})^* = \psi(\mathbf{1}_{ij})$$

for $t \geq 1$ and $1 \leq i, j \leq 2$.

Set $p^{(n)} = \sum_{k=1}^{m} g_{kk}$ for all positive integers $m$. For the induction step, assume that we have continuous paths unitaries $z^{(1)}_i, z^{(2)}_i, \ldots, z^{(n)}_i$ defined on $[0, \infty)$ such that

- $z^{(n)}_i(1) = 1$ for $0 \leq t \leq n - 2$,
- $z^{(n)}_i(p^{(n-1)}) = p^{(n-1)} z^{(n)}_i = p^{(n-1)}$ for all $t \geq 0$,
- $z^{(n-1)}_i u_t \phi_t(\mathbf{1}_{ij}) u_t^*(z^{(1)}_i)^* \cdots (z^{(n)}_i)^* = \psi(\mathbf{1}_{ij})$ for $t \geq n - 1$ and $1 \leq i, j \leq n$.

We must construct a $z^{(n+1)}_i$ with the corresponding properties. Initially, working with $t \in [n, \infty)$, set

$$f_{ij}^{(n)} = z_t^{(n)} \cdots z_1^{(1)} u_t \phi_t(\mathbf{1}_{ij}) u_t^*(z^{(1)}_i)^* \cdots (z^{(n)}_i)^*$$

and let $x_t^{(n)}$ be a continuous path of partial isometries such that $x_t^{(n)}(x_t^{(n)})^* = 1 - \sum_{k=1}^{m} g_{kk} = 1 - p^{(n)}$ and $(x_t^{(n)})^* x_t^{(n)} = 1 - \sum_{k=1}^{m} f_{kk}$. Set $w_t^{(n)} = p^{(n)} + g_{(n+1)1} f_{11}^{(n)} + f_{22}^{(n)}$. This continuous path of unitaries satisfies $w_t^{(n)} p^{(n)} = p^{(n)} w_t^{(n)} = p^{(n)}$ and $w_t^{(n)} f_{ij}^{(n)}(w_t^{(n)})^* = g_{ij}$ for all $t \geq n$ and all $1 \leq i, j \leq n + 1$.

As above, we can find a unitary $c^{(n)}$ such that $z_t^{(n+1)} = c^{(n)} w_t^{(n)}$ is in the connected component of the identity and $c^{(n)} p^{(n+1)} = p^{(n+1)} c^{(n)} = c^{(n)}$. Then extend $z_t^{(n+1)}$ so that it is defined for all $t \geq 0$ and $z_t^{(n+1)} = 1$ for $0 \leq t \leq n - 1$.

Check that this $z^{(n+1)}$ satisfies the corresponding properties listed above.

Now define

$$z_t = \left( \lim_{n \to \infty} z^{(n)}_t \cdots z^{(2)}_t z^{(1)}_t \right) u_t .$$
In a neighborhood of each \( t \), all but finitely many of the \( z_i^{(k)} \) are equal to 1, so this limit of products yields a continuous path of unitaries of \( U_0((K^R \otimes D)^+) \). Moreover, \( z_i \phi_t(\overline{e}_{ij}) z_i^* = \psi(\overline{e}_{ij}) \) whenever \( t \geq i, j \), so that \( \lim_{t \to \infty} z_i \phi_t(\overline{e}_{ij}) z_i^* = \psi(\overline{e}_{ij}) \) for all \( i \) and \( j \), while

\[
\lim_{t \to \infty} z_i \phi_t(e_{11} \otimes a) z_i^* = \lim_{t \to \infty} z_i^{(1)} u_t \phi_t(e_{11} \otimes a) u_t^* z_i^{(1)*} = \psi(e_{11} \otimes a)
\]

for all \( a \in A \). Since the \( \pi_{ij} \) and \( e_{11} \otimes a \) generate \( K^R \otimes A \), this shows that \( t \mapsto z_t \) is an asymptotic unitary equivalence.

\[\square\]

10. Classification of Real Kirchberg Algebras

We now present our main classification theorems for real Kirchberg algebras, analogous to the results of Section 4.2 of [35].

**Theorem 10.1.** Let \( A \) and \( B \) be unital separable nuclear purely infinite c-simple \( C^* \)-algebras.

1. Let \( \eta \) be an invertible element in \( KK(A, B) \). Then there is an isomorphism \( \phi : K^R \otimes A \to K^R \otimes B \) such that \([\phi] = \eta\).
2. Let \( \eta \) be an invertible element in \( KK(A, B) \) such that \([1_A] \times \eta = [1_B] \). Then there is an isomorphism \( \phi : A \to B \) such that \([\phi] = \eta\).

**Proof.** As in the proofs of Theorem 4.2.1 and Corollary 4.2.2 of [35]. \[\square\]

**Theorem 10.2.** Let \( A \) and \( B \) be unital separable nuclear purely infinite c-simple \( C^* \)-algebras that satisfy the universal coefficient theorem.

1. The stable \( C^* \)-algebras \( K^R \otimes A \) and \( K^R \otimes B \) are isomorphic if and only if \( K_{CR}^*(A) \) and \( K_{CR}^*(B) \) are isomorphic \( CRT \)-modules.
2. The unital \( C^* \)-algebras \( A \) and \( B \) are isomorphic if and only if the invariants \( (K_{CR}^*(A), [1_A]) \) and \( (K_{CR}^*(B), [1_B]) \) are isomorphic.
3. The stable \( C^* \)-algebras \( K^R \otimes A \) and \( K^R \otimes B \) are isomorphic if and only if \( K_{CR}^*(A) \) and \( K_{CR}^*(B) \) are isomorphic \( CR \)-modules.
4. The unital \( C^* \)-algebras \( A \) and \( B \) are isomorphic if and only if the invariants \( (K_{cr}^*(A), [1_A]) \) and \( (K_{cr}^*(B), [1_B]) \) are isomorphic.

**Proof.** Parts (1) and (2) are proven as in the proof of Theorem 4.2.4 of [35], using Proposition 2.3. Parts (3) and (4) then follow by Proposition 2.5. \[\square\]

**Corollary 10.3.**

1. The functor \( A \mapsto K_{CR}^*(A) \) is a bijection from isomorphism classes of real stable separable nuclear purely infinite c-simple \( C^* \)-algebras that satisfy the universal coefficient theorem to isomorphism classes of countable acyclic \( CRT \)-modules.
2. The functor \( A \mapsto (K_{CR}^*(A), [1_A]) \) is a bijection from isomorphism classes of real unital separable nuclear purely infinite c-simple \( C^* \)-algebras that satisfy the universal coefficient theorem to isomorphism classes of countable acyclic \( CRT \)-modules \( M \) with distinguished element \( m \in M_0^\Omega \).
Proof. Combine Theorem 10.2 above with Theorem 1 of [10]. □

Definition 10.4.

1. Let $A$ be a complex $C^*$-algebra. A real form of $A$ is a real $C^*$-algebra $B$ such that $B_C \cong A$.

2. Let $G_* = (G_0, G_1)$ be a pair of groups. A real form of $G_*$ is an acyclic $\text{CRT}$-module such that $M_*^U \cong G_*$.

3. Let $G_* = (G_0, G_1, g)$ be a pair of groups with a distinguished element $g \in G_0$. A real form of $G_*$ is a pair $(M, m)$ where $M$ is an acyclic $\text{CRT}$-module and $m$ is a distinguished element of $M_0$ such that $(M_0^U, M_1^U, c(m)) \cong (G_0, G_1, g)$.

Corollary 10.5. Let $A$ be a complex unital separable nuclear purely infinite simple $C^*$-algebra satisfying the universal coefficient theorem.

1. The functor $B \mapsto K^{\text{CRT}}(B)$ is a bijection from isomorphism classes of real forms of $K^R \otimes A$ to isomorphism classes of real forms of $K_* (A)$.

2. The functor $B \mapsto (K^{\text{CRT}}(B), [1_B])$ is a bijection from isomorphism classes of real forms of $A$ to isomorphism classes of real forms of $(K_* (A), [1_A])$.

Proof. If $B$ is a real form of $K^R \otimes A$, then $B$ is necessarily stable separable nuclear purely infinite and $c$-simple. Then $KU_*(B) = K_*(B_C) \cong K_*(A)$, so $K^{\text{CRT}}(B)$ is a real form of $K_* (A)$. Conversely, suppose $M$ is a real form of $K_* (A)$. Since $K_* (A)$ is countable, the exact sequences of Section 2.3 of [14] imply that $M$ is countable. Then by Corollary 10.3, $M \cong K^{\text{CRT}}(B)$ for some real stable separable nuclear purely infinite $c$-simple $C^*$-algebra satisfying the universal coefficient theorem. Since $K_* (B_C) \cong K_* (A)$, it follows from Theorem 4.2.4 of [35] that $B_C \cong A$ hence $B$ is a real form of $A$. Furthermore, Corollary 10.3 also implies that $B$ is unique up to isomorphism.

In the unital case, suppose that $B$ is a real form of $A$. As there is an isomorphism $B_C \cong A$ and the unit of $B_C$ is $1_B$, there is an isomorphism $\phi : KU_*(B) \to K_* (A)$ such that $\phi(c([1_B])) = [1_A]$. Thus $(K^{\text{CRT}}(B), [1_B])$ is a real form of $(K_* (A), [1_A])$. Conversely, if $(M, m)$ is a real form of $K_* (A)$, then let $B$ be a real unital separable nuclear purely infinite $c$-simple $C^*$-algebra such that $(K^{\text{CRT}}(B), [1_B]) \cong (M, m)$. Again, Theorem 4.2.4 of [35], implies that $B$ is a real form of $A$. □

11. REAL FORMS OF CUNTZ ALGEBRAS

In this section, we use Corollary 10.5 to give a complete description of all real forms of the complex Cuntz algebras $\mathcal{O}_n$ for $n \in \{2, \ldots, \infty\}$. The natural real form of $\mathcal{O}_n$ is the real Cuntz algebra $\mathcal{O}_n^R$, but we will find that there are others when $n$ is odd. For reference, we show in Table 1 the groups making up $K^{\text{CRT}}(\mathcal{O}_n^R)$. In the case of $n = \infty$ this arises from the isomorphism $K^{\text{CRT}}(\mathbb{R}) \cong K^{\text{CRT}}(\mathcal{O}_\infty^R)$ of Proposition 2.2; while for finite $n$, these $\text{CRT}$-modules were computed in Section 5.1 of [8].

DOCUMENTA MATHEMATICA 16 (2011) 619–655
Table 1

| 0  | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  |
|----|----|----|----|----|----|----|----|----|
| KO | Z  | Z  | Z  | Z  | 0  | Z  | 0  | Z  |
| KU | Z  | 0  | Z  | 0  | Z  | 0  | Z  | 0  |
| KT | Z  | Z  | Z  | Z  | Z  | 0  | Z  | Z  |

\( K_{\text{CRT}}(O_R) \) for \( n \) even

| 0  | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  |
|----|----|----|----|----|----|----|----|----|
| KO | \( \mathbb{Z}_{n-1} \) | \( \mathbb{Z}_2 \) | \( \mathbb{Z}_2 \) | \( \mathbb{Z}_{n-1} \) | 0  | 0  | 0  | \( \mathbb{Z}_{n-1} \) |
| KU | \( \mathbb{Z}_{n-1} \) | 0  | \( \mathbb{Z}_{n-1} \) | 0  | \( \mathbb{Z}_{n-1} \) | 0  | \( \mathbb{Z}_{n-1} \) | \( \mathbb{Z}_{n-1} \) |
| KT | \( \mathbb{Z}_{n-1} \) | 0  | \( \mathbb{Z}_{n-1} \) | \( \mathbb{Z}_{n-1} \) | \( \mathbb{Z}_2 \) | \( \mathbb{Z}_2 \) | \( \mathbb{Z}_{n-1} \) | \( \mathbb{Z}_{n-1} \) |

\( K_{\text{CRT}}(O_R) \) for \( n - 1 \equiv 2 \pmod{4} \)

| 0  | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  |
|----|----|----|----|----|----|----|----|----|
| KO | \( \mathbb{Z}_{n-1} \) | \( \mathbb{Z}_2 \) | \( \mathbb{Z}_2 \) | \( \mathbb{Z}_{n-1} \) | 0  | 0  | 0  | \( \mathbb{Z}_{n-1} \) |
| KU | \( \mathbb{Z}_{n-1} \) | 0  | \( \mathbb{Z}_{n-1} \) | 0  | \( \mathbb{Z}_{n-1} \) | 0  | \( \mathbb{Z}_{n-1} \) | \( \mathbb{Z}_{n-1} \) |
| KT | \( \mathbb{Z}_{n-1} \) | \( \mathbb{Z}_2 \) | \( \mathbb{Z}_{n-1} \) | \( \mathbb{Z}_{n-1} \) | \( \mathbb{Z}_2 \) | \( \mathbb{Z}_2 \) | \( \mathbb{Z}_{n-1} \) | \( \mathbb{Z}_{n-1} \) |

\( K_{\text{CRT}}(O_R) \) for \( n - 1 \equiv 0 \pmod{4} \)

| 0  | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  |
|----|----|----|----|----|----|----|----|----|
| KO | \( \mathbb{Z}_{n-1} \) | \( \mathbb{Z}_2 \) | \( \mathbb{Z}_2 \) | \( \mathbb{Z}_{n-1} \) | 0  | 0  | 0  | \( \mathbb{Z}_{n-1} \) |
| KU | \( \mathbb{Z}_{n-1} \) | 0  | \( \mathbb{Z}_{n-1} \) | 0  | \( \mathbb{Z}_{n-1} \) | 0  | \( \mathbb{Z}_{n-1} \) | \( \mathbb{Z}_{n-1} \) |
| KT | \( \mathbb{Z}_{n-1} \) | \( \mathbb{Z}_2 \) | \( \mathbb{Z}_{n-1} \) | \( \mathbb{Z}_{n-1} \) | \( \mathbb{Z}_2 \) | \( \mathbb{Z}_2 \) | \( \mathbb{Z}_{n-1} \) | \( \mathbb{Z}_{n-1} \) |

Theorem 11.1.  
(1) For \( n \) even or \( n = \infty \), there is up to isomorphism only one real form of \( O_n \): the real Cuntz algebra \( O_R^n \).

(2) For \( n \) odd, there are up to isomorphism two real forms of \( O_n \): the real Cuntz algebra \( O_R^n \) and an exotic real form \( E_n \).

Proof. First check that for odd integers \( n, n \geq 3 \), the groups and operations shown in Table 2 form an acyclic \( \text{CRT} \)-module. Using Corollary 10.3 (that is, Theorem 1 of [10]), let \( E_n \) be the unique real unital separable nuclear \( C^* \)-simple purely infinite \( C^* \)-algebra satisfying the universal coefficient theorem with united \( K \)-theory as shown in Table 2 and such that \([1_{E_n}]\) corresponds to a generator of the group in the real part in degree 0.

By Corollary 10.5, the problem of classifying real forms of \( O_n \) (for \( n \in \{2,3,\ldots,\infty\} \)) reduces to the algebraic problem of classifying real forms of \( (K_*(O_n),[1_{O_n}]) \). Suppose that \( (M,m) \) is such a real form. For \( n \) even (respectively \( n = \infty \)) we will show that \( (M,m) \) is isomorphic to \( (K_{\text{CRT}}(O_R^n),[1_{O_R^n}]) \) (respectively \( (K_{\text{CRT}}(O_R^\infty),[1_{O_R^\infty}]) \)). For \( n \) odd we will show that \( (M,m) \) is either isomorphic to \( (K_{\text{CRT}}(O_R^n),[1_{O_R^n}]) \) or to \( (K_{\text{CRT}}(E_n),[1_{E_n}]) \). Furthermore, by
Table 2. $K^\text{crit} (\mathcal{E}_n)$, for $n$ odd and $n \geq 3$.

|     | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-----|---|---|---|---|---|---|---|---|---|
| $KO_*$ | $\mathbb{Z}_{2(n-1)}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | 0 | $\mathbb{Z}_{(n-1)/2}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}_{2(n-1)}$ |
| $KU_*$ | $\mathbb{Z}_{n-1}$ | 0 | $\mathbb{Z}_{n-1}$ | 0 | $\mathbb{Z}_{n-1}$ | 0 | $\mathbb{Z}_{n-1}$ | 0 |
| $KT_*$ | $\mathbb{Z}_{n-1}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_{n-1}$ | $\mathbb{Z}_{n-1}$ | $\mathbb{Z}_2$ | $\mathbb{Z}_2$ | $\mathbb{Z}_{n-1}$ | $\mathbb{Z}_{n-1}$ |
| $c_*$ | 1 | 0 | 0 | 0 | 2 | 0 | $\frac{n-1}{2}$ | 0 | 1 |
| $r_*$ | 2 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 2 |
| $\varepsilon_*$ | 1 | 1 | 0 | 0 | 2 | 0 | 1 | $\frac{n-1}{2}$ | 1 |
| $\zeta_*$ | 1 | 0 | $\frac{n-1}{2}$ | 0 | 1 | 0 | $\frac{n-1}{2}$ | 1 | 0 |
| $(\psi_U)_*$ | 1 | 0 | $\frac{n-1}{2}$ | 0 | 1 | 0 | $\frac{n-1}{2}$ | 1 | 0 |
| $(\psi_R)_*$ | 1 | 1 | 1 | $\frac{n-1}{2}$ | 1 | 1 | 1 | $\frac{n-1}{2}$ | 1 |
| $\gamma_*$ | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| $\tau_*$ | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 2 | 1 |

Proposition 2.5 it suffices to restrict our attention to the $CR$-module consisting of the real and complex parts of $M$.

Since $(M, m)$ is a real form of $(K_*(O_n), [1_{O_n}])$ we know that $M^U_0 \cong \mathbb{Z}_{n-1}$ (respectively $M^U_0 \cong \mathbb{Z}$ when $n = \infty$), $M^U_1 = 0$, and $m \in M^O_0$. We further suppose that $c_0(m) \in M^U_0$ is a generator (corresponding to the class of the unit in $K_0(O_n)$).

We will compute the real part of $M$ (and the behavior of the operations $\eta_0, \xi, r, c, \psi_U$) using the long exact sequence

$$
\ldots \to M^O_{n-2} \overset{\eta_0}{\longrightarrow} M^O_{n-1} \overset{\xi}{\longrightarrow} M^U_{n-1} \overset{r\beta^{-1}}{\longrightarrow} M^O_{n-1} \to \ldots
$$

and the $CR$-relations described in Section 2.

Since $M^O_k = 0$ for $k$ odd it follows that $(\eta_0)_k$ is injective for $k$ odd and surjective for $k$ even. Furthermore, our hypothesis that $c_0$ generates $M^U_0$ implies that $c_0$ is surjective, which implies that $r_{-2} = 0$ and that $(\eta_0)_{-2}$ is injective. Thus $(\eta_0)_{-2} : M^O_{-2} \to M^O_1$ is an isomorphism and $\eta^O_3 : M^O_{-3} \to M^O_0$ is injective. Then the relations $\eta^O_3 = 0$ and $2\eta_0 = 0$ imply that $M^O_{-3} = 0$ and that $M^O_{-2}$ consists only of 2-torsion.

Suppose first that $M^O_{-2} \cong M^O_1 = 0$. Then using the long exact sequence above and the relation $rc = 2$, the rest of the groups of $M^O$ can be easily computed; except that in the case that $n$ is odd we encounter an extension problem wherein $M^O_n$ is either isomorphic to $\mathbb{Z}_4$ or to $\mathbb{Z}_2 \oplus \mathbb{Z}_2$. In that case, the same argument as in the computation of $K^\text{crit} (O^E_n)$ in Section 5.1 of [8] shows that $M^O_n \cong \mathbb{Z}_4$ exactly when $n-1 \equiv 0 \pmod{4}$ and $M^O_n \cong \mathbb{Z}_2 \oplus \mathbb{Z}_2$ exactly when $n-1 \equiv 2 \pmod{4}$. Thus we find that the real and complex parts of $M$ (as well as the operations $\eta_0, \xi, r, c, \psi_U$) are isomorphic to the real and complex parts of $K^\text{crit} (O^E_n)$ (respectively $K^\text{crit} (O^E_n)$).
For the remaining case, suppose that \( M_{-3} \neq 0 \). Since this leads to the exotic \( \text{CRT} \)-module \( K^{\text{crt}}(\mathcal{E}_n) \), we will include all of the details of this computation. Since \( c_0 \) is surjective, the relation \( \psi_v c = c \) implies that \( (\psi_v)_0 = 1 \). Then the relation \( \beta_c \psi_v = -\psi_v \beta_c \) implies that \( \psi_v = 1 \) in degrees congruent to 0 (mod 4) and \( \psi_v = -1 \) in degrees congruent to 2 (mod 4).

From \( M_{-3} \neq 0 \) it follows that \( c_{-2} \) is injective. But the only non-trivial 2-torsion subgroup of \( M_{-2} \) is isomorphic to \( \mathbb{Z}_2 \), and that occurs only when \( n \) is finite and odd. Thus \( M_{-2} \cong M_{-2} \cong \mathbb{Z}_2 \) and the complexification map \( c_{-2} : \mathbb{Z}_2 \rightarrow \mathbb{Z}_{n-1} \) is multiplication by \((n-1)/2\) (in terms of chosen generators).

The map \( r_{-4} \) is surjective and has kernel equal to \((n-1)/2\) \( \mathbb{Z}_{n-1} \cong \mathbb{Z}_2 \) so \( M_{-4} \cong \mathbb{Z}_{(n-1)/2} \). The relation \( c_{-4} r_{-4} = 1 + (\psi_v)^{-1} \) implies that the map \( c_{-4} : \mathbb{Z}_{(n-1)/2} \rightarrow \mathbb{Z}_{n-1} \) is multiplication by 2.

Continuing to work our way down, the fact that \( c_{-4} \) is injective implies that \( M_{-8} = 0 \). The fact that the image of \( c_{-4} \) is \( \mathbb{Z}_{n-1} \) implies that \( M_{-6} \cong \mathbb{Z}_2 \) and \( r_{-6} \) is surjective. The relation \( c_{-6} r_{-6} = 1 + (\psi_v)^{-6} \) implies that \( c_{-6} = 0 \) from which we see that \( \eta_{-7} \) is an isomorphism. Thus \( M_{-2} \cong \mathbb{Z}_2 \).

Finally, we compute \( M_{-8} \cong M_0 \). The exact sequence indicates that it is an extension of \( \mathbb{Z}_2 \) by \( \mathbb{Z}_{n-1} \). We will prove that it is isomorphic to \( \mathbb{Z}_{2(n-1)} \). If not, then \( M_0 \cong \mathbb{Z}_2 \oplus \mathbb{Z}_{n-1} \) and we can arrange the direct sum decomposition so that \( \eta_{-1} = (\beta_0 \ 0) \) and \( c_0 = (0 \ 1) \). Then the relation \( r_c = 2 \) implies that \( r_0 = (\frac{1}{2}) \). But then there is no isomorphism from \( M_0 / \text{image}(r_0) \cong \mathbb{Z}_2 \oplus \mathbb{Z}_2 \) to \( M_0 \cong \mathbb{Z}_2 \) as required by the long exact sequence.

Thus in the case that \( M_{-2} \neq 0 \) it must be that \( n \) is odd and it must be that the real and complex parts of \( M \) are isomorphic to the real and complex parts of \( K^{\text{crt}}(\mathcal{E}_n) \) as in Table 2, completing the proof. \( \square \)

We remark that the above result can instead be obtained using the analysis of acyclic \( \text{CRT} \)-modules in [23]. Indeed, let \( M \) be an acyclic \( \text{CRT} \)-module such that \( M_{-4} \) is isomorphic to \( \mathbb{Z}_{k-1} \) or \( \mathbb{Z} \), \( M_0 = 0 \), and \( c_0 : M_0 \rightarrow M_0 \) is surjective (hence \( (\psi_v)_0 = 1 \)). By Lemma 8.3.1, Proposition 8.3.2, and Theorem 8.3.3 of [23], there are isomorphisms

\[
h_k(M) := \ker(1 - (\psi_v)_k) / \text{image}(1 + (\psi_v)_k) \cong \eta_0 M_{k}^{O} \oplus \eta_0 M_{k+4}^{O}
\]

and, furthermore, \( M \) is determined up to isomorphism by \( M_U \), \( \psi_v \), and the resulting decompositions of \( h_k(M) \) for \( k = 0 \) and \( k = 2 \). Using \((\psi_v)_0 = 1 \) and \((\psi_v)_2 = -1 \), we obtain

\[
(h_0(M), h_2(M)) = \begin{cases} 
(\mathbb{Z}_2, 0) & \text{if } M_0^U = \mathbb{Z} \\
(0, 0) & \text{if } M_0^U = \mathbb{Z}_{n-1} \text{ with } n \text{ even} \\
(\mathbb{Z}_2, \mathbb{Z}_2) & \text{if } M_0^U = \mathbb{Z}_{n-1} \text{ with } n \text{ odd}
\end{cases}
\]

The resulting possibilities for \( M \) are realized by the united \( K \)-theory of \( \mathcal{O}_\infty^R \) and \( \mathbb{H} \otimes \mathcal{O}_\infty^R \) in the first case; by that of \( \mathcal{O}_n^R \) in the second case; and by that of \( \mathcal{O}_n^R \), \( \mathbb{H} \otimes \mathcal{O}_n^R \), \( \mathcal{E}_n \), and \( \mathbb{H} \otimes \mathcal{E}_n \) in the third case. The assumption that \( c_0 \) is surjective reduces the possibilities to the united \( K \)-theory of \( \mathcal{O}_\infty^R \), \( \mathcal{O}_n^R \), or \( \mathcal{E}_n \).
References

1. D. W. Anderson, *A new cohomology theory*, Ph.D. thesis, University of California at Berkeley, 1964.
2. B. V. Rajarama Bhat, George A. Elliott, and Peter A. Fillmore (eds.), *Lectures on operator theory*, Fields Institute Monographs, vol. 13, American Mathematical Society, Providence, RI, 1999.
3. Bruce Blackadar, *Rational C*-algebras and nonstable K-theory*, Proceedings of the Seventh Great Plains Operator Theory Seminar (Lawrence, KS, 1987), vol. 20, 1990, pp. 285–316.
4. ———, *K-theory for operator algebras*, second ed., Mathematical Sciences Research Institute Publications, vol. 5, Cambridge University Press, Cambridge, 1998.
5. Bruce Blackadar and David Handelman, *Dimension functions and traces on C*-algebras*, J. Funct. Anal. 45 (1982), no. 3, 297–340.
6. Bruce Blackadar, Alexander Kumjian, and Mikael Rørdam, *Approximately central matrix units and the structure of noncommutative tori*, K-Theory 6 (1992), no. 3, 267–284.
7. Bruce E. Blackadar and Joachim Cuntz, *The structure of stable algebraically simple C*-algebras*, Amer. J. Math. 104 (1982), no. 4, 813–822.
8. Jeffrey L. Boersema, *Real C*-algebras, united K-theory, and the K"unneth formula*, K-Theory 26 (2002), no. 4, 345–402.
9. ———, *Real C*-algebras, united KK-theory, and the universal coefficient theorem*, K-Theory 33 (2004), no. 2, 107–149.
10. ———, *The range of united K-theory*, J. Funct. Anal. 235 (2006), no. 2, 701–718.
11. Jeffrey L. Boersema and Efren Ruiz, *Stability of real C*-algebras*, Canad. Math. Bull., to appear.
12. ———, *Axiomatic K-theory for real C*-algebras*, arXiv.org arXiv:0909.0972v1 (2009).
13. Jeffrey L. Boersema and Peter J. Stacey, *Correction to the paper: “Real structure in purely infinite C*-algebras”*, J. Operator Theory 53 (2005), no. 2, 441–442.
14. A. K. Bousfield, *A classification of K-local spectra*, J. Pure Appl. Algebra 66 (1990), no. 2, 121–163.
15. Ola Bratteli, ErlingStormer, AkitakaKishimoto, and MikaelRørdam, *The crossed product of a UHF algebra by a shift*, Ergodic Theory Dynam. Systems 13 (1993), no. 4, 615–626.
16. Lawrence G. Brown, *Stable isomorphism of hereditary subalgebras of C*-algebras*, Pacific J. Math. 71 (1977), no. 2, 335–348.
17. Joachim Cuntz, *Simple C*-algebras generated by isometries*, Comm. Math. Phys. 57 (1977), no. 2, 173–185.
18. ———, *The structure of multiplication and addition in simple C*-algebras*, Math. Scand. 40 (1977), no. 2, 215–233.
19. Jeffrey L. Boersema, Efren Ruiz, P. J. Stacey, *K*-theory for certain $C^*$-algebras, Ann. of Math. (2) 113 (1981), no. 1, 181–197.
20. Kenneth R. Davidson, *C*-algebras by example, Fields Institute Monographs, vol. 6, American Mathematical Society, Providence, RI, 1996.
21. Thierry Giordano, A classification of approximately finite real *C*-algebras, J. Reine Angew. Math. 385 (1988), 161–194.
22. Paul Halmos, *Finite dimensional vector spaces*, second ed., Van Nostrand, 1958.
23. Beatrice Hewitt, On the homotopical classification of $KO$-module spectra, Ph.D. thesis, University of Illinois at Chicago, 1996.
24. Nigel Higson, A characterization of KK-theory, Pacific J. Math. 126 (1987), no. 2, 253–276.
25. Nhan Bao Ho, Amenability for real $C^*$-algebras, Bull. Aust. Math. Soc. 77 (2008), no. 3, 509–514.
26. Eberhard Kirchberg, the classification purely infinite of $C^*$-algebras using kasparov’s theory, preprint (1994).
27. Eberhard Kirchberg and Mikael Rørdam, Non-simple purely infinite $C^*$-algebras, Amer. J. Math. 122 (2000), no. 3, 637–666.
28. Hua Xin Lin and N. Christopher Phillips, Approximate unitary equivalence of homomorphisms from $O_\infty$, J. Reine Angew. Math. 464 (1995), 173–186.
29. Hua Xin Lin and Shuang Zhang, On infinite simple $C^*$-algebras, J. Funct. Anal. 100 (1991), no. 1, 221–231.
30. N. Christopher Phillips, Approximation by unitaries with finite spectrum in purely infinite $C^*$-algebras, J. Funct. Anal. 120 (1994), no. 1, 98–106.
31. Mikael Rørdam, Classification of inductive limits of Cuntz algebras, J. Reine Angew. Math. 440 (1993), 175–200.
40. Jonathan Rosenberg and Claude Schochet, \textit{The K"unneth theorem and the universal coefficient theorem for Kasparov's generalized $K$-functor}, Duke Math. J. 55 (1987), no. 2, 431–474.
41. Claude Schochet, \textit{Topological methods for $C^\ast$-algebras. II. Geometric resolutions and the K"unneth formula}, Pacific J. Math. 98 (1982), no. 2, 443–458.
42. Peter J. Stacey, \textit{Real structure in direct limits of finite-dimensional $C^\ast$-algebras}, J. London Math. Soc. (2) 35 (1987), no. 2, 339–352.
43. \textit{, Real structure in purely infinite $C^\ast$-algebras}, J. Operator Theory 49 (2003), no. 1, 77–84.

Jeffrey L. Boersema  
Department of Mathematics  
Seattle University  
Seattle, WA  
USA  
boersema@seattlen.edu

Efren Ruiz  
Department of Mathematics  
University of Hawaii  
Hilo, HI  
USA  
ruize@hawaii.edu

P. J. Stacey  
Department of Mathematics  
LaTrobe University  
Melbourne, VIC  
Australia  
P.Stacey@latrobe.edu.au
A SIMPLE CRITERION FOR EXTENDING
NATURAL TRANSFORMATIONS TO HIGHER $K$-THEORY

Gonçalo Tabuada

Received: January 26, 2011
Communicated by Max Karoubi

Abstract. In this article we introduce a very simple and widely applicable criterion for extending natural transformations to higher $K$-theory. More precisely, we prove that every natural transformation defined on the Grothendieck group and with values in an additive theory admits a unique extension to higher $K$-theory. As an application, the higher trace maps and the higher Chern characters originally constructed by Dennis and Karoubi, respectively, can be obtained in an elegant, unified, and conceptual way from our general results.

2010 Mathematics Subject Classification: 19D55, 18G55
Keywords and Phrases: Higher $K$-theory, higher trace maps, higher Chern characters, non-commutative motives

Introduction
In his foundational work, Grothendieck [10] introduced a very simple and elegant construction $K_0$, the Grothendieck group, in order to formulate a far-reaching generalization of the Riemann-Roch theorem. Since then, this versatile construction spawned well-beyond the realm of algebraic geometry to become one of the most important (working) tools in mathematics. Latter, through revolutionary topological techniques, Quillen [23] extended the Grothendieck group to a whole family of higher $K$-theory groups $K_n, n \geq 0$. However, in contrast with $K_0$, these higher $K$-theory groups are rather mysterious and their computation is often out of reach. In order to capture some of its flavour, Connes, Dennis, Karoubi, and others, constructed natural transformations towards simpler theories $E$ making use of a variety of highly involved techniques; see [6, 7, 15]. Typically, the construction of a natural transformation $K_0 \Rightarrow E_0$ is very simple, while its extension $K_n \Rightarrow E_n$ to higher $K$-theory is a real “tour-de-force". For example, the trace map $K_0 \Rightarrow HH_0$ consists simply in taking the trace of an idempotent, while its extension
$K_n \Rightarrow HH_n$ makes use of an array of tools (Hurewicz maps, group homology, assembly maps, etc) coming from topology, algebra, representation theory, etc.

These phenomena motivate the following general questions:

**Questions:** Given a natural transformation $K_0 \Rightarrow E_0$, is it possible to extend it to higher $K$-theory $K_n \Rightarrow E_n$? If so, is such an extension unique?

In this article we prove that if $E$ verifies three very simple conditions, not only such an extension exists, but it is moreover unique. The precisely formulation of our results makes use of the language of Grothendieck derivators, a formalism which allows us to state and prove precise universal properties; see Appendix A.

1. **Statement of results**

A differential graded (=dg) category, over a fixed commutative base ring $k$, is a category enriched over cochain complexes of $k$-modules (morphisms sets are such complexes) in such a way that composition fulfills the Leibniz rule: $d(f \circ g) = (df) \circ g + (-1)^{\deg(f)} f \circ (dg)$. Dg categories extend the classical notion of (dg) $k$-algebra and solve many of the technical problems inherent to triangulated categories; see Keller’s ICM address [16]. In non-commutative algebraic geometry in the sense of Bondal, Drinfeld, Kaledin, Kontsevich, Van den Bergh, and others, they are considered as differential graded enhancements of (bounded) derived categories of quasi-coherent sheaves on a hypothetical non-commutative space; see [1, 8, 9, 14, 17, 18].

Let $E: dgcat \rightarrow Spt$ be a functor, defined on the category of dg categories, and with values in the category of spectra [2]. We say that $E$ is an additive functor if it verifies the following three conditions:

(i) filtered colimits of dg categories are mapped to filtered colimits of spectra;

(ii) derived Morita equivalences (i.e. dg functors which induce an equivalence on the associated derived categories; see [16, §4.6]) are mapped to weak equivalences of spectra;

(iii) split exact sequences (i.e. sequences of dg categories which become split exact after passage to the associated derived categories; see [24, §13]) are mapped to direct sums

$$0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0 \mapsto E(A) \oplus E(C) \simeq E(B)$$

in the homotopy category of spectra.

Examples of additive functors include Hochschild homology ($HH$), cyclic homology ($HC$), and algebraic $K$-theory ($K$); see [16, §5]. Recall from [25] that the category $dgcat$ carries a Quillen model structure whose weak equivalences are the derived Morita equivalences. Given an additive functor $E$, we obtain then an induced morphism of derivators $E: HO(dgcat) \rightarrow HO(Spt)$. Associated to $E$, we have also the composed functors

$$E_n: dgcat \xrightarrow{E} Spt \xrightarrow{\pi^n} Ab \quad n \geq 0,$$
where $\pi_n^s$ denotes the $n^{th}$ stable homotopy group functor and $\mathsf{Ab}$ the category of abelian groups. Our answer to the questions stated in the Introduction is:

**Theorem 1.1.** For any additive functor $E$, the natural map

$$\text{Nat}(K,E) \cong \text{Nat}(K_0,E_0)$$

is bijective. In particular, every natural transformation $\phi : K \Rightarrow E_0$ admits a canonical extension $\phi_n : K_n \Rightarrow E_n$ to all higher $K$-theory groups.

Intuitively speaking, Theorem 1.1 show us that all the information concerning a natural transformation is encoded on the Grothendieck group. Its proof relies in an essential way on the theory of non-commutative motives, a subject envisioned by Kontsevich [17, 19] and whose development was initiated in [3, 4, 24, 25, 27, 28]. In the next section we illustrate the potential of this general result by explaining how the highly involved constructions of Dennis and Karoubi can be obtained as simple instantiations of the above theorem. Due to its generality and simplicity, we believe that Theorem 1.1 will soon be part of the toolkit of any mathematician whose research comes across the above conditions (i)-(iii).

2. Applications

2.1. Higher trace maps. Recall from [16, §5.3] the construction of the Hochschild homology complex $\text{HH}(A)$ of a dg category $A$. This construction is functorial in $A$ and so by promoting it to spectra we obtain a well-defined functor

$$\text{HH} : \text{dgcat} \longrightarrow \text{Spt}.$$  

As explained in loc. cit., this functor verifies conditions (i)-(iii) and hence it is additive. Now, given a $k$-algebra $A$, recall from [20, Example 8.3.6] the construction of the classical trace map

$$K_0(A) \rightarrow \text{HH}_0(A) = A/[A,A].$$

Roughly, it is the map induced by sending an idempotent matrix to the image of its trace (i.e. the sum of the diagonal entries) in the quotient $A/[A,A]$. This construction extends naturally from $k$-algebras to dg categories (see [20]) giving rise to a natural transformation

$$K_0 \Rightarrow \text{HH}_0.$$  

**Proposition 2.3.** In Theorem 1.1 let $E$ be the additive functor (2.1) and let $\phi$ be the natural transformation (2.2). Then, for every $k$-algebra $A$, the canonical extension $\phi_n : K_n(A) \rightarrow \text{HH}_n(A)$ of $\phi$ agrees with the $n^{th}$ trace map constructed originally by Dennis (see [20, §8.4 and §11.4]).
2.2. Higher Chern characters. Recall also from [16, §5.3] the construction of the cyclic homology complex $HC(A)$ of a dg category $A$. By promoting this construction to spectra we obtain a functor

$$HC : \text{dgcat} \rightarrow \text{Spt}$$

which verifies conditions (i)-(iii). Given a $k$-algebra $A$, recall from [20, Theorem 8.3.4] the construction of the Chern characters

$$ch_{0,i} : K_0(A) \rightarrow HC_{2i}(A) \quad i \geq 0.$$ Morally, these are the non-commutative analogues of the classical Chern character with values in even dimensional de Rham cohomology. As shown in [26] this construction extends naturally from $k$-algebras to dg categories giving rise to natural transformations $K_0 \Rightarrow HC_{2i} \quad i \geq 0$.

**Proposition 2.5.** In Theorem 1.1 let $E$ be the additive functor $\Omega^{2i}HC$ (obtained by composing (2.4) with the $(2i)^{th}$ fold looping functor on Spt) and let $\phi$ be the natural transformation $K_0 \Rightarrow (\Omega^{2i}HC)_0 = HC_{2i}$. Then, for every $k$-algebra $A$, the canonical extension $\phi_n : K_n(A) \rightarrow (\Omega^{2i}HC)_n(A) = HC_{n+2i}(A)$ of $\phi$ agrees with the higher Chern character $ch_{n,i}$ constructed originally by Karoubi (see [15, §2.27-2.36]).

3. Proof of Theorem 1.1

We start by describing the natural map (1.2). As mentioned in §1, the category $\text{dgcat}$ carries a Quillen model structure whose weak equivalences are the derived Morita equivalences; see [25, Theorem 5.3]. Let us write $\text{Hmo}$ for the associated homotopy category and $l : \text{dgcat} \rightarrow \text{Hmo}$ for the localization functor. According to our notation the map (1.2) sends a natural transformation $\Phi \in \text{Nat}(K, E)$ to the natural transformation $\pi_0 \circ \Phi(e) \circ l \in \text{Nat}(K_0, E_0)$. Pictorially, we have:

$$\xymatrix{ \text{dgcat} \ar[r]^l & \text{Hmo} \ar[rr]^{\Phi(e)} & & \text{Ho}(\text{Spt}) \ar[r]^{\pi_0} & \text{Ab}.}$$

The functors $K, E : \text{dgcat} \rightarrow \text{Spt}$ are additive and so the following diagrams

$$\xymatrix{ \text{dgcat} \ar[r]^K \ar[d] & \text{Spt} \ar[d] \ar[dl]_{\text{Hmo}} \ar[r]_{\text{Ho}(\text{Spt})} & \text{dgcat} \ar[r]^E \ar[d] & \text{Spt} \ar[d] \ar[dl]_{\text{Hmo}} \ar[r]_{\text{Ho}(\text{Spt})} }$$

are commutative. Moreover, the $0^{th}$ stable homotopy group functor $\pi_0$ descends to the homotopy category $\text{Ho}(\text{Spt})$. These facts show us that the composed horizontal functors in the above diagram (3.1) are in fact $K_0$ and $E_0$.

We now study the set $\text{Nat}(K_0, E_0)$. Recall from [16, §5.1] the notion of additive invariant. Intuitively, it consists of a functor defined on $\text{dgcat}$ and with values
in an additive category which verifies conditions similar to (ii)-(iii). Since by hypothesis $E$ is additive, the composed functor

$$E_0 : \text{dgcat} \xrightarrow{l} \text{Hmo} \xrightarrow{\pi_0^s} \text{Ho}(\text{Spt})$$

is an additive invariant. Hence, as proved in [26, Proposition 4.1], we have the following natural bijection

$$\text{Nat}(K_0, E_0) \xrightarrow{\sim} E_0([k]) \eta \mapsto \eta([k]).$$

(3.2)

Some explanations are in order: $k$ denotes the dg category naturally associated to the base ring $k$, i.e. the dg category with only one object and with $k$ as the dg algebra of endomorphisms (concentrated in degree zero); the symbol $[k]$ stands for the class of $k$ (as a module over itself) in the Grothendieck group $K_0(k) = K_0(k)$.

Let us now turn our attention to $\text{Nat}(K, E)$. Recall from [24, §15] the notion of additive invariant of dg categories. Roughly speaking, it consists of a morphism of derivators defined on $\text{HO}(\text{dgcat})$ and with values in a triangulated derivator which verifies conditions analogous to (i)-(iii). Since the functor $E$ is additive, the induced morphism of derivators

$$E : \text{HO}(\text{dgcat}) \longrightarrow \text{HO}(\text{Spt})$$

is an additive invariant of dg categories. Following [3, Theorem 8.1] we have then a natural bijection

$$\text{Nat}(K, E) \xrightarrow{\sim} \pi_0^s E([k]) = E_0([k]).$$

(3.3)

A careful inspection of the proof of [3, Theorem 8.1] show us that (3.3) sends a natural transformation $\Phi \in \text{Nat}(K, E)$ to the element $\pi_0^s(\Phi([k]))$ of the abelian group $E_0([k])$. Note that this element is simply the image of $[k]$ by the abelian group homomorphism

$$K_0([k]) = \pi_0^s(\text{K}(e)([k])) \xrightarrow{\pi_0^s(\Phi([k]))} \pi_0^s(E(e)([k])) = E_0([k]).$$

We now prove that the following diagram

$$(3.4) \quad \text{Nat}(K, E) \xrightarrow{(1.2)} \text{Nat}(K_0, E_0) \xrightarrow{(3.2)} E_0([k])$$

commutes. Let $\Phi \in \text{Nat}(K, E)$. On the one hand, we observe that the composed map $(3.2) \circ (1.2)$ sends $\Phi$ to the element $(\pi_0^s \circ \Phi(e) \circ l)([k])$ of the abelian group $E_0([k])$. On the other hand, the following equalities hold:

$$(\Phi(e) \circ l)([k]) = \Phi(e)([k]) \quad (\pi_0^s \circ \Phi(e))([k]) = \pi_0^s(\Phi(e)([k])).$$

2In [3] this bijection was established for a localizing invariant $E$. However, the arguments in the additive case are completely similar.
Therefore, we have
\[ (\pi_s^0 \circ \Phi(e) \circ l)([k]) = \pi_s^0(\Phi(e)(k))([k]). \]

Finally, since the right-hand side in this latter equality coincides with the image of \( \Phi \) by the map (3.3), we conclude that (3.3) = (3.2) \circ (1.2).

Theorem 1.1 now follows from diagram (3.4) and the fact that both maps (3.2) and (3.3) are bijective. The canonical extension \( \phi_n : K_n \Rightarrow E_n \) of \( \phi : K_0 \Rightarrow E_0 \) is then the composition \( \pi_s^0 \circ \Phi(e) \circ l \), where \( \Phi \) is the unique natural transformation associated to \( \phi \) under the bijection (1.2).

4. Proof of Proposition 2.3

The essence of the proof consists in describing the unique natural transformation \( \Phi \in \text{Nat}(K, HH) \) which corresponds to (2.2) under the bijection (1.2). Recall from [24, §15] the construction of the universal additive invariant of dg categories
\[ U_A : \text{HO}(dgcat) \longrightarrow \text{Mot}_A. \]

Given any Quillen model category \( \mathcal{M} \) we have an induced equivalence of categories
\[ (U_A)^* : \text{Hom}_A(\text{Mot}_A, \text{HO}(\mathcal{M})) \cong \text{Hom}_A(\text{HO}(dgcat), \text{HO}(\mathcal{M})), \]
where the left-hand side denotes the category of homotopy colimit preserving morphisms of derivators and the right-hand side the category of additive invariants of dg categories. The algebraic \( K \)-theory functor \( K \) is additive and so the induced morphism \( \mathbb{K} \) is an additive invariant of dg categories. Thanks to equivalence (4.1), it factors then uniquely through \( U_A \). Recall from [24, Theorem 15.10] that for every dg category \( \mathcal{A} \) we have a weak equivalence of spectra
\[ \mathbb{R}\text{Hom}(U_A(\mathcal{A}), U_A(\mathcal{A})) \cong K(\mathcal{A}), \]
where \( \mathbb{R}\text{Hom}(\cdot, \cdot) \) denotes the spectral enrichment of \( \text{Mot}_A \) (see [3, §A.3]). Therefore, we conclude that \( \mathbb{K} \) can be expressed as the following composition
\[ \text{HO}(dgcat) \overset{U_A}{\longrightarrow} \text{Mot}_A \overset{\mathbb{R}\text{Hom}(U_A(\cdot), \cdot)}{\longrightarrow} \text{HO}(\text{Spt}). \]

The Hochschild homology functor, with values in the projective Quillen model category \( \mathcal{C}(k) \) of cochain complexes of \( k \)-modules (see [12, Theorem 2.3.11]), verifies conditions (i)-(iii). Hence, it gives rise to an additive additive invariant of dg categories which we denote by
\[ \mathbb{H}H : \text{HO}(dgcat) \longrightarrow \text{HO}(\mathcal{C}(k)). \]

Note that, according to our notation, \( \mathbb{L}L \mathbb{L} \) can be expressed as the following composition
\[ \text{HO}(dgcat) \overset{\mathbb{L}LH}{\longrightarrow} \text{HO}(\mathcal{C}(k)) \overset{\mathbb{R}\text{Hom}(k, \cdot)}{\longrightarrow} \text{HO}(\text{Spt}). \]
Equivalence (4.1) provide us then the following commutative diagram

\[
\begin{array}{ccc}
\text{HO(dgcM)} & \xrightarrow{\mathbb{H}} & \text{HO}(C(k)) \\
\uparrow{\mathcal{U}_A} & & \uparrow{\mathbb{H}} \\
\text{Mot}_A & \xrightarrow{\pi_0} & \text{Mot}_A.
\end{array}
\]

By construction, the morphism \(\mathbb{H}\) maps \(\mathcal{U}_A(k)\) to \(\mathbb{H}(k) = k\). Hence, by making use of the above factorizations (4.2) and (4.3), we conclude that it induces a natural transformation \(\Phi \in \text{Nat}(K, \mathbb{H})\). We now show that the image of this natural transformation \(\Phi\) by the map (1.2) is the natural transformation (2.2). By taking \(E = \mathbb{H}\) in bijection (3.2) we obtain:

(4.4) \[
\text{Nat}(K_0, \mathbb{H}H_0) \simeq \mathbb{H}H_0(k) \simeq k 
\begin{align*}
\eta \mapsto \eta([k])
\end{align*}
\]

Under this bijection, the natural transformation (2.2) corresponds to the unit of the base ring \(k\); see [26, Theorem 1.3]. Hence, it suffices to show that the same holds for the natural transformation \(\pi_0 \circ \Phi(e) \circ l\) associated to \(\Phi\). The class \([k]\) of \(k\) (as a module over itself) in the Grothendieck group \(K_0(k)\) corresponds to the identity morphism in \(\text{Hom}_{\text{Mot}_A}(\mathcal{U}_A(k), \mathcal{U}_A(k)) \simeq K_0(k) \simeq K_0(k)\).

By functoriality, \(\mathbb{H}(e)\) maps this identity morphism to the identity morphism in \(\text{Hom}_{\text{Mot}_A}(\mathbb{H}(k), \mathbb{H}(k))\). Under the natural isomorphisms

\[
\text{Hom}_{\text{Mot}_A}(\mathcal{U}_A(k), \mathcal{U}_A(k)) \simeq \text{K}(k, k) \simeq k
\]

the identity morphism corresponds to the unit of the base ring \(k\) and so we conclude that \(\pi_0 \circ \Phi(e) \circ l\) agrees with (2.2). This implies that \(\Phi\) is in fact the unique natural transformation which corresponds to (2.2) under the bijection (1.2).

Finally, let \(A\) be a \(k\)-algebra. As proved in [27, Theorem 2.8], the canonical extension \(\phi_n : K_n(A) \to \mathbb{H}H_n(A)\) of \(\phi\) (i.e. the abelian group homomorphism \((\pi_0 \circ \Phi(e) \circ l)(A)\)) agrees with the \(n\)-th trace map constructed by Dennis and so the proof is finished.

5. Proof of Proposition 2.5

We prove first the particular case \((i = 0)\). Let us start by describing the unique natural transformation \(\Phi \in \text{Nat}(K, \mathbb{H})\) which corresponds to \(\phi : K_0 \to \mathbb{H}C_0\) under the bijection (1.2). Observe that \(\mathbb{H}C\) can be expressed as the following composition

\[
\text{HO(dgcM)} \xrightarrow{M} \text{HO}(C(A)) \xrightarrow{p} \text{HO}(k[u]-\text{Comod}) \xrightarrow{U} \text{HO}(C(k)) \xrightarrow{\mathbb{H}\text{Hom}(k, -)} \text{HO}(\text{Spt}).
\]

Some explanations are in order: \(C(A)\) is the projective Quillen model category of mixed complexes and \(M\) the morphism induced by the mixed complex construction\(^3\) (see [4, Example 7.10]); \(k[u]-\text{Comod}\) is the Quillen model category of \(k[u]\)-comodules (where \(k[u]\) is the Hopf algebra of polynomials in one variable

\(^3\)Denoted by \(C\) in [4, Example 7.10].
of degree 2) and $P$ the morphism induced by the periodization construction (see [4, Example 7.11]); $U$ is the morphism induced by the natural forgetful construction. Moreover, as explained in [4, Examples 8.10 and 8.11], negative cyclic homology and periodic cyclic homology admit the following factorizations:

\[
\begin{align*}
\mathbb{H}C^- : \text{HO}(\text{dgcat}) & \xrightarrow{M} \text{HO}(\mathcal{C}(\Lambda)) \xrightarrow{\mathbb{R}\text{Hom}(\overline{k},-)} \text{HO}(\text{Spt}) \\
\mathbb{H}P : \text{HO}(\text{dgcat}) & \xrightarrow{(P \circ M)} \text{HO}(k[u]\text{-Comod}) \xrightarrow{\mathbb{R}\text{Hom}(k[u],-)} \text{HO}(\text{Spt}) .
\end{align*}
\]

Therefore, since $P$ maps $k$ to $k[u]$ and $U$ maps $k[u]$ to $k$, we obtain the classical natural transformations

\[
\begin{align*}
\mathbb{H}C^- & \Rightarrow \mathbb{H}P \Rightarrow \mathbb{H}C \\
\text{by construction, the morphism } M \text{ maps } U_A(k) \text{ to } M(k) = k. \end{align*}
\]

Thus, taking the image of $\Phi$ by the map (1.2) is the natural transformation $\phi : K_0 \Rightarrow HC_0$. Recall from [26, Theorem 1.7(ii)] that $\phi$ admits the following factorization

\[
K_0 \xrightarrow{ch_0^-} HC_0^- \Rightarrow HP_0 \Rightarrow HC_0 ,
\]

where $ch_0^-$ is the negative Chern character and the other natural transformations are the ones associated to (5.2). Hence, it suffices to show that $\pi_0^* \circ \Phi_1(e) \circ l$, associated to $\Phi_1 : K \Rightarrow \mathbb{H}C^-$, agrees with $ch_0^-$. This fact is proved in [28, Proposition 4.2] and so we conclude that $\Phi$ is the unique natural transformation which corresponds to $\phi$ under the bijection (1.2). Now, let $A$ be a $k$-algebra. As explained in [20, §11.4.3], Karoubi’s Chern character $ch_{n,0}(A)$ can be expressed as the following composition

\[
K_n(A) \xrightarrow{ch_n^-} HC_n^-(A) \rightarrow HP_n(A) \rightarrow HC_n(A) .
\]

Note that the right-hand side maps coincide the ones associated to (5.2). Therefore, it suffices to show that the abelian group homomorphism

\[
(\pi_0^* \circ \Phi_1(e) \circ l)(A) : K_n(A) \rightarrow HC_n^-(A)
\]

agrees with $ch_n^-(A)$. This latter fact is proved in [27, Theorem 2.8] and so the proof of the particular case ($i = 0$) is finished.
We now prove the case \((i > 0)\). Recall from \([13, \S 1]\) that for any dg category \(\mathcal{A}\) we have a natural periodicity map \(S : \Omega^2 M(\mathcal{A}) \to M(\mathcal{A})\) in the category \(\mathcal{C}(\mathcal{A})\) of mixed complexes. This construction is natural in \(\mathcal{A}\) and so by iterating it we obtain an infinite sequence of maps
\[
\cdots \to \Omega^2 M(\mathcal{A}) \to \cdots \to \Omega^2 M(\mathcal{A}) \to M(\mathcal{A}).
\]
Under the natural equivalences
\[
\mathbb{R}\text{Hom}(k, \Omega^2 M(-)) \cong \Omega^2 \mathbb{H} C^- \\
\mathbb{R}\text{Hom}(k,\mathcal{P}(\Omega^2 M(-))) \cong \Omega^2 \mathbb{H} P \\
\mathbb{R}\text{Hom}(k,\mathcal{U}(\Omega^2 M(-)))) \cong \Omega^2 \mathbb{H} C,
\]
the above sequence of maps \((5.3)\) gives rise to the following commutative diagram of natural transformations
\[
\begin{array}{cccc}
\mathbb{K} & \xrightarrow{\Phi_1} & \mathbb{H} C^- & \xrightarrow{\cong} \mathbb{H} P & \xrightarrow{\cong} \mathbb{H} C \\
\downarrow & & \downarrow & & \downarrow \\
\Omega^2 \mathbb{H} C^- & \xrightarrow{\cong} & \Omega^2 \mathbb{H} P & \xrightarrow{\cong} & \Omega^2 \mathbb{H} C \\
\downarrow & & \downarrow & & \downarrow \\
\cdots & & \cdots & & \cdots \\
\Omega^2 \mathbb{H} C^- & \xrightarrow{\cong} & \Omega^2 \mathbb{H} P & \xrightarrow{\cong} & \Omega^2 \mathbb{H} C \\
\downarrow & & \downarrow & & \downarrow \\
\cdots & & \cdots & & \cdots
\end{array}
\]
The periodicity map \(S\) becomes invertible in periodic cyclic homology and so the middle column in \((5.4)\) consists of natural isomorphisms. Hence, we obtain the classical sequence of natural transformations
\[
\mathbb{H} P \Rightarrow \cdots \Rightarrow \Omega^2 \mathbb{H} C \Rightarrow \cdots \Rightarrow \Omega^2 \mathbb{H} C \Rightarrow \mathbb{H} C
\]
which relates periodic cyclic homology with the even dimensional loopings of cyclic homology; see \([20, \S 5.1.8]\). Let us then take for \(\Phi\) the composed natural transformation
\[
\mathbb{K} \xrightarrow{\Phi} \mathbb{H} C^- \Rightarrow \mathbb{H} P \Rightarrow \Omega^2 \mathbb{H} C.
\]
The fact that its image by the map \((1.2)\) is the natural transformation \(\phi : K_0 \Rightarrow HC_{2i}\) is now an immediate consequence of the following factorization
\[
\phi : K_0 \xrightarrow{ch} HC_0^- \Rightarrow HP_0 \Rightarrow HC_{2i},
\]
see \([26, \text{Theorem 1.7(ii)}]\), and the particular case \((i = 0)\). Similarly, the fact that the canonical extension \(\phi_n : K_n(A) \to HC_{n+2i}(A)\) agrees with Karoubi’s higher Chern character \(ch_{n,i}(A)\) follows from the following factorization
\[
ch_{n,i}(A) : K_n(A) \xrightarrow{ch_n(A)} HC_n^-(A) \Rightarrow HP_n(A) \Rightarrow HC_{n+2i}(A),
\]
see [20, §11.4.3], and the particular case \((i = 0)\). This achieves the proof.

**Appendix A. Grothendieck Derivators**

In order to make this article more self-contained we give a brief introduction to Grothendieck’s theory of derivators [11]; this language can easily be acquired by skimming through [21], [5, §1] or [3, 4, Appendix A].

Derivators originate in the problem of higher homotopies in derived categories. Given a triangulated category \(\mathcal{T}\) and a small category \(I\), it essentially never happens that the diagram category \(\text{Fun}(I^{op}, \mathcal{T})\) remains triangulated; this already fails for the category of arrows in \(\mathcal{T}\). However, our triangulated category \(\mathcal{T}\) often appears as the homotopy category \(\mathcal{T} = \text{Ho}(\mathcal{M})\) of some Quillen model category \(\mathcal{M}\) (see [22]). In this case we can consider the category \(\text{Fun}(I^{op}, \mathcal{M})\) of diagrams in \(\mathcal{M}\) whose homotopy category \(\text{Ho}(\text{Fun}(I^{op}, \mathcal{M}))\) is triangulated and provides a reasonable approximation to \(\text{Fun}(I^{op}, \mathcal{T})\). More importantly, one can let \(I\) vary. This “nebula” of categories \(\text{Ho}(\text{Fun}(I^{op}, \mathcal{M}))\), indexed by small categories \(I\), and the various (adjoint) functors between them is what Grothendieck formalized into the concept of a *derivator*.

A derivator consists of a strict contravariant 2-functor, from the 2-category of small categories to the 2-category of all categories, subject to five natural conditions. We shall not list these conditions here for it would be too long; see [5, §1]. The essential example to keep in mind is the (triangulated) derivator \(\text{Ho}(\mathcal{M})\) associated to a (stable) Quillen model category \(\mathcal{M}\) and defined for every small category \(I\) by

\[
\text{Ho}(\mathcal{M})(I) := \text{Ho}(\text{Fun}(I^{op}, \mathcal{M})).
\]

We will write \(e\) for the 1-point category with one object and one identity morphism. Note that \(\text{Ho}(\mathcal{M})(e)\) is the homotopy category \(\text{Ho}(\mathcal{M})\). Given Quillen model categories \(\mathcal{M}_1\) and \(\mathcal{M}_2\) and weak equivalence preserving functors \(E, F : \mathcal{M}_1 \to \mathcal{M}_2\), we will denote by \(E, F : \text{Ho}(\mathcal{M}_1) \to \text{Ho}(\mathcal{M}_2)\) the induced morphisms of derivators and by \(\text{Nat}(E, F)\) the set of natural transformations from \(E\) to \(F\); see [5, §5]. Note that given \(\Phi \in \text{Nat}(E, F)\), \(\Phi(e)\) is a natural transformation between the induced functors \(E(e), F(e) : \text{Ho}(\mathcal{M}_1) \to \text{Ho}(\mathcal{M}_2)\).

**Acknowledgments:** The author is very grateful to Bruno Kahn for motivating questions, and to Paul Balmer, Denis-Charles Cisinski and Christian Haesemeyer for useful discussions. He would also like to thank the Departments of Mathematics of UCLA, MIT and UIC for its hospitality and the Clay Mathematics Institute, the Midwest Topology Network, and the Calouste Gulbenkian Foundation for financial support.

**References**

[1] A. Bondal and M. Van den Bergh, *Generators and representability of functors in commutative and noncommutative geometry*. Moscow Math. J. 3 (2003), 1–36.
Extending natural transformations to $K$-theory

[2] A. K. Bousfield and D. Kan, *Homotopy limits, completions, and localizations*. Lecture notes in Mathematics 304 (1972).

[3] D.-C. Cisinski and G. Tabuada, *Non-connective $K$-theory via universal invariants*. Compositio Mathematica 147 (2011), 1281–1320.

[4] D.-C. Cisinski, *Symmetric monoidal structure on Non-commutative motives*. Available at arXiv:1001.0228v2. To appear in Journal of $K$-theory.

[5] D.-C. Cisinski and A. Neeman, *Additivity for derivator $K$-theory*. Adv. Math. 217 (2008), no. 4, 1381–1475.

[6] A. Connes, *Non-commutative differential geometry*. Publ. Math. IHES 62 (1985), 257–360.

[7] K. Dennis, *In search of a new homology theory*. Unpublished manuscript (1976).

[8] V. Drinfeld, *DG quotients of DG categories*. J. Algebra 272 (2004), 643–691.

[9] A. Grothendieck, *Les Déivateurs*. Available at http://people.math.jussieu.fr/maltsin/groth/Derivateurs.html.

[10] A. Grothendieck, *Classes de faisceaux et théorème de Riemann-Roch (SGA6)*. Lecture notes in Mathematics 225 (1971), 20–77.

[11] M. Hovey, *Model categories*. Mathematical Surveys and Monographs 63. American Mathematical Society (1999).

[12] J.D.S. Jones and C. Kassel, *Bivariant cyclic theory. K-theory* 3 (1989), 339–365.

[13] D. Kaledin, *Motivic structures in non-commutative geometry*. Available at arXiv:1003.3210. To appear in the Proceedings of the ICM 2010.

[14] M. Karoubi, *Homologie cyclique et $K$-théorie*. Astérisque 149 (1987).

[15] B. Keller, *On differential graded categories*, International Congress of Mathematicians (Madrid), Vol. II, 151–190. Eur. Math. Soc., Zürich (2000).

[16] M. Kontsevich, *Non-commutative motives*. Talk at the Institute for Advanced Study on the occasion of the 61st birthday of Pierre Deligne, October 2005. Video available at http://video.ias.edu/Geometry-and-Arithmetic.

[17] ______, *Triangulated categories and geometry*. Course at the École Normale Supérieure, Paris, 1998. Notes available at www.math.uchicago.edu/mitya/langlands.html.

[18] ______, *Notes on motives in finite characteristic*. Algebra, arithmetic, and geometry: in honor of Yu. I. Manin. Vol. II, 213–247, Progr. Math., 270, Birkhäuser Boston, MA, 2009.

[19] J.-L. Loday, *Cyclic homology*. Grundlehren der Mathematischen Wissenschaften 301 (1992). Springer-Verlag, Berlin.

[20] G. Maltsiniotis, *Introduction à la théorie des dérivateurs (d’après Grothendieck)*. Available at www.math.jussieu.fr/maltsin/textes.html.
[22] D. Quillen, *Homotopical algebra*. Lecture notes in Mathematics 43, Springer-Verlag (1967).
[23] ———, *Higher algebraic K-theory I*. Lecture notes in Mathematics 341 (1973), 85–147.
[24] G. Tabuada, *Higher K-theory via universal invariants*. Duke Math. J. 145 (2008), 121–206.
[25] ———, *Additive invariants of dg categories*. Int. Math. Res. Not. 53 (2005), 3309–3339.
[26] ———, *A universal characterization of the Chern character maps*. Proc. Amer. Math. Soc. 139 (2011), 1263–1271.
[27] ———, *Products, multiplicative Chern characters, and finite coefficients via Non-commutative motives*. Available at arXiv:1101.0731v2.
[28] ———, *Bivariant cyclic cohomology and Connes’ bilinear pairings in Non-commutative motives*. Available at arXiv:1005.2336v2.

Gonçalo Tabuada

Department of Mathematics

MIT

Cambridge, MA 02139

USA

tabuada@math.mit.edu

Dep. de Matemática

FCT-UNL

Quinta da Torre, 2829-516

Caparica

Portugal

tabuada@fct.unl.pt

Documenta Mathematica 16 (2011) 657–668
Ordinarity of Configuration Spaces
and of Wonderful Compactifications

KIRTI JOHSHI

Abstract. We prove the following: (1) if \( X \) is ordinary, the Fulton-MacPherson configuration space \( X[n] \) is ordinary for all \( n \); (2) the moduli of stable \( n \)-pointed curves of genus zero is ordinary. (3) More generally we show that a wonderful compactification \( X_G \) is ordinary if and only if \( (X, G) \) is an ordinary building set. This implies the ordinarity of many other well-known configuration spaces (under suitable assumptions).

2010 Mathematics Subject Classification: 14G17, 14J99
Keywords and Phrases: Ordinary varieties, ordinarity, configuration spaces, wonderful compactification, moduli of \( n \)-pointed, stable curves of genus zero.

1. Introduction
In the past few years a number of configuration spaces have been studied (see \[FM94, DCP95, Uly92, H03, L09, CGK06, KS08\]). This class of schemes also includes the moduli of \( n \)-pointed stable curves of genus zero, denoted here by \( \overline{M}_{0,n} \) (for \( n \geq 3 \)). All these configuration spaces typically arise from an initial datum, which usually consists of a collection of closed non-singular subschemes of a non-singular, projective variety with certain additional properties—like transversal intersection; as well combinatorial data such as an integer or a graph. Given an initial datum, the configuration space associated to it is typically constructed as a sequence of blowups using the subschemes provided in the initial datum. Many configuration schemes constructed in the
above references can also be constructed as wonderful compactifications of suitable open varieties constructed from the initial datum (see [DCP95], [Li09]).

Now suppose that $k$ is an algebraically closed field of characteristic $p > 0$. A smooth, projective variety $X/k$ is said to be ordinary if $H^i(X, BW \Omega^j_X) = 0$ for all $i, j$. Here $H^i(X, BW \Omega^j_X)$ are the groups defined in [IR83] using the de Rham-Witt complex. The vanishing of these is equivalent to the vanishing of the Zariski cohomology groups $H^i(X, B \Omega^j_X)$ for all $i, j$ where for any $j \geq 0$, $B \Omega^j_X = d \Omega^{j-1}_X$ is the sheaf of locally exact $j$-forms (see [IR83]). Ordinarity of a variety is a difficult condition to check in practice as it requires an understanding of crystalline Frobenius. Here are some examples of ordinary varieties: projective spaces, Grassmanians, more generally homogeneous spaces $G/P$ for $G$ semisimple, $P$ parabolic subgroup of $G$; for abelian varieties ordinarity in the above sense is equivalent to ordinarity in the usual sense (invertibility of the Hasse-Witt matrix); that a general abelian variety with a suitable polarization is ordinary is a nontrivial result of Peter Norman and Frans Oort [NO80]; that a general complete intersection in projective space is ordinary is a delicate result of Luc Illusie (see [Ill90]).

Our remark in this note is that a configuration scheme (of the above type), or more generally a wonderful compactification, arising from an initial datum is ordinary if and only if it arises from an ordinary initial datum (see Theorem 3.2 and Corollary 3.3). In particular we prove that the following schemes are ordinary: (1) if $X$ is a smooth, ordinary, projective variety and let $X[n]$ be the configuration space of Fulton-MacPherson (see [FM94]) and its generalizations (see [KS08]). The scheme $X[n]$ is a compactification of stable configurations of $n$-points of $X$. (2) $\overline{M}_{0,n}$, the moduli space of $n$-pointed stable curves of genus zero ([Kee92]). (3) The compactification $X(n)$ of Ulyanov [Uly02]. (4) the compactification of Kuiperberg-Thurston, [Li09]. (5) the spaces $T_{d,n}$ of stable, pointed, rooted trees of [CGK06]. (6) the compactification of open varieties due to Yi Hu (see [Hu03]).

The proof is not difficult but as all of these configuration schemes play an important role in many areas of algebraic geometry, so their properties in positive characteristic are not without interest, and hence worth recording.

This note grew out of our attempt to answer a question raised by Indranil Biswas (unfortunately we cannot answer his question–see Remark 3.4 for more on this). It is a pleasure to thank him for many conversations about his question. We thank Ana-Maria Castravet for many conversations about $\overline{M}_{0,n}$, and especially pointing out the constructions of [Kee92] [Kap03]. We thank the referee for comments and suggestions which have improved the readability of this paper.

2. Preliminaries

Let $k$ be a perfect field of characteristic $p > 0$. Let $W \Omega^*_X$ be the de Rham-Witt complex of $X$. Let $H^i(X, W \Omega^*_X)$ (for $i + j \leq \dim(X)$) be the de Rham-Witt cohomology groups. We say that $X$ is ordinary if $H^i(X, BW \Omega^*_X) = 0$
This is equivalent (see [IR83, Theorem 4.13, Page 209]) to the vanishing of $H^i(X, B\Omega^j_X) = 0$ for $i, j \geq 0$ where $B\Omega^j = d(\Omega^{j-1}_X)$ is the sheaf of locally exact differentials. As we are in characteristic $p > 0$, so $d(f^*\omega) = f^*d\omega$ for any (local) sections $f$ of $O_X$ and $\omega$ of $\Omega^{-1}_X$, hence $B\Omega^j_X$ is naturally a subsheaf of $F_c(\Omega^j_X)$ and it is in fact locally free as $X$ is smooth of finite type and so at any rate the sheaf $B\Omega^j_X$ carries a natural structure of an $O_X$-module (see [Ill79] Proposition 2.2.8, Page 520) for more details). The condition of ordinarity is equivalent (see [IR83] Theorem 4.13, Page 209-210 and its proof) to the condition:

$$F : H^i(X, W\Omega^j_X) \to H^i(X, W\Omega^j_X)$$

is an isomorphism of $W$-modules for all $i, j \geq 0$. We say that $X$ is Hodge-Witt if $H^i(X, W\Omega^j_X)$ is finite type over $W$ for all $i, j \geq 0$. By definition any ordinary variety is Hodge-Witt (see [IR83] Defn. 4.12, Page 208). We will use the following standard results.

The following result is [Eke85] III, Prop 2.1(ii) and Prop 7.2(ii)]

**Proposition 2.2** ([Ekedahl [Eke85]]. Let $X, Y$ be smooth, projective varieties over $k$. Then

1. If $X, Y$ are ordinary then $X \times_k Y$ is ordinary.
2. If $X \times_k Y$ is Hodge-Witt then one of $X$ or $Y$ is ordinary and the other is Hodge-Witt.

The following result is [Ill90] Prop. 1.4.

**Proposition 2.3** ([Illusie [Ill90]]) Let $X$ be a smooth, projective variety over a perfect field $k$. Let $V$ be a vector bundle on $X$. Let $\mathbb{P}(V) \to X$ be the associated projective bundle. Then $X$ is ordinary if and only if $\mathbb{P}(V)$ is ordinary.

For a smooth, projective variety $X$ and $Z \subset X$ a smooth, closed subscheme, let $\text{Bl}_Z(X)$ be the blowup of $X$ along $Z$. For $Y \subset X$ we write $\tilde{Y} \subset \text{Bl}_Z(X)$ for the dominant transform of $Y$ in $\text{Bl}_Z(X)$, defined as $\tilde{Y} = \pi^{-1}(Y)$ if $Y \subset Z$ and the strict transform of $Y$ in $\text{Bl}_Z(X)$ otherwise.

We need the following version of [Ill90] Proposition 1.6):

**Proposition 2.4**. Let $X$ be a smooth, projective scheme over an algebraically closed field $k$. Let $Z \subset X$ be a subscheme of $X$ and let $Y \subset X$ be a smooth, closed subscheme of $X$. Let $\tilde{Y} \subset \text{Bl}_Z(X)$ be the dominant transform of $Y$ in $\text{Bl}_Z(X)$. Then $\tilde{Y}$ is ordinary if and only if $Y$ and $Y \cap Z$ are ordinary.

**Proof.** We write $\pi : \text{Bl}_Z(X) \to X$ for the blowup morphism. Then by [Ill90] Proposition 1.6], $\text{Bl}_Z(X)$ is ordinary if and only if $X, Z$ are ordinary. Next observe that the dominant transform $\tilde{Z}$ of $Z$ is the exceptional divisor and by [Har77] Theorem 8.24(b), page 186], $\tilde{Z} \to Z$ is a projective bundle and so by Proposition 2.3) is $\tilde{Z}$ is ordinary if and only if $Z$ is ordinary.

Now to prove the assertion. Let $Z \subset X$ be a smooth, proper subscheme of a smooth, proper $X$. Let $Y \subset X$ be a smooth, proper subscheme. Let $\tilde{Y} \subset \text{Bl}_Z(X)$ be the dominant transform of $Y$ in $\text{Bl}_Z(X)$. We consider several
subcases. If $Y$ is a subset of $Z$, then the dominant transform $\tilde{Y} \to Y$ is a projective bundle over $Y$ and hence $\tilde{Y}$ is ordinary if and only if $Y$ is ordinary (by Proposition 2.3). If $Y = Z$, then the dominant transform $\tilde{Y}$ is the exceptional divisor $E \subset \text{Bl}_Z(X)$. Since $E$ is a projective bundle over $Z$, we see that $\tilde{Y} = E$ is ordinary if and only if $Z$ is ordinary. If $Y \not\subset Z$ then we proceed as follows. If $Y \cap Z = \emptyset$ then $\tilde{Y} \cong Y$ and hence is ordinary as $Y$ is ordinary. If $Y \cap Z$ is non-empty and by previous considerations, we may assume that $Y \neq Z$. In this case $\tilde{Y}$ is the blowup of $Y$ along $Y \cap Z$ and so the result follows from [Il90, Proposition 1.6]. This proves the claim.

□

3. Building sets and wonderful compactification

Let $X$ be a smooth, projective scheme over an algebraically closed field $k$. Let $\mathcal{S}$ be a finite collection of closed, smooth subschemes of $X$. We say that $\mathcal{S}$ is an arrangement if the scheme theoretic intersection of any elements of $\mathcal{S}$ is either empty or an element of $\mathcal{S}$.

Let $\mathcal{S}$ be an arrangement of subschemes of $X$. We say that $G \subset \mathcal{S}$ is a building set if for all $S \in \mathcal{S}\setminus G$, the minimal elements in $\{G \in G : G \supset S\}$ intersect transversally and their intersection is $S$.

A set of subschemes $\mathcal{G}$ of $X$ is called a building set if the collection of all possible intersections of elements of $\mathcal{G}$ is an arrangement of subschemes of $X$ and $\mathcal{G}$ is a building set of this arrangement.

Let $X$ be a smooth, projective scheme over $k$ and let $\mathcal{G}$ be a building set of $X$. Let $X_G \subset \prod_{G \in G} \text{Bl}_G(X)$ be the closure of $X^o = X \setminus \bigcup_{G \in G} G$. Then we have the following [Li09, Theorem 1.2]:

**Theorem 3.1 (Li09).** Let $X$ be a smooth, projective variety over an algebraically closed field $k$. Let $\mathcal{G}$ be a building set of $X$. Then $X_G$ is a smooth, projective variety over $k$.

The scheme $X_G$ is called the wonderful compactification of $(X, \mathcal{G})$.

We say that a building set $\mathcal{G}$ of $X$ is an ordinary building set if $X$ is ordinary and all the scheme theoretic intersections of any members of $\mathcal{G}$ are ordinary (recall that by our convention empty intersections are also ordinary). We say that an arrangement $\mathcal{S}$ of $X$ is ordinary if $\mathcal{S}$ arises from an ordinary building set.

**Theorem 3.2.** Let $X/k$ be a smooth, projective scheme over a perfect field of characteristic $p > 0$. Let $\mathcal{G}$ be a building set associated to $X$. Then the wonderful compactification $X_G$ associated to $X$ is ordinary if and only if $\mathcal{G}$ is an ordinary building set.

**Corollary 3.3.** Let $X$ be an smooth, projective variety over $k$. Assume that $X$ is ordinary. Then the following schemes associated to $X$ are all ordinary:

1. the scheme $X[n]$ of Fulton-MacPherson (see FM94)
2. the scheme $X(n)$ of Ulyanov (see Uly02)
3. the scheme $X^T$ of Kuiper-Thurston (see Li09)
(4) the generalized Fulton-Macpherson configuration scheme $X_D^{[n]}$, $X_D[n]$ (we assume $D$ is a smooth, ordinary subscheme of $X$) of [KS08].

(5) the moduli, $\overline{M}_{0,n}$ (for $n \geq 3$), of $n$-pointed stable curves of genus zero is ordinary.

(6) the scheme of $T^{d,n}$ of stable, $n$-pointed, rooted trees of $d$-dimensional projective spaces of [CGK06].

Proof of Theorem 3.3. We recall the details of the construction of $X_G$ from [Li09] Definition 2.12, Proposition 2.13. The construction is inductively carried out as follows. Let $S$ be an arrangement of $X$ and $\mathcal{G}$ be a building set of $S$. Then assume that $\mathcal{G} = \{G_1, \ldots, G_N\}$ is indexed so that $G_i \subset G_j$ if $i \leq j$. We define $(X_k, S^{(k)}, \mathcal{G}^{(k)})$ as follows. For $k = 0$, set $X_0 = X, S^{(0)} = S, \mathcal{G}^{(0)} = \mathcal{G}, G_i^{(0)} = G_i$ for $1 \leq i \leq N$. Then $(X_0, S^{(0)}, \mathcal{G}^{(0)})$ is ordinary. Assume by induction that $(X_{k-1}, S^{(k-1)}, \mathcal{G}^{(k-1)})$ has been constructed so that $X_{k-1}$ is ordinary and $\mathcal{G}^{(k-1)}$ is an ordinary building set for $X_{k-1}$. Then $S^{(k-1)}$ consists of ordinary subvarieties of $X_{k-1}$. Define $X_k = \text{Bl}_{\mathcal{G}^{(k-1)}}(X_{k-1})$. Then by Proposition 2.4, $X_k$ is ordinary if and only if $X_{k-1}$ and $G_i^{(k-1)}$ are both ordinary. Now define $G^{(k)}$ be the dominant transform of $G^{(k-1)}$ for $G \in \mathcal{G}$. Define $\mathcal{G}^{(k)} = \{G^{(k)} : G \in \mathcal{G}\}$; by Lemma 2.1, $\mathcal{G}^{(k)}$ is ordinary and define $S^{(k)}$ to be the induced arrangement of $\mathcal{G}^{(k)}$. Since $\mathcal{G}^{(k)}$ is ordinary, we see that $S^{(k)}$ is ordinary. Finally for $k = N$ we get $X_N = X_G$.

We note that the theorem includes the compactification scheme constructed in [Hu03] as a special case. The fact that this scheme arises from a suitable building set is checked in [Li09].

Proof of 3.3. To deduce the Corollary 3.3 from Theorem 3.2 it suffices to produce ordinary building sets to construct $X[n], X(n), X^1$ etc. The building sets for these are constructed in (see the discussion of [Li09] Theorem 1.2 in the paragraph following it). These building sets are building sets of $X^n$. To prove that they are ordinary building sets if $X$ is ordinary, we note that the building sets for (1)-(3) consists of diagonals or polydiagonals, i.e. self-products of $X$ embedded in $X^n$ by various diagonals. Thus the ordinarity of these building sets follows from Proposition 2.2(i) by ordinarity of self-products of ordinary varieties. Conversely if anyone of these configurations spaces is ordinary then as this space is a blowup, by Proposition 2.4 we see that it must be a blowup of an ordinary variety along an ordinary center, descending down the blowup sequence in this fashion, we deduced that some self-product of $X$ is ordinary, and so by applying Proposition 2.2(ii) (which says for us that self-products of $X$ are Hodge-Witt if and only if $X$ is ordinary) we see that $X$ ordinary. For constructing $X_D^{[n]}$, we use a building set which is constructed by [KS08], from $X^n$, by blowing up a suitable subschemes which are self products of $D, X$. By Proposition 2.2 this gives an ordinary building set. The result follows from Theorem 3.2. To construct $X_D[n]$, we start with an ordinary building set in $X_D^{[n]}$, consisting of the proper transform in $X_D^{[n]}$ of the multi-diagonals in $X^n$. This is again an ordinary building set. Again we can see in both these cases,
by repeating our earlier argument, that if any of $X^{[n]}_D$, $X_D[n]$ is ordinary then $X, D$ must be ordinary.

(5) This assertion is strictly part of the formalism of wonderful compactification via $Kap03$, but may be of independent interest and so we give a proof for the sake of completeness using $Kee92$ where $\overline{M}_{0,n}$ is constructed from $\mathbb{P}^1 \times \cdots \times \mathbb{P}^1$ by a suitable sequence of blowups with smooth, ordinary centers which are related to $M_{0,j}$ for $j < n$. In $Kee92$ provided a construction of $\overline{M}_{0,n}$ as a sequence of blowups and products. We will prove Theorem 3.3(5) by induction on $n$. Suppose $n = 3$ then $\overline{M}_{0,3}$ is a point hence is ordinary. Assume that $n = 4$, then $\overline{M}_{0,4} = \mathbb{P}^1$ hence is ordinary. Assume that the ordinarity of $\overline{M}_{0,j}$ has been established for some all $j \leq n$; we will show that $\overline{M}_{0,n}$ is also ordinary. Recall the construction of $Kee92$ (will notations of that paper for this proof). We let $B_1 = \overline{M}_{0,n} \times \overline{M}_{0,4} = \overline{M}_{0,n} \times \mathbb{P}^1$. Then as $\overline{M}_{0,n}$ is ordinary by induction and as $\mathbb{P}^1$ is ordinary, so we deduce that $B_1$ is ordinary (see Proposition 2.2). The construction of $\overline{M}_{0,n}$ shows that for each subset $T \subset \{1, 2, \ldots, n\}$ with $|T^c| \geq 2$, there exists a collection of smooth subschemes $D_T$. For each $T$ these subschemes are isomorphic to $\overline{M}_{0,1} \times \overline{M}_{0,j}$ for suitable $i, j < n$. Thus by our induction hypothesis and Proposition 2.2 we see that $D_T$ are ordinary and hence, by Proposition 2.3 so is the blowup of $B_1$ along these $D_T$ for every $T$. Thus $B_2$ is ordinary. More generally $B_k \rightarrow B_{k-1}$ is the blowup of $B_{k-1}$ along the (disjoint) union of strict transforms of $D_T$ (for $|T^c| = k + 1$) under $B_k \rightarrow B_1$. Then $B_k$ is ordinary as $D_T$ are isomorphic to $\overline{M}_{0,1} \times \overline{M}_{0,j}$ for suitable $i, j < n$. Thus $B_k$ is ordinary and $\overline{M}_{0,n+1} = B_{n-2}$. This proves the assertion.

For (6) this is not immediate from $Li09$ so we recall that $T_{d,n}$ is constructed in $CGK06$ Theorem 3.3.1 in a manner similar to the Fulton-MacPherson configuration scheme $X[n]$. The procedure is inductive and starting from $T_{1,3} = \mathbb{P}^1, T_{1,2} = \mathbb{P}^{d-1}, T_{1,n} = \overline{M}_{0,n+1}$ (note that by the previous results these are all ordinary) we construct $T_{d,n}$ as follows: suppose $T_{d,n}$ has been constructed for some $d, n$. Then $T_{d,n+1}$ is a sequence of blowups of a projective bundle over $T_{d,n}$. Since the later is ordinary by induction, so is the projective bundle over $T_{d,n}$ (by Proposition 2.3). The next blowups are along subschemes of the projective bundle which can be identified with $T_{d} \times T_{d,j}$ for $j < n + 1$ and so these subschemes are ordinary by Proposition 2.2. This proves the assertion. □

Remark 3.4. Indrani Biswas has asked us the following question: if $X$ is a smooth, projective ordinary surface, then is Hilb$_n(X)$ ordinary for all $n \geq 1$? We note that it is known that if $X$ is Frobenius split, smooth, projective surface by then $KT01$ Hilb$_n(X)$ is Frobenius split. By $JR03$ smooth, proper Frobenius split surfaces are ordinary. However by $JR03$ the class of Frobenius split varieties is not a subclass of ordinary varieties in dimensions at least three and we note that the class of ordinary surfaces is much bigger—for instance it includes general type surfaces in $\mathbb{P}^3$ by the result of $H00$. In any case Biswas’ question presents a natural variant of $KT01$. 

Documenta Mathematica 16 (2011) 669–676
Unfortunately we do not know how to answer Biswas’s question. The methods outlined here are not adequate as they require a far better understanding of the geometry of $\text{Hilb}_n(X)$ than we seem to have at the moment. We note however that we can easily deduce the result for $\text{Hilb}_2(X)$ from our result for the Fulton-MacPherson configuration space $X[2]$.

References

[CGK06] Linda Chen, Angela Gibney, and Daniel Krashen, Pointed trees of projective spaces, Preprint, 2006.

[DCP95] C. De Concini and C. Procesi, Wonderful models of subspace arrangements, Selecta Math. (N.S.) 1 (1995), no. 3, 459–494. MR MR1366622 (97k:14013)

[Eke85] Torsten Ekedahl, On the multiplicative properties of the de Rham-Witt complex. II, Ark. Mat. 23 (1985), no. 1, 53–102.

[FM94] William Fulton and Robert MacPherson, A compactification of configuration spaces, Ann. of Math. (2) 139 (1994), no. 1, 183–225.

[Har77] Robin Hartshorne, Algebraic geometry, Springer-Verlag, New York, 1977, Graduate Texts in Mathematics, No. 52.

[Hu03] Yi Hu, A compactification of open varieties, Trans. Amer. Math. Soc. 355 (2003), no. 12, 4737–4753 (electronic).

[Ill79] Luc Illusie, Complexe de de rham-witt et cohomologie cristalline, Annales scientifiques de l’E. N. S. 12 (1979), no. 4, 501–661.

[Ill90] Luc Illusie, Ordinarité des intersections complètes générales, The Grothendieck Festschrift, Vol. II, Progr. Math., vol. 87, Birkhäuser Boston, Boston, MA, 1990, pp. 376–405.

[IR83] Luc Illusie and Michel Raynaud, Les suites spectrales associées au complexe de de Rham-Witt, Inst. Hautes Études Sci. Publ. Math. (1983), no. 57, 73–212.

[JR03] Kirti Joshi and C. S. Rajan, Frobenius splitting and ordinarity, Int. Math. Res. Not. (2003), no. 2, 109–121.

[Kap93] M. M. Kapranov, Veronese curves and Grothendieck-Knudsen moduli space $\overline{M}_{0,n}$, J. Algebraic Geom. 2 (1993), no. 2, 239–262.

[Kee92] Sean Keel, Intersection theory of moduli space of stable $n$-pointed curves of genus zero, Trans. Amer. Math. Soc. 330 (1992), no. 2, 545–574.

[KS08] Bumsig Kim and Fumitoshi Sato, A generalization of Fulton-Macpherson configuration spaces, Preprint, 2008.

[KT01] Shrawan Kumar and Jesper Funch Thomsen, Frobenius splitting of Hilbert schemes of points on surfaces, Math. Ann. 319 (2001), no. 4, 797–808.

[Li09] Li Li, Wonderful compactification of an arrangement of subvarieties, Michigan Math. J. 58 (2009), no. 2, 535–563.

[NO80] Peter Norman and Frans Oort, Moduli of abelian varieties, Ann. of Math. (2) 112 (1980), no. 3, 413–439.
[Uly02] Alexander P. Ulyanov, *Polydiagonal compactification of configuration spaces*, J. Algebraic Geom. 11 (2002), no. 1, 129–159.

Kirti Joshi
Math. department
University of Arizona
617 N Santa Rita
Tucson, 85721-0089
USA
kirti@math.arizona.edu
Let $k$ be a field of characteristic $\neq 2$, and let $L$ be a Galois extension of $k$ with group $G$. Let us denote by $q_L : L \times L \to k$ the trace form, defined by $q_L(x, y) = \text{Tr}_{L/k}(xy)$. Let $(gx)_{g \in G}$ be a normal basis of $L$ over $k$. We say that this is a self–dual normal basis if $q_L(gx, hx) = \delta_{g,h}$. If the order of $G$ is odd, then $L$ always has a self–dual normal basis over $k$ (cf. [1]). This is no longer true in general if the order of $G$ is even; some partial results are given in [2].

If $k$ is a global field, then it is natural to ask whether a local–global principle holds for this problem. In order to make this question precise, we have to consider $G$–Galois algebras and not only field extensions. Moreover, it is useful to note that $q_L$ is a $G$–quadratic form, in other words $q_L(gx, gy) = q_L(x, y)$ for all $x, y \in L$ and $g \in G$. The $G$–Galois algebra has a self–dual normal basis if and only if the $G$–form $q_L$ is isomorphic to the unit $G$–form. This leads to the following question:

**Question.** Suppose that $k$ is a global field, and let $L$ and $L'$ be two $G$–Galois algebras. Assume that for all places $v$ of $k$, the $G$–forms $q_L$ and $q_{L'}$ are isomorphic over $k_v$. Are the $G$–forms $q_L$ and $q_{L'}$ isomorphic over $k$?

Note that a similar Hasse principle does not hold for arbitrary $G$–forms, cf. Morales [5]. In the context of trace forms of $G$–Galois algebras, positive results are obtained in [2] in some special cases. However, the problem is open in general.

The starting point of this paper is to investigate this question. The main tool, which is of independent interest, is to develop induction–restriction methods for arbitrary $G$–forms, generalizing some results of [2] and of Lequeu in [4].
key ingredient is an odd determinant property of the group $G$ (cf. §2) which is shown to hold for instance if the normalizer of a $2$–Sylow subgroup $S$ controls the fusion of $S$ in $G$. We obtain the following:

**Theorem.** Suppose that $k$ is a global field of characteristic $\neq 2$. Let $G$ be a finite group, and suppose that $G$ has the odd determinant property if $\text{char}(k) = 0$. Let $L$ and $L'$ be two $G$–Galois algebras such that for all places $v$ of $k$, the $G$–forms $q_L$ and $q_{L'}$ are isomorphic over $k_v$. Then the $G$–forms $q_L$ and $q_{L'}$ are isomorphic over $k$.

**Corollary.** Suppose that $k$ and $G$ are as above. Then a $G$–Galois algebra has a self–dual normal basis over $k$ if and only if such a basis exists over all the completions of $k$.

We thank Jean–Pierre Serre for many fruitful conversations and correspondence. We also thank Emmanuel Lequeu for very interesting conversations, and for showing us a preliminary version of his paper [4]. The authors acknowledge hospitality of Emory University and the École Polytechnique Fédérale de Lausanne, institutions in which the work was carried out. The first named author was partially supported by the Swiss National Science Foundation, grant 200020-109174/1, and the second named author was partially supported by NSF grant DMS 1001872.

§1. Definitions and basic facts

Let $k$ be a field of characteristic $\neq 2$, let $G$ be a finite group, and let $k[G]$ be the associated group ring. We refer to [7] for basic facts on $k[G]$–modules.

**Group ring and involution**

Let $\iota : k[G] \to k[G]$ be the canonical involution of the group ring, in other words the $k$–linear involution of $k[G]$ characterized by $\iota(g) = g^{-1}$ for all $g \in G$. Let $R$ be the radical of $k[G]$. Then $k[G]/R$ is a semi–simple $k$–algebra, hence we have a decomposition $k[G]/R = \prod_{i=1,\ldots,r} M_{n_i}(D_i)$, where $D_1,\ldots,D_r$ are division algebras. Let us denote by $K_i$ the center of $D_i$, and let $D_i^{\text{op}}$ be the opposite algebra of $D_i$.

Note that $\iota(R) = R$, hence $\iota$ induces an involution $\iota : k[G]/R \to k[G]/R$. Therefore $k[G]/R$ decomposes into a product of involution invariant factors. These can be of two types: either an involution invariant matrix algebra $M_{n_i}(D_i)$, or a product $M_{n_i}(D_i) \times M_{n_i}(D_i^{\text{op}})$, with $M_{n_i}(D_i)$ and $M_{n_i}(D_i^{\text{op}})$ exchanged by the involution. We say that a factor is unitary if the restriction of the involution to its center is not the identity: in other words, either an involution invariant $M_{n_i}(D_i)$ with $\iota|_{K_i}$ not the identity, or a product $M_{n_i}(D_i) \times M_{n_i}(D_i^{\text{op}})$. Otherwise, the factor is said to be of the first kind. In this case, the component
is of the form $M_{n_i}(D_i)$ and the restriction of $\iota$ to $K_i$ is the identity. We say that the component is \textit{orthogonal} if after base change to a separable closure $\iota$ is given by the transposition, and \textit{symplectic} otherwise.

We say that a component $M_{n_i}(D_i)$ is \textit{split} if $D_i$ is a commutative field.

\textbf{G--quadratic forms}

A \textit{G--quadratic form} is a pair $(V,q)$, where $V$ is a $k[G]$--module that is a finite dimensional $k$--vector space, and $q : V \times V \to k$ is a non-degenerate symmetric bilinear form such that 

$$q(gx,gy) = q(x,y)$$

for all $x, y \in V$ and all $g \in G$. We say that two $G$--quadratic forms $(V,q)$ and $(V',q')$ are \textit{isomorphic} if there exists an isomorphism of $k[G]$--modules $f : V \to V'$ such that $q'(f(x),f(y)) = q(x,y)$ for all $x,y \in V$. If this is the case, we write $(V,q) \simeq_G (V',q')$, or $q \simeq_G q'$.

Let $S$ be a subgroup of $G$. We have two operations, induction and restriction (see for instance [2], 1.2 for details):

If $(V,q)$ is an $S$--quadratic form, then $\text{Ind}^G_S(V,q)$ is a $G$--quadratic form;

If $(V,q)$ is a $G$--quadratic form, then $\text{Res}^S_G(V,q)$ is an $S$--quadratic form.

The following result will be used in the sequel

\textbf{Theorem 1.1.} (see [1], th. 4.1) \textit{Let $q$ and $q'$ be two $G$--quadratic forms. If they become isomorphic over an odd degree extension, then they are isomorphic.}

It is well-known that $S$--quadratic forms correspond bijectively to $k[S]$--hermitian forms with respect to the involution $\iota : k[S] \to k[S]$. We will use the same notation for the $S$--quadratic form and the corresponding hermitian form.

\textbf{Trace forms}

Let $L$ be a $G$--Galois algebra, and let 

$$q_L : L \times L \to k, \quad q_L(x,y) = \text{Tr}_{L/k}(xy)$$

be its trace form. Then $q_L$ is a $G$--quadratic form.

Let us recall a result from [2] that will be basic for the proof of the main theorem:

\textbf{Lemma 1.2.} (cf. [2], 2.1.1): \textit{Let $S$ be a 2–Sylow subgroup of $G$. For any $G$–Galois algebra $L$, there exists an odd degree field extension $k'/k$ and an $S$–Galois algebra $M$ over $k'$ such that the $G$–form $(L,q_L) \otimes_k k'$ is isomorphic to the $G$–form $\text{Ind}_S^G(q_M)$.}
§2. The induction-restriction functor and the odd determinant property

The aim of this section is to introduce the odd determinant property, and to state a result (th. 2.2), which will be used in the proof of the Hasse principle result of §3.

Let $G$ be a finite group, let $S$ be a 2–Sylow subgroup of $G$, and let $N = N_G(S)$ be the normalizer of $S$ in $G$. Then $N$ acts on $S$, and we denote by $\Sigma$ the set of orbits of $S$ under the action of $N$.

Let $X$ be the $\mathbb{Z}$–module of $\mathbb{Z}$–valued functions on $S$ invariant under conjugation by $N$, and let $\Phi : X \rightarrow X$ be $\text{Res}_S^G \text{Ind}_S^G$ considered as an endomorphism of $X$ (cf. [7], 7.2).

**Definition 2.1** We say that $G$ has the **odd determinant property** if the determinant of $\Phi : X \rightarrow X$ is an odd integer.

One of the main results of this paper is the following

**Theorem 2.2** Suppose that $G$ has the odd determinant property. Let $(V_1, q_1)$ and $(V_2, q_2)$ be two $S$–quadratic spaces. Suppose that

$$\text{Res}_S^G \text{Ind}_S^G(V_1, q_1) \simeq_S \text{Res}_S^G \text{Ind}_S^G(V_2, q_2).$$

Then

$$\text{Ind}_S^G(V_1, q_1) \simeq_G \text{Ind}_S^G(V_2, q_2).$$

This result is used in the proof of the Hasse principle stated in the introduction, see th. 3.1. The proof relies on an analysis of the odd determinant property, and is the subject matter of sections 4–11. The structure of the proof of th. 2.2 is as follows. Sections 5 and 6 study induction and restriction properties of $S$–quadratic forms. Section 7 is concerned with the odd determinant property in the special case where all the characters of $S$ over $k$ are absolutely irreducible. Using a filtration introduced in §9 and the quadratic descent argument of §8, we obtain a general result (see th. 10.1) based on the case considered in §7. This is then used in §11 to prove th. 2.2.

We next show that the odd determinant property holds if $N$ controls the fusion of $S$ in $G$.

**Definition 2.3** We say that $N$ **controls the fusion of $S$ in $G$** if for all subsets $T$ and $T'$ of $S$, if there exists $g \in G$ with $gTg^{-1} = T'$ then there exists $n \in N$ such that $nTn^{-1} = T'$.
There are many examples of groups $G$ in which the normalizer controls the fusion of the 2-Sylow subgroups; see for instance Thévenaz [8] for a survey.

**Remark.** Note that we only use the following property, which is clearly satisfied if $N$ controls the fusion of $S$ in $G$:

(*) For all $s, t \in S$, if there exists $g \in G$ with $gs^{-1} = t$ then there exists $n \in N$ such that $nsn^{-1} = t$.

It does not seem to be known whether there exist groups $G$ having property (*) where $N$ does not control the fusion of $S$ in $G$.

**Proposition 2.4** Suppose that $N$ controls the fusion of $S$ in $G$. Then $G$ has the odd determinant property.

In order to prove this proposition, we need the following lemma:

**Lemma 2.5** Suppose that $N$ controls the fusion of $S$ in $G$, and let $x \in S$. Then $C_S(x)$ is a 2-Sylow subgroup of $C_G(x)$.

**Proof.** Let $S_0$ be a 2-Sylow subgroup of $C_G(x)$ containing $x$ and let $S_1$ be a 2-Sylow subgroup of $G$ containing $S_0$. Let $g \in G$ be such that $gS_1g^{-1} = S$. In view of the fusion hypothesis, there exists $n \in N$ such that $ngxg^{-1}n^{-1} = x$. Let us consider $\text{Int}(ng) : G \rightarrow G$. Then, as $\text{Int}(ng)(x) = x$, we have $\text{Int}(ng)(C_G(x)) = C_G(x)$. We have $\text{Int}(ng)(S_1) = S$, hence $\text{Int}(ng)(S_0) = \text{Int}(ng)(S_1 \cap C_G(x)) = S \cap C_G(x) = C_S(x)$. This implies that $C_S(x)$ is a 2-Sylow subgroup of $C_G(x)$, as claimed.

**Proof of Prop. 2.4** For $\sigma \in \Sigma$, let $q_\sigma$ be the function on $S$ which is equal to 1 on $\sigma$ and 0 otherwise. Note that the set $(q_\sigma)_{\sigma \in \Sigma}$ is a basis of the $\mathbb{Z}$-module $X$. Let $\sigma, \sigma' \in \Sigma$, and fix $x \in \sigma'$. By definition, the coefficient of $q_\sigma$ in $\Phi(q_{\sigma'})$ is equal to

$$\frac{1}{\#S} \# \{ g \in G \mid gxg^{-1} \in \sigma \}.$$

As $N$ controls the fusion of $S$ in $G$, we have $gxg^{-1} \in \sigma$ if and only if $x \in \sigma$. Therefore the coefficient of $q_\sigma$ in $\Phi(q_{\sigma'})$ is equal to 0 if $\sigma \neq \sigma'$.

The coefficient of $q_\sigma$ in $\Phi(q_\sigma)$ is equal to

$$\frac{1}{\#S} \#C_G(x) \#\sigma = \frac{1}{\#S} \#C_G(x) \frac{\#N \#C_G(x) \#C_S(x)}{\#S \#C_N(x)} = \frac{\#N \#C_G(x) \#C_S(x)}{\#S \#C_N(x)}.$$

Therefore it suffices to check that $\frac{\#C_G(x)}{\#C_S(x)}$ is odd, and this follows from lemma 2.5.
§3. Hasse principle

In this section, we suppose that \( k \) is a global field of characteristic \( \neq 2 \). Let \( G \) be a finite group, and let us denote by \( k[G] \) the associated group ring. One of the main results of this paper is the following:

**Theorem 3.1** Suppose that \( G \) has the odd determinant property if \( \text{char}(k) = 0 \), and let \( L \) and \( L' \) be two \( G \)-Galois algebras. Then \( q_L \simeq_G q_{L'} \) over \( k \) if and only if \( q_L \simeq_G q_{L'} \) over all the completions of \( k \).

As an immediate consequence, we get

**Corollary 3.2** Suppose that \( G \) has the odd determinant property if \( \text{char}(k) = 0 \). Then a \( G \)-Galois algebra has a self-dual normal basis over \( k \) if and only if it has a self-dual normal basis over every completion of \( k \).

By prop. 2.3, we know that \( G \) has the odd determinant property whenever for a 2-Sylow subgroup \( S \), the normalizer \( N_G(S) \) controls the fusion of \( S \) in \( G \). Hence we have

**Corollary 3.3** Suppose that for a 2-Sylow subgroup \( S \) of \( G \), the normalizer \( N_G(S) \) controls the fusion of \( S \) in \( G \). Then the trace forms of two \( G \)-Galois algebras are \( G \)-isomorphic over \( k \) if and only if they are \( G \)-isomorphic over each completion of \( k \). In particular, a \( G \)-Galois algebra has a self-dual normal basis over \( k \) if and only if it has a self-dual normal basis over every completion of \( k \).

**Corollary 3.4** Suppose that \( G \) has a normal 2-Sylow subgroup. Then the trace forms of two \( G \)-Galois algebras are isomorphic over \( k \) if and only if they are isomorphic over each completion of \( k \). In particular, a \( G \)-Galois algebra has a self-dual normal basis over \( k \) if and only if it has a self-dual normal basis over every completion of \( k \).

**Proof.** This is an immediate consequence of 3.3.

The proof of th. 3.1 relies on th. 2.2, and on some properties of group rings and of quadratic and hermitian forms that we recall in this section. Let us first note that the Hasse principle holds for any \( G \)-form provided the orthogonal components of the group ring are split:

**Theorem 3.5** Suppose that all the orthogonal components of \( k[G] \) are split, and let \( q, q' \) be two \( G \)-forms. Then \( q \simeq_G q' \) over \( k \) if and only if \( q \simeq_G q' \) over all the completions of \( k \).

**Proof.** This follows from the Hasse principle for unitary and symplectic forms, as well as the Hasse principle for quadratic forms over global fields (see for instance [6], chap. 10).
Therefore th. 3.1 is new for number fields only – indeed, if char$(k) > 0$, then all the orthogonal components of $k[G]$ are split.

**Proposition 3.6** Let $S$ be a 2–group. Then the orthogonal and unitary components of $k[2]$ are split, and the symplectic components of $k[2]$ are either split, or of the form $M_n(H)$ where $H$ is a quaternion division algebra over its center.

**Proof.** Note that $k[2] = Q[2] \otimes_Q k$ if char$(k) = 0$, and $k[2] = F_p[2] \otimes_{F_p} k$ if char$(k) = p \neq 0$. Therefore it is sufficient to prove the proposition when $k = Q$ or $k = F_p$. As the Brauer group of a finite field is trivial, every component is split if $k = F_p$.

Suppose that $k = Q$. Then each component of $Q[2]$ is invariant under $\iota$ (cf. [6], Chap 8, 13.2.).

Let $M_n(D)$ be a symplectic component of $Q[2]$. This implies that the algebra $M_n(D)$ is of order one or two in the Brauer group of $Q$, and it is well–known that this can only happen if $D$ is a commutative field or a quaternion algebra.

Let us now show that the orthogonal and unitary components of $Q[2]$ are split. Let $v$ be a non–dyadic place of $Q$, and let $O_v$ be the ring of integers of $Q_v$. Since $\#S$ is invertible in $O_v$, it follows that $O_v[2]$ is Azumaya over its center. This implies that this algebra is split mod $\pi$, where $\pi$ is a uniformizer at $v$, therefore it is split over $O_v$. In particular every component of $Q_v[2]$ is split.

If $v$ is the real place of $Q$, then every orthogonal and unitary component of $Q_v[2] = R[2]$ is split (cf. [6], Chap 8, 13.5).

Let $M_n(D)$ be an orthogonal or unitary component of $Q[2]$, and let $Z(D) = K$. As $S$ is a 2–group, $K$ is a subfield of a 2–cyclotomic field, hence $K$ admits a unique dyadic place. Since $D$ is split at all the other places, $D$ is split at the dyadic place as well, hence $D$ is split.

**Corollary 3.7** Let $S$ be a 2–group, and let $q, q'$ be two $S$–forms. Then $q \simeq_S q'$ over $k$ if and only if $q \simeq_S q'$ over all the completions of $k$.

**Proof.** This follows from 3.5 and 3.6.

We are now ready to prove 3.1. The proof uses th. 2.2, which will be proved in section 11.

**Proof of th.** 3.1 Suppose first that char$(k) > 0$. Then all the components of $k[G]$ are split, hence th. 3.5 implies the desired result.

Suppose now that char$(k) = 0$, in other words that $k$ is an algebraic number field. By lemma 1.2, there exists an odd degree field extension $k'/k$ and $S$–Galois algebras $M$ and $M'$ over $k'$ such that $(L, q_L) \otimes_{k'} k' \simeq_G \text{Ind}^G_S(M, q_M)$, and $(L', q_{L'}) \otimes_{k'} k' \simeq_G \text{Ind}^G_S(M', q_{M'})$. Recall that by hypothesis the $G$–forms $(L, q_L)$ and $(L', q_{L'})$ are isomorphic over all the completions of $k$. This implies
that the \(G\)-forms \((L, q_L) \otimes_k k'\) and \((L', q_{L'}) \otimes_k k'\) are isomorphic over all the completions of \(k'\). Hence the \(S\)-forms \(\text{Res}^G_S(L, q_L) \otimes_k k' \simeq_S \text{Res}^G_S(M, q_M)\), and \(\text{Res}^G_S(L', q_{L'}) \otimes_k k' \simeq_S \text{Res}^G_S(M', q_{M'})\) are isomorphic over all the completions of \(k'\). By corollary 3.7, this implies that the \(S\) forms \(\text{Res}^G_S(L, q_L) \otimes_k k' \simeq \text{Res}^G_S(M, q_M)\), and \(\text{Res}^G_S(L', q_{L'}) \otimes_k k' \simeq \text{Res}^G_S(M', q_{M'})\) are isomorphic over \(k'\). As \(G\) has the odd determinant property, th. 2.2 implies that the \(G\)-forms \(\text{Ind}^G_S(M, q_M)\) and \(\text{Ind}^G_S(M', q_{M'})\) are isomorphic. As \((L, q_L) \otimes_k k' \simeq_G \text{Ind}^G_S(M, q_M)\) and \((L', q_{L'}) \otimes_k k' \simeq_G \text{Ind}^G_S(M', q_{M'})\), we get \((L, q_L) \otimes_k k' \simeq_G (L', q_{L'}) \otimes_k k'\). By th. 1.1, this implies that \((L, q_L) \simeq_G (L', q_{L'})\), and this completes the proof of th. 3.1.

§4. Properties of determinants in characteristic 2

This section is concerned with properties of determinants of linear transformations over rings of characteristic 2 that will be needed in the following sections. Let \(F\) be a field of characteristic 2, and let \(R = F[X]/(X^2 + 1)\). We start by recalling a result of linear algebra:

**Proposition 4.1** Let \(M = R^n\) be the free \(R\)-module of rank \(n\), and let \(f: M \to M\) be an \(R\)-linear map. Then

\[
N_{R/F}(\det(f)) = \det(f),
\]

where \(\det(f)\) is the determinant of \(f\) considered as an \(F\)-linear map.

**Corollary 4.2** Let

\[
A = \begin{pmatrix}
    a_{1,1} & b_{1,1} & \cdots & a_{1,n} & b_{1,n} \\
    b_{1,1} & a_{1,1} & \cdots & b_{1,n} & a_{1,n} \\
    \vdots & \vdots & \ddots & \vdots & \vdots \\
    a_{n,1} & b_{n,1} & \cdots & a_{n,n} & b_{n,n} \\
    b_{n,1} & a_{n,1} & \cdots & b_{n,n} & a_{n,n}
\end{pmatrix}
\]

and

\[
B = \begin{pmatrix}
    a_{1,1} + b_{1,1} & \cdots & a_{1,n} + b_{1,n} \\
    \cdots & \cdots & \cdots \\
    a_{n,1} + b_{n,1} & \cdots & a_{n,n} + b_{n,n}
\end{pmatrix}
\]

with \(a_{i,j}, b_{i,j} \in F\). Then

\[
\det(B)^2 = \det(A).
\]

**Proof.** Let \(f: R^n \to R^n\) be defined by \(f(e_j) = \sum_{1 \leq i \leq n} (a_{i,j} + b_{i,j}X)e_i\), where \(e_1, \ldots, e_n\) is the standard basis of \(R^n\). The matrix of \(f\) with respect to the basis \(e_1, \ldots, e_n\) is \((a_{i,j} + b_{i,j}X)\). We have

\[
N_{R/F}(\det(a_{i,j} + b_{i,j}X)) = (\det(a_{i,j} + b_{i,j}))^2,
\]

Documenta Mathematica 16 (2011) 677–707
which is equal to $\det(B)^2$. By 4.1, this is the determinant of $f$ as an $F$–linear map. On the other hand, the determinant of $f$ in the basis $e_1, e_1X, e_2, e_2X, \ldots, e_n, e_nX$ is equal to $\det(A)$; hence we have $\det(B)^2 = \det(A)$, as claimed.

We also need the following observation:

**Lemma 4.3** Let $n \in \mathbb{N}$, and suppose that the group $\{1, \iota\}$ of order 2 acts on the set $\{1, \ldots, n\}$ in such a way that $\{1, \ldots, r\}$ is the set of fixed points. Let $(d_{i,j})_{1 \leq i, j \leq n}$ be an integral matrix such that $d_{\iota(i), \iota(j)} = d_{i,j}$ for all $i, j$. Then

$$\det(d_{i,j})_{1 \leq i, j \leq n} \equiv \det(d_{i,j})_{1 \leq i, j \leq r} \det(d_{i,j})_{r+1 \leq i, j \leq n} \pmod{2}.$$

**Proof.** Let $S$ be the set of permutations of $\{1, 2, \ldots, n\}$. For $s \in S$ and $1 \leq i \leq n$, set $\iota * s(i) = \iota s(i)$. We have

$$\det(d_{i,j})_{1 \leq i, j \leq n} \equiv \sum_{s \in S} \prod_{1 \leq i \leq n} d_{s(i),i} \pmod{2}.$$

Set

$$H = \{s \in S \mid s(i) \leq r \text{ for } i \leq r\}.$$

Then

$$\sum_{s \in S} \prod_{1 \leq i \leq n} d_{s(i),i} = \sum_{s \in H} \prod_{1 \leq i \leq n} d_{s(i),i} + \sum_{s \notin H} \prod_{1 \leq i \leq n} d_{s(i),i}.$$

For $s \notin H$, we have $\iota * s \notin H$ and $s \neq \iota * s$. In view of $d_{i,j} = d_{\iota(i),\iota(j)}$ for all $i, j$, we get

$$\sum_{s \notin H} \prod_{1 \leq i \leq n} d_{s(i),i} \equiv 0 \pmod{2}.$$

Let

$$S^1 = \{s \in S \mid s(i) = i \text{ for } i \geq r+1\},$$

and

$$S^2 = \{s \in S \mid s(i) = i \text{ for } i \leq r\}.$$

Then we have

$$\sum_{s \in H} \prod_{1 \leq i \leq n} d_{s(i),i} = [\sum_{s \in S^1} \prod_{1 \leq i \leq r} d_{s(i),i}] [\sum_{s \in S^2} \prod_{r+1 \leq i \leq n} d_{s(i),i}].$$

This completes the proof of the lemma.

Lemma 4.3 is used in the next sections, in particular in the proofs of 7.1, 8.4 and 8.5.
§5. Group rings of 2–groups and decomposition of \( S \)–quadratic forms

The aim of this section is to introduce some tools and notation that will be used in the sequel. In particular, we set up a decomposition of the quadratic forms invariant by a 2–group, generalizing the approach of [2], §5.

**Group rings of 2–groups**

Let \( k \) be a field of characteristic \( \neq 2 \), and let \( S \) be a 2–group. Recall that \( \iota : k[S] \to k[S] \) is the canonical involution of the group ring.

As the characteristic of \( k \) is not 2, the algebra \( k[S] \) is semi–simple. We have a decomposition of \( k[S] \) into simple factors, corresponding to the irreducible representations of \( S \) over \( k \), hence also to the irreducible characters of \( S \) over \( k \). Let us denote by \( S'_k \) the set of these irreducible characters. Each of them determines a component \( M_{n_s}(\Delta_x) \) of \( k[S] \), where \( \Delta_x \) is a division algebra. Let \( K_x = Z(\Delta_x) \) be the center of \( \Delta_x \). Recall that the orthogonal and unitary components are split, and that the symplectic components are either split, or of the form \( M_n(H) \) where \( H \) is a quaternion division algebra (see prop. 3.6).

Let us denote by \( U_x \) the simple \( k[S] \)–module associated to the irreducible character \( x \in S'_k \). Note that it is isomorphic to \( \Delta_x^{\iota} \). Let \( Y_k \) be the free \( Z \)–module generated by \( S'_k \).

Note that \( \iota \) acts on \( S'_k \) by \( \iota(x)(s) = x(s^{-1}) \) for all \( x \in S'_k \) and \( s \in S \). We say that \( x \in S'_k \) is self–dual if \( \iota(x) = x \). This is equivalent to requiring that the corresponding component \( M_{n_s}(\Delta_x) \) is stable by \( \iota \). If \( x \in S'_k \) is not self–dual, then there exists \( x' \in S'_k \) such that \( x' \neq x \) and \( \iota(x) = x' \). In this case, set \( \overline{x} = (x,x') \). If \( x \) is self–dual, then set \( \overline{x} = x \). Let us denote by \( \overline{S'_k} \) the set of \( \overline{x} \) for \( x \in S'_k \).

Set \( M_{n_s}(\Delta_{\overline{x}}) = M_{n_s}(\Delta_x) \) if \( x \) is self–dual, and \( M_{n_s}(\Delta_{\overline{x}}) = M_{n_s}(\Delta_x) \times M_{n_{s'}}(\Delta_{x'}) \) if \( \iota(x) = x' \neq x \). Similarly, set \( K_{\overline{x}} = K_x \) if \( x \) is self–dual and \( K_{\overline{x}} = K_x \times K_{x'} \) if \( \overline{x} = (x,x') \). Note that \( K_{\overline{x}} \) is an étale algebra, but not necessarily a field. Let \( K_{\overline{x}} \iota = \{ a \in K_{\overline{x}} | \iota(a) = a \} \) be the invariants of \( \iota \) in \( K_{\overline{x}} \). When \( x \) is not self–dual, then we have \( K_x \simeq K_{x'} \simeq K_{\overline{x}} \).

The involution \( \iota \) of \( k[S] \) restricts to the factors \( M_{n_s}(\Delta_{\overline{x}}) \), and it is adjoint to a hermitian or skew–hermitian form, which we fix in the different cases as follows.

If \( x \) is orthogonal, then \( \Delta_{\overline{x}} = K_x \). In this case, we set \( D_{\overline{x}} = K_x \), and we chose the involution \( \tau_{\overline{x}} : D_{\overline{x}} \to D_{\overline{x}} \) to be the identity. The restriction of the involution \( \iota \) to this factor is adjoint to a symmetric form on \( D_{\overline{x}} \), which we
denote by \(\rho_\mathbb{T}\). We define \(m_x = n_x\), and the symmetric form is supported on the simple module \(U_x\).

If \(x\) is symplectic, then \(\Delta_\mathbb{T} = K_\mathbb{T}\) or a quaternion division algebra. We set \(D_{\mathbb{T}} = M_2(K_x)\) in the first case, and \(D_{\mathbb{T}} = \Delta_x\) in the second case. In both cases, we choose the involution \(\tau_\mathbb{T}: D_{\mathbb{T}} \rightarrow D_{\mathbb{T}}\) to be the standard symplectic involution of \(D_{\mathbb{T}}\). In this case, the restriction of the involution \(\iota\) to this factor is adjoint to a hermitian form over \(D_{\mathbb{T}}\), with respect to the involution \(\tau_\mathbb{T}\) which we denote by \(\rho_{\mathbb{T}}\).

The form \(\rho_{\mathbb{T}}\) is supported on the module \(U_x \oplus U_x\) and \(n_x = 2m_x\) if \(D_{\mathbb{T}}\) is not division, it is supported on the module \(U_x\) and \(m_x = n_x\) if \(D_{\mathbb{T}}\) is division.

If \(x\) is unitary, then \(\Delta_\mathbb{T} = K_\mathbb{T}\), and \(K_\mathbb{T}\) is a quadratic algebra over \(K_0^\mathbb{T}\). We set \(D_{\mathbb{T}} = K_\mathbb{T}\), and we fix the involution \(\tau_\mathbb{T}: D_{\mathbb{T}} \rightarrow D_{\mathbb{T}}\) to be the non-trivial automorphism of this quadratic algebra. Then the restriction of the involution \(\iota\) to this factor is adjoint to a hermitian form on \(D_{\mathbb{T}}\), with respect to the involution \(\tau_\mathbb{T}\) which we denote by \(\rho_{\mathbb{T}}\). We set \(m_x = n_x\) in this case. The form \(\rho_{\mathbb{T}}\) is supported on \(U_x\) if \(x\) is self-dual, and on \(U_{x_1} \oplus U_{x_2}\) if \(x = (x_1, x_2)\) with \(x_1 \neq x_2\) and \(\iota(x_1) = x_2\).

Set \(U_{\mathbb{T}} = U_x \oplus U_x\) if \(x\) is symplectic and \(D_{\mathbb{T}}\) not division, \(U_{\mathbb{T}} = U_{x_1} \oplus U_{x_2}\) if \(x\) is unitary with \(x = (x_1, x_2)\) such that \(\iota(x_1) = x_2\) and \(x_1 \neq x_2\), and \(U_{\mathbb{T}} = U_x\) in all other cases. Note that \(U_{\mathbb{T}} \simeq D_{\mathbb{T}}^{0n}\). Therefore in all cases we have a hermitian form \(\rho_{\mathbb{T}}: U_{\mathbb{T}} \times U_{\mathbb{T}} \rightarrow D_{\mathbb{T}}\) which we fix throughout. We denote the hermitian form \((U_{\mathbb{T}}, \rho_{\mathbb{T}})\) by \(\rho_{\mathbb{T}}\).

We also fix a quadratic form \(n_{\mathbb{T}}: D_{\mathbb{T}} \rightarrow K_0^\mathbb{T}\) to be the one-dimensional unit form if \(x\) is orthogonal, the reduced norm form of the quaternion algebra \(D_{\mathbb{T}}\) if \(x\) is symplectic, and the norm form of the quadratic algebra \(D_{\mathbb{T}}\) if \(x\) is unitary.

**Decomposition of \(S\)-quadratic forms**

Let \((V, q)\) be an \(S\)-quadratic form. Then \((V, q)\) decomposes as an orthogonal sum of hermitian forms \((M_{\mathbb{T}}, Q_{\mathbb{T}})\) for \(x \in S_k\), over \(M_{m_x}(D_{\mathbb{T}})\) with respect to the restriction of \(\iota\) to this factor. By Morita theory, fixing \(\rho_{\mathbb{T}}\), the hermitian form \((M_{\mathbb{T}}, Q_{\mathbb{T}})\) is uniquely determined up to isomorphism by a hermitian form \(h_{\mathbb{T}}\) over a free \(D_{\mathbb{T}}\)-module \(W_{\mathbb{T}}\) of finite rank with respect to the involution \(\tau_{\mathbb{T}}\), and conversely the hermitian form \((W_{\mathbb{T}}, h_{\mathbb{T}})\) is uniquely determined up to isomorphism by \((M_{\mathbb{T}}, Q_{\mathbb{T}})\). Moreover, by Jacobson’s theorem the hermitian form \((W_{\mathbb{T}}, h_{\mathbb{T}})\) corresponds to a quadratic form \((V_{\mathbb{T}}, g_{\mathbb{T}})\) over \(K_0^\mathbb{T}\) with the property that \((V_{\mathbb{T}}, g_{\mathbb{T}}) \otimes n_{\mathbb{T}}\) is uniquely determined by \((W_{\mathbb{T}}, h_{\mathbb{T}})\) (cf. [6], 10.1.1 and 10.1.7).

We have \((M_{\mathbb{T}}, Q_{\mathbb{T}}) \simeq \rho_{\mathbb{T}} \otimes K_0^\mathbb{T}(V_{\mathbb{T}}, g_{\mathbb{T}})\), and \((V_{\mathbb{T}}, g_{\mathbb{T}}) \otimes n_{\mathbb{T}}\) is uniquely determined by \((M_{\mathbb{T}}, Q_{\mathbb{T}})\), hence by \((V, q)\). In other words, we have

\[
(V, q) \simeq \bigoplus_{\pi \in S_k} \rho_{\mathbb{T}} \otimes K_0^\mathbb{T}(V_{\mathbb{T}}, g_{\mathbb{T}}),
\]

*Documenta Mathematica* 16 (2011) 677–707
and if \((V_1, q_1)\) and \((V_2, q_2)\) are two \(S\)-quadratic forms with
\[
(V_1, q_1) \simeq \bigoplus_{\pi \in \mathcal{S}_k} \rho_{\pi} \otimes_{K_0} (V_{1, \pi}, g_{1, \pi}) \quad \text{and} \quad (V_2, q_2) \simeq \bigoplus_{\pi \in \mathcal{S}_k} \rho_{\pi} \otimes_{K_0} (V_{2, \pi}, g_{2, \pi}),
\]
then
\[
(V_1, q_1) \simeq_S (V_2, q_2)
\]
if and only if
\[
n_{\pi} \otimes g_{1, \pi} \simeq n_{\pi} \otimes g_{2, \pi}
\]
for all \(\pi \in \mathcal{S}'_k\).

§6. INDUCTION OF \(S\)-FORMS

Let \(k\) be a field of characteristic \(\neq 2\). Let \(G\) be a finite group, and let \(S\) be a 2–Sylow subgroup of \(G\). We use the notation introduced in §5. In particular, \(S'_k\) is the set of irreducible characters of \(S\) over \(k\). Recall that \(\iota : k[S] \to k[S]\) is the standard involution, and that for \(x \in S'_k\) we set \(x = x\) if \(x\) is selfdual, and \(x = (x, x')\) if \(\iota(x) = x' \neq x\).

Let \(N = N_G(S)\) be the normalizer of \(S\) in \(G\). Then \(N\) acts on \(S'_k\) by \(n(x) = x nsn^{-1}\) for all \(n \in N, x \in S'_k\) and \(s \in S\). Note that the actions of \(N\) and \(\iota\) commute. We need the following lemmas:

**Lemma 6.1** The orbits of \(S'_k\) under \(N\) have odd cardinality.

**Proof.** Let \(x \in S'_k\) and let \(\omega\) be the orbit of \(x\) under \(N\). We have \(\sharp(\omega) = \sharp(N/\text{Stab}_N(x))\). As \(S \subset \text{Stab}_N(x)\), we see that \(\sharp(N/\text{Stab}_N(x))\) is odd.

**Lemma 6.2** Let \(x, x' \in S'_k\) such that \(\iota(x) = x' \neq x\). Let \(\omega, \omega'\) be the orbits of \(x, \) respectively \(x'\). Then \(\omega \neq \omega'\).

**Proof.** Indeed, suppose that \(\omega = \omega'\). As the actions of \(N\) and \(\iota\) commute, we see that for every \(n \in N\), we have \(mn(x) \neq n(x)\). This implies that \(\iota y \neq y\) for every \(y \in \omega\), and therefore \(\omega\) has even cardinality, contradicting lemma 6.1. Therefore \(\omega \neq \omega'\).

Let us denote by \(\Omega_k\) the set of orbits of \(S'_k\) under \(N\). There is an induced action of \(N\) on the free \(Z\)-module generated by \(S'_k\) and the set of orbits under this action is the free \(Z\)-module generated by \(\Omega_k\).

Let us define an action of \(\iota\) on \(\Omega_k\) by letting \(\iota \omega\) to be the orbit of \(\iota(x)\) for any \(x \in \omega\); this is well–defined as the actions of \(N\) and \(\iota\) on \(S'_k\) commute. For any
Galois Algebras, Hasse Principle...

Let \( \omega \in \Omega_k \), set \( \overline{\omega} = \omega \) if \( \iota \omega = \omega \), and \( \overline{\omega} = (\omega_1, \omega_2) \) with \( \omega_1 = \omega_2 \) and \( \omega_1 \neq \omega_2 \). Let \( \overline{\Omega}_k \) be the set of all \( \overline{\omega} \) with \( \omega \in \Omega_k \). Let us fix a field extension \( K_0^0 \) of \( k \) such that \( K_0^0 \cong K_0^0 \) for all \( \overline{\tau} \in \overline{\omega} \).

Let \((V,q)\) be an \( S \)-quadratic form. Then we have an orthogonal decomposition

\[
(V,q) \cong \bigoplus_{\overline{\tau} \in \overline{\Omega}_k} \rho_{\overline{\tau}} \otimes K_0^{0,\overline{\tau}} (V_{\overline{\tau}}, g_{\overline{\tau}})
\]

where \((V_{\overline{\tau}}, g_{\overline{\tau}})\) is a quadratic form over \( K_0^{0,\overline{\tau}} \) and \((V_{\overline{\tau}}, g_{\overline{\tau}}) \otimes n_{\overline{\tau}}\) is uniquely determined by \((V,q)\) (cf. §5).

For all \( \overline{\omega} \in \overline{\Omega}_k \), let us consider the orthogonal sum

\[
(V_{\overline{\omega}}, g_{\overline{\omega}}) = \bigoplus_{\overline{\tau} \in \overline{\omega}} (V_{\overline{\tau}}, g_{\overline{\tau}}).
\]

Then \((V_{\overline{\omega}}, g_{\overline{\omega}})\) is a quadratic form over \( K_0^{0,\overline{\omega}} \).

Note that \( \text{Ind}^G_S(\rho_{\overline{\tau}}) \) does not depend on the choice of \( \overline{\tau} \in \overline{\omega} \). Set

\[
I(\overline{\omega}) = \text{Ind}^G_S(\rho_{\overline{\tau}})
\]

where \( \overline{\tau} \) is any element of \( \overline{\omega} \).

Therefore we have

\[
\text{Ind}^G_S(V,q) = \bigoplus_{\overline{\omega} \in \overline{\Omega}_k} I(\overline{\omega}) \otimes K_0^{0,\overline{\omega}} (V_{\overline{\omega}}, g_{\overline{\omega}})
\]

Set

\[
A(V,q) = \text{Res}^G_S \text{Ind}^G_S(V,q).
\]

Then we have

\[
A(V,q) = \bigoplus_{\overline{\tau} \in \overline{\Omega}_k} \text{Res}^G_S(I(\overline{\tau})) \otimes K_0^{0,\overline{\tau}} (V_{\overline{\tau}}, g_{\overline{\tau}}).
\]

Let \( \overline{\eta} \in \overline{\Omega}_k \), and let us take the \( \overline{\eta} \)-component of the equation above. We get

\[
A(V,q)_{\overline{\eta}} = \bigoplus_{\overline{\tau} \in \overline{\Omega}_k} \text{Res}^G_S(I(\overline{\tau}))_{\overline{\eta}} \otimes K_0^{0,\overline{\tau}} (V_{\overline{\tau}}, g_{\overline{\tau}}).
\]

Let \( \overline{\omega} \in \overline{\Omega}_k \) such that \( \overline{\eta} \in \overline{\omega} \). Note that the \( S \)-quadratic spaces \( A(V,q)_{\overline{\eta}} \) and \( \text{Res}^G_S(I(\overline{\tau}))_{\overline{\eta}} \) do not depend on the choice of \( \overline{\eta} \in \overline{\omega} \). Set

\[
A(V,q)_{\overline{\omega}} = A(V,q)_{\overline{\eta}}
\]
and
\[ \text{Res}_G^S(I(\mathcal{W}))_{\mathcal{W}} = \text{Res}_G^S(I(\mathcal{W}'))_{\mathcal{W}} \]
for any \( \mathcal{W} \in \mathcal{W}' \).

Then we have
\[ \text{Res}_G^S(I(\mathcal{W}))_{\mathcal{W}} = \rho_{\mathcal{W}} \otimes_{K_{\mathcal{W}}} F_{\mathcal{W}} \]
for \( \mathcal{W} \in \mathcal{W}' \), where \( F_{\mathcal{W}} \) is a quadratic form over \( K_{\mathcal{W}} \).

Hence
\[ A(V, q)_\mathcal{W} = \rho_{\mathcal{W}} \otimes_{K_{\mathcal{W}}} \bigoplus_{\mathcal{W} \in \mathcal{W}} [F_{\mathcal{W}} \otimes_{K_{\mathcal{W}}} (V_1, g_1)] \]
for all \( \mathcal{W} \in \mathcal{W}' \).

**Notation.** Let \( \omega, \omega' \in \Omega_k \) be such that \( K_{\omega} = K_{\omega'} = k \). We define \( d_{\omega, \omega'} \) to be the dimension of the \( k \)–vector space underlying the quadratic form \( F_{\omega, \omega'} \).

Note that \( d_{\omega, \omega'} \) is the number of times \( \rho_{\mathcal{W}} \) occurs in \( \text{Res}_G^S \text{Ind}_G^S(\rho_{\mathcal{W}}) \) for any \( \mathcal{W} \in \mathcal{W} ', \mathcal{W} \in \mathcal{W}' \). As \( \rho_{\mathcal{W}} \) is the underlying module of \( \rho_{\mathcal{W}} \), the integer \( d_{\omega, \omega'} \) can also be seen as the number of times \( U_{\mathcal{W}} \) occurs in \( \text{Res}_G^S \text{Ind}_G^S(U_{\mathcal{W}}) \) for any \( \mathcal{W} \in \mathcal{W} ', \mathcal{W} \in \mathcal{W}' \).

Let \( (V_1, q_1) \) and \( (V_2, q_2) \) be two \( S \)–quadratic forms. If \( A(V_1, q_1) \simeq A(V_2, q_2) \), then \( A(V_1, q_1)_{\mathcal{W}} \simeq A(V_2, q_2)_{\mathcal{W}} \) for all \( \mathcal{W} \in \Omega_k \). Hence, if we have
\[ (V_1, q_1) \simeq \bigoplus_{\mathcal{W} \in \Omega_k} \rho_{\mathcal{W}} \otimes_{K_{\mathcal{W}}} (V_1, g_1) \quad \text{and} \quad (V_2, q_2) \simeq \bigoplus_{\mathcal{W} \in \Omega_k} \rho_{\mathcal{W}} \otimes_{K_{\mathcal{W}}} (V_2, g_2), \]
then, for each \( \mathcal{W} \in \Omega_k \),
\[ \bigoplus_{\mathcal{W} \in \Omega_k} n_{\mathcal{W}} \otimes_{K_{\mathcal{W}}} [F_{\mathcal{W}} \otimes_{K_{\mathcal{W}}} (V_1, g_1)] \simeq \bigoplus_{\mathcal{W} \in \Omega_k} n_{\mathcal{W}} \otimes_{K_{\mathcal{W}}} [F_{\mathcal{W}} \otimes_{K_{\mathcal{W}}} (V_2, g_2)]. \]

§7. Odd determinant property – a special case

The aim of this section and the next ones is to establish some technical results relative to the odd determinant property. These will be used in §11 to prove th. 2.2.

We keep the notation of the previous sections, and we suppose that all the characters in \( S'_k \) are absolutely irreducible.

Recall that \( \Omega_k \) is the set of \( N \)–orbits of \( S'_k \). The following notation will be important in the sequel:
Notation. Let us define \( d_{\omega,\omega'} \) as being the number of times \( U_y \) occurs in 
\[
\text{Res}_G^S \text{Ind}_G^S(U_x)
\]
for \( x \in \omega, \ y \in \omega' \).

The standard involution \( \iota : k[S] \to k[S] \) acts on \( \Omega_k \). Note that \( d_{\omega,\iota \omega'} = d_{\omega,\omega'} \)
for all \( \omega, \omega' \in \Omega_k \). Let us define
\[
\Omega^1 = \{ \omega \in \Omega_k \mid \iota \omega = \omega \}
\]
and
\[
\Omega^2 = \{ \omega \in \Omega_k \mid \iota \omega \neq \omega \}.
\]

Since all the characters in \( S_k' \) are absolutely irreducible and in view of Lemma 6.2, \( \Omega^1 \) is precisely the set of orbits of irreducible orthogonal and symplectic characters.

**Proposition 7.1** Suppose that \( \det_{\omega,\omega'} \in \Omega_k (d_{\omega,\omega'}) \equiv 1 \pmod{2} \).

Then \( \det_{\omega,\omega'} \in \Omega^1 (d_{\omega,\omega'}) \equiv 1 \pmod{2} \)
and
\( \det_{\omega,\omega'} \in \Omega^2 (d_{\omega,\omega'}) \equiv 1 \pmod{2} \).

**Proof.** Since the group \( \{1, \iota\} \) acts on \( \Omega \) with fixed points precisely \( \Omega^1 \), it follows from lemma 4.3 that
\[
\det_{\omega,\omega'} \in \Omega_k (d_{\omega,\omega'}) \equiv \det_{\omega,\omega'} \in \Omega^1 (d_{\omega,\omega'}) \det_{\omega,\omega'} \in \Omega^2 (d_{\omega,\omega'}) \pmod{2}.
\]
Hence we have
\[
\det_{\omega,\omega'} \in \Omega^1 (d_{\omega,\omega'}) \equiv 1 \pmod{2},
\]
and
\[
\det_{\omega,\omega'} \in \Omega^2 (d_{\omega,\omega'}) \equiv 1 \pmod{2}.
\]
This completes the proof of the proposition.

We define \( \Omega^{1,o} = \{ \omega \in \Omega^1 \mid \omega \text{ orthogonal} \} \), and \( \Omega^{1,s} = \{ \omega \in \Omega^1 \mid \omega \text{ symplectic} \} \).

**Proposition 7.2** Suppose that \( \det_{\omega,\omega'} \in \Omega^1 (d_{\omega,\omega'}) \equiv 1 \pmod{2} \). Then
\[
\det_{\omega,\omega'} \in \Omega^{1,o} (d_{\omega,\omega'}) \equiv 1 \pmod{2},
\]

**Documenta Mathematica** 16 (2011) 677–707
and
\[ \det_{\omega, \omega' \in \Omega^1, \ast}(d_{\omega, \omega'}) \equiv 1 \pmod{2}. \]

**Proof.** Let \( \omega \) be orthogonal and \( \omega' \) symplectic. For \( x \in \omega, y \in \omega' \), recall that \( U_x \) and \( U_y \) are the simple \( k[S] \)-modules associated to \( x \) and \( y \) respectively. Then \( \rho_\tau \) is supported on \( U_\tau = U_y \oplus U_y \) and \( \rho_\tau \) is supported on \( U_\tau = U_x \), hence the \( \gamma \)-component of \( \text{Res}_S^G \text{Ind}_S^G(U_x, \rho_x) \) is isomorphic to \( (U_y \oplus U_y, \rho_y) \otimes_k F_\gamma \). Thus the module \( U_y \) occurs with even multiplicity in \( \text{Res}_S^G \text{Ind}_S^G(U_x, \rho_x) \), so that
\[ d_{\omega, \omega'} \equiv 0 \pmod{2}. \]
Therefore the matrix \( (d_{\omega, \omega'})_{\omega, \omega' \in \Omega^1} \) has the shape
\[ \begin{pmatrix} A & 0 \\ * & B \end{pmatrix}, \]
mod 2, where \( A = \det_{\omega, \omega' \in \Omega^1, \ast}(d_{\omega, \omega'}) \) and \( B = \det_{\omega, \omega' \in \Omega^1, \ast}(d_{\omega, \omega'}) \). This completes the proof of the proposition.

For any \( \omega \in \Omega_k \), recall that \( \omega = \omega \) if \( \iota \omega = \omega \), and \( \omega = (\omega_1, \omega_2) \) with \( \iota \omega_1 = \omega_2 \) and \( \omega_1 \neq \omega_2 \). Let \( \Omega_k \) be the set of all \( \omega \) with \( \omega \in \Omega_k \). Let
\[ \Omega^2 = \{ \omega = (\omega_1, \omega_2) \in \Omega_k \mid \iota \omega_1 = \omega_2 \text{ and } \omega_1 \neq \omega_2 \}. \]

**Proposition 7.3** Suppose that \( \det_{\omega, \omega' \in \Omega^2}(d_{\omega, \omega'}) \equiv 1 \pmod{2} \). Then we have
\[ \det_{\omega, \omega' \in \Omega^2}(d_{\omega, \omega'}) \equiv 1 \pmod{2}. \]

**Proof.** Let \( \omega = (\omega_1, \omega_2) \) and \( \omega' = (\omega'_1, \omega'_2) \), and let \( d_{\omega_1, \omega'_1} = a, d_{\omega_1, \omega'_2} = b \). Then \( d_{\omega, \omega'} = a + b \). For a suitable ordering of the orbits \( \Omega^2 \), and the corresponding ordering of \( \Omega^2 \), the matrices \( d_{\omega, \omega'} \) and \( d_{\omega, \omega'} \) are of the shape \( B \) and \( A \) as in corollary 4.2. Hence \( \det(B)^2 \equiv \det(A) \pmod{2} \). This gives the desired result.

We have the following

**Proposition 7.4** Suppose that \( G \) has the odd determinant property. Then
\[ \det_{\omega, \omega' \in \Omega_k}(d_{\omega, \omega'}) \equiv 1 \pmod{2}. \]

For the proof of prop. 7.4, we need the following lemma.
Lemma 7.5 Let $K$ be a field of characteristic 0, and assume that all the characters in $S'_K$ are absolutely irreducible. Suppose that $G$ has the odd determinant property. Then

$$\det_{\omega,\omega' \in \Omega_K}(d_{\omega,\omega'}) \equiv 1 \pmod{2}.$$ 

Proof. Let $X_K = X \otimes_{\mathbb{Z}} K$ be the vector space of $K$–valued functions on $S$ invariant under conjugation by $N$. For all $\omega \in \Omega_K$, set $p_\omega = \Sigma_{x \in \omega} x$. Note that as all the characters in $S'_K$ are absolutely irreducible, the set $(p_\omega)_{\omega \in \Omega_K}$ is a basis of $X_K$.

Let $\Phi : X_K \to X_K$ be $\text{Res}_K^G \text{Ind}_G^S$ considered as an endomorphism of $X_K$. Note that we have

$$\Phi(p_\omega) = (\sharp\omega)\Sigma_{\omega \in \Omega_K} d_{\omega,\omega'} p_{\omega'}.$$ 

This implies that the matrix of $\Phi$ in the basis $(p_\omega)_{\omega \in \Omega_K}$ is equal to $(\sharp\omega d_{\omega,\omega'})$.

On the other hand, the odd determinant property implies that the determinant of $\Phi : X_K \to X_K$ is odd (cf. §2). Hence the determinant of $\Phi : X_K \to X_K$ is also odd. Note that $\sharp\omega$ is odd for all $\omega \in \Omega$ (see lemma 6.1). This implies that $\det_{\omega,\omega' \in \Omega_K}(d_{\omega,\omega'})$ is odd, hence the lemma is proved.

Proof of prop. 7.4 Note that for any field $E$ and any $\omega, \omega' \in \Omega_E$, we have

$$d_{\omega,\omega'} = \langle x, \text{Res}_E^G \text{Ind}_G^S x' \rangle_S = \langle \text{Ind}_G^S x, \text{Ind}_G^S x' \rangle_G$$

for any $x \in \omega, x' \in \omega'$.

If char($k$) = 0, then the proposition follows from lemma 7.5. Suppose that char($k$) > 0. Let $A$ be a complete discrete valuation ring of characteristic 0 with residue field $k$, and let $\pi$ be a uniformizer of $A$. Let $K$ be the field of fractions of $A$. Then all the characters in $S'_K$ are absolutely irreducible. Indeed, we have $k[S] = \prod_{1 \leq i \leq r} M_{n_i}(k)$, where $r$ is the number of irreducible representations of $S$ over $k$. Since $A[S]$ is complete with respect to the ideal $\pi A[S]$, the isomorphism $A[S]/\pi A[S] \to \prod_{1 \leq i \leq r} M_{n_i}(k)$ can be lifted to an isomorphism $A[S] \simeq \prod_{1 \leq i \leq r} M_{n_i}(A)$, hence we have $K[S] \simeq \prod_{1 \leq i \leq r} M_{n_i}(K)$. Thus every character of $S'_K$ is absolutely irreducible, hence by lemma 7.5 we have

$$\det_{\omega,\omega' \in \Omega_K}(d_{\omega,\omega'}) \equiv 1 \pmod{2}.$$ 

Let us show that the matrices $(d_{\omega,\omega'})_{\omega,\omega' \in \Omega_K}$ are equal for suitable orderings of the sets $\Omega_K$ and $\Omega_K$. As $S$ is a 2–group and char($k$) ≠ 2, every $k[S]$–module is projective. If $P$ is a projective $k[S]$–module, then $\text{Ind}_G^S(P)$ is projective as well.

Let $P$ be a projective $k[S]$–module. Since $A[S]$ is $\pi$–adically complete, there is a projective $A[S]$–module $\tilde{P}$ such that $\tilde{P}/\pi \tilde{P} \simeq P$. Then $\tilde{P}_K = \tilde{P} \otimes_A K$ is...
E. Bayer-Fluckiger and R. Parimala

a projective $K[S]$-module. Moreover, $P$ is simple if and only if $\tilde{P}_K$ is simple. Note that if $P$ and $Q$ are simple $k[S]$-modules, then we have

$$(\text{Ind}_G^S(P), \text{Ind}_G^S(Q))_G = (\text{Ind}_G^S\tilde{P}_K, \text{Ind}_G^S\tilde{Q}_K)_G.$$ 

Therefore the matrices $(d_{\omega,\omega'})_{\omega,\omega'\in\Omega_k}$ and $(d_{\omega,\omega'})_{\omega,\omega'\in\Omega_K}$ are equal for suitable orderings of the sets $\Omega_k$ and $\Omega_K$, and this completes the proof of the proposition.

§8. Odd determinant property–behavior under quadratic extension

This section contains a quadratic descent argument. Together with a filtration introduced in §9, this quadratic descent will enable us to reduce to the case where all the characters are absolutely irreducible, cf. §7. Putting these informations together in §10, we obtain a result (th. 10.1) that will be used in §11 to prove th. 2.2. We start by recalling and introducing some notation that will be needed in this section and the next ones.

Let $G$ be a finite group and let $S$ be a 2-Sylow subgroup of $G$. For any field $E$ with char($E$) $\neq 2$, we denote by $S'_E$ the set of irreducible characters of $S$ over $E$, and by $\Omega_E$ be the set of orbits of $S'_E$ under the action of $N = N_G(S)$. Recall that $\iota: E[S] \to E[S]$ is the standard involution, and that for $x \in S'_E$ we denote $\overline{x} = x$ if $x$ is selfdual, and $\overline{x} = (x, x')$ if $\iota(x) = x' \neq x$.

For any $\omega \in \Omega_E$, recall that $\overline{\omega} = \omega$ if the characters of $\omega$ are invariant under $\iota$, and $\overline{\omega} = (\omega_1, \omega_2)$ if there exist $x_1 \in \omega_1$ and $x_2 \in \omega_2$ such that $\iota(x_1) = x_2$ with $x_1 \neq x_2$. Let $\overline{\Omega}_E$ be the set of all $\overline{\omega}$ with $\omega \in \Omega_E$, and let $K^0_{\overline{\omega}} = K^0_{\omega}$ for $\overline{\omega} \in \overline{\Omega}$.

Let us define $d_{\omega,\omega'}$ as being the number of times $U_y$ occurs in

$$\text{Res}_G^S \text{Ind}_G^S(U_x)$$

for $x \in \omega$, $y \in \omega'$.

Let us recall that for all $\omega, \omega' \in \Omega_k$ such that $K^0_{\omega} = K^0_{\omega'} = k$, we denote by $d_{\overline{\omega}, \overline{\omega'}}$ the dimension of the $k$–vector space underlying the quadratic form $F_{\overline{\omega}, \overline{\omega'}}$ (see §6).

Set

$$\Omega^0_E = \{\omega \in \Omega_E \mid K^0_{\omega} = E\}$$

$$\overline{\Omega}^0_E = \{\overline{\omega} \in \overline{\Omega}_E \mid K^0_{\overline{\omega}} = E\}$$

$$\Omega^0_k = \{\omega \in \Omega^0_E \mid \omega \text{ orthogonal or symplectic}\}$$

Documenta Mathematica 16 (2011) 677–707
$$\Omega^1_o = \{ \omega \in \Omega^0_E \mid \omega \text{ orthogonal} \}$$

$$\Omega^1_s = \{ \omega \in \Omega^0_E \mid \omega \text{ symplectic} \}$$

$$\Omega^2_E = \{ \omega \in \Omega^0_E \mid \omega \text{ unitary} \}$$

$$\Omega^2_E = \{ \overline{\omega} \in \Omega^1_E \mid \overline{\omega} \text{ unitary} \}$$

and

$$\delta^1_E = \det_{\omega, \omega' \in \Omega^1_E} (d_{\omega, \omega'})$$

$$\delta^2_E = \det_{\omega, \omega' \in \Omega^2_E} (d_{\omega, \omega'})$$

$$\delta^2_E = \det_{\overline{\omega}, \overline{\omega'} \in \Omega^2_E} (d_{\overline{\omega}, \overline{\omega'}})$$

$$\delta^1_o = \det_{\omega, \omega' \in \Omega^1_o} (d_{\omega, \omega'})$$

$$\delta^1_s = \det_{\omega, \omega' \in \Omega^1_s} (d_{\omega, \omega'})$$

Let $L/K$ be a quadratic extension, and let $\tau : L \to L$ be the non-trivial automorphism of $L/K$. Then $\tau$ acts on $S'_L$ by $(\tau x)(s) = \tau(x(s))$ for all $s \in S$ and $x \in S'_L$. This induces an action of $\tau$ on $\Omega_L$.

**Proposition 8.1** Let $\omega \in \Omega_L$. Then $\tau \omega = \omega$ if and only if there is a character $x \in S'_L$ with $x \in \omega$ such that $\tau x = x$.

**Proof.** If there exists $x \in \omega$ such that $\tau x = x$, then we have $\tau \omega = \omega$. Conversely, suppose that $\omega \in \Omega_L$ is such that $\tau \omega = \omega$. If we had $\tau x \neq x$ for every $x \in \omega$, then $\tau^n(\omega)$ would be even, contradicting lemma 6.1. This implies that there exists $x \in \omega$ with $\tau x = x$, hence the proposition is proved.

Note that $\tau$ acts on the center of $L[S]$, and that the action of $\tau$ on $S'_L$ can be described in terms of this action. This leads to the following observation, which will be used in the sequel:

**Lemma 8.2** Let $x \in S'_L$ be an orthogonal or symplectic character such that $\tau x = x$. Let $L = K_x$. Then

(i) There exists $x_0 \in S'_K$ such that $K_{x_0} = K$. 

*Documenta Mathematica 16 (2011) 677–707*
Proof. For any field $E$, let us denote by $Z(E[S])$ the center of $E[S]$. The Galois automorphism $\tau : L \to L$ over $K$ acts on $L[S]$, hence also on $Z(L[S])$. Then the subalgebra of $Z(L[S])$ fixed by $\tau$ is equal to $Z(K[S])$. The hypothesis implies that $L$ is one of the factors in the decomposition of $Z(L[S])$. Note that the restriction of $\tau$ to the factor $L$ in $Z(L[S])$ is non–trivial, and that the fixed field is equal to $K$. This corresponds to a factor in the decomposition of $K[S]$, and hence to a character $x_0$ of $S'_K$. This proves (i). Noting that the base change to $L$ of the factor corresponding to $x_0$ in $K[S]$ is the factor corresponding to $(x_0)_L$, points (ii) and (iii) are immediate. Suppose now that $x$ is symplectic. Then the same reasoning proves that if $\Delta_{x_0}$ is a quaternion division algebra and $\Delta_x = L$, then $(x_0)_L = 2x$; if both $\Delta_{x_0}$ and $\Delta_x$ are quaternion division algebras, or if $\Delta_{x_0} = K$ and $\Delta_x = L$, then $(x_0)_L = x$. This proves (iv).

Corollary 8.3 Let $\omega \in \Omega_L^1$ be such that $\tau \omega = \omega$.

(i) If $\omega \in \Omega_L^{1, o}$, there exists $\omega_0 \in \Omega_K^{1, o}$ such that $(\omega_0)_L = \omega$.

(ii) If $\omega \in \Omega_L^{1, s}$, then there exists $\omega_0 \in \Omega_K^{1, s}$ such that $(\omega_0)_L = \omega$ or $(\omega_0)_L = 2\omega$.

Proof. By prop. 8.1 we can choose $x \in \omega$ such that $\tau x = x$. Let $x_0 \in S'_K$ such that $K_{x_0} = K$ (see lemma 8.2 (ii)). Hence we have $K_{(x_0)} \otimes_K L = L = K_x$. (i) Suppose that $\omega \in \Omega_L^{1, o}$. Then $x$ is orthogonal. Hence $x_0$ is orthogonal, and $(x_0)_L = x$ (cf. 8.2 (ii) and (iii)). Let $\omega_0$ be the orbit of $x_0$; then $\omega_0 \in \Omega_K^{1, o}$ and $(\omega_0)_L = \omega$.

(ii) Suppose that $\omega \in \Omega_L^{1, s}$. Then $x$ is symplectic. Hence $x_0$ is symplectic, and $(x_0)_L = x$ or $(x_0)_L = 2x$ (cf. 8.2 (ii) and (iv)). Let $\omega_0$ be the orbit of $x_0$. Then $\omega_0 \in \Omega_K^{1, s}$ has the required property.

Proposition 8.4 Suppose that

$$\delta_{L, o}^1 \equiv 1 \pmod{2}.$$ 

Then

$$\delta_{K, o}^1 \equiv 1 \pmod{2}.$$ 

Proof. The automorphism $\tau$ of $L/K$ induces a permutation of $\Omega_L^{1, o}$. Set

$$\Omega_{L, 1} = \{ \omega \in \Omega_L^{1, o} | \tau \omega = \omega \}.$$ 

Documenta Mathematica 16 (2011) 677–707
and

\[ \Omega_{L}^{1,2} = \{ \omega \in \Omega_{L}^{1,0} | \tau \omega \neq \omega \}. \]

Let \( S_L \) be the group of permutations of \( \Omega_{L}^{1,0} \), and let \( S_L^1 \) respectively \( S_L^2 \) be the group of permutations of \( \Omega_{L}^{1,1} \), respectively \( \Omega_{L}^{1,2} \), regarded as subgroups of \( S_L \).

Set

\[
\alpha = \sum_{s \in S_L^1} \left( \prod_{\omega \in \Omega_{L}^{1,1}} d_{s,\omega},\omega \right),
\]

\[
\beta = \sum_{s \in S_L^2} \left( \prod_{\omega \in \Omega_{L}^{1,2}} d_{s,\omega},\omega \right),
\]

By lemma 4.3, we have

\[ \delta_{L}^{1,o} \equiv \alpha \beta \pmod{2}. \]

On the other hand, we have

\[ \delta_{K}^{1,o} \equiv \alpha \pmod{2}. \]

Indeed, by cor. 8.3 (i) the map \( \omega_0 \mapsto (\omega_0)_L \) induces a bijection between \( \Omega_{K}^{1,o} \) and \( \Omega_{L}^{1,1} \) with \( d_{\omega_0,\omega_0'} = d_{(\omega_0)_L,(\omega_0')_L} \) for \( \omega_0,\omega_0' \in \Omega_{K}^{1,o} \). It follows that \( \delta_{K}^{1,o} \equiv \alpha \pmod{2} \). This completes the proof of the proposition.

**Proposition 8.5** Suppose that

\[ \delta_{L}^{1,s} \equiv 1 \pmod{2}. \]

Then

\[ \delta_{K}^{1,s} \equiv 1 \pmod{2}. \]

**Proof.** The automorphism \( \tau \) of \( L/K \) induces a permutation of \( \Omega^{1,s} \). Let

\[ \Omega_{L}^{1,1} = \{ w \in \Omega_{L}^{1,s} | \tau w = \omega \}, \]

\[ \Omega_{L}^{1,2} = \{ w \in \Omega_{L}^{1,s} | \tau w \neq \omega \}. \]

Let \( S_L \) be the group of permutations of \( \Omega_{L}^{1,s} \), and let \( S_L^1 \) respectively \( S_L^2 \) be the group of permutations of \( \Omega_{L}^{1,1} \), respectively \( \Omega_{L}^{1,2} \), regarded as subgroups of \( S_L \).

Set

\[
\alpha = \sum_{s \in S_L^1} \left( \prod_{\omega \in \Omega_{L}^{1,1}} d_{s,\omega},\omega \right),
\]
\[ \gamma = \Sigma_{s \in S_L^1} \left( \prod_{\omega \in \Omega_{L,2}} d_{s,\omega} \right). \]

Arguing as in 8.4, we get
\[ \delta_{1,s}^1 \equiv \gamma \alpha \pmod{2}. \]

CLAIM. We have
\[ \delta_{1,s}^1 \equiv \alpha \pmod{2}. \]

Let us write
\[ \Omega_{L,1} = \Omega_{L/K}^{s} \cup \Omega_{L/K}^{ns}, \]
where
\[ \Omega_{L/K}^{s} = \{ \omega \in \Omega_{L,1} \mid \text{there exists } \omega_0 \in \Omega_{K}^{1,s} \text{ with } (\omega_0)_L = 2\omega \}, \]
and
\[ \Omega_{L/K}^{ns} = \{ \omega \in \Omega_{L,1} \mid \text{there exists } \omega_0 \in \Omega_{K}^{1,s} \text{ with } (\omega_0)_L = \omega \}. \]

By corollary 8.3 (ii) the above is a disjoint union decomposition.

For \( \omega \in \Omega_{L/K}^{s} \) with \( 2\omega = (\omega_0)_L \) and \( \omega' \in \Omega_{L/K}^{ns} \) with \( \omega' = (\omega'_0)_L \), \( \omega_0, \omega'_0 \in \Omega_K \), we have:
If \( d_{\omega_0,\omega_0} = r \), then \( d_{\omega',\omega} = 2r \).

Thus the matrix \( (d_{\omega,\omega'})_{\omega,\omega' \in \Omega_{L,1}^1} \) is congruent to the matrix
\[ \begin{pmatrix} A & * \\ 0 & B \end{pmatrix} \]
where \( A = (d_{\omega,\omega'})_{\omega,\omega' \in \Omega_{L/K}^{s}} \) and \( B = (d_{\omega,\omega'})_{\omega,\omega' \in \Omega_{L/K}^{ns}} \). Therefore
\[ \alpha \equiv \det(A)\det(B) \pmod{2}. \]

We next determine \( \delta_{1,s}^1 \pmod{2} \). Let us write
\[ \Omega_{K,1}^{s} = \Omega_{K/L}^{s} \cup \Omega_{K/L}^{ns}, \]
with
\[ \Omega_{K/L}^{s} = \{ \omega \in \Omega_K \mid \text{there exists } \omega_0 \in \Omega_{L,1} \text{ with } (\omega)_L = 2\omega_0 \}, \]
and
\[ \Omega_{K/L}^{ns} = \{ \omega \in \Omega_K \mid \text{there exists } \omega_0 \in \Omega_{L,1} \text{ with } (\omega)_L = \omega_0 \}. \]
For $\omega, \omega' \in \Omega^s_{K/L}$, if $\omega_L = 2\omega_0$, $\omega'_L = 2\omega'_0$ for some $\omega_0, \omega'_0 \in \Omega^s_{K/L}$, we have $d_{\omega,\omega'} = d_{\omega_0,\omega'_0}$. Also, for $\omega, \omega' \in \Omega^s_{K/L}$, if $\omega_L = \omega_0$, $\omega'_L = \omega'_0$ for some $\omega_0, \omega'_0 \in \Omega^s_{K/L}$, we have $d_{\omega,\omega'} = d_{\omega_0,\omega'_0}$. Thus the matrix $(d_{\omega,\omega'})_{\omega,\omega' \in \Omega^s_{K/L}}$ is equal to $A$, and the matrix $(d_{\omega,\omega'})_{\omega,\omega' \in \Omega^s_{K/L}}$ is equal to $B$. Further, if $\omega \in \Omega^s_{K/L}$, then $\omega_L = 2\omega_0$ for some $\omega_0 \in \Omega^s_{L}$, and if $\omega' \in \Omega^{s_s}_{K/L}$, then $\omega'_L = \omega'_0$ for some $\omega'_0 \in \Omega^{s_s}_{L}$. If $d_{\omega_0,\omega'_0} = r$, then $d_{\omega,\omega'} = 2r$. Thus the matrix $(d_{\omega,\omega'})_{\omega,\omega' \in \Omega^s_{K/L}}$ is congruent to

$$\begin{pmatrix} A & 0 \\ * & B \end{pmatrix}.$$ mod 2. Therefore

$$\delta_{K}^{s} = \det(A)\det(B) \equiv \alpha \pmod{2},$$

and this completes the proof of the claim. Therefore we see that $\delta_{K}^{s} \equiv 1 \pmod{2}$ implies $\delta_{K}^{s} \equiv 1 \pmod{2}$, hence prop. 8.5 is proved.

Let us recall that for any field $E$ of characteristic $\neq 2$, we denote by $\overline{E}$ the set of $E$ for $x \in S'_K$.

The Galois automorphism $\tau : L \to L$ over $K$ induces an action on $\overline{S}_L$ which we denote by $\overline{\tau} \overline{x}$.

**Lemma 8.6** Let $\overline{y} \in \overline{S}_L$ with $\overline{y}$ unitary such that $\tau \overline{y} = \overline{y}$. Then there exists $\overline{\tau} \in \overline{S}_K$ with $\overline{\tau} \overline{S}_L = \overline{y}$. Moreover, if $K_{\overline{y}} = L$, then $K_{\overline{\tau}} = K$.

**Proof.** Suppose that $\overline{y} = y$ and $\tau y = y$. Then by the method of lemma 8.2 we see that there is a unitary character $x \in S'_K$ such that $x_L = y$. Moreover, $\tau$ restricts to a non-trivial automorphism of $K_y$ which commutes with $\iota$. If $E = (K_y)^{\iota}$, then $\tau E$ is non-trivial and $K_x = E$. Since $Kx_L = EL = K_y$, we have $x_L = y$. Further, if $K_{\overline{y}} = L$, then $K_{\overline{\tau}} = K$.

Suppose that $\overline{y} = (y, y)$ with $\iota y \neq y$ and $\tau \overline{y} = \overline{y}$. Then $\tau y = y$ or $\tau y = y$. Set $M = K_y \times K_{\iota y}$.

Suppose first that $\tau y = y$. Then there is an $x \in S'_K$ such that $x_L = y$. Further, $\tau$ induces an automorphism on $M$ which takes each factor $K_y$ and $K_{\iota y}$ in itself. Moreover, we have $M^\tau = K_x \times K_{\iota x}$. Thus $\overline{\tau} = (x, \iota x) \in \overline{S}_K$ is unitary with $\overline{\tau} \overline{L} = \overline{y}$. Moreover, if $K_{\overline{y}} = M^\iota = L$, then $K_{\overline{\tau}} = K$.

Suppose now that $\tau y = y$. Then $\tau y = y$, and $\tau$ switches the factors $K_y$ and $K_{\iota y}$ of $M$. Let $E = M^\tau$. Then $E$ is a field which is a factor of the center of $K[S]$ and $\iota$ restricted to $E$ is non-trivial. Let $x \in S'_K$ be the character associated...
to \( E \). Then \( x \) is unitary, \( K_x = E \) and \( K_{x_L} = EL = M = K_y \times K_{y'} \). Thus \( x_L = (y, iy) \). Further, if \( K_0^y = M^L = K_y \), then \( K_0^y = E^L = K \), where \( \overline{\pi} = x \).

The automorphism \( \tau \) induces an action on \( \Omega^2_L \) that we denote by \( \overline{\omega} \mapsto \tau \overline{\omega} \).

**Corollary 8.7** Let \( \overline{\omega} \in \Omega^2_L \) be a unitary orbit with \( \tau \overline{\omega} = \overline{\omega} \). Then there is a unitary orbit \( \overline{\omega}_0 \in \Omega^2_K \) such that \( (\overline{\omega}_0)_L = \overline{\omega} \).

**Proof.** Suppose that \( \overline{\omega} = (\omega, \omega') \) with \( \omega' \neq \omega \). Then \( \tau \omega = \omega \) implies that \( \tau' \omega = \omega' \) or \( \tau' \omega = \omega \).

Suppose first that \( \tau' \omega = \omega \). Then by proposition 8.1 there is a \( y \in \Sigma_L^I \) belonging to \( \omega \) with \( \tau y = y \). Further, \( \overline{\omega} = (y, iy) \) is unitary with \( \tau \overline{\omega} = \overline{\omega} \). In this case, we appeal to lemma 8.6 to conclude the proof.

Suppose now that \( \tau' \omega = \omega' \) and that \( \tau' \omega \neq \omega \). Then \( \tau \omega \) induces an action on the characters in \( \omega \). As \( \sharp \omega \) is odd by lemma 6.1, there exists \( y \in \omega \) such that \( \tau y = y \) and \( \tau y \neq y \), since \( \tau y \neq \omega \). Then \( \overline{\omega} = (y, iy) \) is a unitary pair with \( \tau \overline{\omega} = \overline{\omega} \) and the proposition follows from lemma 8.6.

**Proposition 8.8** If \( \det_{\overline{\omega} \in \Omega^2_L} (d_{\overline{\omega} \overline{\omega}'}) \equiv 1 \pmod{2} \), then \( \det_{\overline{\omega} \in \Omega^2_K} (d_{\overline{\omega} \overline{\omega}'}) \equiv 1 \pmod{2} \).

**Proof.** Recall that \( \tau : L \to L \) is the non-trivial automorphism of \( L/K \). Let us write \( \Omega^2_L = \Omega^2_{L,1} \cup \Omega^2_{L,2} \), where

\[
\Omega^2_{L,1} = \{ \overline{\omega} \in \Omega^2_L | \tau \overline{\omega} = \overline{\omega} \},
\]

and

\[
\Omega^2_{L,2} = \{ \overline{\omega} \in \Omega^2_L | \tau \overline{\omega} \neq \overline{\omega} \}.
\]

Arguing as in 8.4 and using 4.3, we get

\[
\det_{\overline{\omega} \in \Omega^2_L} (d_{\overline{\omega} \overline{\omega}'}) = [\det_{\overline{\omega} \in \Omega^2_{L,1}} (d_{\overline{\omega} \overline{\omega}'})][\det_{\overline{\omega} \in \Omega^2_{L,2}} (d_{\overline{\omega} \overline{\omega}'})].
\]

Note that scalar extension induces a bijection between \( \Omega^2_K \) and \( \Omega^2_{L,1} \), and we have \( d_{\overline{\omega} \overline{\omega}'} = d_{\overline{\omega'} \overline{\omega}} \) for \( \overline{\omega}, \overline{\omega}' \in \Omega^2_K \). This proves the proposition.

§9. A filtration

Let \( k \) be a field of characteristic \( \neq 2 \), let \( G \) be a finite group and let \( S \) be a 2-Sylow subgroup of \( G \). In this section, we introduce a quadratic filtration of the field \( k \) that will be needed in the next two sections.
Let \( \kappa \) be the prime field of \( k \), that is, \( \kappa = \mathbb{Q} \) if \( \text{char}(k) = 0 \) and \( \kappa = \mathbb{F}_p \) if \( \text{char}(k) = p > 0 \). Note that \( k[S] = \kappa[S] \otimes_{\kappa} k \), hence it is interesting to investigate the structure of \( \kappa[S] \) in both cases.

Suppose first that \( \kappa = \mathbb{Q} \). We have the following lemma:

**Lemma 9.1** Let \( S \) be a 2–group, and let \( \mathbb{Q}[S] = \prod_{i=1}^{r} M_{n_i}(D_i) \) where the \( D_i \)'s are division algebras, and let \( Z(D_i) = K_i \). Let \( \iota : Q[S] \to Q[S] \) be the standard involution. Then each component of \( Q[S] \) is invariant under \( \iota \). Let us denote by \( K^0_i \) the invariant subfield of \( K_i \) under the restriction of \( \iota \) to \( K_i \). Then there exists \( m \in \mathbb{N} \) such that \( K^0_i \) is a subfield of the real 2–cyclotomic subfield \( \mathbb{Q}(\zeta_{2m} + \zeta_{2m}^{-1}) \).

**Proof.** The fact that each component of \( Q[S] \) is invariant under \( \iota \) follows from [6], Chap. 8, 13.2. We know that as \( S \) is a 2–group, there exists \( m \in \mathbb{N} \) such that for all \( i = 1, \ldots, r \) the field \( K_i \) is a subfield of the cyclotomic field \( \mathbb{Q}(\zeta_{2m}) \). The standard involution \( \iota : Q[S] \to Q[S] \) is positive definite, hence its restriction to each component is positive definite as well. This implies (cf [6], Chap 8, 13.5) that \( K^0_i \subset R \) for all \( i \). Hence for all \( i = 1, \ldots, r \), we have \( K^0_i \subset Q(\zeta_{2m} + \zeta_{2m}^{-1}) \) as claimed.

With the notation of lemma 9.1, let \( L = Q(\zeta_{2m} + \zeta_{2m}^{-1}) \). Since \( L/Q \) is cyclic of degree a power of 2, it has a unique set of subfields which fit into a filtration

\[
L_0 = Q \subset L_1 \subset L_2 \subset \cdots \subset L_s = L
\]

with all inclusions being strict, and \( L_i/L_{i-1} \) of degree 2.

Suppose now that \( \kappa = \mathbb{F}_p \) for some prime number \( p \). We have

\[
F_p[S] = \prod_{i=1}^{r} M_{n_i}(K_i)
\]

where the \( K_i \)'s are finite degree extensions of \( F_p \). As \( S \) is a 2–group, the degrees of these extensions are powers of 2. There exists a finite extension \( L/F_p \) of degree a power of 2 containing all the \( K_i \)'s. Note that as \( F_p \) is a finite field, the extension \( L/F_p \) is cyclic. Hence in this case too, we have a unique set of subfields of \( L \) which fit into a filtration

\[
L_0 = Q \subset L_1 \subset L_2 \subset \cdots \subset L_s = L
\]

with all inclusions being strict, and \( L_i/L_{i-1} \) of degree 2.

Let

\[
k_0 = k \subset k_1 \subset k_2 \subset \cdots \subset k_t = Lk
\]
be the induced strict filtration of $Lk/k$. Note that every subfield of $Lk$ containing $k$ is one of the fields $k_i$. Let $k_r$ be the smallest of these fields containing $K_0^g$ for all $g \in S'_k$.

§10. THE ODD DETERMINANT PROPERTY REVISITED

For any field $E$, set

$$
\overline{\delta}_E^{1,o} = \det_{\omega, \omega' \in \Omega} \omega \left( d_{\omega, \omega'} \right)
$$

$$
\overline{\delta}_E^{1,s} = \det_{\omega, \omega' \in \Omega} \omega \left( d_{\omega, \omega'} \right)
$$

$$
\overline{\delta}_E^{2} = \det_{\omega, \omega' \in \Omega} \omega \left( d_{\omega, \omega'} \right)
$$

$$
d_E^0 = \det_{\omega, \omega' \in \Omega} \omega \left( d_{\omega, \omega'} \right).
$$

The result below will be instrumental in the proof of th. 2.2 in the next section:

**Theorem 10.1** Let $G$ be a finite group having the odd determinant property. Then for any field $K$ of characteristic not 2, we have

$$
d_K^0 \equiv 1 \pmod{2}.
$$

**Proof.** We first treat the case where all the characters in $S'_K$ are absolutely irreducible. The reduction to this case is via the filtration introduced in §9, and the quadratic descent of §8.

Suppose first that all the characters in $S'_K$ are absolutely irreducible. For $x \in S'_K$, the form $\rho_x$ is supported on $U_x$ if $x$ is orthogonal, on $U_x \oplus U_x$ if $x$ is symplectic, and $U_{x_1} \oplus U_{x_2}$ if $x = (x_1, x_2)$ with $\iota(x_1) = x_2$ and $x_1 \neq x_2$. Noting that for a general $K$, the integers $d_{\omega, \omega'}$ can be computed after base changing to an algebraic closure of $K$, we get the following:

1) $d_{\omega, \omega'} = d_{\omega, \omega'}$ if $\omega, \omega' \in \Omega_X^{g,s}$;
2) $d_{\omega, \omega'} = 2d_{\omega, \omega'}$ if $\omega$ is symplectic and $\omega'$ is not symplectic;
3) $d_{\omega, \omega'} = 2d_{\omega, \omega'}$ if $\omega$ is unitary and $\omega'$ is orthogonal;
4) $d_{\omega, \omega'} = d_{\omega, \omega'}$ if $\omega, \omega'$ are orthogonal.

Thus the matrix $(d_{\omega, \omega'})_{\omega, \omega' \in \Omega}$ has the following shape modulo 2.
where

\[
\begin{pmatrix}
A & * & *\\
0 & B & *\\
0 & 0 & C
\end{pmatrix},
\]

Thus \(d^0_K \equiv 1 \pmod{2}\) if and only if \(\delta^1_{K^0} = \text{det}(A) \equiv 1 \pmod{2}\), \(\delta^1_{K^s} = \text{det}(C) \equiv 1 \pmod{2}\), and \(\delta^2_K = \text{det}(B) \equiv 1 \pmod{2}\). We also note that for \(\omega, \omega' \in \Omega_{K^0}\), or for \(\omega, \omega' \in \Omega_{K^s}\), we have \(d_{\omega, \omega'} = d_{\omega', \omega}\). Therefore \(\delta^1_{K^0} = \delta^1_{K^s}\), and \(\delta^1_{K^s} = \delta^1_{K^s}\). Thus \(d^1_K \equiv 1 \pmod{2}\) if and only if \(\delta^1_{K^0} \equiv 1 \pmod{2}\), \(\delta^1_{K^s} \equiv 1 \pmod{2}\), and \(\delta^2_K \equiv 1 \pmod{2}\).

There exists a field extension \(L/K\) and a filtration by quadratic extensions

\[
K \subset K_2 \subset \ldots \subset K_n = L
\]
such that all characters in \(S'_L\) are absolutely irreducible (cf. §9). By prop. 7.2 and 7.3, we have \(\delta^1_{K^0} \equiv 1 \pmod{2}\), \(\delta^1_{K^s} \equiv 1 \pmod{2}\), and \(\delta^2_K \equiv 1 \pmod{2}\). By the quadratic descent results 8.4, 8.5 and 8.8, we get \(\delta^1_{K^0} = \text{det}(A) \equiv 1 \pmod{2}\), \(\delta^1_{K^s} = \text{det}(C) \equiv 1 \pmod{2}\), and \(\delta^2_K = \text{det}(B) \equiv 1 \pmod{2}\). Therefore \(d^0_K \equiv 1 \pmod{2}\).

§11. Proof of the induction–restriction result

The aim of this section is to prove th. 2.2. Let

\[
k_0 = k \subset k_1 \subset k_2 \subset \ldots \subset k_r
\]

be the filtration introduced in §9, \(k_r\) being the smallest of these fields containing \(K_{\pi}^n\) for all \(\pi\).

Proof of theorem 2.2 Let \((V, h)\) be an \(S\)–quadratic form. We have a decomposition (cf. §5)

\[
(V, h) \simeq \bigoplus_{\pi} \pi \otimes K_{\pi}^n (V_{\pi}, h_{\pi}),
\]

Documenta Mathematica 16 (2011) 677–707
where \((V_\varpi, g_\varpi)\) is a quadratic form over \(K_\varpi^0\) (cf. §5). Recall that the Witt class

\[ (V_\varpi, g_\varpi) \otimes n_\varpi \in W(K_\varpi^0) \]

is uniquely determined by \((V, h)\), where \(n_\varpi\) is the reduced norm of \(D_\varpi\) over \(K_\varpi^0\) if \(D_\varpi\) is a quaternion algebra, the norm of \(K_\varpi\) over \(K_\varpi^0\) if \(K_\varpi\) is a quadratic algebra, and \(n_\varpi = 1\) otherwise. We have

\[ \text{Ind}_G^S(V, h) = \bigoplus_{\varpi \in \Omega} I(\varpi) \otimes_{K_\varpi^0} (V_\varpi, g_\varpi), \]

where

\[ (V_\varpi, g_\varpi) = \bigoplus_{\varpi \in \Omega} (V_\varpi, g_\varpi) \]

is a quadratic space determined up to multiplication by \(n_\varpi = n_\varpi\). We have \(I(\varpi) = \text{Ind}_G^S(\rho_\varpi)\), which does not depend on the choice of \(\varpi \in \Omega\).

We have

\[ \text{Res}_G^S \text{Ind}_G^S(V, h) = \bigoplus_{\varpi \in \Omega} \text{Res}_S^G(I(\varpi)) \otimes_{K_\varpi^0} (V_\varpi, g_\varpi). \]

For \(\overline{\varpi} \in \overline{S}_k\), the \(\overline{\varpi}\)-component of \(\text{Res}_G^S(I(\varpi))\) is \(\rho_\varpi \otimes_{K_\varpi^0} F_{\overline{\varpi}^0} \overline{\varpi}\), where \(\overline{\varpi} \in \overline{\varpi}'\), and where \(F_{\overline{\varpi}^0} \overline{\varpi}\) is a quadratic space over \(K_\varpi^0\), determined up to multiplication by \(n_\varpi\).

Let \((V_1, h_1)\) and \((V_2, h_2)\) be two \(S\)-quadratic forms such that \(\text{Res}_G^S \text{Ind}_G^S(V_1, h_1) \simeq_S \text{Res}_G^S \text{Ind}_G^S(V_2, h_2)\).

Let

\[ (V_1, h_1) \simeq \bigoplus_{\varpi \in S} \rho_\varpi \otimes_{K_\varpi^0} (V_1^\varpi, g_1^\varpi) \text{ and } (V_2, h_2) \simeq \bigoplus_{\varpi \in S} \rho_\varpi \otimes_{K_\varpi^0} (V_2^\varpi, g_2^\varpi) \]

and

\[ (V_\varpi, g_\varpi) = \bigoplus_{\varpi \in \Omega} (V_\varpi, g_\varpi), \]

for \(i = 1, 2\).

Note that as the \(k[S]\)-modules \(\text{Res}_S^G \text{Ind}_G^S(V_1)\) and \(\text{Res}_S^G \text{Ind}_G^S(V_2)\) are isomorphic, the \(k[G]\)-modules \(\text{Ind}_G^S(V_1)\) and \(\text{Ind}_G^S(V_2)\) are also isomorphic (see for instance [3], cor. 6.8). This implies that \(\dim(V_\varpi^1) = \dim(V_\varpi^2)\) for all \(\varpi \in \Omega\).

**Claim.** We have

\[ n_\varpi \otimes_k (V_\varpi^1, g_\varpi^1) \simeq n_\varpi \otimes_k (V_\varpi^2, g_\varpi^2). \]
For the proof, we distinguish two cases

**Case 1.** Suppose that \( K^0_{\widehat{\mathfrak{m}}} = k \) for all \( \widehat{\mathfrak{m}} \in S'_k \).

Then we have \( \text{Res}_{S}^G(I(\overline{\mathfrak{m}}))_{\overline{\mathfrak{m}}} = \rho_{\overline{\mathfrak{m}}} \otimes_k F_{\overline{\mathfrak{m}}} \mathbb{W}, \) where \( \overline{\mathfrak{m}} \in \overline{\mathfrak{m}}, \) where \( F_{\overline{\mathfrak{m}}} \mathbb{W} \) is a quadratic form over \( k, \) and \( n_{\overline{\mathfrak{m}}} \otimes_k F_{\overline{\mathfrak{m}}} \mathbb{W} \) is determined by \( \text{Res}_{S}^G(I(\overline{\mathfrak{m}}))_{\overline{\mathfrak{m}}} \). Hence

\[
\text{Res}_{S}^G \text{Ind}_{S}^G(V_i, h_i) = \bigoplus_{\overline{\mathfrak{m}} \in \overline{\mathfrak{m}}} \rho_{\overline{\mathfrak{m}}} \otimes_k \left[ \bigoplus_{\overline{\mathfrak{m}} \in \overline{\mathfrak{m}}} F_{\overline{\mathfrak{m}}} \mathbb{W} \otimes_k (V_{\overline{\mathfrak{m}}}^0, g_{\overline{\mathfrak{m}}}) \right]
\]

Suppose that \( \text{Res}_{S}^G \text{Ind}_{S}^G(V_1, h_1) \simeq \text{Res}_{S}^G \text{Ind}_{S}^G(V_2, h_2), \) and set \( g_{\overline{\mathfrak{m}}}^i = (V_{\overline{\mathfrak{m}}}^0, g_{\overline{\mathfrak{m}}}) \) for \( i = 1, 2. \) Then

\[
n_{\overline{\mathfrak{m}}} \otimes_k \left[ \bigoplus_{\overline{\mathfrak{m}} \in \overline{\mathfrak{m}}} F_{\overline{\mathfrak{m}}} \mathbb{W} \otimes_k g_{\overline{\mathfrak{m}}} \right] \simeq n_{\overline{\mathfrak{m}}} \otimes_k \left[ \bigoplus_{\overline{\mathfrak{m}} \in \overline{\mathfrak{m}}} F_{\overline{\mathfrak{m}}} \mathbb{W} \otimes_k g_{\overline{\mathfrak{m}}} \right].
\]

Let us denote by \( f_{\overline{\mathfrak{m}}} \mathbb{W} \) the element of \( W(k) \) determined by the quadratic form \( F_{\overline{\mathfrak{m}}} \mathbb{W}, \) and let \( (f_{\overline{\mathfrak{m}}} \mathbb{W})_i \) be the matrix of cofactors of the matrix \( (f_{\overline{\mathfrak{m}}} \mathbb{W}) \) in the Witt ring \( W(k). \) Then the product \( (f_{\overline{\mathfrak{m}}} \mathbb{W})(n_{\overline{\mathfrak{m}}} \otimes_k f_{\overline{\mathfrak{m}}} \mathbb{W}) \) is equal to

\[
\varphi \begin{pmatrix} n_{\overline{\mathfrak{m}}_1} & 0 & \cdots & 0 \\
0 & n_{\overline{\mathfrak{m}}_2} & \cdots & 0 \\
0 & \cdots & \cdots & n_{\overline{\mathfrak{m}}_n} \end{pmatrix},
\]

a diagonal matrix with diagonal entries \( \varphi, n_{\overline{\mathfrak{m}}}, \) where \( \varphi \in W(k) \) is the determinant of the matrix \( (f_{\overline{\mathfrak{m}}} \mathbb{W}). \) Let \( \psi_{\overline{\mathfrak{m}}} \) be the element of \( W(k) \) determined by the quadratic form \( g_{\overline{\mathfrak{m}}}^i = (V_{\overline{\mathfrak{m}}}^0, g_{\overline{\mathfrak{m}}}) \) for \( i = 1, 2. \) Then we get

\[
\varphi, n_{\overline{\mathfrak{m}}} \otimes_k \left( \psi_{\overline{\mathfrak{m}}}^1 - \psi_{\overline{\mathfrak{m}}}^2 \right) = 0
\]

in \( W(k), \) for every \( \overline{\mathfrak{m}} \in \overline{\mathfrak{m}}. \) Note that \( \det(\text{dim}(f_{\overline{\mathfrak{m}}} \mathbb{W})) = \det(\text{dim}(f_{\overline{\mathfrak{m}}} \mathbb{W})), \) and that

\[
\det(\text{dim}(f_{\overline{\mathfrak{m}}} \mathbb{W})) = \det(d_{\overline{\mathfrak{m}}} \mathbb{W}) = d_k^0.
\]

Since \( G \) has the odd determinant property, by prop. 10.1 we have \( d_k^0 \equiv 1 \pmod{2}. \) Therefore \( \dim(\varphi) \) is odd, hence \( \varphi \) is not a zero divisor in \( W(k) \) (see for instance [6], 2.6.5). Therefore we have

\[
n_{\overline{\mathfrak{m}}} \otimes_k \left( \psi_{\overline{\mathfrak{m}}}^1 - \psi_{\overline{\mathfrak{m}}}^2 \right) = 0
\]

in \( W(k), \) for all \( \overline{\mathfrak{m}} \in \overline{\mathfrak{m}}, \) and hence \( n_{\overline{\mathfrak{m}}} \otimes_k (V_{\overline{\mathfrak{m}}}^1, g_{\overline{\mathfrak{m}}}^1) \) and \( n_{\overline{\mathfrak{m}}} \otimes_k (V_{\overline{\mathfrak{m}}}^2, g_{\overline{\mathfrak{m}}}^2) \) are in the same Witt class. Recall that \( \dim(V_{\overline{\mathfrak{m}}}^1) = \dim(V_{\overline{\mathfrak{m}}}^2) \) for all \( \overline{\mathfrak{m}} \in \overline{\mathfrak{m}}. \) Hence the quadratic forms \( n_{\overline{\mathfrak{m}}} \otimes_k (V_{\overline{\mathfrak{m}}}^1, g_{\overline{\mathfrak{m}}}^1) \) and \( n_{\overline{\mathfrak{m}}} \otimes_k (V_{\overline{\mathfrak{m}}}^2, g_{\overline{\mathfrak{m}}}^2) \) have the same dimension and are in the same Witt class, therefore we have
for all $\omega \in \Omega$. This completes the proof of the claim in case 1.

**General case.** Let us consider $(V_i, h_i) \otimes_k k_r$. We have

$$\text{Res}^G_S \text{Ind}^G_S(V_1, h_1) \otimes_k k_r \simeq \text{Res}^G_S \text{Ind}^G_S(V_2, h_2) \otimes_k k_r.$$  

Moreover, $K^0 \otimes_k k_r \simeq \prod_{\alpha \in \text{Gal}(K^0/k)} k^0_{\alpha}$. The orbit $\overline{\omega}$ splits into distinct conjugate orbits over $k_r$. Each $\overline{\omega} \in \Omega_k$ with $K^0_{\omega} = k_r$ occurs as one of the conjugate orbits over $k_r$. Using case 1, we get, for orbits $\overline{\omega}$ with $K^0_{\omega} = k_r$,

$$n_{\overline{\omega}} \otimes (V_{1, \overline{\omega}}, g_{1, \overline{\omega}}) \simeq n_{\overline{\omega}} \otimes (V_{2, \overline{\omega}}, g_{2, \overline{\omega}}).$$

Cancelling these factors, we may assume that

$$\text{Ind}^G_S(V, h) = \bigoplus_{\overline{\omega} \in \Omega} I(\overline{\omega}) \otimes_{K^0_{\overline{\omega}}} (V_{\overline{\omega}}, g_{\overline{\omega}})$$

with $K^0_{\overline{\omega}} \subset k_r - 1$ for all $\overline{\omega}$ in the above decomposition. Inductively we get, for all $\overline{\omega}$, that

$$n_{\overline{\omega}} \otimes (V_{1, \overline{\omega}}, g_{1, \overline{\omega}}) \simeq n_{\overline{\omega}} \otimes (V_{2, \overline{\omega}}, g_{2, \overline{\omega}}).$$

This completes the proof of the theorem.

**Bibliography**

[1] E. Bayer–Fluckiger and H.W. Lenstra, Jr., Forms in odd degree extensions and self-dual normal bases, *Amer. J. Math.* 112 (1990), 359-373.

[2] E. Bayer–Fluckiger and J–P. Serre, Torsions quadratiques et bases normales autoduales, *Amer. J. Math.* 116 (1994) 1–64.

[3] D.J. Benson, R.A. Parker, The Green ring of a finite group, *J. Algebra* 87 (1984), 290-331.

[4] E. Lequeu, Hermitian G-Forms and Induction, in preparation.

[5] J. Morales, Integral bilinear forms with a group action, *J. Algebra* 98 (1986), 470 - 484.

[6] W. Scharlau, *Quadratic and Hermitian Forms*, Grundlehren der Math. Wiss, Springer–Verlag (1985).

[7] J–P. Serre, *Linear representations of finite groups*, Graduate Texts in Mathematics, Vol. 42. Springer–Verlag (1977).
[8] J. Thévenaz, Most finite groups are $p$-nilpotent. *Exposition. Math.* 11 (1993), 359 - 363.

E. Bayer–Fluckiger
École Polytechnique
Fédérale de Lausanne
EPFL/Fsb/Mathgeom/CSAG
Station 8
1015 Lausanne, Switzerland
eva.bayer@epfl.ch

R. Parimala
Department of Mathematics
& Computer Science
Emory University
Atlanta, GA 30322, USA.
parimala@mathcs.emory.edu
Irreducible Modules over the Virasoro Algebra

Rencai Lü, Xiangqian Guo, and Kaiming Zhao

Received: April 11, 2011
Revised: September 22, 2011
Communicated by Edward Frenkel

Abstract. In this paper, we construct two different classes of Virasoro modules from twisting Harish-Chandra modules over the twisted Heisenberg-Virasoro algebra by an automorphism of the twisted Heisenberg-Virasoro algebra. Weight modules in the first class are some irreducible highest weight modules over the twisted Heisenberg-Virasoro algebra. The non-weight modules in the first class are irreducible Whittaker modules over the Virasoro algebra. We obtain concrete bases for all irreducible Whittaker modules (instead of a quotient of modules). This generalizes known results on Whittaker modules. The second class of modules are non-weight modules which are not Whittaker modules. We determine the irreducibility and isomorphism classes of these modules.

2010 Mathematics Subject Classification: 17B10, 17B20, 17B65, 17B66, 17B68
Keywords and Phrases: Virasoro algebra, Whittaker module, irreducible module.

1 Introduction

Throughout this paper, we will use \( \mathbb{C}, \mathbb{C}^*, \mathbb{Z}, \mathbb{Z}_+, \) and \( \mathbb{N} \) to denote the sets of complex numbers, nonzero complex numbers, integers, nonnegative integers and positive integers respectively.

The theory of weight modules with finite-dimensional weight spaces over the Heisenberg algebra, the Virasoro algebra and the twisted Heisenberg-Virasoro algebra are fairly well developed. We refer the readers to [1], [4] [5], [11], [12], [13] and the references therein. For weight modules with infinite dimensional weight spaces, see [3], [7], [17]. Recently Whittaker modules over those algebras
were studied by many authors, see for example [2], [6], [10], [14], [16]. Besides Whittaker modules, some new non-weight modules over the Virasoro algebra were just constructed in [8].

We will use modules over the twisted Heisenberg-Virasoro algebra to study modules over the Virasoro algebra. Now we first recall the twisted Heisenberg-Virasoro algebra.

The twisted Heisenberg-Virasoro algebra $\mathcal{L}$ is the universal central extension of the Lie algebra {\(f(t)\frac{d}{dt} + g(t)f, g \in \mathbb{C}[t, t^{-1}]\)} of differential operators of order at most one on the Laurent polynomial algebra $\mathbb{C}[t, t^{-1}]$. More precisely, the twisted Heisenberg-Virasoro algebra $\mathcal{L}$ is a Lie algebra over $\mathbb{C}$ with the basis {\(d_n, t^n, z_1, z_2, z_3 \mid n \in \mathbb{Z}\)} and the Lie bracket given by

\[
[d_n, d_m] = (m-n)d_{m+n} + \delta_{n-m} \frac{n^3 - n}{12} z_1, \quad (1.1)
\]
\[
[d_n, t^m] = mt^{m+n} + \delta_{n-m} (n^2 + n) z_2, \quad (1.2)
\]
\[
[t^n, t^m] = n\delta_{n-m} z_3, \quad (1.3)
\]
\[
[L, z_1] = [L, z_2] = [L, z_3] = 0. \quad (1.4)
\]

The Lie algebra $\mathbb{L}$ has a Virasoro subalgebra $\text{Vir}$ with basis \(\{d_i, z_1 \mid i \in \mathbb{Z}\}\), and a Heisenberg subalgebra $H$ with basis \(\{t^i, z_3 \mid i \in \mathbb{Z}\}\).

Let $\sigma$ be an endomorphism of $\mathcal{L}$, and $V$ be any weight module of $\mathbb{L}$. We can make $V$ into another $\mathcal{L}$-module, by defining the new action of $\mathcal{L}$ on $V$ as

\[
x \circ v = \sigma(x) v, \forall x \in \mathcal{L}, v \in V. \quad (1.5)
\]

We will call the new module as the twisted module of $V$ by $\sigma$, and denote it by $V^\sigma$.

To avoid any ambiguity, we will not omit the circ for the new action.

The module $V^\sigma$ can be regarded as the Vir module by restriction to the Virasoro subalgebra. One important fact is that we can get a lot of new irreducible modules over Vir in this simple way, which include some new irreducible Whittaker modules. Since these modules are generally not weight modules, it is not trivial to determine isomorphism classes and irreducibility for these modules.

The paper is organized as follows. In section 2, we collect some known results for later use. In section 3, we construct our first class of Virasoro modules by twisting a highest weight $\mathbb{L}$-module (oscillator representation) with automorphisms of $\mathcal{L}$, then we obtain some new irreducible Whittaker modules $L_{\psi_m, z_1}$, where $m > 0$, over the Virasoro algebra. This concrete realization allows us to give concrete bases for all irreducible Whittaker modules (not only as a quotient of modules). Our bases for irreducible Whittaker modules $L_{\psi_m, z_1}$ with $m = 1$ generalize those results in [14] where it was required: $\psi_1(d_1)\psi_1(d_2) \neq 0$, and those results in [16] where an explicit formula for the Whittaker vector was give.
only for $\psi_1(d_1) \neq 0$ and $\psi_1(d_2) = 0$ in terms of Jack symmetric polynomial.

In section 4, we construct our second class of Virasoro modules by twisting $L$ modules of intermediate series with automorphisms of $L$. Then we determine the isomorphism classes and irreducibility of these Virasoro modules.

2 Preliminaries

In this section, we collect some notations and known facts for later use. For details, we refer the readers to [9], [12], [15], and the references therein.

Let us recall the definition of weight modules and highest weight modules over $L$.

It is well-known that $L$ has a natural $\mathbb{Z}$-gradation: $\deg d_n = \deg t = n$ and $\deg z_i = 0$ for $i = 1, 2, 3$. Set

$$L_+ = \sum_{n>0} (Cd_n + Ct^n), \quad L_- = \sum_{n<0} (Cd_n + Ct^n),$$

and

$$L_0 = \mathbb{C}d_0 + \mathbb{C}t^0 + \mathbb{C}z_1 + \mathbb{C}z_2 + \mathbb{C}z_3.$$

Then we have the triangular decomposition $L = L_+ \oplus L_0 \oplus L_-.$

For any $L$-module $V$ and $(\lambda, \lambda_H, c_1, c_2, c_3) \in \mathbb{C}^5$, set

$$V(\lambda, \lambda_H, c_1, c_2, c_3) = \{ v \in V \mid d_0v = \lambda v, t^0v = \lambda_H v, \text{ and } z_i v = c_i v \text{ for } i = 1, 2, 3 \},$$

which we generally call the weight space of $V$ corresponding to the weight $(\lambda, \lambda_H, c_1, c_2, c_3) \in \mathbb{C}^5$. When $t^0, z_1, z_2, z_3$ act as scalars on the whole space $V$, we shall simply write $V_\lambda$ instead of $V(\lambda, \lambda_H, c_1, c_2, c_3)$.

An $L$-module $V$ is called a weight module if $V$ is the sum of all its weight spaces. A weight $L$-module $V$ is called a highest weight module with highest weight $(\lambda, \lambda_H, c_1, c_2, c_3) \in \mathbb{C}^5$, if there exists a nonzero weight vector $v \in V(\lambda, \lambda_H, c_1, c_2, c_3)$ such that

1) $V$ is generated by $v$ as an $L$-module;
2) $L_+ v = 0$.

It is well known that, up to isomorphism, there exists a unique irreducible highest weight module $V(\lambda, \lambda_H, c_1, c_2, c_3)$ over $L$ with the highest weight $(\lambda, \lambda_H, c_1, c_2, c_3) \in \mathbb{C}^5$.

For any $a, b \in \mathbb{C}$, we have the Vir module $A_{a, b}$, called the module of intermediate series, which has basis $\{ t^k \mid k \in \mathbb{Z} \}$ such that $z_1$ acts trivially and

$$d_i t^k = (a + k + bi)t^{i+k}, \forall i, k \in \mathbb{Z}. \quad (2.1)$$

It is well-known that $A_{a, b}$ is irreducible if and only if $a \not\in \mathbb{Z}$ or $b \not\in \{0, 1\}$. We put $A'_{a, b} = A_{a, b}$ if $A_{a, b}$ is irreducible; otherwise $A'_{a, b}$ be the unique nontrivial irreducible sub-quotient of $A_{a, b}$.
Let us summarize some well-known results for the modules of intermediate series.

**Theorem 1.** Let $a, b, a_1, b_1 \in \mathbb{C}$.

1. If $b \notin \{0, 1\}$, then $A_{a,b} \cong A_{a_1,b_1}$ if and only if $b = b_1$ and $a - a_1 \in \mathbb{Z}$;

2. If $b \in \{0, 1\}$ and $a \notin \mathbb{Z}$, then we have $A_{a,b} \cong A_{a_1,b_1}$ if and only if $b_1 \in \{0, 1\}$ and $a - a_1 \in \mathbb{Z}$;

3. If $b \in \{0, 1\}$ and $a \in \mathbb{Z}$, then we have $A_{a,b} \cong A_{a_1,b_1}$ if and only if $b_1 \in \{0, 1\}$ and $a - a_1 \in \mathbb{Z}$;

4. If $a \in \mathbb{Z}$, then $A'_a(0,0) \cong A'_a(0,1) \cong A'(a,0)$.

We also need the following result from [15], and we will write it in a slightly different form for later use.

For any $\alpha = \sum_{i \in \mathbb{Z}} a_i t^i \in \mathbb{C}[t, t^{-1}], b \in \mathbb{C},$ (2.2)

we have the $\sigma = \sigma_{\alpha,b} \in \text{Aut}(L)$ defined as

$$\sigma(d_n) = d_n + t^n (\alpha + nb) - (n + 1)a_{-n}z_2$$

$$-\frac{1}{2} \left( \sum_i a_i a_{-n-i} - a_{-n}nb \right)z_3 + \delta_{n,0}bz_3 \left( z_2 + \frac{b}{2}z_3 \right)$$

$$\sigma(t^n) = t^n + \delta_{n,0}bz_3 - a_{-n}z_3, \sigma(z_1) = z_1 - 24bz_2 - 12b^2z_3,$$

$$\sigma(z_2) = z_2 + bz_3, \sigma(z_3) = z_3.$$ (2.3)

This can be verified directly, but one has to use the following formula

$$\sum_i (m+i)a_i a_{-m-i}z_3 = \frac{m-n}{2} \sum_i a_i a_{-m-i}z_3.$$ (2.4)

It is clear that

$$\sigma_{\alpha,b} \sigma_{\alpha_1,b_1} = \sigma_{\alpha+\alpha_1,b+b_1}, \forall \alpha, \alpha_1 \in \mathbb{C}[t, t^{-1}], b, b_1 \in \mathbb{C}. \quad (2.5)$$

### 3 Irreducible Whittaker Modules over the Virasoro Algebra

Let us recall the oscillator representation of the Heisenberg-Virasoro algebra $L$ on the Fock space $\mathcal{F} = \mathbb{C}[x_1, x_2, \ldots, x_n, \ldots]$. The action of $L$ is defined as (see Prop.2.3, Lemma 2.2 in [9])

$$t^n = \frac{\partial}{\partial x_n}, t^{-n} = nx_n, \forall n \in \mathbb{N}. \quad (3.1)$$
\[ t^0 = 0, z_3 = 1; z_2 = 0, z_1 = 1, \quad (3.2) \]
\[ d_k = -\frac{1}{2} \sum_{i \in \mathbb{Z}} t^{-i} \cdot t^{i+k}, \forall k \in \mathbb{Z}. \quad (3.3) \]

Actually, \( B \) is isomorphic to the irreducible highest weight module \( V(0,0,1,0,1) \) over \( \mathbb{L} \) as in [12].

For any homogenous polynomial \( u = x_1^{i_1} \ldots x_k^{i_k} \in B \), define \( \deg(u) = \sum_j i_j \lambda_j \), and denote \( B_i = \text{span}\{u = x_1^{i_1} \ldots x_k^{i_k} \in B| \deg(u) = i\} \) for all \( i \in \mathbb{N} \). Then we have the weight space decomposition \( B = \oplus_{i \in \mathbb{N}} B_i \), where \( B_i \) has the weight \(-i\).

From now on in this section we fix \( B \) and denote \( V = h^t((\alpha,b) \in \mathbb{C}) = \mathbb{C}[t] \).

Actually, \( B \) is a basis of \( B^\alpha \). It is easy to verify (or from results in [1]) that: for any \( B \),
\[ f_\alpha, \beta \in B \text{ is irreducible} \]
\[ \text{if and only if} \]
\[ f_\alpha, \beta \text{ is isomorphic to the irreducible highest weight module } V(\alpha, \beta). \]

The following set
\[ (3.8) \]
\[ \text{is isomorphic to } V(0,0,1,0,1) \] over \( \mathbb{L} \) as in [12].

**Lemma 2.** In \( B^\alpha \), we have
\[ d_n \circ 1 = 0, \forall n \geq 2m + 1, \quad (3.4) \]
\[ d_n \circ 1 = -\frac{1}{2} \sum_{i \in \mathbb{Z}} a_i a_{n-i} + a_{n} b, \forall m \leq n \leq 2m, \quad (3.5) \]
\[ z_1 \circ 1 = 1 - 2b^2, \quad (3.6) \]
\[ \text{htm}(d_n \circ (x_{i_1}^{k_1} \ldots x_{i_k}^{k_k})) = (m-n)a_m a_{m-n} x_{i_1}^{k_1} \ldots x_{i_k}^{k_k}, \forall n < m. \quad (3.7) \]
\[ \text{htm}((d_{i_1} d_{i_2} \ldots d_{i_k}) \circ 1) = a x_{m-i_1} x_{m-i_2} \ldots x_{m-i_k} \quad (3.8) \]

for all \( i_1, i_2, \ldots, i_k < m \), where \( a = \prod_{j=1}^{k} (a_{m-i_j}) \in \mathbb{C}^* \).

**Proof.** These follow from straightforward computations by using the formulas (2.3)-(2.5).

From (3.8) we know the following

**Lemma 3.** The following set
\[ B = \{(d_{i_1} d_{i_2} \ldots d_{i_k}) \circ 1 | i_1 \leq i_2 \leq \ldots \leq i_k < m\} \]
is a basis of \( B^\alpha \).

**Theorem 4.** For any \( \alpha \in A \setminus \mathbb{C} \) and \( b \in \mathbb{C} \), the module \( B^\alpha \) is irreducible over \( \text{Vir} \).
Proof. Recall that we have assumed that \( \alpha = \sum_{i=-m}^{m'} a_i t^i \in A \setminus \mathbb{C}[t] \) with \( m > 0 \) and \( a_{-m} \neq 0 \). Let \( V \) be a nonzero Vir submodule of \( B^{\alpha,0} \), and \( 0 \neq f \in V \) with lowest degree. Suppose that \( f \not \in \mathbb{C} \), and \( \deg(f) = n \). Say \( f_n = \text{htm}(f) \) with \( \frac{\partial}{\partial x_i}(f_n) \neq 0 \). Then we have

\[
g = (d_{i_1 + m} + t^{i_1 + m}(\alpha + (i_1 + m)b) \cdot f = d_{i_1 + m} \circ f + af \in V,
\]

where \( a = \sum_{i=-m}^{m'} \frac{a_i a_{m+i}}{2} \in \mathbb{C} \). It is straightforward to compute that

\[
\text{htm}(g) = \text{htm}(a_{-m} t^{i_1} \cdot f) = a_{-m} \frac{\partial}{\partial x_i}(f_n) \neq 0.
\]

And \( \deg(g) = n - i_1 < n \), which contradicts the choice of \( f \). So \( 1 \in V \). Thus \( V = B^{\alpha,0} \) by Lemma 3. Therefore \( B^{\alpha,0} \) is irreducible as a Vir module. \( \square \)

For any \( m \in \mathbb{N} \), denote \( \text{Vir}_{\geq m} = \bigoplus_{i \geq m} \mathbb{C}d_i \). Let \( \psi_m : \text{Vir}_{\geq m} \to \mathbb{C} \) be any nonzero homomorphism of Lie algebras and \( \dot{z}_1 \in \mathbb{C} \). Defined the one dimensional \( \text{Vir}_{\geq m} + \mathbb{C}z_1 \) module \( \mathbb{C}v \) by \( d_i v = \psi_m(d_i)v \) and \( z_1 v = \dot{z}_1 v \). Then we have the induced Vir module

\[
L_{\psi_m, z_1} = \text{Ind}_{U(\text{Vir}_{\geq m} + \mathbb{C}z_1)}^{U(\text{Vir})} \mathbb{C}v.
\]

Note that \( L_{\psi_m, z_1} \) is a Whittaker module with respect to the Whittaker pair \( (\text{Vir}, \text{Vir}_{\geq m} + \mathbb{C}z_1) \), and \( w = 1 \otimes v \) be a cyclic Whittaker vector in the sense of [2].

From the PBW theorem, \( L_{\psi_m, z_1} \) has a basis

\[
\{(d_{i_1} d_{i_2} \ldots d_{i_k}) w | i_1 \leq i_2 \leq \ldots \leq i_k < m \}. \tag{3.9}
\]

For \( m = 1 \), the irreducibility of \( L_{\psi_1, z_1} \) with \( \psi_1(d_1) \psi_1(d_2) \neq 0 \) was studied in [14] (see Proposition 4.8 and 6.1 in [14]), and an explicit formula for the Whittaker vector was given in [16] for \( \psi_1(d_1) \neq 0 \) and \( \psi_1(d_2) = 0 \) in terms of Jack symmetric polynomial.

**Theorem 5.** Let \( b \in \mathbb{C} \) and \( \alpha = \sum_{i=-m}^{m'} a_i t^i \in A \) with \( a_{-m} \neq 0 \) and \( m > 0 \). Then \( B^{\alpha,0} \cong L_{\psi_m, 1-12b} \) with \( \psi_m(d_n) = -\left( \sum_{i=0}^{m-1} \frac{a_i a_{-n-i}}{2} + a_{-n} b \right) \), for all \( n \geq m \).

**Proof.** It follows by Lemma 2, Lemma 3 and (3.9). \( \square \)

**Theorem 6.** Suppose \( \dot{z}_1 \in \mathbb{C} \) and \( \psi_m(d_{2m}) \neq 0 \). Then the Whittaker module \( L_{\psi_m, \dot{z}_1} \) over \( \text{Vir} \) is irreducible.

**Proof.** Denote \( \psi_m(d_n) = d_n \) for \( n = m, \ldots , 2m \). It is easy to see that there exist \( b \in \mathbb{C} \) and some \( \alpha(t) = \sum_{i=-m}^{m'} a_i t^i \in \mathbb{C}[t^{-1}] \) with \( a_{-m} \neq 0 \) satisfying the following equations:
Irreducible Modules over the Virasoro Algebra

\[ \begin{align*}
\dot{z}_1 &= 1 - 12b^2, \\
d_{2m} &= -\frac{a_m^2}{2}, \\
d_{2m-1} &= - (a_{m-1} - a_{m+1} + a_{2m+1}(2m - 1)b), \\
&\vdots \\
d_m &= - (\sum_{i=1}^{m} a_{m-i} + a_{m}b), \\
&\vdots \\
d_1 &= - (\sum_{i=1}^{m-1} a_{m-i} + a_{m}b).
\end{align*} \]

From Theorem 5 we know that \( L_{\psi, z_1} \cong B^{m, k} \). Using Theorem 4 we see that \( L_{\psi, z_1} \) is irreducible over Vir.

Now we can give the main result in this section.

**Theorem 7.** Suppose that \( z_1 \in \mathbb{C}, m \in \mathbb{N}, \) and \( \psi : \text{Vir}_{\geq m} \rightarrow \mathbb{C} \) is a Lie algebra homomorphism. Then the Whittaker module \( L_{\psi, z_1} \) over Vir is irreducible if and only if \( \psi_m(d_{2m}) \neq 0 \) or \( \psi_m(d_{2m-1}) \neq 0 \).

**Proof.** \( \Rightarrow \): Suppose that \( \psi_m(d_{2m}) = 0 \), and \( \psi_m(d_{2m-1}) = 0 \). Then it is straightforward to see that \( L_{\psi, z_1} \) has a proper submodule generated by \( d_{m-1} \).

\( \Leftarrow \): From Theorem 8, we need only to consider the case where \( \psi_m(d_{2m}) = 0 \) and \( \psi_m(d_{2m-1}) \neq 0 \).

Let \( \text{Vir}[\frac{1}{2}\mathbb{Z}] \) be the Virasoro algebra with the basis \( \{ d_k, z | k \in \frac{1}{2}\mathbb{Z} \} \) and subject to the relations:

\[ [d_m, d_n] = (m - n) d_{n+m} + \delta_{n,-m} \frac{n^3 - n}{12} z_1, \ \forall \ m, n \in \frac{1}{2}\mathbb{Z}. \]

Then \( \text{Vir}[\mathbb{Z}] \) is a subalgebra of \( \text{Vir}[\frac{1}{2}\mathbb{Z}] \). Now we define \( \psi_m^{-1/2} \) on \( \text{Vir}[\frac{1}{2}\mathbb{Z}] \) as

\[ \psi_m^{-1/2}(d_k) = \psi_m(d_k), \ \psi_m^{-1/2}(d_{k-1/2}) = 0, \ \forall \ k \in \mathbb{N}. \]

From Theorem 8 we know that the Whittaker module \( L_{\psi, z_1} \) with respect to the Whittaker pair \( (\text{Vir}[\frac{1}{2}\mathbb{Z}], \text{Vir}_{\geq m-1/2} + \mathbb{C}z_1) \) is irreducible with a basis:

\[ (d_{k_1}d_{k_2} \cdots d_{k_r})(d_{p_1}d_{p_2} \cdots d_{p_t})1 \]

where \( k_1, k_2, \ldots, k_r \in \frac{1}{2} + \mathbb{Z} \) with \( k_1 \leq k_2 \leq \cdots \leq k_r \leq m - 1 \), and \( p_1, p_2, \ldots, p_t \in \mathbb{Z} \) with \( p_1 \leq p_2 \leq \cdots \leq p_t \leq m - 1 \). Clearly, \( W = \text{span}\{d_{p_1}d_{p_2} \cdots d_{p_t}1 | p_1, p_2, \ldots, p_t \in \mathbb{Z} \text{ with } p_1 \leq p_2 \leq \cdots \leq p_t \leq m - 1 \} \) is a \( \text{Vir}[\mathbb{Z}] \)-module which is a Whittaker module isomorphic to \( L_{\psi, z_1} \) with respect to the Whittaker pair \( (\text{Vir}[\mathbb{Z}], \text{Vir}_{\geq m} + \mathbb{C}z_1) \).
We want to show that $W$ is irreducible as a $\text{Vir}[\mathbb{Z}]$-module. To the contrary, we assume that $W$ is not irreducible. Take a nonzero proper submodule $V$ of $W$. Let $V'$ be the span of the following subspaces

$$(d_{k_1}d_{k_2} \cdots d_{k_r})V; k_1, k_2, \ldots, k_r \in 1/2 + \mathbb{Z} \text{ with } k_1 \leq k_2 \leq \cdots \leq k_r \leq m - 1.$$ 

Then $V'$ is a proper subspace of $L_{\psi_{m-1/2}; z_1}$. Using PBW Theorem one can easily show that $V'$ is a submodule of $L_{\psi_{m-1/2}; z_1}$ over $\text{Vir}[\frac{1}{2}\mathbb{Z}]$, which is a contradiction. Thus $W$ is irreducible as a $\text{Vir}[\mathbb{Z}]$-module. Therefore the Whittaker module $L_{\psi_{m}; z_1}$ is irreducible over $\text{Vir}[\mathbb{Z}]$ if $\psi_{m}(d_{2m-1}) \neq 0$. 

\[ \blacksquare \]

We like to mention that the results in [14] and [16] are the special case of the above theorem with $m = 1$ and $\psi_1(d_1)\psi_1(d_2) \neq 0$ or $\psi_1(d_1) \neq 0$.

4 IRREDUCIBLE MODULES OVER $\text{Vir}$ WITH $z_1 = 0$

We can make $A = \mathbb{C}[t, t^{-1}]$ into an $L$-module by defining $d_i = t^{i+1} \frac{d}{dt}$, $t^i$ acting as multiplication by $t^i$, and $z_1, z_2, z_3$ acting as zero, i.e.,

$$d_i \cdot t^j = j^i t^{i+j}, t^i \cdot t^j = t^{i+j}, z_k \cdot A = 0, \forall k = 1, 2, 3.$$ 

The module on $A$ is isomorphic to $V(0, 0; 1)$ defined on Page 187 of [12] which is an irreducible module over $L$. We will use this module instead of the most general case $V(a, b; F)$, where $a, b, F \in \mathbb{C}$, since we will essentially obtain isomorphic Virasoro modules.

For any $\alpha, \beta \in \mathbb{C}$, we have the $L$-module $A_{\alpha, \beta} = A^{\sigma_{\alpha, \beta}}$. The action of $L$ on $A_{\alpha, \beta}$ is

$$d_n \circ t^i = (\alpha + i + nb)t^{n+i}, \forall i, n \in \mathbb{Z}, \quad (4.1)$$

$$t^i \circ t^j = t^{i+j}, z_k \circ A = 0, \forall i, j \in \mathbb{Z}, k = 1, 2, 3. \quad (4.2)$$

In this section, the irreducibility and the isomorphism classes of such modules are completely determined.

Note that if $\alpha \in \mathbb{C}$, then $A_{\alpha, b}$ is simply a weight module of intermediate series in [12].

**Lemma 8.** (1) Let $V$ be an irreducible $L$ module, then $V^{\sigma_{\alpha, \beta}}$ is irreducible for any $\alpha \in A, b \in \mathbb{C}$.

(2) $A_{\alpha, b}$ is irreducible as an $L$ module for any $\alpha \in A, b \in \mathbb{C}$.

**Proof.** The statements in this Lemma are obvious. \[ \blacksquare \]

**Lemma 9.** For any $k \in \mathbb{Z}$, let

$$w_k = -\frac{1}{2}d_{k-1}d_1 - \frac{1}{2}d_{k+1}d_{-1} + d_kd_0 \in U(\text{Vir}). \quad (4.3)$$

Then $w_k \circ g = b(b - 1)t^k g$, for all $k \in \mathbb{Z}$ and $g \in A_{\alpha, b}$.
Proof. For any \( k \in \mathbb{Z} \) and \( t^j \in A_{\alpha,b} \), we compute
\[
(d_{k-1}d_i) \circ t^j = d_{k-1} \circ (\alpha + j + ib)t^{k+j} = (\alpha + j + ib)(\alpha + i + j + (k - i)b)t^{k+j}.
\]
Taking \( i = -1, 0, 1 \) respectively, we get
\[
w_k \circ t^j = \left[-\frac{1}{2}(\alpha + j + b)(\alpha + 1 + j + (k - 1)b) - \frac{1}{2}(\alpha + j - b)(\alpha - 1 + j + (k + 1)b) + (\alpha + j)(\alpha + j + kb)\right]t^{k+j} = b(b-1)t^{k+j}.
\]
Thus \( w_k \circ g = b(b-1)t^k g \) for all \( k \in \mathbb{Z} \) and \( g \in A_{\alpha,b} \).

Corollary 10. Suppose that \( b \notin \{0, 1\}, 0 \neq g \in A_{\alpha,b} \). Then \( \text{span}_C\{w_i \circ g | i \in \mathbb{Z}\} = A_{\alpha,b} \) if and only if \( g = ct^i \) for some \( c \in \mathbb{C}^* \) and \( i \in \mathbb{Z} \).

Proof. From Lemma 9, \( \text{span}_C\{w_i \circ g | i \in \mathbb{Z}\} = A_{\alpha,b} \) if and only if \( g = ct^i \) for some \( c \in \mathbb{C}^* \) and \( i \in \mathbb{Z} \).

Lemma 11. If \( \alpha \in A \setminus C \), then \( E_\alpha = \text{span}_C\{\alpha t^i + it^i | i \in \mathbb{Z}\} \neq A \).

Proof. For any \( 0 \neq f = \sum_{i=s}^{n} b_i t^i \in A \) with \( b_s, b_r \neq 0 \), define \( \text{deg}(f) = (s, r) \) and \( l(f) = r - s \). Suppose that
\[
\alpha = \sum_{i=m}^{n} \alpha_i t^i \in A \setminus C, \quad \text{with deg} \alpha = (m, n).
\]

Case 1. \( m < 0 < n \).
It is easy to see that \( l(f) \geq n - m \) for all \( f \in E_\alpha \), hence \( E_\alpha \neq A \) in this case.
Case 2. \( m \geq 0 \) or \( n \leq 0 \).
Without loss of generality, we may assume that \( m \geq 0 \). If \( m = 0 \) then \( n > 0 \) since \( \alpha \notin C \); if \( m > 0 \) then \( n \geq m > 0 \). If \( a_0 \notin \mathbb{Z} \) or \( m > 0 \), then it is easy to check that \( l(f) \geq 1 \) for all \( f \neq f \in E_\alpha \). So \( E_\alpha \neq A \). If \( a_0 \in \mathbb{Z} \) and \( m = 0 \), then \( n > 0 \) and \( a_0 \neq 0 \). It is not hard to see that \( t^{-a_0} \notin E_\alpha \).

Theorem 12. Let \( \alpha \in A, b \in \mathbb{C} \).

1. If \( b \notin \{0, 1\} \), then \( A_{\alpha,b} \) is irreducible as a Vir module with action defined as in (4.1).
2. The Vir module \( A_{\alpha,1} \) is irreducible if and only if \( \alpha \notin C \setminus \mathbb{Z} \). If \( \alpha \notin C \setminus \mathbb{Z} \), then
\[
d_0 \circ A_{\alpha,1} = \oplus_{i \in \mathbb{Z}} C(\alpha + i)t^i
\]
is the unique irreducible Vir submodule of \( A_{\alpha,b} \), and Vir acts on \( A_{\alpha,1}/(d_0 \circ A_{\alpha,1}) \) as zero.
3. The Vir module \( A_{\alpha,0} \) is irreducible if and only if \( \alpha \notin \mathbb{Z} \).
Proof. (1). Suppose that \(b \neq 0, 1\). Let \(M\) be a nonzero Vir submodule of \(A_{\alpha,b}\). Then by Lemma 9, we have \(p \cdot M \subset M\) for all \(i \in Z\). Hence \(M\) is also a \(L\) submodule of \(A_{\alpha,b}\). Thus \(M = A_{\alpha,b}\) by Lemma 8. So \(A_{\alpha,b}\) is irreducible as Vir module in this case.

(2). Note that \(d_j \circ g = d_0 \circ (t^j g)\) for all \(g \in A_{\alpha,1}, i \in Z\). So \(d_0 \circ A_{\alpha,1}\) is a Vir submodule of \(A_{\alpha,1}\), and Vir acts trivially on \(A_{\alpha,1}/(d_0 \circ A_{\alpha,1})\). For any nonzero Vir submodule \(M\) of \(A_{\alpha,1}\). Let \(AM = \text{span}\{t^j g | i \in Z, g \in M\}\). Since \(d_0 \circ t^j g = d_i \circ g \in M\) for all \(g \in M\), then

\[
\text{d}_0 \circ AM \subset M.
\] (4.4)

Noting that \(\text{d}_j \circ (t^i g) = \text{d}_{i+j} \circ g\) we see that \(AM\) is a Vir submodule. Then \(AM\) is also an \(L\) module, and \(AM = A\) by Lemma 8. Combining with (4.4), we have \(\text{d}_0 \circ A_{\alpha,1} \subset M\), i.e., \(\text{d}_0 \circ A_{\alpha,1}\) is the unique minimum nonzero submodule of \(A_{\alpha,1}\). Note that \(\text{d}_0 \circ A_{\alpha,1} = \text{span}_C\{\alpha + it^i | i \in Z\}\). Using Lemma 11 we see that \(A_{\alpha,1}\) is not irreducible if \(\alpha \notin C\). For \(\alpha \in C\) it is well-known that \(A_{\alpha,1}\) is not irreducible if and only if \(\alpha \in Z\). This proves (2).

(3) and (4). It is straightforward to check that \(\eta : A_{\alpha,0} \rightarrow \text{d}_0 \circ A_{\alpha,1}\) with \(\eta(g) = \text{d}_0 \circ g\) is a Vir-module epimorphism. From (2), we know that \(A_{\alpha,0}\) is irreducible if and only if \(\eta\) is injective. Note that if \(\text{d}_0 \circ g = 0\) for some \(0 \neq g \in A_{\alpha,1}\), then \( \frac{d_0}{dt}\) \(g\) + \(ag\) = 0. By comparing the terms of highest and lowest degree in \(t\) respectively, we have \(\alpha \in C\). Using well-known results on \(A_{\alpha,0}\) for \(\alpha \in C\), we see that the statements in (3) and (4) are true. So we have proved the theorem.

**Lemma 13.** If \(\alpha_1 - \alpha_2 \in Z\), then \(A_{\alpha_1,b} \cong A_{\alpha_2,b}\) as Vir-modules.

**Proof.** Suppose that \(\alpha_1 = \alpha_2 + k\) for some \(k \in Z\). Then it is straightforward to check that \(\eta : A_{\alpha_1,b} \rightarrow A_{\alpha_2,b}\) with \(\eta(t^i) = t^{i+k}\) is a Vir-module isomorphism.

**Lemma 14.** For any \(\alpha \notin Z\) there exist finitely many nonzero \(v^{(i)}_{\alpha} \in U(Vir)\) for \(i \in Z\) such that for any \(k \in Z\), the element

\[
u_{\alpha,k} = \sum_i d_{i+k}v^{(i)}_{\alpha} \in U(Vir)
\] (4.5)
satisfies that \(u_{\alpha,k} \circ g = t^k u_{\alpha,0} \circ g\) in \(A_{\alpha,0}\) for all \(g \in A_{\alpha,0}\).

**Proof.** Note that \(A_{\alpha,0}\) is irreducible as a Vir module. Since \(d_0 \circ 1 = \alpha \neq 0\), there exists \(u_{\alpha} \in U(Vir)\) such that \(u_{\alpha} \circ d_0 \circ 1 = 1\). Let \(u_{\alpha,0} = u_{\alpha} \circ d_0 \in U(Vir)\). Then we can write \(u_{\alpha,0}\) as in (4.5) with \(k = 0\), and we have the definition for \(u_{\alpha,k}\). Note that \(d_{i+k} \circ g = t^k(d_i \circ g)\) for all \(g \in A_{\alpha,0}\) and \(k \in Z\). Then we can easily verify that \(u_{\alpha,k} \circ g = t^{k} u_{\alpha,0} \circ g\) in \(A_{\alpha,0}\) for all \(g \in A_{\alpha,0}\).
Theorem 15. Let \( \alpha_1, \alpha_2 \in A, b_1, b_2 \in \mathbb{C} \).

1. If \( b_1 \notin \{0, 1\} \), then \( A_{\alpha_1, b_1} \cong A_{\alpha_2, b_2} \) as Vir modules if and only if \( \alpha_1 - \alpha_2 \in \mathbb{Z} \) and \( b_1 = b_2 \).

2. If \( b \in \{0, 1\} \), then \( A_{\alpha_1, b} \cong A_{\alpha_2, b} \) as Vir modules if and only if \( \alpha_1 - \alpha_2 \in \mathbb{Z} \).

3. \( A_{\alpha_1, 0} \cong A_{\alpha_2, 1} \) as Vir modules if and only if \( \alpha_1 - \alpha_2 \in \mathbb{Z} \) and \( \alpha_1 \in \mathbb{C} \setminus \mathbb{Z} \).

Proof. From Theorem 1, Theorem 12, and Lemma 13, we see that all the sufficient conditions in (1)-(3) are satisfied. So we need only to prove the necessity of the conditions in (1)-(3).

(1). From Lemma 9, we have \( b_1(b_1-1) = b_2(b_2-1) \neq 0 \). So \( b_2 \notin \{0, 1\} \). Suppose that \( \sigma : A_{\alpha_1, b_1} \to A_{\alpha_2, b_2} \) is a Vir module isomorphism. By Lemma 9 again, we know that \( \{w_k \circ 1 \mid k \in \mathbb{Z}\} \) is a basis for \( A_{\alpha_1, b_1} \). Then \( \{w_k \circ \sigma(1) \mid k \in \mathbb{Z}\} \) is a basis for \( A_{\alpha_2, b_2} \). By Corollary 10, we know that \( \sigma(1) = ct^{iu} \) for some \( c \in \mathbb{C}^* \) and \( i_0 \in \mathbb{Z} \). Computing \( \sigma(d_j \circ 1) = d_j \circ \sigma(1) \), we have \( (i_0 + \alpha_2 - \alpha_1) + (b_2 - b_1)j = 0 \) for all \( j \in \mathbb{Z} \). Thus \( \alpha_1 = \alpha_2 + i_0 \) and \( b_1 = b_2 \).

(2). Case 1. \( b = 0 \).

If \( \alpha_1 \in \mathbb{Z} \), then we have \( \alpha_2 \in \mathbb{Z} \) by Theorem 12 (3).

Now suppose that \( \alpha_1, \alpha_2 \notin \mathbb{Z} \), and that \( \sigma : A_{\alpha_1, 0} \to A_{\alpha_2, 0} \) is a Vir-module isomorphism. Take \( u_{\alpha_1, k} \) as in Lemma 14. Then \( u_{\alpha_1, k} \circ 1 = t^k \), for all \( k \in \mathbb{Z} \). Note that
\[
\sigma(t^k) = \sigma(u_{\alpha_1, k} \circ 1) = u_{\alpha_1, k} \circ \sigma(1) = t^k(u_{\alpha_1, k} \circ \sigma(1)) = t^k \sigma(u_{\alpha_1, 0} \circ 1) = t^k \sigma(1) \in A_{\alpha_2, 0}.
\]
Then
\[
A_{\alpha_2, 0} = \sigma(A_{\alpha_1, 0}) = \text{span}_\mathbb{C}\{\sigma(u_{\alpha_1, k} \circ 1) \mid k \in \mathbb{Z}\} = \text{span}_\mathbb{C}\{t^k \sigma(1) \mid k \in \mathbb{Z}\} = A\sigma(1).
\]
So \( \sigma(1) = ct^{iu} \) for some \( c \in \mathbb{C}^*, i_0 \in \mathbb{Z} \), and
\[
\sigma(t^i) = ct^{i+iu}, \quad \forall i \in \mathbb{Z}.
\]
By a similar computation as in the last step in (1), we deduce that \( \alpha_1 = \alpha_2 + i_0 \) and the result follows in this case.

Case 2. \( b = 1 \).

Suppose that \( A_{\alpha_1, 1} \cong A_{\alpha_2, 1} \) as Vir modules. If \( \alpha_1 \in \mathbb{Z} \), then \( A_{\alpha_1, 1} \) is a weight module with respect to \( d_0 \). This forces \( A_{\alpha_2, 1} \) to be a weight module with respect to \( d_0 \). So \( \alpha_2 \in \mathbb{C} \). Theorem 14 (2) ensures that \( \alpha_1 - \alpha_2 \in \mathbb{Z} \).

Now suppose that \( \alpha_1, \alpha_2 \notin \mathbb{Z} \). From Theorem 14 (4) we know that \( A_{\alpha_1, 0} \cong d_0 \circ A_{\alpha_1, 1} \cong d_0 \circ A_{\alpha_2, 1} \cong A_{\alpha_2, 0} \) as irreducible Vir modules. Thus \( \alpha_1 - \alpha_2 \in \mathbb{Z} \) by Case 1.

(3). If \( \alpha_2 \in \mathbb{C} \), then \( A_{\alpha_2, 1} \) is a weight module with respect to \( d_0 \). This forces \( A_{\alpha_1, 0} \) to be a weight module with respect to \( d_0 \). So \( \alpha_1 \in \mathbb{C} \). From Theorem 1 we know that \( \alpha_1 - \alpha_2 \in \mathbb{Z} \) and \( \alpha_1 \in \mathbb{C} \setminus \mathbb{Z} \).

Documenta Mathematica 16 (2011) 709–721
Now suppose that $\alpha_2 \notin \mathbb{C}$, then $A_{\alpha_2,1}$ is a non-weight module with respect to $d_0$. This forces $A_{\alpha_1,0}$ to be a non-weight module with respect to $d_0$. So $\alpha_1 \notin \mathbb{C}$. From Theorem 12 (2) and (4) we know that $A_{\alpha_1,0}$ is irreducible while $A_{\alpha_2,1}$ is not irreducible. So $A_{\alpha_1,0}$ and $A_{\alpha_2,1}$ cannot be isomorphic in this case.

Acknowledgement. We are grateful to the referee for giving nice suggestions. X.G. is partially supported by NSF of China (11101380 and 11026155), R.L. is partially supported by NSF of China, and K.Z. is partially supported by NSERC and NSF of China.

References

[1] E. Arbarello, C. De Concini, V. G. Kac and C. Procesi, Moduli Spaces of Curves and Representation Theory, *Commun. Math. Phys.*; 117 (1988), 1-36.

[2] P. Batra and V. Mazorchuk, Blocks and modules for Whittaker pairs, *J. Pure Appl. Algebra*, 215 (2011), no.7, 1552-1568.

[3] V. Bеккерт, G. Benkart and V. Futorny, Weight modules for Weyl algebras, Kac-Moody Lie algebras and related topics, *Contemp. Math.*, vol.343, 2004, 17-42.

[4] Y. Billig, Representations of the twisted Heisenberg-Virasoro algebra at level zero, *Canad. Math. Bull.*, 46 (2003), no.4, 529-533.

[5] V. Chari, Integrable representations of affine Lie algebras, *Invent. Math.* 85(2) (1986), 317-335.

[6] K. Christodouloupoulou, Whittaker modules for Heisenberg algebras and imaginary Whittaker modules for affine Lie algebras, *J. Algebra*, 320 (2008), 2871-2890.

[7] C. Conley and C. Martin, A family of irreducible representations of the Witt Lie algebra with infinite-dimensional weight spaces, *Compos. Math.*, 128(2)(2001), 153-175.

[8] X. Guo, R. Lu and K. Zhao, Fraction modules of the Virasoro algebra, preprint, 2010.

[9] V. Kac and A. Raina, *Bombay lectures on highest weight representations of infinite dimensional Lie algebras*, World Sci., Singapore, 1987.

[10] D. Liu, Y. Wu and L. Zhu, Whittaker modules for the twisted Heisenberg-Virasoro algebra, *J. Math. Phys.*, 51(1) (2010), 023524, 12pp.

[11] D. Liu and C. Jiang, Harish-Chandra modules over the twisted Heisenberg-Virasoro algebra, *J. Math. Phys.*, 49(1) (2008), 012901, 13pp.
[12] R. Lu and K. Zhao, Classification of irreducible weight modules over the twisted Heisenberg-Virasoro algebras, *Comm. Contemp. Math.*, Vol. 12, No.2 (2010) 183-205.

[13] O. Mathieu, Classification of Harish-Chandra modules over the Virasoro Lie algebra, *Invent. Math.*, 107 (1992), no. 2, 225-234.

[14] M. Ondrus and E. Wiesner, Whittaker Modules for the Virasoro Algebra, *J. Algebra Appl.*, 8(2009), no.3, 363-377.

[15] R. Shen and C. Jiang, The Derivation algebra and automorphism group of the twisted Heisenberg-Virasoro algebra, *Commun. Alg.*, 34 (2006), 2547-2558.

[16] S. Yanagida, Whittaker vectors of the Virasoro algebra in terms of Jack symmetric polynomial, *J. Algebra*, 333 (2011), 273-294.

[17] H. Zhang, A class of representations over the Virasoro algebra, *J. Algebra*, 190 (1997), 1-10.

---

R. Lü  
Department of Mathematics  
Soochow University  
Suzhou 215006, Jiangsu  
P. R. China  
rhu@suda.edu.cn

X. Gu  
Department of Mathematics  
Zhengzhou University  
Zhengzhou 450001, Henan  
P. R. China  
guoxq@amss.ac.cn

K. Zhao  
Department of Mathematics  
Wilfrid Laurier University  
Waterloo, ON, Canada N2L 3C5  
and College of Mathematics  
and Information Science  
Hebei Normal (Teachers)  
University Shijiazhuang  
Hebei, 050016  
P. R. China.  
kzhao@whu.ca
Theta Series and Function Field Analogue of Gross Formula

Fu-Tsun Wei and Jing Yu

Received: October 14, 2010
Revised: July 6, 2011
Communicated by Peter Schneider

Abstract. Let $k = \mathbb{F}_q(t)$, with $q$ odd. In this article we introduce “definite” (with respect to the infinite place of $k$) Shimura curves over $k$, and establish Hecke module isomorphisms between their Picard groups and the spaces of Drinfeld type “new” forms of corresponding level. An important application is a function field analogue of Gross formula for the central critical values of Rankin type $L$-series coming from automorphic cusp forms of Drinfeld type.

2010 Mathematics Subject Classification: 11R58 11G18 11F41 11F67
Keywords and Phrases: function field, quaternion algebra, Shimura curve, automorphic form, Hecke operator, special value of $L$-series.

Introduction

We present here a theory of “definite” quaternion algebras over the rational function field $k := \mathbb{F}_q(t)$ with $q$ odd, “definite” means that the place $\infty$ at infinity ramifies for the quaternion algebra in question. Following Gross [5], we first give a geometric translation of Eichler’s arithmetic theory of definite quaternion algebra by introducing the so-called “definite” Shimura curves. The geometry of these curves is simple and easy to manipulate. Basing on Eichler’s trace computation, one is lead (via Jacquet–Langlands) to an explicit Hecke module isomorphism between the Picard groups of definite Shimura curves and spaces of automorphic forms of Drinfeld type over the function field $k$.

1Both authors were supported by National Science Council grant (100)-2119-M002-007. The first author was also supported by the Graduate Students Study Abroad Program of National Science Council.
Automorphic forms of Drinfeld type are very useful tools for function fields arithmetic (cf. [7], [12] and [17] for more details and applications), which can be viewed as an analogue of classical modular forms of weight 2. To illustrate our approach to quaternion algebras over function field, we give an application to the study of central critical values of certain $L$-series of “Rankin type” in the global function field setting. These $L$-series include, among others, $L$-series coming from elliptic curves over $k$ with square free conductor supported at even number of places and having split multiplicative reduction at $\infty$. Having the extensive calculations done in [12], we obtain in particular a function field analogue of Gross formula for the central critical values of these $L$-series over “imaginary” quadratic extensions of $k$ (with respect to $\infty$).

The structure of this article is modelled on [8]. Let $D$ be a “definite” quaternion algebra over $k$ and let $N_0$ be the product of finite ramified primes of $D$. We introduce the definite Shimura curve $X = X_{N_0}$ over $k$ (for maximal orders) in §1 which is a finite union of genus zero curves. Also introduced are the Gross points, which are special points on these curves associated to orders in imaginary quadratic extensions of $k$. With a natural choice of basis on the Picard group $\text{Pic}(X)$, the Hecke correspondences can be expressed by Brandt matrices.

From the entries of Brandt matrices we introduce certain theta series. Taking into account the Gross height pairing on the $\text{Pic}(X)$ (defined in §1.2), we then have at hand a construction of automorphic forms of Drinfeld type for the congruence subgroup $\Gamma_0(N_0)$ of $\text{GL}_2(\mathbb{F}_q[t])$. The main theorem of this article in §2.3 is:

**Theorem.** There is a map $\Phi : \text{Pic}(X) \times \text{Pic}(X)^\vee \longrightarrow M_{\text{new}}(\Gamma_0(N_0))$ such that for all monic polynomials $m$ of $\mathbb{F}_q[t]$

$$T_m \Phi(e, e') = \Phi(t_m e, e') = \Phi(e, t_m e').$$

Here $\text{Pic}(X)^\vee$ is the dual group $\text{Hom}(\text{Pic}(X), \mathbb{Z})$, $M_{\text{new}}(\Gamma_0(N_0))$ is the space of Drinfeld type “new” forms for $\Gamma_0(N_0)$, $t_m$ are Hecke correspondences on $X$, and $T_m$ are Hecke operators on $M_{\text{new}}(\Gamma_0(N_0))$. Moreover, this map induces an isomorphism (as Hecke modules)

$$(\text{Pic}(X) \otimes_{\mathbb{Z}} \mathbb{C}) \otimes_{\mathbb{T}_c} (\text{Pic}(X)^\vee \otimes_{\mathbb{Z}} \mathbb{C}) \cong M_{\text{new}}(\Gamma_0(N_0)).$$

This theorem in fact tells us that all automorphic “new” forms of Drinfeld type come from our theta series. The special case of our theorem when $N_0$ is single prime is also obtained in Papikian [10] §3, by a different geometric method using Néron models of jacobians of Drinfeld modular curve $X_0(N_0)$. In our proof of the above theorem, we use the explicit construction of theta series and claim the equality of the trace of the $m$-th Brandt matrix $B(m)$ and the trace of the Hecke operator $T_m$ on $M_{\text{new}}(\Gamma_0(N_0))$ for each monic polynomial $m$ in $\mathbb{F}_q[t]$. This claim is essentially the Jacquet-Langlands correspondence (cf.
Theta series and function field analogue . . .

[9]) between automorphic representations of quaternion algebras over $k$ and automorphic cuspidal representations of $GL_2$ over $k$. Another crucial step in the proof is to show that the Hecke module $M_{\text{new}}(\Gamma_0(N_0))$ is free of rank one, which follows from the multiplicity one theorem (cf. [3]). For the sake of completeness, we recall these results in Appendix.

Let $D$ be an irreducible polynomial in $\mathbb{F}_q[t]$ such that $K = k(\sqrt{D})$ is imaginary and $P$ is inert in $K$ if the prime $P$ divides $N_0$. For each ideal class $A$ of $\mathbb{F}_q[t][\sqrt{D}] = O_K$, we construct in §2.4 an automorphic form $g_A$ of Drinfeld type with its Fourier coefficients worked out. In §3.1 we recall Rankin product of $L$-series $\Lambda(f, A, s)$ associated to Drinfeld type new form $f$ for $\Gamma_0(N_0)$ and partial zeta function $\zeta_A$. In §3.2 we express the central critical value $\Lambda(f, A, 0)$ as the Petersson inner product of $f$ and $g_A$. Furthermore, when $f$ is a “normalized” Hecke eigenform and $\chi$ is a character of ideal class group $\text{Pic}(O_K)$ of $O_K$, we give the twisted critical value $\Lambda(f, \chi, 0)$ explicitly in terms of the Gross height of a special divisor class $\epsilon_{f, \chi}$ on the definite Shimura curve $X_{N_0}$. This is our analogue of Gross formula.

Let $E$ be an elliptic curves over $k$ with conductor $N_0\infty$ and split multiplicative reduction at $\infty$. From the work of Weil, Jacquet-Langlands, and Deligne, it is well known that there exists a Drinfeld type cusp form $f_E$ such that

$$L(E/k, s + 1) = L(f_E, s).$$

Here $L(E/k, s)$ is the Hasse-Weil $L$-series of $E$ over $k$. After doing base change to the quadratic field $K$, one gets

$$L(E/K, s + 1) = \Lambda(f, 1_D, s)$$

where $1_D$ is the trivial character of $\text{Pic}(O_K)$. Our formula can certainly be applied to these elliptic curves. An example is given in §3.4.

Notation

We fix the following notations:

- $k$: the rational function field $\mathbb{F}_q(t)$, $q = p^{\ell_0}$ where $p$ is an odd prime.
- $A$: the polynomial ring $\mathbb{F}_q[t]$.
- $\infty$: the infinite place, which corresponds to degree valuation $v_\infty$.
- $\pi_\infty$: $t^{-1}$, a fixed uniformizer of $\infty$.
- $k_\infty$: $\mathbb{F}_q((t^{-1}))$, i.e. the completion of $k$ at $\infty$.
- $\mathcal{O}_\infty$: $\mathbb{F}_q[[t^{-1}]]$, i.e. the valuation ring in $k_\infty$.
- $P$: a finite prime (place) of $k$.
- $kp$: the completion of $k$ at the finite prime $P$.
Fu-Tsun Wei and Jing Yu

\( A_P \): the closure of  in \( k_P \).
\( k_k \): the adele ring of \( k \).
\( \mathring{k} \): \( \prod_P k_P \), the finite adele ring of \( k \).
\( \mathring{A} \): \( \prod_P A_P \).

\( \psi_\infty \): a fixed additive character on \( k_{\infty} \): for \( y = \sum a_i \pi_i^{1} \in k_{\infty} \), we define
\[
\psi_\infty(y) := \exp \left( \frac{2\pi i}{p} \cdot \Tr_{F_q/F_p}(-a_1) \right).
\]

We identify non-zero ideals of \( A \) with the monic polynomials in \( A \) by using the same notation.

1 Definite Shimura curves

Let \( D \) be a quaternion algebra over \( k \) ramified at \( \infty \) (call \( D \) “definite”). Before introducing the definite Shimura curve for \( D \), we start with a genus 0 curve \( Y \) over \( k \) associated with the quaternion algebra \( D \), which is defined by the following: the points of \( Y \) over any \( k \)-algebra \( M \) are
\[
Y(M) = \{ x \in D \otimes_k M : x \neq 0, \Tr(x) = \Nr(x) = 0 \}/M^\times,
\]
where the action of \( M^\times \) on \( D \otimes_k M \) is by multiplication on \( M \), \( \Tr \) and \( \Nr \) are respectively the reduced trace and the reduced norm of \( D \). More precisely, if \( D = k + ku + kv + kuv \) where \( u^2 = \alpha \), \( v^2 = \beta \), \( \alpha \) and \( \beta \) are in \( k^\times \), and \( uv = -vu \), then \( Y \) is just the conic
\[
\alpha y^2 + \beta z^2 = \alpha \beta w^2
\]
in the projective plane \( \mathbb{P}^2 \). The group \( D^\times \) acts on \( Y \) (from the right) by conjugation. If \( K \) is a quadratic extension of \( k \), \( Y(K) \) is canonically identified with the set \( \text{Hom}(K, D) \) of embeddings: for each embedding \( f : K \rightarrow D \), let \( y = y_f \) be the image of the unique \( K \)-line on the quadric \( \{ x \in D \otimes_k K : \Tr(x) = \Nr(x) = 0 \} \) on which conjugation by \( f(K^\times) \) acts by multiplication by the character \( a \mapsto a/\bar{a} \). Note that \( y_f \) is one of the two fixed points of \( f(K^\times) \) acting on \( Y(K) \); another one is the image of the line where conjugation acts by the character \( a \mapsto \bar{a}/a \).

Let \( N_0 \) be the product of the finite ramified primes of \( D \). Choose a maximal \( A \)-order \( R \) of \( D \). For any finite prime \( P \) let \( R_P := R \otimes_A A_P \), \( D_P := D \otimes_k k_P \), and
\[
\mathring{R} := R \otimes_A \mathring{A}, \mathring{D} := D \otimes_k \mathring{k}.
\]

DEFINITION 1.1. (cf. [2] and [8]) The definite Shimura curve \( X_{N_0} \) is defined as
\[
X_{N_0} = \left( \mathring{R}^\times \backslash \mathring{D}^\times \times Y \right)/D^\times.
\]

We will use the notation \( X \) instead of \( X_{N_0} \) when \( N_0 \) is fixed.

LEMMA 1.2. \( X_{N_0} \) is a finite union of curves of genus 0.
Proof. Let $g_1, \ldots, g_n$ be representatives for the finite double coset space $\hat{\mathcal{R}}^\times \backslash \hat{\mathcal{D}}^\times /\hat{\mathcal{D}}^\times$, i.e.

$$\hat{\mathcal{D}}^\times = \coprod_{i=1}^n \hat{\mathcal{R}}^\times g_i \hat{\mathcal{D}}^\times.$$ 

Then each coset of $X_{N_0}$ has a representative $(\hat{\mathcal{R}}^\times g_i, y)$ mod $\hat{\mathcal{D}}^\times$ and the map

$$X_{N_0} \to \coprod_{i=1}^n Y/\Gamma_i \quad (\hat{\mathcal{R}}^\times g_i, y) \mapsto y \mod \Gamma_i$$

is a bijection, where $\Gamma_i = g_i^{-1} \hat{\mathcal{R}}^\times g_i \cap \hat{\mathcal{D}}^\times$ is a finite group for $i = 1, \ldots, n$. \hfill \Box

**Definition 1.3.** Let $K$ be an imaginary quadratic extension of $k$ (i.e. $\infty$ is not split in $K$). We call $x = (g, y) \in \text{Image} \left[ \hat{\mathcal{R}}^\times \backslash \hat{\mathcal{D}}^\times \times Y(K) \to X_{N_0}(K) \right]$ a **Gross point** on $X_{N_0}$ over $K$.

Let $f : K \to \mathcal{D}$ be the embedding corresponding to $y$. Then

$$f(K) \cap g^{-1} \hat{\mathcal{R}} \hat{\mathcal{D}} = f(O_d)$$

for some quadratic order $O_d := A[\sqrt{d}]$ where $d$ is an element in $A$ with $d \notin k_\infty^2$. In this case, we say $x$ is of **discriminant** $d$. Note that the discriminant of a Gross point is well-defined up to multiplying with elements in $(\mathbb{F}_q^\times)^2$. Set $X_i := Y/\Gamma_i$.

If the component $g$ of a Gross point $x$ is congruent to $g_i$ in $\hat{\mathcal{R}}^\times \backslash \hat{\mathcal{D}}^\times /\hat{\mathcal{D}}^\times$, then $x$ lies on the component $X_i(K) = (Y/\Gamma_i)(K)$.

1.1 **Actions on Gross Points**

Let $a \in \hat{K}^\times$ where $\hat{K} := K \otimes_k \hat{k}$ and $x = (g, y)$ be a Gross point of discriminant $d$. Let $f : K \to \mathcal{D}$ be the embedding corresponding to $y$. This induces a homomorphism $\hat{f} : \hat{K} \to \mathcal{D}$ and we define

$$x_a := (g\hat{f}(a), y).$$

Note that $x_a$ is also of discriminant $d$, and it is easy to check that $x = x_a$ if and only if $a \in \hat{O}_d^\times K^\times$ where $\hat{O}_d := O_d \otimes_A \hat{A}$. Hence $\hat{O}_d^\times \backslash \hat{K}^\times /K^\times \cong \text{Pic}(O_d)$ acts freely on the set $G_d$ of Gross points of discriminant $d$.

The orbit space $G_d / \text{Pic}(O_d)$ is identified with the space of double cosets

$$\hat{\mathcal{R}}^\times \backslash \hat{\mathcal{E}} /\hat{f}(\hat{K}^\times),$$
where \( f : K \rightarrow \mathcal{D} \) is a fixed embedding (if any exist) and

\[
\mathcal{E} := \{ g \in \hat{\mathcal{D}}^\times : f(K) \cap g^{-1} \hat{R}g = f(O_d) \}.
\]

Note that

\[
\hat{R}^\times \backslash \mathcal{E} / f(\hat{K}^\times) = \prod_P R_P^\times \backslash \mathcal{E}_P / f(K_P^\times)
\]

where \( \mathcal{E}_P := \{ g_P \in \mathcal{D}_P^\times : f(K_P) \cap g_P^{-1} R_P g_P = f(O_{d,P}) \} \) and \( O_{d,P} \) is the closure of \( O_d \) in \( K_P := K \otimes_k k_P \).

**Lemma 1.4.** (cf. [16] or [17])

\[
\#(R_P^\times \backslash \mathcal{E}_P / f(K_P^\times)) = \begin{cases} 
1 & \text{if } P \nmid N_0, \\
1 - \left\{ \frac{d}{P} \right\} & \text{if } P \mid N_0.
\end{cases}
\]

Here \( \left\{ \frac{d}{P} \right\} \) is the *Eichler quadratic symbol*, i.e.

\[
\left\{ \frac{d}{P} \right\} = \begin{cases} 
1 & \text{if } P^2 \mid d \text{ or } d \bmod P \in \left( (A/P)^\times \right)^2, \\
-1 & \text{if } d \bmod P \in (A/P)^\times - \left( (A/P)^\times \right)^2, \\
0 & \text{if } P \mid d \text{ but } P^2 \nmid d.
\end{cases}
\]

**Remark.** The above lemma tells us that the number \( \#(G_d) \) is equal to

\[
h(d) \prod_{P \mid N_0} \left( 1 - \left\{ \frac{d}{P} \right\} \right)
\]

where \( h(d) \) is the class number of \( O_d \).

There is a natural action of \( \text{Gal}(K/k) \) on Gross points in the following way: let \( x = (g, y) \) be a Gross point and \( f_y : K \hookrightarrow \mathcal{D} \) be the embedding corresponding to \( y \). Define

\[
x^\sigma = (g, y^\sigma) = (g, y_\sigma)
\]

where \( \sigma \in \text{Gal}(K/k) \) and \( y_\sigma \) corresponds to the embedding \( f_y \circ \sigma \). If \( x \) is a Gross point of discriminant \( d \) in \( X_d \), then so is \( x^\sigma \). Moreover, let \( a \in \hat{O}_d^\times \backslash \hat{K}^\times / K^\times \cong \text{Pic}(O_d) \) and \( \sigma \in \text{Gal}(K/k) \) one has

\[
(x^\sigma)_a = (x_{\sigma(a)})^\sigma.
\]

Therefore we have an action of \( \text{Pic}(O_d) \rtimes \text{Gal}(K/k) \) on the set \( G_d \) of Gross points of discriminant \( d \).
1.2 Hecke correspondences and Gross height pairing

Let $P$ be a prime of $A$. Let $\mathcal{T}$ be the Bruhat-Tits tree of $\text{PGL}_2(k_P)$ as defined in [14]. The vertices are the equivalence classes of $A_P$-lattices $L$ in $k_P^2$, and two such vertices $[L]$ and $[L']$ are adjacent if there exists an integer $r$ such that

$$P^{r+1}L \subseteq L' \subseteq P^rL.$$ 

This is a tree where each vertex has degree $q^{\deg P} + 1$. For a vertex $v$, the Hecke correspondence $t_P$ sends $v$ to the formal sum of its $q^{\deg P} + 1$ neighbors in the tree. Identifying $\text{PGL}_2(A_P) \setminus \text{PGL}_2(k_P)$ with the Bruhat-Tits tree, we can write the Hecke correspondence for $g \in \text{PGL}_2(A_P) \setminus \text{PGL}_2(k_P)$:

$$t_P(g) := \sum_{\deg(u) < \deg P} \left( \begin{array}{cc} 1 & u \\ 0 & P \end{array} \right) g + \left( \begin{array}{cc} P & 0 \\ 0 & 1 \end{array} \right) g.$$

Note that $X_{N_0}$ can be written as

$$\left( \hat{R}^\times \backslash \hat{\mathcal{D}}^\times / \hat{k}^\times \right) \times Y/\mathcal{D}^\times$$

and

$$\hat{R}^\times \backslash \hat{\mathcal{D}}^\times / \hat{k}^\times = \prod_P \hat{R}_P^\times \backslash \hat{\mathcal{D}}_P^\times / k_P^\times.$$ 

When $(P, N_0) = 1$,

$$\hat{R}_P^\times \backslash \hat{\mathcal{D}}_P^\times / k_P^\times \cong \text{PGL}_2(A_P) \setminus \text{PGL}_2(k_P)$$

and so we have the Hecke correspondence $t_P$ on $X_{N_0}$.

Now suppose $P$ divides $N_0$, then $\hat{R}_P^\times \backslash \hat{\mathcal{D}}_P^\times / k_P^\times$ has two elements and define the Atkin-Lehner involution

$$w_P(g, y) := (g', y)$$

where $g'$ is another double coset in $\hat{R}_P^\times \backslash \hat{\mathcal{D}}_P^\times / k_P^\times$.

From the construction, these correspondences commute with each other. Therefore we can define Hecke correspondence $t_m$ for every non-zero ideal $(m)$ of $A$ in the following way:

$$\begin{cases} t_{mm'} = t_m t_{m'} & \text{if } m \text{ and } m' \text{ are relatively prime}, \\ t_{pt} = t_{pt-1} t_P - q^{\deg P} t_{pt-2} & \text{for } P \mid N_0, \\ t_{pt} = w_P^t & \text{for } P \mid N_0. \end{cases}$$

Note that $X = X_{N_0} = \coprod_{i=1}^n X_i$, where $n$ is the left ideal class number of $R$. Consider the Picard group $\text{Pic}(X)$, which is isomorphic to $\mathbb{Z}^n$ and is generated
by the classes $e_i$ of degree 1 corresponding to the component $X_i$. Then the correspondences $t_m$ induce endomorphisms of the group $\text{Pic}(X)$. In fact, with respect to the basis $\{e_1, ..., e_n\}$, these endomorphisms can be represented by Brandt matrices.

Let $\{I_1, ..., I_n\}$ be a set of left ideals of $R$ representing the distinct ideal classes, with $I_1 = R$. Let $w_i := \#(R^i)/(q - 1)$ where $R_i$ is the right order of $I_i$. Consider $M_{ij} := I_i^{-1} I_j$, which is a left ideal of $R_j$ with right order $R_i$. Choose a generator $N_{ij} \in k$ of the reduced ideal norm $\text{Nr}(M_{ij}) := \langle \text{Nr}(b) : b \in M_{ij} \rangle$ of $M_{ij}$. For each monic polynomial $m$ in $A$, define

$$B_{ij}(m) := \frac{\# \{ b \in M_{ij} : \text{Nr}(b)/N_{ij} = (m) \}}{(q - 1)w_j}$$

and the $m$-th Brandt matrix

$$B_{ij}(m) := \left( B_{ij}(m) \right)_{1 \leq i, j \leq n} \in \text{Mat}_n(\mathbb{Z}).$$

**Proposition 1.5.** For all non-zero ideal $(m)$ in $A$ and $i = 1, 2, ..., n$,

$$t_m e_i = \sum_{j=1}^{n} B_{ij}(m) e_j.$$ 

**Proof.** From the definition of $t_m$ and the recurrence relations of $B(m)$ (cf. [10]), we can reduce the proof to the case when $m = P$ is a prime. From the following bijection

$$\hat{R}^x \setminus \mathfrak{D}^x \cong \{ \text{left ideals of } R \}$$

$$\hat{R}^x \mathfrak{g} \leftrightarrow I_g := \hat{R}_g \cap \mathfrak{D},$$

for any element $g$ in $\hat{D}^x$ we can identify the following set

$$\left\{ \hat{R}^x \begin{pmatrix} 1 & u \\ 0 & P \end{pmatrix} g : \deg u < \deg P \right\} \cup \left\{ \hat{R}^x \begin{pmatrix} P & 0 \\ 0 & 1 \end{pmatrix} g \right\}$$

with

$$\{ \text{left ideal } I \text{ of } R \text{ contained in } I_g \text{ with } \text{Nr}(I) = P \text{Nr}(I_g) \}.$$ 

According to the definition of $t_P$, $t_P e_i = \sum_j \alpha_j e_j$ where $\alpha_j$ is the number of left ideals $I$ of $R$ equivalent to $I_j$ which are contained in $I_i$ with $\text{Nr}(I) = P \text{Nr}(I_j)$. It is easy to see that $\alpha_j = B_{ij}(P)$ and so the proposition holds.

We define the Gross height pairing $< \cdot, \cdot >$ on $\text{Pic}(X)$ with values in $\mathbb{Z}$ by setting

$$< e_i, e_j > := \begin{cases} 0 & \text{if } i \neq j, \\ w_i & \text{if } i = j, \end{cases}$$

and extending bi-additively. Therefore $\text{Pic}(X)^\vee := \text{Hom}(\text{Pic}(X), \mathbb{Z})$ can be viewed as a subgroup of $\text{Pic}(X) \otimes_{\mathbb{Z}} \mathbb{Q}$ with basis $\{ e_i^* := e_i/w_i : i = 1, ..., n \}$ via this pairing. Since $w_j B_{ij}(m) = w_i B_{ji}(m)$, one has the following proposition.
Proposition 1.6. For all classes $e$ and $e'$ in $\text{Pic}(X)$, we have
\[ \langle t_m e, e' \rangle = \langle e, t_m e' \rangle. \]

Proof. Since $w_j B_{ij}(m) = w_i B_{ji}(m)$, we have
\[ \langle t_m e_i, e_j \rangle = \langle e_i, t_m e_j \rangle \]
for all $i, j$ and the result holds.

Let $d \in A$ with $d \not\in k_2^\infty$. Assume every prime factor $P$ of $N_0$ is not split in $K$ and $P^2$ does not divides $d$ (i.e. the set $G_d$ of Gross points of discriminant $d$ is not empty). For any prime $P \mid N_0$, one has $w_P(G_d) = G_d$. Suppose $P_1, \ldots, P_r$ are primes dividing $N_0$ and inert in $K$. We have in fact a free action of $\text{Pic}(O_d) \times \prod_{i=1}^r \langle w_{P_i} \rangle$ on $G_d$. Since $w_{P_i}$ are of order 2 for all $i$, $\text{Pic}(O_d) \times \prod_{i=1}^r \langle w_{P_i} \rangle$ acts simply transitively on $G_d$.

Let $a \in A$ with $a \not\in k_2^\infty$. Consider the rational divisor
\[ c_a := \sum_{a = df^2, f \text{ monic}} \frac{1}{2u(d)} \sum_{x_d \in G_d} x_d. \]
Here $u(d) = \#(O_d^*)$. By calculation one has
\[ \deg(c_a) = \frac{1}{2} \sum_{a = df^2, f \text{ monic}} \left[ \frac{h(d)}{u(d)} \cdot \prod_{P \mid N_0} \left( 1 - \frac{d}{P} \right) \right]. \]

Let $e_a \in \text{Pic}(X) \otimes \mathbb{Q}$ be the class of the divisor $c_a$. It can be shown that

Proposition 1.7. The class $e_a$ lies in $\text{Pic}(X)^\vee$, which is considered as a subgroup of $\text{Pic}(X) \otimes \mathbb{Q}$.

Note that we can extend the Gross height pairing to $\text{Pic}(X) \otimes \mathbb{C}$ which is linear in the first term and conjugate linear in the second. In the next section this pairing gives us a construction of automorphic forms of Drinfeld type.

2 Automorphic forms of Drinfeld type and main theorem

2.1 Automorphic forms of Drinfeld type

Consider the open compact subgroup $\mathcal{K}_0(N\infty) := \prod_P \mathcal{K}_{0,P} \times \Gamma_\infty$ of $\text{GL}_2(A_k)$, where
\[ \mathcal{K}_{0,P} := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{GL}_2(A_P) : c \in NA_P \right\} \]
for finite prime $P$, and
\[ \Gamma_\infty := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{GL}_2(O_\infty) : c \in \pi_\infty O_\infty \right\}. \]
An automorphic form \( f \) on \( \text{GL}_2(\mathbb{A}_k) \) for \( \mathcal{K}_0(N\infty) \) (with trivial central character) is a \( \mathbb{C} \)-valued function on the double coset space

\[
\text{GL}_2(k) \backslash \text{GL}_2(\mathbb{A}_k) / \mathcal{K}_0(N\infty)k_\infty^\times.
\]

Note that by strong approximation theorem (cf. [16]) we have the following bijection

\[
\text{GL}_2(k) \backslash \text{GL}_2(\mathbb{A}_k) / \mathcal{K}_0(N\infty)k_\infty^\times \cong \Gamma_0(N) \backslash \text{GL}_2(k_\infty) / \Gamma_\infty k_\infty^\times
\]

where

\[
\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{GL}_2(\mathbb{A}) : c \equiv 0 \mod N \right\}.
\]

Therefore \( f \) can be viewed as a \( \mathbb{C} \)-valued function on \( \Gamma_0(N) \backslash \text{GL}_2(k_\infty) / \Gamma_\infty k_\infty^\times \). From now on, we call \( f \) an automorphic form for \( \Gamma_0(N) \) if \( f \) is a function on the space of double cosets \( \Gamma_0(N) \backslash \text{GL}_2(k_\infty) / \Gamma_\infty k_\infty^\times \). An automorphic form \( f \) for \( \Gamma_0(N) \) is called a cusp form if for every \( g_\infty \in \text{GL}_2(k_\infty) \) and \( \gamma \in \text{GL}_2(\mathbb{A}) \)

\[
\int_{\text{A}\backslash k_\infty} f \left( \gamma \begin{pmatrix} 1 & h_\gamma x \\ 0 & 1 \end{pmatrix} g_\infty \right) dx = 0.
\]

Here \( du \) is a Haar measure with \( \int_{\text{A}\backslash k_\infty} du = 1 \) and \( h_\gamma \) is a generator of the ideal of \( A \) which is maximal for the property that

\[
\gamma \begin{pmatrix} 1 & h_\gamma A \\ 0 & 1 \end{pmatrix} \gamma^{-1} \subset \Gamma_0(N).
\]

Note that the coset space \( \text{GL}_2(k_\infty) / \Gamma_\infty k_\infty^\times \) can be represented by the two disjoint sets

\[
\mathcal{T}_+ := \left\{ \begin{pmatrix} \pi^r_\infty & u \\ 0 & 1 \end{pmatrix} : r \in \mathbb{Z}, u \in k_\infty / \pi^r_\infty \mathcal{O}_\infty \right\}
\]

and

\[
\mathcal{T}_- := \left\{ \begin{pmatrix} \pi^r_\infty & u \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ \pi_\infty & 0 \end{pmatrix} : r \in \mathbb{Z}, u \in k_\infty / \pi^r_\infty \mathcal{O}_\infty \right\}.
\]

DEFINITION 2.1. An automorphic form \( f \) on \( \text{GL}_2(k_\infty) \) is of Drinfeld type if it satisfies the following harmonic properties: for any \( g_\infty \in \text{GL}_2(k_\infty) \) we have

\[
\tilde{f}(g_\infty) := f(g_\infty \begin{pmatrix} 0 & 1 \\ \pi_\infty & 0 \end{pmatrix}) = -f(g_\infty) \quad \text{and} \quad \sum_{\kappa \in \text{GL}_2(\mathcal{O}_\infty) / \Gamma_\infty} f(g_\infty \kappa) = 0.
\]
Suppose a function $f : \begin{pmatrix} 1 & A \\ 0 & 1 \end{pmatrix} \backslash \mathrm{GL}_2(k_\infty)/\Gamma_\infty k_\infty \to \mathbb{C}$ is given. Recall the Fourier expansion of $f$ (cf. [18]): for $r \in \mathbb{Z}$ and $u \in k_\infty$,

$$f \left( \begin{pmatrix} \pi^r & u \\ 0 & 1 \end{pmatrix} \right) = \sum_{\lambda \in A} f^*(r, \lambda) \psi_\infty(\lambda u)$$

where

$$f^*(r, \lambda) := \int_{A \backslash k_\infty} f \left( \begin{pmatrix} \pi^r & u \\ 0 & 1 \end{pmatrix} \right) \psi_\infty(-\lambda u) du.$$ 

Here $\psi_\infty$ is the fixed additive character on $k_\infty$ in the notation table. Since $f(g \gamma_\infty) = f(g)$ for all $\gamma_\infty \in \Gamma_\infty$, $f^*(r, \lambda) = 0$ if $\deg \lambda + 2 > r$. Moreover, if $f$ satisfies harmonic properties, then

$$f^*(r, \lambda) = q^{-r + \deg \lambda + 2} f^*(\deg \lambda + 2, \lambda)$$

if $\deg \lambda + 2 \leq r$.

2.1.1 Example: Theta series

Fix a definite quaternion algebra $\mathcal{D} = \mathcal{D}(N_0)$ where $N_0$ is the product of finite ramified primes of $\mathcal{D}$. Let $R$ be a maximal order and $n$ be the class number.

With representatives of left ideal classes fixed in §1.2, we have introduced for each $(i, j)$, the ideal $M_{ij}$ of $\mathcal{D}$ and chose a generator $N_{ij}$ of the fractional ideal $\text{Nr}(M_{ij})$. For $1 \leq i, j \leq n$ and $(x, y) \in k_\infty \times k_\infty$, define

$$\theta_{ij}(x, y) := \sum_{b \in M_{ij}} \phi_\infty \left( \frac{\text{Nr}(b)}{N_{ij}} x t^2 \right) \cdot \psi_\infty \left( \frac{\text{Nr}(b)}{N_{ij}} y \right),$$

where $\phi_\infty$ is the characteristic function of $\mathcal{O}_\infty$. It is easy to obtain the following properties:

1. $\theta_{ij}(x, y) = \sum_{\lambda \in A} B'_{ij}(\lambda) \psi_\infty(\lambda y)$

where for each $\lambda \in A$,

$$B'_{ij}(\lambda) = \# \{ b \in M_{ij} : \text{Nr}(b)/N_{ij} = \lambda \}.$$

2. $\theta_{ij}(x, y + h) = \theta_{ij}(x, y)$ for $h \in A$.
3. $\theta_{ij}(ax, bx + y) = \theta_{ij}(x, y)$ for $a \in \mathcal{O}_\infty^\times, \beta \in \mathcal{O}_\infty$.

Basing on Poisson summation formula, we have the following transformation law for $\theta_{ij}$ (cf. Appendix B):
Proposition 2.2. Let \((x, y) \in k_\infty^x \times k_\infty\) and \(\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}_2(A)\). Suppose \(v_\infty(cx) > v_\infty(cy + d)\) and \(c \equiv 0 \mod N_0\). Then for \(1 \leq i, j \leq n\),

\[
\theta_{ij} \left( \frac{x}{(cy + d)^2} \bigg(\begin{array}{c} ay + b \\ cy + d \end{array}\right) \right) = q^{-2v_\infty(cy + d)} \cdot \theta_{ij}(x, y).
\]

For \(g_\infty \in \text{GL}_2(k_\infty)\), write \(g_\infty\) as \(\gamma \begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix} \gamma_\infty z_\infty\), where \(\gamma\) is in \(\Gamma_0(N_0)\), \((x, y)\) is in \(k_\infty^x \times k_\infty\), \(\gamma_\infty\) is in \(\Gamma_\infty\), and \(z_\infty\) is in \(k_\infty^x\). We introduce the theta series \(\Theta_{ij}\) for \(M_{ij}\):

\[
\Theta_{ij}(g_\infty) := \frac{1}{(q - 1)w_j} \cdot q^{-v_\infty(x)} \cdot \left( \sum_{\epsilon \in F_q^\times} \theta_{ij}(x, \epsilon y) \right)
= q^{-v_\infty(x)} \cdot \left[ \frac{1}{w_j} + \sum_{m \in A \text{ monic}, \deg m + 2 \leq v_\infty(x)} B_{ij}(m) \left( \sum_{\epsilon \in F_q^\times} \psi_\infty(\epsilon m y) \right) \right].
\]

The last equality follows from \(B'_{ij}(0) = 1\) and for each monic polynomial \(m \in A\),

\[(q - 1)w_j \cdot B_{ij}(m) = \sum_{\epsilon \in F_q^\times} B'_{ij}(\epsilon m)\]

The transformation law of \(\theta_{ij}\) tells us that

Lemma 2.3. \(\Theta_{ij}\) is a well-defined \(\mathbb{Q}\)-valued function on the double coset space \(\Gamma_0(N_0) \backslash \text{GL}_2(k_\infty) / \Gamma_\infty k_\infty^x\).

Proof. Let \(g_\infty\) be an element in \(\text{GL}_2(k_\infty)\). Suppose

\[
g_\infty = \gamma_1 \begin{pmatrix} x_1 & y_1 \\ 0 & 1 \end{pmatrix} \gamma_{\infty,1} z_1 = \gamma_2 \begin{pmatrix} x_2 & y_2 \\ 0 & 1 \end{pmatrix} \gamma_{\infty,2} z_2,
\]

where for \(i = 1, 2\), \(\gamma_i \in \Gamma_0(N_0)\), \((x_i, y_i) \in k_\infty^x \times k_\infty^x\), \(\gamma_{\infty,i} \in \Gamma_\infty\), \(z_i \in k_\infty^x\). We need to show that

\[
q^{-v_\infty(x_1)} \cdot \left( \sum_{\epsilon \in F_q^\times} \theta_{ij}(x_1, \epsilon y_1) \right) = q^{-v_\infty(x_2)} \cdot \left( \sum_{\epsilon \in F_q^\times} \theta_{ij}(x_2, \epsilon y_2) \right).
\]

Set \(\gamma = \gamma_2^{-1} \gamma_1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, z = z_1^{-1} z_2\), and \(\gamma_\infty = \gamma_{\infty,1}^{-1} \gamma_{\infty,2}\). Then one has \(v_\infty(cx_1) > v_\infty(cy_1 + d)\) and

Documenta Mathematica 16 (2011) 723–765
\[ \gamma \left( \begin{array}{cc} x_1 & y_1 \\ 0 & 1 \end{array} \right) = \left( \begin{array}{cc} \det \gamma \cdot x_1 & ay_1 + b \\ (cy_1 + d)^2 & cy_1 + d \end{array} \right) \left( \begin{array}{cc} 1 & 0 \\ cy_1 + d & 1 \end{array} \right) \left( \begin{array}{cc} c_y & d \\ cy_1 + d & 1 \end{array} \right) \left( \begin{array}{cc} c_y & d \\ cy_1 + d & 1 \end{array} \right) \right) \]

\[ = \left( \begin{array}{cc} x_2 & y_2 \\ 0 & 1 \end{array} \right) \gamma_{\infty z}. \]

Therefore \( v_{\infty}(x_2) = v_{\infty}(x_1) - 2v_{\infty}(cy_1 + d) \), and the properties of \( \theta_{ij} \) implies

\[ \theta_{ij}(x_2, \epsilon y_2) = \theta_{ij} \left( \frac{\det \gamma \cdot x_1}{(cy_1 + d)^2}, \epsilon cy_1 + d \right). \]

for each \( \epsilon \in \mathbb{F}_q^{\times} \). Hence the transformation law of \( \theta_{ij} \) in Proposition 2.2 shows

\[ q^{-v_{\infty}(x_2)} \cdot \left( \sum_{\epsilon \in \mathbb{F}_q^{\times}} \theta_{ij}(x_2, \epsilon y_2) \right) = q^{-v_{\infty}(x_1)} \cdot \left( \sum_{\epsilon \in \mathbb{F}_q^{\times}} \theta_{ij}(\det \gamma \cdot x_1, \epsilon \det \gamma \cdot y_1) \right) \]

\[ = q^{-v_{\infty}(x_1)} \cdot \left( \sum_{\epsilon \in \mathbb{F}_q^{\times}} \theta_{ij}(x_1, \epsilon y_1) \right). \]

\[ \square \]

The Fourier coefficients of \( \Theta_{ij} \) can be easily read off from Brandt matrices: for each \( r \in \mathbb{Z} \) and \( \lambda \in A \) with \( \deg \lambda + 2 \leq r \) the Fourier coefficients

\[ \Theta^*_{ij}(r, \lambda) = \begin{cases} q^{-r} B_{ij}(m) & \text{if } (\lambda) = (m) \neq 0, \\ q^{-r} / w_j & \text{if } \lambda = 0. \end{cases} \]

Therefore \( \Theta^*_{ij}(r + 1, \lambda) = q^{-1} \Theta^*_{ij}(r, \lambda) \) for all \( \lambda \in A \) with \( \deg \lambda + 2 \leq r \).

In fact, \( \Theta_{ij} \) are of Drinfeld type for all \( 1 \leq i, j \leq n \). To show the harmonicity of \( \Theta_{ij} \), by Lemma 2.13, it is enough to prove that for all \( g_{\infty} \in \text{GL}_2(k_{\infty}) \)

\[ \tilde{\Theta}_{ij}(g_{\infty}) = -\Theta_{ij}(g_{\infty}). \]

Let \( \pi_{\infty}^c \in k_{\infty}^c \) and \( u \in k_{\infty} \). Choose \( c, d \in A \) with \( c \equiv 0 \mod N_0 \), \( (c, d) = 1 \), \( v_{\infty}(u + \bar{u}) \geq r + 1 \), and find \( a, b \in A \) with \( ad - bc = 1 \). Then for \( \ell \in \mathbb{Z} \) with \( \ell \leq r + 1 \) the following two matrices:

\[ \left( \begin{array}{cc} \pi_c^\ell & u \\ 0 & \pi_{\infty} \end{array} \right) \left( \begin{array}{cc} 0 & 1 \\ \pi_{\infty} & 0 \end{array} \right) \quad \text{and} \quad \left( \begin{array}{cc} d & -b \\ -c & a \end{array} \right) \left( \begin{array}{cc} s_{1-\ell} & \bar{a} \\ \bar{c} & \bar{e} \end{array} \right) \left( \begin{array}{cc} a & \bar{c} \\ 0 & 1 \end{array} \right) \]

\[ \text{Documenta Mathematica 16 (2011) 723–765} \]
represent the same coset in \( \text{GL}_2(k_\infty)/\Gamma_\infty k_\infty^\times \). Using this fact for \( \ell = r \) and \( \ell = r + 1 \) one obtains

\[
\tilde{\Theta}_{ij} \left( \begin{array}{cc} \pi_\infty^r & u \\ 0 & 1 \end{array} \right) - q^{-1} \tilde{\Theta}_{ij} \left( \begin{array}{cc} \pi_\infty^{r+1} & u \\ 0 & 1 \end{array} \right) = \sum_{\deg \mu + 2 = 1-r+2 \deg c} \Theta_{ij}^* (1 - r + 2 \deg c, \mu) \psi_\infty (\mu \frac{a}{c}).
\]

Set \( u_\epsilon := -\frac{d}{c} + \epsilon \pi_\infty^r \) for \( \epsilon \in \mathcal{F}_q^\times \). From the identity

\[
\left( \frac{\pi_\infty^r}{c} \right) \left( \begin{array}{cc} a u_\epsilon + b \\ c u_\epsilon + d \end{array} \right) = \frac{a}{c} - \frac{1}{c^2 \epsilon \pi_\infty^r},
\]

and summing over all \( \epsilon \) we get:

\[
(q-1) \tilde{\Theta}_{ij} \left( \begin{array}{cc} \pi_\infty^r & 0 \\ 0 & 1 \end{array} \right) = \sum_{\epsilon \in \mathcal{F}_q^\times} \Theta_{ij} \left( \begin{array}{cc} \frac{\pi_\infty^{1-r}}{c} \\ 0 \end{array} \right) \left( \begin{array}{cc} a u_\epsilon + b \\ c u_\epsilon + d \end{array} \right)
\]

\[
= q \sum_{\deg \mu + 2 = 1-r+2 \deg c} \Theta_{ij}^* (1 - r + 2 \deg c, \mu) \psi_\infty (\mu \frac{a}{c}).
\]

Note that

\[
\left( \begin{array}{cc} \frac{\pi_\infty^{1-r}}{c} & 0 \\ 0 & 1 \end{array} \right) \text{ and } \left( \begin{array}{cc} a & b \\ c & d \end{array} \right) \left( \begin{array}{cc} \pi_\infty^{r+1} & u_\epsilon \\ 0 & 1 \end{array} \right)
\]

represent the same coset in \( \text{GL}_2(k_\infty)/\Gamma_\infty k_\infty^\times \). Thus one has

\[
\tilde{\Theta}_{ij} \left( \begin{array}{cc} \pi_\infty^{r+1} & u \\ 0 & 1 \end{array} \right) - \tilde{\Theta}_{ij} \left( \begin{array}{cc} \pi_\infty^r & u \\ 0 & 1 \end{array} \right) = \sum_{\epsilon \in \mathcal{F}_q^\times} \Theta_{ij} \left( \begin{array}{cc} \pi_\infty^{r+1} & u + \epsilon \pi_\infty^r \\ 0 & 1 \end{array} \right).
\]

From the Fourier expansion of \( \tilde{\Theta}_{ij} \) and \( \Theta_{ij} \) we have that for \( \lambda \in A \) with \( \deg \lambda + 2 \leq r \),

\[
\tilde{\Theta}_{ij}^* (r + 1, \lambda) - \tilde{\Theta}_{ij}^* (r, \lambda) = (q-1) \Theta_{ij}^* (r + 1, \lambda),
\]

and for \( \deg \lambda + 2 = r + 1 \),

\[
\tilde{\Theta}_{ij}^* (\deg \lambda + 2, \lambda) = -\Theta_{ij}^* (r + 1, \lambda).
\]

Therefore \( \tilde{\Theta}_{ij}^* (r, \lambda) = -\Theta_{ij}^* (r, \lambda) \) for \( \lambda \in A \) with \( \lambda \neq 0 \) and \( r \geq \deg \lambda + 2 \).

To compute \( \tilde{\Theta}_{ij}^* (r, 0) \), note that

\[
\tilde{\Theta}_{ij} \left( \begin{array}{cc} \pi_\infty^r & 0 \\ 0 & 1 \end{array} \right) = \sum_{\deg \lambda \leq r-2} \tilde{\Theta}_{ij}^* (r, \lambda)
\]

\[
= \tilde{\Theta}_{ij}^*(r, 0) + \sum_{\lambda \neq 0, \deg \lambda \leq r-2} -\Theta_{ij}^* (r, \lambda).
\]
On the other hand, for any $\epsilon \in F_q^\times$ and $\ell \geq 0$ the following two matrices
\[
\begin{pmatrix}
\pi^{\deg N_0 + \ell} & 0 \\
0 & 1
\end{pmatrix}
\begin{pmatrix}
0 & 1 \\
\pi_\infty & 0
\end{pmatrix},
\begin{pmatrix}
\epsilon^{-1} & -1 \\
-t^\ell N_0 & \epsilon(t^\ell N_0 + 1)
\end{pmatrix}
\begin{pmatrix}
\pi^{1-\deg N_0 - \ell} N_0 & \epsilon(t^\ell N_0 + 1) \\
0 & 1
\end{pmatrix}
\]
represent the same coset in $GL_2(k_{\infty})/\Gamma_{\infty}$. Therefore
\[
\tilde{\Theta}_{ij} \begin{pmatrix}
\pi^{\deg N_0 + \ell} & 0 \\
0 & 1
\end{pmatrix} = \sum_{\deg \lambda \leq \deg N_0 + \ell - 1} \Theta^*_ij(\deg N_0 + \ell + 1, \lambda) \psi(\lambda \frac{\epsilon}{t^\ell N_0})
\]
\[
= \sum_{\deg \lambda \leq \deg N_0 + \ell - 2} \Theta^*_ij(\deg N_0 + \ell + 1, \lambda) - \frac{1}{q-1} \sum_{\deg \lambda = \deg N_0 + \ell - 1} \Theta^*_ij(\deg N_0 + \ell + 1, \lambda).
\]
This gives
\[
\tilde{\Theta}^*_ij(\deg N_0 + \ell, 0) = \Theta^*_ij(\deg N_0 + \ell + 1, 0) + \frac{1}{q - 1} \left[ q \Theta_{ij} \begin{pmatrix}
\pi^{\deg N_0 + \ell} & 0 \\
0 & 1
\end{pmatrix} - \Theta_{ij} \begin{pmatrix}
\pi^{\deg N_0 + \ell + 1} & 0 \\
0 & 1
\end{pmatrix} \right]
\]
Using the fact that $M_{ij}$ is discrete and cocompact in $D_{\infty} = D \otimes_k k_{\infty}$, it can be deduced that for sufficiently large $s$ one has
\[
q \Theta_{ij} \begin{pmatrix}
\pi^s & 0 \\
0 & 1
\end{pmatrix} = \Theta_{ij} \begin{pmatrix}
\pi^{s+1} & 0 \\
0 & 1
\end{pmatrix}.
\]
Thus from the equality $\tilde{\Theta}^*_ij(r + 1, 0) - \tilde{\Theta}^*_ij(r, 0) = (q - 1)\Theta^*_ij(r + 1, 0)$ for all $r \in Z$ one has
\[
\tilde{\Theta}^*_ij(r, 0) = -\Theta^*_ij(r, 0).
\]
Comparing the Fourier coefficients we obtain $\tilde{\Theta}_{ij} = -\Theta_{ij}$ and hence $\Theta_{ij}$ is of Drinfeld type for any $1 \leq i, j \leq n$. 

Documenta Mathematica 16 (2011) 723–765
2.2 Hecke operators

Let $f$ be an automorphic form on $GL_2(k_{\infty})$ for $\Gamma_0(N)$. For each prime $P$ of $A$, the Hecke operator $T_P$ is defined by:

$$T_P f(g) := \sum_{\deg u < \deg P} f\left( \begin{pmatrix} 1 & u \\ 0 & P \end{pmatrix} \cdot g \right) + f\left( \begin{pmatrix} P & 0 \\ 0 & 1 \end{pmatrix} \cdot g \right) \quad \text{if } P \nmid N,$$

$$T_P f(g) := \sum_{\deg u < \deg P} f\left( \begin{pmatrix} 1 & u \\ 0 & P \end{pmatrix} \cdot g \right) \quad \text{if } P \mid N.$$

Note that the Fourier coefficients of $T_P f$ are of the form:

$$(T_P f)^*(r, \lambda) = q^{\deg(P)} \cdot f^*(r + \deg(P), \lambda) + f^*(r - \deg(P), \frac{P}{\lambda}) \quad \text{if } P \nmid N,$$

$$(T_P f)^*(r, \lambda) = q^{\deg(P)} \cdot f^*(r + \deg(P), \lambda) \quad \text{if } P | N.$$

Here $f^*(\pi_{\infty}, \frac{1}{\lambda}) = 0$ if $P \nmid \lambda$. Since $T_P$ and $T_P'$ commute, we can define Hecke operators $T_m$ for monic polynomial $m$ in $A$ as follows:

\[
\begin{cases}
T_{mm'} = T_m T_{m'} & \text{if } m \text{ and } m' \text{ are relatively prime}, \\
T_{P^t} = T_{P^{-t} P} - q^{\deg(P)} T_{P^{-2}} & \text{for } P \nmid N, \\
T_{P^t} = T_P & \text{for } P | N.
\end{cases}
\]

We point out that if $f$ is of Drinfeld type, then so is $T_m f$ for any monic polynomial $m$ (cf. [7] Section 4.9).

When $T_m$ acts on $\Theta_{ij}$, we get:

**Proposition 2.4.** For any monic polynomial $m$ in $A$,

$$T_m \Theta_{ij} = \sum_{\ell} B_{\ell}(m) \Theta_{ij} = \sum_{\ell} B_{\ell}(m) \Theta_{\ell i}.$$

**Proof.** The second identity will follow from the first, as

$$w_i \Theta_{ij} = w_i \Theta_{ji} \text{ and } w_i B_{\ell}(m) = w_i B_{\ell i}(m).$$

Note that the Hecke operators $T_m$ satisfy the same relations as the matrices $B(m)$. Moreover, from the recurrence relations of Brandt matrices (cf. [10]) we have

$$\sum_{\ell} B_{\ell}(P) B_{\ell j}(m) = B_{ij}(mP) + q^{\deg(P)} B_{ij}(m/P) \quad \text{if } P \nmid N_0,$$

$$\sum_{\ell} B_{\ell}(P) B_{\ell j}(m) = B_{ij}(mP) \quad \text{if } P | N_0.$$

Comparing the Fourier coefficients the result holds. \qed
Remark. Let $E_{N_0} := \sum_{j=1}^n \Theta_{ij}$ (which is independent of the choice of $i$). For $r \in \mathbb{Z}$ and $\lambda \in A$ with $\deg \lambda + 2 \leq r$ the Fourier coefficients are

$$E_{N_0}^*(r, \lambda) = q^{-r} \sigma(\lambda)_{N_0}$$

where

$$\sigma(\lambda)_{N_0} = \sum_{(m, N_0) = 1}^n q^{\deg m},$$

and

$$E_{N_0}^*(r, 0) = q^{-r} \sum_{j=1}^n \frac{1}{w_j}.$$

Moreover, from Proposition 2.4 we have

$$T_m E_{N_0} = \sigma(m)_{N_0} E_{N_0}$$

for all monic polynomials $m$ in $A$. This tells us that the function $E_{N_0}$, which is an analogue of Eisenstein series, generates a one-dimensional eigenspace for all Hecke operators. We point out that suppose $N_0 = \prod_{i=1}^\ell P_i$, by comparing the Fourier coefficients one gets

$$q^2 E_{N_0}(g_\infty) = E(g_\infty) + \sum_{i=1}^\ell (-1)^i \left[ \sum_{1 \leq j_1 < \cdots < j_i \leq \ell} E \left( \begin{pmatrix} P_{j_1} \cdots P_{j_i} & 0 \\ 0 & 1 \end{pmatrix} g_\infty \right) \right]$$

for $g_\infty \in \text{GL}_2(k_\infty)$ where $E$ is the improper Eisenstein series introduced in [5].

For each non-zero ideal $N$ of $A$, recall the Petersson inner product, which is a non-degenerate pairing on the finite dimensional $\mathbb{C}$-vector space $S(\Gamma_0(N))$ of automorphic cusp forms of Drinfeld type for $\Gamma_0(N)$,

$$(f, g) := \int_{G_0(N)} f \cdot \overline{g}.$$

Here $G_0(N) = \Gamma_0(N) \setminus \text{GL}_2(k_\infty)/\Gamma_\infty k_\infty^\times$. The measure on $G_0(N)$ is taken by counting the size of the stabilizer of an element (cf. [7] §4.8). More precisely, let $\Gamma$ be a congruence subgroup and $e \in \text{GL}_2(k_\infty)/\Gamma_\infty k_\infty^\times$. We denote $\text{Stab}_\Gamma(e)$ the stabilizer of $e$ in $\Gamma$, which is a finite subgroup in $\Gamma$. One takes the measure $d([e])$ of each double coset $[e]$ in $\Gamma \setminus \text{GL}_2(k_\infty)/\Gamma_\infty k_\infty^\times$ where

$$d([e]) := \frac{\#(Z(\Gamma))}{\#(\text{Stab}_\Gamma(e))}.$$

Here $Z(\Gamma)$ is the subgroup of scalar matrices in $\Gamma$. When $\Gamma = \Gamma_0(N)$, for $f$ and $g$ in $S(\Gamma_0(N))$,

$$(f, g) = \sum_{[e] \in G_0(N)} f(e) \overline{g(e)} d([e]).$$
2.5. An old form is a linear combinations of forms

\[ f' \left( \begin{pmatrix} d & 0 \\ 0 & 1 \end{pmatrix} g_\infty \right) \]

for \( g_\infty \in \text{GL}_2(k_\infty) \), where \( f' \) is an automorphic cusp form of Drinfeld type for \( \Gamma_0(M), M|N, M \neq N \), and \( d|(N/M) \). An automorphic cusp form \( f \) of Drinfeld type for \( \Gamma_0(N) \) is called a new form if for any old form \( f' \) one has

\[ (f, f') = 0. \]

If \( f \) is a new form which is also a Hecke eigenform, then \( f \) is called a newform.

It is known that the dimension of Drinfeld type cusp forms for \( \Gamma_0(N) \) is equal to the genus of the Drinfeld modular curve \( X_0(N) \) (cf. [7]). Let \( S^\text{new}(\Gamma_0(N)) \) be the space of new forms for \( \Gamma_0(N) \) and \( h_N \) be the number of left ideal classes of the maximal order \( R \). As in the classical case, we can deduce that

\[ h_N = \frac{1}{q^2 - 1} \prod_{P|N} (q^{\deg P} - 1) - \frac{q}{2(q + 1)} \prod_{P|N} (1 - (-1)^{\deg P}). \]

From the genus formula of \( X_0(N) \) in [5], the dimension of \( S^\text{new}(\Gamma_0(N)) \) is equal to \( h_N - 1 \).

In the next subsection we will give our main theorem, which is essentially a construction of the space \( S^\text{new}(\Gamma_0(N)) \) of new forms for \( \Gamma_0(N) \) via the theta series \( \Theta_{ij} \).

2.3 Main theorem

Consider the definite Shimura curve \( X = X_{N_0} \), introduced in [4]. Recall the height pairing

\[ <e, e'> = \sum_i a_i a'_i \]

where \( e \in \text{Pic}(X) \) with \( e = \sum_i a_i e_i \) and \( e' \in \text{Pic}(X)^\vee \) with \( e' = \sum_i a'_i e_i \).

Let \( M(\Gamma_0(N_0)) \) be the space of automorphic forms of Drinfeld type for \( \Gamma_0(N_0) \). Define \( \Phi : \text{Pic}(X) \times \text{Pic}(X)^\vee \to M(\Gamma_0(N_0)) \) by

\[ \Phi(e, e') := q^2 \sum_{i,j} a_i a'_j \Theta_{ij} \]

for any \( e \in \text{Pic}(X) \) with \( e = \sum_i a_i e_i \) and \( e' \in \text{Pic}(X)^\vee \) with \( e' = \sum_i a'_i e_i \). Then for \( r \in \mathbb{Z} \) and \( u \in k_\infty \) we have the following Fourier expansion

\[ \Phi(e, e') \left( \begin{pmatrix} \pi_\infty^r & u \\ 0 & 1 \end{pmatrix} \right) = q^{-r+2} \left( \deg e \cdot \deg e' + \sum_{m \text{ prime}, \deg m \leq r-2} <e, e'> \sum_{\tau \in \mathbb{P}^1_q} \psi_\infty(\tau m u) \right). \]
Since
\[ < t_m e, e' > = < e, t_m e' > \]
for any monic polynomial \( m \in A \), by Proposition 2.4 one has
\[ T_m (\Phi(e, e')) = \Phi(t_m e, e') = \Phi(e, t_m e'). \]
In fact, the image of \( \Phi \) is in \( M_{\text{new}}(\Gamma_0(N_0)) := S_{\text{new}}(\Gamma_0(N_0)) \oplus \mathbb{C}E_{N_0} \). To see this, we need the following claim.

**Claim:** for any monic \( m \), consider \( t_m \) as in \( \text{End}(\text{Pic}(X)) \) and restrict \( T_m \) to the subspace \( M_{\text{new}}(\Gamma_0(N_0)) \). We have
\[ \text{Tr} t_m = \text{Tr} T_m. \]
This claim tells us that the \( \mathbb{C} \)-algebra \( T_{\mathbb{C}} \) generated by all \( t_m \) is isomorphic to the \( \mathbb{C} \)-algebra generated by all Hecke operators \( T_m \). Moreover, \( \text{Pic}(X) \otimes_{\mathbb{Z}} \mathbb{C} \) and \( M_{\text{new}}(\Gamma_0(N_0)) \) are isomorphic as \( T_{\mathbb{C}} \)-modules.

According to multiplicity one theorem, which will be recalled in the Appendix 3.2, \( M_{\text{new}}(\Gamma_0(N_0)) \) is a free rank one \( T_{\mathbb{C}} \)-module. More precisely, \( M_{\text{new}}(\Gamma_0(N_0)) \) is generated by the element \( f \) whose Fourier coefficients are \( f^*(r, \lambda) = q^{-r+2} \cdot \text{Tr}(T_m) \) for all \( 0 \neq \lambda \in A, (\lambda) = (m), \deg \lambda + 2 \leq r \). Therefore \( \text{Pic}(X) \otimes_{\mathbb{Z}} \mathbb{C} \) is also a free rank one \( T_{\mathbb{C}} \)-module. This shows
\[ \dim_{\mathbb{C}} M_{\text{new}}(\Gamma_0(N_0)) = \dim_{\mathbb{C}} \left( (\text{Pic}(X) \otimes_{\mathbb{Z}} \mathbb{C}) \otimes_{T_{\mathbb{C}}} (\text{Pic}(X)^{\vee} \otimes_{\mathbb{Z}} \mathbb{C}) \right). \]
Moreover, since
\[ \sum_{i=1}^{n} < e_i, t_m e_i > = \text{Tr}(B(m)) = \text{Tr}(t_m), \]
we get \( \sum_{i=1}^{n} \Phi(e_i, e'_i) = f \), which generates \( M_{\text{new}}(\Gamma_0(N_0)) \). This also tells us that \( \sum_{i=1}^{n} e_i \otimes e'_i \) is a generator of the cyclic \( T_{\mathbb{C}} \)-module \( (\text{Pic}(X) \otimes_{\mathbb{Z}} \mathbb{C}) \otimes_{T_{\mathbb{C}}} (\text{Pic}(X)^{\vee} \otimes_{\mathbb{Z}} \mathbb{C}) \).

The above argument gives us the main result:

**Theorem 2.6.** There is a map \( \Phi : \text{Pic}(X) \times \text{Pic}(X)^{\vee} \rightarrow M_{\text{new}}(\Gamma_0(N_0)) \) satisfying that for \( r \in \mathbb{Z} \) and \( u \in k_{\infty} \)
\[ \Phi(e, e') \left( \begin{array}{cc} \pi_{\infty}^r & u \\ 0 & 1 \end{array} \right) = q^{-r+2} \left( \deg e \cdot \deg e' + \sum_{\substack{m \text{ monic} \\deg m \leq r-2}} < e, t_m e' > \sum_{(\lambda) = (m)} \psi_{\infty}(\lambda u) \right). \]
and for all monic polynomials $m$ in $A$

$$T_m \Phi(e,e') = \Phi(t_m e,e') = \Phi(e,t_m e').$$

Moreover, this map induces an isomorphism

$$(\text{Pic}(X) \otimes \mathbb{Z} \mathbb{C}) \otimes_{T_{\mathbb{C}}} (\text{Pic}(X)^{\vee} \otimes \mathbb{Z} \mathbb{C}) \cong M_{\text{new}}^{\text{new}}(\Gamma_0(N_0))$$

as $T_{\mathbb{C}}$-modules.

**Remark.** 1. When $N_0$ is a prime, $M_{\text{new}}^{\text{new}}(\Gamma_0(N_0)) = M(\Gamma_0(N_0))$ and so the theta series $\Theta_{i,j}$ gives us a construction of all automorphic forms of Drinfeld type for $\Gamma_0(N_0)$. This case was proven by Papikian [10] via a geometric approach.

2. Since the theta series $\Theta_{i,j}$ are $\mathbb{Q}$-valued, the map $\Phi$ in Theorem 2.6 in fact induces an isomorphism

$$(\text{Pic}(X) \otimes \mathbb{Z} \mathbb{Q}) \otimes_{T_{\mathbb{Q}}} (\text{Pic}(X)^{\vee} \otimes \mathbb{Z} \mathbb{Q}) \cong M_{\text{new}}^{\text{new}}(\Gamma_0(N_0), \mathbb{Q})$$

where $T_{\mathbb{Q}}$ is the $\mathbb{Q}$-algebra generated by $t_m$ for all monic $m$ in $A$ and $M_{\text{new}}^{\text{new}}(\Gamma_0(N_0), \mathbb{Q})$ is the space of $\mathbb{Q}$-valued functions in $M_{\text{new}}^{\text{new}}(\Gamma_0(N_0))$.

3. The Claim above is essentially Jacquet-Langlands correspondence over the function field $k$, which will be recalled in the Appendix §A.1.

### 2.4 Example: The function $g_A$.

Having Theorem 2.6, we exhibit automorphic forms of Drinfeld type with nice arithmetic properties.

Let $D \in A - k^2$ be a square-free element with the quadratic Legendre symbol $(\frac{D}{P}) \neq 1$ for all $P | N_0$. Let $K$ be the imaginary quadratic field $k(\sqrt{D})$ and $O_K$ be the integral closure of $A$ in $K$. Recall that in §1.1 one has a free action of $\text{Pic}(O_K)$ on the set $G_D$ of Gross points of discriminant $D$ in the definite Shimura curve $X = X_{N_0}$:

$$G_D \times \text{Pic}(O_K) \rightarrow G_D$$
$$x \cdot (x,A) \mapsto x_A.$$ 

Suppose a Gross point $x$ of discriminant $D$ in $X$ is given. For each ideal class $A$ in $\text{Pic}(O_K)$, denote $e_A$ to be the divisor class $(x_A)$ in $\text{Pic}(X)$. Define

$$g_A := \sum_{B \in \text{Pic}(O_K)} \Phi(e_B, e_{AB}).$$

We have a nice formula for the Fourier coefficients of $g_A$ in terms of Hecke actions: for monic $m \in A$ with $\text{deg} \ m + 2 \leq r$,

$$g_A^*(r,m) = q^{-r+2} \sum_{B \in \text{Pic}(O_K)} < e_B, t_m e_{AB} >,$$ 

$$g_A^*(r,0) = q^{-r+2} h_{O_K}.$$
Here \( h_{O_K} = \# \text{Pic}(O_K) \). Note that \( g_A \) is independent of the choice of the Gross point \( x \).

From now on we assume \( D \) is irreducible with \( \left( \frac{D}{P} \right) = -1 \) for all primes \( P \mid N_0 \). According to Dirichlet’s theorem there exists a monic irreducible polynomial \( Q \) prime to \( N_0 \) and \( \epsilon_0 \in \mathbb{F}_q^\times - \mathbb{F}_q^2 \) such that \( \deg N_0 Q D \) is odd and \( \epsilon_0 N_0 Q \equiv 1 \mod D \). Then there exists \( j \in D \) with \( j^2 \equiv \epsilon_0 N_0 Q \) so that \( D = K + Kj \) and \( j^{-1} \alpha j = \bar{\alpha} \) for \( \alpha \in \mathfrak{D} \).

Let \( \mathfrak{d} = (\sqrt{D}) \) be the different of \( O_K \), which is a prime ideal in \( O_K \). Since \( \epsilon_0 N_0 Q \equiv 1 \mod D \), one has \( \left( \frac{\epsilon_0 N_0 Q}{\mathfrak{d}} \right) = 1 \). From the reciprocity law we get \( \left( \frac{\mathfrak{d}}{\mathfrak{d}} \right) = 1 \) and so the prime ideal \( (Q) \) is split in \( K \). Suppose \( (Q) = q \mathfrak{q} \) and set

\[
R := \{ \alpha + \beta j : \alpha \in \mathfrak{d}^{-1}, \beta \in \mathfrak{d}^{-1} q^{-1}, \alpha - \beta \in O_\mathfrak{d} \}.
\]

Here \( O_\mathfrak{d} \) is the localization of \( O_K \) at \( \mathfrak{d} \). It is clear that \( R \) is an \( A \)-lattice in \( \mathfrak{D} \) containing 1. In fact, \( R \) is a maximal \( A \)-order and \( K \cap R = O_K \). To show \( R \) is an \( A \)-order, let \( \alpha_1 + \beta_1 j \) and \( \alpha_2 + \beta_2 j \) be two elements in \( R \). Then

\[
(\alpha_1 + \beta_1 j)(\alpha_2 + \beta_2 j) = (\alpha_1 \alpha_2 + \beta_1 \beta_2 \epsilon_0 N_0 Q) + (\alpha_1 \beta_2 + \beta_1 \alpha_2)j.
\]

For \( i = 1, 2 \), write \( \beta_i \) as \( \alpha_i + \delta_i \) with \( \delta_i \in O_\mathfrak{d} \). Then

\[
\alpha_1 \alpha_2 + \beta_1 \beta_2 \epsilon_0 N_0 Q = \alpha_1 (\alpha_2 + \bar{\alpha}_2) + (\delta_1 \bar{\beta}_2 + \beta_1 \bar{\delta}_2 + \delta_1 \bar{\beta}_2) \epsilon_0 N_0 Q.
\]

Since \( \alpha_2 \in \mathfrak{d}^{-1} = A + \sqrt{D}^{-1} A \), one has

\[
\alpha_2 + \bar{\alpha}_2 \in A.
\]

Hence

\[
\alpha_1 \alpha_2 + \beta_1 \bar{\beta}_2 \epsilon_0 N_0 Q \in \mathfrak{d}^{-2} \cap \sqrt{D}^{-1} O_\mathfrak{d} = \mathfrak{d}^{-1}.
\]

Similarly,

\[
\alpha_1 \beta_2 + \beta_1 \alpha_2 \in \mathfrak{d}^{-2} q^{-1} \cap \sqrt{D}^{-1} O_\mathfrak{d} = \mathfrak{d}^{-1} q^{-1}.
\]

From the condition that \( \epsilon_0 N_0 Q \equiv 1 \mod D \), one can check that

\[
\alpha_1 \alpha_2 + \beta_1 \bar{\beta}_2 \epsilon_0 N_0 Q - (\alpha_1 \beta_2 + \beta_1 \alpha_2) \in O_\mathfrak{d}.
\]

Therefore \( R \) is an \( A \)-order. The discriminant of \( R \) is \( (N_0)^2 \), which can be checked locally. This implies that \( R \) is maximal.

Let \( x \) be the Gross point in the definite Shimura curve \( X = X_{N_0} \) which corresponds to the trivial ideal \( R \) and the embedding \( K \hookrightarrow \mathfrak{D} \). Then \( x \) is of discriminant \( D \). Using this particular Gross point we can get an explicit formula for the Fourier coefficients of \( g_A \).

Documenta Mathematica 16 (2011) 723–765
We can express the Fourier coefficients of $g$ for ideals $(N)$. Assume Proposition 2.7. Let $\alpha$, $\beta$, and $\delta_2$ be primes respectively. Therefore

\[ \bar{\alpha} \bar{\beta} = \{ \alpha + \beta j : \alpha \in \mathfrak{a}^{-1} \mathfrak{a}, \beta \in \mathfrak{b}^{-1} \mathfrak{b} \mathfrak{a}, \alpha \in (1) \} \].

We can express the Fourier coefficients of $g_A$ in terms of sums of the counting numbers $r_A((\delta)): = \{ a \in A : a \text{ integral with } \text{Nr}(a) = (\delta) \}$, for ideals $(\delta)$ of $A$, by the following proposition:

**Proposition 2.7.** Suppose $D \in A - k_\infty^2$ is irreducible with $(\mathfrak{B}) = -1$ for all primes $P | N_0$. Then for any monic polynomial $m$ in $A$,

\[ \sum_{\delta \in \text{Pic}(O_K) \setminus \{0\}} r_A((\delta)) = \frac{1}{2(q-1)} \left[ 2r_A((mD)(q-1))b_{O_K} + \sum_{\deg(\mu N_0 - mD) \leq \deg(mD)} r_A((\mu N_0 - mD))(t(\mu, D) + 1)(1 - \delta_{\mu N_0 (\mu N_0 - mD)})(\sum_{c \in \mu} \left( \frac{D}{c} \right) ) \right] . \]

Here $t(\mu, D) = 1$ if $D$ divides $\mu$ and $0$ otherwise, and $\delta_2$ is the norm residue symbol of $\alpha$ for $z \in k_\infty^2 : \delta_2 = 1$ if $z \in \text{Nr}(K_\infty^2)$ and $-1$ otherwise.

**Proof.** Let $\mathfrak{a} \in A$ which is a proper ideal of $O_K$ and prime to $N_0 \mathfrak{a}$. Fix a generator $\lambda_0$ of $\text{Nr}(\mathfrak{a}) = \mathfrak{a} \mathfrak{a}$. Given $\mathfrak{B} \in \text{Pic}(O_K)$. Let $b \in \mathfrak{B}$. For $b = \alpha + \beta j \in \mathfrak{b}^{-1} \mathfrak{b} \mathfrak{a}$, i.e. $\alpha \in \mathfrak{a}^{-1} \mathfrak{a}$, $\beta \in \mathfrak{b}^{-1} \mathfrak{b}^{-1} \mathfrak{b} \mathfrak{a}$, $\alpha \in (1) \text{ord}_b b \beta \in O_\mathfrak{B}$, define:

1. $e := (\beta) \mathfrak{b}^{-1} \mathfrak{b} \mathfrak{a} \in [\mathfrak{B}]^2 A$,
2. $\nu := - \text{Nr}(\alpha) D \lambda_0^{-1} \in A$,
3. $\mu := - e_0 \text{Nr}(\beta) D Q \lambda_0^{-1} \in A$.

Here $[\mathfrak{B}] \in \text{Pic}(O_K)$ is the ideal class containing $\mathfrak{B}$. Then $e$ is integral and

\[ \text{Nr}(\alpha + \beta j) = \text{Nr}(\alpha) - e_0 N_0 Q \text{Nr}(\beta) = (-\nu + N_0 \mu) D^{-1} \lambda_0 . \]

Thus $(\text{Nr}(\alpha + \beta j)) = (m \lambda_0)$ if and only if $\nu = N_0 \mu - emD$ for a uniquely determined $e \in \mathfrak{B}^\times$.
Since $\beta = 0$ if and only if $b = \alpha \in A$, one has
\[
\#\{b \in b^{-1}Rb\alpha : \text{Nr}(b) = (m\lambda_0)\}
= \#\{b = \alpha + \beta j \in b^{-1}Rb\alpha : \beta \neq 0, \text{Nr}(b) = (m\lambda_0)\}
+ \#\{\alpha \in A : \text{Nr}(\alpha) = (m\lambda_0)\}.
\]

It can be shown that $\#\{\alpha \in A : \text{Nr}(\alpha) = (m\lambda_0)\} = (q - 1)r_d((mD))$. Note that $\beta \neq 0$ if and only if $\mu \neq 0$. In this case, $\beta$ is uniquely determined by the integral ideal $c$ up to multiplying elements in $O_K^\times$.

Conversely, given $0 \neq \mu \in A$ and $\epsilon \in \mathbb{F}_q^\times$ and set $\nu = N_0\mu - \epsilon mD$. The number of elements $\alpha \in \mathfrak{d}^{-1}a$ with $\text{Nr}(\alpha) = -\nu D^{-1}\lambda_0$ is $r_{a,\lambda_0}(N_0\mu - \epsilon mD)$. Here
\[
r_{a,\lambda_0}(\lambda) := \#\{\alpha \in A : \text{Nr}(\alpha) = \lambda\lambda_0\} \text{ for } \lambda \in A.
\]

In the case of $r_{a,\lambda_0}(N_0\mu - \epsilon mD) \neq 0$, choose an element $\alpha \in \mathfrak{d}^{-1}a$ with $\text{Nr}(\alpha) = -\nu D^{-1}\lambda_0$. Let $c$ be an integral ideal which lies in a class differing from the ideal class $A[\mathfrak{q}]$ by a square $[b]^2$ in the class group $\text{Pic}(O_K)$ and with ideal norm $(\mu)$. Then
\[
c = b^{-1}b\alpha q^{-1}\mathfrak{d}^{-1} = (\beta)
\]
for some $\beta \in K^\times$. Suppose we can find $\beta$ so that $\mu = -\epsilon_0\text{Nr}(\beta)D\lambda_0^{-1} \in A$. Since $\epsilon_0N_0\mathfrak{q} \equiv 1 \mod D$, the equality $\epsilon m\lambda_0 = \text{Nr}(\alpha) - \epsilon_0N_0\mathfrak{q} \text{Nr}(\beta) \in A$ implies $\alpha \pm \beta \in O_\mathfrak{d}$.

Choose $\ell \in \{0, 1\}$ and replace $b$ by $bd^\ell$ so that $\alpha - (-1)^{ord_b(b)}\beta \in O_\mathfrak{d}$. Therefore $b = \alpha + \beta j \in b^{-1}Rb\alpha$ with $\text{Nr}(b) = \epsilon m\lambda_0$. Note that if $\beta$ is not in $O_\mathfrak{d}$ (i.e. $D \nmid \mu$), then $\ell$ is uniquely determined. If $\beta \in O_\mathfrak{d}$ (i.e. $D \mid \mu$), then we have two choices $\pm \beta$. The existence of $\beta$ is equivalent to that $-\epsilon_0D\mu \mathfrak{q}^{-1}\lambda_0$ is in $\text{Nr}(K^\times)$. Since $\text{Nr}(\sqrt{D}) = -D$ and $(\epsilon_0^{-1}\mu \mathfrak{q}^{-1}\lambda_0) = \text{Nr}(\mathfrak{q}^{-1}a)$, we have $\epsilon_0^{-1}\mu \mathfrak{q}^{-1}\lambda_0 \in \text{Nr}(K^\times)$ if and only if $\delta_0^{-1}\mu \mathfrak{q}^{-1}\lambda_0 = 1$. Therefore combining the above arguments we have
\[
\sum_{b \in \text{Pic}(O_K)} \#\{b = \alpha + \beta j \in b^{-1}Rb\alpha : \beta \neq 0, \text{Nr}(b) = (m\lambda_0)\}
= \sum_{\theta \neq \mu \in A} \sum_{\epsilon \in \mathbb{F}_q^\times} r_{a,\lambda_0}(N_0\mu - \epsilon mD) \cdot (t(\mu, D) + 1) \cdot \mathcal{R}_{\{A[\mathfrak{q}]\}}((\mu)) \cdot \frac{1 + \delta_0^{-1}\mu \mathfrak{q}^{-1}\lambda_0}{2}.
\]

Here $\mathcal{R}_{\{A[\mathfrak{q}]\}}((\mu))$ is the number of integral ideals $\mathfrak{c}$, which lie in a class differing from the class $A[\mathfrak{q}]$ by a square in the class group $\text{Pic}(O_K)$ and with ideal norm $(\mu)$. Following the proof of Lemma 3.4.9 in [12] one has

**Lemma 2.8.** For $0 \neq \mu \in A$,
\[
\mathcal{R}_{\{A[\mathfrak{q}]\}}((\mu)) \cdot \frac{1 + \delta_0^{-1}\mu \mathfrak{q}^{-1}\lambda_0}{2} = \frac{1}{q - 1} \sum_{c|\mu} \left(\frac{D}{c}\right) \cdot \frac{1 + \delta_0^{-1}\mu \mathfrak{q}^{-1}\lambda_0}{2}.
\]
Since $\delta_{\eta, Q^\lambda_0} = 1$ if and only if $\delta_{N_0 \mu \lambda_0} = -1$, with Lemma 2.6 we have

$$
\sum_{B \in \text{Pic}(O_K)} \# \{ b = a + \beta j \in b^{-1} \mathfrak{R}b a : \beta \neq 0, \text{Nr}(b) = (m_0) \} = \sum_{0 \neq \mu \in A, \epsilon \in \mathbb{F}_q^\times} \{ \begin{array}{c} \epsilon \in \mathbb{F}_q \\ \mu = \frac{N_0 \mu}{ \epsilon^2} \end{array} \} r_{\epsilon, \lambda_0}((N_0 \mu - \epsilon m D)(t(\mu, D) + 1) \cdot \frac{1 - \delta_{N_0 \mu \lambda_0}}{2} q^{-1} \sum_{c \mid \mu} \frac{D}{c} \\
= \sum_{\mu \in A, \epsilon \in \mathbb{F}_q^\times} \{ \begin{array}{c} \epsilon \in \mathbb{F}_q \\ \mu = \frac{N_0 \mu}{ \epsilon^2} \end{array} \} r_{\mu}(\mu N_0 - m D)(t(\mu, D) + 1) \cdot \frac{1 - \delta_{N_0 \mu \lambda_0}}{2} q^{-1} \sum_{c \mid \mu} \frac{D}{c}.
$$

Therefore

$$
\sum_{B \in \text{Pic}(O_K)} < e_B, t_m e_{A_B} > = \frac{1}{2(q^{-1})} \left[ 2 r_{A}((m D))(q - 1) h_{O_K} + \sum_{\mu \in A, \epsilon \in \mathbb{F}_q^\times} \{ \begin{array}{c} \epsilon \in \mathbb{F}_q \\ \mu = \frac{N_0 \mu}{ \epsilon^2} \end{array} \} r_{\mu}(\mu N_0 - m D)(t(\mu, D) + 1)(1 - \delta_{N_0 \mu \lambda_0}) \sum_{c \mid \mu} \frac{D}{c} \right].
$$

3 Special values of $L$-series

3.1 Rankin product

To an automorphic cusp form $f$ of Drinfeld type for $\Gamma_0(N)$ one can attach an $L$-series $L(f, s)$: let $m$ be an effective divisor of $k$, which can be written as $\text{div}(\lambda_0) + (r - \deg \lambda) \infty$ for a nonzero polynomial $\lambda (= \lambda_0)$ in $A$, with

$$
\text{div}(\lambda_0) := \sum_{\text{finite prime } P} \text{ord}_P(\lambda) P.
$$

Denote

$$
f^*(m) := \int_{A(k_\infty)} f \left( \begin{array}{cc} \pi^{r+2} & u \\ 0 & 1 \end{array} \right) \psi(\lambda u) du = f^*(r + 2, \lambda).
$$

The $L$-series $L(f, s)$ attached to $f$ is

$$
L(f, s) := \sum_{m \geq 0} f^*(m) q^{-\deg(m)s}, \quad \text{Re } s > 1.
$$

Let $D \in A - A_\infty$ be a square-free element. Consider the imaginary field $K = k(\sqrt{D})$. Let $O_K$ be the integral closure of $A$ in $K$ and $\text{Pic}(O_K)$ be the ideal class group of $O_K$. Given an ideal class $A \in \text{Pic}(O_K)$ and a polynomial $\lambda$ in
The number of integral ideals $\mathfrak{a}$ in the class $\mathcal{A}$ with $N_{K/k}(\mathfrak{a}) = (\lambda)$ leads to the partial zeta function attached to $\mathcal{A}$:

$$\zeta_\mathcal{A}(s) := \sum_{m \geq 0} r_\mathcal{A}(m) q^{-\deg(m)s}, \quad \Re s > 1.$$  

Here for each effective divisor $m = \text{div}(\lambda)_0 + (r - \deg \lambda)\infty$,

$$r_\mathcal{A}(m) := \# \{ \mathfrak{a} \in \mathcal{A} : \mathfrak{a} \text{ integral with } N_{K/k}(\mathfrak{a}) = (\lambda) \}.$$  

Let $f$ be an automorphic cusp form of Drinfeld type for $\Gamma_0(N)$. For each ideal class $\mathcal{A} \in \text{Pic}(O_K)$, we are interested in the Rankin product:

$$L(f, \mathcal{A}, s) := \sum_{m \geq 0} f^*(m) r_\mathcal{A}(m) q^{-\deg(m)s}, \quad \Re(s) > 1.$$  

To study the analytic continuation and the functional equation of $L(f, \mathcal{A}, s)$, consider the function $\Lambda(f, \mathcal{A}, s)$ which is defined by:

$$\Lambda(f, \mathcal{A}, s) := \begin{cases} L^{(N,D)}(2s + 1)L(f, \mathcal{A}, s) & \text{when } \deg D \text{ is odd}, \\ \frac{1}{1 + q^{-s-1}}L^{(N,D)}(2s + 1)L(f, \mathcal{A}, s) & \text{when } \deg D \text{ is even}. \end{cases}$$  

Here $L^{(N,D)}(s)$ is the following $L$-series indexed by effective divisors supported outside $\infty$

$$L^{(N,D)}(s) := \frac{1}{q - 1} \sum_{d \in A, (d,N)=1} \left( \frac{D}{d} \right) q^{-s \deg d}, \quad \Re(s) > 1,$$  

where $\left( \frac{D}{d} \right)$ denotes the Legendre symbol for the polynomial ring $A$. Note that

$$L^{(N,D)}(s) = L_D(s) \cdot \prod_{\text{prime ideals } P \mid N} \left( 1 - \left( \frac{D}{P} \right) q^{-s \deg P} \right)^{-1}$$  

where $L_D(s)$ is the Dirichlet $L$-series:

$$L_D(s) := \frac{1}{q - 1} \sum_{d \in A, d \neq 0} \left( \frac{D}{d} \right) q^{-s \deg d}, \quad \Re(s) > 1.$$  

It is known that $L_D(s)$ can be extended to a polynomial in $q^{-s}$ with the functional equation (cf. [1]):

$$L_D(2s + 1) = q^s(-2 \deg D + 2)q^{-\frac{1}{2} \deg D + \frac{1}{2}}L_D(-2s)$$  

if $\deg D$ is odd, and

$$L_D(-2s + 1) = \frac{1 + q^{1-2s}}{1 + q^{2s}} q^{\deg D(2s-\frac{1}{2})}L_D(2s)$$  

Documenta Mathematica 16 (2011) 723–765
if $\deg D$ is even.

When $f$ is a new form and $D$ is irreducible, Rück and Tipp (12) prove the following functional equation of $\Lambda(f, A, s)$:

$$\Lambda(f, A, s) = -\left(\frac{D}{N}\right) q^{(5-2\deg D-2\deg N)s} \Lambda(f, A, -s)$$

when $\deg D$ is odd, and

$$\Lambda(f, A, s) = -\left(\frac{D}{N}\right) q^{(6-2\deg D-2\deg N)s} \Lambda(f, A, -s)$$

when $\deg D$ is even.

3.2 Central critical values of $\Lambda(f, A, s)$

We are interested in the special value of $\Lambda(f, A, s)$ at $s = 0$. Note that if $\left(\frac{D}{N}\right) = 1$, then $\Lambda(f, A, s)$ has a zero at $s = 0$. We focus here on the special case when $\left(\frac{D}{N}\right) = -1$ for all primes $P \mid N_0$. Adapting Rankin’s method (cf. [12]), we can establish the following theorem.

**Theorem 3.1.** Let $f$ be a Drinfeld type new form for $\Gamma_0(N_0)$ and let $D$ be an irreducible polynomial in $A - k_{\infty}^2$ with $\left(\frac{D}{P}\right) = -1$ for all primes $P \mid N_0$. One has

$$\Lambda(f, A, 0) = \begin{cases} (f, g_A) & \text{when } \deg D \text{ is odd}, \\ \frac{1}{2q^{\deg D}} (f, g_A) & \text{when } \deg D \text{ is even}. \end{cases}$$

Here $(\cdot, \cdot)$ is the Petersson inner product and $g_A$ is the Drinfeld type automorphic form for $\Gamma_0(N_0)$ canonically attached to $A$ in [2,4].

3.2.1 Review of Rankin’s method

Given $A \in \text{Pic}(O_K)$. Choose $a_0 \in A^{-1}$ and $\lambda_0 \in k$ such that $N_{K/k}(a_0) = (\lambda_0)$.

Recall the counting number

$$r_{a_0, \lambda_0}(\lambda) = \# \{ \mu \in a_0 : N_{K/k}(\mu) = \lambda_0 \lambda \}.$$

Note that $r_{a_0, \lambda_0}(\lambda) = r_{a_0^{-1}\lambda_0^{-1}}(\lambda)$, and for effective divisor $m = \text{div}(\lambda)_0 + (\deg m - \deg \lambda)_{\infty}$ we have

$$r_A(m) = \frac{1}{q-1} \sum_{\epsilon \in \mathbb{F}_q^\times} r_{a_0, \lambda_0}(\epsilon \lambda).$$

We consider the following theta series $\theta_{a_0, \lambda_0}$ (introduced in [11]) defined on $k_\infty^\times \times k_\infty$:

$$\theta_{a_0, \lambda_0}(\pi_\infty^r, u) := \sum_{\deg \lambda + 2 \leq r} r_{a_0, \lambda_0}(\lambda) \psi_{\infty}(\lambda u).$$
It satisfies the following transformation law:
\[ \theta_{a_0,\lambda_0} \left( \frac{\pi^r_{\infty}}{(cu+d)^2} \frac{au+b}{cu+d} \right) = \delta_{cu+d} \left( \frac{d}{D} \right) q^{-\nu_{\infty}(cu+d)} \theta_{a_0,\lambda_0}(\pi^r_{\infty}, u) \]

for all \( \gamma = \left( \begin{array}{cc} a & b \\ c & d \end{array} \right) \in \Gamma_0^{(1)}(N) := \Gamma_0(N) \cap \text{SL}_2(A) \) with \( \nu_{\infty}(c\pi^r_{\infty}) > \nu_{\infty}(cu+d) \).

Here \( \delta \) is the local norm symbol at \( \infty \), i.e. \( \delta_z = 1 \) if \( z \in k \times \infty \) is a norm of an element in \( K_{\infty} = k_{\infty}(\sqrt{D}) \) and \( -1 \) otherwise.

Viewing \( \theta_{a_0,\lambda_0} \) as a function on \( \mathbb{H}_{\infty} := \left( \frac{1}{A} \right) \setminus \left( \frac{k^{\times}_{\infty}}{k_{\infty}} \right) \left/ \left( \frac{O^{\times}_{\infty}}{O_{\infty}} \right) \right. \),

one can write
\[ L(f, A, s) = \frac{q}{q-1} \sum_{r=2}^{\infty} \left[ \sum_{u \in \pi_{\infty} O_{\infty}/\pi^r_{\infty} O_{\infty}} f \cdot \theta_{a_0,\lambda_0}(\pi^r_{\infty}, u) q^{-r(s+1)+2s} \right] \]

\[ = \frac{q}{q-1} \int_{\mathbb{H}_{\infty}} f(h) \theta_{a_0,\lambda_0}(h) q^{-r(s+1)+2s} dh. \]

For every monic polynomial \( M \) in \( A \), the canonical map \( \mathbb{H}_{\infty} \to G(M) := \Gamma_0^{(1)}(M) \setminus \text{GL}_2(k_{\infty})/\Gamma_{\infty} k^{\times}_{\infty} \)

is surjective. Following [12], we consider the “Eisenstein series”
\[ E_s \left( \frac{\pi^r_{\infty}}{0 1} \right) := \sum_{\substack{c,d \in A, cd \equiv 0 \mod D \\ \nu_{\infty}(c\pi^r_{\infty}) > \nu_{\infty}(cu+d)}} \left( \frac{d}{D} \right) \delta_{cu+d} q^{-\nu_{\infty}(cu+d)(2s+1)} \]

and let \( H_s \left( \frac{\pi^r_{\infty}}{0 1} \right) := \)
\[ \begin{cases} \rho^{-r(s+1)+2s} E_s \left( \frac{N\pi^r_{\infty}}{0 1} \right) & \text{when } \deg D \text{ is odd,} \\
\left( \frac{-1}{2} \right)^{r-\deg \lambda_0+1} \cdot q^{-r(s+1)+2s} E_s \left( \frac{N\pi^r_{\infty}}{0 1} \right) & \text{when } \deg D \text{ is even.} \end{cases} \]

Then \( \theta_{a_0,\lambda_0} H_s \) can be viewed as a function on \( G(ND) \). By [12] Proposition 2.2.2 and Proposition 2.3.2
\[ \Lambda(f, A, s) = \frac{q}{2(q-1)} \int_{G(ND)} f \cdot \theta_{a_0,\lambda_0} H_s. \]
Given $M \in A$. Let $\mathcal{F}(M)$ be the space of functions on $G(M)$. The trace map from $\mathcal{F}(ND)$ to $\mathcal{F}(N)$ is given by

$$f \mapsto \text{Tr}^{ND}_{N} f(g) := \sum_{\gamma \in \Gamma_0^{(1)}(ND) \setminus \Gamma_0^{(1)}(N)} f(\gamma g).$$

Set $\Phi_s := \text{Tr}^{ND}_{N}(\theta_{\alpha_0, \lambda_0} H_s)$. Then

$$\Lambda(f, A, s) = q^2(q-1) \int_{G(N)} f \cdot \Phi_s.$$ 

From the harmonicity of $f$ one has

$$\Lambda(f, A, s) = q^4(q-1) \int_{G(N)} f \cdot F_s$$

where for $g \in \text{GL}_2(k_\infty)$,

$$F_s(g) := \frac{q}{q+1} (\Phi_s(g) - \Phi_s(g)) - \frac{1}{q+1} \sum_{\beta \in \text{GL}_2(O_\infty) / \Gamma_\infty} (\Phi_s(g\beta) - \Phi_s(g\beta)).$$

Note that $F_s$ depends on the choice of $\alpha_0$ and $\lambda_0$.

### 3.2.2 Proof of Theorem 3.1

Let $\Psi$ be the average map from functions $F$ on $G(N)$ to functions on $G_0(N)$:

$$\Psi(F)(g) := \frac{1}{q-1} \sum_{\epsilon \in \mathbb{F}_q^*} F(\begin{pmatrix} \epsilon & 0 \\ 0 & 1 \end{pmatrix}) g).$$

Define

$$\Psi_A := \Psi(F_0).$$

Note that $\Psi_A$ now depends only on $A$.

Taking the formulas in Proposition 2.7.2 and Proposition 2.7.5 in [12] and specializing at $s = 0$ we deduce that for any $\lambda \in A$ with $\deg \lambda + 2 \leq r$

$$\Psi^*_A(r, \lambda) = \frac{3 - (-1)^{\deg D}}{4} \cdot q^{-r+1-\lfloor \frac{\deg D}{2} \rfloor} \cdot \left[ 2r_A((\lambda D))(q-1)L_D(0) + \sum_{\mu \in A, \mu \neq 0, \deg(\mu N) \leq \deg(\lambda D)} r_A((\mu N - \lambda D))(t(\mu, D) + 1)(1 - \delta_{\mu N(\mu N - \lambda D)}) \sum_{c | \mu} \left( \frac{D}{c} \right) \right].$$

Moreover, one has
PROPOSITION 3.2.

\[ \Lambda(f, A, 0) = \frac{q}{2(q-1)} \int_{G_0(N)} f \cdot \Psi_A. \]

Let \( N = N_0 \). Note that \( L_D(0) = h_{O_K} \). Comparing the Fourier coefficients of \( \Psi_A \) with that of \( g_A \) we obtain

\[ \Psi_A = g_A \cdot \begin{cases} q^{-\frac{1}{2} \deg D + \frac{1}{2}} \cdot \frac{q-2}{2} \cdot (q-1) \cdot 2 & \text{when } \deg D \text{ is odd}, \\ q^{-1} \cdot q^{-\frac{1}{2} \deg D} \cdot (q-1) & \text{when } \deg D \text{ is even}. \end{cases} \]

Therefore Theorem 3.1 holds.

3.3 A function field analogue of Gross formula

Now given a character \( \chi : \text{Pic}(O_K) \to \mathbb{C}^\times \), define

\[ \Lambda(f, \chi, s) := \sum_{A \in \text{Pic}(O_K)} \chi(A) \Lambda(f, A, s). \]

When \( \chi \) is the trivial character and \( f \) is a newform which is “normalized” so that the Fourier coefficient \( f^*(0) = 1 \), one has

\[ \Lambda(f, \chi, s) = L(f, s)L(f \otimes \varepsilon_D, s) \]

where \( \varepsilon_D \) is the following quadratic character on divisors of \( k \):

\[ \varepsilon_D(P) = \left( \frac{D}{P} \right) \text{ and } \varepsilon_D(\infty) = \begin{cases} -1 & \text{if } \deg D \text{ is even,} \\ 0 & \text{if } \deg D \text{ is odd}; \end{cases} \]

and \( L(f \otimes \varepsilon_D, s) \) is the twisted \( L \)-series of \( f \) by \( \varepsilon_D \):

\[ L(f \otimes \varepsilon_D, s) := \sum_{m \geq 0} f^*(m) \varepsilon_D(m) q^{-\deg ms}. \]

From the definition of \( \Lambda(f, \chi, s) \) and Theorem 3.1 one has

\[ \Lambda(f, \chi, 0) = \left( \sum_{A \in \text{Pic}(O_K)} \chi(A)(f, g_A) \right) \cdot \begin{cases} \frac{1}{q^{\frac{1}{2} \deg D} + 1} & \text{if } \deg D \text{ is odd,} \\ \frac{1}{2q^{\frac{1}{2} \deg D}} & \text{if } \deg D \text{ is even.} \end{cases} \]

Note that

\[ \sum_{A \in \text{Pic}(O_K)} \chi(A)^{-1} g_A = \sum_{A \in \text{Pic}(O_K)} \left( \sum_{B \in \text{Pic}(O_K)} \chi(A)^{-1} \Phi(e_B, e_{AB}) \right) = \Phi(e_{O_K}, e_{O_K}) \]
where $\Phi$ is the map in Theorem 2.6 and

$$e_\chi = \sum_{A \in \text{Pic}(O_K)} \chi(A) e_A.$$ 

Suppose $f$ is a normalized newform. Then from Theorem 2.6 $f$ corresponds to a particular element $e_f \in \text{Pic}(X) \otimes \mathbb{Z} \mathbb{R}$ such that

$$f = \Phi(e_f, e_f).$$

Let $e_{f,X}$ be the projection of $e_\chi$ to the $e_f$-isotypical component in the space $\text{Pic}(X) \otimes \mathbb{Z} \mathbb{C}$ with respect to the Gross height pairing. Then the $f$-eigencomponent of $\Phi(e_\chi, e_\chi)$ is equal to

$$\Phi(e_{f,X}, e_\chi) = \Phi(e_\chi, e_{f,X}) = <e_{f,X}, e_f>.$$ 

The last equality holds as $f$ is normalized (i.e. $f^*(0) = 1$) and the Fourier coefficient $\Phi(e_{f,X}, e_{f,X}) = <e_{f,X}, e_{f,X}>$. Therefore we obtain

**Theorem 3.3.** Let $f$ be an automorphic cusp form of Drinfeld type for $\Gamma_0(N_0)$ which is also a normalized newform. Then

$$\Lambda(f, \chi, 0) = \begin{cases} 
\frac{(f, f)}{q^{\frac{1}{2} \deg D + 1}} <e_{f,X}, e_f> & \text{if } \deg D \text{ is odd,} \\
\frac{(f, f)}{2q^{\frac{1}{2} \deg D}} <e_{f,X}, e_f> & \text{if } \deg D \text{ is even.}
\end{cases}$$

**Remark.** 1. If $\chi$ is non-trivial, then $\deg e_\chi = 0$ and so $\Phi(e_\chi, e_\chi)$ is a cusp form. 
2. When $\chi$ is trivial, then

$$\sum_{\text{monic } m | N_0} t_m e_\chi = 2e_D$$

where $e_D$ is the divisor class introduced in Proposition 1.7.
3. The special case when $N_0$ is a prime and $\deg D$ is odd, the above formula coincides with the result in [10] §4 (be aware of the different choices of measures for the Petersson inner product).
4. When irreducible $D \in A - k^*_\infty$ satisfies $\left(D_{N_0}\right) = 1$, the derivative of $\Lambda(f, \chi, s)$ at $s = 0$ is given by Néron-Tate height of Heegner points on the Drinfeld modular curve $X_0(N_0)$, and an analogue of Gross-Zagier formula has been proved by Rück and Tipp in the case $D$ is irreducible (cf. [12]).

**3.4 Example and application to elliptic curves**

Let $E$ be a non-iso-trivial elliptic curve over $k$ (i.e. $E$ is not defined over the constant field $\mathbb{F}_q$). From the work of Weil, Jacquet-Langlands, and Deligne, one knows that there exists an automorphic cusp form $f_E$ such that

$$L(E/k, s + 1) = L(f_E, s).$$
Consider the Hasse-Weil $L$-series $L(E/K, s)$ of $E$ over the imaginary quadratic field $K = k(\sqrt{D})$ where $D \in \mathcal{A}$ with $(\frac{D}{t^2}) = -1$ for all primes $P \mid N_0$. One has

$$L(E/K, s + 1) = L(f_E, s)L(f_E \otimes \varepsilon_D, s)$$

where $L(f_E \otimes \varepsilon_D, s)$ is the twisted $L$-series of $f_E$ by the quadratic character $\varepsilon_D$. Since

$$L(f_E, s)L(f_E \otimes \varepsilon_D, s) = \Lambda(f_E, 1_D, s)$$

where $1_D$ is the trivial character on $\text{Pic}(\mathcal{O}_K)$, from Theorem 3.3 we obtain a formula for the special value of $L(E/K, s)$ at $s = 1$ when $D$ is irreducible.

Now, let $k = \mathbb{F}_q(t)$ (i.e. $q = 3$). Let $E$ be the following elliptic curve over $k$:

$$E : y^2 = x^3 + (t^2 + 1)x^2 + t^2x = x(x + 1)(x + t^2).$$

The conductor of $E$ is $(t)(t + 1)(t - 1)\infty$. More precisely, $E$ has split multiplicative reduction at $(t)$ and $\infty$, and has non-split multiplicative reduction at $(t + 1)$ and $(t - 1)$. Let $N_0 = t(t + 1)(t - 1) = t^3 - t$. Let $f_E$ be the normalized Drinfeld type cusp form for $\Gamma(N_0)$ associated to $E$. Since the $L$-series $L(E/k, s)$ of $E$ over $k$ is a polynomial in $q^{-s}$ of degree $(\deg N_0 + 1) - 4$ with constant term 1, this implies that $L(E/k, s) = L(f_E, s - 1) = 1$.

Let $D = t^3 - t - 1$ and $K = k(\sqrt{D})$. Then

$$\left(\frac{D}{t}\right) = \left(\frac{D}{t + 1}\right) = \left(\frac{D}{t - 1}\right) = -1.$$

The twist $E_D$ of $E$ by $D$ is the following elliptic curve over $k$:

$$y^2 = x^3 + (t^2 + 1)Dx^2 + t^2D^2x.$$

The conductor of $E_D$ is $(D)^2(t)(t + 1)(t - 1)\infty^2$, and the $L$-series $L(E_D/k, s)$ is

$$1 + q^{-s} + 4q^{-2s} + 108q^{-5s} + 243q^{-6s} + 2187q^{-7s}.$$

Since $L(E/K, s) = L(E/k, s) \cdot L(E_D/k, s)$, we have

$$L(E/K, s) = 1 + q^{-s} + 4q^{-2s} + 108q^{-5s} + 243q^{-6s} + 2187q^{-7s}$$

and $L(E/K, 1) = \frac{42}{9}$.

On the other hand, from a formula of Gekeler (cf. [13] Theorem 1.1) we immediately get

$$(f_E, f_E) = 32.$$
We point out that our choice of the measure is twice of the one in [13]. Such computation can be also checked via the algorithm in [15].

The only remaining term is the Gross height of the corresponding point $e_{f_E}$ in $\text{Pic}(X_{N_0}) \otimes \mathbb{Z}$. Let $\mathcal{D}$ be the definite quaternion algebra over $k$ ramified at $(t)$, $(t+1)$, and $(t-1)$. Then

$$\mathcal{D} = k + k\alpha + k\beta + k\alpha\beta$$

where $\alpha^2 = -1$, $\beta^2 = N_0 = t^3 - t$, and $\beta\alpha = -\alpha\beta$. Let $R = A + A\alpha + A\beta + A\alpha\beta$, which is a maximal order in $\mathcal{D}$. The cardinality of $R^\times$ is 8, and the class number (of left ideal classes of $R$) is 4. We choose the following representatives of left ideal classes of $R$:

- $I_1 = R$,
- $I_2 = At + At\alpha + A\beta + A\alpha\beta$,
- $I_3 = A(t+1) + A(t+1)\alpha + A\beta + A\alpha\beta$,
- $I_4 = A(t-1) + A(t-1)\alpha + A\beta + A\alpha\beta$.

Note that these ideals are in fact two-sided, and the norm form on each of them can be easily written down. We calculate the following Brandt matrices:

$$B(t) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$B(t+1) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$B(t-1) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Since we have $T_tf_E = f_E$, $T_{t+1}f_E = -f_E$, $T_{t-1}f_E = -f_E$, and the Gross height $<e_{f_E}, e_{f_E}> = f_E(0) = 1$, the corresponding point $e_{f_E}$ in $\text{Pic}(X_{N_0}) \otimes \mathbb{Z}$ can only be

$$\pm [1/4, 1/4, -1/4, -1/4].$$

The class number of $O_K (= A[\sqrt{D}])$ is 1. Choose the Gross point $x$ in the first component of $X_{N_0}$ corresponding to the embedding $K \hookrightarrow \mathcal{D}$ which maps $\sqrt{D}$ to $\alpha + \beta$. Then $e_x = [1, 0, 0, 0]$ in $\text{Pic}(X_{N_0}) \otimes \mathbb{Z}$. Therefore

$$<e_{f_E;1_D}, e_{f_E;1_D}> = <e_{f_E}, e_x>^2 = (4 \cdot 1/4)^2 = 1$$

and

$$\frac{(f_E, f_E)}{q_{1_D}^{\deg \mathcal{D}+1}} <e_{f_E;1_D}, e_{f_E;1_D}> = \frac{32}{9} = L(E/K, 1).$$

Appendix
A Jacquet-Langlands correspondence and multiplicity one theorem

Let \( \varpi \) be a Hecke character on \( k^\times \backslash \mathbb{A}_k^\times \). Let \( \mathcal{D} \) be a quaternion algebra over \( k \) and set \( \mathcal{D}_{\mathbb{A}_k} := \mathcal{D} \otimes_k \mathbb{A}_k \). We embed \( \mathbb{A}_k \) into \( \mathcal{D}_{\mathbb{A}_k} \) by \( a \mapsto 1 \otimes a \). A \( \mathbb{C} \)-valued function \( f \) on \( \mathcal{D}_+ \backslash \mathcal{D}_{\mathbb{A}_k}^+ \) is called an automorphic form on \( \mathcal{D}_+^\times \mathbb{A}_k^\times \) (for \( \mathcal{K} \)) with central character \( \varpi \) if \( f \) is a function on the double coset space \( \mathcal{D}_+ \backslash \mathcal{D}_{\mathbb{A}_k}^+ / \mathcal{K} \) for an open compact subgroup \( \mathcal{K} \) of \( \mathcal{D}_+ \backslash \mathcal{D}_{\mathbb{A}_k}^+ \) satisfying that for all \( g \) in \( \mathcal{D}_+ \backslash \mathcal{D}_{\mathbb{A}_k}^+ \) and \( a \) in \( \mathbb{A}_k^\times \),
\[
 f(ag) = \varpi(a)f(g).
\]

Suppose \( \mathcal{D} = \text{Mat}_2(k) \). Then \( \mathcal{D}_+^\times = \text{GL}_2(k) \) and \( \mathcal{D}_{\mathbb{A}_k}^\times = \text{GL}_2(\mathbb{A}_k) \). \( f \) is called a cusp form if for all \( g \) in \( \text{GL}_2(\mathbb{A}_k) \),
\[
 \int_{\mathbb{A}_k^\times} f \left( \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} g \right) du = 0.
\]

We denote \( \mathbf{A}_0(\varpi) \) to be the space of automorphic cusp forms on \( \text{GL}_2(\mathbb{A}_k) \) with central character \( \varpi \).

We recall Jacquet-Langlands correspondence in \( \S A.1 \) and use newform theory to explain the claim in \( \S 2.3 \). In \( \S A.2 \) we use multiplicity one theorem to show that the space \( \mathbf{M}^{\text{new}}(\Gamma_0(N_0)) \) in \( \S 2.3 \) is a free \( \mathbb{T}_{\mathbb{C}} \)-module of rank one.

A.1 Jacquet-Langlands correspondence

Let \( \mathcal{D} = \mathcal{D}(N_0) \) be a definite quaternion algebra over \( k \) where \( N_0 \) is the product of finite ramified primes of \( \mathcal{D} \). Let \( \mathbf{A}'(\varpi) \) be the space of automorphic forms on \( \mathcal{D}_{\mathbb{A}_k}^\times \) with central character \( \varpi \). Jacquet-Langlands correspondence describes the connection between \( \mathbf{A}'(\varpi) \) and \( \mathbf{A}_0(\varpi) \):

\[ L(s, \varpi' \otimes \rho) = L(s, \varpi' \otimes \rho') \]

for all Hecke characters \( \varpi' \).

Note that \( \rho = \otimes_v \rho_v \) where \( \rho_v = \rho'_v \) for finite primes \( v \) not dividing \( N_0 \). Moreover, for the ramified primes \( v \) of \( \mathcal{D} \), \( \rho_v \) is determined from \( \rho'_v \) via theta correspondence.

Conversely, suppose \( \rho = \otimes_v \rho_v \) is a constituent of \( \mathbf{A}_0(\varpi) \). If for every ramified primes \( v \) of \( \mathcal{D} \) the representation \( \rho_v \) is special or supercuspidal, then there is a
constituent \( \rho' = \otimes \rho'_v \) of \( A'(\varpi) \) such that \( \rho_v = \rho'_v J_L \). In particular, \( \rho'_v \) is one dimensional for ramified prime \( v \) if and only if \( \rho_v \) is special.

Let \( R \) be a fixed maximal order of \( \mathcal{D} \). From Jacquet-Langlands correspondence one has an isomorphism \( \Psi \) between

\[
\{ \text{\( \mathbb{C} \)-valued non-constant functions on} \ \hat{R}^\times \backslash \hat{\mathcal{D}}^\times /\mathcal{D}^\times \} \quad \text{and} \quad \{ \text{Drinfeld type new forms on} \ \Gamma_0(N_0) \backslash \text{GL}_2(k_\infty)/\Gamma_\infty k_\infty^\times \}
\]

which satisfies

\[
\Psi(t_m f) = T_m \Psi(f)
\]

for all non-constant functions \( f \) on \( \hat{R}^\times \backslash \hat{\mathcal{D}}^\times /\mathcal{D}^\times \) and monic polynomials \( m \) in \( A \). We briefly recall the argument in the following and refer the reader to [9] for further details.

Fix \( \varpi = \otimes_v \varpi_v \) to be the trivial Hecke character on \( k^\times \backslash k_\infty^\times \). Let \( v \) be a prime of \( k \), \( \mathcal{O}_v \) be the valuation ring in \( k_v \), and \( \pi_v \) a uniformizer in \( \mathcal{O}_v \). Recall that an irreducible admissible infinite-dimensional representation \( (\rho_v, V_v) \) of \( \text{GL}_2(k_v) \) with central character \( \varpi_v \) has conductor \( v^{c_v} \) if \( \pi_v^{c_v} \mathcal{O}_v \) is the largest ideal of \( \mathcal{O}_v \) such that the space of elements \( u \in V_v \) with

\[
\rho_v(g_v)u = u \quad \text{for all} \quad g_v \in K_c(v)
\]
is non-empty. In fact, it is one dimensional. Here

\[
\mathcal{K}_0^{c(v)} = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{GL}_2(\mathcal{O}_v) : c \in \pi_v^{c(v)} \mathcal{O}_v \right\}.
\]

It is known that

\[
c(v) = \begin{cases} 0 & \text{if} \ \rho_v \text{ is an unramified principal series}, \\ 1 & \text{if} \ \rho_v \text{ is an unramified special representation}, \\ \geq 2 & \text{if} \ \rho_v \text{ is supercuspidal or ramified}. \end{cases}
\]

Let \( (\rho, V) = \otimes_v (\rho_v, V_v) \) be a constituent of \( A_0(\varpi) \). The conductor of \( \rho \) is:

\[
\prod_v v^{c_v(v)}.
\]

The space of elements \( f \in V \) with

\[
\rho(g) f = f \quad \text{for all} \quad g \in \prod_v \mathcal{K}_0^{c_v(v)}
\]
is one dimensional, and called the space of new-forms of \( \rho \). Any new-form \( f \) of \( \rho \) is a Hecke eigenform, i.e. \( T_v f = a_v f \) for all \( v \) where \( a_v \in \mathbb{C} \).

Recall that \( L(s, \rho) = \prod_v L(s, \rho_v) \), where

\[
L(s, \rho_v) = \left( 1 - \chi_v(\pi_v)q^{-s \deg v} \right)^{-1} \cdot \left( 1 - \chi_v(\pi_v)q^{-s \deg v} \right)^{-1}
\]

if \( \rho_v \) is an unramified principal series \( \pi(\chi_{v,1}, \chi_{v,2}) \);

\[
L(s, \rho_v) = \left( 1 - \chi_v(\pi_v)q^{-(s+1/2) \deg v} \right)^{-1}
\]

if \( \rho_v \) is an unramified special representation \( \sp(\chi_v \cdot \lfloor v \rfloor, \chi_v \cdot \lfloor v \rfloor^{-1/2}) \);

\[
L(s, \rho_v) = 1
\]

if \( \rho_v \) is supercuspidal or ramified. Here \( \chi_{v,1}, \chi_{v,2} \), and \( \chi_v \) are unramified characters of \( k_v^\times \) with \( \chi_{v,1} \cdot \chi_{v,2} = 1 = \chi_v^2 \). It is known that

\[
a_v = \begin{cases} 
q^\frac{\deg v}{2}(\chi_{v,1}(\pi_v) + \chi_{v,2}(\pi_v)) & \text{if } \rho_v \cong \pi(\chi_{v,1}, \chi_{v,2}), \\
\chi_v(\pi_v) & \text{if } \rho_v \cong \sp(\chi_v \cdot \lfloor v \rfloor, \chi_v \cdot \lfloor v \rfloor^{-1/2}).
\end{cases}
\]

Suppose \( \rho = \otimes_v \rho_v \) is of conductor \( N_0\infty \) and \( \rho_\infty \cong \sp(\lfloor \cdot \rfloor_{\infty}, \chi_{\infty} \cdot \lfloor \cdot \rfloor_{\infty}^{-1/2}) \). Then new-forms of \( \rho \) are functions on

\[
\GL_2(k) \setminus \GL_2(\mathcal{A}) / \mathcal{K}_{0}(N_0\infty)k_\infty^\times.
\]

From the bijection in §2.1

\[
\GL_2(k) \setminus \GL_2(\mathcal{A}) / \mathcal{K}_{0}(N_0\infty)k_\infty^\times \cong \Gamma_0(N_0) \setminus \GL_2(k_0) / \Gamma_0 k_\infty^\times,
\]

new-forms of such \( \rho \) can be viewed as newforms of Drinfeld type for \( \Gamma_0(N_0) \). In fact, the space \( S_{\text{new}}(\Gamma_0(N_0)) \) of Drinfeld type new forms for \( \Gamma_0(N_0) \) is spanned by the new-forms of such \( \rho \) with conductor \( N_0\infty \).

Since \( \rho \) is of conductor \( N_0\infty \), \( \rho_P \cong \sp(\chi_P \cdot \lfloor P \rfloor, \chi_P \cdot \lfloor P \rfloor^{-1/2}) \) for all \( P \mid N_0 \) where \( \chi_P \) is an unramified character of \( k_P^\times \) with \( \chi_P^2 = 1 \). By Jacquet-Langlands correspondence we can find an irreducible constituent \( (\rho', V') = \otimes_v \rho'_v \) of \( \mathcal{A}'(\pi) \) so that \( \rho = \rho'_{\mathcal{A}} \). In this case, \( \rho'_P = \chi_P \circ \text{Nr} \) for \( P \mid N_0 \) and \( \rho'_\infty \) is the trivial representation. Therefore we can find a subspace of elements \( f' \in V' \) which are non-constant functions on

\[
\mathcal{D}^\times \setminus \hat{\mathcal{D}}^\times / \hat{R}^\times.
\]

This subspace is also one dimensional, called the space of new-forms of \( \rho' \). Any new-form \( f' \) of \( \rho' \) is also a Hecke eigenform, i.e. \( t_v f' = a'_v f' \), where \( a'_v \) appears in the local factor \( L_v(s, \rho'_v) \). Since for any place \( v \)

\[
L(s, \rho_v) = L(s, \rho'_v),
\]
we have $a_v = a'_v$. 

In fact, the space of non-constant functions on $\mathcal{D}^x \setminus \hat{\mathcal{D}}^x / \hat{\mathcal{R}}^x$ is generated by new-forms such that $\rho' = \otimes_v \rho'_v$ where $\rho'_\infty$ is trivial and for $P \mid N_0$, $\rho'_P = \chi_P \circ \text{Nr}$ for an unramified character $\chi_P$ of $k_P$ with $\chi_P^2 = 1$. By taking conjugate, we identify functions on $\mathcal{D}^x \setminus \hat{\mathcal{D}}^x / \hat{\mathcal{R}}^x$ with functions on $\hat{\mathcal{R}}^x \setminus \hat{\mathcal{D}}^x / \mathcal{D}^x$. From the dimension formula at the end of §2.2 we have a bijective map $\Psi$ from

$$\{ \text{C-valued non-constant functions on } \hat{\mathcal{R}}^x \setminus \hat{\mathcal{D}}^x / \mathcal{D}^x \}$$

to

$$\{ \text{Drinfeld type new forms on } \Gamma_0(N_0) \setminus \text{GL}_2(k_{\infty}) / \Gamma_{\infty} k_{\infty}^{*} \}$$

so that for each monic polynomial $m$ in $A$,

$$\Psi(t_m f) = T_m \Psi(f).$$

Since constant functions on $\hat{\mathcal{R}}^x \setminus \hat{\mathcal{D}}^x / \mathcal{D}^x$ are eigenfunctions of $t_m$ with eigenvalue $\sigma(m)_{N_0}$, we extend $\Psi$ by mapping constant functions into the one dimensional subspace $\mathcal{C} \mathcal{E}_{N_0}$ of $M_{\text{new}}(\Gamma_0(N_0))$.

Consider the definite Shimura curve $X = X_{N_0}$. We have a canonical bijection between components of $X$ and ideal classes of $\mathcal{R}$ and this gives the canonical isomorphism

$$\{ \text{(C-valued) functions on } \hat{\mathcal{R}}^x \setminus \hat{\mathcal{D}}^x / \mathcal{D}^x \} \cong \text{Hom}(\text{Pic}(X), \mathbb{C}) \cong \text{Pic}(X)^{\vee} \otimes_{\mathbb{Z}} \mathbb{C}.$$

Therefore one has:

**Theorem A.1.** $\Psi : \text{Pic}(X)^{\vee} \otimes_{\mathbb{Z}} \mathbb{C} \cong M_{\text{new}}(\Gamma_0(N_0))$ is an isomorphism so that $\Psi(t_m f) = T_m \Psi(f)$ for any monic polynomial $m$ in $A$. Moreover,

$$\text{Tr}(t_m) = \text{Tr}(T_m)$$

and so the $\mathbb{C}$-algebra $\mathbb{T}_{\mathbb{C}}$ generated by Hecke correspondences $t_m$ on $X$ is isomorphic to the $\mathbb{C}$-algebra generated by Hecke operators $T_m$ on $M_{\text{new}}(\Gamma_0(N_0))$.

**A.2 Multiplicity one theorem**

Let $\varpi : \mathbb{A}^{1}_{k}/k^{*}$ be a Hecke character. Let $\rho_1 = \otimes_v \rho_{1,v}$ and $\rho_2 = \otimes_v \rho_{2,v}$ be two irreducible admissible representations which are constituents of $\mathcal{A}_0(\varpi)$. The **multiplicity one theorem** (cf. [3]) tells us that $\rho_1 = \rho_2$ if and only if

$$\rho_{1,v} \cong \rho_{2,v}$$

for all place $v$.

Fix $\varpi$ to be trivial. Choose two irreducible admissible representations $\rho_1 = \otimes_v \rho_{1,v}$ and $\rho_2 = \otimes_v \rho_{2,v}$ of conductor $N_{0\infty}$ which are constituents of $\mathcal{A}_0(\varpi)$ satisfying

$$\rho_{1,\infty} \cong \rho_{2,\infty} \cong \text{sp}(| \cdot |_{\infty}^{-1/2}, | \cdot |_{\infty}^{1/2})$$
and \( \rho_{1,P} \) and \( \rho_{2,P} \) are unramified special representations for \( P \mid N_0 \). Let \( f_1 \) and \( f_2 \) be new-forms of \( \rho_1 \) and \( \rho_2 \) respectively. Then \( T_P f_i = a_{P,i} f_i \) where \( a_{P,i} \in \mathbb{C} \) for \( i = 1, 2 \) and all prime \( P \) in \( \mathbb{A} \). If \( a_{P,1} = a_{P,2} \) for all \( P \), then \( L_P(s, \rho_{1,P}) = L_P(s, \rho_{2,P}) \) and so

\[
\rho_{1,P} \cong \rho_{2,P}
\]

for all \( P \). By multiplicity one theorem we have \( \rho_1 = \rho_2 \) and so \( f_1, f_2 \) are linearly dependent.

Recall that \( M_{\text{new}}(\Gamma_0(N_0)) = S_{\text{new}}(\Gamma_0(N_0)) \oplus \mathbb{C} E_{N_0} \). for \( \Gamma_0(P_0) \). As a \( \mathbb{T}_C \)-module, the space \( M_{\text{new}}(\Gamma_0(N_0)) \) is a direct sum \( (\oplus_i C f_i) \oplus \mathbb{C} E_{N_0} \) of one dimensional submodules and each \( f_i \) is a new-form of an irreducible admissible representation \( \rho_i = \otimes_v \rho_{i,v} \) which is a constituent of \( A_0(\mathfrak{a}) \) with

\[
\rho_{i,\infty} \cong sp(|\cdot|_{\infty}^{1/2}, |\cdot|_{\infty}^{-1/2})
\]

and \( \rho_{i,P} \) is an unramified special representation for \( P \mid N_0 \). According to multiplicity one theorem, each pair of these one dimensional submodules are non-isomorphic. Therefore \( M_{\text{new}}(\Gamma_0(N_0)) \) is a cyclic \( \mathbb{T}_C \)-module, which is generated by \( E_{N_0} + \sum_i f_i \). Viewing \( \mathbb{T}_C \) as a subring of \( \text{End}_C(M_{\text{new}}(\Gamma_0(N_0))) \), we have

\[
\dim_{\mathbb{C}} \mathbb{T}_C \leq \dim_{\mathbb{C}} M_{\text{new}}(\Gamma_0(N_0)).
\]

Therefore

**Proposition A.2.** The space \( M_{\text{new}}(\Gamma_0(N_0)) \) is a free \( \mathbb{T}_C \)-module of rank one.

### B Transformation law of theta series

Fix a definite quaternion algebra \( \mathcal{D} = \mathcal{D}(N_0) \) where \( N_0 \) is the product of finite ramified primes of \( \mathcal{D} \). Let \( R \) be a maximal order and \( n \) be the class number. In this section we deduce the transformation law of the theta series \( \theta_{ij} \) for \( 1 \leq i, j \leq n \) introduced in §2.1.1. Recall that for each \((i, j)\), theta series \( \theta_{ij} \) is a function on \( k_{\infty} \times k_{\infty} \times k_{\infty} \):

\[
\theta_{ij}(x, y) = \sum_{b \in M_{ij}} \phi_{\infty}(\frac{Nr(b)}{N_{ij}} x^2) \cdot \psi_{\infty}(\frac{Nr(b)}{N_{ij}} y),
\]

where \( \phi_{\infty} \) is the characteristic function of \( O_{\infty} \) and \( \psi_{\infty} \) is the fixed additive character on \( k_{\infty} \).

#### B.1 Fourier Transform

Let \( \mathcal{D}_{\infty} = \mathcal{D} \otimes_k k_{\infty} \). For \( \alpha, \beta \in k_{\infty}^\times \) with \( v_\infty(\alpha) > v_\infty(\beta) - 2 \), let

\[
\Phi_{\alpha, \beta} : \mathcal{D}_{\infty} \longrightarrow \mathbb{C}
\]

\[
w \longmapsto \phi_{\infty}(\text{Nr}(w)\alpha) \psi_{\infty}(\text{Nr}(w)\beta).
\]
Define \([\cdot,\cdot]: \mathcal{D}_\infty \times \mathcal{D}_\infty \to \mathbb{C}^\times\) by \([w,w^*] := \psi_\infty(\text{Tr}(ww^*)).\) The Fourier transform of \(\Phi_{\alpha,\beta}\) is given by:

\[
\Phi_{\alpha,\beta}^*(w^*) := \int_{\mathcal{D}_\infty} \Phi_{\alpha,\beta}(w,w^*)dw,
\]
for all \(w^* \in k_\infty\)

where \(dw\) is a Haar measure on \(\mathcal{D}_\infty\).

We define

\[
S(\alpha,\beta,dw) := \int_{\mathcal{D}_\infty} \phi_\infty(\text{Nr}(w)\alpha)\psi_\infty(\text{Nr}(w)\beta)dw.
\]

Then \(\Phi_{\alpha,\beta}^*(w^*)\) is equal to

\[
S(\alpha,\beta,dw)\phi_\infty(\text{Nr}(w^*)\frac{\alpha}{\beta})\psi_\infty(\text{Nr}(w^*)\frac{-1}{\beta} - h\mu^*).
\]

More generally, take \(h \in k_\infty^\times\), \(\rho \in \mathcal{D}_\infty\). For \(\alpha,\beta \in k_\infty^\times\) with \(v_\infty(\alpha) > v_\infty(\beta) - 2\), let \(\Psi_{\alpha,\beta}(w) := \Phi_{\alpha,\beta}(\rho + hw)\). Then \(\Phi_{\alpha,\beta}^*(w^*)\) is equal to

\[
q^{4v_\infty(h)}S(\alpha,\beta,dw)\phi_\infty(\text{Nr}(\frac{w^*}{h})\frac{\alpha}{\beta})\psi_\infty(\text{Nr}(\frac{w^*}{h})\frac{-1}{\beta} - h\mu^*)\psi_\infty(\text{Tr}(\frac{\rho w^*}{h})).
\]

B.2 Poisson summation

Let \(\mathcal{O}_{\mathcal{D}_\infty}\) be the maximal order of \(\mathcal{D}_\infty\). For \(v_\infty(\alpha) > v_\infty(\beta) - 2\), we have

\[
S(\alpha,\beta,dw) = -q^{2v_\infty(\beta) - 3}dw(\mathcal{O}_{\mathcal{D}_\infty}).
\]

For the pair \((i,j), 1 \leq i,j \leq n\), we choose Haar measure \(dw\) with \(dw(\mathcal{D}_\infty/M_{ij}) = 1\) and denote the integral \(S(\alpha,\beta,dw)\) by \(S(\alpha,\beta,M_{ij})\). Then

\[
S(\alpha,\beta,M_{ij}) = -q^{2v_\infty(\beta) - \text{deg}(N_0)}q^{2v_\infty(N_{ij})}.
\]

Let \(\tilde{M}_{ij}\) be the dual lattice of \(M_{ij}\), i.e.,

\[
\tilde{M}_{ij} = \{w \in \mathcal{D}_\infty : \text{Tr}(w\mu) \in A \text{ for all } \mu \in M_{ij}\}.
\]

We apply the Poisson summation formula

\[
\sum_{\mu \in M_{ij}} \Psi_{\alpha,\beta}(\mu) = \sum_{\mu^* \in \tilde{M}_{ij}} \Psi_{\alpha,\beta}^*(\mu^*)
\]

and get

**Proposition B.1.** Let \(\alpha,\beta \in k_\infty^\times\) with \(v_\infty(\alpha) > v_\infty(\beta) - 2\), \(h \in k_\infty^\times\), \(\rho \in \mathcal{D}_\infty\). Then

\[
\sum_{\mu \in M_{ij}} \phi_\infty(\text{Nr}(\rho + h\mu)\alpha)\psi_\infty(\text{Nr}(\rho + h\mu)\beta)
\]

\[
= q^{4v_\infty(h)}S(\alpha,\beta,M_{ij}) \sum_{\mu^* \in \tilde{M}_{ij}} \phi_\infty(\text{Nr}(\frac{\mu^*}{h})\frac{\alpha}{\beta})\psi_\infty(\text{Nr}(\frac{\mu^*}{h})\frac{-1}{\beta} - h\mu^*)\psi_\infty\left(\text{Tr}(\frac{\rho \mu^*}{h})\right).
\]
Let \( x \in k_\infty^x, y \in k_\infty, M \subset D_\infty \) a discrete \( A \)-lattice, \( N_M \in k \) such that \( N_M \cdot A \) is the fractional ideal of \( A \) generated by \( \text{Nr}(\mu) \) for \( \mu \in M \). For \( h \in A \) with \( h \neq 0, \rho \in M \), define "partial theta" series:

\[
\theta(x, y, M, N_M, h, \rho) := \sum_{\mu \in M, \mu \equiv \rho \mod hM} \phi_{\infty} \left( \frac{\text{Nr}(\mu)x^2}{N_M h} \right) \psi_{\infty} \left( \frac{\text{Nr}(\mu)y}{N_M h} \right).
\]

Note that \( \theta(x, y) = \theta(x, y, M_1, N_1, 1, 0) \), and

\[
\theta(x, y, M, N_M, h, \rho) = \sum_{\mu \in M} \phi_{\infty}(\text{Nr}(\rho + h\mu)\alpha)\psi_{\infty}(\text{Nr}(\rho + h\mu)\beta)
\]

where \( \alpha = \frac{x^2}{N_M h}, \beta = \frac{y}{N_M h} \).

**Proposition B.2.** Let \( x, y \in k_\infty^x, v_\infty(x) > v_\infty(y), 0 \neq h \in A, \kappa \in \tilde{M}_{ij} \). Then

\[
\theta(x, y, M_{ij}, N_{ij}, h, \rho) = S' \left( \frac{xt^2}{N_{ij}N_0h}, \frac{yt}{N_{ij}N_0h}, M_{ij} \right) - 1 \sum_{\rho \in M_{ij}/hM_{ij}} \psi_{\infty}(\text{Tr}(\frac{\rho K}{h}))\theta(\frac{x}{N_0}, \frac{y}{N_0}, M_{ij}, N_{ij}, h, \rho).
\]

**Proof.** By Proposition B.1 we have

\[
\theta(x, y, M_{ij}, N_{ij}, h, \rho)
= q^{\nu_{\infty}(h)}S(\alpha, \beta, M_{ij}) \sum_{\mu^* \in M_{ij}} \phi_{\infty} \left( \text{Nr}(\frac{\mu^*}{h})\frac{\alpha}{\beta^2} \right) \psi_{\infty} \left( \text{Nr}(\frac{\mu^*}{h})\frac{1}{\beta} \right) \psi_{\infty} \left( \text{Tr}(\frac{-\mu^*}{h}) \right).
\]

Multiply this by \( \psi_{\infty}(\text{Tr}(\frac{\rho K}{h})) \) for \( \kappa \in \tilde{M}_{ij} \) and sum over \( \rho \in M_{ij}/hM_{ij} \), we obtain

\[
\sum_{\rho \in M_{ij}/hM_{ij}} \psi_{\infty}(\text{Tr}(\frac{\rho K}{h})) \cdot \theta(x, y, M_{ij}, N_{ij}, h, \rho)
= q^{\nu_{\infty}(h)}S(\alpha, \beta, M_{ij}) \sum_{\mu^* \in M_{ij}} \left\{ \phi_{\infty} \left( \text{Nr}(\frac{\mu^*}{h})\frac{\alpha}{\beta^2} \right) \psi_{\infty} \left( \text{Nr}(\frac{\mu^*}{h})\frac{1}{\beta} \right) \right. \\
\left. \psi_{\infty} \left( \text{Tr}(\frac{\rho K}{h}(\kappa - \mu^*)) \right) \right\}.
\]

Since

\[
\sum_{\rho \in M_{ij}/hM_{ij}} \psi_{\infty}(\text{Tr}(\frac{\rho K}{h}(\kappa - \mu^*))) = \begin{cases} 0 & \text{if } \mu^* - \kappa \notin h\tilde{M}_{ij}, \\ q^{-\nu_{\infty}(h)} & \text{if } \mu^* - \kappa \in h\tilde{M}_{ij}, \end{cases}
\]

The proposition follows by replacing \( x \) with \( \frac{x}{N_0} \), and \( y \) with \( \frac{y}{N_0} \). \( \square \)
B.3 Transformation law

Let \((x, y) \in \mathbb{R}^\times \times \mathbb{R}^\infty\). Suppose a matrix \(\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{GL}_2(A)\) is given such that \(cy + d \neq 0\). We define

\[
\gamma \circ (x, y) := \begin{pmatrix} x(ad - bc) + ay & by \\ cy + d \end{pmatrix}.
\]

Lemma B.3. Suppose \(\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}_2(A)\), \(c \equiv 0 \mod N_0\), \(v_\infty(x) > v_\infty(y)\), and \(v_\infty(cx) > v_\infty(cy + d)\). Let \(1 \leq i, j \leq n\). Then

\[
\theta_{ij}(\gamma \circ (x, y)) = S(N_{ij} x^2, -N_{ij} (cy + d), \tilde{M}_{ij})^{-1} \cdot \left( \sum_{\kappa \in M_{ij}/dM_{ij}} \psi_\infty \left( \frac{N_{ij} b}{N_{ij} d} \right) \right) \theta_{ij}(x, y).
\]

Proof. Put \(u = \frac{x}{y}, v = \frac{d}{y}\). Then

\[
\theta_{ij}(\gamma \circ (x, y)) = \theta \left( \frac{u}{(c - dv)^2} \cdot \frac{b}{d} + \frac{1}{d(c - dv)}, M_{ij}, N_{ij}, 1, 0 \right)
\]

\[
= \sum_{\kappa \in M_{ij}/dM_{ij}} \theta \left( \frac{du}{c - dv}, b + \frac{1}{c - dv}, M_{ij}, N_{ij}, d, \kappa \right)
\]

\[
= \sum_{\kappa \in M_{ij}/dM_{ij}} \psi_\infty \left( \frac{N_{ij} b}{N_{ij} d} \right) \theta \left( \frac{du}{(dv - c)^2}, -1 \cdot \frac{dv}{dv - c}, M_{ij}, N_{ij}, d, \kappa \right).
\]

Since \(v_\infty(cx) > v_\infty(cy + d)\), we have \(v_\infty(du) > v_\infty(dv - c)\) and

\[
\theta_{ij}(\gamma \circ (x, y)) = S(N_{ij} x^2, N_{ij}(v - c/d), \tilde{M}_{ij})^{-1} \cdot \sum_{\kappa \in M_{ij}/dM_{ij}} \left[ \psi_\infty \left( \frac{N_{ij} b}{N_{ij} d} \right) \right]
\]

\[
\cdot \sum_{\rho \in M_{ij}/dM_{ij}} \psi_\infty \left( \text{Tr} \left( \frac{\rho d}{d} \right) \right) \theta \left( \frac{dv}{N_0}, \frac{dv - c}{N_0}, \tilde{M}_{ij}, N_{ij}^{-1} N_0^{-1}, d, \rho \right).
\]

Since \(-c/N_0 \in A\), we have

\[
\theta_{ij}(\gamma \circ (x, y)) = \left[ \sum_{\rho \in M_{ij}/dM_{ij}} \psi_\infty \left( \frac{N_{ij} b}{N_{ij} d} \right) \right]
\]

\[
\cdot \left[ \sum_{\rho \in M_{ij}/dM_{ij}} \psi_\infty \left( \text{Tr} \left( \frac{\rho d}{d} \right) \right) \theta \left( \frac{dv}{N_0}, \frac{dv - c}{N_0}, \tilde{M}_{ij}, N_{ij}^{-1} N_0^{-1}, d, \rho \right) \right]
\]

\[
\cdot \left[ \sum_{\kappa \in M_{ij}/dM_{ij}} \psi_\infty \left( \frac{N_{ij} b}{N_{ij} d} \right) + \frac{\text{Tr} (\rho \kappa)}{d} - \frac{N_{ij} c}{d} \right] \right).
\]
Note that $cN_{ij}\bar{\rho} \in M_{ij}$. Replacing $\kappa$ by $\kappa + cN_{ij}\bar{\rho}$ the last summand equals to
\[
\frac{\Nr(\kappa)b}{N_{ij}d} + a \Tr(\rho\kappa) + N_{ij}ac\Nr(\rho).
\]
Since $a \Tr(\rho\kappa) + N_{ij}ac\Nr(\rho) \in A$, we have
\[
\theta_{ij}(\gamma \circ (x, y)) = S(N_{ij}ut^2, N_{ij}(v - c/d), \tilde{M}_{ij})^{-1} \cdot \sum_{\kappa \in M_{ij}/dM_{ij}} \psi_\infty(\frac{\Nr(\kappa)b}{N_{ij}d}) \theta_{ij}(x, y).
\]
Recall that $u = \frac{x}{y^2}$, $v = \frac{1}{y}$. By Proposition B.2 we have
\[
\theta_{ij}(g \circ (x, y)) = S\left(\frac{N_{ij}xt^2}{y^2}, \frac{-N_{ij}(cy + d)}{dy}, \tilde{M}_{ij}\right)^{-1} \cdot \sum_{\kappa \in M_{ij}/dM_{ij}} \psi_\infty(\frac{\Nr(\kappa)b}{N_{ij}d}) \theta_{ij}(x, y).
\]
Note that
\[
S\left(\frac{N_{ij}xt^2}{y^2}, \frac{-N_{ij}(cy + d)}{dy}, \tilde{M}_{ij}\right) \cdot S\left(\frac{xt^2}{N_{ij}}, \frac{y}{N_{ij}}, M_{ij}\right) = q^{2v_\infty(cx + d) + 2 \deg d}.
\]
By standard argument we get $\sum_{\kappa \in M_{ij}/dM_{ij}} \psi_\infty(\frac{\Nr(\kappa)b}{N_{ij}d}) = q^{2 \deg d}$. Since $\theta_{ij}(x, y) = \theta_{ij}(x, y + h)$ for any $h \in A$, we can drop the assumption $v_\infty(x) > v_\infty(y)$ and obtain the transformation law of $\theta_{ij}$:

**Theorem B.4.** For $1 \leq i, j \leq n$. Let $x \in k_\infty, y \in k_\infty$, $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \SL_2(A)$.
Assume $v_\infty(cx) > v_\infty(cy + d)$, and $c \equiv 0 \mod N_0$. Then
\[
\theta_{ij}(\gamma \circ (x, y)) = q^{-2v_\infty(cy + d)} \cdot \theta_{ij}(x, y).
\]

**Acknowledgements**
We would like to thank the referee for referring us to the work of Papikian [10].
References

[1] Artin, E., *Quadratische Körper der komplexen Multiplikation*, Enzyklopädie der Math. Wiss. Band I, 2. Teil, Heft 10, Teil II.

[2] Bertolini, M. & Darmon, H., *Heegner points on Mumford-Tate curves*, Inv. Math., 126 (1996) 413-456.

[3] Bump, D., *Automorphic Forms and Representations*, Cambridge studies in advanced mathematics 55, (1996).

[4] Eichler, M., *Zur Zahlentheorie der Quaternionen-Algebren*, Crelle J. 195 (1955), 127-151.

[5] Gekeler, E.-U., *Invariants of Some Algebraic Curves Related to Drinfeld Modular Curves*, J. Number Theory 90 (2001) 166-183.

[6] Gekeler, E.-U., *Improper Eisenstein Series on Bruhat-Tits Trees*, manuscripta math. 86 (1995), 367-391.

[7] Gekeler, E.-U. & Reversat, M., *Jacobians of Drinfeld modular curves*, J. reine angew. Math. 476 (1996), 27-93.

[8] Gross, B. H., *Heights and the Special Values of L-series*, CMS Conference Proceedings, H. Kisilevsky, J. Labute, Eds., 7 (1987) 116-187.

[9] Jacquet, H. & Langlands, R., *Automorphic Forms on GL(2)*, LNM 114, Springer 1970.

[10] Papikian, M., *On the variation of Tate-Shafarevich groups of elliptic curves over hyperelliptic curves*, J. Number Theory 115 (2005) 249-283.

[11] Rück, H.-G., *Theta Series of Imaginary Quadratic Function Fields*, manuscripta math. 88 (1995), 387-407.

[12] Rück, H.-G. & Tipp, U., *Heegner Points and L-series of Automorphic Cusp Forms of Drinfeld Type*, Documenta Mathematica 5 (2000) 365-444.

[13] Schweizer, A., *Strong Weil curves over \( \mathbb{F}_q(T) \) with small conductor*, J. Number Theory 131 (2011) 285-299.

[14] Serre, J.-P., *Tree*, Springer, Berlin-Heidelberg-New York 1980.

[15] Tan, K.-S. & Reckmore, D., *Computation of L-series for elliptic curves over function fields*, J. reine angew. Math. 424 (1992), 107-135.

[16] Vignéras, M.-F., *Arithmétique des Algèbres de Quaternions*, LNM 800, Springer 1980.

[17] Wei, F.-T., *On arithmetic of curves over function fields*. Ph.D. thesis, National Tsing Hua University, 2010.
[18] Weil, A., *Dirichlet Series and Automorphic Forms*, LNM 189, Springer 1971.
K-HOMOLOGY CLASS OF THE DIRAC OPERATOR ON A COMPACT QUANTUM GROUP

SERGEY NESHVEYEV AND LARS TUSET

Received: March 29, 2011
Revised: September 29, 2011
Communicated by Joachim Cuntz

Abstract. By a result of Nagy, the C*-algebra of continuous functions on the q-deformation $G_q$ of a simply connected semisimple compact Lie group $G$ is KK-equivalent to $C(G)$. We show that under this equivalence the K-homology class of the Dirac operator on $G_q$, which we constructed in an earlier paper, corresponds to that of the classical Dirac operator. Along the way we prove that for an appropriate choice of isomorphisms between completions of $U_q g$ and $U g$ a family of Drinfeld twists relating the deformed and classical coproducts can be chosen to be continuous in $q$.

2000 Mathematics Subject Classification: 58B34, 58B32, 46L80
Keywords and Phrases: Quantum groups, Dirac operator, K-homology

Introduction

In [10] we constructed a Dirac operator $D_q$ on the q-deformation $G_q$ of any simply connected semisimple compact Lie group $G$. The construction involved a special unitary element $F^q$ in the von Neumann algebra $W^*(G) \hat{\otimes} W^*(G)$, which relates the coproducts in $W^*(G_q)$ and $W^*(G)$. The existence of such an element, called a unitary Drinfeld twist, is a consequence of a fundamental and highly nontrivial result in quantum group theory due to Kazhdan and Lusztig [5] (see also [11]). Since the construction of a Drinfeld twist is involved and not particularly explicit, certain properties of the operators $D_q$ are not immediate. In particular, even though it is intuitively clear that $D_q$ is a deformation of the classical Dirac operator and therefore should in some sense define the same index map on K-theory, it is not even obvious that the K-homology

*Supported by the Research Council of Norway.
class of $D_q$ is always nonzero. The goal of this note is to show that it is indeed nonzero. In fact we show that the $K$-homology class of $D_q$ corresponds exactly to that of the classical Dirac operator via the KK-equivalence of $C(G_q)$ and $C(G)$ established by Nagy [9]. Therefore, upon identifying the $K$-theories of $G_q$ and $G$, the index map defined by $D_q$ does not depend on $q$, as expected. We should remark that the question of invariance of the index map under deformation has been studied in a recent paper by Yamashita [15] in the context of Connes-Landi $\theta$-deformations.

The paper is organized as follows. In Section 1 we show that the family of $C^*$-algebras $C(G_q)$ has a canonical continuous field structure. The result is more or less known [9], but is usually formulated in terms of standard generators of $\mathbb{C}[G_q]$. We propose a simpler approach based on a natural notion of a continuous family of isomorphisms $W^*(G_q) \cong W^*(G)$.

In Section 2 we prove that once a continuous family of isomorphisms $W^*(G_q) \cong W^*(G)$ is fixed, the corresponding family of Drinfeld twists $F_q$ can be chosen to be continuous in $q$. This result is not strictly speaking necessary for our main result on $D_q$, for which it suffices to know that $D_q$ does not depend on $F_q$ for a fixed isomorphism $W^*(G_q) \cong W^*(G)$ (see [12]), but it simplifies the arguments and is of independent interest. Both results, continuity of $F_q$ and uniqueness of $D_q$, depend crucially on the fact that any two unitary Drinfeld twists differ by the coboundary of a central unitary element, a result we proved in [12].

In Section 3 we prove our main result. For this we show that the family of operators $D_q$ define a Kasparov module for the algebra of continuous sections of $(C(G_q))_{q \in [a,b]}$ and the algebra $C[a,b]$. With the preparation in the previous two sections the proof essentially boils down to observing that estimates in our paper [10] are uniform in $q$.

Acknowledgement. The authors are grateful to Christian Voigt and Makoto Yamashita for stimulating discussions.

1. Continuous field of function algebras

Let $G$ be a simply connected semisimple compact Lie group, $\mathfrak{g}$ its complexified Lie algebra, $\mathfrak{h} \subset \mathfrak{g}$ the Cartan subalgebra defined by a maximal torus in $G$. Fix a system $\{\alpha_1, \ldots, \alpha_r\}$ of simple roots. Let $(a_{ij})_{1 \leq i,j \leq r}$ be the Cartan matrix defined by $\alpha_1, \ldots, \alpha_r$, and $d_1, \ldots, d_r$ be the coprime positive integers such that $(d_i a_{ij})_{i,j}$ is symmetric.

For every $q > 0$, $q \neq 1$, consider the quantized universal enveloping algebra $U_q \mathfrak{g}$ with generators $E_i^q, F_i^q, K_i^q$ ($1 \leq i \leq r$); we follow the conventions in [11, 12]. For $q = 1$ we let $U_1 \mathfrak{g} = U \mathfrak{g}$ and denote by $E_i, F_i, h_i$ the standard generators of $U \mathfrak{g}$. We will often omit various indices corresponding to $q = 1$. Consider the category $\mathcal{C}_q(\mathfrak{g})$ of finite dimensional admissible $U_q \mathfrak{g}$-modules and denote by $\mathcal{U}(G_q)$ the endomorphism ring of the forgetful functor $\mathcal{C}_q(\mathfrak{g}) \to Vec$. We think of $\mathcal{U}(G_q)$ as a completion of $U_q \mathfrak{g}$. For $q \neq 1$ denote by $h_i^q$ the unique self-adjoint element of $\mathcal{U}(G_q)$ such that $K_i^q = q^{d_i h_i^q}$. 
Since simple objects of \( C_q(g) \) are classified by dominant integral weights, for every \( q > 0 \) we have a canonical identification of the centers of \( U(G_q) \) and \( U(G) \). It extends to a \(*\)-isomorphism \( \varphi^q : U(G_q) \to U(G) \). For every dominant integral weight \( \lambda \in P_+ \) and \( q > 0 \) fix an irreducible \(*\)-representation \( \pi^q_\lambda : U_q g \to B(V^q_\lambda) \) with highest weight \( \lambda \) and a highest weight unit vector \( \xi^q_\lambda \in V^q_\lambda \). Then to define \( \varphi^q \) is the same as to fix an isomorphism \( B(V^q_\lambda) \cong B(V_\lambda) \) for every \( \lambda \).

We say that a family \( \{ \varphi^q \}_{q > 0} \) of \(*\)-isomorphisms extending the canonical identifications of the centers is continuous, if for every finite dimensional representation \( \pi \) of \( U \mathfrak{g} \) the operators \( \pi(\varphi^q(E_i^I)) \), \( \pi(\varphi^q(F_i^I)) \) and \( \pi(\varphi^q(h_i^I)) \) depend continuously on \( q \).

**Lemma 1.1.** A continuous family of \(*\)-isomorphisms \( \varphi^q : U(G_q) \to U(G) \) always exists.

**Proof.** It suffices to show that for every \( \lambda \in P_+ \) there exist unitaries \( u^q_\lambda : V^q_\lambda \to V_\lambda \) such that the operators \( u^q_\lambda \pi^q_\lambda(X^q) u^q_\lambda \) depend continuously on \( q \) for \( X^q = E^q_i, F^q_i, h^q_i \). For this, in turn, it is enough to show that such unitaries exist locally.

Therefore fix \( \lambda \) and \( q_0 > 0 \). For every multi-index \( I = (i_1, \ldots, i_k) \) \((1 \leq i_j \leq r)\) define

\[
e^q_I = \pi^q_\lambda(F^q_i \cdots F^q_1) \xi^q_\lambda \in V^q_\lambda.
\]

We can choose multi-indices \( I_1, \ldots, I_n \) such that the vectors \( e^{q_0}_{I_1}, \ldots, e^{q_0}_{I_n} \) form a basis in \( V^{q_0}_\lambda \). The quantum Serre relations, together with the identities \( F^{q_0}_i = (K^{q_0}_i)^{-1} E^{q_0}_i \), imply that the scalar products \( (e^q_I, e^q_J) \) depend continuously on \( q \).

Hence, for some \( \varepsilon > 0 \), the vectors \( e^q_{I_1}, \ldots, e^q_{I_n} \) form a basis in \( V^q_\lambda \) for all \( q \in (q_0 - \varepsilon, q_0 + \varepsilon) \). Applying the Gram-Schmidt orthogonalization we get an orthonormal basis \( \zeta^q_1, \ldots, \zeta^q_n \) in \( V^q_\lambda \). Let \( v^q : V^q_\lambda \to V^{q_0}_\lambda \) be the unitary mapping \( \zeta^q_i \) into \( \zeta^{q_0}_i \). By construction, for every multi-index \( I \), the coefficients of \( e^q_I \) in the basis \( \zeta^q_1, \ldots, \zeta^q_n \) depend continuously on \( q \), hence the matrix coefficients of \( \pi^q_\lambda(F^q_i) \) in this basis also depend continuously on \( q \). It follows that the operators \( e^q \pi^q_\lambda(F^q_i) e^q \) depend continuously on \( q \). The same is clearly true for \( h^q_i \) in place of \( F^q_i \) (in fact, we even have \( e^q \pi^q_\lambda(h^q_i) e^q = h^{q_0}_i \)), hence also for \( E^q_i = g^q a h^q_i F^q_i \).

Now take an arbitrary unitary \( u : V^{q_0}_\lambda \to V_\lambda \). Then the unitaries \( u^q_\lambda = u v^q \), \( q \in (q_0 - \varepsilon, q_0 + \varepsilon) \), have the required properties. \( \square \)

3. From now on we will fix a continuous family of \(*\)-isomorphisms \( \varphi^q : U(G_q) \to U(G) \) such that \( \varphi^1 = \iota \).

For every \( q > 0 \) denote by \( C[G_q] \subset U(G_q)^* \) the Hopf \(*\)-algebra of matrix coefficients of finite dimensional admissible \( U_q \mathfrak{g} \)-modules, and by \( C(G_q) \) its \( C^* \)-completion.

**Theorem 1.2.** The family of \( C^* \)-algebras \( C(G_q) \), \( q > 0 \), has a unique structure of a continuous field of \( C^* \)-algebras such that for every \( a \in C[G] \) the section \( q \mapsto a \varphi^q \in C(G_q) \) is continuous. This structure does not depend on the choice of a continuous family of \(*\)-isomorphisms \( \varphi^q \).
It is known that the $\mathcal{C}^*$-algebras $C(SU_q(N))$ have a continuous field structure such that the matrix coefficients of the fundamental representation form continuous sections. For $N = 2$ this was proved by Bauval [2] and for all $N \geq 2$ by Nagy [9]. As mentioned in [9] the same proof as for $SU_q(N)$ works for all other classical simple compact Lie groups. In principle the same result is also true for exceptional groups once explicit generators of $C$ other classical simple compact Lie groups. In principle the same result is also true for exceptional groups once explicit generators of $C$ have been found. The point of the above theorem is that there is actually no need to do this, it is enough to know that there exists a ‘continuous’ choice of generators. Although we will not need this here, we note that the theorem and its proof also imply that the families of function algebras on $q$-deformations of homogeneous spaces of $G$ can be given a continuous field structure without working out explicit generators and relations in those algebras. This, as well as the relation of the above result to Rieffel’s notion of strict deformation quantization, will be discussed in a subsequent paper [13].

Proof of Theorem 1.2. First consider the dependence of the continuous field structure on the isomorphisms $\varphi^q$. Assume we have another continuous family of isomorphisms $\psi^q$. For every $\lambda \in P_+$ denote by $\gamma^q_\lambda$ the unique automorphism of $B(V_\lambda)$ such that $\pi_\lambda \psi^q = \gamma^q_\lambda \pi_\lambda \varphi^q$. Then the map $q \mapsto \gamma^q_\lambda \in \text{Aut}(B(V_\lambda))$ is continuous. It follows that for any linear functional $\omega$ on $B(V_\lambda)$ the elements $\omega \gamma^q_\lambda \pi_\lambda \in \mathbb{C}[G]$ decompose into finite linear combinations of elements $\nu \pi_\lambda$ with continuous coefficients, so that the sections $q \mapsto \omega \pi_\lambda \varphi^q$ are finite linear combinations of sections $q \mapsto \nu \pi_\lambda \psi^q$ with continuous coefficients. Therefore if the latter sections are continuous, the former are continuous as well.

Since $\mathbb{C}[G_q]$ is dense in $C(G_q)$, it is also clear that the continuous field structure is unique if it exists.

To prove existence, first consider the case $G = SU(2)$. As usual identify the weight lattice $P$ with the half-integers. For every $s \in \frac{1}{2} \mathbb{Z}_+$ consider the orthonormal basis in $V^q$ consisting of the vectors $\|((F^q)^k \xi^q, \xi^q)^{-1}(F^q)^k \xi^q, k = 0, \ldots, 2s$. Let $u^q_{ij} \in \mathbb{C}[SU_q(2)]$ be the matrix coefficients of $\pi^q_\xi$ in this basis. We also use these bases to construct the isomorphisms $\varphi^q$, so that $u^q_{ij} = u^q_{ij} \varphi^q$. By [2], see also [3], there exists a continuous field structure on the $\mathcal{C}^*$-algebras $C(SU_q(2))$ such that the sections $q \mapsto u^q_{ij}$ are continuous. To prove that the sections $q \mapsto u^q_{ij}$ are continuous in this structure we just have to show that $u^q_{ij}$ can be expressed as polynomials of the elements $u^{1/2q}_{ij}$ with continuous coefficients. That this is indeed possible is easy to see using the unique embedding of $V^q$ into $(V^q_{1/2})^{\otimes 2s}$ mapping $\xi^q$ into $(\xi^q_{1/2})^{\otimes 2s}$; in fact an explicit expression for these polynomials is known [14].

Turning to the general case, for every simple root $\alpha$, consider the $*$-homomorphism $\sigma^q_\alpha : C(G_q) \to C(SU_q(2))$ which is dual to the embedding $\rho^q_\alpha : U(SU_q(2)) \hookrightarrow U(G_q)$ defined by $E^q_{\alpha^*} \mapsto E^q_{\alpha^*}$, $F^q_{\alpha^*} \mapsto F^q_{\alpha^*}$ and $h^q_{\alpha^*} \mapsto h^q_{\alpha^*}$. Let $w = s_1 \cdots s_{n_\alpha}$ be the longest element in the Weyl group of $G$ written in reduced form. Put $A_q = C(SU_q(s_{n_\alpha}(2)) \otimes \cdots \otimes C(SU_q(s_1(2)))$. Since $C(SU_q(2))$
is a C*-algebra of type I, the C*-algebra of continuous sections of the field \((C(SU_q(2)))_{q > 0}\) vanishing at infinity is of type I as well, hence exact. By [7, Theorem 4.6] it follows that the field \((A_q)_{q > 0}\) has a continuous field structure such that the tensor product of continuous sections is a continuous section. Define a *-homomorphism

\[
\sigma^q : C(G_q) \to A_q \quad \text{by} \quad \sigma^q(a) = (\sigma^q_1 \otimes \cdots \otimes \sigma^q_n) \Delta^q_n(a).
\]

It follows from the description of irreducible representations of \(C(G_q)\), see e.g. [8, Theorem 6.2.7], that \(\sigma^q\) is injective for every \(q\). Therefore the field \((C(G_q))_{q > 0}\) embeds into \((A_q)_{q > 0}\), so to prove existence of the required continuous field structure on \((C(G_q))_{q > 0}\) it suffices to show that for every \(a \in C[G]\) the section \(q \mapsto \sigma^q(a \varphi^q) \in A_q\) is continuous.

Since \(\varphi^q\) is an algebra homomorphism, the dual map \(C[G] \to C[G_q], a \mapsto a \varphi^q\), is a coalgebra homomorphism. Therefore, using Sweedler’s sumless notation, \(\Delta^q_{(n-1)}(a \varphi^q) = a_{(0)} \varphi^q \otimes \cdots \otimes a_{(n-1)} \varphi^q\). Hence to prove that the sections \(q \mapsto \sigma^q(a \varphi^q) \in A_q\) are continuous for all \(a \in C[G]\) it suffices to show that the sections \(q \mapsto \sigma^q_i(a \varphi^q)\) of the field \((C(SU_{q^i}(2)))_{q > 0}\) are continuous for all \(a \in C[G]\) and \(1 \leq i \leq r\).

Fix \(i\) and a continuous family of *-isomorphisms \(\theta^q : \mathcal{U}(SU_q(2)) \to \mathcal{U}(SU(2))\) such that \(\theta^1 = \iota\). To simplify the notation assume \(d_i = 1\), so \(\sigma^q_i\) is defined by an embedding \(\rho^q_i : U_q \mathfrak{sl}_2 \to U_q \mathfrak{b}\). To finish the proof it suffices to show that there exists a continuous family of *-automorphisms \(\gamma^q\) of \(U(G)\) such that \(\gamma^q \varphi^q \rho^q_i = \rho^q_i \varphi^q\). Indeed, then exactly as in the first part of the proof, a section \(q \mapsto \sigma^q_i(a \varphi^q)\) is continuous if and only if \(q \mapsto \sigma^q_i(a \gamma^q \varphi^q)\) is continuous.

\[
\sigma^q_i(a \gamma^q \varphi^q) = a \gamma^q \varphi^q \rho^q_i = a \rho^q_i \varphi^q,
\]

the latter section is indeed continuous by definition of the continuous field structure on the C*-algebras \(C(SU_q(2))\).

The automorphisms \(\gamma^q\) will be defined by a family of automorphisms \(\gamma^q_\lambda\) of \(B(V_\lambda)\). Fix \(\lambda \in P_+\). Let \(N \in \mathbb{N}\) be such that \(\omega(h_1) \leq N\) for every \(\omega \in P\) such that \(V_\lambda(\omega) \neq 0\). For every \(q > 0\) consider the direct sum \(\oplus_{\lambda \in N/2} V^q_\lambda\) of \(U_q \mathfrak{sl}_2\)-modules and the corresponding surjective homomorphism \(\alpha^q\) from \(U_q \mathfrak{sl}_2\) into the algebra \(B_q = \oplus_{\lambda \in N/2} B(V^q_\lambda)\). Then the representation \(\pi^q \rho^q : U_q \mathfrak{sl}_2 \to B(V^q)\) factors through \(B_q\), so \(\pi^q \rho^q = \beta^q \alpha^q\) for a unique \(\beta^q : B_q \to B(V^q)\). We summarize all the maps involved in the following diagram, which is commutative along solid lines:

\[
\begin{array}{ccc}
\mathcal{U}(G_q) & \xrightarrow{\pi^q} & B(V^q_\lambda) \xrightarrow{\varphi^q_\lambda} B(V_\lambda) \xleftarrow{\pi_\lambda} \mathcal{U}(G) \\
\mathcal{U}(SU_q(2)) & \xrightarrow{\rho^q} & B_q \\
\end{array}
\]

\[
\begin{array}{ccc}
B_q & \xrightarrow{\beta^q} & B \\
\mathcal{U}(SU_q(2)) & \xrightarrow{\theta^q} & \mathcal{U}(SU(2)) \\
\end{array}
\]
where \( \varphi^q_\lambda: B(V^q_\lambda) \to B(V_\lambda) \) and \( \tilde{\theta}^q: B_q \to B \) are the isomorphisms defined by the isomorphisms \( \varphi^q: \mathcal{U}(G_q) \to \mathcal{U}(G) \) and \( \theta^q: \mathcal{U}(SU_q(2)) \to \mathcal{U}(SU(2)) \), respectively. Consider the family of homomorphisms \( \varphi^q_\lambda \beta^q(\tilde{\theta}^q)^{-1}: B \to B(V_\lambda) \).

Since the families of homomorphisms \( \tilde{\theta}^q \alpha^q = \alpha \theta^q: U_q \rightarrow B \) and \( \varphi^q_\lambda \beta^q \alpha^q = \varphi^q_\lambda \pi^q_\lambda \rho^q_\lambda = \pi^q_\lambda \varphi^q_\lambda \rho^q_\lambda: U_q \rightarrow B(V_\lambda) \) are continuous in the sense defined earlier, and the homomorphisms in the first family are surjective, it follows that the homomorphisms \( \varphi^q_\lambda \beta^q(\tilde{\theta}^q)^{-1}: B \to B(V_\lambda) \) depend continuously on \( q \). Furthermore, for \( q = 1 \) we get the homomorphism \( \beta \). Hence, by a standard result on homomorphisms of finite dimensional \( C^* \)-algebras, we can choose a continuous family of \( * \)-automorphisms \( \gamma^q_\lambda \) of \( B(V_\lambda) \) such that \( \gamma^q_\lambda \varphi^q_\lambda \beta^q(\tilde{\theta}^q)^{-1} = \beta \) for all \( q > 0 \). In other words, if we replace \( \varphi^q_\lambda \) by \( \gamma^q_\lambda \varphi^q_\lambda \) in (1.1), we get a commutative diagram, which is what we need.

\[ \square \]

**Remark 1.3.** The space of sections of the form \( q \mapsto a \varphi^q \), \( a \in \mathbb{C}[G] \), is not an algebra and is not closed under involution. But the space of finite sums of sections of the form \( q \mapsto f(q) a \varphi^q \), where \( a \in \mathbb{C}[G] \) and \( f \) is a continuous function, is a \( * \)-algebra. Indeed, assume \( a \) is a matrix coefficient of a finite dimensional representation of \( G \) and \( \{ a_1 \} \), \( 1 \) is a basis in the space spanned by the matrix coefficients of the contragradient representation. Then \( (a \varphi^q)^* = \sum f_i(q) a_i \varphi^q \) for uniquely defined functions \( f_i \). Since \( q \mapsto (a \varphi^q)^* \) is a continuous section, the functions \( f_i \) must be continuous. Therefore the space is closed under involution. Similarly we check that the space is closed under multiplication. This of course can also be checked without relying on the above theorem. Note also that this space does not depend on the choice of \( \varphi^q \).

For \( b > a > 0 \) denote by \( C(G_{[a,b]} \) the \( C^* \)-algebra of continuous sections of the field \( (C(G_q))_q \in [a,b] \).

**Theorem 1.4 ([9]).** For any \( b > a > 0 \) and \( q \in [a,b] \) the evaluation map \( ev_q: C(G_{[a,b]} \to C(G_q) \) is a KK-equivalence.

Since this is not exactly how the result is formulated in [9, Corollaries 3.8 and 3.11], some comments are in order. What is proved in [9], is that \( C(SU_q(N)) \) is KK-equivalent to \( C(SU(N)) \) for \( q \in (0,1) \), and it is mentioned that the same can be proved for the other classical simple compact Lie groups. The proof, however, shows that the above more precise result holds for \( b \leq 1 \), and once the family of \( C^* \)-algebras \( C(G_q) \) is given a continuous field structure as described above, the general case is essentially identical to \( G = SU(N) \), see [13, Section 6] for details. To deal with the case \( b > 1 \) we can argue as follows. It is easy to see that it suffices to prove the theorem for the evaluations at the end points, see the proof of [13, Lemma 6.3]. We therefore have to show that the \( C^* \)-algebras \( I_{a,b} \) (resp. \( J_{a,b} \)) of continuous sections of \( (C(G_q))_q \in [a,b] \) vanishing at \( a \) (resp., at \( b \)) are KK-contractible, knowing already that this is true when \( b \leq 1 \). If \( a \geq 1 \), this follows from the canonical isomorphisms \( G_q \cong G_{q-1} \). If \( a < 1 < b \) then the KK-contractibility of \( I_{a,b} \) and \( J_{a,b} \) follows from the exact sequences \( 0 \to I_{1,b} \to I_{a,b} \to I_{a,1} \to 0 \) and \( 0 \to J_{a,1} \to J_{a,b} \to J_{1,b} \to 0 \).
2. Continuous family of Drinfeld twists

As in the previous section, fix a continuous family of $*$-isomorphisms $\varphi^q : U(G_q) \to U(G)$ with $\varphi^1 = \iota$.

For $q > 0$, let $h_q \in i\mathbb{R}$ be such that $q = e^{\pi i h_q}$. Denote by $t \in g \otimes g$ the $g$-invariant element defined by the ad-invariant symmetric form on $g$ such that the induced form on $h^*$ satisfies $(\alpha_i, \alpha_j) = d_{ij}$. By a result of Kazhdan and Lusztig [5], see [11] for details, for every $q > 0$ there exists a unitary element $F^q \in U(G \times G)$, which we call a unitary Drinfeld twist, such that

(i) $(\varphi^q \otimes \varphi^q)(\hat{\Delta} q) = F^q \hat{\Delta} q \varphi^q(\cdot) F^q *$

(ii) $(\varphi^q \otimes \varphi^q)(R_q) = F^q_{21} q^1 R^q_{q2} F^q_{q2}$, where $R_q \in U(G_q \times G_q)$ is the universal $R$-matrix;

(iii) $(\varphi^q \otimes \varphi^q)(\hat{\Delta} q) = F^q_{21} q^1 \hat{\Delta} q \varphi^q(\cdot) F^q_{q2} = \Phi_{KZ}(h_{q1}, h_{q2})$, where $\Phi_{KZ}$ is Drinfeld’s KZ-associator.

Such an element is not unique, but by [12] any other unitary Drinfeld twist (for the same isomorphism $\varphi^q$) has the form $(c \otimes c) F^q \hat{\Delta} (c) ^*$ for a unitary element $c$ in the center of $U(G)$.

We say that a family $\{F^q\}_{q > 0}$ of unitary Drinfeld twists is continuous, if the map

$q \mapsto F^q \in W^*(G \times G)$

is continuous in the strong operator topology on the von Neumann algebra $W^*(G \times G) \subset U(G \times G)$ of $G \times G$. In other words, the map $q \mapsto (\pi_\lambda \otimes \pi_\mu)(F^q)$ is continuous for all $\lambda, \mu \in P_+$.

**Theorem 2.1.** There exists a continuous family of unitary Drinfeld twists $F^q$ such that $F^1 = 1$. Furthermore, if $\{\psi^q : U(G_q) \to U(G)\}_{q > 0}$ is another continuous family of $*$-isomorphisms such that $\psi^1 = \iota$, and $\{E^q\}_{q > 0}$ is a corresponding continuous family of unitary Drinfeld twists with $E^1 = 1$, then there exists a unique continuous family of unitary elements $u^q \in U(G)$ such that

$u^1 = 1$, and $\psi^q = u^q \varphi^q(\cdot) u^q *$ and $E^q = (u^q \otimes u^q) F^q \hat{\Delta} (u^q) ^*$ for all $q > 0$.

**Proof.** To prove existence, consider the set $\Omega$ of pairs $(q, F)$, where $q > 0$ and $F$ is a unitary Drinfeld twist for $\varphi^q$. It is a closed subset of the direct product of $\mathbb{R}_+$ and the unitary group of $W^*(G \times G)$ (this is used already in the proof of [11, Lemma 3.2]), so it is a locally compact space. Let $p : \Omega \to \mathbb{R}_+$ be the projection onto the first coordinate. The compact abelian group of elements of the form $(c \otimes c) \hat{\Delta} (c)^*$, where $c$ is a unitary element in the center of $U(G)$, acts freely by multiplication on the right on $\Omega$, and by [12, Theorem 5.2] this action is transitive on each fiber of the map $p$. Therefore if this group were a compact Lie group, then by a theorem of Gleason [4], $p : \Omega \to \mathbb{R}_+$ would be a fiber bundle, hence $p$ would have a continuous section. Since the group of elements of the form $(c \otimes c) \hat{\Delta} (c)^*$ is not a Lie group, we cannot apply Gleason’s theorem directly and will proceed as follows.

*Documenta Mathematica 16 (2011) 767–780*
Choose an increasing sequence of finite subsets $P_n \subset P_+$ such that $P_1 = \{0\}$ and $\cup_n P_n = P_+$. For every $q > 0$ we will construct a sequence of unitary Drinfeld twists $\mathcal{F}_n^q$ such that $\mathcal{F}_n^q = 1$, the map $q \mapsto (\pi_\lambda \otimes \pi_\nu)(\mathcal{F}_n^q)$ is continuous for all $\lambda, \nu \in P_n$ and $n \geq 1$, and $(\pi_\lambda \otimes \pi_\nu)(\mathcal{F}_n^q) = (\pi_\lambda \otimes \pi_\nu)(\mathcal{F}_n^0)$ for all $\lambda, \nu \in P_n$ and $n \geq 1$. Then, for every $q > 0$, the sequence $\{\mathcal{F}_n^q\}_n$ converges to a unitary Drinfeld twist $\mathcal{F}^q$ with the required properties.

For $n = 1$ and $q \neq 1$ we take $\mathcal{F}_1^q$ to be any unitary Drinfeld twist, and we take $\mathcal{F}_1^1 = 1$.

Assume the Drinfeld twists $\mathcal{F}_n^q$ are already constructed for some $n$. Denote by $\Omega_{n+1}$ the set of pairs $(q, U)$, where $U = (U_{\lambda, \nu})_{\lambda, \nu}$ is a unitary element in $\prod_{(\lambda, \nu) \in P_{n+1} \times P_{n+1}} B(V_\lambda \otimes V_\nu)$ such that there exists a unitary Drinfeld twist $\mathcal{F}$ for $\varphi^q$ satisfying

$$(\pi_\lambda \otimes \pi_\nu)(\mathcal{F}) = U_{\lambda, \nu} \quad \text{for all} \quad (\lambda, \nu) \in P_{n+1} \times P_{n+1} \backslash P_n \times P_n,$$

$$(\pi_\lambda \otimes \pi_\nu)(\mathcal{F}) = (\pi_\lambda \otimes \pi_\nu)(\mathcal{F}_n^q) \quad \text{for all} \quad \lambda, \nu \in P_n.

Let $p_{n+1} : \Omega_{n+1} \to \mathbb{R}_+^*$ be the projection onto the first coordinate. The set $\Omega_{n+1}$ is a closed subset of the direct product of $\mathbb{R}_+^*$ and the unitary group of $\prod_{(\lambda, \nu) \in P_{n+1} \times P_{n+1}} p_{n+1} B(V_\lambda \otimes V_\nu)$. For every $q > 0$ the fiber $p_{n+1}^{-1}(q)$ is nonempty, as it contains the element $((\pi_\lambda \otimes \pi_\nu)(\mathcal{F}_n^q))_{\lambda, \nu}$.

Let $S_{n+1}$ be the set of weights $\lambda \in P_+$ such that either $\lambda \in P_{n+1}$, or $V_\lambda$ is equivalent to a subrepresentation of $V_\nu \otimes V_\eta$ for some $\nu, \eta \in P_{n+1}$, in which case we write $V_\lambda \prec V_\nu \otimes V_\eta$. Let $K_{n+1} = \prod_{\lambda \in S_{n+1}} T$. We have a homomorphism $p_{n+1}$ from $K_{n+1}$ to the unitary group of $\prod_{(\lambda, \nu) \in P_{n+1} \times P_{n+1}} p_{n+1} B(V_\lambda \otimes V_\nu)$: $p_{n+1}(c)$ acts on the isotypic component of $V_\nu \otimes V_\eta$ of type $V_\lambda$ as the multiplication by $c_{\nu, \eta} \mathcal{F}_n^q$. We also have a similar homomorphism $\theta_{n+1}$ from $K_{n+1}$ into the unitary group of $\prod_{(\lambda, \nu) \in P_+} B(V_\lambda \otimes V_\nu)$.

The group $\ker \theta_{n+1}$ acts on $\Omega_{n+1}$ by multiplication by $\rho_{n+1}(c)$ on the right. On every fiber of $p_{n+1}$ this action is transitive, and the stabilizer of every point is $\ker \rho_{n+1} \cap \ker \theta_{n+1}$. Since $\ker \theta_{n+1}$ is a compact Lie group, by Gleason's theorem we conclude that $p_{n+1} : \Omega_{n+1} \to \mathbb{R}_+^*$ is a fiber bundle, hence it is a trivial bundle. Choosing a continuous section of this bundle, by definition of $\Omega_{n+1}$ we conclude that there exist unitary Drinfeld twists $\mathcal{E}^q$ such that the map $q \mapsto (\pi_\lambda \otimes \pi_\nu)(\mathcal{E}^q)$ is continuous for all $\lambda, \nu \in P_{n+1}$ and $(\pi_\lambda \otimes \pi_\nu)(\mathcal{E}^q) = (\pi_\lambda \otimes \pi_\nu)(\mathcal{E}^0)$ for all $\lambda, \nu \in P_n$. There exists a unitary central element $c \in \mathcal{U}(G)$ such that $\mathcal{E}^1 = (c^* \otimes c^*) \Delta(c)$. We can then set $\mathcal{F}_{n+1}^q = \mathcal{E}^q(c \otimes c) \Delta(c)^*$.

This finishes the proof of existence.

Assume now that $\{\psi^q : \mathcal{U}(G_\lambda) \to \mathcal{U}(G)\}_{q \geq 0}$ is another continuous family of $*$-isomorphisms such that $\psi^q = \iota$, and $\{\mathcal{E}^q\}_{q \geq 0}$ is a corresponding continuous family of unitary Drinfeld twists with $\mathcal{E}^1 = 1$. For every $\lambda \in P_+$, let $\varphi^q_\lambda, \psi^q_\lambda : B(V_\lambda^q) \to B(V_\lambda)$ be the isomorphisms defined by $\varphi^q$ and $\psi^q$. The set of unitaries $v \in B(V_\lambda)$ such that $\psi^q_\lambda = v \varphi^q_\lambda(\iota) v^*$ forms a circle bundle over $\mathbb{R}_+^*$, so it has a continuous section $v^q_\lambda$. Since $\psi^q_\lambda = \varphi^q_\lambda = \iota$, we may assume that $v^q_\lambda = 1$. The unitaries $v^q_\lambda$ define a continuous family of unitaries $v^q \in \mathcal{U}(G)$.
For every $q > 0$, the element $(v^q \otimes v^q)\mathcal{F}q\bar{\Delta}(v^q)^*$ is a unitary Drinfeld twist for $\psi^q$. Hence, for every $q$, there exists a unitary central element $c \in \mathcal{U}(G)$ such that
\begin{equation}
\mathcal{E}^q = (v^q \otimes v^q)\mathcal{F}q\bar{\Delta}(v^q)^*(c \otimes c)\bar{\Delta}(c)^*.
\end{equation}
Furthermore, the element $c$ is defined up to a group-like unitary element in the center of $\mathcal{U}(G)$, that is, up to an element of the center $Z(G)$ of $G$. Therefore, applying once again Gleason’s theorem (which in this case is quite obvious as $Z(G)$ is finite), we see that the set of pairs $(q, c)$ with $c$ satisfying (2.1) is a principal $Z(G)$-bundle over $\mathbb{R}_+^\times$, hence it has a continuous section $q \mapsto (q, c^q)$. The element $c^1$ is group-like, so replacing $c^q$ by $c^q c^1^{-1}$ we may assume that $c^1 = 1$. Letting $u^q = c^q v^q$, we get the required continuous family of unitary elements.

Finally, if $\tilde{u}^q$ is another continuous family of unitary elements with the same properties, then $c^q = \tilde{u}^q u^q$ is a unitary central group-like element in $\mathcal{U}(G)$, hence $c^q \in Z(G)$. Since $c^q$ depends continuously on $q$, $Z(G)$ is finite and $c^1 = 1$, we conclude that $c^q = 1$ for all $q$. \hfill \Box

Another way of formulating the above result is to say that the set of triples $(q, \varphi, \mathcal{F})$ such that $q > 0$, $\varphi: \mathcal{U}(G_q) \to \mathcal{U}(G)$ is a $*$-isomorphism extending the canonical identification of the centers and $\mathcal{F}$ is a unitary Drinfeld twist for $\varphi$, has a structure of a principal $U(W^*(G))/Z(G)$-bundle over $\mathbb{R}_+^\times$, where $U(W^*(G))$ is the unitary group of $W^*(G)$.

Note also that by analyzing the proof of Kazhdan and Lusztig [6] one could hope to prove a stronger result: the family of unitary Drinfeld twists $\mathcal{F}q$ can be chosen to be real-analytic for an appropriate choice of $\varphi^q$.

3. FAMILY OF DIRAC OPERATORS

We continue by fixing a continuous family of $*$-isomorphisms $\varphi^q: \mathcal{U}(G_q) \to \mathcal{U}(G)$ with $\varphi^1 = \iota$ and a continuous family of unitary Drinfeld twists $\mathcal{F}q$ for $\varphi^q$ with $\mathcal{F}^1 = 1$.

For every $q > 0$ we have a Dirac operator $D_q$ on $G_q$ defined in [10] as follows. Consider a basis $\{x_i\}_i$ of $\mathfrak{g}$ such that $(x_i, x_j) = -\delta_{ij}$, and let $\gamma: \mathfrak{g} \to \text{Cl}(\mathfrak{g})$ denote the inclusion of $\mathfrak{g}$ into the complex Clifford algebra with the convention that $\gamma(x_i)^2 = -1$. Identifying $so(\mathfrak{g})$ with $\text{spin}(\mathfrak{g})$, the adjoint action is defined by the representation $\text{ad}: \mathfrak{g} \to \text{spin}(\mathfrak{g}) \subset \text{Cl}(\mathfrak{g})$ given by
\begin{equation}
\text{ad}(x) = \frac{1}{4} \sum_i \gamma(x_i)\gamma([x, x_i]).
\end{equation}

We denote by the same symbol $\tilde{\text{ad}}$ the corresponding homomorphism $\mathcal{U}(G) \to \text{Cl}(\mathfrak{g})$.

Let $s: \text{Cl}(\mathfrak{g}) \to \text{End}(\mathbb{S})$ be an irreducible representation. Denote by $\partial$ the representation of $U\mathfrak{g}$ by left-invariant differential operators. Identifying the sections $\Gamma(S)$ of the spin bundle $S$ over $G$ with $C^{\infty}(G) \otimes \mathbb{S}$, the Dirac operator...
Let \((L^2(G_q), \pi_{r,q}, \xi^q_h)\) be the GNS-triple defined by the Haar state on \(C(G_q)\). The right regular representation of the von Neumann algebra \(W^*(G_q) \subset U(G_q)\) of \(G_q\) on \(L^2(G_q)\), denoted by \(\hat{\pi}_{r,q}\), is defined by

\[
\hat{\pi}_{r,q}(a)\xi^q_h = (\pi_{r,q} \otimes \omega) \Delta_q(a)\xi^q_h = a(1)(\omega)\pi_{r,q}(a(0))\xi^q_h.
\]

The Dirac operator \(D_q\) on \(G_q\) is the unbounded operator on \(L^2(G_q) \otimes \mathbb{S}\) defined by

\[
D_q = (\partial_q \otimes s)(\mathcal{D}_q),
\]

where \(\mathcal{D}_q \in U(G_q) \otimes \text{Cl}(g)\) is given by

\[
\mathcal{D}_q = (\varphi^q \otimes \iota)^{-1}(\iota \otimes \text{ad})(\mathcal{F}^q)\mathcal{D}(\iota \otimes \text{ad})(\mathcal{F}^q)^*.
\]

Our goal is to show that the family \(\{D_q\}_q\) is continuous in the sense that it defines a Kasparov \((C(G_{q,a,b}), C[a,b])\)-module.

**Lemma 3.1.** The family \((L^2(G_q))_{q>0}\) has a unique structure of a continuous field of Hilbert spaces such that the vector field \(q \mapsto \pi_{r,q}(a^q)\xi^q_h\) is continuous for every continuous section \(a^q\) of the field \((C(G_q))_{q>0}\).

**Proof.** It suffices to show that the function \(q \mapsto (\pi_{r,q}(a^q)\xi^q_h, \xi^q_h)\) is continuous. But this is clear, since if \(a^q = a\varphi^q\), where \(a\) is a matrix coefficient of a nontrivial irreducible representation of \(G\), then the function is zero by the orthogonality relations. \(\square\)

It is easy to see that the continuous field \((L^2(G_q))_{q>0}\) is trivial. To formulate a more precise result, recall the exact form of the orthogonality relations.

First let us introduce some notation. For a weight \(\beta = \sum_c c_i \alpha_i\) (\(c_i \in \mathbb{C}\)) put \(h^q_\beta = \sum_c c_i d_i h^q_i \in U(G_q)\) and \(K^q_\lambda = q^h^q_\beta\); note that for \(q = 1\) the element \(h_{\beta}\) is characterized by \(\lambda(h_{\beta}) = (\lambda, \beta)\) for any weight \(\lambda\), and for \(q \neq 1\) we have \(K^q_\lambda = K^q_\lambda\). Let \(\rho\) be half the sum of positive roots. Then \(K^q_\rho\) is the Woronowicz character \(f_\rho\) for \(G_q\), in particular, for the square of the antipode \(\tilde{S}_q\) on \(U(G_q)\) we have \(\tilde{S}_q(\omega) = K^q_{2\rho, \omega}K^q_{2\rho}\). Put \(\text{dim}_q(V^q_\lambda) = \text{Tr}(\pi^q_\lambda(K^q_{2\rho})) = \text{Tr}(\pi^q_\lambda(K^q_{2\rho} \omega))\).

For \(\xi, \zeta \in V^q_\lambda\) define \(a_{\xi, \zeta}^q \in \mathbb{C}[G_q]\) by

\[
a_{\xi, \zeta}^q(\omega) = (\pi^q_\lambda(\omega)\zeta, \xi) \quad \text{for} \quad \omega \in U(G_q).
\]

Then the orthogonality relations state that the vectors \(\pi_{r,q}(a_{\xi, \zeta}^q)\xi^q_h\) are mutually orthogonal for different \(\lambda\), and

\[
(\pi_{r,q}(a_{\xi, \zeta}^q)\xi^q_h, \pi_{r,q}(a_{\zeta', \zeta'}^q)\xi^q_h) = \frac{1}{\text{dim}_q(V^q_\lambda)}(\pi^q_\lambda(K^q_{2\rho}))_{\xi, \zeta'}(\zeta, \zeta').
\]
Let $d^q \in \mathcal{U}(G_q)$ be the element such that
\[
\pi^q(\lambda^{1/2}) = \frac{\dim(V_{\lambda}^q)}{\dim(V_{\lambda}^{q^2})^{1/2}} \pi^q_{K^q}(K^q_{\lambda}) \text{ for all } \lambda \in P_+.
\]

**Lemma 3.2.** The linear operator $W_q : \pi_{r,q}(\mathbb{C}[G_q])\xi_h \rightarrow \pi_{r}(\mathbb{C}[G])\xi_h$ defined by
\[
W_q \pi_{r,q}(a\varphi^q)\xi_h = a(0)(\varphi^q(d^q)) \pi_{r}(a(1))\xi_h \text{ for } a \in \mathbb{C}[G]
\]
is unitary. It has the property
\[
W_q \partial_q(\omega) = \partial(\varphi^q(\omega)) W_q \text{ for all } \omega \in \mathcal{U}(G_q).
\]

**Proof.** By the orthogonality relations we have a decomposition $L^2(G_q) = \bigoplus_{\lambda \in P_+} V_{\lambda}^q \otimes V_{\lambda}^q$, with $\pi_{r,q}(a^{\lambda,q}_{\lambda})(\xi_{\lambda}) = \xi_{\lambda}$ corresponding to $\dim_{q}(V_{\lambda}^q)^{-1/2} \pi_{r,q}(K^q_{\lambda}) \otimes \xi \in V_{\lambda}^q \otimes V_{\lambda}^q$. The isomorphism $\varphi^q : \mathcal{U}(G_q) \rightarrow \mathcal{U}(G)$ is implemented by unitaries $\xi_{\lambda}: V_{\lambda}^q \rightarrow V_{\lambda}$. Then the unitaries
\[
\overline{u}_{\lambda}^q \otimes \xi_{\lambda}^q: V_{\lambda}^q \otimes V_{\lambda}^q \rightarrow V_{\lambda} \otimes V_{\lambda}
\]
define a unitary $L^2(G_q) \rightarrow L^2(G)$. This is exactly the unitary $W_q$, since if $a = a^{\lambda}_{\lambda} \xi_{\lambda} \in \mathbb{C}[G]$ then $a \varphi^q = a^{\lambda,q}_{\lambda} \pi_{r,q}(\varphi^q)\xi_h$ and hence the vector $\pi_{r,q}(a\varphi^q)\xi_h$ is mapped onto
\[
\frac{\dim(V_{\lambda}^q)^{1/2}}{\dim(V_{\lambda}^{q^2})^{1/2}} \pi_{r}(a^{\lambda,q}_{\lambda}(K^q_{\lambda})\pi_{r,q}(\varphi^q)\xi_h),
\]
which gives the formula in the formulation, as
\[
da^{\lambda,q}_{\lambda}(K^q_{\lambda})\pi_{r,q}(\varphi^q)\xi_h = (a^{\lambda,q}_{\lambda}(\cdot)\pi_{r,q}(\varphi^q(\lambda^{1/2})))\xi_h = a(0)(\varphi^q(K^q_{\lambda}))\xi_h
\]

The last statement in the formulation follows either by a direct computation or by observing that $\partial_q(\omega)$ acts on $V_{\lambda}^q \otimes V_{\lambda}^q \subset L^2(G_q)$ as $1 \otimes \pi^q_{K^q}(\omega)$. \hfill \Box

The extension of $W_q$ to $L^2(G_q)$ we continue to denote by the same symbol. Therefore the unitaries $W_q$ define an isomorphism of the continuous field $(L^2(G_q)_{q>0})$ onto the trivial field with fiber $L^2(G)$. For $b > a > 0$ denote by $L^2(G_{[a,b]})$ the right Hilbert $C[a,b]$-module of continuous sections of $(L^2(G_q))_{q \in [a,b]}$. The $C^*$-algebra $C(G_{[a,b]})$ acts on $L^2(G_{[a,b]})$ via GNS-representations $\pi_{r,q}$; we denote by $\pi$ the corresponding homomorphism from $C(G_{[a,b]})$ into the algebra of adjointable operators on $L^2(G_{[a,b]})$. The operators $D_q$ define an unbounded operator $D_{[a,b]}$ on $L^2(G_{[a,b]}) \otimes \mathbb{S}$ with domain of definition consisting of continuous vector fields $\xi$ such that $\xi(q) \in \text{Dom}(D_q)$ for all $q$ and the vector field $q \mapsto D_q \xi(q)$ is continuous. When $G$ is even-dimensional we also define a grading on $L^2(G_{[a,b]}) \otimes \mathbb{S}$ using the chirality element $\chi \in C(G)$. 

**Theorem 3.3.** For any $b > a > 0$ the triple $(L^2(G_{[a,b]}), \pi(\cdot) \otimes 1, D_{[a,b]})$ is an unbounded Kasparov $(C(G_{[a,b]}), C[a,b])$-module of the same parity as the dimension of $G$.

**Proof.** By definition of an unbounded Kasparov module [1] it suffices to check that

---

*Documenta Mathematica 16 (2011) 767–780*
(i) $D_{[a,b]}$ is a regular self-adjoint operator such that $(1 + D_{[a,b]}^2)^{-1}$ is generalized compact (recall that regularity means that the operator $1 + D_{[a,b]}^2$ is surjective);
(ii) there exists a dense $*$-subalgebra $A$ of $C(G_{[a,b]})$ and an $A$-invariant core of $D_{[a,b]}$ such that the commutators $[D_{[a,b]}, \pi(c) \otimes 1]$ are bounded on this core for all $c \in A$.

By definition of $D_q$ and Lemma 3.2 we have
\begin{equation}
(\partial \otimes \text{ad})(\mathcal{F}^q)^*(W_q \otimes 1)D_q = D(\partial \otimes \text{ad})(\mathcal{F}^q)^*(W_q \otimes 1).
\end{equation}
Therefore the unitaries $(\partial \otimes \text{ad})(\mathcal{F}^q)^*(W_q \otimes 1)$ define an isomorphism of the continuous field $(L^2(G_q) \otimes \mathbb{S})_q$ onto the constant field with fiber $L^2(G) \otimes \mathbb{S}$, which maps $D_{[a,b]}$ onto the operator which acts as $D$ on every fiber. This immediately gives (i).

To prove (ii) consider the space $\mathcal{A} \subset C(G_{[a,b]})$ of finite sums of sections of the form $q \mapsto f(q)c\varphi^t$ with $f$ a continuous function and $c \in \mathbb{C}[G]$. By Remark 1.3 this is a $*$-algebra. It is dense in $C(G_{[a,b]})$ since it is dense in every fiber and is closed under multiplication by continuous functions. The linear span of vectors of the form $q \mapsto \pi_{r,q}(c^q)\xi_h \otimes \zeta$, where $c \in \mathcal{A}$ and $\zeta \in \mathbb{S}$, is an $\mathcal{A}$-invariant core for $D_{[a,b]}$. This follows e.g. from (3.1) and the fact that vectors of the form $\pi_r(c)\xi_h \otimes \zeta$, where $c \in \mathbb{C}[G]$ and $\zeta \in \mathbb{S}$, span a core for $D$ which is a union of $D$-invariant finite dimensional subspaces.

It remains to check boundedness of commutators. For this it is enough to show that for every $c \in \mathbb{C}[G]$ the commutators $[D_q, \pi_{r,q}(c\varphi^q) \otimes 1]$ are uniformly bounded on $[a,b]$. By [10, Proposition 3.1] we have
\[ [D_q, \pi_{r,q}(c\varphi^q) \otimes 1] = -(\pi_{r,q}(c^q\varphi^q) \otimes 1)(\partial_q \circ (\varphi^q)^{-1} \otimes s)(c(1) \otimes t \otimes t)(U_qT_qU_q^*) \]
where
\[ U_q = (t \otimes t \otimes \text{ad})(\mathcal{F}^q \otimes 1)(\hat{\Lambda} \otimes \iota)(\mathcal{F}^q) \] and $T_q \in U(G \times G) \otimes \text{Cl}(\mathfrak{g})$ is defined by
\[ T_q = (t \otimes t \otimes \gamma)(t_{13}) + (t \otimes t \otimes \gamma)(t_{23}) - (t \otimes t \otimes \text{ad})(\Phi^q)(t \otimes t \otimes \gamma)(t_{23})(t \otimes t \otimes \text{ad})(\Phi^q), \]
with $\Phi^q = \Phi_{KZ}(h_qt_{12}, h_qt_{23})$. It is therefore enough to prove that the operators
\[ (c \otimes t \otimes t)(U_qT_qU_q^*) \] are uniformly bounded on $[a,b]$ for all $c \in \mathbb{C}[G]$. Equivalently, the operators $(\pi \otimes t \otimes t)(T_q)$ are uniformly bounded for any finite dimensional unitary representation $\pi$ of $G$, which, in turn, is the same as uniform boundedness of
\[ [(\pi \otimes t \otimes \gamma)(t_{23}), (\pi \otimes t \otimes \text{ad})(\Phi_{KZ}(h_qt_{12}, h_qt_{23}))]. \]
By [10, Proposition 3.6] the latter property indeed holds: the norm of the above commutator is bounded by $6||[(\pi \otimes \gamma))(t)||$ independently of $q > 0$. □

Recall that by Theorem 1.4 the evaluation map $ev_q : C(G_{[a,b]}) \rightarrow C(G_q)$ is a KK-equivalence for any $q \in [a,b]$. For $q, q' \in [a,b]$ define an invertible element $\gamma_{q,q'}$ in $KK(C(G_q), C(G_{q'}))$ by $\gamma_{q,q'} = [ev_{q'}]^{-1}[ev_q]$. It does not depend
on the segment \([a, b]\) containing \(q\) and \(q'\). Denote also by \([D_q]\) the element of \(KK_i(C(G,q), \mathbb{C})\) defined by \(D_q\), where \(i = \dim G \pmod 2\).

**Corollary 3.4.** For any \(q, q' > 0\) we have \(\gamma_{q,q'}[D_{q'}] = [D_q]\).

**Proof.** Denote by \([D_{[a,b]}]\) the element of \(KK_i(C(G_{[a,b]}), C_{[a,b]})\) the class of the Kasparov module

\[
(L^2(G_{[a,b]}) \otimes \mathcal{S}, \pi(\cdot) \otimes 1, D_{[a,b]}).
\]

If \(\tilde{ev}_q : C_{[a,b]} \to \mathbb{C}\) is the evaluation at \(q\) then clearly \([D_{[a,b]}][\tilde{ev}_q] = [ev_q][D_q]\).

Since \([\tilde{ev}_q]\) does not depend on \(q\), we thus see that the class \([ev_q][D_q] \in KK_i(C(G_{[a,b]}), \mathbb{C})\) does not depend on \(q \in [a, b]\) either, which is what we need. \(\square\)

**References**

[1] S. Baaj and P. Julg, *Théorie bivariante de Kasparov et opérateurs non bornés dans les C∗-modules hilbertiens*, C. R. Acad. Sci. Paris Sér. I Math. 296 (1983), 875–878.

[2] A. Bauval, *Quantum group- and Poisson-deformation of \(SU(2)\)*, in: “Recent advances in operator algebras” (Orléans, 1992), Astérisque No. 232 (1995), 49–65.

[3] É. Blanchard, *Déformations de C∗-algèbres de Hopf*, Bull. Soc. Math. France 124 (1996), 141–215.

[4] A.M. Gleason, *Spaces with a compact Lie group of transformations*, Proc. Amer. Math. Soc. 1 (1950), 35–43.

[5] D. Kazhdan and G. Lusztig, *Tensor structures arising from affine Lie algebras. III*, J. Amer. Math. Soc. 7 (1994), 335–381.

[6] D. Kazhdan and G. Lusztig, *Tensor structures arising from affine Lie algebras. IV*, J. Amer. Math. Soc. 7 (1994), 383–453.

[7] E. Kirchberg and S. Wassermann, *Operations on continuous bundles of C∗-algebras*, Math. Ann. 303 (1995), 677–697.

[8] L.I. Korogodski and Ya.S. Soibelman, *Algebras of functions on quantum groups. Part I*, Mathematical Surveys and Monographs, 56, American Mathematical Society, Providence, RI, 1998.

[9] G. Nagy, *Deformation quantization and K-theory*, in: “Perspectives on quantization” (South Hadley, MA, 1996), 111–134, Contemp. Math., 214, Amer. Math. Soc., Providence, RI, 1998.

[10] S. Neshveyev and L. Tuset, *The Dirac operator on compact quantum groups*, J. Reine Angew. Math. 641 (2010), 1–20.

[11] S. Neshveyev and L. Tuset, *Notes on the Kazhdan-Lusztig theorem on equivalence of the Drinfeld category and the category of \(U_q(g)\)-modules*, Algebr. Represent. Theory 14 (2011), no. 5, 897–948.

[12] S. Neshveyev and L. Tuset, *Symmetric invariant cochains on the duals of \(q\)-deformations*, Adv. Math. 227 (2011), no. 1, 146–169.
[13] S. Neshveyev and L. Tuset, *Quantized algebras of functions on homogeneous spaces with Poisson stabilizers*, preprint arXiv:1103.4346v1 [math.OA].

[14] L.L. Vaksman and Ya.S. Soibelman, *An algebra of functions on the quantum group SU(2)*, Funct. Anal. Appl. 22 (1988), no. 3, 170–181 (1989).

[15] M. Yamashita, *Connes-Landi deformation of spectral triples*, Lett. Math. Phys. 94 (2010), no. 3, 263–291.
Dilation Theory, Commutant Lifting and Semicrossed Products

Kenneth R. Davidson\textsuperscript{1} and Elias G. Katsoulis\textsuperscript{2}

Received: July 14, 2011
Revised: September 16, 2011
Communicated by Joachim Cuntz

Abstract. We take a new look at dilation theory for nonself-adjoint operator algebras. Among the extremal (co)extensions of a representation, there is a special property of being fully extremal. This allows a refinement of some of the classical notions which are important when one moves away from standard examples. We show that many algebras including graph algebras and tensor algebras of C*-correspondences have the semi-Dirichlet property which collapses these notions and explains why they have a better dilation theory. This leads to variations of the notions of commutant lifting and Ando’s theorem. This is applied to the study of semicrossed products by automorphisms, and endomorphisms which lift to the C*-envelope. In particular, we obtain several general theorems which allow one to conclude that semicrossed products of an operator algebra naturally imbed completely isometrically into the semicrossed product of its C*-envelope, and the C*-envelopes of these two algebras are the same.

2010 Mathematics Subject Classification: 47L55, 47L40, 46L05, 37B20

Keywords and Phrases: dilation, extremal coextension, semicrossed product, commutant lifting, Fuglede property, C*-envelopes, Ando’s theorem

\textsuperscript{1}First author partially supported by an NSERC grant.
\textsuperscript{2}Second author was partially supported by a grant from ECU.
1 Introduction

This paper is a general study of dilation theory for arbitrary nonself-adjoint operator algebras. It began with an attempt to formalize those properties need to obtain dilation theorems for covariant representations of an operator algebra and an endomorphism, in order to understand the semicrossed product and its C*-envelope. In this paper, we discuss versions of commutant lifting and Ando’s theorem and consider when they allow us to determine the structure of a semicrossed product and its C*-envelope. This forced us to revisit basic notions in dilation theory, and to introduce a notion stronger than that of extremal (co)extensions. We feel that certain notions in dilation theory are too closely modelled on what happens for the disk algebra. This algebra has been shown to have many very strong properties, and they are often not perfectly reflected in the general case. Certain refinements should be considered to clarify the various dilation properties in a general context.

Dilation theory. Dilation theory for a single operator has its roots in the seminal work of Sz.Nagy [57] which is developed in the now classical book that he wrote with Foiaş [58]. Dilation theory for more general operators was initiated by the deep work of Arveson [5, 6]. The ideas have evolved over the past six decades. The basic ideas are nicely developed in Paulsen’s book [45].

In formulating general properties related to commutant lifting and Ando’s theorem, we were strongly motivated, in part, by the general module formulation expounded by Douglas and Paulsen [26] and the important study by Muhly

References

782

Documenta Mathematica 16 (2011) 781–868
Dilation Theory. . .

and Solel [40]. The language used there is a module theoretic approach, while we will mostly talk about representations instead. But the general constructs can, of course, be formulated in either language. Douglas and Paulsen focus on Shilov modules as a primary building block. Muhly and Solel adopt this view, but focus more on a somewhat stronger property of orthoprojective modules. They may have gone further, as we do, had they known what we do today. We will argue that these are more central to the theory.

Another important influence is the Dritschel–McCullough [27] proof of the existence of Arveson’s C*-envelope [5, 6], first established using different methods by Hamana [31]. They provide a proof strongly influenced by ideas of Agler [1]. What they show is: given a completely contractive representation of a unital operator algebra \( \mathcal{A} \), that among all dilations of this representation, there are always certain representations which are maximal in the sense that any further dilation can only be obtained by appending a direct sum of another representation. These dilations always exist, as they show, and they are precisely those representations which extend to a \( * \)-representation of the C*-envelope. It is in this manner that they establish that the existence of the C*-envelope.

This fact was anticipated by Muhly and Solel in [41] where they show, assuming Hamana’s theorem, that every representation has a dilation which is both orthoprojective and orthoinjective. It is easy to see that this is a reformulation of the maximal dilation property. Indeed, one can see that a representation \( \rho \) is orthoprojective if and only if it is maximal as a coextension (called an extremal coextension)—meaning that any coextension can be obtained only by appending a direct sum of another representation. Dritschel and McCullough proved that these exist as well. The dual version shows that orthoinjective representations coincide with the extremal extensions.

An extremal (co)extension of a representation \( \rho \) on \( \mathcal{H} \) is called minimal provided that the whole space is the smallest reducing subspace containing \( \mathcal{H} \).

This is a weaker notion than saying that \( \mathcal{H} \) is cyclic. However, there can be many extremal coextensions which are minimal but \( \mathcal{H} \) is not cyclic. Among extremal (co)extensions, there are some preferred (co)extensions which we call fully extremal because they satisfy a stronger maximality property. While in many classical cases, this notion reduces to the usual extremal property, we argue that in general they are preferred. The existence of fully extremal (co)extensions is established by an argument similar to Arveson’s proof [8] of the existence of maximal dilations.

Commutant lifting. The classical commutant lifting theorem was established by Sz.Nagy and Foiaș [59]. Many variations on this theorem have been established in various contexts for a variety of operator algebras. Douglas and Paulsen [26] formulate a version for arbitrary operator algebras, and we propose a modification of their definition.

Shilov representations of an operator algebra \( \mathcal{A} \) are those which are obtained by
taking a \( * \)-representation of the C*-envelope and restricting it to an invariant subspace for the image of \( \mathcal{A} \). All extremal coextensions (orthoprojective representations) are Shilov. The converse holds in some of the classical situations, but is not valid in general. As we will argue, the notions of commutant lifting are better expressed in terms of fully extremal coextensions rather than Shilov coextensions. Limiting the family of coextensions for which lifting occurs increases the family of algebras with this property. Indeed the \textit{strong} version of commutant lifting can only hold when there is a unique minimal fully extremal coextension (of \( \rho \)).

The Douglas-Paulsen formulation of commutant lifting starts with a (completely contractive) representation \( \rho \) of an operator algebra \( \mathcal{A} \), an operator \( X \) in the commutant of \( \rho(\mathcal{A}) \), and a Shilov coextension \( \sigma \) of \( \rho \); and they ask for a coextension of \( X \) to an operator \( Y \) of the same norm commuting with \( \sigma(\mathcal{A}) \). As remarked in the previous paragraph, this only holds when the minimal Shilov extension is unique. We show that this holds when \( \mathcal{A} \) is \textit{semi-Dirichlet}, meaning that
\[
\mathcal{A}^* \mathcal{A} \subset \overline{\mathcal{A} + \mathcal{A}^*},
\]
such as the disk algebra, the non-commutative disk algebras, and all tensor algebras of graphs and C*-correspondences. The fact that this large class of popular algebras has this remarkable property has perhaps kept us from looking further for an appropriate definition of commutant lifting in other contexts.

We were also influenced by a different approach of Paulsen and Power \cite{46,47} and subsequent work of theirs with the first author \cite{22,15}. In this version, the coextension \( \sigma \) is not specified, and one looks for common coextensions \( \sigma \) and \( Y \) which commute. We will use extremal coextensions only, rather than arbitrary Shilov coextensions, with the obvious parallel definitions. The first version will be called \textit{strong commutant lifting}, and the latter \textit{commutant lifting}. A crucial point is that strong commutant lifting turns out to be equivalent commutant lifting plus uniqueness of the minimal fully extremal coextension.

The intertwining version of commutant lifting proved to be challenging in this context. The resolution of this problem was critical to obtaining good dilation theorems for semicrossed products.

\textbf{Ando’s theorem.} Ando’s Theorem \cite{2} states that if \( A_1 \) and \( A_2 \) are commuting contractions, then they have coextensions \( V_i \) which are commuting isometries. For us, an Ando theorem for an operator algebra \( \mathcal{A} \) will be formulated as follows: given a (completely contractive) representation \( \rho \) of an operator algebra \( \mathcal{A} \) and a contraction \( X \) in the commutant of \( \rho(\mathcal{A}) \), there is a fully extremal coextension \( \sigma \) of \( \rho \) and an isometric coextension \( V \) of \( X \) which commutes with it. Even in the case of the disk algebra, our formulation is stronger than the original, as it asks that one of the isometries, say \( V_1 \), should have the form
\[
V_1 \simeq V_{A_1} \oplus U
\]
where $V_{A_1}$ is the minimal isometric coextension of $A_1$ and $U$ is unitary (see Corollary 7.11).

In the classical case of the disk algebra, the universal algebra of a contraction, the generator of a representation, $A = \rho(z)$, plays a role parallel to the operator $X$ which commutes with it. For this reason, commutant lifting can be applied recursively to $A$ and $X$, alternating between them, in order to obtain Ando’s Theorem. So in this context, the Sz.Nagy-Foiaș Commutant Lifting Theorem [59] is equivalent to Ando’s Theorem. But for other algebras, there are two distinct aspects, dilating $\rho$ to an extremal coextension and simultaneously coextending $X$ to a commuting contraction, and on the other hand coextending $X$ to an isometry and simultaneously coextending $\rho$ to a commuting representation.

Paulsen and Power [47] formulate Ando’s theorem as a dilation result for $A \otimes_{\min} A(\mathbb{D})$, or equivalently that

$$A \otimes_{\min} A(\mathbb{D}) = A \otimes_{\max} A(\mathbb{D}).$$

Such a result holds for a wide class of CSL algebras [46, 22, 15]. The stronger version of commutant lifting only holds in a restricted class [40]. See [40, chapter 5] for a discussion of the differences. In our language, they start with a representation $\rho$ and a commuting contraction $X$, and seek a maximal dilation $\pi$ and a simultaneous dilation of $X$ to a unitary $U$ commuting with $\pi(A)$. We show that this is equivalent to the weaker property of obtaining some coextension $\sigma$ of $\rho$ and an isometric coextension $V$ of $X$ which commute. This is only ‘half’ of Ando’s theorem in our formulation.

Another property that we will consider is an analogue of the Fuglede theorem: that the commutant of a normal operator is self-adjoint. We formulate this for an operator algebra $A$ with $C^*$-envelope $C^*_e(A)$ as saying that for any $\ast$-representation $\pi$ of $C^*_e(A)$, the commutant of $\pi(A)$ coincides with the commutant of $\pi(C^*_e(A))$. We show that a number of operator algebras have this property including all function algebras, the non-commutative disk algebras and more generally the tensor algebras of all finite directed graphs.

**Semicrossed products.** If $A$ is a unital operator algebra and $\alpha$ is a completely isometric endomorphism, then the semicrossed product

$$A \times_\alpha \mathbb{Z}_+$$

is the operator algebra that encodes the system $(A, \alpha)$ in the sense that its (completely contractive) representations are in bijective correspondence with the covariant representations of the dynamical system. Concrete versions of these algebras occur in work or Arveson [4, 11]. When $A$ is a $C^*$-algebra, the abstract semicrossed product was defined by Peters [48]. The extension to arbitrary nonself-adjoint operator algebras is straightforward.
The structure of these semicrossed products can often be better understood by showing that the $C^*$-envelope is a full $C^*$-algebra crossed product. Peters [49] does this for the semicrossed product that encodes a discrete dynamical system. The operator algebras of multivariable dynamical systems is developed in [16]. The $C^*$-envelope is further explained in [25], extending Peter’s analysis to this context. More recently, the second author and Kakariadis [32] develop an important generalization of these techniques to very general operator algebras. They show that for nonself-adjoint operator algebras, one first should try to imbibe a general semicrossed product into a $C^*$-semicrossed product. They show how to accomplish this, and demonstrate that often the two operator algebras have the same $C^*$-envelope.

When $\alpha$ is the identity map, the semicrossed product is closely tied to commutant lifting. What we show here is that commutant lifting theorems can be sufficient to understand other semicrossed products provided the algebra has some other nice properties. We concern ourselves only with endomorphisms that extend to $*$-endomorphisms of the $C^*$-envelope. When $\mathcal{A}$ satisfies the Ando property, every semicrossed product by a completely isometric automorphism is isometrically isomorphic to a subalgebra of the semicrossed product of $C^*_\alpha(\mathcal{A})$. These general techniques recover various results in the literature about the structure of crossed products, especially of the non-commutative disk algebras [17] and tensor algebras of $C^*$-correspondences [32]. To our knowledge, all of these results used the strong commutant lifting property (SCLT), which implies uniqueness of fully extremal extensions. Indeed, the theorems relate to algebras with a row contractive condition, the most general of which are tensor algebras of $C^*$=correspondences. Our new result requires only commutant lifting, and applies much more widely.

With a stronger commutant lifting property and the Fuglede property, we can do the same for endomorphisms which lift to the $C^*$-envelope. This applies, in particular, for the disk algebra (which has all of the good properties studied here). This recovers our results [18] for the semicrossed product of $A(D)$ by an endomorphism of the form $\alpha(f) = f \circ b$, in the case where $b$ is a non-constant finite Blaschke product. These general results that imbibe a semicrossed product into a $C^*$-algebra crossed product are actually dilation theorems. Typically one proves a dilation theorem first, and then deduces the structure of the $C^*$-envelope. However the papers [32, 18] actually compute the $C^*$-envelope first and deduce the dilation theorem afterwards. One of the original motivations for this paper was an attempt to identify the $C^*$-envelope of a semicrossed product using general dilation properties such as commutant lifting. Three such theorems are obtained in section 12.
2 A review of dilations

In this paper, an operator algebra will be a unital abstract operator algebra \( \mathcal{A} \) in the sense of Blecher, Ruan and Sinclair [12]. A representation of \( \mathcal{A} \) will mean a unital completely contractive representation \( \rho \) on some Hilbert space \( \mathcal{H} \). An extension of \( \rho \) is a representation \( \sigma \) on a Hilbert space \( \mathcal{K} = \mathcal{H}^\perp \oplus \mathcal{H} \) which leaves \( \mathcal{H} \) invariant, and thus has the form

\[
\sigma(a) = \begin{bmatrix}
\sigma_{11}(a) & 0 \\
\sigma_{12}(a) & \rho(a)
\end{bmatrix}.
\]

Dually, a coextension of \( \rho \) is a representation \( \sigma \) on a Hilbert space \( \mathcal{K} = \mathcal{H} \oplus \mathcal{H}^\perp \) which leaves \( \mathcal{H}^\perp \) invariant, and thus has the form

\[
\sigma(a) = \begin{bmatrix}
\rho(a) & 0 \\
\sigma_{12}(a) & \sigma_{22}(a)
\end{bmatrix}.
\]

A dilation of \( \rho \) is a representation \( \sigma \) on a Hilbert space \( \mathcal{K} \) containing \( \mathcal{H} \) so that \( \rho(a) = \text{Proj}\mathcal{H}\sigma(a)|_{\mathcal{H}} \). A familiar result of Sarason [55] shows that \( \mathcal{K} \) decomposes as \( \mathcal{K} = \mathcal{H}^- \oplus \mathcal{H} \oplus \mathcal{H}^+ \) so that

\[
\sigma(a) = \begin{bmatrix}
\sigma_{11}(a) & 0 & 0 \\
\sigma_{21}(a) & \rho(a) & 0 \\
\sigma_{31}(a) & \sigma_{32}(a) & \sigma_{33}(a)
\end{bmatrix}.
\]

A representation \( \rho \) is an extremal coextension if whenever \( \sigma \) is a coextension of \( \rho \), it necessarily has the form \( \sigma = \rho \oplus \sigma' \). That is, if \( \mathcal{H} \) is a subspace of \( \mathcal{K} \) and \( \sigma \) is a representation of \( \mathcal{A} \) on \( \mathcal{K} \) which leaves \( \mathcal{H}^\perp \) invariant and \( \text{Proj}\mathcal{H}\sigma(a)|_{\mathcal{H}} = \rho(a) \) for \( a \in \mathcal{A} \), then \( \mathcal{H} \) reduces \( \sigma \). Similarly, a representation \( \rho \) is an extremal extension if whenever \( \sigma \) is an extension of \( \rho \), it necessarily has the form \( \sigma = \rho \oplus \sigma' \). That is, if \( \mathcal{H} \) is a subspace of \( \mathcal{K} \) and \( \sigma \) is a representation of \( \mathcal{A} \) on \( \mathcal{K} \) which leaves \( \mathcal{H} \) invariant and \( \text{Proj}\mathcal{H}\sigma(a)|_{\mathcal{H}} = \rho(a) \) for \( a \in \mathcal{A} \), then \( \mathcal{H} \) reduces \( \sigma \). Finally, a representation \( \rho \) is an extremal representation or a maximal representation if whenever \( \sigma \) is a dilation of \( \rho \), it necessarily has the form \( \sigma = \rho \oplus \sigma' \). That is, if \( \mathcal{H} \) is a subspace of \( \mathcal{K} \) and \( \sigma \) is a representation of \( \mathcal{A} \) on \( \mathcal{K} \) so that \( \text{Proj}\mathcal{H}\sigma(a)|_{\mathcal{H}} = \rho(a) \) for \( a \in \mathcal{A} \), then \( \mathcal{H} \) reduces \( \sigma \). A dilation \( \sigma \) of \( \rho \) is an extremal dilation or a maximal dilation of \( \rho \) if it is a maximal representation.

Hilbert modules. In the module language espoused by Douglas and Paulsen in [26], a representation \( \rho \) makes the Hilbert space \( \mathcal{H} \) into a left \( \mathcal{A} \) module \( \mathcal{H}_\rho \) by \( a \cdot h := \rho(a)h \) for \( a \in \mathcal{A} \) and \( h \in \mathcal{H} \). If \( \mathcal{K} = \mathcal{M} \oplus \mathcal{H} \) and \( \sigma \) is a representation of \( \mathcal{A} \) on \( \mathcal{K} \) which leaves \( \mathcal{M} \) invariant, so that with respect to the decomposition \( \mathcal{K} = \mathcal{H} \oplus \mathcal{M} \) of \( \sigma \) is

\[
\sigma(a) = \begin{bmatrix}
\sigma_{11}(a) & 0 \\
\sigma_{12}(a) & \sigma_{22}(a)
\end{bmatrix}.
\]
then $K_\sigma$ is an $\mathcal{A}$-module with $M_{\sigma_{22}}$ as a submodule and $H_{\sigma_{11}}$ as a quotient module, and there is a short exact sequence
\[ 0 \to M_{\sigma_{22}} \to K_\sigma \to H_{\sigma_{11}} \to 0. \]
Here all module maps are completely contractive. So an extension $\sigma$ of $\sigma_{22}$ on $\mathcal{M}$ corresponds to larger Hilbert module $K_\sigma$ containing $M_{\sigma_{22}}$ as a submodule; and a coextension $\sigma$ of $\sigma_{11}$ corresponds to the Hilbert module $K_\sigma$ having $H_{\sigma_{11}}$ as a quotient module.

A module $\mathcal{P}_\rho$ is **orthoprojective** if whenever there is an **isometric** short exact sequence of module maps
\[ 0 \to M_{\sigma_{22}} \xrightarrow{\iota} K_\sigma \xrightarrow{q} \mathcal{P}_\rho \to 0, \]
meaning that $\iota$ is isometric and $q$ is coisometric, then there is an isometric module map $\varphi : \mathcal{P} \to K$ so that $K_\sigma = \mathcal{M} \oplus \varphi(\mathcal{P})$ as an $\mathcal{A}$-module. It is not difficult to see that this is equivalent to saying that $\rho$ is an extremal coextension. The term orthoprojective was coined by Muhly and Solel [40], and we think that it is superior to the Douglas-Paulsen terminology of hypo-projective because of its more positive aspect. Similarly, one can define **orthoinjective** modules, and observe that they are equivalent to extremal extensions. A maximal dilation corresponds to a module which is both orthoprojective and orthoinjective.

**The C*-envelope.** Every unital operator algebra $\mathcal{A}$ has a completely isometric representation $\iota$ on a Hilbert space $\mathcal{H}$ so that the C*-algebra $C^*(\iota(\mathcal{A})) =: C^*_\iota(\mathcal{A})$ is minimal in the sense that if $\sigma$ is any other completely isometric representation on a Hilbert space $\mathcal{H}'$, then there is a unique $*$-homomorphism $\pi$ of $C^*(\sigma(\mathcal{A}))$ onto $C^*_\iota(\mathcal{A})$ so that the following diagram commutes:
\[
\begin{array}{ccc}
\mathcal{A} & \xrightarrow{\iota} & C^*_\iota(\mathcal{A}) \\
\sigma \downarrow & & \pi \downarrow \\
C^*(\sigma(\mathcal{A})) & & C^*(\sigma(\mathcal{A}))
\end{array}
\]

The C*-envelope was described by Arveson [5, 6] in his seminal work on non-commutative dilation theory. Its existence was established by Hamana [31].

Muhly and Solel [41] show that maximal dilations exist by invoking Hamana’s theorem. They accomplish this by showing:

**Theorem 2.1 (Muhly-Solel).** A representation is maximal if and only if it is both orthoprojective and orthoinjective. Equivalently, a representation is maximal if and only if it is both an extremal coextension and an extremal extension.

Dritschel and McCullough [27] establish the existence of maximal dilations directly based on ideas of Agler [1]. In this way, they provide a new and more
revealing proof of the existence of the $C^*$-envelope. In fact, they show that every representation has an extension which is extremal; and dually also has a coextension which is extremal. In particular, the maximal representations of $\mathcal{A}$ are precisely those representations which extend to $*$-representations of $C^*_e(\mathcal{A})$.

Arveson [10] provides a refinement of this result in the separable case by showing that there are sufficiently many irreducible maximal representations, which are the boundary representations that Arveson introduced in [5] as an analogue of the Choquet boundary of a function algebra. We will not require this strengthened version.

**Extremal versus Shilov.** Douglas-Paulsen [26] and Muhly-Solel [40] focus on Shilov modules. One starts with a $*$-representation $\pi$ of $C^*_e(\mathcal{A})$ on a Hilbert space $\mathcal{K}$. Consider $\mathcal{K}_\pi$ as an $\mathcal{A}$-module. A submodule $\mathcal{H}$ of $\mathcal{K}_\pi$ is a Shilov module. It is easy to deduce from the discussion above that every orthoprojective module is Shilov. Unfortunately, the converse is false. We provide an example below. In the language of representations, a Shilov module corresponds to a representation which has an extension to a maximal representation. However it may still have proper coextensions.

Shilov modules are useful because every completely contractive $\mathcal{A}$-module $\mathcal{M}$ has a finite resolution of the form

$$0 \to \mathcal{S}_1 \to \mathcal{S}_2 \to \mathcal{M} \to 0,$$

where $\mathcal{S}_1$ and $\mathcal{S}_2$ are Shilov. Using orthoprojective modules, one can obtain

$$\mathcal{S}_2 \to \mathcal{M} \to 0$$

with $\mathcal{S}_2$ orthoprojective. But since submodules do not inherit this extremal property, one does not obtain a short exact sequence. Indeed, while this procedure can be iterated, there need be no finite resolution. This occurs, for example, in the theory of commuting row contractions due to Arveson [9, §9]. However Arveson also argues that, in his context, these are the natural resolutions to seek.

Our view is that it is the extremal coextensions rather than Shilov coextensions which play the role in dilation theory that best models the classical example of the unilateral shift as an isometric model of the disc algebra.

**Example 2.2.** Consider the non-commutative disk algebra $\mathfrak{A}_n$. It is the unital subalgebra of the Cuntz algebra $\mathcal{O}_n$ generated as a unital non-self-adjoint subalgebra by the canonical isometric generators $s_1, \ldots, s_n$ of $\mathcal{O}_n$. A representation $\rho$ of $\mathfrak{A}_n$ is determined by $A_i = \rho(s_i)$, and it is completely contractive if and only if

$$A = [A_1 \ldots A_n]$$

is a contraction as an operator from $\mathcal{H}^{(n)}$ to $\mathcal{K}$ [51]. The Frazho-Bunce-Popescu dilation theorem [30, 13, 50] states that $A$ has a coextension to a row isometry.
Conversely, it is clear that any coextension of a row isometry must be obtained as a direct sum. Thus these row isometric representations are precisely the extremal coextensions and correspond to orthoprojective modules. The Wold decomposition [23] shows that this row isometry decomposes as a direct sum of a Cuntz row unitary and a multiple of the left regular representation of the free semigroup $F_n^+$ on Fock space. This representation generates the Cuntz-Toeplitz $C^*$-algebra, and thus is not a maximal representation. It can be extended to a maximal dilation in many explicit ways [23]. It is clear in this case that every $*$-representation of $O_n$ sends

$$s = [s_1 \ldots s_n]$$

to a row unitary, and the restriction to any invariant subspace is a row isometry. Thus every Shilov module is orthoprojective.

Example 2.3. Let $A_n$ be the universal algebra of a row contraction with commuting coefficients. This algebra was studied extensively by Arveson beginning in [8]. The basic von Neumann inequality was proven much earlier by Drury [28], but the full version of the dilation theorem was due to Müller and Vasilescu [43] and later, Arveson. Arveson further showed that the multipliers $S_1, \ldots, S_n$ on symmetric Fock space $H^2_n$ in $n$ variables form a canonical model for $A_n$. Also $H^2_n$ is a reproducing kernel Hilbert space, and $A_n$ is the algebra of continuous multipliers. The $C^*$-algebra generated by these multipliers is the $C^*$-envelope of $A_n$ [8].

The dilation theorem shows that every commuting row contraction has a coextension to a direct sum $S_1^{\alpha} \oplus U_i$ where $\alpha$ is some cardinal and $U_i$ are commuting normal operators satisfying

$$\sum_{i=1}^n U_i U_i^* = I.$$ 

These are precisely the extremal coextensions and determine the orthoprojective modules. Surprisingly they are also the maximal representations. So while one can dilate in both directions to obtain a maximal dilation of a representation $\rho$, only coextensions are required.

However, no submodule of the symmetric Fock space is orthoprojective. They are all Shilov, but none model the algebra in a useful way. Davidson and Le [21, Example 4.1] provide an explicit example of this phenomenon in their paper on the commutant lifting theorem for $A_n$.

3 Fully Extremal Coextensions

There is a natural partial order $\prec$ on dilations: say that $\rho \prec \sigma$ if $\sigma$ acts on a Hilbert space $K$ containing a subspace $H$ so that $P_H \sigma|_H$ is unitarily equivalent to $\rho$. There is also a partial order on extensions $\prec_e$: say that $\rho \prec_e \sigma$ if $\sigma$
acts on a Hilbert space $K$ containing an invariant subspace $H$ so that $\sigma|_H$ is unitarily equivalent to $\rho$. Similarly, for coextensions, say that $\rho \prec_c \sigma$ if $\sigma$ acts on a Hilbert space $K$ containing a co-invariant subspace $H$ so that $P_H\sigma|_H$ is unitarily equivalent to $\rho$.

Dritschel and McCullough [27] establish the existence of extremals dominating $\rho$ in each of these classes. We want something a little bit stronger. It is possible for an extremal coextension $\sigma$ of $\rho$ to have a proper extension which is also a coextension of $\rho$, so that $\sigma$ is not extremal in the partial order $\prec$. We provide an example shortly. We will require knowing that $\rho$ has a coextension which is extremal with respect to $\prec$.

**Definition 3.1.** If $\rho$ is a representation of $A$, say that a coextension $\sigma$ of $\rho$ is **fully extremal with respect to $\rho$** if whenever $\sigma \prec \tau$ and $\rho \prec_c \tau$, then $\tau = \sigma \oplus \tau'$. Similarly we define an extension $\sigma$ of $\rho$ to be fully extremal with respect to $\rho$ if whenever $\sigma \prec \tau$ and $\rho \prec_e \tau$, then $\tau = \sigma \oplus \tau'$.

**Example 3.2.** Fix an orthonormal basis $e_1, \ldots, e_n$ for $\mathbb{C}^n$, and let $E_{ij}$ be the canonical matrix units. Consider the subalgebra $A$ of $\mathbb{M}_n$ spanned by the diagonal algebra

$$D_n = \text{span}\{E_{ii} : 1 \leq i \leq n\}$$

and

$$\text{span}\{E_{ij} : |i - j| = 1, \ j \text{ odd} \}.$$ 

This is a reflexive operator algebra with invariant subspaces

$$Ce_{2i}, \text{ for } 1 < 2i \leq n$$

and

$$L_{2i+1} = \text{span}\{e_{2i}, e_{2i+1}, e_{2i+2}\} \text{ for } 1 \leq 2i + 1 \leq n,$$

where we ignore $e_0$ and $e_{n+1}$ if they occur. The elements of $A$ have the form

$$A = \begin{bmatrix}
a_{11} & 0 & 0 & 0 & 0 & \cdots \\
a_{21} & a_{22} & a_{32} & 0 & 0 & \cdots \\
0 & 0 & a_{33} & 0 & 0 & \cdots \\
0 & 0 & a_{43} & a_{44} & a_{45} & \cdots \\
0 & 0 & 0 & 0 & a_{55} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \\
\end{bmatrix}$$

Consider the representation $\rho(A) = a_{11}$, the 1,1 matrix entry of $A$. Since $Ce_1$ is coinvariant, this is a representation. The compression $\sigma_2$ of $A$ to $M_2 = \text{span}\{e_1, e_2\}$ is a coextension of $\rho$ given by

$$\sigma_2(A) = P_{M_2}A|_{M_2} = L_1 \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix}$$
This is readily seen to be an extremal coextension of $\rho$. It is minimal in the sense we use: it contains no proper reducing subspace containing $\mathcal{H}_\rho = Cc_1$, and is also minimal in the sense that $\mathcal{H}_\sigma = \sigma(A)\mathcal{H}_\rho$.

However $\sigma_2$ is not fully extremal. Let

$$M_k = \text{span}\{e_i : 1 \leq i \leq k\}$$

and set

$$\sigma_k(A) = P_{M_k}A|_{M_k}.$$ 

Then $\sigma_{2i+1}$ is an extension of $\sigma_{2i}$ and $\sigma_{2i+2}$ is a coextension of $\sigma_{2i+1}$. All are coextensions of $\rho$. Each $\sigma_{2i}$ is an extremal coextension of $\rho$, as is $\sigma_n = \text{id}$ even if $n$ is even. Moreover all are minimal in that they have no proper reducing subspace containing $Cc_1$. Only $\sigma_n$ is fully extremal. Note that to get from $\rho$ to $\sigma_n$, one must alternately coextend and extend $n - 1$ times if at each stage, you take a classical minimal extension or coextension.

One can also define an infinite version of this algebra where it takes a countable number of steps to attain the fully extremal coextension.

**Example 3.3.** Let $A(D)$ be the disk algebra. A representation of $A(D)$ is determined by $T = \rho(z)$, and it is completely contractive if and only if $\|T\| \leq 1$. Every contraction coextends to a unique minimal isometry. So the extremal coextensions must be isometries. But conversely, it is easy to see that any contractive coextension of an isometry is obtained by adding a direct sum. So when $T$ is an isometry, $\rho$ is an extremal coextension. The minimal isometric dilation $V$ of $T$ yields a fully extremal coextension because the range of $V$ together with $\mathcal{H}_\rho$ spans the whole space. Any (contractive) dilation of $V$ must map the new subspace orthogonal to the range of $V$. So if it is not a summand, the range will not be orthogonal to $\mathcal{H}_\rho$, so it won’t be a coextension of $\rho$.

The extremal coextensions of $\rho$ correspond to all isometric coextensions of $T$, namely $V \oplus W$ where $V$ is the minimal isometric dilation and $W$ is any isometry. But if $W$ isn’t unitary, it can be extended to a unitary. This extension is still a coextension of $T$. So the fully extremal coextensions correspond to $V \oplus U$ where $U$ is unitary.

Similarly, the maximal dilations of $\rho$ correspond to unitary dilations of $A$. The restriction of a unitary to an invariant subspace is an isometry. So a Shilov representations are extremal coextensions. In particular, a minimal Shilov dilation of $\rho$ is a fully extremal coextension.

**Example 3.4.** Let $A = A(D^2)$ with generators $z_1$ and $z_2$. Then a completely contractive representation is determined by a pair of commuting contractions $A_i = \rho(z_i)$. By Ando’s Theorem [2], every commuting pair coextends to a pair of commuting isometries. It is clear that any coextension of isometries to a larger space can only be obtained by adding a direct summand. So the
extremal coextensions are the commuting isometries. It is also clear that any restriction to an invariant subspace is still isometric.

Moreover every pair of commuting isometries extends to a pair of commuting unitaries. These are the maximal dilations, and determine a $*$-representation of

$$C_\ell(A) = C(T^2).$$

The restriction of a unitary to an invariant subspace is an isometry. So every Shilov module is orthoprojective (an extremal coextension).

What we wish to point out is that extremal coextensions of a representation $\rho$ of $A(D^2)$ need not be fully extremal. Let

$$A_1 = A_2 = 0$$

acting on $H = \mathbb{C}$.

Identify $H$ with $\mathbb{C}e_{0,0}$ in

$$K = \text{span}\{e_{m,n} : m, n \geq 0\},$$

where $\{e_{m,n} : m, n \geq 0\}$ is an orthonormal basis. Then it is clear that there is a coextension of $A_i$ to the commuting isometries

$$S_1 = S \otimes I \text{ and } S_2 = I \otimes S,$$

where $S$ is the unilateral shift. Let $\sigma$ be the corresponding coextension of $\rho$. This is an extremal coextension because the $S_i$ are isometries.

Enlarge this orthonormal basis further to obtain a space

$$L = \text{span}\{e_{k,l} : \max\{k, l\} \geq 1 \text{ or } k = l = 0\}$$

containing $K$. Let $T_i$ be the commuting isometries given by

$$T_1 e_{m,n} = e_{m+1,n} \text{ and } T_2 e_{m,n} = e_{m,n+1}.$$ 

Let $\tau$ be the induced representation of $A(D^2)$. It is clear by inspection that $H = Ce_{00}$ is coinvariant, and hence $\rho \prec_e \tau$. Moreover, $\tau$ is extremal because $T_i$ are isometries. The subspace $K$ is invariant for $T_1$ and $T_2$, and $T_i|_K = S_i$. Therefore $\sigma \prec_e \tau$. So $\sigma$ is not fully extremal with respect to $\rho$.

We claim that $\tau$ is fully extremal with respect to $\rho$. Since it is extremal, it can only fail to be fully extremal if there is a larger space $M \supset L$ and commuting isometries $V_i$ on $M$ extending $T_i$ so that $L$ is not coinvariant, but $Ce_{00}$ is. Hence one of the isometries, say $V_1$, has $P_L V_1 P_L \neq 0$. Let

$$N = (\text{Ran} T_1 \lor Ce_{00})^\perp = \text{span}\{e_{1,l} : l < 0\}.$$ 

There must be a vector $x \in N$ so that $V_1^* x \neq 0$. Equivalently, there are vectors $y, z \in L^\perp$ so that $V_1 y = z + x$. Write $x = \sum_{l<0} a_l e_{1,l}$, and let $l_0$ be the least integer so that $a_{-l_0} \neq 0$. Let

$$x' = T_2^{l_0-1} x = \sum_{l<0} a_{l+1-l_0} e_{1,l} := \sum_{l<0} a'_l e_{1,l};$$
so that \( a'_{-1} \neq 0 \). Also set
\[
y' = V_2^{l_0-1}y \quad \text{and} \quad z' = V_2^{l_0-1}z.
\]
Then
\[
V_1y' = V_1V_2^{l_0-1}y = V_2^{l_0-1}V_1y = V_2^{l_0-1}(z + x) = z' + x'.
\]
Moreover, \( z' \) is orthogonal to the range of \( T_2^{l_0-1} \), which contains \( \mathcal{N} \). Hence
\[
\langle V_2V_1y', e_{1,0} \rangle = \langle V_2(z' + x'), e_{1,0} \rangle = \langle z' + x', e_{1,-1} \rangle = a'_{-1} \neq 0.
\]
Therefore
\[
0 \neq \langle V_1V_2y', e_{1,0} \rangle = \langle V_2y', e_{0,0} \rangle.
\]
This contradicts the fact that \( \tau \) is a coextension of \( \rho \). Thus \( \tau \) must be fully extremal relative to \( \rho \).

Now we turn to the issue of establishing that fully extremal coextensions (and extensions) always exist.

**Theorem 3.5.** Let \( A \) be a unital operator algebra, and let \( \rho \) be a representation of \( A \) on \( \mathcal{H} \). Then \( \rho \) has a fully extremal coextension \( \sigma \).

If \( A \) and \( \mathcal{H} \) are separable, then one can take \( \sigma \) acting on a separable Hilbert space.

**Proof.** Our argument is based on Arveson’s proof [10, Theorem 2.5] that maximal dilations exist. He works with the operator system generated by \( A \), which is self-adjoint. As we will work directly with \( A \), we need to consider adjoints as well. The goal is to construct a coextension \( \sigma \) of \( \rho \) on a Hilbert space \( K \) so that for every \( a \in A \) and \( k \in K \),
\[
\|\sigma(a)k\| = \sup\{\|\tau(a)k\| : \tau \succ \sigma, \tau \succ c\rho\}
\]
and
\[
\|\sigma(a)^*k\| = \sup\{\|\tau(a)^*k\| : \tau \succ \sigma, \tau \succ c\rho\}.
\]
Once this is accomplished, it is evident that any dilation \( \tau \) of \( \sigma \) which is a coextension of \( \rho \) must have \( K \) as a reducing subspace, as claimed.

To this end, choose a dense subset of \( A \times \mathcal{H} \), and enumerate it as
\[
\{(a_\alpha, h_\alpha) : \alpha \in \Lambda\}
\]
where $\Lambda$ is an ordinal. Suppose that we have found coextensions $\sigma_\alpha$ of $\rho$ for all $\alpha < \alpha_0 < \Lambda$ acting on $K_\alpha$, where $K_\beta \subset K_\alpha$ when $\beta < \alpha$, so that
\[ \|\sigma_\alpha(a_\beta)h_\beta\| = \sup\{\|\tau(a_\beta)h_\beta\| : \tau \succ \sigma_\alpha, \tau \succ_c \rho\} \] (1')
and
\[ \|\sigma_\alpha(a_\beta)^*h_\beta\| = \sup\{\|\tau(a)^*h_\beta\| : \tau \succ \sigma_\alpha, \tau \succ_c \rho\}. \] (2')
for all $\beta < \alpha$. This latter condition is automatic because each $\tau(a)^*$ leaves $H$ invariant, and agrees with $\rho(a)^*$ there. But we carry this for future use.

If $\alpha_0$ is a limit ordinal, we just form the natural direct limit of the $\sigma_\alpha$ for $\alpha < \alpha_0$, and call it $\sigma_{\alpha_0}$. Note that it will now satisfy (1') and (2') for $\beta < \alpha_0$.

Otherwise $\alpha_0 = \beta_0 + 1$ is a successor ordinal. Choose a dilation $\tau_1 \succ \sigma_{\beta_0}$ on $M_1 \supset K_{\beta_0}$ such that $\tau_1 \succ_c \rho$ and satisfies
\[ \|\tau_1(a_{\beta_0})h_{\beta_0}\| \geq \sup\{\|\tau(a_{\beta_0})h_{\beta_0}\| : \tau \succ \sigma_{\alpha}, \tau \succ_c \rho\} - 2^{-1} \] and
\[ \|\tau_1(a_{\beta_0})^*h_{\beta_0}\| \geq \sup\{\|\tau(a)^*h_{\beta_0}\| : \tau \succ \sigma_{\alpha}, \tau \succ_c \rho\} - 2^{-1}. \]

Then choose recursively dilations $\tau_{n+1}$ of $\tau_n$ on $M_{n+1} \supset M_n$ which are all coextensions of $\rho$ so that
\[ \|\tau_{n+1}(a_{\beta_0})h_{\beta_0}\| \geq \sup\{\|\tau(a_{\beta_0})h_{\beta_0}\| : \tau \succ \sigma_{\alpha}, \tau \succ_c \rho\} - 2^{-n-1} \] and
\[ \|\tau_{n+1}(a_{\beta_0})^*h_{\beta_0}\| \geq \sup\{\|\tau(a)^*h_{\beta_0}\| : \tau \succ \sigma_{\alpha}, \tau \succ_c \rho\} - 2^{-n-1}. \]

The inductive limit is a representation $\sigma_{\alpha_0}$ with the desired properties.

Once we reach $\Lambda$, we have constructed a representation $\tilde{\sigma}_1$ on $\tilde{K}_1$ coextending $\rho$ and satisfying (1) and (2) for vectors $h \in H$. Now repeat this starting with $\tilde{\sigma}_1$ and a dense subset of $A \times \tilde{K}_1$, but still considering dilations which are coextensions of $\rho$. This time, the equations involving the adjoint are important. The result is a representation $\tilde{\sigma}_2$ on $\tilde{K}_2$ dilating $\tilde{\sigma}_1$ and coextending $\rho$ satisfying (1) and (2) for all vectors $k \in \tilde{K}_1$. Repeat recursively for all $n \geq 1$ and in the end, we obtain the desired coextension.

If $A$ and $K$ are separable, a countable sequence of points suffices, and at each stage of this countable process, one obtains separable spaces. So the result is a separable representation.

**Remark 3.6.** It easily follows from the proof of existence of fully extremal coextensions that if $\sigma$ is a coextension of $\rho$, then there is a dilation $\tau$ of $\sigma$ which is a fully extremal coextension of $\rho$. 

**Documenta Mathematica 16 (2011) 781–868**
Remark 3.7. A proof of existence of extremal coextensions can be made along the same lines. It is only necessary to achieve \( \sigma \succ_c \rho \) on \( K \) such that:

\[
\| \sigma(a)k \| = \sup \{ \| \tau(a)k \| : \tau \succ_c \sigma \}.
\]

One can always achieve this by repeated coextension, and in this way one obtains an extremal coextension \( \sigma \) of \( \rho \) with the additional property that \( \mathcal{H} \) is cyclic, i.e. \( \mathcal{K} = \sigma(A)\mathcal{H} \). This is evidently not the case in general for extremal coextensions, and in particular, for fully extremal coextensions. See the preceding examples and Remark 3.13.

The same result for extensions follows by duality.

Corollary 3.8. Let \( A \) be a unital operator algebra, and let \( \rho \) be a representation of \( A \) on \( \mathcal{H} \). Then \( \rho \) has a fully extremal extension \( \sigma \).

Corollary 3.9. If \( \rho_1 \) and \( \rho_2 \) are representations of \( A \), then \( \sigma_1 \) and \( \sigma_2 \) are fully extremal coextensions of \( \rho_1 \) and \( \rho_2 \), respectively, if and only if \( \sigma_1 \oplus \sigma_2 \) is a fully extremal coextension of \( \rho_1 \oplus \rho_2 \).

In particular, \( \sigma_1 \) and \( \sigma_2 \) are extremal coextensions of \( A \) if and only if \( \sigma_1 \oplus \sigma_2 \) is an extremal coextension.

Proof. First suppose that \( \sigma_1 \) and \( \sigma_2 \), acting on \( \mathcal{K}_1 \) and \( \mathcal{K}_2 \), are fully extremal coextensions of \( \rho_1 \) and \( \rho_2 \), respectively. Suppose that \( \tau \) is a representation on \( \mathcal{P} = \mathcal{K}_1 \oplus \mathcal{K}_2 \oplus \mathcal{P}' \) such that

\[
\tau \succ \sigma_1 \oplus \sigma_2 \quad \text{and} \quad \tau \succ_c \rho_1 \oplus \rho_2.
\]

Then as \( \tau \succ \sigma_1 \) and \( \tau \succ_c \rho_1 \), we deduce that \( \tau \) reduces \( \mathcal{K}_i \) and hence reduces \( \mathcal{K}_1 \oplus \mathcal{K}_2 \). So \( \sigma_1 \oplus \sigma_2 \) is a fully extremal coextension of \( \rho_1 \oplus \rho_2 \).

Conversely, if \( \sigma_1 \oplus \sigma_2 \) is a fully extremal coextension of \( \rho_1 \oplus \rho_2 \), suppose that \( \tau \) is a representation on \( \mathcal{P} = \mathcal{K}_1 \oplus \mathcal{P}' \) satisfies \( \tau \succ \sigma_1 \) and \( \tau \succ_c \rho_1 \). Then

\[
\tau \oplus \sigma_2 \succ \sigma_1 \oplus \sigma_2 \quad \text{and} \quad \tau \oplus \sigma_2 \succ_c \rho_1 \oplus \rho_2.
\]

It follows that \( \tau \oplus \sigma_2 \) reduces \( \mathcal{K}_1 \oplus \mathcal{K}_2 \). So \( \tau \) reduces \( \mathcal{K}_1 \). Whence \( \sigma_1 \) is fully extremal.

Applying this to \( \rho_i = \sigma_i \) yields the last statement.

If one starts with a representation \( \rho \) and alternately forms extremal extensions and coextensions, it may require a countable sequence of alternating extensions and coextensions in order to obtain a maximal dilation as in Example 3.2. One advantage of fully extremal extensions and coextensions is that only one is required to obtain a maximal dilation.
Proposition 3.10. Let $\rho$ be a representation of $A$. If $\sigma$ is an extremal co-extension of $\rho$, and $\tau$ is a fully extremal extension of $\sigma$, then $\tau$ is a maximal dilation.

Proof. Since $\tau$ is an extremal extension, it suffices to show that it is also an extremal coextension. For then the Muhly-Solel result, Theorem 2.1, will show that $\tau$ is a maximal dilation.

Say that $\rho$, $\sigma$ and $\tau$ act on $\mathcal{H}$, $\mathcal{K}$ and $\mathcal{L}$ respectively. Suppose that $\pi$ is a coextension of $\tau$ acting on $\mathcal{P}$. Decompose

$$\mathcal{P} = (\mathcal{L} \oplus \mathcal{K}) \oplus \mathcal{H} \oplus (\mathcal{K} \oplus \mathcal{H}) \oplus (\mathcal{P} \oplus \mathcal{L}).$$

Then we have

$$\pi = \begin{bmatrix}
\tau_{11} & 0 & 0 & 0 \\
\tau_{21} & \rho & 0 & 0 \\
\tau_{31} & \sigma_{32} & \sigma_{33} & 0 \\
\pi_{41} & \pi_{42} & \pi_{43} & \pi_{44}
\end{bmatrix},$$

where $\sigma$ is represented by the middle $2 \times 2$ square, and $\tau$ is represented by the upper left $3 \times 3$ corner. The lower right $3 \times 3$ corner is a coextension of $\sigma$. Since $\sigma$ is an extremal coextension, we obtain

$$\pi_{42} = 0 = \pi_{43}.$$

Thus we can rearrange the decomposition moving $\mathcal{P} \oplus \mathcal{L}$ to the first coordinate to obtain

$$\pi \simeq \begin{bmatrix}
\pi_{44} & \pi_{41} & 0 & 0 \\
0 & \tau_{11} & 0 & 0 \\
0 & \tau_{21} & \rho & 0 \\
0 & \tau_{31} & \sigma_{32} & \sigma_{33}
\end{bmatrix}.$$ 

This is a coextension of $\tau$ which is an extension of $\sigma$. By the fact that $\tau$ is a fully extremal extension of $\sigma$, we deduce that $\pi_{41} = 0$ and so

$$\pi \simeq \pi_{44} \oplus \tau.$$ 

Therefore $\tau$ is also an extremal coextension.

The dual result is obtained the same way.

Corollary 3.11. Let $\rho$ be a representation of $A$. If $\sigma$ is an extremal extension of $\rho$, and $\tau$ is a fully extremal coextension of $\sigma$, then $\tau$ is a maximal dilation.

The classical notion of minimal coextension is that the space is cyclic for $A$. However, it seems more natural that the original space merely generate the whole space as a reducing subspace. This is because fully extremal coextensions do not generally live on the cyclic subspace generated by the original space.
Definition 3.12. An extremal coextension $\sigma$ on $K$ of a representation $\rho$ of $A$ on $H$ is minimal if the only reducing subspace of $K$ containing $H$ is $K$ itself. Likewise we define minimality for fully extremal coextensions, extremal extensions and fully extremal extensions. This minimal (fully) extremal (co)extension is unique if any two of these objects are unitarily equivalent via a unitary which is the identity on $H$.

Say that a coextension $\sigma$ on $K$ of a representation $\rho$ of $A$ on $H$ is cyclic if $K = \sigma(A)H$, i.e. $H$ is cyclic for $\sigma(A(D^2))$. However the extension $\tau$ of $\sigma$ is also an extremal coextension of $\rho$. While it is no longer true that $\tau(A(D^2))H$ is the whole space, it is nevertheless the smallest reducing subspace containing $H$, and so it is also minimal. Thus it is a minimal fully extremal coextension.

It is important to note that there are minimal fully extremal coextensions obtained in the natural way.

Proposition 3.14. Let $\rho$ be a representation of $A$ on $H$. Let $\sigma$ be a fully extremal (co)extension of $\rho$ on $K$. Let $\sigma_0$ be the restriction of $\sigma$ to the smallest reducing subspace $K_0$ for $\sigma(A)$ containing $H$. Then $\sigma_0$ is fully extremal. Moreover, $\sigma = \sigma_0 \oplus \sigma_1$ where $\sigma_1$ is a maximal representation. Conversely, every (co)extension of this form is fully extremal.

Proof. The proof is straightforward. Since $K_0$ reduces $\sigma$, we can write

$$\sigma = \sigma_0 \oplus \sigma_1$$

acting on $K = K_0 \oplus K_0^\perp$. Suppose that $\tau$ is a dilation of $\sigma_0$ which is a coextension of $\rho$. Then $\tau \oplus \sigma_1$ is a dilation of $\sigma$ which is a coextension of $\rho$. Since $\sigma$ is fully extremal, we have a splitting

$$\tau \oplus \sigma_1 \simeq \sigma \oplus \tau_1 = \sigma_0 \oplus \sigma_1 \oplus \tau_1.$$ 

Hence

$$\tau = \sigma_0 \oplus \tau_1.$$ 

It follows that $\sigma_0$ is a fully extremal coextension of $\rho$. 

Documenta Mathematica 16 (2011) 781–868
Any dilation of $\sigma_1$ yields a dilation of $\sigma$ which is a coextension of $\rho$. As $\sigma$ is fully extremal, this must be by the addition of a direct summand. Hence $\sigma_1$ is a maximal representation. Conversely, if $\sigma_0$ is a (minimal) fully extremal coextension of $\rho$ and $\sigma_1$ is a maximal representation, then $\sigma = \sigma_0 \oplus \sigma_1$ is a fully extremal coextension because any dilation of $\sigma$ is a dilation of $\sigma_0$ direct summed with $\sigma_1$.

The same argument works for extensions.

We refine Proposition 3.10. In light of Remark 3.7, we know that the coextensions asked for in the following proposition always exist.

**Proposition 3.15.** Let $\rho$ be a representation of $A$ on $H$. Let $\sigma$ be a cyclic extremal coextension of $\rho$ on $K$. Let $\pi$ be a minimal fully extremal extension of $\sigma$. Then $\pi$ is a minimal maximal dilation of $\rho$.

**Proof.** It suffices to show that the whole space, $L$, is the smallest reducing subspace for $\pi(A)$ containing $H$. In particular, it contains

$$\pi(A)H = \sigma(A)H = K.$$

But the minimality of $\pi$ as a fully extremal extension of $\sigma$ ensures that there is no proper reducing subspace containing $K$. So $\pi$ is minimal as a maximal dilation.

We require a result which is more subtle than Proposition 3.15 but is valid for fully extremal coextensions.

**Theorem 3.16.** Let $\rho$ be a representation of $A$ on $H$. Let $\sigma$ be a minimal fully extremal coextension of $\rho$ on $K$. Let $\pi$ be a minimal fully extremal extension of $\sigma$. Then $\pi$ is a minimal maximal dilation of $\rho$.

Moreover, the representation $\pi$ determines $\sigma$, and thus two inequivalent minimal fully extremal coextensions of $\rho$ yield inequivalent minimal maximal dilations of $\rho$.

**Proof.** Let $\pi$ act on the Hilbert space $L$. Note that $\pi$ is a maximal dilation of $\rho$ by Proposition 3.10. Let

$$M = \overline{\pi(A)H} \oplus H.$$

Then $M^\perp$ is the largest invariant subspace for $\pi(A)$ in which $H$ is coinvariant. Let $\tau$ denote the restriction of $\pi$ to $M^\perp$. Since $M^\perp$ contains $K$, we have

$$\tau \succ \sigma \quad \text{and} \quad \tau \succ_c \rho.$$

Hence by the fully extremal property of $\sigma$, we deduce that

$$\tau = \sigma \oplus \tau' \quad \text{on} \quad M^\perp = K \oplus (M + K)^\perp.$$
Now the smallest reducing subspace for $\pi(A)$ containing $\mathcal{H}$ clearly contains $\mathcal{M}$. Thus it contains the smallest $\tau(A)$ reducing subspace of $\mathcal{M}^\perp$ containing $\mathcal{H}$. But since $\tau = \sigma \oplus \tau'$ and $\sigma$ is minimal as a fully extremal coextension, the smallest $\tau(A)$ reducing subspace containing $\mathcal{H}$ is $\mathcal{K}$. Then since $\pi$ is a minimal fully extremal extension of $\sigma$, we see that $\mathcal{L}$ is the smallest reducing subspace containing $\mathcal{K}$. So $\pi$ is minimal.

From the arguments above, we see that $\sigma$ is recovered from $\pi$ by forming

$$\mathcal{M} = \pi(A)^* \mathcal{H} \ominus \mathcal{H}.$$  

restricting $\pi$ to $\mathcal{M}^\perp$ to get $\tau$, and taking the smallest $\tau$ reducing subspace of $\mathcal{M}^\perp$ containing $\mathcal{H}$. The restriction to this subspace is $\sigma$. Hence $\pi$ determines $\sigma$. Consequently, two inequivalent fully extremal coextensions of $\rho$ yield inequivalent minimal maximal dilations of $\rho$.

The following is immediate by duality.

**Corollary 3.17.** Let $\rho$ be a representation of $A$ on $\mathcal{H}$. Let $\sigma$ be a minimal fully extremal extension of $\rho$ on $\mathcal{K}$. Let $\pi$ be a minimal fully extremal coextension of $\sigma$. Then $\pi$ is a minimal maximal dilation of $\rho$.

Moreover, the dilation $\pi$ determines $\sigma$. Thus two inequivalent minimal fully extremal extensions of $\rho$ yield inequivalent minimal maximal dilations of $\rho$.

## 4 Semi-Dirichlet algebras

In this section, we consider a class of algebras where the theory is more like the classical one. The semi-Dirichlet property is a powerful property that occurs often in practice. From the point of view of dilation theory, these algebras are very nice.

**Definition 4.1.** Say that an operator algebra $A$ is *semi-Dirichlet* if

$$A^*A \subset A + A^*$$

when $A$ is considered as a subspace of its C*-envelope.

A unital operator algebra (not necessarily commutative) is called *Dirichlet* if $A + A^*$ is norm dense in $C_e^*(A)$.

Notice that since $A$ is unital, we always have $A + A^* \subset \text{span}(A^*A)$, so semi-Dirichlet means that

$$\text{span}(A^*A) = A + A^*.$$  

The interested reader can note that in the case of w*-closed algebras, the proofs below can be modified to handle the natural w*-closed condition which we call *semi-$\sigma$-Dirichlet* if

$$A^*A \subset A + A^{\sigma-\ast}.$$  

**Documenta Mathematica** 16 (2011) 781–868
Free semigroup algebras and free semigroupoid algebras of graphs and nest algebras all are semi-$\sigma$-Dirichlet.

The following simple proposition establishes a few elementary observations.

**Proposition 4.2.**

(i) $\mathcal{A}$ is Dirichlet if and only if $\mathcal{A}$ and $\mathcal{A}^*$ are semi-Dirichlet.

(ii) If $\sigma$ is a completely isometric representation of $\mathcal{A}$ on $\mathcal{H}$, and

$$\sigma(\mathcal{A})^*\sigma(\mathcal{A}) \subset \sigma(\mathcal{A}) + \sigma(\mathcal{A})^*,$$

then $\mathcal{A}$ is semi-Dirichlet.

(iii) If $\sigma$ is a Shilov representation of a semi-Dirichlet algebra $\mathcal{A}$, then $\sigma(\mathcal{A})$ is semi-Dirichlet.

**Proof.** (i) It is obvious that if $\mathcal{A}$ is Dirichlet, then both $\mathcal{A}$ and $\mathcal{A}^*$ are semi-Dirichlet. For the converse, notice that if $\mathcal{A}$ is semi-Dirichlet, then an easy calculation shows that $\text{span}(\mathcal{A}\mathcal{A}^*)$ is a C*-algebra [8]. Since $\mathcal{A}$ generates $C^*_{\text{e}}(\mathcal{A})$, this is the C*-algebra $\text{span}(\mathcal{A}\mathcal{A}^*)$. Thus the semi-Dirichlet property for $\mathcal{A}$ now shows that $\mathcal{A} + \mathcal{A}^*$ is norm dense in $C^*_{\text{e}}(\mathcal{A})$.

(ii) If $\sigma$ is completely isometric, then $\mathcal{A} = C^*_{\text{e}}(\sigma(\mathcal{A}))$ is a C*-cover of $\mathcal{A}$. By the minimal property of the C*-envelope, there is a quotient map $q : \mathcal{A} \to C^*_{\text{e}}(\mathcal{A})$ so that $q\sigma|_{\mathcal{A}}$ is the identity map. If $\sigma(\mathcal{A})^*\sigma(\mathcal{A})$ is contained in $\sigma(\mathcal{A}) + \sigma(\mathcal{A})^*$, then passing to the quotient yields the semi-Dirichlet property.

(iii) If $\sigma$ is Shilov, then there is a $*$-representation $\pi$ of $C^*_{\text{e}}(\mathcal{A})$ on $\mathcal{K}$ and an invariant subspace $\mathcal{H}$ so that $\sigma(a) = \pi(a)|_{\mathcal{H}}$. The map

$$\tilde{\sigma}(x) = P_{\mathcal{H}}\pi(x)|_{\mathcal{H}} \quad \text{for } x \in C^*_{\text{e}}(\mathcal{A})$$

is a completely positive map extending $\sigma$. In particular,

$$\tilde{\sigma}(a^*) = \sigma(a)^* \quad \text{for } a \in \mathcal{A}.$$

For $a, b \in \mathcal{A}$, we calculate

$$\pi(a^*b) = \left[ \begin{array}{cc} * & * \\ * & \tilde{\sigma}(a^*b) \end{array} \right] = \pi(a)^*\pi(b)$$

$$= \left[ \begin{array}{cc} * & * \\ 0 & \sigma(a)^* \end{array} \right] \left[ \begin{array}{cc} * & 0 \\ * & \sigma(b) \end{array} \right] = \left[ \begin{array}{cc} * & * \\ * & \sigma(a)^*\sigma(b) \end{array} \right]$$

Hence

$$\tilde{\sigma}(a^*b) = \sigma(a)^*\sigma(b) \quad \text{for all } a, b \in \mathcal{A}.$$

Since $\mathcal{A}$ is semi-Dirichlet, we can write

$$a^*b = \lim_n c_n^* + d_n$$
where \( c_n, d_n \in \mathcal{A} \). Thus,
\[
\sigma(a)^\ast \sigma(b) = \bar{\sigma}(a^* b) = \lim \sigma(c_n)^* + \sigma(d_n).
\]
That is,
\[
\sigma(\mathcal{A})^\ast \sigma(\mathcal{A}) \subset \overline{\sigma(\mathcal{A})^\ast + \sigma(\mathcal{A})^\ast}.
\]
It now follows from (ii) that \( \sigma(\mathcal{A}) \) is semi-Dirichlet.

**Example 4.3.** Observe that if \( \mathcal{A} \) is a function algebra with Shilov boundary \( X = \partial \mathcal{A} \), then \( \text{span}(\mathcal{A}^*, \mathcal{A}) \) is a norm closed self-adjoint algebra which separates points. So by the Stone-Weierstrass Theorem, it is all of \( \text{C}(X) \). So the semi-Dirichlet property is just the Dirichlet property for function algebras.

**Example 4.4.** The non-commutative disk algebras \( \mathfrak{A}_n \) are semi-Dirichlet. This is immediate from the relations 
\[
s_j^* s_i = \delta_{ij} I.
\]
Indeed, it is easy to see that all tensor algebras of directed graphs and tensor algebras of C*-correspondences are semi-Dirichlet. For those familiar with the terminology for the tensor algebra of a C*-correspondence \( E \) over a C*-algebra \( \mathfrak{A} \), the algebra \( \mathcal{T}^+(E) \) is generated by
\[
\sigma(\mathfrak{A}) \quad \text{and} \quad \{ T(\xi) : \xi \in E \},
\]
where \( \sigma \) and \( T \) are the canonical representations of \( \mathfrak{A} \) and \( E \), respectively, on the Fock space of \( E \). The relation
\[
T(\xi)^* T(\eta) = \sigma(\langle \xi, \eta \rangle)
\]
yields the same kind of cancellation as for the non-commutative disk algebra to show that
\[
\mathcal{T}^+(E)^* \mathcal{T}^+(E) \subset \overline{\mathcal{T}^+(E) + \mathcal{T}^+(E)^*}.
\]

**Example 4.5.** There is no converse to Proposition 4.2(ii). Consider the disk algebra \( \mathcal{A}(\mathbb{D}) \). The Toeplitz representation on \( H^2 \) given by \( \sigma(f) = T_f \), the Toeplitz operator with symbol \( f \), is completely isometric. This is Shilov, and so has the semi-Dirichlet property. This is also readily seen from the identity
\[
T_{f}^* T_{g} = T_{ef} \quad \text{for all } f, g \in \mathcal{A}(\mathbb{D}).
\]
However the representation
\[
\rho(f) = T_{f(z)}
\]
generated by \( \rho(z) = T_z^* \) is also completely isometric. However
\[
\rho(z)^* \rho(z) = T_z T_z^* = I - e_0 e_0^*.
\]
This is not a Toeplitz operator, and so is a positive distance from
\[
\rho(\mathcal{A}(\mathbb{D})) + \rho(\mathcal{A}(\mathbb{D}))^* = \overline{\{ T_{f+g} : f, g \in \mathcal{A}(\mathbb{D}) \} = \{ T_f : f \in \text{C}(\mathbb{T}) \}.}
\]

Documenta Mathematica 16 (2011) 781–868
Dilation Theory...

We will establish the following theorem.

**Theorem 4.6.** Suppose that $\mathcal{A}$ is a semi-Dirichlet unital operator algebra. Let $\rho$ be a representation of $\mathcal{A}$. Then $\rho$ has a unique minimal extremal coextension $\sigma$, it is fully extremal and cyclic (i.e. $\mathcal{K} = \sigma(\mathcal{A})\mathcal{H}$). Moreover, every Shilov representation is an extremal coextension.

We begin with a couple of lemmas.

**Lemma 4.7.** Suppose that $\mathcal{A}$ is a semi-Dirichlet unital operator algebra. Let $\rho$ be a representation of $\mathcal{A}$, and let $\sigma$ be a cyclic extremal coextension of $\rho$ on $\mathcal{K}$. Then $\sigma$ is fully extremal.

**Proof.** Suppose that $\tau$ is an extremal coextension of $\rho$ which is a dilation of $\sigma$. Say $\tau$ acts on $\mathcal{L} \supset \mathcal{K}$. Let $\pi$ be a fully extremal extension of $\tau$. Then $\pi$ is a maximal dilation of $\rho$ by Proposition 3.10. Moreover $\mathcal{L}$ is invariant for $\pi(\mathcal{A})$ because $\tau$ is extremal, as is $\mathcal{K}$ for the same reason. Therefore

$$\tau(a) = \pi(a)|_{\mathcal{L}} \quad \text{and} \quad \sigma(a) = \pi(a)|_{\mathcal{K}} = \tau(a)|_{\mathcal{K}}.$$

Also $\mathcal{H}$ is semi-invariant for $\pi(\mathcal{A})$ and coinvariant for $\tau(\mathcal{A})$ in $\mathcal{L}$.

If $\sigma$ is not a direct summand of $\tau$, then $\mathcal{K}$ is not coinvariant for $\tau(\mathcal{A})$. Thus there is a vector $x \in \mathcal{L} \ominus \mathcal{K}$ and $a \in \mathcal{A}$ so that

$$P_{\mathcal{K}}\tau(a)x \neq 0.$$

This vector in $\mathcal{K}$ can be approximated by a vector $\sigma(b)h$ for some $b \in \mathcal{A}$ and $h \in \mathcal{H}$ sufficiently well so that

$$\langle \tau(a)x, \sigma(b)h \rangle \neq 0.$$

Now $a^*b \in \mathcal{A}^*\mathcal{A}$ can be written as

$$a^*b = \lim_{n \to \infty} c_n + d_n^* \quad \text{where} \quad c_n, d_n \in \mathcal{A}.$$

Therefore

$$0 \neq \langle \tau(a)x, \sigma(b)h \rangle = \langle \pi(a)x, \pi(b)h \rangle = \lim_{n \to \infty} \langle x, \pi(c_n)h + \pi(d_n)^*h \rangle = \lim_{n \to \infty} \langle x, \tau(d_n)^*h \rangle.$$

Here we used the fact that

$$\pi(c_n)h = \sigma(c_n)h \in \mathcal{K},$$

which is orthogonal to $x$, and then the fact that the compression of $\pi(d_n)^*$ to $\mathcal{L}$ is $\tau(d_n)^*$. This calculation shows that $\mathcal{H}$ is not coinvariant for $\tau$, contrary to our hypothesis. This means that $\tau$ does indeed have $\sigma$ as a direct summand. So $\sigma$ is fully extremal.

\[ \square \]
K.R. Davidson and E.G. Katsoulis

Lemma 4.8. Suppose that $\mathcal{A}$ is a semi-Dirichlet unital operator algebra. Let $\rho$ be a representation of $\mathcal{A}$. Then any two cyclic Shilov coextensions $\sigma_i$ of $\rho$, $i = 1, 2$, on $\mathcal{K}_i$ are equivalent. Hence a cyclic Shilov coextension of $\rho$ is fully extremal.

Proof. Let $\sigma_i$, $i = 1, 2$, be two minimal cyclic Shilov coextensions of $\rho$ on $\mathcal{K}_i$; so that $\mathcal{K}_i = \sigma_i(\mathcal{A})|\mathcal{H}$. Let $\pi_i$ be the maximal dilations of $\rho$ on $\mathcal{L}_i \supset \mathcal{K}_i$ such that $\mathcal{K}_i$ is invariant and $\pi_i(a)|\mathcal{K}_i = \sigma_i(a)$ for $a \in \mathcal{A}$. The idea is to follow the standard proof by showing that there is a map $U \in B(\mathcal{K}_1, \mathcal{K}_2)$ given by

$$U\sigma_1(a)h = \sigma_2(a)h$$

which is a well defined isometry of $\mathcal{K}_1$ onto $\mathcal{K}_2$. To this end, it suffices to verify that

$$\langle \sigma_1(a_1)h_1, \sigma_1(a_2)h_2 \rangle = \langle \sigma_2(a_1)h_1, \sigma_2(a_2)h_2 \rangle$$

for all $a_1, a_2 \in \mathcal{A}$ and $h_1, h_2 \in \mathcal{H}$.

By hypothesis, we can find $b_n, c_n \in \mathcal{A}$ so that

$$a_2^*a_1 = \lim_{n \to \infty} b_n + c_n^*.$$

We calculate

$$\langle \sigma_1(a_1)h_1, \sigma_1(a_2)h_2 \rangle = \langle \pi_i(a_1)h_1, \pi_i(a_2)h_2 \rangle$$

$$= \langle \pi_i(a_2^*a_1)h_1, h_2 \rangle$$

$$= \lim_{n \to \infty} \langle \pi_i(b_n) + \pi_i(c_n)^* \rangle h_1, h_2 \rangle$$

$$= \lim_{n \to \infty} \langle \rho(b_n) + \rho(c_n)^* \rangle h_1, h_2 \rangle.$$

This quantity is independent of the dilation, and thus $U$ is a well-defined isometry.

Since $\mathcal{K}_i = \sigma_i(\mathcal{A})|\mathcal{H}$, it follows that $U$ is unitary. It is also evident that $U|\mathcal{H}$ is the identity map. So $U$ is the desired unitary equivalence of $\sigma_1$ and $\sigma_2$.

Since $\mathcal{A}$ always has a cyclic extremal coextension $\sigma$, it must be the unique cyclic Shilov coextension. By Lemma 4.7, $\sigma$ is fully extremal. \hfill \blacksquare

Proof of Theorem 4.6. Let $\tau$ be any minimal extremal coextension of $\rho$ on $\mathcal{L} \supset \mathcal{H}$. Set

$$\mathcal{K} = \overline{\tau(\mathcal{A})|\mathcal{H}} \quad \text{and} \quad \sigma = \tau|\mathcal{K}.$$

Also let $\pi$ be a fully extremal extension of $\tau$. By Proposition 3.10, $\pi$ is a maximal dilation of $\rho$. Since $\mathcal{L}$ is invariant for $\pi(\mathcal{A})$ and $\mathcal{K}$ is invariant for $\tau(\mathcal{A})$, it follows that $\mathcal{K}$ is invariant for $\pi(\mathcal{A})$. Hence $\sigma$ is Shilov. By Lemma 4.8, $\sigma$ is fully extremal. It follows that $\tau = \sigma \oplus \tau'$. However $\tau$ is minimal. So

$$\tau = \sigma \quad \text{and} \quad \mathcal{L} = \mathcal{K} = \overline{\tau(\mathcal{A})|\mathcal{H}}.$$

Documenta Mathematica 16 (2011) 781–868
Hence $\tau$ is cyclic. By Lemma 4.8, $\tau$ is unique.

Now let $\sigma$ be a Shilov representation of $A$. Let $\tau$ be a cyclic extremal coextension of $\sigma$. By Lemma 4.8, $\sigma$ and $\tau$ are equivalent coextensions of $\sigma$. Therefore $\tau = \sigma$. Thus $\sigma$ is extremal.

The consequences for Dirichlet algebras are apparent.

**Corollary 4.9.** If $A$ is a Dirichlet operator algebra, then every Shilov extension and every Shilov coextension is fully extremal; and the minimal extremal (co-)extension of a representation is unique. Moreover the minimal maximal dilation of a representation is unique.

**Proof.** The first statement is immediate from Theorem 4.6 and its dual. Let $\pi$ be a minimal maximal dilation of a representation $\rho$ on a Hilbert space $\mathcal{L}$. Let $K = \pi(A)\mathcal{H}$. This is the minimal Shilov subspace containing $\mathcal{H}$. Thus by Theorem 4.6, it is the unique minimal fully extremal coextension of $\rho$. Let $\mathcal{L}_0 = \pi(A)^*K$. This is the minimal Shilov extension of $\sigma$. Hence by the dual of Theorem 4.6, this coincides with the unique minimal fully extremal extension of $\sigma$. By Corollary 3.10, the restriction of $\pi$ to $\mathcal{L}_0$ is a maximal dilation. Since $\pi$ is minimal, $\mathcal{L}_0 = \mathcal{L}$. So $\pi$ is obtained by taking the unique minimal extremal coextension of $\rho$ to get $\sigma$, followed by the unique minimal extremal extension of $\sigma$. So $\pi$ is uniquely determined.

While semi-Dirichlet algebras behave exceptionally well for coextensions, they are not nearly so well behaved for extensions.

**Example 4.10.** We consider extensions for representations of the non-commutative disk algebra $\mathfrak{A}_n$. Denote the generators by $s_1, \ldots, s_n$, and write $s = [s_1 \ldots s_n]$. A representation $\rho$ on $\mathcal{H}$ is determined by a row contraction $A = \rho(s) = [\rho(s_1) \ldots \rho(s_n)] =: [A_1 \ldots A_n]$, where $\|A\| = \left\| \sum A_i A_i^* \right\|^{1/2} \leq 1$. We have seen that $A$ has a unique minimal coextension to a row isometry, and this is the unique minimal fully extremal coextension.

Now consider an extension $\sigma$ of $\rho$ acting on $K$. This correspond to simultaneous extensions of $A_i$ to

$$B_i = \sigma(s_i) = \begin{bmatrix} A_i & B_{i,12} \\ 0 & B_{i,22} \end{bmatrix}$$

such that $B = [B_1 \ldots B_n]$ is a row contraction. It is straightforward to verify that it is extremal if and only if $B$ is a coisometry. We claim that: an extension $\sigma$ of $\rho$ is fully extremal if and only if $B$ is a coisometry such that

$$\text{Ran } B^* \vee \mathcal{H}(n) = \mathcal{K}^{(n)}.$$
Indeed, if this condition holds, then there is no proper extension of $B$; so consider any row contractive coextension $C$ of $B$ which is an extension of $A_i$. Then $C_i^*$ are extensions of $B_i^*$ which are coextensions of $A_i^*$. So

$$C^* = \begin{bmatrix} B_i^* & X \\ 0 & Y \end{bmatrix} = \begin{bmatrix} A_i^* & 0 & 0 \\ B_{i2} & B_{22} & X_2 \\ 0 & 0 & Y \end{bmatrix}.$$ 

Since $B_i^*$ is an isometry, we require that Ran $X$ be orthogonal to Ran $B_i^*$. And since $C$ is an extension of $A$, we have Ran $X$ is orthogonal to $\mathcal{H}^{(n)}$. Therefore by hypothesis, Ran $X$ is orthogonal to $\mathcal{K}^{(n)}$, and thus $X = 0$. Therefore $C = B \oplus Y$ is a direct sum.

Conversely, suppose that there is a unit vector $x = (x_1, \ldots, x_n)^t$ in $\mathcal{K}^{(n)}$ which is orthogonal to Ran $B_i^* \vee \mathcal{H}^{(n)}$. Define an extension of $B_i^*$ to $\mathcal{K} \oplus \mathcal{C}$ by

$$C_i^* = \begin{bmatrix} B_i^* & x_i \\ 0 & 0 \end{bmatrix}.$$ 

Since $x_i \in \mathcal{H}^\perp$, this is a coextension of $A_i^*$. So $C$ determines an extension of $A$ which is a coextension of $B$. Clearly it does not split as a direct sum. Finally, $C$ is a coisometry because $C^* = \begin{bmatrix} B_i^* & 0 \\ 0 & 0 \end{bmatrix}$ is an isometry. In particular, $C$ is a row contraction.

Next we observe that the minimal fully extremal extensions are far from unique in general by showing how to construct a fully extremal coextension.

Start with $A$ which is not coisometric. Then

$$D = (I - \sum A_i A_i^*)^{1/2} \neq 0.$$ 

Consider a fully extremal extension $B$ as above. Then $B$ is a coisometry on $\mathcal{K} = \mathcal{H} \oplus \mathcal{K}_0$; whence $[A B_{12}]$ is a coisometry in $B(\mathcal{K}, \mathcal{H})$. Therefore

$$I_{\mathcal{H}} = [A B_{12}] [A B_{12}]^* = AA^* + B_{12}B_{12}^*.$$ 

Hence

$$B_{12}B_{12}^* = D^2,$$ 

and therefore $B_{12} = DX$ where $X = [X_1 \ldots X_n]$ is a coisometry in $B(\mathcal{K}_0^{(n)}, \mathcal{H})$.

Let $\mathcal{R} = \text{Ran } X^*$. Then to be fully extremal, we have that $B_{22}^*$ is an isometry from $\mathcal{K}_0$ onto $\mathcal{R}^\perp \subset \mathcal{K}_0^{(n)}$. Now let $V$ be any isometry in $B(\mathcal{K}_0^{(n)})$ with Ran $V = \mathcal{R}^\perp$. Then $V^* B_{22}^*$ is a unitary in $B(\mathcal{K}_0, \mathcal{K}_0^{(n)})$. Decompose the unitary

$$S := B_{22}V = [S_1 \ldots S_n]$$ 

where $S_i \in B(\mathcal{K}_0)$. Observe that $S_i$ are isometries such that

$$\sum_i S_i S_i^* = I;$$
in other words they are Cuntz isometries. Since
\[ B_{22} = B_{22}VV^* = SV^* \]
in \( B(K_0^{(n)}, K_0) \), we decompose this as
\[ B_{22} = SV^* = [T_1 \ldots T_n]. \]
We obtain
\[ B_i = [A_i B_{i,12}] = [A_i DX_i^*]. \]
Conversely, if we choose any coisometry \( X \) in \( B(K_0^{(n)}, \mathcal{H}) \), we may define \( \mathcal{R} = \text{Ran } X^* \), choose an isometry \( V \) in \( B(K_0^{(n)}) \) with \( \text{Ran } V = \mathcal{R}^+ \), and a set of Cuntz isometries \( S_i \) in \( B(K_0) \), then the formulae above yield a fully extremal extension. This may not be minimal in general, but the restriction to the smallest reducing subspace containing \( H \) is a minimal fully extremal extension. This restriction will not change \( X \). So if two minimal fully extremal extensions are equivalent, then at the very least, there is a unitary \( U \in B(K_0) \) so that \( XU = X' \). It is easy to see that there are many inequivalent choices for \( X \) even if \( D \) is rank one.

5 Commutant Lifting

Many variants of the commutant lifting theorem have been established for a wide range of operator algebras. They differ somewhat in the precise assumptions and conclusions. The general formulation in Douglas-Paulsen [26] and Muhly-Solel [40] uses Shilov modules. But we will formulate it using only fully extremal coextensions. The second definition is motivated by the lifting results of Paulsen-Power [47].

**Definition 5.1.** An operator algebra \( \mathcal{A} \) has the **strong commutant lifting property** (SCLT) if whenever \( \rho \) is a completely contractive representation of \( \mathcal{A} \) on \( \mathcal{H} \) with a fully extremal coextension \( \sigma \) on \( \mathcal{K} \supset \mathcal{H} \), and \( X \) commutes with \( \rho(\mathcal{A}) \), then \( X \) has a coextension \( Y \) in \( B(\mathcal{K}) \) with \( \|Y\| = \|X\| \) which commutes with \( \sigma(\mathcal{A}) \).

An operator algebra \( \mathcal{A} \) has the **commutant lifting property** (CLT) if whenever \( \rho \) is a completely contractive representation of \( \mathcal{A} \) on \( \mathcal{H} \) and \( X \) commutes with \( \rho(\mathcal{A}) \), then \( \rho \) has a fully extremal coextension \( \sigma \) on \( \mathcal{K} \supset \mathcal{H} \) and \( X \) has a coextension \( Y \) in \( B(\mathcal{K}) \) with \( \|Y\| = \|X\| \) which commutes with \( \sigma(\mathcal{A}) \).

An operator algebra \( \mathcal{A} \) has the **weak commutant lifting property** (WCLT) if whenever \( \rho \) is a completely contractive representation of \( \mathcal{A} \) on \( \mathcal{H} \) and \( X \) commutes with \( \rho(\mathcal{A}) \), then \( \rho \) has an extremal coextension \( \sigma \) on \( \mathcal{K} \supset \mathcal{H} \) and \( X \) has a coextension \( Y \) in \( B(\mathcal{K}) \) with \( \|Y\| = \|X\| \) which commutes with \( \sigma(\mathcal{A}) \).
The important distinction is that in SCLT, the coextension is prescribed first, while in CLT, it may depend on $X$.

It is clear that the more that we restrict the class of coextensions for which we have strong commutant lifting, the weaker the property. Thus SCLT using only fully extremal coextensions is asking for less than using all extremal coextensions, which in turn is weaker than using all Shilov extensions. As we will want a strong commutant lifting theorem, it behooves us to limit the class of extensions. On the other hand, as we limit the class of coextensions, the property CLT becomes stronger.

Observe that for the SCLT and CLT, it suffices to consider minimal fully extremal extensions. This is because any fully extremal extension decomposes as $\sigma = \sigma_0 \oplus \tau$ where $\sigma_0$ is minimal. Any operator $Y$ commuting with $\sigma(A)$ will have a 1,1 entry commuting with $\sigma_0(A)$.

**Example 5.2.** The disk algebra has SCLT by the Sz.Nagy–Foiaş Commutant Lifting Theorem [59]. In fact, as noted above, any isometric dilation is an extremal coextension, but not all are fully extremal. So $A(D)$ has the SCLT with respect to the larger class of all extremal coextensions, and these are all of the Shilov extensions. The reason it works is that every isometric coextension splits as $\sigma_0 \oplus \tau$ where $\sigma_0$ is the unique minimal isometric coextension.

**Example 5.3.** The non-commutative disk algebra $A_n$ also has SCLT by Popescu’s Commutant Lifting Theorem [50]. As noted in Example 2.2, the extremal coextensions are the row isometric ones, and these are Shilov. As in the case of the disk algebra, it is only fully extremal if it is the direct sum of the minimal isometric coextension with a row unitary.

More generally, the tensor algebra of any C*-correspondence has SCLT by the Muhly-Solel Commutant Lifting Theorem [42].

**Example 5.4.** The bidisk algebra $A(D^2)$ does not have WCLT, since commutant lifting implies the simultaneous unitary dilation of three commuting contractions [60, 44].

**Example 5.5.** The algebra $A_n$ of continuous multipliers on symmetric Fock space (see Example 2.3) has SCLT [21]. The extremal extensions are in fact maximal dilations, and so in particular are fully extremal.

The relationship between SCLT and CLT is tied to uniqueness of minimal coextensions. We start with an easy lemma.

**Lemma 5.6.** Let $A$ be a unital operator algebra. Suppose that $\sigma$ is a minimal dilation on $K$ of a representation $\rho$ on $\mathcal{H}$, in the sense that $K$ is the smallest reducing subspace for $\sigma(A)$ containing $\mathcal{H}$. If $X$ is a contraction commuting with $\sigma(A)$ such that $P_{\mathcal{H}}X|_{\mathcal{H}} = I$, then $X = I$. 

*Documenta Mathematica* 16 (2011) 781–868
Proof. Since $\|X\| = 1$ and $P_HX|_H = I$, $X$ reduces $\mathcal{H}$. Note that for all $h \in \mathcal{H}$ and $a \in \mathcal{A}$,

$$X \sigma(a)h = \sigma(a)Xh = \sigma(a)h$$

and

$$X^* \sigma(a)^*h = \sigma(a)^*X^*h = \sigma(a)^*h.$$ 

So the restriction of $X$ to $\sigma(\mathcal{A}) \mathcal{H}$ is the identity. As $X$ is a contraction, it reduces this space. Similarly, the restriction of $X^*$ to $\sigma(\mathcal{A})^* \mathcal{H}$ is the identity; and $X$ reduces this space as well. Recursively we may deduce that $X$ is the identity on the smallest reducing subspace containing $\mathcal{H}$, which is $K$.

**Theorem 5.7.** Let $\mathcal{A}$ be a unital operator algebra. Then $\mathcal{A}$ has SCLT if and only if it has CLT and unique minimal fully extremal coextensions.

**Proof.** Assume first that $\mathcal{A}$ has CLT and unique minimal fully extremal coextensions. Let $\rho$ be a representation of $\mathcal{A}$ on $\mathcal{H}$ with a fully extremal coextension $\sigma$ on $K \supset \mathcal{H}$, and suppose that $X$ commutes with $\rho(\mathcal{A})$. By CLT, there is a fully extremal coextension $\tau$ on $L \supset \mathcal{H}$ and $X$ has a coextension $Z$ in $B(L)$ with $\|Z\| = \|X\|$ which commutes with $\tau(\mathcal{A})$. By uniqueness of minimal fully extremal coextensions, there is a fully extremal coextension $\mu$ on $K_0$ so that

$$\sigma \simeq \mu \oplus \sigma' \quad \text{and} \quad \tau \simeq \mu \oplus \tau'.$$

With respect to the latter decomposition, $Z$ can be written as a $2 \times 2$ matrix commuting with $\mu(a) \oplus \tau'(a)$ for all $a \in \mathcal{A}$. Moreover, the corner entry $Z_{11}$ is a coextension of $X$. A simple calculation of the commutator shows that $\mu(a)$ commutes with $Z_{11}$. Thus $Y \simeq Z_{11} \oplus 0$ is the desired coextension of $X$ commuting with $\sigma(\mathcal{A})$.

Conversely, suppose that $\mathcal{A}$ has SCLT. A fortiori, it will have CLT. Suppose that a representation $\rho$ on $\mathcal{H}$ has two minimal fully extremal coextensions $\sigma_1$ and $\sigma_2$ on $K_1 = \mathcal{H} \oplus K'_1$ and $K_2 = \mathcal{H} \oplus K'_2$, respectively. Then $\rho \oplus \rho$ has $\sigma_1 \oplus \sigma_2$ as a fully extremal coextension. This can be seen, for example, because of the identities (1) and (2) in the proof of Theorem 3.5. The operator $X = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$ commutes with $(\rho \oplus \rho)(\mathcal{A})$. So by SCLT, $X$ has a coextension $Y$ on $K_1 \oplus K_2$ of norm 1 which commutes with $(\sigma_1 \oplus \sigma_2)(\mathcal{A})$. Since $P_{H \oplus H}Y|_{H \oplus H} = X$ is unitary, $Y$ reduces $\mathcal{H} \oplus \mathcal{H}$.

Now $Y^2$ commutes with $\sigma_1 \oplus \sigma_2(\mathcal{A})$ and its restriction to $\mathcal{H} \oplus \mathcal{H}$ is $X^2 = I$. Thus by Lemma 5.6, $Y^2 = I$. In particular, $Y$ is unitary. Let

$$Y_{12} = P_{K_1}Y|_{K_2} \quad \text{and} \quad Y_{21} = P_{K_2}Y|_{K_1}.$$ 

Observe that for $a \in \mathcal{A}$,

$$Y_{21} \sigma_1(a) = \sigma_2(a)Y_{21} \quad \text{and} \quad Y_{12} \sigma_2(a) = \sigma_1(a)Y_{12}.$$ 

Documenta Mathematica 16 (2011) 781–868
Moreover the restriction of $Y_{21}$ to $\mathcal{H}$ is $X$ restricted to $\mathcal{H} \oplus \{0\}$, which is the identity map if we identify $\mathcal{H} \oplus \{0\}$ and $\{0\} \oplus \mathcal{H}$ with $\mathcal{H}$. We deduce that $Y_{12}Y_{21}$ commutes with $\sigma_1(A)$ and coincides with $I$ on $\mathcal{H}$. So by Lemma 5.6, $Y_{12}Y_{21} = I$.

Similarly, $Y_{21}Y_{12} = I$.

Since they are contractions, $Y_{12}$ is unitary and $Y_{21} = Y_{12}^*$. The identities above now show that $Y_{21}$ implements a unitary equivalence between $\sigma_1$ and $\sigma_2$ fixing $\mathcal{H}$. Hence the minimal fully extremal coextension of $\rho$ is unique.

We can weaken CLT to WCLT if we strengthen the uniqueness hypothesis to minimal extremal coextensions. This seems a fair bit stronger in comparison however.

**Corollary 5.8.** If $A$ has WCLT and unique minimal extremal coextensions, then $A$ has SCLT.

*Proof.* If $\rho$ is a representation of $A$, then the unique minimal extremal coextension $\sigma$ of $\rho$ must be fully extremal, since by Theorem 3.5, there are fully extremal coextensions and hence there are minimal ones. These are also minimal as extremal coextensions. Thus $A$ has unique minimal fully extremal coextensions. Moreover, as in the proof above, if a contraction $X$ commutes with $\rho(A)$, then WCLT provides a coextension to a contraction $Y$ commuting with an extremal coextension $\tau$. But $\tau = \sigma \oplus \tau'$. So arguing as before, the compression $Z$ of $Y$ to $K_{\sigma}$ commutes with $\sigma(A)$ and is a coextension of $X$. Now if $\varphi$ is an arbitrary extremal coextension of $\rho$, again split $\varphi = \sigma \oplus \varphi'$. One extends $Z$ to $Z \oplus 0$ to commute with $\varphi(A)$.

It is common to look for a version of commutant lifting for intertwining maps between two representations. In the case of WCLT and SCLT, this is straightforward. Such a version for CLT is valid here too, but some care must be taken.

**Proposition 5.9.** Suppose that $A$ has SCLT. Let $\rho_i$ be representations of $A$ on $\mathcal{H}_i$ for $i = 1, 2$ with fully extremal coextensions $\sigma_i$ on $K_{\mathcal{H}_i}$. Suppose that $X$ is a contraction in $B(\mathcal{H}_2, \mathcal{H}_1)$ such that $\rho_1(a)X = X\rho_2(a)$ for all $a \in A$.

Then there is a contraction $Y$ in $B(K_2, K_1)$ so that $P_{\mathcal{H}_1}Y = XP_{\mathcal{H}_2}$ and $\sigma_1(a)Y = Y\sigma_2(a)$ for all $a \in A$.  

*Documenta Mathematica* 16 (2011) 781–868
Proof. Let $\rho = \rho_1 \oplus \rho_2$. By Corollary 3.9, $\sigma = \sigma_1 \oplus \sigma_2$ is a fully extremal coextension of $\rho$. Observe that

$$
\tilde{X} = \begin{bmatrix} 0 & X \\ 0 & 0 \end{bmatrix}
$$

commutes with $\rho(A)$. Hence by SCLT, there is a coextension $\tilde{Y}$ of $\tilde{X}$ which commutes with $\sigma(A)$.

Write $Y$ as a matrix with respect to

$$
K = K_1 \oplus K_2 = H_1 \oplus H_2 \oplus (K_1 \oplus H_1) \oplus (K_2 \oplus H_2)
$$

and rearrange this to the decomposition

$$
K = H_1 \oplus (K_1 \oplus H_1) \oplus H_2 \oplus (K_2 \oplus H_2)
$$

We obtain the unitary equivalence

$$
\tilde{Y} = 
\begin{bmatrix} \tilde{X} & 0 \\ * & * \end{bmatrix} = 
\begin{bmatrix} 0 & X & 0 & 0 \\ 0 & 0 & 0 & 0 \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \simeq 
\begin{bmatrix} 0 & 0 & X & 0 \\ 0 & 0 & 0 & 0 \\ * & * & * & * \\ * & * & * & * \end{bmatrix}
$$

Restricting to the upper right $2 \times 2$ corner, we obtain

$$
Y := P_{K_1} \tilde{Y}|_{K_2} = 
\begin{bmatrix} X & 0 \\ * & * \end{bmatrix}
$$

Then $Y$ is a contraction, and as an operator in $B(K_2, H_1)$ we have

$$
P_{H_1} Y = \begin{bmatrix} X & 0 \end{bmatrix} = X P_{H_2}.
$$

Finally the commutation relations show that

$$
\sigma_1(a) Y = Y \sigma_2(a) \quad \text{for all } a \in A.
$$

A similar argument shows the following:

**Proposition 5.10.** Suppose that $A$ has WCLT. Let $\rho_i$ be representations of $A$ on $H_i$ for $i = 1, 2$. Suppose that $X$ is a contraction in $B(H_2, H_1)$ such that

$$
\rho_1(a) X = X \rho_2(a) \quad \text{for all } a \in A.
$$

Then there are extremal coextensions $\sigma_i$ of $\rho_i$ acting on $K_i \supset H_i$ for $i = 1, 2$ and a contraction $Y$ in $B(K_2, K_1)$ so that

$$
P_{H_1} Y = X P_{H_2}
$$

and

$$
\sigma_1(a) Y = Y \sigma_2(a) \quad \text{for all } a \in A.
$$
Proof. Again form $\rho = \rho_1 \oplus \rho_2$ and $\tilde{X}$ as above. Use WCLT to coextend $\rho$ to an extremal coextension $\sigma$ and $\tilde{X}$ to a contraction $\tilde{Y}$ commuting with $\sigma(\mathcal{A})$ on $\mathcal{K} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{K}'$. Now notice that $\sigma$ is an extremal coextension of both $\rho_i$. Considering $\tilde{Y}$ as a map from $\mathcal{H}_1 \oplus (\mathcal{H}_2 \oplus \mathcal{K}')$ to $\mathcal{H}_2 \oplus (\mathcal{H}_1 \oplus \mathcal{K}')$, one finds that it has a matrix form

$$\tilde{Y} = \begin{bmatrix} X & 0 & 0 \\ 0 & 0 & 0 \\ \ast & \ast & \ast \end{bmatrix}.$$ 

This has the desired form.

Remark 5.11. The issue with CLT and fully extremal coextensions is that a fully extremal coextension of $\rho = \rho_1 \oplus \rho_2$ need not even contain a fully extremal coextension of $\rho_i$ as a summand. Consider the subalgebra $\mathcal{A} \subset \mathfrak{M}_5$ given by $\mathcal{A} = \text{span}\{E_{21}, E_{32}, E_{31}, E_{45}, E_{35}, E_{ii} : 1 \leq i \leq 5\}$. Let $\rho_1(a) = E_{22}a|_{C_{e_2}}$ and $\rho_2(a) = E_{44}a|_{C_{e_4}}$.

The minimal fully extremal coextensions of $\rho_i$ are $\sigma_1(a) = E_{11}^+a|_{(C_{e_1})^\perp}$ and $\sigma_2(a) = E_{55}^+a|_{(C_{e_5})^\perp}$.

However the minimal fully extremal coextension of $\rho_1 \oplus \rho_2$ is $\sigma(a) = (E_{22} + E_{33} + E_{44})a|_{\text{span}\{e_2, e_3, e_4\}}$.

Thus a proof of the following result must follow different lines. This proof has its roots in the work of Sz.Nagy and Foiaş. Notice that it allows a specification of one of the coextensions. Normally we will use a fully extremal coextension $\sigma_2$ of $\rho_2$.

Theorem 5.12. Suppose that $\mathcal{A}$ has CLT. Let $\rho_i$ be representations of $\mathcal{A}$ on $\mathcal{H}_i$ for $i = 1, 2$. Suppose that $X$ is a contraction in $\mathcal{B}(\mathcal{H}_2, \mathcal{H}_1)$ such that $\rho_1(a)X = X\rho_2(a)$ for all $a \in \mathcal{A}$.

Let $\sigma_2$ be an extremal coextension of $\rho_2$ on $\mathcal{K}_2$. Then there is a fully extremal coextension $\sigma_1$ of $\rho_1$ acting on $\mathcal{K}_1$ and a contraction $Y$ in $\mathcal{B}(\mathcal{K}_2, \mathcal{K}_1)$ so that $P_{\mathcal{H}_1}Y = XP_{\mathcal{H}_2}$ and $\sigma_1(a)Y = Y\sigma_2(a)$ for all $a \in \mathcal{A}$.
Proof. Let $K_2 = H_2 \oplus K'_2$, and decompose

$$\sigma_2(a) = \begin{bmatrix} \rho_2(a) & 0 \\ \sigma_{21}(a) & \sigma_{22}(a) \end{bmatrix}$$

Observe that $[X \ 0] \in B(K_2, H_1)$ satisfies

$$\rho_1(a)[X \ 0] = [\rho_1(a)X \ 0] = [X \ 0] \begin{bmatrix} \rho_2(a) & 0 \\ \sigma_{21}(a) & \sigma_{22}(a) \end{bmatrix}.$$ 

Therefore, $\tilde{X} = \begin{bmatrix} 0 & X & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ commutes with the range of $\rho = \rho_1 \oplus \sigma_2$.

Now we apply the CLT property to $\rho$ and $\tilde{X}$ to obtain a fully extremal coextension $\tau$ of $\rho$ and contraction $\tilde{Y}$ coextending $\tilde{X}$ on $\mathcal{L} = H_1 \oplus K_2 \oplus \mathcal{L}'$ which commute. We may write

$$\tau(a) = \begin{bmatrix} \rho_1(a) & 0 & 0 \\ 0 & \sigma_2(a) & 0 \\ \tau_{31}(a) & \tau_{32}(a) & \tau_{33}(a) \end{bmatrix}.$$ 

Observe that the lower right $2 \times 2$ corner is a coextension of $\sigma_2$. Since $\sigma_2$ is extremal, we see that $\tau_{32} = 0$. Define

$$\sigma_1(a) = \begin{bmatrix} \rho_1(a) & 0 \\ \tau_{31}(a) & \tau_{33}(a) \end{bmatrix}.$$ 

To complete the proof, we need to establish that $\sigma_1$ is a fully extremal coextension of $\rho_1$. Suppose that $\gamma$ is a representation of $\mathcal{A}$ which dilates $\sigma_1$ and coextends $\rho_1$. Then $\gamma \oplus \sigma_2$ dilates $\sigma_1 \oplus \sigma_2 = \tau$ and coextends $\rho_1 \oplus \sigma_2$. Since $\tau$ is fully extremal, $\gamma = \sigma_1 \oplus \gamma'$ as desired.

We make a few more definitions. (Apologies for all the acronyms.)

**Definition 5.13.** If $\mathcal{A}^*$ has SCLT, CLT or WCLT, we say that $\mathcal{A}$ has SCLT*, CLT* or WCLT*.

Say that $\mathcal{A}$ has maximal commutant lifting (MCLT) if for every representation $\rho$ on $\mathcal{H}$ and contraction $X$ commuting with $\rho(\mathcal{A})$, there is a maximal dilation $\pi$ of $\rho$ on a Hilbert space $K \supset \mathcal{H}$ and a contraction $Y$ commuting with $\pi(\mathcal{A})$ such that

$$P_H(\pi(a)Y^n)|_H = \rho(a)X^n$$

for all $a \in \mathcal{A}$ and $n \geq 0$.

If the maximal dilation $\pi$ can be specified a priori, then say that $\mathcal{A}$ has strong maximal commutant lifting (SMCLT).
It is clear that the commutant lifting properties for $A^*$ can be interpreted as lifting commutants to (fully) extremal extensions instead of coextensions. On rare occasions, one gets both. For example, the disk algebra $A(D)$ is completely isometrically isomorphic to its adjoint algebra. Hence it has both SCLT and SCLT*.

The property MCLT for $A$ is equivalent to MCLT for $A^*$. So we will not have a property MCLT*.

The definition of MCLT contains the information that the compression of the algebra generated by $\pi(A)$ and $Y$ to $H$ is an algebra homomorphism which sends $\pi$ to $\rho$ and $Y$ to $X$. It follows that $H$ is semi-invariant for this algebra; i.e. $H$ is the difference of two subspaces which are invariant for both $\pi(A)$ and $Y$.

**Theorem 5.14.** If $A$ has WCLT and WCLT*, then $A$ has MCLT.

**Proof.** One uses WCLT to coextend $\rho$ on $H$ to an extremal coextension $\sigma_1$ on $K_1$ and coextend $X$ to a contraction $Z_1 \in B(K_1)$ commuting with $\sigma_1(A)$. Then use WCLT* extend $\sigma_1$ to an extremal extension $\tau_1$ on $L_1$, and lift $Y_1$ to a contraction $Z_1 \in B(L_1)$ commuting with $\tau_1(A)$. Alternate these procedures, obtaining an extremal coextension $\sigma_{n+1}$ of $\tau_n$ on $K_{n+1}$ and a contractive coextension $Y_{n+1} \in B(K_{n+1})$ in the commutant of $\sigma_{n+1}(A)$; and then extending $\sigma_{n+1}$ to an extremal $\tau_{n+1}$ on $L_{n+1}$ and extending $Y_{n+1}$ to a contraction $Z_{n+1} \in B(L_{n+1})$ in the commutant of $\tau_{n+1}(A)$. It is easy to see that at every stage, the original space $H$ is semi-invariant for both the representation and the contraction—so that these are always simultaneous dilations of the representation and the contraction.

Moreover, we can write $\sigma_{n+1}$ as a dilation of $\sigma_n$ in the matrix form relative to

$$K_{n+1} = (L_n \oplus K_n) \oplus K_n \oplus (K_{n+1} \oplus L_n)$$

as

$$\sigma_{n+1} = \begin{bmatrix} * & 0 & 0 \\ * & \sigma_n & 0 \\ * & * & * \end{bmatrix}$$

where the upper left $2 \times 2$ corner represents $\tau_n$. The lower right $2 \times 2$ corner is a coextension of $\sigma_n$. Since $\sigma_n$ is an extremal coextension, the $3,2$ entry is 0. Rearranging this as

$$K_{n+1} = K_n \oplus (L_n \oplus K_n) \oplus (K_{n+1} \oplus L_n),$$

we have

$$\sigma_{n+1} = \begin{bmatrix} \sigma_n & * & 0 \\ 0 & * & 0 \\ 0 & * & * \end{bmatrix}.$$

**Documenta Mathematica** 16 (2011) 781–868
A similar analysis holds for the $\tau_n$. Therefore, with respect to 
$$K_1 \oplus (L_1 \ominus K_1) \ominus (K_2 \ominus L_1) \ominus (L_2 \ominus K_2) \ominus \ldots,$$
these representations have a tridiagonal form
$$\begin{bmatrix}
\sigma_1 & * & 0 & 0 & 0 & 0 & \ldots \\
0 & * & 0 & 0 & 0 & 0 & \ldots \\
0 & * & * & 0 & 0 & 0 & \ldots \\
0 & 0 & 0 & * & 0 & 0 & \ldots \\
0 & 0 & 0 & * & * & 0 & \ldots \\
0 & 0 & 0 & 0 & * & 0 & \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{bmatrix}.$$ 

It follows that the direct limit $\pi$ exists as a sot- limit. Moreover, as $\pi$ is a limit of extremal coextensions, it is an extremal coextension; and similarly it is an extremal extension. Thus it is a maximal dilation.

The operators $Y_n$ and $Z_n$ each leave the subspaces $\mathcal{H}$ semi-invariant, and the restriction of $Z_n$ to $K_n$ is $Y_n$, and the compression (actually a co-restriction) of $Y_{n+1}$ to $L_n$ is $Z_n$. Therefore the direct limit $Y$ exists as a wot limit. It follows that $Y$ is a contraction that commutes with $\pi(A)$. To see this, let $x_n \in K_n$ and $y_n \in L_n$. Then for $a \in A$,
$$\pi(a)x_n = \sigma_k(a)x_n \quad \text{for all } k \geq n.$$ 

Similarly,
$$\pi(a)^*y_n = \sigma_k(a)^*y_n \quad \text{for all } k \geq n + 1.$$ 

So we can compute:
$$\langle (\pi(a)Y - Y\pi(a))x_n, y_n \rangle = \langle x_n, Y^*\pi(a)^*y_n \rangle - \langle Y\pi(a)x_n, y_n \rangle$$
$$= \lim_{k \to \infty} \langle x_n, Y_k^*\pi(a)^*y_n \rangle - \langle Y_k\pi(a)x_n, y_n \rangle$$
$$= \lim_{k \to \infty} \langle x_n, Y_k^*\sigma_k(a)^*y_n \rangle - \langle Y_k\sigma_k(a)x_n, y_n \rangle$$
$$= \lim_{k \to \infty} \langle (\sigma_k(a)Y_k - Y_k\sigma_k(a))x_n, y_n \rangle = 0.$$ 

So we have obtained the desired commutant lifting. 

We can modify the proof of Theorem 5.7 characterizing SCLT to characterize SMCLT.

**Theorem 5.15.** The following are equivalent for $A$:

(i) $A$ has SMCLT.

(ii) $A$ has MCLT and unique minimal maximal dilations.
Proof. Suppose that $\mathcal{A}$ has SMCLT. Then evidently it has MCLT. Moreover, suppose that a representation $\rho$ has two minimal maximal dilations $\pi_1$ and $\pi_2$. Then $\pi_1 \oplus \pi_2$ is a maximal dilation of $\rho \oplus \rho$. Now $(\rho \oplus \rho)(\mathcal{A})$ commutes with $X = \begin{bmatrix} I & I \\ I & I \end{bmatrix}$. By SMCLT, this dilates to a contraction $Y$ which commutes with $(\pi_1 \oplus \pi_2)(\mathcal{A})$ and $\mathcal{H} \oplus \mathcal{H}$ is jointly semi-invariant for $Y$ and $(\pi_1 \oplus \pi_2)(\mathcal{A})$. Arguing exactly as in the proof of Theorem 5.7, we deduce that $\pi_1$ and $\pi_2$ are unitarily equivalent via a unitary which fixes $\mathcal{H}$. So $\mathcal{A}$ has unique maximal dilations.

Conversely, it is routine to see that unique maximal dilations and MCLT yields SMCLT. So (i) and (ii) are equivalent.

Remark 5.16. One might suspect, as we did, that SMCLT is also equivalent to the following:

(iii) $\mathcal{A}$ has SCLT and SCLT*.

(iv) $\mathcal{A}$ has MCLT and unique minimal fully extremal extensions and coextensions.

But this is not the case.

If (ii) holds, then by Theorem 3.16 there is uniqueness of minimal fully extremal coextensions; and Corollary 3.17 yields uniqueness of minimal fully extremal extensions. So (iv) holds.

Also if (iii) holds, then there are unique minimal fully extremal extensions and coextensions by Theorem 5.7 and its dual result for SCLT*. Also MCLT holds by Theorem 5.14. So (iv) holds.

However, it is possible that SCLT and SCLT* hold, yet $\mathcal{A}$ does not have unique minimal maximal dilations. See the example of $2 \times 2$ matrices developed in the next section. Thus SMCLT fails to hold.

We do not know if SMCLT implies (iii).

In the case of semi-Dirichlet algebras, we have something extra. We do not know if all semi-Dirichlet algebras have SCLT. However, Muhly and Solel [42] show that the tensor algebra over any C*-correspondence has SCLT.

Proposition 5.17. Suppose that $\mathcal{A}$ is semi-Dirichlet and has MCLT. Then $\mathcal{A}$ has SCLT.

Proof. Let $\rho$ be a representation of $\mathcal{A}$ on $\mathcal{H}$ commuting with a contraction $X$. Use MCLT to obtain a simultaneous dilation of $\rho$ to a maximal dilation $\pi$ and $X$ to a contraction $Y$ commuting with $\pi(\mathcal{A})$. Let $\mathcal{K}$ be the common invariant subspace for $\pi(\mathcal{A})$ and $Y$ containing $\mathcal{H}$. Since $\mathcal{H}$ is semi-invariant, the restriction of $\pi$ to $\mathcal{K}$ is a coextension $\sigma$ of $\rho$. The compression $Z$ of $Y$ to $\mathcal{K}$ is a contraction commuting with $\sigma(\mathcal{A})$. 

Documenta Mathematica 16 (2011) 781–868
By Theorem 4.6, there is a unique minimal extremal coextension $\sigma_0$ of $\rho$, and it must coincide with $\sigma|_{K_0}$ where $K_0 = \sigma(A)\mathcal{H}$. Thus $\sigma = \sigma_0 \oplus \sigma'$. It follows that the compression $Z_0$ of $Z$ to $K_0$ commutes with $\sigma_0(A)$. Moreover, since $\mathcal{H}$ is coinvariant for $Z$, it is also coinvariant for $Z_0$. By Theorem 5.7, $A$ has SCLT.

**Corollary 5.18.** If $A$ is a Dirichlet algebra, the following are equivalent:

(i) $A$ has MCLT
(ii) $A$ has SMCLT
(iii) $A$ has SCLT and SCLT*
(iv) $A$ has WCLT and WCLT*.

Finally we point out that there is also an intertwining version for MCLT.

**Proposition 5.19.** Suppose that $A$ has MCLT. Let $\rho_i$ be representations of $A$ on $\mathcal{H}_i$ for $i = 1, 2$. Suppose that $X$ is a contraction in $\mathcal{B}(\mathcal{H}_2, \mathcal{H}_1)$ such that

$$\rho_1(a)X = X\rho_2(a) \quad \text{for all } a \in A.$$  

Then there are maximal dilations $\pi_i$ of $\rho_i$ acting on $K_i \supset \mathcal{H}_i$ for $i = 1, 2$ and a contraction $Y$ simultaneously dilating $X$ in $\mathcal{B}(K_2, K_1)$ so that

$$\pi_1(a)Y = Y\pi_2(a) \quad \text{for all } a \in A.$$  

**Proof.** Again form $\rho = \rho_1 \oplus \rho_2$ and $\tilde{X}$ as before. Use MCLT to dilate $\rho$ to a maximal dilation $\pi$ and $\tilde{X}$ to a contraction $\tilde{Y}$ commuting with $\pi(A)$ on

$$K = \mathcal{K} = \mathcal{K}_- \oplus \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{K}_+.$$  

Write

$$\tilde{Y} = \begin{bmatrix} * & 0 & 0 & 0 \\ * & 0 & X & 0 \\ * & 0 & 0 & 0 \\ * & * & * & * \end{bmatrix}.$$  

Now notice that $\pi$ is an maximal dilation of both $\rho_i$. Considering $\tilde{Y}$ as a map from

$$K_- \oplus \mathcal{H}_1 \oplus (\mathcal{H}_2 \oplus \mathcal{K}_+) \quad \text{to} \quad (K_- \oplus \mathcal{H}_1) \oplus \mathcal{H}_2 \oplus \mathcal{K}_+$$

one finds that it has a matrix form

$$\tilde{Y} = \begin{bmatrix} * & 0 & 0 & 0 \\ * & 0 & X & 0 \\ * & 0 & 0 & 0 \\ * & * & * & * \end{bmatrix}.$$  

This is a dilation of $X$ which commutes with $\pi(A)$, where $\pi$ is considered as a dilation of both $\rho_1$ and $\rho_2$ using the two decompositions of $\mathcal{K}$.  

---

**Documenta Mathematica 16 (2011) 781–868**
6 A 2 × 2 matrix example.

Consider the algebra $\mathcal{A} = \text{span}\{I_2, n\} \subset \mathcal{M}_2$ where

$$n = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$ 

Observe that $\mathcal{A}$ is unitarily equivalent to $\mathcal{A}^*$. It is not semi-Dirichlet. We will show that it has unique minimal extremal (co)extensions, which are always maximal dilations. But it does not have unique maximal dilations. It also has SCLT, SCL* and MCLT, but does not have SMCLT.

Observe that a representation $\rho$ of $\mathcal{A}$ is determined by $N := \rho(n)$, which satisfies $N^2 = 0$ and $\|N\| \leq 1$. Conversely, any such $N$ yields a completely contractive representation. It is easy to check that $N$ is unitarily equivalent to an operator of the form

$$\begin{bmatrix} 0 & 0 \\ 0 & B \end{bmatrix}$$

where $B$ has dense range by setting $\mathcal{H}_2 = \text{Ran} N$ and decomposing $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$. This can be refined by using the polar decomposition of $B$ to the form

$$\begin{bmatrix} 0 & 0 \\ A & 0 \end{bmatrix} \oplus 0$$
on $\mathcal{H} = \text{Ran} N^* \oplus \text{Ran} N \oplus (\ker N \cap \ker N^*)$,

where $A$ is a positive injective operator.

Clearly $C^*_e(\mathcal{A}) = \mathcal{M}_2$. Thus a maximal representation $\pi$ extends to a $*$-representation of $\mathcal{M}_2$. Hence $\pi(n) = N$ is a partial isometry such that

$$NN^* + N^*N = I \quad \text{and} \quad N^2 = 0;$$
or equivalently,

$$(N + N^*)^2 = I \quad \text{and} \quad N^2 = 0.$$ 

In other words, there is a unitary $U$ so that

$$N \simeq \begin{bmatrix} 0 & 0 \\ U & 0 \end{bmatrix} \simeq \begin{bmatrix} 0 & 0 \\ I & 0 \end{bmatrix}.$$ 

Geometrically, this says that $N$ is a partial isometry such that

$$\text{Ran} N + \text{Ran} N^* = \mathcal{H}.$$ 

**Proposition 6.1.** The algebra $\mathcal{A}$ of $2 \times 2$ matrices of the form

$$\begin{bmatrix} a & 0 \\ b & a \end{bmatrix}$$

has unique minimal extremal coextensions (extensions). Moreover they are fully extremal coextensions (extensions), and in fact, are maximal dilations; and the original space is cyclic (cocyclic).
Proof. We first show that a representation \( \rho \) which is not maximal has a proper coextension. Use the matrix form
\[
\rho(n) = N = \begin{bmatrix} 0 & 0 \\ B & 0 \end{bmatrix}
\]
on \( \mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2 \) where \( B \) has dense range. Suppose that \( B \) is not an isometry, and let
\[
D_B = (I - B^*B)^{1/2} \quad \text{and} \quad \mathcal{D}_B = \text{Ran}D_B.
\]
Consider \( P_{\mathcal{D}_B}D_B \) as an operator from \( \mathcal{H}_1 \) to \( \mathcal{D}_B \). Then \( B \) may be coextended to an isometry
\[
V = \begin{bmatrix} B \\ P_{\mathcal{D}_B}D_B \end{bmatrix}
\]
mapping \( \mathcal{H}_1 \) to \( \mathcal{K}_2 = \mathcal{H}_2 \oplus \mathcal{D}_B \). So \( \rho \) coextends to \( \sigma \) where
\[
\sigma(n) = S = \begin{bmatrix} 0 & 0 \\ V & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ B & 0 & 0 \\ P_{\mathcal{D}_B}D_B & 0 & 0 \end{bmatrix}
\]
on
\( \mathcal{K} = \mathcal{H}_1 \oplus \mathcal{K}_2 = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{D}_B \).

Hence when \( \rho \) is extremal, \( B \) is an isometry with dense range, so it is unitary. Thus \( \rho \) is a maximal representation.

The coextension constructed above is generally not extremal because \( V \) is not unitary. So one can coextend \( S \) again using the same procedure. \( \mathcal{K}_2 \) splits as \( \mathcal{R}_V \oplus \mathcal{D} \), where
\[
\mathcal{R}_V = \text{Ran}V \quad \text{and} \quad \mathcal{D} = \text{Ran}(I_{\mathcal{K}_2} - VV^*)
\]
We now decompose
\[
\mathcal{K} = \mathcal{H}_1 \oplus \mathcal{D} \oplus \mathcal{R}_V,
\]
and write
\[
S = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ P_{\mathcal{R}_V}V & 0 & 0 \end{bmatrix}
\]
Using the same dilation as the previous paragraph, but noting that
\[
D_{[P_{\mathcal{R}_V}V,0]} = (I_{\mathcal{K}_2} - VV^*)
\]
we obtain the coextension \( \tau \) of \( \sigma \) on a Hilbert space
\[
\mathcal{L} = \mathcal{H}_1 \oplus \mathcal{D} \oplus \mathcal{R}_V \oplus \mathcal{D}
\]
by
\[
\tau(n) = W = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ P_{\mathcal{R}_V}V & 0 & 0 & 0 \\ 0 & P_{\mathcal{D}}(I_{\mathcal{K}_2} - VV^*) & 0 & 0 \end{bmatrix}
\]
The bottom left $2 \times 2$ corner is now a surjective isometry. So this is a maximal dilation.

To see that $\tau$ is minimal as a coextension, we need to verify that

$$\mathcal{L} = A\mathcal{H} = \mathcal{H} \vee W\mathcal{H}.$$  

To see this, we rewrite $W$ with respect to the decomposition

$$\mathcal{L} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{D}_B \oplus \mathcal{D}$$

to get

$$W = \begin{bmatrix}
B & 0 & 0 & 0 \\
0 & B & 0 & 0 \\
P_{\mathcal{D}_B}D_B & 0 & 0 & 0 \\
0 & P_D(I_{K_2} - VV^*) & 0 & 0
\end{bmatrix}$$

Since the range of $D_B$ is dense in $\mathcal{D}_B$, this is contained in $\mathcal{H} \vee W\mathcal{H}$. For the space $\mathcal{D}$, we expand the expression for $I_{K_2} - VV^*$ on $K_2 = \mathcal{H}_2 \oplus \mathcal{D}_B$:

$$I_{K_2} - VV^* = \begin{bmatrix}
I & 0 \\
0 & I
\end{bmatrix} - \begin{bmatrix}
BB^* & BD_B \\
DB_B^* & I - B^*B
\end{bmatrix} = \begin{bmatrix}
I - BB^* & -BD_B \\
-D_BB^* & B^*B
\end{bmatrix} = \begin{bmatrix}
D_B^2 & -BD_B \\
-BD_B & B^*B
\end{bmatrix}$$

Thus one sees that

$$(I_{K_2} - VV^*)P_{\mathcal{H}_2} = \begin{bmatrix}
D_{B^*} \\
-B
\end{bmatrix} D_{B^*}.$$  

Observe that $D_{B^*}$ maps $\mathcal{H}_2$ onto a dense subspace of $\ker(I - BB^*)^\perp$, and

$$\ker(I - BB^*) = \text{Ran} V \cap \mathcal{H}_2 \subset \ker(I - VV^*) \cap \mathcal{H}_2.$$  

The range of

$$\begin{bmatrix}
D_{B^*} \\
-B
\end{bmatrix}$$

is easily seen to be the orthogonal complement of $\text{Ran} V$, so this is $\mathcal{D}$. Restricting this map to the range of $D_{B^*}$ does not affect the closed range, since we only miss some of the kernel. Thus $W\mathcal{H}_2$ is dense in $\mathcal{D}$. Therefore this is a minimal extremal coextension.

Now we consider uniqueness. To this end, suppose that $\tau'$ is a minimal extremal coextension of $\rho$ on

$$\mathcal{L}' = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{L}_3.$$
Then we can write
\[ \tau'(n) = W' = \begin{bmatrix} 0 & 0 & 0 \\ B & 0 & 0 \\ X & Y & Z \end{bmatrix}. \]

This is a partial isometry satisfying
\[ W'^2 = 0 \quad \text{and} \quad W'W'^* + W'^*W' = I. \]

In particular, \( \mathcal{H}_1 \) is orthogonal to the range of \( W' \), so the restriction of \( W' \) to \( \mathcal{H}_1 \) is an isometry. Therefore
\[ X = U P_D B, \]
where \( U \) is an isometry of \( D_B \) into \( L_3 \). Split
\[ L_3 = U D_B \oplus L_4 \cong D_B \oplus L_4. \]

By identifying the range of \( U \) with \( D_B \), we have the refined form
\[ W' = \begin{bmatrix} 0 & 0 & 0 & 0 \\ B & 0 & 0 & 0 \\ P_D B X_1 & Y_1 & Z_1 \\ 0 & X_2 & Y_2 & Z_2 \end{bmatrix}. \]

The range of \( W'H_1^\perp \) is orthogonal to
\[ W'H_1 \vee \mathcal{H}_2 = \mathcal{H}_2 \oplus D_B. \]

So
\[ X_1 = Y_1 = Z_1 = 0. \]

Next note that minimality ensures that \( X_2 \) has dense range in \( L_4 \). So \( L_4 \) is in
\[ \text{Ran} \ W' = \ker W'^*. \]

Hence \( Z_2 = 0 \). Observe that
\[ \text{Ran} \ W' = \text{Ran} \ V \oplus L_4, \]
and hence
\[ \text{Ran} \ W'^* = (\text{Ran} \ W')^\perp = H_1 \oplus D. \]

Moreover, the operator
\[ \begin{bmatrix} X_2 \\ Y_2 \end{bmatrix} \]
is an isometry of \( \mathcal{D} \) onto \( L_4 \). Hence we may identify \( L_4 \) with \( \mathcal{D} \) in such a way that we obtain
\[ \begin{bmatrix} X_2 \\ Y_2 \end{bmatrix} \simeq P_D (I_{K_2} - V V^*). \]

This shows that \( W' \) is unitarily equivalent to \( W \) via a unitary which fixes \( \mathcal{H} \). Therefore the minimal extremal coextension is unique.

The proof for extensions follows immediately since \( \mathcal{A}^* \) is unitarily equivalent to \( \mathcal{A} \).
Corollary 6.2. Every representation of the algebra $\mathcal{A}$ is Shilov.

Proof. By the previous theorem, one can extend $\rho$ to a maximal dilation $\pi$. Thus $\rho$ is obtained as the restriction of a maximal representation to an invariant subspace; i.e. it is Shilov. \(\blacksquare\)

Example 6.3. Take $\rho$ to be the character representation on $\mathcal{H} = \mathbb{C}$ given by

$$\rho(aI_2 + bn) = a.$$  

This coextends to a maximal representation on $\mathcal{K} = \mathcal{H} \oplus \mathbb{C}$ as

$$\sigma(aI_2 + bn) = \begin{bmatrix} a & 0 \\ b & a \end{bmatrix}.$$  

It also extends to a maximal representation $\tau$ on $\mathcal{K} = \mathbb{C} \oplus \mathcal{H}$ where

$$\tau(aI_2 + bn) = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}.$$  

Note that $\sigma$ and $\tau$ are not unitarily equivalent by a unitary which fixes $\mathcal{H}$. So these are inequivalent minimal maximal dilations.

Corollary 6.4. The minimal maximal dilation of a representation $\rho$ of $\mathcal{A}$ is not unique except when $\rho$ is already maximal.

Proof. Let $\tau$ be the minimal extremal coextension of $\rho$, and let $\sigma$ be the minimal extremal extension of $\rho$. Then in the first case, $\mathcal{H}$ is identified with a coinvariant subspace and in the latter with an invariant subspace. If these two dilations are unitarily equivalent via a unitary which fixes $\mathcal{H}$, then $\mathcal{H}$ is reducing, and therefore by minimality, $\rho = \tau$ is maximal. \(\blacksquare\)

Next we show that $\mathcal{A}$ has commutant lifting.

Theorem 6.5. The algebra $\mathcal{A}$ of $2 \times 2$ matrices of the form

$$\begin{bmatrix} a & 0 \\ b & a \end{bmatrix}$$  

has SCLT, SCLT* and MCLT, but not SMCLT.

Proof. It is enough to verify CLT. Since the minimal fully extremal coextensions are unique, it then has SCLT by Theorem 5.7. Since $\mathcal{A}^* \simeq \mathcal{A}$, it has SCLT* as well. Thus by Theorem 5.14, it has MCLT. But by Theorem 5.15, it does not have SMCLT.

We make use of the construction of the minimal extremal extension in the proof of Theorem 6.1. Write

$$\rho(n) = N = \begin{bmatrix} 0 & 0 \\ B & 0 \end{bmatrix}$$  

Documenta Mathematica 16 (2011) 781–868
as before on
\[ \mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2 \quad \text{where} \quad \mathcal{H}_2 = \overline{\text{Ran} \ N}. \]
Suppose that it commutes with a contraction \( T \). Then it is routine to check that
\[ T = \begin{bmatrix} X & 0 \\ Y & Z \end{bmatrix} \]
such that
\[ BX = ZB. \]
Coextend \( \rho \) to the coextension \( \sigma \) on
\[ \mathcal{K} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{D}_B \]
where
\[ \sigma(n) = S = \begin{bmatrix} 0 & 0 & 0 \\ B & 0 & 0 \\ D_B & 0 & 0 \end{bmatrix}. \]
We first find a coextension of \( T \) to \( \tilde{T} \) which commutes with \( S \) and has norm one.
Consider the isometric dilation of \( N \). Observe that
\[ D_N = (I - N^*N)^{1/2} = \begin{bmatrix} D_B & 0 \\ 0 & I \end{bmatrix}. \]
So the minimal isometric dilation acts on
\[ (\mathcal{H}_1 \oplus \mathcal{H}_2) \oplus (\mathcal{D}_B \oplus \mathcal{H}_2)^{(\infty)}, \]
and has the form
\[ U = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ B & 0 & 0 & 0 & 0 & 0 & \cdots \\ D_B & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & I & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & I & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & I & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \]
Notice that \( S \) is the upper left \( 3 \times 3 \) corner. By the Sz.Nagy-Foiaş Commutant Lifting Theorem, we can coextend \( T \) to a contraction \( R \) commuting with \( U \). It has the form
\[ R = \begin{bmatrix} X & 0 & 0 & \cdots \\ Y & Z & 0 & \cdots \\ C_1 & C_2 & C_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \]
It is routine to verify that
\[
\tilde{T} = \begin{bmatrix}
X & 0 & 0 \\
Y & Z & 0 \\
C_1 & C_2 & C_3
\end{bmatrix}
\]
commutes with \( S \).

Now we repeat the argument with \( S \) and \( \tilde{T} \). The same procedure was shown in Theorem 6.1 to yield the minimal (fully) extremal coextension \( \tau \) of \( \rho \). The operator \( \tilde{T} \) is coextended once again to obtain a contraction commuting with \( \tau(n) = W \). This establishes SCLT.

Remark 6.6. We will show that in the commutant lifting theorem for \( \mathcal{A} \), it is not possible to coextend so that \( \tilde{T} \) is an isometry. In the language of the next section, this will show that \( \mathcal{A} \) does not have ICLT (isometric commutant lifting) nor the Ando property.

To see this, consider the identity representation
\[
\text{id}(n) = N = \begin{bmatrix}
0 & 0 \\
1 & 0
\end{bmatrix}
\]
on \( \mathcal{H} = \mathbb{C}^2 \). Then \( \text{id}(A) \) commutes with \( T = N \). Suppose that there were a coextension of \( \text{id} \) and \( T \) to \( \sigma \) and an isometry \( V \) on \( K \) so that \( \sigma(A) \) commutes with \( V \). Since \( \text{id} \) is maximal, \( \sigma = \text{id} \oplus \tau \). So
\[
\sigma(n) = M = N \oplus M_0
\]
where \( M_0^2 = 0 \). Let the canonical basis for \( \mathcal{H} \) be \( e_1, e_2 \). Since \( Te_2 = 0 \), we have \( Ve_2 = v \) is a unit vector in \( \mathcal{H}^\perp \); while
\[
Ve_1 = Te_1 = e_2.
\]
Therefore
\[
v = Ve_2 = VMe_1 = MVe_1 = Me_2 = 0.
\]
This contradiction shows that no such coextension is possible.

7 Isometric Commutant Lifting and Ando’s Theorem

Paulsen and Power [47] formulate commutant lifting and Ando’s theorem in terms of tensor products. In doing so, they are also able to discuss lifting commuting relations between two arbitrary operator algebras. They are interested in dilations which extend to the enveloping C*-algebra, which are the maximal dilations when this C*-algebra is the C*-envelope. The Paulsen-Power version of Ando’s theorem involves maximal dilations and a commuting unitary. The
Dilation Theory.

classical Ando Theorem, from our viewpoint, states that two commuting contractions have coextensions to commuting isometries. We will give a similar definition using extremal co-extensions instead which is actually stronger than the Paulsen-Power version.

**Definition 7.1.** We say that \( \mathcal{A} \) has *isometric commutant lifting* (ICLT) if whenever \( \rho \) is a representation of \( \mathcal{A} \) on \( \mathcal{H} \) commuting with a contraction \( X \), then there is a coextension \( \sigma \) of \( \rho \) and an isometric coextension \( V \) of \( X \) on a common Hilbert space \( \mathcal{K} \) so that \( \sigma(\mathcal{A}) \) and \( V \) commute.

**Remark 7.2.** We can see ICLT as a commutant lifting result for \( \mathcal{A}(\mathcal{D}) \) relative to \( \mathcal{A} \). More generally, we can formulate a relative commutant lifting property for two operator algebras. We define this here, but will not pursue it except in the case above. Let \( \mathcal{A} \) and \( \mathcal{B} \) be unital operator algebras. Say that \( \mathcal{A} \) has \( \mathcal{B} \)-CLT (or *commutant lifting with respect to \( \mathcal{B} \)*) if whenever \( \alpha \) and \( \beta \) are (completely contractive) representations of \( \mathcal{A} \) and \( \mathcal{B} \) on a common Hilbert space \( \mathcal{H} \) which commute:

\[
\alpha(a)\beta(b) = \beta(b)\alpha(a) \quad \text{for all } a \in \mathcal{A} \text{ and } b \in \mathcal{B},
\]

then there exists an extremal coextension \( \sigma \) of \( \alpha \) on a Hilbert space \( \mathcal{K} \) and a coextension \( \tau \) of \( \beta \) on \( \mathcal{K} \) which commute.

Because there is no uniqueness for the classical Ando’s theorem, we do not seek a strong form. So we do not use the adjective weak either. Note that ICLT is not stronger than WCLT because the extremal condition is on the isometry, not on the coextension of \( \mathcal{A} \). Nevertheless it does imply a much stronger conclusion, as we show in Theorem 7.3 below.

It is often observed that Ando’s Theorem is equivalent to commutant lifting. However neither direction is completely trivial. From Ando’s theorem, one easily gets WCLT. So the uniqueness of the minimal isometric coextension, and the fact that this is fully extremal, then yields SCLT. Conversely, in deducing Ando’s theorem from WCLT, one is really using WCLT for both contractions. One iteratively dilates one contraction to an isometry and lifts the other to commute. The inductive limit is a pair of commuting isometries. We will see this more clearly for operator algebras other than the disk algebra.

The next result shows that the Paulsen-Power version of Ando’s theorem is equivalent to ICLT. It also shows why we consider \( \mathcal{A} \)-CLT for \( \mathcal{A}(\mathcal{D}) \) as a strong property for \( \mathcal{A} \), and makes it worthy of the term ICLT.

**Theorem 7.3.** For a unital operator algebra \( \mathcal{A} \), the following are equivalent:

(i) \( \mathcal{A} \) had ICLT; i.e. if \( \rho \) is a representation of \( \mathcal{A} \) on \( \mathcal{H} \) commuting with a contraction \( X \), then there is a coextension \( \sigma \) of \( \rho \) and an isometric coextension \( V \) of \( X \) on a common Hilbert space \( \mathcal{K} \) which commute.

(ii) If \( \rho \) is a representation of \( \mathcal{A} \) on \( \mathcal{H} \) commuting with a contraction \( X \), then there is a Shilov coextension \( \sigma \) of \( \rho \) and an isometric coextension \( V \) of \( X \) on a common Hilbert space \( \mathcal{K} \) which commute.
(iii) If $\rho$ is a representation of $A$ on $\mathcal{H}$ commuting with a contraction $X$, then there is a simultaneous dilation of $\rho$ to a maximal dilation $\pi$ on $K$ and of $X$ to a unitary $U$ commuting with $\pi(A)$; i.e. there is a Hilbert space $K \supset \mathcal{H}$, a $*$-representation $\pi$ of $C^*_e(A)$ on $K$ and a unitary operator $U$ on $K$ commuting with $\pi(C^*_e(A))$ so that
\[ P_H \pi(a) U^n |_H = \rho(a) X^n \]
for all $a \in A$ and $n \geq 0$.

Proof. It is evident that (iii) implies (ii) by restriction to the smallest invariant subspace containing $\mathcal{H}$. And (ii) clearly implies (i). So assume that (i) holds. We will establish (iii).

First we dilate $\sigma$ and $V$ to $\tau$ and $W$ so that $W$ is unitary and commutes with $\tau(A)$. To accomplish this, consider the system with $K_n = K$ and $V$ considered as a map from $K_n$ into $K_{n+1}$:

\[
\begin{array}{ccccccc}
K_1 & \overset{V}{\rightarrow} & K_2 & \overset{V}{\rightarrow} & K_3 & \overset{V}{\rightarrow} & \cdots & \overset{V}{\rightarrow} & P \\
\downarrow V & & \downarrow V & & \downarrow V & & \downarrow V & & \downarrow W \\
K_1 & \overset{V}{\rightarrow} & K_2 & \overset{V}{\rightarrow} & K_3 & \overset{V}{\rightarrow} & \cdots & \overset{V}{\rightarrow} & P
\end{array}
\]

Then $P$ is the Hilbert space direct limit of copies of $K$ under $V$. Let $J_n$ denote the canonical injection of $K_n$ into $K_{n+1}$:
\[ J_n = J_{n+1} V \text{ for } n \geq 1. \]

The map $V$ also determines isometries acting on each $K_n$, which we also denote by $V$. The direct limit of this system of maps is a unitary $W$ on $P$ such that its restriction to each $K_n$ coincides with $V$. Hence
\[ J_n V = W J_n \text{ for } n \geq 1. \]

In particular, $W$ is an extension of $V$ acting on $K_1$, which we identify with $K$.

We define a representation $\tau$ of $A$ on $P$ by
\[ \tau(a) J_n k = J_n \sigma(a) k \text{ for } a \in A, \ k \in K_n \text{ and } n \geq 1. \]

Clearly each subspace $J_n K_n$ is invariant for $\tau$ and the restriction of $\tau$ to $K_n$ is equivalent to $\sigma$. In particular, $\tau$ is an extension of $\sigma$, where we identify $K$ with $K_1$. Additionally, since $\tau$ is completely contractive when restricted to each $J_n K_n$, we see that $\tau$ is completely contractive. Finally, for $a \in A$ and $k \in K_n$ for $n \geq 2$,
\[
\begin{align*}
\tau(a) W J_n(a) k &= \tau(a) J_n V k = \tau(a) J_{n-1} k \\
&= J_{n-1} \sigma(a) k = J_n V \sigma(a) k \\
&= W J_n \sigma(a) k = W \tau(a) J_n k.
\end{align*}
\]
Therefore, $W$ commutes with $\tau(\mathcal{A})$.

The completely contractive map $\tau$ extends to a unique completely positive unital map on the operator system

$$\mathcal{A} + \mathcal{A}^* \subset C^*_e(\mathcal{A}).$$

By Fuglede’s Theorem, $W$ commutes with $\tau(\mathcal{A}) + \tau(\mathcal{A})^*$. The commutant $\mathcal{H}$ of $W$ is a type I von Neumann algebra, and therefore it is injective. Therefore by Arveson’s Extension Theorem, there is a completely positive extension of $\tau$ to $C^*_e(\mathcal{A})$ with range in $\mathcal{H}$. By Stinespring’s Dilation Theorem, there is a minimal dilation to a $*$-representation $\pi$ of $C^*_e(\mathcal{A})$ on a larger Hilbert space. Now a commutant lifting result of Arveson [5, Theorem 1.3.1] shows that there is a unique extension of $W$ to an operator $U$ commuting with $\pi(C^*_e(\mathcal{A}))$. This extension map is a $*$-homomorphism, so $U$ is unitary. Moreover the fact that the restriction of $\pi|_{\mathcal{A}}$ to the space $\mathcal{P}$ is a homomorphism means that $\pi$ is a maximal dilation of $\tau$, and hence of $\rho$.

The following corollary is a consequence of (i) implies (iii) above.

**Corollary 7.4.** Property ICLT implies MCLT for $\mathcal{A}$.

Another easy corollary is a consequence of the fact that (iii) is symmetric.

**Corollary 7.5.** Property ICLT is equivalent to ICLT*. So if $\mathcal{A}$ has ICLT, so does $\mathcal{A}^*$.

**Example 7.6.** Finite dimensional nest algebras have ICLT by Paulsen and Power [46, 47]. They actually prove variant (iii). They also claim that the minimal $*$-dilation is unique. This follows because finite dimensional nest algebras are Dirichlet. So by Theorem 5.15, they have SMCLT.

Dirichlet implies semi-Dirichlet for the algebra and its adjoint. So there are unique minimal fully extremal (co)extensions. The proof of ICLT in fact first coextends to an isometry in the commutant. Hence finite dimensional nest algebras have SCLT and SCLT*.

Nest algebras have the SCLT, SCLT* and MCLT for weak-$*$ continuous completely contractive representations.

The first part of the following proposition uses exactly the same proof as Proposition 5.19 with the exception that the dilation $\hat{Y}$ obtained can be taken to be a unitary operator when ICLT is invoked.

**Proposition 7.7.** Suppose that $\mathcal{A}$ has ICLT. Let $\rho_i$ be representations of $\mathcal{A}$ on $\mathcal{H}_i$ for $i = 1, 2$. Suppose that $X$ is a contraction in $B(\mathcal{H}_2, \mathcal{H}_1)$ such that

$$\rho_1(a)X = X\rho_2(a) \quad \text{for all } a \in \mathcal{A}.$$
Then there are maximal dilations $\pi_i$ of $\rho_i$ acting on $K_i \supset H_i$ for $i = 1, 2$ and a unitary operator $U$ simultaneously dilating $X$ in $\mathcal{B}(K_2, K_1)$ so that

$$\pi_1(a)U = U\pi_2(a) \quad \text{for all } a \in \mathcal{A}.$$  

Consequently, there exist Shilov coextensions $\sigma_i$ of $\rho_i$ on $\mathcal{L}_i$ and a coextension of $X$ to an isometry $V \in \mathcal{B}(\mathcal{L}_2, \mathcal{L}_1)$ so that

$$\sigma_1(a)V = V\sigma_2(a) \quad \text{for all } a \in \mathcal{A}.$$  

Proof. We only discuss the second statement. Let $\pi_i$ act on $K_i = K_i^- \oplus H_i \oplus K_i^+$, where $\mathcal{L}_i = H_i \oplus K_i^+$ and $K_i^+$ are invariant subspaces for $\pi_i(\mathcal{A})$. With respect to these decompositions, we have the matrix forms

$$\pi_i(a) = \begin{bmatrix} * & 0 & 0 \\ * & \rho_i(a) & 0 \\ * & \pi_{32}(a) & \pi_{33}(a) \end{bmatrix} \quad \text{and} \quad U = \begin{bmatrix} * & 0 & 0 \\ * & X & 0 \\ * & U_{32} & U_{33} \end{bmatrix}.$$  

Set

$$\sigma_i(a) = P_{\mathcal{L}_i}\pi_i(a)|_{\mathcal{L}_i} = \begin{bmatrix} \rho_i(a) \\ \pi_{32}(a) \\ \pi_{33}(a) \end{bmatrix} \quad \text{for } i = 1, 2.$$  

and

$$V = P_{\mathcal{L}_1}U|_{\mathcal{L}_2} = \begin{bmatrix} X \\ U_{32} \\ U_{33} \end{bmatrix}.$$  

Then $\sigma_i$ are Shilov coextensions of $\rho_i$, $V$ is an isometry, and

$$\sigma_1(a)V = P_{\mathcal{L}_1}\pi_1(a)|_{\mathcal{L}_1}P_{\mathcal{L}_1}U|_{\mathcal{L}_2} = P_{\mathcal{L}_1}\pi_1(a)U|_{\mathcal{L}_2} = P_{\mathcal{L}_1}\pi_2(a)P_{\mathcal{L}_2} = P_{\mathcal{L}_1}U|_{\mathcal{L}_1}\pi_2(a)|_{\mathcal{L}_1}P_{\mathcal{L}_2} = VP_{\mathcal{L}_2}\pi_2(a)|_{\mathcal{L}_2} = V\sigma_2(a).$$

We ask a bit more for what we will call the Ando property. This is stronger than the classical Ando Theorem for $\mathcal{A}(\mathcal{D})$. The weak version for $\mathcal{A}(\mathcal{D})$ is just Ando’s Theorem.

Definition 7.8. A unital operator $\mathcal{A}$ satisfies the Ando property if whenever $\rho$ is a representation of $\mathcal{A}$ on $\mathcal{H}$ and $X \in \mathcal{B}(\mathcal{H})$ is a contraction commuting with $\rho(\mathcal{A})$, there is a fully extremal coextension $\sigma$ of $\rho$ on a Hilbert space $\mathcal{K}$ and a coextension of $X$ to an isometry on $\mathcal{K}$ which commute.

Likewise say that $\mathcal{A}$ satisfies the weak Ando property if whenever $\rho$ is a representation of $\mathcal{A}$ on $\mathcal{H}$ and $X \in \mathcal{B}(\mathcal{H})$ is a contraction commuting with $\rho(\mathcal{A})$, there is an extremal coextension $\sigma$ of $\rho$ on a Hilbert space $\mathcal{K}$ and a coextension of $X$ to an isometry on $\mathcal{K}$ which commute.

If $\mathcal{A}^*$ has the (weak) Ando property, say that $\mathcal{A}$ has the (weak) Ando* property.
It is apparent that the Ando property implies CLT and ICLT for $A$; and the weak Ando property implies WCLT and ICLT. The converse of the latter fact follows the same lines as the classical deduction of Ando’s theorem from CLT. But the converse for the full Ando property is more difficult. The difference is that an extremal coextension of a coextension is extremal, but a fully extremal coextension of a coextension is generally not fully extremal. So more care has to be taken, and a Schaeffer type construction makes it work.

**Proposition 7.9.** A unital operator algebra $A$ has the weak Ando property if and only if $A$ has WCLT and ICLT.

**Proof.** Let $\rho$ be a representation of $A$ which commutes with a contraction $X$. Assume WCLT and ICLT. Coextend $\rho$ and $X$ to an extremal coextension $\sigma_1$ and a commuting contraction $Y_1$ using WCLT. Then coextend $\sigma_1$ to $\rho_1$ and $Y_1$ to dilate to a commuting isometry by ICLT. Iterate these procedures recursively. The inductive limit has the desired properties. □

**Theorem 7.10.** A unital operator algebra $A$ has the Ando property if and only if $A$ has CLT and ICLT.

**Proof.** Let $\rho$ be a representation of $A$ which commutes with a contraction $X$. Assume CLT and ICLT. Use CLT to coextend $\rho$ and $X$ to a fully extremal coextension $\sigma$ and a commuting contraction $Y$ on $K \supset H$. Then use ICLT to coextend this to a Shilov dilation $\tau$ and commuting isometry $V$ on $L = K \oplus L'$. Since $\sigma$ is extremal, $\tau = \sigma \oplus \tau'$. Write

$$V = \begin{bmatrix} Y & 0 \\ Z & V' \end{bmatrix}$$

with respect to this decomposition of $L$. Then

$$\tau'(a)Z = Z\sigma(a) \quad \text{for } a \in A.$$ 

Since $\tau$ is Shilov, there is a maximal dilation $\pi$ which is an extension of $\tau$ on a Hilbert space $M$ containing $L$ as an invariant subspace. So $L'$ is also invariant, as it reduces $\pi(A)|_L = \tau(A)$. Let $P$ be the projection of $M$ onto $L'$. Then

$$\pi(a)P = \tau'(a)$$

and thus

$$\pi(a)(PZ) = P\tau'(a)Z = (PZ)\sigma(a) \quad \text{for all } a \in A.$$ 

The representation $\sigma \oplus \pi(\infty)$ is a fully extremal coextension of $\rho$. Moreover, $X$
coextends to $Y$ which coextends to

\[
W = \begin{bmatrix}
  Y & 0 & 0 & 0 & \cdots \\
  PZ & 0 & 0 & 0 & \cdots \\
  0 & I & 0 & 0 & \cdots \\
  0 & 0 & I & 0 & \cdots \\
  \vdots & \vdots & \vdots & \ddots & \ddots \\
\end{bmatrix}.
\]

It is easy to see that $W$ is an isometry. The relations established in the previous paragraph ensure that $W$ commutes with $(\sigma \oplus \pi(\infty))(A)$. Thus we have verified that $A$ has the Ando property.

This yields a strengthening of the classical Ando Theorem. The usual Ando Theorem verifies the weak Ando property, and hence ICLT. But the disk algebra has SCLT. So by Theorem 7.10, it has the Ando property. We provide a direct proof that is of independent interest.

**Corollary 7.11.** The disk algebra has the Ando property; i.e. if $A_1$ and $A_2$ are commuting contractions on $\mathcal{H}$, then they have commuting isometric coextensions $V_i$ on a common Hilbert space $\mathcal{K}$. Moreover, we can arrange for $V_2$ to be a fully extremal coextension (i.e. $V_{A_2} \oplus U$, where $V_{A_2}$ is the minimal isometric coextension and $U$ is unitary).

**Proof.** If $A$ is a contraction, let $V_A$ denote the unique minimal isometric coextension of $A$. Let $V_i$ be isometric dilations of $A_i$, $i = 1, 2$, acting on a common Hilbert space $\mathcal{K} = \mathcal{H} \oplus \mathcal{K}'$. (The minimal dilations may not have additional subspaces of the same dimension, for example if one is already an isometry. In this case, we just add a unitary summand to one of them.) Let $V = V_{A_1}, A_2$.

Note that both $V_1V_2$ and $V_2V_1$ are isometries of the form

\[
\begin{bmatrix}
  A_1A_2 & 0 \\
  * & *
\end{bmatrix}.
\]

So by the uniqueness of the minimal dilation, we can write

\[
V_1V_2 \simeq V \oplus W_1 \quad \text{and} \quad V_2V_1 \simeq V \oplus W_2,
\]

where $W_i$ is an isometry acting on a Hilbert space $\mathcal{K}_i$ (possibly of dimension 0).

Now we dilate $A_1$ to

\[
S_1 := V_1 \oplus W_1(\infty) \oplus W_2(\infty)
\]

and dilate $A_2$ to

\[
S_2 := V_2 \oplus I_{\mathcal{K}_1}(\infty) \oplus I_{\mathcal{K}_2}(\infty).
\]
on
\[ \mathcal{K} \oplus \mathcal{K}_1^{(\infty)} \oplus \mathcal{K}_2^{(\infty)}. \]

So
\[ S_1S_2 \simeq V \oplus W \oplus W_1^{(\infty)} \oplus W_2^{(\infty)} \simeq V \oplus W_1^{(\infty)} \oplus W_2^{(\infty)} \]
and
\[ S_2S_1 \simeq V \oplus W \oplus W_1^{(\infty)} \oplus W_2^{(\infty)} \simeq V \oplus W_1^{(\infty)} \oplus W_2^{(\infty)}. \]

These unitary equivalences both fix \( H \). Therefore there is a unitary operator \( U \) that fixes \( H \) so that
\[ S_2S_1 = U^*S_1S_2U. \]

Now define isometric dilations \( U^*S_1 \) of \( A_1 \) and \( S_2U \) of \( A_2 \). Then
\[ (U^*S_1)(S_2U) = U^*S_1S_2U = S_2S_1 = (S_2U)(U^*S_1). \]

This yields a commuting isometric dilation.

Moreover, if
\[ VA_2 = \begin{bmatrix} A_2 & 0 \\ D_2 & J \end{bmatrix}, \]
then \( S_2U \) has the form
\[ S_2U \simeq \begin{bmatrix} A_2 & 0 & 0 \\ D_2 & J & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} I & 0 & 0 \\ 0 & U_{22} & U_{23} \\ 0 & U_{32} & U_{33} \end{bmatrix} \simeq \begin{bmatrix} A_2 & 0 & 0 \\ D_2 & J & 0 \\ 0 & 0 & U' \end{bmatrix}. \]

The basic observation is that \( J \oplus I \) is an isometry on \( H^\perp \) with range equal to the orthocomplement of \( \text{Ran} \ D_2 \). The same is therefore true for the lower \( 2 \times 2 \) corner of \( S_2U \). By the uniqueness of the minimal isometric dilation, this corner splits as a direct sum \( J \oplus U' \) where \( U' \) must map onto the complement of the range of \( D \) and \( J \). So \( U' \) is unitary as claimed.

**Example 7.12.** One might ask to dilate both \( A_1 \) and \( A_2 \) to commuting isometries of the form \( VA_i \oplus U_i \) with \( U_i \) unitary. This is not possible, as the following example due to Orr Shalit shows. Let
\[ A_1 = 0 \quad \text{and} \quad A_2 = S, \]
where \( S \) is the unilateral shift on \( H = \ell^2 \). Then
\[ VA_2 = A_2 = S \]
and
\[ VA_1 = I \otimes S \]
acting on \( H \otimes \ell^2 \).

Suppose that \( U \in \mathcal{B}(\mathcal{H}_0) \) is unitary, and that
\[ W_1 = U \oplus VA_1. \]
K.R. Davidson and E.G. Katsoulis

commutes with

\[ W_2 = A_2 \oplus X = S \oplus X \quad \text{(with } H \text{ appropriately identified).} \]

Then they can be written as

\[
W_1 = \begin{bmatrix}
U & 0 & 0 & 0 & \ldots \\
0 & 0 & 0 & 0 & \ldots \\
0 & I & 0 & 0 & \ldots \\
0 & 0 & I & 0 & \ldots \\
\vdots & \vdots & \vdots & \ddots & \ddots
\end{bmatrix} \quad \text{and } W_2 = \begin{bmatrix}
X_{00} & 0 & X_{02} & X_{03} & \ldots \\
0 & S & 0 & 0 & \ldots \\
X_{20} & 0 & X_{22} & X_{23} & \ldots \\
X_{30} & 0 & X_{32} & X_{33} & \ldots \\
\vdots & \vdots & \vdots & \ddots & \ddots
\end{bmatrix}.
\]

Computing the two products yields

\[
\begin{bmatrix}
U X_{00} & 0 & U X_{02} & U X_{03} & \ldots \\
0 & 0 & 0 & 0 & \ldots \\
0 & S & 0 & 0 & \ldots \\
X_{20} & 0 & X_{22} & X_{23} & \ldots \\
X_{30} & 0 & X_{32} & X_{33} & \ldots \\
\vdots & \vdots & \vdots & \ddots & \ddots
\end{bmatrix} = \begin{bmatrix}
X_{00} U & X_{02} & X_{03} & X_{04} & \ldots \\
0 & 0 & 0 & 0 & \ldots \\
0 & S & 0 & 0 & \ldots \\
X_{20} U & X_{22} & X_{23} & X_{24} & \ldots \\
X_{30} U & X_{32} & X_{33} & X_{34} & \ldots \\
\vdots & \vdots & \vdots & \ddots & \ddots
\end{bmatrix}.
\]

Equating terms shows that

\[ X_{00} U = U X_{00}, \quad X_{ij} = S \text{ for } j \geq 2 \quad \text{and } X_{ij} = 0 \text{ otherwise.} \]

Thus

\[ W_2 \simeq X_{00} \oplus (S \otimes I), \]

which is not \( V_{A_2} \) direct sum a unitary.

Katsoulis and Kakariadis [32, Theorem 3.5] (in the special case of the identity automorphism) show that every tensor algebra of a C*-correspondence has the weak Ando property. As noted in Example 5.3, the Muhly-Solel Commutant Lifting Theorem [42] shows that the tensor algebra \( T^* (E) \) of a C*-correspondence \( E \) satisfies SCLT. Thus tensor algebras have the Ando property by Theorem 7.10. This forms a large class of algebras with this property. This includes all tensor algebras of graphs. The case of the non-commutative disk algebra follows from [17]. We would like to know if all semi-Dirichlet algebras have this property.

**Theorem 7.13 (Katsoulis-Kakariadis).** The tensor algebra of a C*-correspondence has the Ando property.

**Corollary 7.14.** The tensor algebra of a directed graph has the Ando property. In particular, the non-commutative disk algebras have this property.
The proof in [32] proves more, providing a lifting for relations that intertwine an automorphism. More will be said about this later when we discuss semi-crossed products.

Another proof can be based on an Ando type theorem of Solel [56, Theorem 4.4]. He shows that any representation of a product system over $\mathbb{N}^2$ coextends to an isometric representation. One can think of a product system over $\mathbb{N}^2$ as a pair of C*-correspondences over $\mathbb{N}$ together with commutation relations. Here we only need a special case, where there is one C*-correspondence $E$ over $\mathbb{N}$, and take the second correspondence to be $F = C$ with the relations that $F$ commutes with $E$. Then a representation $\rho$ of $\mathcal{T}^+(E)$ which commutes with a contraction $X$ determines a representation of the product system. Applying Solel’s Theorem yields the desired isometric lifting. This verifies ICLT and, in fact, the weak Ando property. Now Theorem 7.10 shows that $\mathcal{T}^+(E)$ has the Ando property.

We finish this section by giving the intertwining version of the Ando property.

**Theorem 7.15.** Suppose that $\mathcal{A}$ has the Ando property. Suppose also that $\rho_1$ and $\rho_2$ are representations of $\mathcal{A}$ on $\mathcal{H}_1$ and $\mathcal{H}_2$ and $X$ is a contraction in $\mathcal{B}(\mathcal{H}_2, \mathcal{H}_1)$ satisfying

$$\rho_1(a)X = X\rho_2(a) \quad \text{for all } a \in \mathcal{A}.$$  

Given a fully extremal coextension $\sigma_2$ of $\rho_2$, there exist a fully extremal coextension $\tilde{\sigma}_1$ of $\rho_1$ on $\mathcal{K}_1$, a maximal representation $\pi$ determining a fully extremal coextension $\tilde{\sigma}_2 = \sigma_2 \oplus \pi$ of $\rho_2$ on $\mathcal{K}_2$, and an isometry $W \in \mathcal{B}(\mathcal{K}_2, \mathcal{K}_1)$ so that

$$P_{\mathcal{H}_1}W = XP_{\mathcal{H}_2},$$

and

$$\tilde{\sigma}_1(a)W = W\tilde{\sigma}_2(a) \quad \text{for all } a \in \mathcal{A}.$$  

**Proof.** The proof parallels the proof of Theorem 7.10 using the intertwining versions of CLT and ICLT. So we just sketch the plan. One starts with the fully extremal coextension $\sigma_2$ of $\rho_2$ on $\mathcal{K}_2$. By Theorem 5.12, coextend $X$ to $Y$ and $\rho_1$ to a fully extremal $\sigma_1$ on $\mathcal{K}_1$ so that

$$\sigma_1(a)Y = Y\sigma_2(a) \quad \text{for all } a \in \mathcal{A}.$$  

Then use Proposition 7.7 to coextend $Y$ to an isometry $V$ and $\sigma_1$ to a Shilov representation $\tau_1$ on $\mathcal{L}_1$ so that

$$\tau_1(a)V = V\tau_2(a) \quad \text{for all } a \in \mathcal{A}.$$  

Since $\sigma_i$ are fully extremal, we have $\tau_i = \sigma_i \oplus \tau_i'$. If $Z = P_{\mathcal{L}_1 \oplus \mathcal{K}_1}V|_{\mathcal{K}_2}$ is the 2,1 entry of $V$ with respect to the decompositions $\mathcal{L}_1 = \mathcal{K}_1 \oplus (\mathcal{L}_1 \oplus \mathcal{K}_1)$, then

$$\tau_i'(a)Z = Z\sigma_2(a) \quad \text{for all } a \in \mathcal{A}.$$
Since \( \tau_i \) are Shilov, so are \( \tau_i' \). So choose maximal representations \( \pi_i \) on \( P_i \) which have invariant subspaces \( M_i \) so that \( \pi_i|_{M_i} \simeq \tau_i' \). Set
\[
\tilde{\sigma}_1 = \sigma_1 \oplus (\pi_1 \oplus \pi_2)^{(\infty)} \quad \text{and} \quad \tilde{\sigma}_2 = \sigma_2 \oplus (\pi_2 \oplus \pi_1)^{(\infty)}.
\]
Then the isometry
\[
W \in B(K_2 \oplus (P_2 \oplus P_1)^{(\infty)}, K_1 \oplus (P_1 \oplus P_2)^{(\infty)}
\]
described in the proof of Theorem 7.10 is the desired intertwiner. The details are left to the reader.

8 Incidence Algebras

An incidence algebra is a subalgebra \( \mathcal{A} \) of the \( n \times n \) matrices, \( \mathbb{M}_n \), containing the diagonal algebra \( \mathbb{D}_n \) with respect to a fixed orthonormal basis. Clearly, \( \mathcal{A} \) is spanned by the matrix units \( E_{ij} \) that it contains. One can define a partial order \( R \) on \( \{1, 2, \ldots, n\} \) by
\[
i \prec j \ (\text{or} \ (i, j) \in R) \quad \text{if} \quad E_{ij} \in \mathcal{A}.
\]
This can be identified with a directed graph, but note that generally the algebra is not the same as the tensor algebra of the graph. There is a reduced partial order obtained by identifying equivalent indices
\[
i \equiv j \quad \text{if} \quad i \prec j \quad \text{and} \quad j \prec i.
\]
The representation theory of the algebra of a partial order and its reduced partial order are related simply by multiplicity.

The algebra \( \mathcal{A} \cap \mathcal{A}^* \) is a C*-algebra containing the diagonal \( \mathbb{D}_n \), and is spanned by \( \{E_{ij} : i \equiv j\} \). A representation \( \rho \) restricts to a completely contractive representation of \( \mathcal{A} \cap \mathcal{A}^* \), and thus is a \( * \)-representation. So each diagonal matrix unit \( E_{ii} \) is sent to an orthogonal projection \( P_i = \rho(E_i) \) onto a subspace \( \mathcal{H}_i \). Since \( \rho \) is unital,
\[
\mathcal{H} = \sum_{1 \leq i \leq n} \mathcal{H}_i.
\]
In general, there are contractions \( T_{ij} \in B(\mathcal{H}_j, \mathcal{H}_i) \) so that
\[
\rho(E_{ij}) = P_i T_{ij} P_j.
\]
When \( i \equiv j \), \( T_{ij} \) is a unitary and \( T_{ji} = T_{ij}^* \). The homomorphism property shows that
\[
T_{ik} = T_{ij} T_{jk} \quad \text{when} \quad i \prec j \prec k.
\]
These relations are sufficient to determine an algebraic homomorphism.
Not all choices of contractions \( \{ T_{ij} : (i, j) \in R \} \) yield a completely contractive representation in general. However this does hold in some situations. Paulsen and Power [46] establish this for nest algebras. Davidson, Paulsen and Power [22] establish this for bilateral tree algebras. These are the algebras where the reduced relation is generated as a transitive relation by a directed bilateral tree (a directed graph with no loops). Finally the class of all such algebras was determined by Davidson in [15] as the interpolating graphs.

Muhly and Solel consider unilateral tree algebras, which are the incidence algebras \( A \) which are semi-Dirichlet. They show that graphically means that the relation is generated by a directed unilateral tree (each vertex is the range of at most one edge and there are no loops). These incidence algebras actually coincide with the tensor algebra of the unilateral tree because there is always at most one edge into each vertex. For example, the algebra

\[ A = \text{span}\{E_{11}, E_{22}, E_{33}, E_{13}, E_{23}\}, \]

which is determined by the graph formed by edges from \( v_3 \) to each of \( v_1 \) and \( v_2 \), is a unilateral tree algebra. However \( A^* \), which is determined by edges from \( v_1 \) and \( v_2 \) into \( v_3 \) is not a unilateral tree algebra, but it is a bilateral tree algebra. See [40, Chapter 5] for a discussion of “trees and trees”.

If \( \rho \) is a representation of an incidence algebra \( A \) as above, then a coextension \( \sigma \) will act on a Hilbert space

\[ \mathcal{K} = \sum_{1 \leq i \leq n} \oplus \mathcal{K}_i, \]

where \( \mathcal{K}_i \supset \mathcal{H}_i \) is the range of the projection \( \sigma(E_{ii}) \), and is determined by coextensions \( V_{ij} \) of \( T_{ij} \) in \( B(K_j, K_i) \) of the form

\[ V_{ij} = \begin{bmatrix} T_{ij} & 0 \\ D_{ij} & S_{ij} \end{bmatrix} \]

with respect to the decompositions \( \mathcal{K}_i \supset \mathcal{H}_i \oplus \mathcal{K}'_i \). The homomorphism property requires that

\[ V_{ij}V_{jk} = V_{ik} \quad \text{whenever} \quad i \prec j \prec k. \]

In general, these are complicated relations to dilate. One of the simplest examples where things get complicated is the \( 2k \)-cycle graph \( C_{2k} \) for \( k \geq 2 \). This graph has vertices \( \{1, 2, \ldots, 2k\} \) and edges

\[ 2i + 1 \succ 2i, \quad 2i + 1 \succ 2i + 2 \quad \text{and} \quad 1 \succ 2k. \]

The algebra for this graph has representations such that \( \rho(e_{ij}) = T_{ij} \) are all contractions, but \( \|\rho\|_{cb} > 1 \) [15, Theorem 2.2]. This is related to the famous example of Parrott [44] for \( \Lambda(D^3) \) and a similar example due to Paulsen and Power [47, Theorem 3.1] for the incidence algebra \( T_2 \otimes T_2 \otimes T_2 \), where \( T_2 \) is
the algebra of $2 \times 2$ upper triangular matrices. Also see the exposition in [14, Example 20.27].

The case of bilateral tree algebras is more conducive to analysis.

**Theorem 8.1.** Let $A$ be a bilateral tree incidence algebra. Then a representation $\rho$ is an extremal coextension if and only if each edge $E_{ij}$ is mapped to a partial isometry $V_{ij}$ such that $V_{ij}^* V_{ij} = \rho(E_{ii})$.

If $\rho$ is a representation, then a coextension $\sigma$ of $\rho$ is fully extremal if and only if it is extremal and $K_i = H_i \vee V_{ij} K_j$ for all edges of the tree.

**Proof.** The key observation from [22] is that every matrix unit in $A$ factors uniquely as a product of matrix units in the tree $T$, corresponding to the combinatorial fact that every edge in the transitive relation corresponds to the unique path on the tree from one vertex to another. Thus if for each matrix unit in the tree, $T_{ij} = \rho(E_{ij})|H_j$ is coextended to $V_{ij}$, then we can extend this definition to every matrix unit in a unique way; and the homomorphism property guarantees that each is a coextension of $T_{ij}$. It is possible to coextend each $T_{ij}$ in the tree to an isometry from $K_j$ to $K_i$. If more than one edge is entering a single vertex $i$, then one has to ensure that $K_i$ is large enough to accommodate all $T_{ij}$. (Of course, the ranges can overlap or even coincide.) By [22], this representation is still completely contractive. Thus to be extremal, each $\rho(E_{ij})$ needs to be isometric from $K_j$ into $K_i$.

Conversely, if each $T_{ij} = \rho(E_{ij})|H_j$ is an isometry on $H_j$, consider any coextension $\sigma$ of $\rho$. Then each $\sigma(E_{ii})$ has range $K_i \supset H_i$, and each isometry $T_{ij}$ coextends to a contraction $S_{ij} = \sigma(E_{ij})|K_j$. Therefore $S_{ij}|H_j = T_{ij}$ and $S_{ij}|K_j \cap H_j$ is a contraction with range orthogonal to $H_i$. Hence $\mathcal{K}' = \mathcal{K} \ominus H$ is reducing. Therefore $\sigma$ decomposes as a direct sum of $\rho$ and $\sigma' = \sigma|\mathcal{K}'$. Fully extremal coextensions of $\rho$ are more complicated. If we start with $\rho$, and coextend to an extremal $\sigma$, so that each $T_{ij}$ is coextended to an isometry $V_{ij}$, it may be possible to dilate $\sigma$ to $\tau \succ c \rho$ if there is some ‘room’ left. More precisely, if for some edge $E_{ij}$ in the tree, we have

$$K_i \neq H_i \vee V_{ij} K_j,$$

then one can extend $\sigma$ to $\tau$, a coextension of $\rho$, to use this extra space. Pick a vector $e \in K_j = K_i \ominus H_i$ which is orthogonal to the range of $V_{ij}$. Form

$$\tilde{K}_j = K_j \oplus C e.$$

Extend $V_{ij}$ to

$$\tilde{V}_{ij} := V_{ij} + e f^*.$$

It is apparent that this is indeed a coextension of $\rho$ and an extension of $\sigma$ which is not obtained as a direct sum. So $\sigma$ is not fully extremal.
Conversely, if for all edges $i ≺ j$ of the tree,

$$K_i = H_i \vee V_{ij}H_j$$

then for any extension $\tau$ of $\sigma$, the operators $\tilde{V}_{ij} = \tau(E_{ij})$ will have to map any new summand of $\tilde{K}_j$, namely $\tilde{K}_j \oplus K_i$, to a space orthogonal to $H_i$ because it is a coextension of $\rho$, and orthogonal to $V_{ij}K_j$, because $\tilde{V}_{ij}$ is a contraction. Hence by hypothesis, it maps into $\tilde{K}_i \oplus K_i$. This makes it apparent that $\tau$ splits as a direct sum of $\sigma$ and another representation. Therefore $\sigma$ is fully extremal.

**Corollary 8.2.** Every Shilov extension of a bilateral tree algebra $A$ is extremal.

**Proof.** A maximal representation $\pi$ of $A$ extends to a $*$-representation of $\mathcal{M}_n$. In particular, each $\pi(E_{ij})$ is a unitary from $H_j$ onto $H_i$. Thus any restriction $\sigma$ to an invariant subspace sends each vertex to a projection onto a subspace $K_i \subset H_i$ and each edge $E_{ij}$ to an isometry of $K_j$ into $K_i$. By Theorem 8.1, $\sigma$ is extremal.

**Remark 8.3.** Since $\rho(E_{ii})$ can be 0, an isometry can be vacuous. So for the algebra $A_n$ of Example 3.2, the coextensions $\sigma_{2i}$ are all extremal. But to be fully extremal, each edge has to be mapped to an isometry with maximal range, so only $\sigma_n$ is fully extremal.

**Uniqueness.** Let us explain why only the unilateral tree algebras have unique minimal fully extremal coextensions. The unilateral tree algebras are semi-Dirichlet. So this property is a consequence of Theorem 4.6.

A typical example is

$$A = \text{span}\{E_{21}, E_{31}, D_3\} \subset \mathcal{M}_3.$$  

Consider a representation $\rho$ on $H = H_1 \oplus H_2 \oplus H_3$ where $\rho(E_{ii}) = T_i$ for $i = 2, 3$. Let

$$D_i = P_{D_i}(I - T_i^*T_i)^{1/2} \in B(H_i, D_i),$$

where

$$D_i = \text{Ran}(I - T_i^*T_i)$$

is the closed range of $(I - T_i^*T_i)^{1/2}$. Coextend $\rho$ to $\sigma$ on $K = K_1 \oplus K_2 \oplus K_3$ where

$$K_1 = H_1 \quad \text{and} \quad K_i = H_i \oplus D_i \quad \text{for} \quad i = 2, 3$$

by setting

$$V_{ii} = \begin{bmatrix} T_i \\ D_i \end{bmatrix}.$$  

This is easily seen to be a fully extremal coextension by the previous proposition.
Any other isometric coextension $\tau$ will act on a Hilbert space $\mathcal{L}$ where $\mathcal{L}_i = \mathcal{H}_i \oplus \mathcal{L}'_i$, and

$$\tau(E_{1i}) = \begin{bmatrix} T_i & 0 \\ U_i D_i & S_i \end{bmatrix}$$

for $i = 2, 3$.

Here $U_i$ is an isometric embedding of $\mathcal{D}_i$ into $\mathcal{L}'_i$ and $S_i$ is an isometry of $\mathcal{L}'_i$ into $\mathcal{L}'_i$ with range orthogonal to the range of $\begin{bmatrix} T_i & 0 \\ U_i D_i \end{bmatrix}$.

A bit of thought shows that this splits as a direct sum of a representation on

$$\mathcal{H}_1 \oplus (\mathcal{H}_2 \oplus U_2 \mathcal{D}_2) \oplus (\mathcal{H}_3 \oplus U_3 \mathcal{H}_3)$$

which is unitarily equivalent to $\sigma$ and another piece.

On the other hand, any graph which is a bilateral tree but not a unilateral tree will have two edges mapping into a common vertex. The compression to this three dimensional space yields the algebra $\mathcal{A}^*$. We explain why $\mathcal{A}^*$ has non-unique minimal fully extremal coextensions.

Fix $\theta \in (0, \pi/4)$, and consider a representation $\rho$ on

$$\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3$$

where $\mathcal{H}_i = \mathbb{C}^2$ given by

$$\rho(E_{1i}) = T = \begin{bmatrix} \cos \theta & 0 \\ 0 & \sin \theta \end{bmatrix}$$

for $i = 2, 3$.

Let

$$D = (I - T^*T)^{1/2} = \begin{bmatrix} \sin \theta & 0 \\ 0 & \cos \theta \end{bmatrix}.$$ 

Then for any unitaries $U_i$ in the $2 \times 2$ unitary group $\mathfrak{U}_2$, we can coextend $\rho$ to an isometric representation $\sigma$ on

$$\mathcal{K} = \mathbb{C}^4 \oplus \mathbb{C}^2 \oplus \mathbb{C}^2$$

by setting

$$V_i = \begin{bmatrix} T_i \\ U_i D_i \end{bmatrix}$$

for $i = 2, 3$.

These are all fully extremal coextensions of $\rho$ by Theorem 8.1. Moreover they are evidently minimal.

Consider when two such representations will be unitarily equivalent via a unitary $W$ which is the identity on $\mathcal{H}$, and thus has the form

$$(I_2 \oplus V) \oplus I_2 \oplus I_2.$$ 

Conjugation by $W$ carries $\sigma$ to the representation which replaces $U_i D$ by $V U_i D$ for $i = 1, 2$. It is clear that one can arrange to match up the 1, 2 entry by appropriate choice of $V$. But that leaves no control over the 1, 3 entry. The possible minimal fully extremal coextensions of $\rho$ are parameterized by $\mathfrak{U}_2$.
Commutant lifting. Davidson, Paulsen and Power [22] showed that bilateral tree algebras have ICLT, and hence MCLT. They do not generally have SMCLT because of failure of unique dilations. Among finite dimensional incidence algebras, these are precisely the algebras with ICLT [15, Theorem 4.6].

Muhly and Solel show that unilateral tree algebras satisfy SCLT. This now can be seen from the fact that they have ICLT, whence MCLT. So by Theorem 4.6, we can obtain unique minimal fully extremal coextensions, and that every Shilov extension is fully extremal. So this implies SCLT.

The connected graphs which are unilateral trees and have adjoints which are unilateral trees as well are evidently chains. So the incidence algebras with this property are just direct sums of finite dimensional nest algebras. Since these algebras are Dirichlet, they have many good properties from section 4.

Proposition 8.4. Bilateral tree algebras have WCLT and WCLT*, and well as the weak Ando and Ando* properties.

Proof. If $A$ is a bilateral tree algebra, then so is $A^*$. So it suffices to prove WCLT. Let $\rho$ be a representation of $A$, and let $X$ be a contraction commuting with $\rho(A)$. By [22], $A$ has ICLT. Hence by Theorem 7.3 (ii), there is a Shilov coextension $\sigma$ and an isometric coextension $V$ of $X$ on $K \supset H$ which commute. By Corollary 8.2, $\sigma$ is extremal. Thus $A$ has WCLT and the weak Ando property.

The goal now is to refine this construction to obtain fully extremal coextensions to obtain CLT and hence the Ando property. We begin by establishing the Ando property for $T_2$, the $2 \times 2$ upper triangular matrices. Since

$$T_2 = \text{span}\{E_{11}, E_{12}, E_{22}\},$$

a representation $\rho$ is determined by $\rho(E_{ii}) = P_i = P_{H_i}$, where $H = H_1 \oplus H_2$, and a contraction $X \in B(H_2, H_1)$ where $\rho(E_{12}) = P_1 XP_2$. A contraction $A$ commuting with $\rho(T_2)$ commutes with $P_i$, and so has the form $A = A_1 \oplus A_2$; plus we have $A_1 X = X A_2$. Thus Ando’s Theorem for $T_2$ can be reformulated as a commutant lifting theorem:

Lemma 8.5. Suppose that $A_i \in B(H_i)$ for $i = 1, 2$ and $X \in B(H_2, H_1)$ are contractions such that

$$A_1 X = X A_2.$$ 

Then there are coextensions of $A_i$, $i = 1, 2$ and $X$ to isometries $\tilde{A}_i$ in $B(K_i)$ and $\tilde{X}$ in $B(K_2, K_1)$ so that

$$\tilde{A}_1 \tilde{X} = \tilde{X} \tilde{A}_2 \quad \text{and} \quad K_1 = H_1 \vee \tilde{X} K_2.$$
Proof. The algebra $T_2$ is a tree algebra, and so has the weak Ando property by the previous proposition. Hence there are isometric coextensions $B_i$ of $A_i$ and $Y$ of $X$ so that

$$B_1 Y = Y B_2,$$

acting on spaces $\mathcal{L}_1 \supset \mathcal{H}_i$. We will modify this to obtain the desired form.

Observe that the commutation relation implies that the range of $Y$ is invariant for $B_1$. Let

$$\mathcal{L} = \mathcal{H}_1 \vee Y \mathcal{L}_2; \quad \text{and} \quad B'_1 = P_\mathcal{L} B_1 |_\mathcal{L}.$$ 

Let $Y' \in \mathcal{B}(\mathcal{L}_2, \mathcal{L})$ be $Y$ considered as an operator into $\mathcal{L}$. Then

$$B'_1 Y' = P_\mathcal{L} B_1 P_\mathcal{L} Y = P_\mathcal{L} B_1 Y = P_\mathcal{L} Y B_2 = Y' B_2.$$ 

Also since $Y'$ is an isometry,

$$B_2 = Y'^* B'_1 Y'.$$

In particular, the commutation relations hold, and

$$\mathcal{H}_1 \vee Y \mathcal{L}_2 = \mathcal{L}.$$ 

The contraction $B'_1$ may no longer be an isometry, but it is a coextension of $A_1$.

Let $\tilde{A}_1$ be the minimal isometric dilation of $B'_1$ on $\mathcal{K}_1 \supset \mathcal{L}$. Write $\mathcal{K}'_1 = \mathcal{K}_1 \ominus \mathcal{L}$. With respect to $\mathcal{K}_1 = \mathcal{L} \oplus \mathcal{K}'_1$, we have

$$\tilde{A}_1 = \begin{bmatrix} B'_1 & 0 \\ C & D \end{bmatrix}.$$ 

Define

$$\mathcal{K}_2 = \mathcal{L}_2 \oplus \mathcal{K}'_1 \quad \text{and} \quad \tilde{X} = Y' \oplus I \in \mathcal{B}(\mathcal{K}_2, \mathcal{K}_1).$$ 

Set

$$\tilde{A}_2 = \tilde{X}^* \tilde{A}_1 \tilde{X}$$ 

$$= \begin{bmatrix} Y'^* & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} B'_1 & 0 \\ C & D \end{bmatrix} \begin{bmatrix} Y' & 0 \\ 0 & I \end{bmatrix}$$ 

$$= \begin{bmatrix} Y'^* B'_1 Y' & 0 \\ CY' & D \end{bmatrix}$$ 

$$= \begin{bmatrix} B_2 & 0 \\ CY' & D \end{bmatrix}.$$ 

Thus $A_2$ is a coextension of $B_2$, and hence of $A_2$. It is easy to verify that

$$\tilde{A}_1 \tilde{X} = \tilde{X} \tilde{A}_2.$$ 

Since $\tilde{A}_1$ and $\tilde{X}$ are isometries, so is $\tilde{A}_2$. Moreover, we now have

$$\mathcal{H}_1 \vee \tilde{X} \mathcal{K}_2 = (\mathcal{H}_1 \vee Y' \mathcal{L}_2) \oplus \mathcal{K}'_1 = \mathcal{K}_1.$$ 

\[ \text{\hfill \blacksquare} \]
THEOREM 8.6. Bilateral tree algebras have the Ando and Ando* properties.

Proof. Again it suffices to establish the Ando property. We first assume that the \( A \cap A^* = D_n \), so that the relation and reduced relation coincide.

Before proceeding, we make a few observations and set some notation. Suppose that the tree \( T \) has vertices \( v_i \) for \( 1 \leq i \leq n \). Let \( \rho \) be a representation of the algebra \( A \) commuting with a contraction \( A \). Then since

\[
\rho(v_i) = P_i = P_{H_i}
\]

are pairwise orthogonal projections summing to the identity, we see that

\[
A = \sum_{1 \leq i \leq n} \oplus A_i
\]

where

\[
A_i = A|_{H_i} \in B(H_i).
\]

If \( e_{ij} \) is an edge of the tree, let

\[
T_{ij} = \rho(e_{ij})|_{H_j} \in B(H_j, H_i).
\]

We have

\[
A_i T_{ij} = T_{ij} A_j,
\]

and conversely any \( A \) with these two properties commutes with \( A \).

A finite bilateral tree has an elimination scheme, in the sense that every bilateral tree has a vertex \( v \) which has at most one edge \( e \) such that either \( s(e) = v \) or \( r(v) = v \). This allows a proof by induction. So proceed by induction on the number of vertices.

If the graph has a single vertex, then it has no edges and \( A = C \). It is trivial to verify Ando’s property in this case.

Now suppose that the result holds for every bilateral tree on fewer than \( n \) vertices, and let \( T \) be a bilateral tree on \( n \) vertices. Let \( \rho \) and \( A \) be as above. We may assume that \( T \) is connected; for otherwise we may dilate each component by the induction hypothesis. So every vertex has an edge. Let \( v_{i_0} \) be a vertex with one edge \( e \). Restrict the representation to \( T \setminus \{v_{i_0}, e\} \) acting on \( H_{i_0} \), called \( \rho' \), commuting with \( A' = A|_{H_{i_0}} \). Use the induction hypothesis to coextend \( \rho' \) to a fully extremal coextension \( \sigma' \) commuting with an isometric coextension \( B' \) of \( A' \). Let

\[
\text{Ran} \sigma'(E_{ii}) =: \mathcal{L}'_i \quad \text{for} \quad i \neq i_0.
\]

Then

\[
B' = \sum_{i \neq i_0} \oplus B'_i.
\]
There are two cases: either \( s(e) = v_{i_0} \) and \( r(e) = v_{j_0} \) or \( s(e) = v_{j_0} \) and \( r(e) = v_{i_0} \). Assume the former. Let
\[
\rho(e) = X \in \mathcal{B}(\mathcal{H}_{i_0}, \mathcal{H}_{j_0}).
\]
Then
\[
A_{j_0}X = XA_{i_0}.
\]
Use Lemma 8.5 to coextend \( A_{i_0}, A_{j_0} \) and \( X \) to isometries \( \tilde{A}_{i_0}, \tilde{A}_{j_0} \) and \( \tilde{X} \) so that
\[
\tilde{A}_{j_0}\tilde{X} = \tilde{X}\tilde{A}_{i_0} \quad \text{and} \quad \mathcal{K}_{j_0} = \mathcal{H}_{j_0} \vee \tilde{X}\mathcal{K}_{i_0}.
\]
We can decompose
\[
\tilde{A}_{j_0} = V_{j_0} \oplus W_{j_0},
\]
where \( V_{j_0} \) is the minimal isometric coextension of \( A_{j_0} \), with respect to
\[
\mathcal{K}_{j_0} = \mathcal{K}_{j_0}^0 \oplus \mathcal{K}_{j_0}^1.
\]
Similarly, decompose the isometry
\[
B'_{j_0} = V_{j_0} \oplus W'_{j_0}
\]
with respect to
\[
\mathcal{L}'_{j_0} \simeq \mathcal{K}_{j_0}^0 \oplus \mathcal{L}_{j_0}^1.
\]
Define the coextension \( \sigma \) of \( \rho \) and isometric coextension \( \mathcal{B} \) of \( A \) as follows. Set
\[
\mathcal{L}_i = \mathcal{L}_i' \oplus \mathcal{K}_{j_0}^1 \quad \text{for} \quad i \neq i_0 \quad \text{and} \quad \mathcal{L}_{i_0} = \mathcal{K}_{i_0} \oplus \mathcal{L}_{j_0}^1.
\]
Define
\[
\sigma(e_{ij}) = \sigma'(e_{ij}) \oplus (E_{ij} \oplus I_{\mathcal{K}_{j_0}^1}) \quad \text{for} \quad i \neq i_0 \quad \text{and} \quad \sigma(e) = \tilde{X} \oplus I_{\mathcal{L}_{j_0}^1}.
\]
\[
B_i = B'_i \oplus W_{i_0} \quad \text{for} \quad i \neq i_0 \quad \text{and} \quad B_{i_0} = \tilde{A}_{i_0} \oplus W_{j_0}.
\]
Here \( E_{ij} \oplus I_{\mathcal{K}_{j_0}^1} \) is interpreted as the unitary that maps the copy of \( \mathcal{K}_{j_0}^1 \) contained in \( \mathcal{L}_i \) to the corresponding copy in \( \mathcal{L}_i' \). One needs only verify that each \( \sigma(e_{ij}) \) intertwines \( B_j \) with \( B_i \) and
\[
\mathcal{H}_i \vee \sigma(e_{ij})\mathcal{L}_j = \mathcal{L}_i.
\]
Both of these facts are routine. Thus a fully extremal coextension of \( \rho \) is produced that commutes with an isometric coextension of \( A \). This verifies Ando’s property.

The second case, in which the edge \( e \) maps \( v_{j_0} \) to \( v_{i_0} \) is handled similarly by first dilating the graph on \( n - 1 \) vertices, and producing a dilation of the one edge \( e \) using Lemma 8.5. Then as above, split the two isometries over the vertex \( v_{j_0} \) into the minimal isometric coextension direct summed with another isometry; and then define an explicit coextension with the desired properties.

The following consequence is immediate.

**Corollary 8.7.** Bilateral tree algebras have CLT and CLT*.
9 The Fuglede Property

We introduce another property of an abstract unital operator algebra reminiscent of the classical Fuglede Theorem that the commutant of a normal operator is self-adjoint.

**Definition 9.1.** Let \( \mathcal{A} \) be a unital operator algebra with C*-envelope \( C^*_e(\mathcal{A}) \). Say that \( \mathcal{A} \) has the Fuglede Property (FP) if for every faithful unital \(*\)-representation \( \pi \) of \( C^*_e(\mathcal{A}) \), one has

\[
\pi(\mathcal{A})' = \pi(C^*_e(\mathcal{A}))'.
\]

It is easy to characterize this property among abelian algebras.

**Proposition 9.2.** If \( \mathcal{A} \) is an abelian operator algebra, then the following are equivalent:

(i) \( \mathcal{A} \) has the Fuglede property.
(ii) \( \mathcal{A} \) is a function algebra.
(iii) \( C^*_e(\mathcal{A}) \) is abelian.

**Proof.** If (1) holds, then for every \( a \in \mathcal{A} \), \( \pi(a) \) lies in \( \pi(\mathcal{A})' \) and hence in \( \pi(C^*_e(\mathcal{A}))' \). Thus \( \pi(C^*_e(\mathcal{A})) \) is abelian. So (3) holds. If (3) holds, then \( \mathcal{A} \) is a function algebra since \( \mathcal{A} \) separates points by the definition of the C*-envelope. Finally if (2) holds, then \( \pi(\mathcal{A}) \) is contained in \( \pi(C(X)) \) which is an algebra of commuting normal operators. So the FP property follows from the usual Fuglede Theorem. \( \blacksquare \)

The following is a useful class of operator algebras which has the FP property.

**Proposition 9.3.** Suppose that there is a family \( \{U_k = [a^{(k)}_{ij}]\} \) of unitary matrices \( U_k \in M_{m_k,n_k}(\mathcal{A}) \) such that the set of matrix coefficients \( \{a^{(k)}_{ij}\} \) generate \( \mathcal{A} \) as an operator algebra. Then \( \mathcal{A} \) has FP.

**Proof.** If \( B \) commutes with \( \pi(\mathcal{A}) \), then

\[
B^{(m)}\pi(U_k) = \pi(U_k)B^{(n)}.
\]

By the Fuglede–Putnam Theorem, we obtain

\[
B^{*(m)}U_k = U_kB^{*(n)}.
\]

Therefore \( B^* \) commutes with each \( a^{(k)}_{ij} \). As these generate \( \mathcal{A} \), we deduce that \( \pi(\mathcal{A})' \) is self-adjoint. \( \blacksquare \)
Example 9.4. The non-commutative disk algebras of Popescu, $\mathfrak{A}_n$, are generated by a row isometry $S = [S_1 \ldots S_n]$. The C*-envelope is the Cuntz algebra $\mathcal{O}_n$. As an element of $\mathfrak{M}(\mathcal{O}_n)$, $S$ is a unitary operator. Thus $\mathfrak{A}_n$ has the FP property. This property has been explicitly observed in [19, Proposition 2.10].

Example 9.5. The algebra $\mathfrak{A}_\infty$ generated by a countable family of isometries with pairwise orthogonal ranges does not have the Fuglede property. This is because $*$-representations of the C*-envelope, $\mathcal{O}_\infty$, are determined by any countably infinite family of isometries with pairwise orthogonal ranges, and does not force the sum of these ranges to be the whole space. In the left regular representation, the commutant of $\mathfrak{A}_\infty$ is the wot-closed algebra generated by the right regular representation, which is not self-adjoint.

Example 9.6. The algebra $\mathcal{A}_d$ of continuous multipliers on the Drury-Arveson space $H^2_d$ is abelian, but the norm is not the sup norm. So this is not a function algebra. Arveson [8] identifies the C*-envelope, which contains the compact operators; so it is clearly not abelian.

Since $\mathcal{A}_d$ is a quotient of $\mathfrak{A}_d$, one sees that FP does not pass to quotients. One can also see this by noting that there are quotients of functions algebras which are not themselves functions algebras.

Example 9.7. The tensor algebra of any finite graph has FP. It does not follow immediately from Proposition 9.3, but does follow by a simple modification. A finite graph $G = (V,E,r,s)$ consists of a finite set $V$ of vertices, a finite set $E$ of edges, and range and source maps $r, s : E \to V$. The graph C*-algebra $C^*(G)$ is the universal C*-algebra generated by pairwise orthogonal projections $p_v$ for $v \in V$ and partial isometries $u_e$ for $e \in E$ such that

$$\sum_{v \in V} p_v = 1, \quad u_e^* u_e = p_{s(e)} \quad \text{and} \quad \sum_{r(e) = v} u_e u_e^* = p_v$$

unless $v$ is a source, meaning that there are no edges with range $v$. The tensor algebra of the graph $T^+(G)$ is the non-self-adjoint subalgebra of $C^*(G)$ generated by

$$\{p_v, u_e : v \in V, e \in E\}.$$ 

Then $C^*_s(T(G)) = C^*(G)$ [29, 33]. Suppose that $\pi$ is a $*$-representation of $C^*(G)$ and $T \in \pi(T(G))'$. Then $T$ commutes with $\pi(p_v) =: P_v$, and thus $T = \oplus_{v \in V} T_v$ where $T_v \in B(P_v \mathcal{H})$. If there are edges with $r(e) = v$, say $e_1, \ldots, e_k$, then let $s(e_i) = v_i$ and consider

$$U = [\pi(u_{e_1}) \ldots \pi(u_{e_k})] \in B(P_{v_1} \mathcal{H} \oplus \ldots P_{v_k} \mathcal{H}, P_v \mathcal{H}).$$

This is a unitary element, and

$$T_v U = U (T_{v_1} \oplus \cdots \oplus T_{v_k}).$$

Documenta Mathematica 16 (2011) 781–868
By the Fuglede-Putnam Theorem, we also obtain
\[ T^*_v U = U(T^*_v \oplus \cdots \oplus T^*_v) \]
If there are no edges with range \( v \), there is nothing to check. We deduce as in Proposition 9.3 that \( \mathcal{T}^+(G) \) has FP.

**Example 9.8.** The algebra of any finite \( k \)-graph has FP. This is established as in the case of a 1-graph.

**Example 9.9.** Let \( \mathcal{N} \) be a nest (a complete chain of closed subspaces of a Hilbert space). Set
\[ \mathcal{A} = \mathcal{T}(\mathcal{N}) \cap \mathcal{R}^+ \]
where \( \mathcal{T}(\mathcal{N}) \) is the nest algebra [14] and \( \mathcal{R}^+ = CI + \mathcal{R} \) is unitization of the space of compact operators. By the Erdos Density Theorem, \( \mathcal{T}(\mathcal{N}) \cap \mathcal{R} \) contains a norm 1 approximate identity; and thus \( \mathcal{A} \) is weak-\( * \) dense in \( \mathcal{T}(\mathcal{N}) \). Therefore its commutant is trivial, and coincides with the commutant of \( \mathcal{R}^+ \), the enveloping C*-algebra. Moreover, \( \mathcal{R}^+ \) is the C*-envelope because \( \mathcal{A} \) is the only ideal, and the quotient \( q \) of \( \mathcal{R}^+ \) onto \( C \) is clearly not isometric on \( \mathcal{A} \). A \( * \)-representation of \( \mathcal{R}^+ \) has the form
\[ \pi(A) = q(A)I_{K_0} \oplus A \otimes I_{K_1} \]
on a Hilbert space \( K = K_0 \oplus (H \otimes K_1) \). By the earlier remark, the commutant of \( \pi(A) \) is seen to be
\[ B(K_0) \oplus (CI_H \otimes B(K_1)) \]
which is the commutant of \( \pi(\mathcal{R}^+) \). So \( \mathcal{A} \) has FP.

In particular, any finite dimensional nest algebra \( \mathcal{T} \subset \mathcal{M}_n \) has FP.

**Example 9.10.** More generally, let \( \mathcal{L} \) be a completely distributive commutative subspace lattice (see [14]); and let \( \text{Alg}(\mathcal{L}) \) be the corresponding CSL algebra. Let \( \mathfrak{M} \) be a masa containing (the projections onto) \( \mathcal{L} \). Also let
\[ \mathfrak{M} = \text{Alg}(\mathcal{L} \cap \mathcal{L}^\perp) \]
Observe that \( \mathcal{L} \cap \mathcal{L}^\perp \) is a completely distributive Boolean algebra, and thus is atomic. Therefore
\[ \mathfrak{M} = \oplus \sum_i B(\mathcal{H}_i) \]
is an \( \ell^\infty \) direct sum with respect to the decomposition \( \mathcal{H} = \oplus \sum_i \mathcal{H}_i \), where \( \mathcal{H}_i \) are the ranges of the atoms of \( \mathcal{L} \cap \mathcal{L}^\perp \).

Also by complete distributivity, \( \text{Alg}(\mathcal{L}) \cap \mathfrak{M} \) has a norm one approximate identity. So again \( \mathcal{A} = \text{Alg}(\mathcal{L}) \cap \mathcal{R}^+ \) is weak-\( * \) dense in \( \text{Alg}(\mathcal{L}) \). It is straightforward to see that
\[ C_0^*(\mathcal{A}) = \{ M + \sum_i A_i : A_i \in \mathcal{R}^+(\mathcal{H}_i) \text{ and } \lim_i A_i = 0 \} \]
The irreducible representations of $C^*_e(A)$ are unitarily equivalent to compression to some $H_i$ and the character that evaluates $\lambda$. So every representation is a direct sum of these irreducible ones with multiplicity. As in the nest case, it is straightforward to show that

$$\pi(A)' = \pi(C^*_e(A))'.$$

So $A$ has FP.

We have one minor result relating FP with commutant lifting.

**Proposition 9.11.** If an operator algebra $A$ has FP and MCLT, then it has ICLT.

*Proof.* Suppose that $\rho$ is a representation of $A$ on $H$ and $X \in B(H)$ is a contraction commuting with $\rho(A)$. Then by MCLT, there is a maximal dilation $\pi$ of $\rho$ and a contractive dilation $Y$ of $X$ which commutes with $\pi(A)$ and has $H$ as a common semi-invariant subspace. Since $\pi$ is maximal, it extends to a $*$-representation of $C^*_e(A)$ which we also denote by $\pi$. By the Fuglede property, $Y$ commutes with $\pi(C^*_e(A))$. Hence $C^*(Y)$ is contained in $\pi(C^*_e(A))'$. The standard Schaeffer dilation of $Y$ to a unitary $U$ on $K(\infty)$ has coefficients in $C^*(Y)$. So $U$ commutes with $\pi^{(\infty)}$. This establishes ICLT.

**Corollary 9.12.** If an operator algebra $A$ has properties FP, WCLT and WCLT*, then it has the weak Ando property.

If an operator algebra $A$ has properties FP, CLT and CLT*, then it has the Ando property.

*Proof.* Theorem 5.14 shows that WCLT and WCLT* imply MCLT. So by the preceding proposition, we obtain ICLT. By Proposition 7.9 and Theorem 7.10, WCLT and ICLT imply the weak Ando property and CLT and ICLT imply the Ando property.

10 **Completely Isometric Endomorphisms**

In the category of operator algebras, the natural morphisms are completely bounded maps. Among the endomorphisms, those that work best for semi-crossed products are the completely isometric ones. These are the analogue of the faithful $*$-endomorphisms of C*-algebras. In this section, we investigate lifting completely isometric endomorphisms of operator algebras to $*$-endomorphisms of some C*-cover.

Let $\text{Aut}(A)$ denote the completely isometric automorphisms of an operator algebra $A$. If $\mathfrak{A}$ is a C*-algebra, then (completely) isometric automorphisms
are automatically $*$-automorphisms. If $\mathcal{A} \subset \mathfrak{A}$, let
\[
\text{Aut}_{\mathcal{A}}(\mathfrak{A}) = \{ \alpha \in \text{Aut}(\mathfrak{A}) : \alpha(\mathcal{A}) = \mathcal{A} \}.
\]
Similarly, let $\text{End}(\mathcal{A})$ denote the completely isometric unital endomorphisms of $\mathcal{A}$. Again, for a C*-algebra, these are faithful unital $*$-endomorphisms. When $\mathcal{A} \subset \mathfrak{A}$, we let
\[
\text{End}_{\mathcal{A}}(\mathfrak{A}) = \{ \alpha \in \text{End}(\mathfrak{A}) : \alpha(\mathcal{A}) \subset \mathcal{A} \}.
\]

First we begin with an easy result.

**Proposition 10.1.** Let $\mathcal{A}$ be a unital operator algebra. Then every completely isometric automorphism of $\mathcal{A}$ lifts to a $*$-automorphism of $C^*_e(\mathcal{A})$ which fixes $\mathcal{A}$ (as a set). Thus
\[
\text{Aut}_{\mathcal{A}}(C^*_e(\mathcal{A})) \cong \text{Aut}(\mathcal{A})
\]
by restriction to $\mathcal{A}$.

**Proof.** Let $\alpha \in \text{Aut}(\mathcal{A})$. Consider $\mathcal{A}$ as a subalgebra of $C^*_e(\mathcal{A})$. The map $\alpha$ takes $\mathcal{A}$ completely isometrically and isomorphically onto itself, and the image generates $C^*_e(\mathcal{A})$ as a C*-algebra. By the basic property of C*-envelopes, $\alpha$ extends to a $*$-homomorphism $\tilde{\alpha}$ of $C^*_e(\mathcal{A})$ onto itself. The kernel of $\tilde{\alpha}$ is a boundary ideal because the map is completely isometric on $\mathcal{A}$, and hence is $\{0\}$. Thus $\tilde{\alpha}$ is an automorphism which fixes $\mathcal{A}$ as a set. The converse is evident.

The restriction of $\alpha \in \text{Aut}_{\mathcal{A}}(C^*_e(\mathcal{A}))$ to $\mathcal{A}$ is injective since $\mathcal{A}$ generates $C^*_e(\mathcal{A})$ as a C*-algebra. Thus the restriction map is an isomorphism. \[\square\]

**Example 10.2.** It is not true that every $\alpha \in \text{End}(\mathcal{A})$ lifts to an endomorphism of $C^*_e(\mathcal{A})$. Take $\mathcal{A} = \Lambda(D)$ and let $\tau \in \Lambda(D)$ be the composition of a conformal map of $D$ onto the rectangle
\[
\{x + iy : -1 < x < 0 \text{ and } |y| \leq 10 \}
\]
followed by the exponential map. Thus $\tau$ maps $D$ onto the annulus
\[
\mathcal{A} := \{z : e^{-1} < |z| < 1 \}.
\]
It follows that
\[
\alpha(f) = f \circ \tau
\]
is a completely isometric endomorphism. However, since $\tau$ maps parts of $T$ into the interior of $D$, this map does not extend to an endomorphism of $C(T)$.

Observe that $\alpha$ lifts to an endomorphism of $C(\overline{D})$ by
\[
\tilde{\alpha}(f) = f \circ \tau.
\]
Unfortunately this map is not faithful, as its kernel is
\[ \ker \tilde{\alpha} = I(A) = \{ f \in C(D) : f|_A = 0 \}. \]

The remedy is a bit subtle. Let
\[ X = \bigcap_{n \geq 0} \tau^n(D). \]
This is a connected compact set with two key properties:
\[ \tau(X) = X \quad \text{and} \quad T \subset X. \]
The latter holds because \( \tau(T) \) contains \( T \). Now consider \( A(D) \) as a subalgebra of the C*-algebra \( C(X) \). The embedding is isometric because \( T \subset X \). This is a C*-cover by the Stone-Weierstrass Theorem. Here \( \alpha \) extends to
\[ \tilde{\alpha}(f) = f \circ \tau. \]
This is a faithful \(*\)-endomorphism because \( \tau \) is surjective on \( X \).

**Example 10.3.** Here is a different example. Take
\[ A = \Lambda(D) \oplus (c \otimes T) \]
where \( c \) is the space of convergent sequences and \( T = C^*(T_z) \) is the Toeplitz algebra generated by the shift \( T_z \) on \( H^2 \). It is easy to see that
\[ C^*_c(A) = C(T) \oplus (c \otimes T). \]
Write an element of \( A \) as \( f \oplus (T_n)_{n \geq 1} \), where \( \lim_{n \to \infty} T_n =: T_0 \) exists. Fix \( \lambda \in D \) and consider the map
\[ \alpha(f \oplus (T_n)_{n \geq 1}) = f(\lambda)I \oplus (T_f, T_{n-1})_{n \geq 2}. \]
This is evidently a completely isometric unital endomorphism. However one can restrict \( \alpha \) to a map \( \beta \) which takes \( \Lambda(D) \) to a subalgebra of \( T \) by \( \beta(f) = T_f \). The range of \( \beta \) generates \( T \) as a C*-algebra, which is non-abelian. Therefore there is no extension of \( \beta \) to a homomorphism of \( C(T) \) into \( T \). Thus \( \alpha \) does not lift to a \(*\)-endomorphism of \( C^*_c(A) \).

If \( |\lambda| = 1 \), we can embed \( A \) into
\[ \mathfrak{A} = T \oplus (c \otimes T) \]
in the natural way and extend \( \alpha \) to the endomorphism
\[ \tilde{\alpha}(T \oplus (T_n)_{n \geq 1}) = qT(\lambda)I \oplus (T, T_{n-1})_{n \geq 2}. \]
where \( q \) is the quotient map of \( \mathcal{T} \) onto \( C(T) \).

However if \(|\lambda| < 1\), evaluation at \( \lambda \) is not multiplicative on \( \mathcal{T} \), so \( \alpha \) does not lift to an endomorphism of \( \mathfrak{A} \). We can instead let

\[
\mathfrak{B} = \mathbb{C} \oplus \mathcal{T} \oplus (c \otimes \mathcal{T})
\]

and embed \( \mathcal{A} \) by

\[
j(f \oplus (T_n)_{n \geq 1}) = f(\lambda) \oplus T_f \oplus (T_n)_{n \geq 1}.
\]

Clearly \( j \) is completely isometric. The C*-algebra generated by \( j(\mathcal{A}(D)) \) is all of \( \mathfrak{B} \) because

\[
j(1) - j(z^*)j(z) = (1 - |\lambda|^2) \oplus 0
\]

shows that \( \mathbb{C} \oplus 0 \) is contained in this algebra.

Observe that evaluation at \( \lambda \) is now a character of \( \mathfrak{B} \). Moreover

\[
j(\alpha(f \oplus (T_n)_{n \geq 1})) = j(f(\lambda)I \oplus (T_f, T_{n-1})_{n \geq 2}) = f(\lambda) \oplus f(\lambda)I \oplus (T_f, T_{n-1})_{n \geq 2}.
\]

So we may extend \( \alpha \) to \( \tilde{\alpha} \in \text{End}(\mathfrak{B}) \) by

\[
\tilde{\alpha}(w \oplus T \oplus (T_n)_{n \geq 1}) = w \oplus wI \oplus (T, T_{n-1})_{n \geq 2}.
\]

A modification of Peters’ argument \([48, \text{Prop.I.8}]\) shows that we can extend completely isometric endomorphisms to automorphisms of a larger algebra.

**Proposition 10.4.** If \( \mathcal{A} \) is a unital operator algebra and \( \alpha \in \text{End}(\mathcal{A}) \), then there is a unital operator algebra \( \mathcal{B} \) containing \( \mathcal{A} \) as a unital subalgebra and \( \beta \in \text{Aut}(\mathcal{B}) \) such that \( \beta|_{\mathcal{A}} = \alpha \). Moreover, \( \mathcal{B} \) is the closure of \( \bigcup_{n \geq 0} \beta^{-n}(\mathcal{A}) \).

**Proof.** First observe that the standard orbit representation makes sense for \( \mathcal{A} \). Let \( \sigma \) be a completely isometric representation of \( \mathcal{A} \) on a Hilbert space \( \mathcal{H} \) so that \( C^*(\sigma(\mathcal{A})) \cong C^*_c(\mathcal{A}) \). Form the Hilbert space \( \tilde{\mathcal{H}} = \mathcal{H} \otimes \ell^2 \) and define

\[
\pi(a) = \sum_{n=0}^{\infty} \oplus \sigma(\alpha^n(a)) \quad \text{and} \quad V = I \otimes S,
\]

where \( S \) is the unilateral shift. Then it is evident that \( (\pi, V) \) is an isometric covariant representation of \( (\mathcal{A}, \alpha) \). In particular, \( \pi(\mathcal{A}) \) is completely isometric to \( \mathcal{A} \). Define the corresponding endomorphisms \( \tilde{\alpha} \) on \( \pi(\mathcal{A}) \) by

\[
\tilde{\alpha}(\pi(a)) = \tilde{\alpha}\left(\sum_{n=0}^{\infty} \oplus \sigma(\alpha^n(a))\right) = \sum_{n=0}^{\infty} \oplus \sigma(\alpha^{n+1}(a)) = \pi(\alpha(a))
\]

for \( a \in \mathcal{A} \).

**Documenta Mathematica** 16 (2011) 781–868
Form the injective system

\[
\begin{array}{ccccccc}
\pi(A) & \overset{\alpha}{\longrightarrow} & \pi(A) & \overset{\alpha}{\longrightarrow} & \pi(A) & \overset{\alpha}{\longrightarrow} & \cdots & B \\
\downarrow \alpha & & \downarrow \alpha & & \downarrow \alpha & & \downarrow \alpha & \\
\pi(A) & \overset{\alpha}{\longrightarrow} & \pi(A) & \overset{\alpha}{\longrightarrow} & \pi(A) & \overset{\alpha}{\longrightarrow} & \cdots & B \\
\end{array}
\]

Then \( B \) is a unital operator algebra containing \( A \) as a unital subalgebra and \( \beta \) is a completely isometric endomorphism. However it is evident that \( \beta \) is surjective, so \( \beta \) is an automorphism.

Finally, observe that \( B \) is the closure of \( \bigcup_{n \geq 0} \beta^{-n}(A) \).

Now we can use this to lift endomorphisms.

**Theorem 10.5.** If \( A \) is a unital operator algebra and \( \alpha \in \text{End}(A) \), then there is a \( C^* \)-cover \( \mathfrak{A} \) of \( A \) and an endomorphisms \( \bar{\alpha} \) of \( \mathfrak{A} \) such that \( \bar{\alpha}|_A = \alpha \).

**Proof.** Use Proposition 10.4 to lift \( \alpha \) to an automorphism \( \beta \) of a larger algebra \( B \). Then apply Proposition 10.1 to lift it again to an automorphism \( \bar{\beta} \) of the \( C^* \)-algebra \( B = C^*_e(B) \). Let \( \mathfrak{A} \) be the \( C^* \)-subalgebra of \( B \) generated by \( A \). Since \( \bar{\beta}|_A = \alpha \), we see that \( A \) is invariant under \( \bar{\beta} \). Hence so is \( \mathfrak{A}^* \). Since \( \mathfrak{A} \) is generated by \( A \) and \( \mathfrak{A}^* \), it is also invariant under \( \bar{\beta} \). So \( \bar{\alpha} = \bar{\beta}|_A \) is the desired \( * \)-endomorphism.

While not all endomorphisms of \( A \) lift to \( C^*_e(A) \), those that do lift behave well when lifted to larger algebras. This simplifies the hypotheses in some of the results in [32] as explained in the next section.

**Proposition 10.6.** Let \( A \) be a unital operator algebra and let \( \mathfrak{A} \) be a \( C^* \)-cover. Suppose that \( \alpha \in \text{End}_A(C^*_e(A)) \) and that \( \beta \in \text{End}_A(\mathfrak{A}) \) such that \( \beta|_A = \alpha|_A \). Then \( \alpha q = q \beta \), and hence \( \beta \) fixes the Shilov ideal \( \mathfrak{J}_A \), where \( q \) is the canonical quotient map of \( \mathfrak{A} \) onto \( C^*_e(A) \).

**Proof.** Observe that \( \alpha q \) and \( q \beta \) are \( * \)-homomorphisms of \( \mathfrak{A} \) into \( C^*_e(A) \) which agree on the generating set \( A \). Thus they are equal. Hence

\[ \mathfrak{J}_A = \ker \alpha q = \ker q \beta. \]

It follows that

\[ \mathfrak{J}_A = \{ a \in \mathfrak{A} : \beta(a) \in \mathfrak{J}_A \} = \beta^{-1}(\mathfrak{J}_A). \]

In particular,

\[ \beta(\mathfrak{J}_A) \subset \mathfrak{J}_A. \]

Extremal and fully extremal coextensions behave well under automorphisms, but not for endomorphisms.
Proposition 10.7. Let \( \rho \) be a representation of a unital operator algebra \( \mathcal{A} \) with (fully) extremal coextension \( \sigma \). If \( \alpha \in \text{Aut}(\mathcal{A}) \), then \( \sigma \circ \alpha \) is a (fully) extremal coextension of \( \rho \circ \alpha \).

Proof. Let \( \rho \) act on \( \mathcal{H} \) and \( \sigma \) act on \( \mathcal{K} \supset \mathcal{H} \). Suppose first that \( \sigma \) is extremal. Since \( \mathcal{H} \) is coinvariant for \( \sigma(\mathcal{A}) \), it is also coinvariant for \( \sigma(\alpha(\mathcal{A})) \). Thus \( \sigma \circ \alpha \) is a coextension of \( \rho \circ \alpha \). Suppose that \( \tau \) is a coextension of \( \sigma \circ \alpha \). Then \( \tau \circ \alpha^{-1} \) is a coextension of \( \sigma \). Hence it splits as

\[
\tau \circ \alpha^{-1} = \sigma \oplus \varphi.
\]

Thus

\[
\tau = \sigma \circ \alpha \oplus \varphi \circ \alpha.
\]

So \( \sigma \circ \alpha \) is extremal.

A similar proof works for fully extremal coextensions of \( \rho \circ \alpha \). \qed

Example 10.8. Let

\[
\mathcal{A} = \Lambda(\mathcal{D}) \oplus (c \otimes C(\overline{\mathcal{D}}))
\]

with elements \( (f, (g_n)_{n \geq 1}) \) where \( f \in \Lambda(\mathcal{D}), \) \( g_n \in C(\overline{\mathcal{D}}) \) for \( n \geq 1 \) and \( \lim_{n \to \infty} g_n =: g_0 \) exists. Define

\[
\alpha(f, (g_n)_{n \geq 1}) = (f(0), (f, g_{n-1})_{n \geq 2}).
\]

This is a completely isometric endomorphism.

A representation of \( \mathcal{A} \) has the form

\[
\rho(f, (g_n)_{n \geq 1}) = f(T) \oplus \sum_{n \geq 1} \oplus \rho_n(g_n)
\]

where \( T \) is a contraction and \( \rho_n \) are \( \ast \)-representations of \( C(\overline{\mathcal{D}}) \). It is straightforward to show that a representation of \( \mathcal{A} \) is an extremal coextension if and only if \( T \) is an isometry.

Taking \( T \) to be an isometry and \( \rho_n \) all vacuous (the zero representation on zero dimensional space), we have an extremal coextension \( \rho \) such that

\[
\rho \circ \alpha(f, (g_n)_{n \geq 1}) = f(0)I.
\]

Since 0 is not an isometry, this is not extremal.

11 A Review of Semicrossed Products

Nonself-adjoint crossed products were introduced by Arveson [4, 11] as certain concretely represented operator algebras that encoded the action of a subsemigroup of a group acting on a measure space, and later a topological space.

Documenta Mathematica 16 (2011) 781–868
K.R. Davidson and E.G. Katsoulis

McAsey, Muhly and Sato [37, 36] further analyzed such algebras, again relying on a concrete representation to define the algebra. In [48], Peters gave a more abstract and universal definition of the semicrossed product of a C*-algebra by a single endomorphism. Actually he provides a concrete definition, but then proves that it has the universal property which has become the de facto definition of a semicrossed product.

One can readily extend Peter’s definition of the semicrossed product of a C*-algebra by a ∗-endomorphism to unital operator algebras and unital completely isometric endomorphisms.

**Definition 11.1.** Let \( \mathcal{A} \) be a unital operator algebra and \( \alpha \in \text{End}(\mathcal{A}) \). A **covariant representation** of \((\mathcal{A}, \alpha)\) is a pair \((\rho, T)\) consisting of a completely contractive representation \( \rho : \mathcal{A} \to \mathcal{B}(\mathcal{H}) \) and a contraction \( T \in \mathcal{B}(\mathcal{H}) \) such that

\[
\rho(a)T = T\rho(\alpha(a)) \quad \text{for all } a \in \mathcal{A}.
\]

The **semicrossed product** \( \mathcal{A} \times_\alpha \mathbb{Z}_+ \) is the universal operator algebra for covariant representations. This is the closure of the algebra \( \mathcal{P}(\mathcal{A}, t) \) of formal polynomials of the form \( p = \sum_{i=0}^n t^i a_i \), where \( a_i \in \mathcal{A} \), with multiplication determined by

\[
\alpha t = \alpha(a)
\]

and the norm

\[
\|p\| = \sup_{(\rho, T) \text{ covariant}} \left\| \sum_{i=0}^n T^i \rho(a_i) \right\|.
\]

This supremum is clearly dominated by \( \sum_{i=0}^n \|a_i\| \); so this norm is well defined. Since this is the supremum of operator algebra norms, it is also easily seen to be an operator algebra norm. By construction, for each covariant representation \((\rho, T)\), there is a unique completely contractive representation \( \rho \times T \) of \( \mathcal{A} \times_\alpha \mathbb{Z}_+ \) into \( \mathcal{B}(\mathcal{H}) \) given by

\[
\rho \times T(p) = \sum_{i=0}^n T^i \rho(a_i).
\]

This is the defining property of the semicrossed product.

When \( \mathfrak{A} \) is a C*-algebra, the completely isometric endomorphisms are the faithful ∗-endomorphisms. Peters [48] shows that when \( \alpha \) is a faithful ∗-endomorphism of \( \mathfrak{A} \), the norm of \( \mathfrak{A} \times_\alpha \mathbb{Z}_+ \) can be computed by using orbit representations. Let \( \sigma \) be a faithful ∗-representation of \( \mathfrak{A} \) on \( \mathcal{H} \). Form the ∗-representation \( \pi \) on \( \mathcal{H} \otimes \ell^2 \) by

\[
\pi(a) = \sum_{n=0}^\infty \sigma(a^n(a))
\]

and let \( V = I \otimes S \), where \( S \) is the unilateral shift on \( \ell^2 \). It is evident that \((\pi, V)\) is a covariant representation of \((\mathfrak{A}, \alpha)\). The corresponding representation \( \pi \times V \) of \( \mathfrak{A} \times_\alpha \mathbb{Z}_+ \) is known as the **orbit representation** of \( \sigma \).
Theorem 11.2 (Peters). If \( \alpha \) is a faithful \(*\)-endomorphism of a C*-algebra \( \mathfrak{A} \), and \( \sigma \) is a faithful \(*\)-representation of \( \mathfrak{A} \), then the orbit representation \( \sigma \times V \) provides a completely isometric representation of \( \mathfrak{A} \times_\alpha \mathbb{Z}_+ \).

Moreover, Peters [48, Prop.I.8] establishes Proposition 10.4 in the case where \( \mathcal{A} \) is a C*-algebra. It follows [48, Prop.II.4] that \( \mathfrak{A} \times_\alpha \mathbb{Z}_+ \) is a C*-algebra. It follows [48, Prop.II.4] that \( \mathfrak{A} \times_\alpha \mathbb{Z}_+ \) is completely isometrically isomorphic to the subalgebra of the crossed product algebra \( \mathfrak{B} \times_\beta \mathbb{Z} \) generated as a non-self-adjoint algebra by \( j(\mathfrak{A}) \) and the unitary \( u \) implementing \( \beta \) in the crossed product. Kakariadis and the second author [32, Thm.2.5] show that this crossed product is the C*-envelope:

Theorem 11.3 (Kakariadis-Katsoulis). Let \( \alpha \) be a faithful \(*\)-endomorphism of a C*-algebra \( \mathfrak{A} \) and let \( (\mathfrak{B}, \beta) \) be the lifting of \( \alpha \) to an automorphism described above. Then

\[
C^*_\varepsilon(\mathfrak{B} \times_\beta \mathbb{Z}) = \mathfrak{B} \times_\beta \mathbb{Z}.
\]

Since dilation theorems fail in many classical cases, such as commuting triples of contractions [60, 44], one can circumvent the issue by considering only isometric covariant relations. This semicrossed product was introduced by Kakariadis and the second author [32]. The results there have immediate application for us.

Definition 11.4. Let \( \mathcal{A} \) be an operator algebra and let \( \alpha \in \text{End}(\mathcal{A}) \). A covariant representation \((\rho, T)\) of \((\mathcal{A}, \alpha)\) is called **isometric** if \( \rho \) is a complete isometry and \( T \) is an isometry. The **isometric semicrossed product** \( \mathcal{A} \times_\alpha \mathbb{Z}_+ \) is the universal operator algebra for isometric covariant representations. This is the closure of the algebra \( \mathcal{P}(\mathcal{A}, t) \) of formal polynomials of the form \( p = \sum_{i=0}^n t^i a_i \), where \( a_i \in \mathcal{A} \), under the norm

\[
\|p\| = \sup_{(\rho, T) \text{ isometric covariant}} \left\| \sum_{i=0}^n T^i \rho(a_i) \right\|.
\]

Theorem 11.5 (Kakariadis-Katsoulis). If \( \mathcal{A} \) is a unital operator algebra and \( \alpha \in \text{End}_\mathcal{A}(C^*_\varepsilon(\mathcal{A})) \), then \( \mathcal{A} \times_\alpha \mathbb{Z}_+ \) is (completely isometrically isomorphic to) a subalgebra of \( C^*_\varepsilon(\mathcal{A}) \times_\alpha \mathbb{Z}_+ \). Moreover,

\[
C^*_\varepsilon(\mathcal{A} \times_\alpha \mathbb{Z}_+) = C^*_\varepsilon(C^*_\varepsilon(\mathcal{A}) \times_\alpha \mathbb{Z}_+).
\]

More generally, they consider an arbitrary C*-cover \( \mathfrak{A} \) of \( \mathcal{A} \). Let \( \mathcal{J}_\mathcal{A} \) denote the Shilov boundary, i.e. the kernel of the unique homomorphism of \( \mathfrak{A} \) onto \( C^*_\varepsilon(\mathcal{A}) \) which is the identity on \( \mathcal{A} \). Suppose that \( \alpha \in \text{End}_\mathcal{A}(\mathfrak{A}) \) also leaves \( \mathcal{J}_\mathcal{A} \) invariant. Then it is easy to see that this induces an endomorphism \( \hat{\alpha} \in \text{End}_\mathcal{A}(C^*_\varepsilon(\mathcal{A})) \). Hence \( \mathcal{A} \times_{\hat{\alpha}} \mathbb{Z}_+ \) is (canonically completely isometrically isomorphic to) a subalgebra of \( C^*_\varepsilon(\mathcal{A}) \times_{\hat{\alpha}} \mathbb{Z}_+ \). They show that the same norm is also induced on \( \mathcal{A} \times_\alpha \mathbb{Z}_+ \) as a subalgebra of \( \mathfrak{A} \times_\alpha \mathbb{Z}_+ \). This result should be considered in conjunction with Proposition 10.6.
Theorem 11.6 (Kakariadis-Katsoulis). If $A$ is a unital operator algebra with C*-cover $\mathfrak{A}$ and $\alpha \in \text{End}_A(\mathfrak{A})$ fixes $J_A$, then $A \times_{\alpha} Z_+$ is also (canonically completely isometrically isomorphic to) a subalgebra of $\mathfrak{A} \times_{\alpha} Z_+$.

In conjunction with Proposition 10.6, we obtain:

Corollary 11.7. If $\alpha \in \text{End}_A(C^*_e(A))$, $\mathfrak{A}$ is a C*-cover of $A$ and $\beta \in \text{End}_A(A)$ such that $\beta|_A = \alpha|_A$, then $A \times_{\beta} Z_+$ is completely isometrically isomorphic to $A \times_{\alpha} Z_+$.

12 Dilating Covariance Relations

We consider the following problem: suppose that $\alpha \in \text{End}(A)$ lifts to a ∗-endomorphism of $C^*_e(A)$. When is $A \times_{\alpha} Z_+$ canonically completely isometrically isomorphic to the subalgebra of $C^*_e(A) \times_{\alpha} Z_+$ generated by $A$ and the element $t$ inducing the action $\alpha$. To simplify statements, we will just say, in this case, that $A \times_{\alpha} Z_+$ is a subalgebra of $\mathfrak{A} \times_{\alpha} Z_+$. We are seeking general properties of $A$ which make this true.

Commutant lifting properties can be seen as special cases of dilation theorems for semicrossed products in the case of the identity automorphism. The goal of this section is to provide several theorems which allow one to obtain results about general semicrossed products from the various commutant lifting properties.

The literature contains a number of dilation theorems for endomorphisms of operator algebras. Ling and Muhly [35] establish an automorphic version of Ando’s theorem, which is a lifting theorem for an action of $\mathbb{Z}_2$. Peters [48] and Muhly and Solel [38, 39] dilate actions of $\mathbb{Z}_+$ on C*-algebras. Our first result in this section uses SCLT and models our dilation theorem for the non-commutative disk algebras [17]. We wish to contrast the explicit dilation obtained here with the more inferential one obtained in Theorem 12.3 which requires only CLT.

Theorem 12.1. Suppose that $A$ is a unital operator algebra satisfying SCLT and ICLT. Let $\alpha \in \text{Aut}(A)$. Then every covariant representation $(\rho,T)$ of $(A,\alpha)$ has a coextension $(\sigma,V)$ such that $\sigma$ is a fully extremal coextension of $\rho$ and $V$ is an isometry.

Proof. Let $\sigma$ be a fully extremal coextension of $\rho$. Then by Proposition 10.7, $\sigma \circ \alpha$ is also fully extremal. By Corollary 3.9,

$$\sigma \oplus (\sigma \circ \alpha)$$

is a fully extremal coextension of $\rho \oplus (\rho \circ \alpha)$. Now the covariance relation
implies that for $a \in \mathcal{A}$,
\[
\begin{bmatrix}
\rho(a) & 0 \\
0 & \rho(\alpha(a))
\end{bmatrix}
\begin{bmatrix}
0 & T \\
0 & 0
\end{bmatrix}
= 
\begin{bmatrix}
0 & T \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
\rho(a) & 0 \\
0 & \rho(\alpha(a))
\end{bmatrix}
\]

Using SCLT, we obtain a contractive coextension of $[\begin{bmatrix}
0 & T \\
0 & 0
\end{bmatrix}$ which commutes with $\sigma \oplus (\sigma \circ \alpha)(\mathcal{A})$. The 1, 2 entry is a contractive coextension $T_1$ of $T$ such that
\[
\sigma(a)T_1 = T_1\sigma(\alpha(a)) \quad \text{for all } a \in \mathcal{A}.
\]

So
\[
\begin{bmatrix}
0 & T_1 \\
0 & 0
\end{bmatrix}
\]
also commutes with $\sigma \oplus (\sigma \circ \alpha)(\mathcal{A})$.

Now use ICLT to coextend this operator to an isometry $V$ commuting with a Shilov coextension $\tau$ of $\sigma \oplus (\sigma \circ \alpha)$. As $\sigma \oplus (\sigma \circ \alpha)$ is an extremal coextension, $\tau$ decomposes as
\[
\tau = \sigma \oplus (\sigma \circ \alpha) \oplus \tau_1.
\]

With respect to this decomposition $\mathcal{K} \oplus \mathcal{K} \oplus \mathcal{P}$, we can write
\[
V = 
\begin{bmatrix}
0 & T_1 \\
0 & 0 \\
* & D & *
\end{bmatrix}
.
\]

In particular, the commutation relation shows that
\[
\tau_1(a)D = D\sigma(\alpha(a)) \quad \text{for all } a \in \mathcal{A}.
\]

A direct summand of a Shilov extension is Shilov, so $\tau_1$ is Shilov. Let $\pi_1$ be a maximal representation on $\mathcal{L} \supset \mathcal{P}$ so that $\mathcal{P}$ is invariant, and $\pi_1|\mathcal{P} = \tau_1$. Then consider $P_\mathcal{P}D$ as an operator in $\mathcal{B}(\mathcal{K}, \mathcal{L})$, and note that
\[
\pi_1(a)P_\mathcal{P}D = P_\mathcal{P}D\sigma(\alpha(a)) \quad \text{for all } a \in \mathcal{A}.
\]

Define a coextension of $\rho$ by
\[
\pi = \sigma \oplus \sum_{n \geq 0} \pi_1 \circ \alpha^n
\]
acting on $\mathcal{M} = \mathcal{K} \oplus \mathcal{L}^{(\infty)}$. Since $\sigma$ is a fully extremal coextension of $\rho$ and each $\pi_1 \circ \alpha^n$ is maximal, it follows that $\pi$ is a fully extremal coextension of $\rho$.

Now define an isometry $W$ on $\mathcal{M}$ by
\[
W = 
\begin{bmatrix}
T_1 & 0 & 0 & 0 & \ldots \\
P_\mathcal{P}D & 0 & 0 & 0 & \ldots \\
0 & I & 0 & 0 & \ldots \\
0 & 0 & I & 0 & \ldots \\
\vdots & \ddots & \ddots & \ddots
\end{bmatrix}
\]
Then one readily verifies that
\[ \pi(a)W = W\pi(\alpha(a)) \quad \text{for all } a \in A. \]
This is the desired isometric coextension of the covariance relations.

Observe that Theorem 12.1 shows that if \( A \) has SCLT and ICLT, then
\[ A \times_\alpha Z_+ = A \times_{\alpha_0} Z_+. \]
Thus by applying Theorem 11.5 to see that it imbeds into \( C_\alpha^*(A) \times_\alpha Z_+ \). Since \( \alpha \) is an automorphism, the C*-envelope of this is just the full crossed product \( C_\alpha^*(A) \times_\alpha Z \).

**Theorem 12.2.** Suppose that \( A \) has properties SCLT and ICLT, and \( \alpha \in \text{Aut}(A) \). Then \( A \times_\alpha Z_+ \) is (canonically completely isometrically isomorphic to) a subalgebra of \( C_\alpha^*(A) \times_\alpha Z_+ \). Moreover,
\[ C_\alpha^*(A \times_\alpha Z_+) = C_\alpha^*(A) \times_\alpha Z. \]

We wish to improve this theorem so that we require only CLT, not SCLT. This requires a different approach, and does not yield a direct construction of the isometric coextension of a covariant representation.

**Theorem 12.3.** Suppose that an operator algebra \( A \) has the Ando property, and \( \alpha \in \text{Aut}(A) \). Then \( A \times_\alpha Z_+ \) is (canonically completely isometrically isomorphic to) a subalgebra of \( C_\alpha^*(A) \times_\alpha Z_+ \). Moreover,
\[ C_\alpha^*(A \times_\alpha Z_+) = C_\alpha^*(A) \times_\alpha Z. \]

**Proof.** Suppose that \((\rho, T)\) is a covariant representation of \((A, \alpha)\). Let \( \sigma_0 \) be a fully extremal coextension of \( \rho \) on a Hilbert space \( K_0 \). Then by Proposition 10.7, \( \sigma_0 \circ \alpha \) is also fully extremal. Theorem 7.15 yields a fully extremal coextension \( \sigma_1 \) of \( \rho_1 \) on \( K_1 \), an isometry \( V_0 \) and a maximal dilation \( \pi_0 \) so that
\[ \sigma_1(a)V_0 = V_0((\sigma_0 \circ \alpha) \oplus \pi_0)(a). \]
Recursively, we obtain fully extremal coextension \( \sigma_{n+1} \) of \( \rho_1 \) on \( K_n \), an isometry \( V_n \) and a maximal dilation \( \pi_n \) so that
\[ \sigma_{n+1}(a)V_n = V_n((\sigma_n \circ \alpha) \oplus \pi_n)(a). \]
Let
\[ \pi = \sum_{n \geq 0} \sum_{k \in \mathbb{Z}} \oplus(\pi_n \circ \alpha^k)(\infty). \]
Then set $\tilde{\sigma}_n = \sigma_n \oplus \pi$ acting on $\tilde{K}_n$. Extending $V_n$ appropriately to an isometry $\tilde{V}_n$, we obtain

$$\tilde{\sigma}_{n+1}(a)\tilde{V}_n = \tilde{V}_n(\tilde{\sigma}_n \circ \alpha)(a).$$

Now define a representation on

$$\tilde{\mathcal{K}} = \sum_{n \geq 0} \oplus \tilde{K}_n$$

by

$$\tilde{\sigma}(a) = \begin{bmatrix}
\tilde{\sigma}_0(a) & 0 & 0 & 0 & \ldots \\
0 & \tilde{\sigma}_1(a) & 0 & 0 & \ldots \\
0 & 0 & \tilde{\sigma}_2(a) & 0 & \ldots \\
0 & 0 & 0 & \tilde{\sigma}_3(a) & \ddots \\
\vdots & \vdots & \ddots & \ddots & \ddots
\end{bmatrix}$$

and

$$\tilde{V} = \begin{bmatrix}
\tilde{V}_0 & 0 & 0 & 0 & \ldots \\
0 & \tilde{V}_1 & 0 & 0 & \ldots \\
0 & 0 & \tilde{V}_2 & 0 & \ddots \\
\vdots & \vdots & \ddots & \ddots & \ddots
\end{bmatrix}$$

This is an isometric covariant representation which coextends the contractive covariant representation on $\mathcal{H} \otimes \ell^2$:

$$\rho(\infty)(a) = \begin{bmatrix}
\rho(a) & 0 & 0 & 0 & \ldots \\
0 & \rho(a) & 0 & 0 & \ldots \\
0 & 0 & \rho(a) & 0 & \ldots \\
0 & 0 & 0 & \rho(a) & \ddots \\
\vdots & \vdots & \ddots & \ddots & \ddots
\end{bmatrix}$$

and

$$T \otimes S = \begin{bmatrix}
0 & 0 & 0 & 0 & \ldots \\
T & 0 & 0 & 0 & \ldots \\
0 & T & 0 & 0 & \ldots \\
0 & 0 & T & 0 & \ddots \\
\vdots & \vdots & \ddots & \ddots & \ddots
\end{bmatrix}$$

This latter representation induces the same seminorm on $\mathcal{A} \times_{\alpha} \mathbb{Z}_+$ as the covariant pair $(\rho(\infty), T \otimes U)$ on $\mathcal{H} \otimes \ell^2(\mathbb{Z})$, where $U$ is the bilateral shift, because this representation is an inductive limit of copies of $(\rho(\infty), T \otimes S)$. However $(\rho(\infty), T \otimes U)$ has a seminorm which clearly dominates the seminorm arising from $(\rho, T)$. 

Documenta Mathematica 16 (2011) 781–868
K.R. Davidson and E.G. Katsoulis

It follows that
\[ \mathcal{A} \times_\alpha \mathbb{Z}^+ = \mathcal{A} \times_\alpha^i \mathbb{Z}^+. \]

The rest follows as in Theorem 12.2.

In particular, one gets a dilation theorem that we cannot find by a direct construction. Indeed, even the isometric dilation is not explicitly obtained, unlike the proof of Theorem 12.1. This result considerably expands the class of operator algebras for which we can obtain these results.

**Corollary 12.4.** Suppose that an operator algebra \( \mathcal{A} \) has the Ando property, and \( \alpha \in \text{Aut}(\mathcal{A}) \). Then every covariant representation \((\rho, T)\) of \((\mathcal{A}, \alpha)\) dilates to a covariant representation \((\pi, U)\) of \((C^*_\beta(\mathcal{A}), \alpha)\) where \( \pi \) is a \(*\)-representation of \(C^*_\beta(\mathcal{A})\) and \( U \) is unitary.

The following result applies to endomorphisms, not just automorphisms. This result was only recently established for the disk algebra [18]. The hypotheses are quite strong.

**Theorem 12.5.** Suppose that \( \mathcal{A} \) is a unital operator algebra with SMCLT and FP. Let \( \alpha \in \text{End}_\mathbb{A}(C^*_\beta(\mathcal{A})) \). Then \( \mathcal{A} \times_\alpha \mathbb{Z}^+ \) is (canonically completely isometrically isomorphic to) a subalgebra of \( C^*_\beta(\mathcal{A}) \times_\alpha \mathbb{Z}^+ \).

**Proof.** To establish that \( \mathcal{A} \times_\alpha \mathbb{Z}^+ \) is completely isometric to a subalgebra of \( C^*_\beta(\mathcal{A}) \times_\alpha \mathbb{Z}^+ \), it suffice to show that if \((\rho, T)\) is a covariant representation of \((\mathcal{A}, \alpha)\), then \( \rho \) has a \(*\)-dilation of \( C^*_\beta(\mathcal{A}) \) on a Hilbert space \( K \supseteq H \) and a contraction \( S \) dilating \( T \) to \( K \) such that \((\pi, S)\) is a covariant representation of \((C^*_\beta(\mathcal{A}), \alpha)\) with \( H \) as a coinvariant subspace. For this then shows that

\[ \| \rho \times T(p) \| \leq \| \pi \times S(p) \| \leq \| p \|_{C^*_\beta(\mathcal{A}) \times_\alpha \mathbb{Z}^+}. \]

The reverse inequality is evident. Hence the norm on \( \mathcal{A} \times_\alpha \mathbb{Z}^+ \) will coincide with the norm as a subalgebra of \( C^*_\beta(\mathcal{A}) \times_\alpha \mathbb{Z}^+ \).

First dilate \( \rho \) to a maximal dilation \( \pi \). This extends to a \(*\)-representation of \( C^*_\beta(\mathcal{A}) \) which we also denote by \( \pi \). We may write:

\[ \pi(a) = \begin{bmatrix} * & 0 & 0 \\ 0 & \rho(a) & 0 \\ * & * & * \end{bmatrix} \quad \text{for } a \in \mathcal{A}. \]

Observe that the covariance condition is equivalent to

\[ \begin{bmatrix} \rho(a) & 0 \\ 0 & \rho(\alpha(a)) \end{bmatrix} \begin{bmatrix} 0 & T \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & T \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \rho(a) & 0 \\ 0 & \rho(\alpha(a)) \end{bmatrix}. \]

Now \( \rho \circ (\rho \circ \alpha) \) dilates to a \(*\)-dilation

\[ \pi \oplus (\pi \circ \alpha), \]

**Documenta Mathematica** 16 (2011) 781–868
So by SMCLT, there is a contraction dilating

\[
\begin{bmatrix}
0 & T \\
0 & 0
\end{bmatrix}
\]

which commutes with \((\pi \oplus (\pi \circ \alpha))(A)\) of the form

\[
\tilde{S} = \begin{bmatrix}
* & S \\
S^* & *
\end{bmatrix}
\]

so that \(\mathcal{H} \oplus \mathcal{H}\) is simultaneously semi-invariant for \((\pi \oplus \pi \circ \alpha)(A)\) and \(\tilde{S}\). There is no loss in assuming that the * entries are all 0. Commutation again means that \((\pi, S)\) is a covariant representation of \(A\).

Now by the Fuglede property, the adjoint

\[
\begin{bmatrix}
0 & S^* \\
S & 0
\end{bmatrix}
\]

also commutes with \((\pi \oplus (\pi \circ \alpha))(A)\). That means that

\[
\pi(\alpha(a))S^* = S^*\pi(a) \quad \text{for all } a \in A.
\]

Equivalently, since \(\pi\) is a *-representation,

\[
\pi(a^*)S = S\pi(a^*) \quad \text{for all } a \in A.
\]

But the set of operators in \(C^*_e(A)\) satisfying the covariance relations is a closed algebra, and contains \(A\) and \(A^*\), whence it is all of \(C^*_e(A)\). Thus we have obtained the desired dilation to covariance relations for \((C^*_e(A), \alpha)\). Hence \(A \times_{\alpha} Z_+\) is (canonically completely isometrically isomorphic to) a subalgebra of \(C^*_e(A) \times_{\alpha} Z_+\).

The following is immediate from the dilation theory for \(C^*_e(A) \times_{\alpha} Z_+\) and Theorem 11.5. Combining this with Theorem 11.3, one obtains an explicit description of this C*-envelope.

**Corollary 12.6.** Suppose that a unital operator algebra \(A\) has FP and SMCLT, and \(\alpha \in \text{End}_A(C^*_e(A))\). Then every covariant representation \((\rho, T)\) of \((A, \alpha)\) dilates to a covariant representation \((\pi, U)\) of \((C^*_e(A), \alpha)\) where \(\pi\) is a *-representation of \(C^*_e(A)\) and \(U\) is unitary. Moreover,

\[
C^*_e(A \times_{\alpha} Z_+) = C^*_e(C^*_e(A) \times_{\alpha} Z_+).
\]
13 Further Examples

The disk algebra. The first application yields a recent result about semi-crossed products by completely isometric endomorphisms for the disk algebra [18]. We note that endomorphisms which are not completely isometric are also treated there, but our results do not apply in that case.

The C*-envelope of the disk algebra $A(D)$ is $C(T)$, which is generated by the unitary element $z$. The classical Fuglede Theorem shows that $A(D)$ has FP. Also the classical Sz.Nagy–Foiaş Commutant Lifting Theorem yields the properties SCLT and SMCLT. Ando’s property is Corollary 7.11, which was a strengthening of Ando’s theorem. As $A(D)$ is Dirichlet, we have uniqueness of extremal coextensions and of extremal extensions, which are also consequences of the original Sz.Nagy theory. As $A(D) ≃ A(D)^*$, we have SCLT* as well. Indeed, $A(D)$ has all of the properties studied in this paper.

Suppose that $α ∈ \text{End}_{A(D)}(C(T))$. Then $b = α(z) ∈ A(D)$; and has spectrum

$$\sigma_{A(D)}(b) = \sigma_{A(D)}(z) = D \quad \text{and} \quad \sigma_{C(T)}(b) = \sigma_{C(T)}(z) = T.$$ 

Thus Ran$(b) = D$ and Ran$(b|_T) = T$. It follows that $b$ is a non-constant finite Blaschke product. We have $α(f) = f \circ b$ for all $f ∈ C(T)$.

**Theorem 13.1.** Let $b$ be a non-constant finite Blaschke product, and let $α(f) = f \circ b$ in $\text{End}_{A(D)}(C(T))$. Then $A(D) ×_α Z_+$ is (canonically completely isometrically isomorphic to) a subalgebra of $C(T) ×_α Z_+$; and

$$C^*_ε(A(D) ×_α Z_+) = C^*_ε(C(T) ×_α Z_+).$$

This is explicitly described as $C(S_α) ×_β Z$ where $S_α$ is the solenoid

$$S_α = \{(z_n)_{n ≥ 1} : z_n = b(z_{n+1}), z_n ∈ T, n ≥ 1 \}$$

and $β$ is the backward shift on $S_α$.

**Proof.** The first statement follows from Theorems 12.5. The detailed description of the C*-envelope comes from the Kakariadis-Katsoulis Theorem 11.3. □

It is worth restating this theorem as a dilation result.

**Corollary 13.2.** Let $b$ be a non-constant finite Blaschke product and suppose that $S$ and $T$ are contractions satisfying $ST = Tb(S)$. Then there exist unitary operators $U$ and $V$, dilating $S$ and $T$ respectively, so that $UV = V b(U)$.

The non-commutative disk algebras. For $n ≥ 2$ finite, the non-commutative disk algebra $A_n$ has the Cuntz algebra $O_n$ as its C*-envelope. The Frazho-Bunce-Popescu dilation theorem [30, 13, 50] shows that the minimal external coextension of a representation is unique. This also follows because
\(A_n\) is semi-Dirichlet. Popescu [52] proves the SCLT property in a similar manner to the original proof of Sz.Nagy and Foiaş. The FP property follows from Proposition 9.3.

There are many distinct ways to extend the left regular representation to a maximal representation (see [23, §3]). In particular, the minimal fully extremal extensions are not unique. Nevertheless, \(A_n\) has ICLT and MCLT. This follows from our paper [17] specialized to the identity automorphism.

The completely isometric automorphisms of \(A_n\) are the analogues of the conformal automorphisms of the ball \(B_n\) of \(\mathbb{C}^n\). These were first described by Voiculescu [61] as \(*\)-automorphisms of \(O_n\) which fix the analytic part. These are the only such automorphisms of \(A_n\) [24]. See also [53]. Thus we recover our results on semicrossed products of \(A_n\) in [17] as a consequence of Theorem 12.2.

**Theorem 13.3 (Davidson-Katsoulis).** If \(\alpha \in \text{Aut}(A_n) = \text{Aut}_{A_n}(O_n)\), then

\[
C^*_\alpha(A_n \times_{\alpha} \mathbb{Z}_+^1) = O_n \times_{\alpha} \mathbb{Z}.
\]

It is also easy to determine \(\text{End}(O_n)\). Every \(n\)-tuple of isometries \(t_i \in O_n\) such that \(\sum_{i=1}^n t_i t_i^* = 1\) determines an endomorphism with \(\alpha(s_i) = t_i\) by the universal property of the Cuntz algebra. For the endomorphism \(\alpha\) to leave \(A_n\) invariant, it is then necessary and sufficient that \(t_i\) belong to \(A_n\). Given that \(\text{End}(O_n)\) is so rich, the following result seems surprising.

**Theorem 13.4.** For \(n \geq 2\) finite,

\[
\text{End}_{A_n}(O_n) = \text{Aut}_{A_n}(O_n).
\]

**Proof.** We represent \(A_n\) on \(\ell^2(\mathbb{F}_+^n)\) by the left regular representation \(\lambda\) with generators \(L_i = \lambda(s_i)\), where

\[
L_i \xi_w = \xi_{iw}.
\]

Note that

\[
C^*(\lambda(A_n)) = \mathcal{E}_n
\]

is the Cuntz-Toeplitz algebra, and that

\[
g: \mathcal{E}_n \rightarrow \mathcal{E}_n / K = O_n
\]

is the quotient by the compact operators. Let \(\mathcal{R}_n\) denote the wot-closed right regular representation algebra generated by \(R_i\), \(1 \leq i \leq n\), where

\[
R_i \xi_w = \xi_{wi}.
\]

Then \(\lambda(A_n)' = \mathcal{R}_n\) [3, 23]. We use the notation \(R_v \xi_v = \xi_{wv}\) for words \(v \in \mathbb{F}_+^n\).

Suppose that \(\alpha \in \text{End}_{A_n}(O_n)\). Then \(t_i = \alpha(s_i)\) are isometries in \(A_n\) such that

\[
\sum_{i=1}^n t_i t_i^* = 1.
\]
We want to clarify when this is possible. Let
\[ T_i = \lambda(t_i) \quad \text{and} \quad T = [T_1 \ldots T_n]. \]
Then \( q(T_i) = t_i \), and thus \( T \) is an essential isometry, as are each \( T_i \). However since \( q \) is a complete isometry on \( \mathfrak{A}_n \), we have \( \|T\| = 1 \).

We claim that each \( T_i \) is an isometry. Indeed, if \( \zeta \in \ell^2(F^n) \) with \( \|\zeta\| = 1 \) and \( \|T_i\zeta\| \neq 1 \), then
\[ \|T_i(R_n\zeta)\| = \|R_nT_i\zeta\| = \|T_i\zeta\|. \]
Since \( R_n^i\zeta \) tends to 0 weakly, we see that \( T_i \) is not an essential isometry, contrary to fact. Since \( \|T\| = 1 \), the \( T_1, \ldots, T_n \) are isometries in \( \lambda(\mathfrak{A}_n) \) with pairwise orthogonal range. So \( T \) is a row isometry. Since
\[ q(\sum_{i=1}^n T_iT_i^*) = 1, \]
we deduce that
\[ P = I - \sum_{i=1}^n T_iT_i^* \]
is a finite rank projection.

The range of each \( T_i \) is a cyclic invariant subspace for \( R_n \), with cyclic vector \( \zeta_i = T_i\zeta_\emptyset \). Let \( N = \text{Ran} \, P \). Then \( N^\perp \) is the sum of the ranges of the \( T_i \), and so it is invariant for \( \mathfrak{A}_n \) with wandering space \( W = \text{span}\{\zeta_i : 1 \leq i \leq n\} \).

Thus \( N \) is coinvariant. Let
\[ A_i = P_NR_i|_N \quad \text{and} \quad A = [A_1 \ldots A_n]. \]
Then \( A \) is a row contraction with a row isometric dilation
\[ R = [R_1 \ldots R_n]. \]
The minimal row isometric dilation is unique [50], and any other is the direct sum of the minimal dilation with another row isometry. Since \( R \) is irreducible, this is the minimal dilation of \( A \).

By [20], the wandering space \( W \) of \( N^\perp \) is given by
\[ W = (N + \sum_{i=1}^n R_iN) \cap N. \]
Note that
\[ I_N - \sum_{i=1}^n A_iA_i^* = P_N(I - \sum_{i=1}^n R_iR_i^*)|_N = (P_N\xi_\emptyset)(P_N\xi_\emptyset)^*. \]
This is non-zero because if $N$ were orthogonal to $\xi$, then $\xi$ would also be orthogonal to the invariant subspace it generates, which is the whole space. Thus $N$ is not contained in $\sum_{i=1}^{n} R_i N$ because this space is orthogonal to $\xi$. So now we compute

$$n = \dim W = \dim(N + \sum_{i=1}^{n} R_i N) - \dim N \geq (1 + n \dim N) - \dim N = 1 + (n - 1) \dim N.$$ 

Therefore $\dim N \leq 1$; whence $\dim N = 1$ because no $n$-tuple of isometries $T_i$ in $L_n$ is of Cuntz type.

The only coinvariant subspaces of dimension one are $C_{\nu_{\lambda}}$, where $\nu_{\lambda}$ is an eigenvector of $R^n$ [3, 23]. These are indexed by points $\lambda$ in the open unit ball $B_n$ of $C^n$. It now follows from the analysis in [24] that $\alpha$ is an automorphism. Briefly, one can compose $\alpha$ with an automorphism $\theta_{\lambda}$ so that $\lambda = 0$ and so $N = C_{\xi}$. Then

$$W = \text{span}\{\xi_i : 1 \leq i \leq n\}.$$ 

So $\{\zeta_i\}$ form an orthonormal basis for $W$. The unitary $U \in U_n$ which takes $\xi_i$ to $\zeta_i$ induces a gauge unitary $\tilde{U}$ which takes each $L_i$ to $T_i$, as this is the unique element of $A_n$ with $T_i \xi = \zeta_i$. Hence $\theta_{\lambda} \alpha = \text{ad} \tilde{U}$ is an automorphism; whence so is $\alpha$.

**Finite dimensional nest algebras.** A finite dimensional nest algebra can be described as the block upper triangular matrices with respect to a decomposition $H = H_1 \oplus \cdots \oplus H_k$ of a finite dimensional Hilbert space into a direct sum of subspaces. These are the incidence algebras which are Dirichlet. They have SCLT, SCLT*, ICLT, ICLT*, MCLT and the Ando property. By Example 9.9, finite dimensional nest algebras have FP.

The only isometric endomorphisms are isometric automorphisms. These are unitarily implemented, and the unitary preserves the nest. (Ringrose [54] characterizes the isomorphisms between nest algebras in infinite dimensions, and includes the more elementary finite dimensional case. See [14].) Hence the unitary has the form $U = U_1 \oplus \cdots \oplus U_k$ with respect to the decomposition of $H$. Clearly $\text{ad} U$ extends to a $*$-automorphism of the $C^*$-envelope $B(H) \simeq M_n$, where $n = \dim H$.

**Graph algebras and tensor algebras of $C^*$-correspondences.** The tensor algebra $T^+ (E)$ of a $C^*$-correspondence $E$ over a $C^*$-algebra $A$ is semi-Dirichlet. Thus every Shilov coextension of a representation $\rho$, and in particular every extremal coextension of $\rho$, is fully extremal; and the minimal extremal coextension of $\rho$ unique. So in particular, the minimal fully extremal
coextension is a cyclic coextension. Muhly and Solel [42] show that the tensor algebra of a C*-correspondence has SCLT. The C*-envelope is the Cuntz-Pimsner algebra \( \mathcal{O}(E) \) [42, 29, 34]. Kakariadis and Katsoulis [32] establish that for every \( \alpha \in \text{Aut}_{\mathcal{T}}(\mathcal{O}(E)) \) such that \( \alpha|_{\mathcal{A}} = \text{id} \), the semi-crossed product \( \mathcal{T}^{+}(E) \otimes_{\alpha} \mathbb{Z}_+ \) imbeds canonically, completely isometrically as a subalgebra of \( \mathcal{O}(E) \otimes_{\alpha} \mathbb{Z}_+ \); and this is its C*-envelope. In particular, taking \( \alpha = \text{id} \), one obtains the Ando property, so it has ICLT and SCLT.

Thus, by circular reasoning, Theorems 12.1 and 12.2 apply. The point however is that the dilation theorems for automorphisms follow immediately once one has the appropriate commutant lifting theorems, which basically deal with the identity automorphism. In principle, and often in practice, this is much easier.

An important special case of a tensor algebra is the tensor algebra \( \mathcal{T}^{+}(G) \) of a directed graph \( G \). Some of the properties are somewhat easier to see here. In addition, by Example 9.7, finite graph algebras have FP.

**Bilateral Tree Algebras.** In the case of a bilateral tree algebra \( \mathcal{A} \), one readily sees that \( C^*_e(\mathcal{A}) \) is a direct sum of full matrix algebras corresponding to the connected components of the graph. The automorphisms of finite dimensional C*-algebras are well understood. Modulo inner automorphisms, one can only permute subalgebras of the same size. Automorphisms of the tree algebra are more restrictive, and modulo those inner automorphisms from unitaries in \( \mathcal{A} \cap \mathcal{A}^* \), they come from automorphisms of the associated directed graph.

Bilateral tree incidence algebras have the Ando property by Theorem 8.6. Hence by Theorem 12.3, we obtain:

**Corollary 13.5.** Let \( \mathcal{A} \) be a bilateral tree algebra, and let \( \alpha \in \text{Aut}(\mathcal{A}) \). Then

\[
C^*_e(\mathcal{A} \times_{\alpha} \mathbb{Z}_+) = C^*_e(C^*_e(\mathcal{A}) \times_{\alpha} \mathbb{Z}_+).
\]

Hence if \( (\rho, T) \) is a covariant representation of \( (\mathcal{A}, \alpha) \), there is a maximal dilation \( \pi \) of \( \rho \) and a unitary dilation \( U \) of \( T \) so that \( (\pi, U) \) is a covariant representation of \( (C^*_e(\mathcal{A}), \alpha) \).

It is an interesting question to look at the infinite dimensional WOT-closed versions. A commutative subspace lattice (CSL) is a strongly closed lattice of commuting projections. A CSL algebra is a reflexive algebra whose invariant subspace lattice is a CSL. Since every CSL is contained in a masa, one can instead define a CSL algebra to be a reflexive algebra containing a masa. The seminal paper, which provides a detailed structure theory for these algebras, is due to Arveson [7]. See also [14].

When dealing with weak-*-closed operator algebras, the class of all representations is generally too large. Instead one restrict attention to weak-*-continuous (completely contractive) representations. To apply the results from this paper, a weak-* version needs to be developed.
A CSL algebra is a bilateral tree algebra if the lattice satisfies an
measure
theoretic version of the discrete bilateral tree property. We will not define this
precisely here, but refer the reader to [22] for the full story. The approximation
results from [22] show that every bilateral tree algebra can be approximated in
two very strong ways by a sequence of finite dimensional subalgebras which are
completely isometrically isomorphic to bilateral tree incidence algebras. These
results should be a crucial step towards deducing similar dilation results for
semicrossed products of these infinite dimensional bilateral tree algebras by
weak-* continuous endomorphisms.

References

[1] J. Agler, An abstract approach to model theory, in Surveys of some recent
results in operator theory, Vol. II, pp. 1–23, Longman Sci. Tech., Harlow,
1988.
[2] T. Ando, On a pair of commuting contractions, Acta Sci. Math. (Szeged)
24 (1963), 88–90.
[3] A. Arias and G. Popescu, Factorization and reflexivity on Fock spaces,
Integral Equations and Operator Theory 23 (1995), 268–286.
[4] W. Arveson, Operator algebras and measure preserving automorphisms,
Acta Math. 118, (1967), 95–109.
[5] W. Arveson, Subalgebras of C*-algebras, Acta Math. 123 (1969), 141–224.
[6] W. Arveson, Subalgebras of C*-algebras II, Acta Math. 128 (1972), 271–
308.
[7] W. Arveson, Operator algebras and invariant subspaces, Ann. Math. 100
(1974), 433–532.
[8] W. Arveson, Subalgebras of C*-algebras III, Acta Math. 181 (1998), 159–
228.
[9] W. Arveson, The curvature invariant of a Hilbert module over
\( \mathbb{C}[z_1, \ldots, z_d] \), J. Reine Angew. Math. 522 (2000), 173–236.
[10] W. Arveson, The noncommutative Choquet boundary, J. Amer. Math. Soc.
21 (2008), 1065–1084.
[11] W. Arveson and K. Josephson, Operator algebras and measure preserving
automorphisms II, J. Funct. Anal. 4 (1969), 100–134.
[12] D. Blecher, Z.-J. Ruan and A. Sinclair, A characterization of operator
algebras, J. Funct. Anal. 89 (1990), 188–201.
[13] J. Bunce, Models for \( n \)-tuples of non-commuting operators, J. Funct. Anal.
57 (1984), 21–30.
[14] K.R. Davidson, Nest Algebras, Pitman Research Notes in Mathematics
Series 191, Longman Scientific and Technical Pub. Co., London, New
York, 1988.

Documenta Mathematica 16 (2011) 781–868
K.R. Davidson and E.G. Katsoulis

[15] K.R. Davidson, *When locally contractive representations are completely contractive*, J. Funct. Anal. 128 (1995), 186–225.

[16] K.R. Davidson and E. Katsoulis, *Operator algebras for multivariable dynamics*, Memoirs Amer. Math. Soc. 209 (2011), no. 982, 53+viii pages.

[17] K.R. Davidson and E. Katsoulis, *Dilating covariant representations of the non-commutative disc algebras*, J. Funct. Anal. 259 (2010), Pages 817–831.

[18] K.R. Davidson and E. Katsoulis, *Semicrossed products of the disk algebra*, Proc. Amer. Math. Soc., to appear (arXiv:1104.1398).

[19] K.R. Davidson, E. Katsoulis and D.R. Pitts, *The structure of free semigroup algebras*, J. reine angew. Math. 533 (2001), 99–125.

[20] K.R. Davidson, D.W. Kribs and M.E. Shpigel, *Isometric Dilations of non-commuting finite rank n-tuples*, Canad. J. Math. 53 (2001), 506–545.

[21] K.R. Davidson and Trieu Le, *Commutant lifting for commuting row contractions*, Bull. London Math. Soc. 42 (2010), 506–516.

[22] K.R. Davidson, V.I. Paulsen, and S.C. Power, *Tree algebras, semidiscreteness, and dilation theory*, Proc. London Math. Soc. (3) 68 (1994), 178–202.

[23] K.R. Davidson and D.R. Pitts, *Invariant subspaces and hyper-reflexivity for free semi-group algebras*, Proc. London Math. Soc. 78 (1999), 401–430.

[24] K.R. Davidson and D.R. Pitts, *The algebraic structure of non-commutative analytic Toeplitz algebras*, Math Ann. 311 (1998), 275–303.

[25] K.R. Davidson and J. Roydor, *C*-envelopes of tensor algebras for multivariable dynamics*, Proc. Edinburgh Math. Soc. 53 (2010), 333–351.

[26] R.G. Douglas and V.I. Paulsen, *Hilbert modules over function algebras*, Pitman Research Notes in Math. 217, Longman Scientific & Technical, Harlow, 1989.

[27] M. Dritschel and S. McCullough, *Boundary representations for families of representations of operator algebras and spaces*, J. Operator Theory 53 (2005), 159–167.

[28] S. Drury, *A generalization of von Neumann’s inequality to the complex ball*, Proc. Amer. Math. Soc. 68 (1978), 300–304.

[29] N. Fowler, P. Muhly and I. Raeburn, *Representations of Cuntz-Pimsner algebras*, Indiana Univ. Math. J. 52 (2003), 569–605.

[30] A. Frazho, *Models for non-commuting operators*, J. Funct. Anal. 48 (1982), 1–11.

[31] M. Hamana, *Injective envelopes of operator systems*, Publ. Res. Inst. Math. Sci. 15 (1979), 773-785.

[32] E. Kakariadis and E. Katsoulis, *Semicrossed products of operator algebras and their C*-envelopes*, preprint (arXiv:1101.1482).

[33] E. Katsoulis and D. Kribs, *The C*-envelope of the tensor algebra of a directed graph*, Integral Equations Operator Theory 56 (2006), 401–414.
Dilation Theory...

[34] E. Katsoulis and D. Kribs, Tensor algebras of C*-correspondences and their C*-envelopes, J. Funct. Anal. 234 (2006), 226–233.

[35] K. Linag and P. Muhly, An automorphic form of Ando’s theorem, Integral Equations Operator Theory 12 (1989), 424–434.

[36] M. McAsey and P. Muhly, Representations of nonself-adjoint crossed products, Proc. London Math. Soc. 47(1983), 128–144.

[37] M. McAsey, P. Muhly and K. Sato, Nonself-adjoint crossed products (invariant subspaces and maximality), Trans. Amer. Math. Soc. 248(1979), 381–409.

[38] P. Muhly and B. Solel, A lifting theorem for analytic crossed products, Houston J. Math. 18 (1992), 153–166.

[39] P. Muhly and B. Solel, Extensions and dilations for C*-dynamical systems, Operator theory, operator algebras, and applications, 375–381, Contemp. Math. 414, Amer. Math. Soc., Providence, RI, 2006.

[40] P. Muhly and B. Solel, Hilbert modules over operator algebras, Mem. Amer. Math. Soc. 117 (1995), no. 559.

[41] P. Muhly and B. Solel, An algebraic characterization of boundary representations, Nonselfadjoint operator algebras, operator theory, and related topics, 189–196, Oper. Theory Adv. Appl. 104, Birkhäuser, Basel, 1998.

[42] P. Muhly and B. Solel, Tensor algebras over C*-correspondences (Representations, dilations and C*-envelopes), J. Funct. Anal. 158 (1998), 389–457.

[43] V. Müller and F. Vasilescu, Standard models for some commuting multioperators, Proc. Amer. Math. Soc. 117 (1993), 979–989.

[44] S. Parrott, Unitary dilations for commuting contractions, Pacific J. Math. 34 (1970), 481–490.

[45] V. Paulsen, Completely bounded maps and operator algebras, Cambridge Studies in Advanced Mathematics 78, Cambridge University Press, Cambridge, (2002), xii+300 pp.

[46] V. Paulsen and S. Power, Lifting theorems for nest algebras, J. Operator Theory 20 (1988), 311–327.

[47] V. Paulsen and S. Power, Tensor products of non-self-adjoint operator algebras, Rocky Mountain J. Math. 20 (1990), 331–350.

[48] J. Peters, Semicrossed products of C*-algebras, J. Funct. Anal. 59 (1984), 498–534.

[49] J. Peters, The C*-envelope of a semicrossed product and nest representations, in Operator Structures and Dynamical Systems, M. de Jeu, S. Silvestrov, C. Skau, and J. Tomiyama, eds., Contemporary Math. 503 (2009), pp. 197–215.

[50] G. Popescu, Isometric dilations for infinite sequences of noncommuting operators, Trans. Amer. Math. Soc. 316 (1989), 523–536.
[51] G. Popescu, von Neumann inequality for \((B(H)^n)_1\), Math. Scand. 68 (1991), 292–304.

[52] G. Popescu, On intertwining dilations for sequences of noncommuting operators, J. Math. Anal. Appl. 167 (1992), 382–402.

[53] G. Popescu, Free holomorphic functions on the unit ball of \(B_n\), J. Funct. Anal. 241 (2006), 268–333.

[54] J. Ringrose, On some algebras of operators II, Proc. London Math. Soc. (3) 16 (1966), 385–402.

[55] D. Sarason, Invariant subspaces and unstarred operator algebras, Pacific J. Math. 17 (1966), 511–517.

[56] B. Solel, Representations of product systems over semigroups and dilations of commuting CP maps, J. Funct. Anal., 235 (2006), 593–618.

[57] Sz. Nagy, B., Sur les contractions de l’espace de Hilbert, Acta Sci. Math. (Szeged) 15 (1953), 87–92.

[58] B. Sz. Nagy and C. Foiaş, Harmonic analysis of operators on Hilbert space, North Holland Pub. Co., London, 1970.

[59] B. Sz. Nagy and C. Foiaş, Dilatation des commutants d’operateurs, Comptes Rendus Paris 266 (1968), A493–495.

[60] N. Varopoulos, On an inequality of von Neumann and an application of the metric theory of tensor products to operators theory, J. Funct. Anal. 16 (1974), 83–100.

[61] D. Voiculescu, Symmetries of some reduced free product \(C^*\)-algebras, Operator algebras and their connections with topology and ergodic theory (Busteni, 1983), Lecture Notes in Math. 1132 (1985), 556–588, Springer Verlag, New York, 1985.
GREEN FUNCTIONS VIA HYPERBOLIC LOCALIZATION

Pramod N. Achar

Received: April 26, 2010
Communicated by Wolfgang Soergel

Abstract. Let $G$ be a reductive algebraic group, with nilpotent cone $\mathcal{N}$ and flag variety $\mathcal{B}$. We construct an exact functor from perverse sheaves on $\mathcal{N}$ to locally constant sheaves on $\mathcal{B}$, and we use it to study Ext-groups and stalks of simple perverse sheaves on $\mathcal{N}$ in terms of the cohomology of $\mathcal{B}$.

2010 Mathematics Subject Classification: Primary 20G05; Secondary 14F05, 32S60.
Keywords and Phrases: nilpotent cone; perverse sheaves; hyperbolic localization.

1. Introduction

Let $G$ be a connected reductive algebraic group over an algebraically closed field $k$ of good characteristic. Let $\mathcal{N}$ denote the nilpotent cone in its Lie algebra $\mathfrak{g}$, and let $W$ denote its Weyl group. An explicit description of the stalks of simple perverse sheaves on $\mathcal{N}$ has been given by Lusztig [L], building on earlier ideas of Shoji [S1, S2]. For most such perverse sheaves (those appearing in the Springer correspondence), this description involves the representation theory of $W$, and specifically its coinvariant algebra. The coinvariant algebra of $W$ is also isomorphic to the cohomology ring $\text{H}^\bullet(\mathcal{B})$ of the flag variety $\mathcal{B}$. However, that cohomology ring does not appear in the proofs in [L], which rely instead on orthogonality properties of character sheaves coming from the geometry of semisimple classes.

The present paper is an attempt to understand Lusztig’s results directly in terms of the geometry of $\mathcal{B}$. Consider the diagram

\[
\begin{array}{c}
\mathcal{N} \\
\mu \downarrow \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \ quad
where $D^b_m(X)$ denotes the category of mixed complexes of $\bar{Q}_\ell$-sheaves (for some $\ell \neq \text{char } k$) that are constructible with respect to the $G$-orbits on $X$.

The main results, proved in Sections 3–4, are summarized in the statement below. This statement involves the following categories: $P_G(N) \subset D^b_m(N)$ is the abelian category of perverse sheaves; $\text{Spr} \subset P_G(N)$ is the Serre subcategory containing the simple perverse sheaves appearing in the Springer correspondence; and $\text{Sh}_G(\mathcal{B}) \subset D^b_m(\mathcal{B})$ is the abelian category of locally constant sheaves.

Theorem 1.1. The functor $\Phi$ restricts to give an exact functor of abelian categories $\Phi|_{P_G(N)} : P_G(N) \to \text{Sh}_G(\mathcal{B})$. Moreover, for $F, F' \in \text{Spr}$, the objects $\Phi(F)$ and $\Phi(F')$ and the vector space $\text{Hom}^i(\Phi(F), \Phi(F'))$ each carry a natural action of the Weyl group $W$, and $\Phi$ induces an isomorphism

$$
\text{Hom}^i_{D^b_m(N)}(F, F') \cong \text{Hom}^i_{\text{Sh}_G}(\Phi(F), \Phi(F'))^W.
$$

Here, the notation “$\text{Hom}$” denotes a Hom-group equipped with an action of Frobenius; along the way to the theorem above, we show that $\text{Hom}^i_{D^b_m(N)}(F, F')$ is pure. However, weights and purity are not used in any essential way; the main results are also valid in the unmixed setting.

The $W$-action on $\Phi(F)$ induces one on the space of global sections $\Gamma(\Phi(F))$, and the composition $\Gamma \circ \Phi|_{\text{Spr}} : \text{Spr} \to \text{Rep}(W)$ turns out to be an equivalence of categories that may be regarded as a categorical version of the Springer correspondence. On the other hand, the $W$-action on $\text{Hom}^i(\Phi(F), \Phi(F'))$ is a generalization of the usual action of $W$ on $H^\bullet(\mathcal{B})$. Indeed, (1.3) can be used together with known formulas for the fake degrees of $W$ to carry out explicit calculations of Ext-groups.

As an application, in Sections 5–6, we use Theorem 1.1 to give new proofs of two results from [L]: a decomposition of $D^b_m(N)$ into orthogonal subcategories, and the algorithmic description of stalks of perverse sheaves on $N$ mentioned above.

ACKNOWLEDGMENTS. The author is grateful to J.M. Douglass, A. Henderson, and D. Treumann for helpful conversations.

2. Preliminaries

2.1. General conventions. Throughout the paper, $G$ and all related varieties will be assumed to be defined over the algebraic closure $k$ of a finite field $\mathbb{F}_q$ and equipped with an $\mathbb{F}_q$-rational structure. For any $G$-variety $X$, $D^b_m(X)$ will denote the category of mixed étale $\bar{Q}_\ell$-complexes that are constructible with respect to some $G$-stable stratification. Let $1_X$ denote the constant sheaf with value $\bar{Q}_\ell$, and let $pt = \text{Spec } k$. Let $\omega_X = a^! 1_{pt}$ (where $a : X \to pt$ is the constant map) denote the dualizing complex, and let $\mathbb{D} = R\text{Hom}(\cdot, \omega_X)$ denote the Verdier duality functor.

For $F, F' \in D^b_m(X)$, we let $\text{Hom}^i(F, F') = H^i(a, R\text{Hom}(F, F'))$. This is a mixed $\bar{Q}_\ell$-sheaf on a point, i.e., a $\bar{Q}_\ell$-vector space equipped with an action
of Frobenius. Forgetting that action yields the \( \mathbb{Q}_\ell \)-vector space of morphisms \( F \to F' [i] \) over \( k \).

Nearly all results (those not explicitly involving purity) are also valid in the setting of unmixed sheaves over an arbitrary algebraically closed field \( k \) of good characteristic, or for \( \mathbb{C} \)-sheaves in the classical topology when \( k = \mathbb{C} \).

2.2. Further notation. Let \( d = \dim B \). We will frequently encounter shifts and Tate twists related to \( \dim B \), so we adopt the following shorthand notation: if \( F \) is any sheaf, morphism, or functor, we put \( F^\flat = F[2d](d) \) and \( F^\sharp = F[-2d](-d) \).

Because \( B \) is a smooth variety of dimension \( d \), we have \( \omega_B \cong 1_B^\flat \), and because \( \pi \) is a smooth map of relative dimension \( d \), we also have \( \pi^! \cong (\pi^*)^\flat \) and \( \omega_{\tilde{N}} \cong 1_{\tilde{N}}^\sharp \).

Throughout the paper, \( W \) will denote the universal Weyl group of \( G \), cf. \([CG, \S 3.1]\). This group does not depend on the choice of a maximal torus or a Borel subgroup, and we do not fix any such choice in this paper. Let \( \text{Irr}(W) \) denote the set of isomorphism classes of irreducible representations of \( W \) on \( \mathbb{Q}_\ell \)-vector spaces. For each \( \chi \in \text{Irr}(W) \), choose a representative \( V_\chi \). We will sometimes regard \( V_\chi \) as a pure object of weight 0 in \( D^b_{m}(\text{pt}) \), by letting the Frobenius act on it as the identity. It is known that

\[
V_\chi \cong V_\chi^*
\]

for all \( \chi \in \text{Irr}(W) \). However, this isomorphism is not canonical.

It is well known that \( W \) acts naturally on the cohomology ring \( H^\bullet(B) \), and that under this action, \( H^\bullet(B) \) can be identified with the coinvariant algebra of \( W \). Let

\[
\kappa : W \to \text{Aut}(H^\bullet(B))
\]

denote this action.

Let \( Z \) denote the Steinberg variety \( Z = \tilde{N} \times_{\mathbb{C}^*} \tilde{N} \). Finally, let \( i : B \to \tilde{N} \) be the inclusion of the zero section, and let \( i_0 : \text{pt} \to N \) denote the inclusion of the point 0. We then have a cartesian square

\[
\begin{array}{ccc}
B & \to & \tilde{N} \\
\downarrow i & & \downarrow i \\
\text{pt} & \to & N \\
\end{array}
\]

2.3. Springer correspondence. Let \( A = \mu_* 1_{\tilde{N}} = \mu_* 1_{\tilde{N}}[2d](d) \). This is a semisimple perverse sheaf on \( N \), known as the Springer sheaf. One approach to studying \( A \), developed in detail in \([CG]\), involves Borel–Moore homology, which is defined in terms of the hypercohomology of the dualizing complex. For our purposes, it is convenient to adopt a slightly nonstandard normalization and put

\[
\tilde{H}^i(X) = H^i(X, \omega_X[-4d](-2d)).
\]
The Borel–Moore homology of the Steinberg variety $\tilde{\mathcal{H}}^\bullet(Z)$ is equipped with a “convolution product,” making it into a graded algebra. Two key results are that there are natural algebra isomorphisms

$$\tilde{\mathcal{H}}^0(Z) \simeq \mathbb{Q}_\ell[W] \quad \text{and} \quad \tilde{\mathcal{H}}^\bullet(Z) \simeq \text{Hom}^\bullet(A, A),$$

and that the latter is an isomorphism of graded algebras. In particular, we have a natural isomorphism $\mathbb{Q}_\ell[W] \simeq \text{End}(A)$, and so an action

$$\sigma : W \to \text{Aut}(A).$$

Any action of $W$ on $A$ would allow us to decompose $A$ into isotypic components, but since $\mathbb{Q}_\ell[W] \simeq \text{End}(A)$, we actually have

$$A \simeq \bigoplus_{\chi \in \text{Irr}(W)} \text{IC}_\chi \otimes V_\chi,$$

where the IC$\chi$ are various distinct simple perverse sheaves on $\mathcal{N}$. This labeling of certain simple perverse sheaves by $\text{Irr}(W)$ is what is usually known as the Springer correspondence. For $\chi \in \text{Irr}(W)$, let $\mathcal{O}_\chi \subset \mathcal{N}$ be the unique open $G$-orbit in the support of IC$\chi$, and let $L_\chi$ be the irreducible global system given by $\text{IC}_\chi|_{\mathcal{O}_\chi} \cong -\dim \mathcal{O}_\chi(-\frac{1}{2} \dim \mathcal{O}_\chi)$. Thus, $\text{IC}_\chi \simeq \text{IC}(\mathcal{O}_\chi, L_\chi)$.

2.4. Modules for $\tilde{\mathcal{H}}^\bullet(Z)$. The convolution product construction also makes the Borel–Moore homology of any subvariety of $\mathcal{N}$ into a graded $\tilde{\mathcal{H}}^\bullet(Z)$-module. The convolution action of $\tilde{\mathcal{H}}^0(Z)$ on $\tilde{\mathcal{H}}^\bullet(\mathcal{B}) \simeq \tilde{\mathcal{H}}^\bullet(\mathcal{B})F$ is none other than the action $\kappa$. It is clear by base change in (2.2) that $i_0^* A \simeq \tilde{\mathcal{H}}^\bullet(\mathcal{B})F \simeq \tilde{\mathcal{H}}^\bullet(\mathcal{B})F$. The functor $i_0^*$ therefore induces a map $i_0^* : \text{End}(A) \to \text{End}(\tilde{\mathcal{H}}^\bullet(\mathcal{B}))$ that is a homomorphism of $\tilde{\mathcal{H}}^\bullet(Z)$-modules. In particular, it is $W$-equivariant, so we have

$$i_0^*(\sigma(w)f) = \kappa(w)i_0^*(f).$$

Finally, consider $\tilde{\mathcal{H}}^\bullet(\mathcal{N})$. This is a convolution algebra in its own right. Since $\tilde{\mathcal{H}}^\bullet(\mathcal{N}) \simeq \tilde{\mathcal{H}}^\bullet(\mathcal{N})$, it has another algebra structure given by cup product in ordinary cohomology, but it follows from [CG, Theorem 8.6.7] (cf. [DR, Theorem 2.3]) that these two algebra structures coincide. The following theorem of Douglass–Rörle [DR] relates the $W$-action on $\tilde{\mathcal{H}}^\bullet(\mathcal{N})$ to that on $\tilde{\mathcal{H}}^\bullet(Z)$.

**Theorem 2.1** (Douglass–Rörle). Let $\delta : \mathcal{N} \to Z$ be the diagonal embedding. The induced map $\delta_* : \tilde{\mathcal{H}}^\bullet(\mathcal{N}) \to \tilde{\mathcal{H}}^\bullet(Z)$ in Borel–Moore homology satisfies

$$\delta_*(w \ast f) = w \ast \delta_*(f) \ast w^{-1},$$

for any $w \in W$, where $\ast$ denotes the convolution product.

2.5. Weakly $G_m$-equivariant objects. Let $X$ be a variety endowed with an action of $G_m$. An object $F \in D^b_{m}(X)$ is said to be weakly equivariant if it is in the image of the forgetful functor $U : D^b_{m,G_m}(X) \to D^b_m(X)$, where $D^b_{m,G_m}(X)$ denotes the $G_m$-equivariant derived category of $X$ in the sense of Bernstein–Lunts [BL]. Weakly equivariant objects have the following useful property.
Lemma 2.2 (Springer [Sp, Proposition 1], cf. Braden [B, Lemma 6]). Let \( p : V \to Z \) be a vector bundle, and suppose \( \mathbb{G}_m \) acts linearly on the fibers of \( p \) with strictly positive weights (or strictly negative weights). Let \( i : Z \to V \) be the inclusion of the zero section. For a weakly equivariant object \( S \in D^b_{\mathbb{G}_m}(V) \), there are natural isomorphisms \( i^!S \to p_!S \) and \( p_*S \to i^*S \). \( \square \)

Remark 2.3. Any object obtained by pullback or push-forward of a weakly equivariant object along a \( \mathbb{G}_m \)-equivariant map is automatically weakly equivariant, and the constant sheaf is always weakly equivariant. Therefore:

- If \( \mathbb{G}_m \) acts on \( N \) by scaling along the fibers, and on \( \tilde{N} \) by scaling, then the objects \( 1_{\tilde{N}}, A, \mu^*A, \mu^!A \), and all direct summands of the last three are weakly equivariant.
- If \( \mathbb{G}_m \) acts on \( N \) by an action that factors through \( \tilde{G} \), where \( \tilde{G} \) is a group isogenous to \( G \) with simply-connected derived subgroup, then every semisimple perverse sheaf on \( N \) is weakly equivariant, since any local system on any nilpotent orbit is \( \tilde{G} \)-equivariant.

Remark 2.4. Suppose \( \phi \) is a morphism of functors that is an isomorphism on weakly equivariant objects. If the domain category of \( \phi \) is generated as a triangulated category by weakly equivariant objects, then a standard dévissage argument shows that \( \phi \) actually induces isomorphisms for all objects; in other words, \( \phi \) is an isomorphism of functors outright. This observation will be used when we apply Lemma 2.2 and other results to \( D^b_{\mathbb{G}_m}(N) \), which is generated by the objects of \( P_G(N) \).

3. Exactness of \( \Phi \)

In this section, we use hyperbolic localization to prove exactness results for the functor \( \Phi = \pi_!\mu^* : D^b_{\mathbb{G}_m}(N) \to D^b_{\mathbb{G}_m}(\mathcal{B}) \) of (1.2), as well as for the dual functor

\[
\Phi' = (\pi_*\mu!)^\vee : D^b_{\mathbb{G}_m}(N) \to D^b_{\mathbb{G}_m}(\mathcal{B}).
\]

To study these functors, we will make use of the additional functor

\[
\Psi = \mu_*\pi^! \simeq (\mu\pi^!)^\vee : D^b_{\mathbb{G}_m}(\mathcal{B}) \to D^b_{\mathbb{G}_m}(N).
\]

It is clear that \( \Psi \) is left-adjoint to \( \Phi' \) and right-adjoint to \( \Phi \). In addition, we have

\[
\Phi \simeq i^!\mu^*, \quad \Phi' \simeq (i^*\mu^!)^\vee, \quad \Psi(1_{\mathcal{B}}) \simeq A,
\]

with the first two assertions relying on Remark 2.4. Moreover, since all objects of \( \text{Sh}_G(\mathcal{B}) \) are direct sums of copies of \( 1_{\mathcal{B}} \), it follows that \( \Psi \) restricts to an exact functor of abelian categories \( \Psi : \text{Sh}_G(\mathcal{B}) \to \text{P}_G(N) \). The main result of this section is the following.

Theorem 3.1. The functors \( \Phi, \Phi' : D^b_{\mathbb{G}_m}(N) \to D^b_{\mathbb{G}_m}(\mathcal{B}) \) restrict to give isomorphic exact functors of abelian categories \( \Phi \simeq \Phi' : \text{P}_G(N) \to \text{Sh}_G(\mathcal{B}) \).

We first require the following preliminary result.

Lemma 3.2. Let \( e : \text{pt} \to \mathcal{B} \) be the inclusion of a point. There is a natural isomorphism of functors \( e^*\pi_*\mu^! \simeq e^*\pi_!\mu^* : D^b_{\mathbb{G}_m}(N) \to D^b_{\mathbb{G}_m}(\text{pt}) \).
Proof. For this proof, we fix a choice of Borel subgroup $B \subset G$ and maximal torus $T \subset B$. Recall that these choices yield a canonical identification $W = N_G(T)/T$. Let $g = u \oplus t \oplus u$ be the corresponding triangular decomposition of $g$. That is, $t = \text{Lie}(T)$, $u$ is the nilpotent radical of $\text{Lie}(B)$, and $\bar{u}$ is the nilpotent radical of the Lie algebra of the opposite Borel subgroup.

Choose a regular dominant cocharacter $\lambda : G_m \to T$, and let $G_m$ act on $g$ by composing $\lambda$ with the adjoint action of $T$ on $g$. Clearly, the triangular decomposition of $g$ is stable under this action. Moreover, $G_m$ acts on $u$ with positive weights and on $\bar{u}$ with negative weights, and it acts trivially on $t$. From these observations, it is easy to see that the point $0 \in N$ is the unique fixed point for the action of $G_m$ on $N$, and that

$$u = \{ x \in N | \lim_{z \in G_m} z x = 0 \} \quad \text{and} \quad \bar{u} = \{ x \in N | \lim_{z \to \infty} z x = 0 \}.$$

Consider the following diagram of inclusion maps:

$$\begin{array}{ccc}
pt & \xrightarrow{\iota} & u \\
\downarrow & & \downarrow \\
\bar{u} & \xrightarrow{g} & \tilde{N}
\end{array}$$

This is a setting in which we may apply the formalism of hyperbolic localization, following [B]. The main theorem of [B] states that there is a natural morphism of functors $i^* g^! \to i^! g^*$ that is an isomorphism on weakly equivariant objects. By Remark 2.4, this is an isomorphism of functors in our situation.

Next, consider the constant maps $p : u \to pt$ and $\bar{p} : \bar{u} \to pt$. Using Lemma 2.2, we obtain a natural isomorphism $\bar{p}_* g^! \to \bar{p}^! g^*$. Finally, let $e : pt \to \mathcal{B}$ (resp. $\bar{e} : pt \to \mathcal{B}$) be the inclusion of the point corresponding to the Borel subgroup $B$ (resp. the opposite Borel subgroup to $B$). Forming pullbacks over $\pi$, we obtain the diagrams

$$
\begin{array}{ccc}
u & \xrightarrow{\mu} & N \\
\downarrow & & \downarrow \pi \\
\tilde{N} & \xrightarrow{\bar{e}} & \mathcal{B}
\end{array}
\begin{array}{ccc}
u & \xleftarrow{\mu} & N \\
\downarrow & & \downarrow \pi \\
\tilde{N} & \xleftarrow{\bar{e}} & \mathcal{B}
\end{array}
$$

It is clear that $\mu \bar{e} = g$ and $\mu \bar{e} = \bar{g}$. By base change, we have $p \bar{e}^* \simeq e^* \mu$, and $\bar{p} \bar{e}^! \simeq \bar{e}^! \bar{\mu}_* \bar{\mu}$. Combining with our earlier observations, we obtain an isomorphism $e^* \pi_* \mu^! \simeq \bar{e}^! \pi_* \bar{\mu}_* \bar{\mu}$. It is clear that on $\mathcal{B}^n \equiv \mathcal{B}$, we have $e^* \simeq e^*$ and $e^! \simeq e^!$, so we now have the isomorphism $e^* \pi_* \mu^! \simeq e^* \pi_* \bar{\mu}_*$, as desired. $\square$

Proof of Theorem 3.1. We begin by showing that $\Phi|_{P_G(N)}$ is exact. Let $F \in P_G(N)$. We wish to show that $H^j(\Phi(F)) = 0$ for $j \neq 0$. First, observe that for $n < 0$, we have

$$\text{Hom}(\Phi(F), 1_{\mathcal{B}}[n]) \simeq \text{Hom}(F, \Psi(1_{\mathcal{B}})[n]) \simeq \text{Hom}(F, A[n]) = 0.$$

This shows that $H^j(\Phi(F)) = 0$ for $j > 0$. 

Documenta Mathematica 16 (2011) 869–884
By duality, we have $H^j(\pi_*\mu^*F) \simeq H^j(D\pi_*\mu^*DF) \simeq (DH^{-2d-j}(\pi_*\mu^*DF))[-2d]$. Thus, $H^j(\pi_*\mu^*F) = 0$ if $j < -2d$. For the inclusion of a point $e : \text{pt} \to \mathcal{B}$, we know that $e^*$ is an exact functor on $\text{Sh}_G(\mathcal{B})$, so it follows that $H^j(e^*\pi_*\mu^*F^j) = 0$ if $j < 0$. By Lemma 3.2, this implies that $H^j(e^*\pi_*\mu^*F) = 0$ for $j < 0$ as well. Now, $e^*$ is also faithful on the abelian category $\text{Sh}_G(\mathcal{B})$ (though not, of course, on $D^b(\mathcal{B})$), so that vanishing implies that $H^j(\pi_*\mu^*F) = 0$ for $j < 0$, as desired. Thus, $\Phi|_{P_G(N)}$ is exact.

Finally, note that the exact functor $e^* : \text{Sh}_G(\mathcal{B}) \to \text{Sh}(\text{pt})$ is an equivalence of categories. The isomorphism $\Phi|_{P_G(N)} \simeq \Phi|_{P_G(\mathcal{N})}$ then follows from Lemma 3.2.

**Corollary 3.3.** The functor $\Phi \simeq \Phi' : P_G(N) \to P_G(\mathcal{B})$ preserves purity. That is, it takes pure objects of weight $w$ to pure objects of weight $w$.

**Proof.** Let $F \in P_G(\mathcal{N})$ be pure of weight $w$. By the well-known rules [BBD, 5.1.14] for behavior of weights under various sheaf functors, we see that $\Phi = \pi_*\mu^*$ takes $F$ to an object with weights $\leq w$, whereas $\Phi' = \pi_*\mu^*$ takes it to one with weights $\geq w$. \hfill \square

### 4. Action of the Weyl Group

In Section 2, we considered the action $\sigma$ of $W$ on $\text{End}(A)$, and the action $\kappa$ on $H^*(\mathcal{B})$. In this section, we discuss several additional actions, and prove a $W$-equivariance result for $\Psi$. There are two natural commuting actions $\lambda, \rho : W \to \text{Aut}(\text{Hom}(A, A))$, given by

$$
\lambda(w)(f) = \sigma(w) \circ f \quad \text{and} \quad \rho(w)(f) = f \circ \sigma(w^{-1}).
$$

The exactness result of Section 3 allows us to construct a new action as follows.

**Proposition 4.1.** For any $F \in P_G(N)$, the sheaf $\Phi(F)$ carries a natural action of $W$. If $F$ is simple, then we have

$$
\Phi(F) \simeq \begin{cases} 0 & \text{if } F \notin \text{Spr}, \\ 1_\mathfrak{g} \otimes V_\chi & \text{if } F \simeq IC_\chi. \end{cases}
$$

**Proof.** The following general principle is easy to see: if $\mathfrak{g}$ is a semisimple $\mathbb{k}$-linear abelian category containing a unique simple object $S$ up to isomorphism, and $\text{End}(S) \simeq \mathbb{k}$, then any object $A$ is canonically isomorphic to $S \otimes \text{Hom}(S, A)$. Applying this to $\text{Sh}_G(\mathcal{B})$, we have $\Phi(F) \simeq 1_\mathfrak{g} \otimes \text{Hom}(1_\mathfrak{g}, F)$. By adjunction, $\text{Hom}(1_\mathfrak{g}, \Phi(F)) \simeq \text{Hom}(\Psi(1_\mathfrak{g}), F) \simeq \text{Hom}(1_\mathfrak{g}, F)$. The $W$-action on $A$ induces one on $\text{Hom}(A, F)$ for any $F$, and therefore on $\Phi(F) \simeq 1_\mathfrak{g} \otimes \text{Hom}(A, F)$. For simple $F$, it is clear from (2.4) that $\text{Hom}(A, F) = 0$ if $F \notin \text{Spr}$, and that $\text{Hom}(A, IC_\chi) \simeq V_\chi$. \hfill \square

The action described in this proposition will be denoted $\nu : W \to \text{Aut}(\Phi(F))$. This action gives rise to a $W$-action on the vector space $\Gamma(\Phi(F))$ for any $F \in P_G(\mathcal{N})$. In other words, the functor $\Gamma \circ \Phi$ may be regarded as taking values in $\text{Rep}(W)$. Since $P_G(\mathcal{N})$ and $\text{Rep}(W)$ are both semisimple categories, the following result is immediate from Proposition 4.1.
Theorem 4.2. For a simple perverse sheaf $F \in \mathcal{P}_G(N)$, the functor $\Gamma \circ \Phi : \mathcal{P}_G(N) \to \text{Rep}(W)$ is given by

$$(\Gamma \circ \Phi)(F) \simeq \begin{cases} 0 & \text{if } F \notin \text{Spr}, \\ V^*_\chi & \text{if } F \simeq \text{IC}_\chi. \end{cases}$$

In particular, $\Gamma \circ \Phi|_{\text{Spr}} : \text{Spr} \to \text{Rep}(W)$ is an equivalence of categories. □

Corollary 4.3. The category $\text{Spr} \subset \mathcal{P}_G(N)$ is stable under Verdier duality $\mathbb{D}$. In fact, for each simple perverse sheaf $\text{IC}_\chi \in \text{Spr}$, we have $\mathbb{D}\text{IC}_\chi \simeq \text{IC}_\chi$.

Proof. It follows from Theorem 3.1 and the fact that $\mathcal{B}$ is a projective variety that $\Gamma \circ \Phi$ commutes with $\mathbb{D}$. For a simple perverse sheaf $F \in \mathcal{P}_G(N)$, we see that $(\Gamma \circ \Phi)(F) \neq 0$ if and only if $(\Gamma \circ \Phi)(\mathbb{D}F) \neq 0$, so $\mathbb{D}$ preserves $\text{Spr}$. Moreover, for $F \simeq \text{IC}_\chi$, we have $(\Gamma \circ \Phi)(\mathbb{D}\text{IC}_\chi) \simeq \mathbb{D}(V^*_\chi) \simeq V^*_\chi$. The result follows using the noncanonical isomorphism (2.1). □

Note that when $F = A$, $\nu$ is obtained via the adjunction isomorphism (4.1)

$$\theta : \text{Hom}(1_{\mathcal{B}}, \Phi(A)) \sim \text{Hom}(A, A)$$

from the action on $\text{Hom}(A, A)$ that we have called $\rho$. The other action $\lambda$ on $\text{Hom}(A, A)$ also induces an action on $\Phi(A)$, which we denote $\hat{\lambda} : W \to \text{Aut}(\Phi(A))$.

Since $\rho$ and $\lambda$ commute, $\nu$ and $\hat{\lambda}$ commute as well. By an abuse of notation, we will also write $\nu$ and $\hat{\lambda}$ for the corresponding actions on the vector space $\text{Hom}_i(1_{\mathcal{B}}, \Phi(A))$. By definition, we have

$$\theta(\hat{\lambda}(\nu)f) = \lambda(\nu)\rho(f) \theta(f) = \sigma(\nu) \circ \theta(f) \circ \sigma(w^{-1}).$$

Lemma 4.4. The map $\Psi : \text{Hom}^1(1_{\mathcal{B}}, 1_{\mathcal{B}}) \to \text{Hom}^1(A, A)$ has the property that $\Psi(\nu)f = \sigma(\nu) \circ \Psi(f) \circ \sigma(w^{-1})$.

Proof. This is essentially a restatement of Theorem 2.1 due to Douglass–Röhrle. Note first that the functor $\pi^!$ induces a $W$-equivariant isomorphism $H^*(\mathcal{B}) \sim H^*(\tilde{N})$, so it suffices to study the $W$-equivariance of $\mu_* : D_m^b(\tilde{N}) \to D_m^b(N)$. Let $q : Z \to N$ be the natural projection map. We then have a commutative diagram:

Recall [CG, Section 8.6] that the second isomorphism in (2.3) arises from the fact that $q_*\omega^\sharp_Z \simeq R\text{Hom}(A, A)$. There is also a natural adjunction map $\delta_*\omega^\natural_N \to \omega_Z$. Consider the composition

$$\mu_*\omega^\natural_N \xrightarrow{q_*\omega^\sharp_Z} R\text{Hom}(A, A).$$
Applying $R\Gamma$ to the first map yields the induced map $\delta_* : \hat{H}^*(\mathcal{N}) \to \hat{H}^*(Z)$ in Borel–Moore homology. On the other hand, we can identify $\omega^\mathcal{P}_\mathcal{N}^Z \cong 1_{\mathcal{N}} \cong R\text{Hom}(1_{\mathcal{N}}, 1_{\mathcal{N}})$, and then the composition above becomes the canonical morphism

$$\mu_* R\text{Hom}(1_{\mathcal{N}}, 1_{\mathcal{N}}) \to R\text{Hom}(A, A),$$

and applying $R\Gamma$ gives us the map $\text{Hom}^*(1_{\mathcal{N}}, 1_{\mathcal{N}}) \to \text{Hom}^*(A, A)$ induced by $\mu$. To summarize, we have the following commutative diagram:

$$\begin{array}{ccc}
\hat{H}^*(\mathcal{N}) & \xrightarrow{\delta_*} & \text{Hom}^*(\mathcal{N}, \mathcal{N}) \\
\downarrow & & \downarrow \\
\hat{H}^*(Z) & \xrightarrow{\mu} & \text{Hom}^*(A, A)
\end{array}$$

Since the top isomorphism is $W$-equivariant, and the bottom one is an algebra isomorphism sending $w \in \mathcal{Q}_W W \cong \hat{H}^0(Z)$ to $\sigma(w) \in \text{End}(A)$, the result follows from Theorem 2.1.

\begin{proposition}
There is a natural isomorphism

$$\alpha : H^*(\mathcal{D}) \otimes \text{Hom}(1_{\mathcal{D}}, \Phi(A)) \cong \text{Hom}^*(1_{\mathcal{D}}, \Phi(A)).$$

Its composition with the adjunction $\theta$, denoted

$$\Theta = \theta \circ \alpha : H^*(\mathcal{D}) \otimes \text{Hom}(1_{\mathcal{D}}, \Phi(A)) \cong \text{Hom}^*(A, A),$$

is $W$-equivariant in the following way: for $u, v, w \in W$, we have

$$\Theta(\kappa(u)f \otimes \hat{\lambda}(v)\nu(w)\theta) = \lambda(v)\rho(w)\Theta(\kappa(w)^{-1}u)f \otimes \theta g.$$  

\end{proposition}

\begin{proof}
Recall from the proof of Proposition 4.1 that $\Phi(A) \simeq 1_{\mathcal{D}} \otimes \text{Hom}(1_{\mathcal{D}}, \Phi(A))$. It follows that

$$\text{Hom}^*(1_{\mathcal{D}}, \Phi(A)) \simeq \text{Hom}^*(1_{\mathcal{D}}, 1_{\mathcal{D}} \otimes \text{Hom}(1_{\mathcal{D}}, \Phi(A))) \simeq \text{Hom}^*(1_{\mathcal{D}}, 1_{\mathcal{D}}) \otimes \text{Hom}(1_{\mathcal{D}}, \Phi(A)).$$

Since $H^*(\mathcal{D}) \simeq \text{Hom}^*(1_{\mathcal{D}}, 1_{\mathcal{D}})$, we obtain the desired isomorphism $\alpha$. Note that $\alpha$ is given by composition: that is, if $f \in \text{Hom}^*(1_{\mathcal{D}}, 1_{\mathcal{D}})$ and $g \in \text{Hom}(1_{\mathcal{D}}, \Phi(A))$, then

$$\alpha(f \otimes g) = g \circ f.$$  

The adjunction isomorphism $\theta$ behaves on compositions according to the rule $\theta(g \circ f) = \theta(g) \circ \Psi(f)$. Using (4.2) and Lemma 4.4, we find

$$\Theta(\kappa(u)f \otimes \hat{\lambda}(v)\nu(w)\theta) = \lambda(v)\rho(w)\theta(\kappa(w)^{-1}u)f \otimes \theta g.$$  

The result follows.

\end{proof}
THEOREM 4.6. For \( \chi, \psi \in \text{Irr}(W) \), there is a natural isomorphism

\[
\text{Hom}^i(\Phi(IC_{\chi}), \Phi(IC_{\psi})) \simeq V_{\chi} \otimes H^i(\mathcal{R}) \otimes V_{\psi}^*,
\]

and thus \( \text{Hom}^i(\Phi(IC_{\chi}), \Phi(IC_{\psi})) \) is endowed with a natural action of \( W \). Moreover, \( \Phi \) induces isomorphisms

\[
\text{Hom}^i(IC_{\chi}, IC_{\psi}) \simeq \text{Hom}^i(\Phi(IC_{\chi}), \Phi(IC_{\psi}))^W \simeq (V_{\chi} \otimes H^i(\mathcal{R}) \otimes V_{\psi}^*)^W.
\]

Proof. The isomorphism (4.3) is immediate from Proposition 4.1. Next, using (2.4), we can decompose \( \text{Hom}^i(A, A) \) as

\[
\text{Hom}^i(A, A) \simeq \bigoplus_{\chi, \psi} \text{Hom}^i(IC_{\chi}, IC_{\psi}) \otimes V_{\chi}^* \otimes V_{\psi}.
\]

Thus, in terms of the action of \( W \times W \) on \( \text{Hom}^i(A, A) \) by \( \lambda \boxtimes \rho \), we can find \( \text{Hom}^i(IC_{\chi}, IC_{\psi}) \) by picking out the \( \chi^* \boxtimes \psi \)-isotypic component:

\[
\text{Hom}^i(IC_{\chi}, IC_{\psi}) \simeq \text{Hom}_{W \times W}(V_{\chi}^* \boxtimes V_{\psi}, \text{Hom}^i(A, A)).
\]

Using Proposition 4.5, this is isomorphic to

\[
\text{Hom}_{W \times W}(V_{\chi}^* \boxtimes V_{\psi}, H^i(1_{\mathcal{R}}) \otimes \text{Hom}(1_{\mathcal{R}}, \Phi(A))),
\]

where \( W \times W \) acts on \( H^i(1_{\mathcal{R}}) \otimes \text{Hom}(1_{\mathcal{R}}, \Phi(A)) \) by \( \lambda \boxtimes \nu \). That is, for \( v, w \in W \) and \( f \otimes g \in H^i(1_{\mathcal{R}}) \otimes \text{Hom}(1_{\mathcal{R}}, \Phi(A)) \), we put

\[
(v, w) \cdot (f \otimes g) = \kappa(w)f \otimes \lambda(v)\nu(w)g.
\]

Using the adjunction (4.1) and the isomorphism (2.3), we see that \( \text{Hom}(1_{\mathcal{R}}, \Phi(A)) \) decomposes under \( \lambda \boxtimes \nu \) as \( \text{Hom}(1_{\mathcal{R}}, \Phi(A)) \simeq \bigoplus_{\nu} V_{\nu} \otimes V_{\nu}^* \). Picking off the \( \chi^* \)-isotypic component for the first factor of \( W \), we find that

\[
\text{Hom}^i(IC_{\chi}, IC_{\psi}) \simeq \text{Hom}(V_{\psi}, H^i(\mathcal{R}) \otimes V_{\chi}) \simeq (V_{\psi}^* \otimes H^i(\mathcal{R}) \otimes V_{\chi})^W.
\]

Corollary 4.7. For \( \chi, \psi \in \text{Irr}(W) \), \( \text{Hom}^i(IC_{\chi}, IC_{\psi}) \) vanishes if \( i \) is odd, and is pure of weight \( i \) if \( i \) is even.

Proof. This follows from the previous theorem and the well-known fact that \( H^i(\mathcal{R}) \) vanishes if \( i \) is odd and is pure of weight \( i \) if \( i \) is even.

5. ORTHOGONAL DECOMPOSITION OF \( D_m^b(A) \)

For a \( G \)-stable locally closed subvariety \( Y \subseteq N \), and let \( D_{\text{Spr}}(Y) \subseteq D_m^b(Y) \) be the full triangulated subcategory generated by objects \( IC_{\chi}|_Y \) with \( \mathcal{O}_Y \subseteq Y \). On the other hand, let \( D_{\mathcal{O}}(Y) \subseteq D_m^b(Y) \) be the full triangulated subcategory generated by simple perverse sheaves \( IC(\mathcal{O}, L)|_Y \) with \( \mathcal{O} \subseteq Y \) but \( IC(\mathcal{O}, L) \notin \text{Spr} \).

THEOREM 5.1. For any \( G \)-stable locally closed subvariety \( u : Y \hookrightarrow N \), we have

\[
D_m^b(Y) \simeq D_{\text{Spr}}(Y) \oplus D_{\mathcal{O}}(Y).
\]
Moreover, if $s : Z \hookrightarrow Y$ is the inclusion of a smaller $G$-stable locally closed subvariety, the functors $s^*$ and $s^!$ respect this decomposition: we have

\begin{align}
& s^*(D_{Spr}(Y)), \ s^!(D_{Spr}(Y)) \subset D_{Spr}(Z), \\
& s^*(D_{Spr}^+(Y)), \ s^!(D_{Spr}^+(Y)) \subset D_{Spr}^+(Z).
\end{align}

(5.2)

Proof. If $s : Z \to Y$ is an open embedding, then (5.2) is obvious. Since the inclusion of any locally closed subvariety can be factored as a closed embedding followed by an open embedding, we henceforth treat (5.2) only in the case where $Z$ is closed in $Y$.

Let $n_Y$ denote the number of nilpotent orbits in $Y \setminus Y$. We will prove (5.1) and (5.2) simultaneously by induction on $n_Y$. Note that (5.1) is equivalent to the assertion that for $F \in D_{Spr}(Y)$ and $F' \in D_{Spr}(Y)$, we have $\text{Hom}(F, F') = \text{Hom}(F', F) = 0$. We can further reduce to the case where $F$ and $F'$ are shifts of simple perverse sheaves. That is, (5.1) is equivalent to the statement that if $F$ and $F'$ are simple perverse sheaves with $F \in D_{Spr}(Y)$ and $F' \in D_{Spr}^+(Y)$, then

\begin{equation}
\text{Hom}^i(F, F') = \text{Hom}^i(F', F) = 0 \quad \text{for all } i \geq 0.
\end{equation}

(5.3)

We begin by proving (5.3) in the case where $n_Y = 0$, i.e., when $Y$ is closed in $N$. In fact, since $u_s : D_m(Y) \to D_m(N)$ is faithful for any closed $Y \subset N$, we may reduce to the case where $Y = N$. Since $F \in \text{Spr}$, $F$ is a direct summand of $A$, and it suffices to show that $\text{Hom}^i(A, F') = \text{Hom}^i(F', A) = 0$. Since $A \simeq \Psi(1_{\mathfrak{g}})$, we have by adjunction that

\begin{equation}
\text{Hom}^i(A, F') \simeq \text{Hom}^i(1_{\mathfrak{g}}, \Phi'(F')) \quad \text{and} \quad \text{Hom}^i(F', A) \simeq \text{Hom}^i(\Phi(F), 1_{\mathfrak{g}}).
\end{equation}

Since $\Phi(F') = \Phi(F') = 0$ by Theorem 3.1 and Proposition 4.1, the desired vanishing holds.

Suppose now that (5.1) is known to hold whenever $n_Y \leq k$. Let us prove (5.2). Since $Z$ is assumed to be a closed subvariety of $Y$, we clearly have $n_Z \leq n_Y$; in particular, we know that $D_{Spr}^+(Z) \simeq D_{Spr}(Z) \oplus D_{Spr}^+(Z)$. Therefore, given $F \in D_{Spr}(Y)$, we have a canonical decomposition $s^*F \simeq (s^*F)_{Spr} \oplus (s^*F)_{Spr}^+$ with $(s^*F)_{Spr} \in D_{Spr}(Z)$ and $(s^*F)_{Spr}^+ \in D_{Spr}^+(Z)$. We wish to prove that $(s^*F)_{Spr}^+ = 0$. If that is not the case, there certainly exists some simple perverse sheaf $F' \in D_{Spr}^+(Z)$ and some $i \in \mathbb{Z}$ such that $\text{Hom}^i((s^*F)_{Spr}^+, F') \neq 0$. We also know that $\text{Hom}^i((s^*F)_{Spr}, F') = 0$, so it follows that

\begin{equation}
0 \neq \text{Hom}^i((s^*F)_{Spr}^+, F') \simeq \text{Hom}^i(s^*F, F') \simeq \text{Hom}^i(F, s^*F')
\end{equation}

But $s^*F'$ is clearly a simple perverse sheaf in $D_{Spr}^+(Y)$, and since $F \in D_{Spr}(Y)$, we have a contradiction. Thus, $s^*F \in D_{Spr}(Z)$. The proofs of the other assertions in (5.2) are parallel.

Now, suppose that (5.1) and (5.2) are both known for $n_Y \leq k$, and let us prove (5.3) when $n_Y = k + 1$. Let $F$ and $F'$ be simple perverse sheaves on $Y$ with $F \in D_{Spr}(Y)$ and $F' \in D_{Spr}^+(Y)$. Choose an orbit $O$ that is open in $Y \setminus Y$, and let $\hat{Y} = Y \cup O$. Let $s : O \to \hat{Y}$ and $j : Y \to \hat{Y}$ be the inclusion...
maps, and consider the objects \( j_\ast F, j_\ast F' \in D^b_m(\tilde{Y}) \). These are simple perverse sheaves on \( \tilde{Y} \); moreover, we clearly have \( j_\ast F \in D_{\text{Spr}}(\tilde{Y}) \) and \( j_\ast F' \in D^b_{\text{Spr}}(\tilde{Y}) \). Form the distinguished triangle

\[
s_\ast s'_\ast j_\ast F' \rightarrow j_\ast F' \rightarrow j_\ast j^\ast j_\ast F' \rightarrow.
\]

Note that \( j^\ast j_\ast F' \simeq F' \). Next, form the long exact sequence

\[
\cdots \rightarrow \text{Hom}(j_\ast F, j_\ast F') \rightarrow \text{Hom}(j_\ast F, j_\ast F') \rightarrow \text{Hom}^{i+1}(j_\ast F', s_\ast s'_\ast j_\ast F') \rightarrow \cdots.
\]

Note that \( n_Y = n_Y - 1 \), so (5.3) holds on \( \tilde{Y} \) by assumption. In particular, we have \( \text{Hom}^\bullet(j_\ast F, j_\ast F') = 0 \). Since (5.2) also holds by assumption, we have \( s'_\ast j_\ast F \in D_{\text{Spr}}(O) \) and \( s'_\ast j_\ast F' \in D^b_{\text{Spr}}(O) \), so

\[
\text{Hom}^\bullet(j_\ast F', s_\ast s'_\ast j_\ast F') = \text{Hom}^\bullet(s'_\ast j_\ast F', s'_\ast j_\ast F') = 0.
\]

We conclude that \( \text{Hom}^i(j_\ast F, j_\ast F') = 0 \). But that means \( \text{Hom}^i(j^\ast j_\ast F, F') \simeq \text{Hom}^i(F, F') = 0 \) as well, as desired. \( \Box \)

One concrete consequence of this theorem is the following.

**Corollary 5.2.** If \( L \) is an irreducible local system on an orbit \( O \subset \mathcal{N} \) that occurs as a composition factor of some cohomology sheaf \( H^i(\text{IC}_X|_O) \) with \( \chi \in \text{Irr}(W) \), then \( L \simeq L_\psi \) for some \( \psi \in \text{Irr}(W) \). \( \Box \)

### 6. Green Functions

The aim of this section is to study the restrictions \( \text{IC}_X|_O \) as \( O \) varies over the \( G \)-orbits in \( \mathcal{N} \). Specifically, we encode information about these restrictions in a family of polynomials \( p^\chi_\psi(t) \), sometimes called *Green functions*. The main result, Theorem 6.1, gives a way to compute these polynomials, by relating them to the known groups \( \text{Hom}^i(\text{IC}_X, \text{IC}_\psi) \).

We begin with some notation. For a variety \( X \), let \( K(X) \) denote the quotient of the Grothendieck group of \( D^b_m(X) \) obtained by identifying the classes of simple perverse sheaves of the same weight that become isomorphic when the Weil structure is forgotten. (Thus, \( K(X) \) does not detect twists in the Weil structure by a root of unity.) For an object \( F \in D^b_m(X) \), let \( [F] \) denote its class in \( K(\mathcal{N}) \). This Grothendieck group is naturally a module over the Laurent polynomial ring \( \mathbb{Z}[t, t^{-1}] \), where multiplication by \( t \) corresponds to Tate twist: \( [F(1)] = t^{-1}[F] \). By choosing a square root of the Tate sheaf, we can regard \( K(X) \) as a module over \( \mathbb{Z}[t^{1/2}, t^{-1/2}] \), with \( [F(\frac{1}{2})] = t^{-1/2}[F] \). For instance, the group \( K(\text{pt}) \) is a free \( \mathbb{Z}[t^{1/2}, t^{-1/2}] \)-module of rank 1, generated by the class \( [1_{\text{pt}}] \) of a 1-dimensional vector space of weight 0.

For an orbit \( O \subset \mathcal{N} \), the group \( K(O) \) is a free \( \mathbb{Z}[t^{1/2}, t^{-1/2}] \)-module generated by the classes of irreducible local systems on \( O \). In view of Corollary 5.2, we may write

\[
(6.1) \quad [\text{IC}_X|_O] = \sum_{\psi|_{\mathcal{O}_x} = \psi} p^\chi_\psi(t)[L_\psi] \quad \text{for some } p^\chi_\psi(t) \in \mathbb{Z}[t^{1/2}, t^{-1/2}].
\]
Another description of these polynomials is as follows. Each cohomology sheaf $O$ series. We then have

\[ p_{\chi, \psi}(t) = \begin{cases} t^{-(\dim O_{\chi})/2} & \text{if } \psi = \chi, \\ 0 & \text{if } O_{\chi} \not\subset O_{\psi}, \text{ or if } O_{\psi} = O_{\chi} \text{ and } \psi \neq \chi. \end{cases} \]

Our goal is to determine the polynomials $p_{\chi, \psi}(t)$. Another description of these polynomials is as follows. Each cohomology sheaf $H^i(IC_{\chi}(\sigma))$ is, of course, a finite-length object in the category of local systems on $O$. Let $(H^i(IC_{\chi}(\sigma) : L_{\psi}(j)))$ denote the multiplicity of $L_{\psi}(j)$ in any composition series. We then have

\[ p_{\chi, \psi}(t) = \sum_{i \in \mathbb{Z}, j \in \mathbb{Z}} (-1)^i (H^i(IC_{\chi}(\sigma)) : L_{\psi}(j)) t^{-j}. \]

A result of Springer leads to a tremendous simplification of this formula; see Remark 6.2. Actually, $p_{\chi, \psi}(t)$ lies in $\mathbb{Z}[t^{-1}]$ and has nonnegative coefficients (see (6.7)), but we will not require these facts.

To state the main result, we require two additional families of polynomials. First, define $\lambda_{\chi, \psi}(t) \in \mathbb{Z}[t^{1/2}, t^{-1/2}]$ as follows:

\[ [R\Gamma_c(O, L_{\chi} \otimes L_\psi)] = \lambda_{\chi, \psi}(t)\mathbb{1}_{\text{pt}} \quad \text{if } O_{\chi} = O_{\psi}, \]

\[ \lambda_{\chi, \psi}(t) = 0 \quad \text{if } O_{\chi} \neq O_{\psi}. \]

Second, define polynomials $\omega_{\chi, \psi}(t)$ by

\[ [\mathbb{D} R\text{Hom}(IC_{\chi}, IC_{\psi})] = \omega_{\chi, \psi}(t)\mathbb{1}_{\text{pt}}. \]

Recall from Corollary 4.7 that $R\text{Hom}(IC_{\chi}, IC_{\psi})$ is pure of weight 0 and has vanishing cohomology in odd degrees. The same statements then hold for its dual as well. As with $p_{\chi, \psi}(t)$ in (6.7), it follows that

\[ \omega_{\chi, \psi}(t) = \sum_{i \in \mathbb{Z}} \dim H^i(\mathbb{D} R\text{Hom}(IC_{\chi}, IC_{\psi})) t^i = \sum_{i \in \mathbb{Z}} \dim \text{Hom}_{\mathbb{D}}^{-2i}(IC_{\chi}, IC_{\psi}) t^i. \]

The coinvariant algebra of $W$ has the property that the $W$-action in complementary degrees is related by tensoring with the sign character $\varepsilon$: that is, $H^i(\mathbb{R}) \simeq H^{2d-i}(\mathbb{R}) \otimes \varepsilon$ as $W$-representations. Using the noncanonical isomorphism (2.1) together with Theorem 4.6, we can rewrite the above formula as

\[ \omega_{\chi, \psi}(t) = t^{-2d} \sum_{i \in \mathbb{Z}} \dim \text{Hom}_W(V_{\chi} \otimes V_{\psi} \otimes \varepsilon, H^{2i}(\mathbb{R})) t^i. \]

The main result of this section is the following.

**THEOREM 6.1.** The matrices $P = (p_{\chi, \psi})$, $\Lambda = (\lambda_{\chi, \psi})$, and $\Omega = (\omega_{\chi, \psi})$ satisfy

\[ P \Lambda P^t = \Omega, \]

\[ \text{Documenta Mathematica 16 (2011) 869–884} \]
where $P^t$ is the transpose of $P$. In other words, given $\chi, \psi \in \text{Irr}(W)$, we have
\[ \sum_{\phi, \phi' \in \text{Irr}(W)} p_{\chi, \phi}(t) \lambda_{\phi, \phi'}(t) p_{\phi, \phi'}(t) = \omega_{\chi, \psi}(t). \]
Moreover, $P$ and $\Lambda$ are the unique matrices with entries in $\mathbb{Q}(t^{1/2})$ satisfying (6.6), (6.2), and (6.4).

Remark 6.2. This theorem is essentially equivalent to the part of [L, Theorem 24.8] relevant to Spr. The most substantial difference is that in loc. cit., the polynomials $p_{\chi, \psi}(t)$ are defined in a slightly different way. Correcting for different normalization conventions (see Remark 6.3), the definition in [L] is
\[ (6.7) \quad p_{\chi, \psi}(t) = \sum_{i \in \mathbb{Z}} (H^2(\text{IC}_\chi|_\mathcal{O}) : L_\psi(-i)) t^i. \]
The equivalence of this formula with (6.3) is implied by an important result of Springer [Sp], which states that for any $\chi \in \text{Irr}(W)$ and any orbit $\mathcal{O} \subset \mathcal{N}$, the object $\text{IC}_\chi|_\mathcal{O} \in \text{D}^b_m(\mathcal{O})$ is pure of weight 0, and that $H^i(\text{IC}_\chi|_\mathcal{O}) = 0$ if $i$ is odd. The proof of (6.6) in [L] also relies on Springer’s purity theorem.

Remark 6.3 (Lusztig–Shoji algorithm). The uniqueness asserted in Theorem 6.1 is proved by Lusztig [L] in a very explicit constructive way. This proof, which will not be repeated here, consists primarily of a description of an algorithm for finding $P$ and $\Lambda$ from knowledge of $\Omega$. Since $\Omega$ can be described as in (6.5) using only the representation theory of $W$, this algorithm can be effectively used to compute the $p_{\chi, \psi}(t)$. Generalizations of this algorithm, sometimes called the Lusztig–Shoji algorithm, have been studied in [AA, AH, GM, S3, S4, S5], and a computer implementation is available at [A].

The reader should be aware that Lusztig originally used polynomials $p'_{\chi, \psi}(t)$ and $\omega'_{\chi, \psi}(t)$ following different normalization conventions, while the recent works mentioned above involve polynomials $p''_{\chi, \psi}(t)$ and $\omega''_{\chi, \psi}(t)$ following a third convention. The relationship among these is as follows:
\[ p'_{\chi, \psi}(t) = t^{\dim \mathcal{O}} p_{\chi, \psi}(t) \quad \quad p''_{\chi, \psi}(t) = t^d p_{\chi, \psi}(t) \]
\[ \omega'_{\chi, \psi}(t) = t^{\dim \mathcal{O} + \dim \mathcal{O}} \omega_{\chi, \psi}(t) \quad \quad \omega''_{\chi, \psi}(t) = t^{2d} \omega_{\chi, \psi}(t). \]

For the next three lemmas, let $j_\mathcal{O} : \mathcal{O} \to \mathcal{N}$ denote the inclusion of an orbit.

Lemma 6.4. For any $F, F' \in \text{D}^b_m(\mathcal{N})$, we have
\[ [\mathbb{D} \text{RHom}(F, F')] = \sum_{\mathcal{O} \subset \mathcal{N}} [\mathbb{D} \text{RHom}(j_\mathcal{O}^* F, j_\mathcal{O}^* F')]. \]

Proof. Let $X \subset \mathcal{N}$ denote the closure of the support of $F'$, and let $\mathcal{O}_0 \subset X$ be an orbit that is open in $X$. Let $h : Y \to \mathcal{N}$ be the inclusion of the closed subset $Y = X \setminus \mathcal{O}_0$, and let $s : (\mathcal{N} \setminus Y) \to \mathcal{N}$ be the inclusion of its open complement. Consider the distinguished triangle $h_* h^! F' \to F' \to s_* s^* F' \to$. Since $s^* F' \simeq s^! F'$ is supported on $\mathcal{O}$, we see that $s_* s^* F'$ is naturally isomorphic.
to $j_{O^\prime}j_{O_n}^J F'$. Applying $\mathbb{D} R\text{Hom}(F, \cdot)$ and the usual adjunction properties, we obtain a distinguished triangle

$$
\mathbb{D} R\text{Hom}(j_{O_n}^J F, j_{O_n}^J F') \to \mathbb{D} R\text{Hom}(F, F') \to \mathbb{D} R\text{Hom}(h^* F, h^* F') \to \mathbb{D} R\text{Hom}(F, F')
$$

so $[\mathbb{D} R\text{Hom}(F, F')] = [\mathbb{D} R\text{Hom}(h^* F, h^* F')] + [\mathbb{D} R\text{Hom}(j_{O_n}^J F, j_{O_n}^J F')]$. The result then follows by induction on the number of orbits in the support of $F'$.

**Lemma 6.5.** For any orbit $O \subset N$, we have

$$
[j_{O}^{\dagger}\text{IC}_{\chi}] = t^{-\text{dim} O} \sum_{\{\psi \mid O_\psi = O\}} p_{\chi, \psi}(t^{-1})[L_\psi].
$$

**Proof.** Using Corollary 4.3, we have $j_{O}^{\dagger}\text{IC}_{\chi} \simeq \mathbb{D} j_{O}^{\dagger}(\mathbb{D}\text{IC}_{\chi}) \simeq \mathbb{D} (\text{IC}_{\chi}|_O)$. We can therefore obtain a formula for $[j_{O}^{\dagger}\text{IC}_{\chi}]$ by applying $\mathbb{D}$ to the righthand side of (6.1). For any local system $L_\psi$ on $O$, we have $\mathbb{D}(L_\psi(-i)) \simeq L_\psi[2 \text{dim } O](\text{dim } O + i)$, so the map $[\mathbb{D}(\cdot)] : K(O) \to K(O)$ sends $t^i[L_\psi] \mapsto t^{-\text{dim } O - i}[L_\psi]$. The result follows.

**Lemma 6.6.** We have

$$
[\mathbb{D} R\text{Hom}(j_{O}^{\dagger}\text{IC}_{\chi}, j_{O}^{\dagger}\text{IC}_{\psi})] = \sum_{\{\phi, \phi' \mid O_\phi = O_{\phi'} = O\}} p_{\chi, \phi}(t)\lambda_{\phi, \phi'}(t)[\mathbb{D} R\text{Hom}(L_\phi, L_{\phi'})].
$$

**Proof.** Observe that $\mathbb{D} R\text{Hom}$ transforms Tate twists according to the formula $\mathbb{D} R\text{Hom}(F(n), F'(m)) \simeq \mathbb{D} R\text{Hom}(F, F')(n - m)$. That means that the homomorphism $[\mathbb{D} R\text{Hom}(\cdot, \cdot)] : K(O) \times K(O) \to K(pt)$ is $\mathbb{Z}[1/2, t^{-1/2}]$-linear in the first variable, but antilinear with respect to the involution $t^{1/2} \mapsto t^{-1/2}$ in the second variable. Using (6.1) and Lemma 6.5, we find that

$$
[\mathbb{D} R\text{Hom}(j_{O}^{\dagger}\text{IC}_{\chi}, j_{O}^{\dagger}\text{IC}_{\psi})] = t^{\text{dim } O} \sum_{\{\phi, \phi' \mid O_\phi = O_{\phi'} = O\}} p_{\chi, \phi}(t)p_{\psi, \phi'}(t)[\mathbb{D} R\text{Hom}(L_\phi, L_{\phi'})].
$$

It suffices to show that $[\mathbb{D} R\text{Hom}(L_\phi, L_{\phi'})] = t^{-2 \text{dim } O} \lambda_{\phi, \phi'}(t)[1_{\text{pt}}]$. Using Corollary 4.3, we have that $\mathbb{D} R\text{Hom}(L_\phi, L_{\phi'}) \simeq \mathbb{D} R\text{Hom}(1_O, L_\phi^* \otimes L_{\phi'}) \simeq R\Gamma(O, L_\phi \otimes L_{\phi'})$, where $L_\phi^*$ denotes the dual local system: $L_\phi^* = (\mathbb{D} L_\phi)[-2 \text{dim } O](\text{dim } O)$. Therefore,

$$
\mathbb{D} R\text{Hom}(L_\phi, L_{\phi'}) \simeq R\Gamma_c(O, D(L_\phi \otimes L_{\phi'})) \simeq R\Gamma(O, L_\phi \otimes L_{\phi'})[2 \text{dim } O](\text{dim } O),
$$

so $[\mathbb{D} R\text{Hom}(L_\phi, L_{\phi'})] = t^{-\text{dim } O}[R\Gamma_c(O, L_\phi \otimes L_{\phi'})]$, as desired.

**Proof of Theorem 6.1.** The equation (6.6) follows from Lemmas 6.4 and 6.6, and the uniqueness assertion has been addressed in Remark 6.3. □

**References**

[A] P. Achar, *An implementation of the generalized Lusztig–Shoji algorithm*, software available for download from http://www.math.lsu.edu/~pramod/, 2008.
[AA] P. Achar and A.-M. Aubert, Springer correspondences for dihedral groups, Transform. Groups 13 (2008), 1–24.

[AH] P. Achar and A. Henderson, Orbit closures in the enhanced nilpotent cone, Adv. Math. 219 (2008), 27–62.

[BBD] A. Beilinson, J. Bernstein, and P. Deligne, Faisceaux pervers, Analyse et topologie sur les espaces singuliers, I (Luminy, 1981), Astérisque, vol. 100, Soc. Math. France, Paris, 1982, pp. 5–171.

[BL] J. Bernstein and V. Lunts, Equivariant sheaves and functors, Lecture Notes in Mathematics, vol. 1578, Springer-Verlag, Berlin, 1994.

[B] T. Braden, Hyperbolic localization of intersection cohomology, Transform. Groups 8 (2003), 209–216.

[CG] N. Chriss and V. Ginzburg, Representation theory and complex geometry, Birkhäuser Boston Inc., Boston, MA, 1997.

[DR] J. M. Douglass and G. Röhrle, Homology of the Steinberg variety and Weyl group coinvariants, Documenta Math. 14 (2009), 339–357.

[GM] M. Geck and G. Malle, On special pieces in the unipotent variety, Experiment. Math. 8 (1999), 281–290.

[L] G. Lusztig, Character sheaves. V, Adv. in Math. 61 (1986), 103–155.

[S1] T. Shoji, On the Green polynomials of a Chevalley group of type $F_4$, Comm. Algebra 10 (1982), 505–543.

[S2] T. Shoji, On the Green polynomials of classical groups, Invent. Math. 74 (1983), 239–267.

[S3] T. Shoji, Green functions associated to complex reflection groups, J. Algebra 245 (2001), 650–694.

[S4] T. Shoji, Green functions associated to complex reflection groups. II, J. Algebra 258 (2002), 563–598.

[S5] T. Shoji, Green functions attached to limit symbols, Representation theory of algebraic groups and quantum groups, Adv. Stud. Pure Math., vol. 40, Math. Soc. Japan, Tokyo, 2004, pp. 443–467.

[Sp] T. A. Springer, A purity result for fixed point varieties in flag manifolds, J. Fac. Sci. Univ. Tokyo Sect. IA Math. 31 (1984), 271–282.

Pramod N. Achar
Department of Mathematics
Louisiana State University
Baton Rouge
LA 70803
USA
pramod@math.lsu.edu
Abstract. The goal of this note is to prove, under some assumptions, a formula relating the Selmer groups of isogenous Galois representations. Local and global Euler-Poincaré characteristic formulas are key tools in the proof. With additional hypotheses, we use the isogeny formula to study how the formation of Selmer groups interacts with normalization of the coefficient ring and discuss how a main conjecture for a big Galois representation over a non-normal ring follows from a corresponding conjecture over the normalization.

2010 Mathematics Subject Classification: 11R23, 11R34
Keywords and Phrases: Iwasawa theory, Selmer group, Galois cohomology

1 Introduction

1.1. Set $G_{\mathbb{Q}} = \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ and suppose given a continuous Galois representation

$$\rho : G_{\mathbb{Q}} \to \text{Aut}_R(T),$$

where $R$ is a ring finite and free over the power series ring $\mathcal{O}[[X_1,\ldots,X_n]]$, with $\mathcal{O}$ the integer ring of a $p$-adic field, and $T$ is a finitely-generated $R$-module. One can, under suitable hypotheses, attach a Selmer group $\text{Sel}(\rho)$ to such $\rho$. This Selmer group is a finitely-generated $R$-module which is canonically defined in terms of the Galois cohomology of $\rho$.

The basic question we investigate below is the following. Given representations $\rho_1$ and $\rho_2$ as above on $R$-modules $T_1$ and $T_2$ which are isogenous, i.e., such that there is an $R[G_{\mathbb{Q}}]$-linear homomorphism $T_1 \to T_2$ with $R$-torsion cokernel, how are the Selmer groups $\text{Sel}(\rho_1)$ and $\text{Sel}(\rho_2)$ related? We prove the following...
formula relating the support divisors of \( \text{Sel}(\rho_1) \) and \( \text{Sel}(\rho_2) \) in terms of local and global invariants of the quotient \( Q = T_2/\phi(T_1) \) (see Theorem ?? for the precise statement).

**Theorem.** If \( T_1 \) and \( T_2 \) satisfy certain natural hypotheses (cf. ??), then

\[
\text{div}(\text{Sel}(\rho_1)) - \text{div}(\text{Sel}(\rho_2)) = \\
= \sum_{v \text{ real}} \text{div}(Q_{K_v}) - (r_1 + r_2) \text{div}(Q) + \sum_{v \mid p} [K_v : Q_p] \text{div}(F_v^+ Q).
\]

1.2. Our main motivation (and a key example of this type of representation) comes from Hida theory. Let \( f \) be a \( p \)-ordinary cuspidal newform. By work of Hida [?], such \( f \) belongs to a \( p \)-adic family \( \mathcal{F} \) of newforms, which can be viewed as a formal power series with coefficients in a ring \( R \) finite and free over \( \mathcal{O}[X] \), where \( \mathcal{O} \) is a suitable finite extension of \( \mathbb{Z}_p \). The specializations of \( \mathcal{F} \) at appropriate values of \( T \) are power series expansions of classical \( p \)-stabilized newforms of varying weight, level, and character. One can attach a Galois representation \( \rho_F \) to \( \mathcal{F} \) on a rank 2 module \( T \) over the ring \( R \) interpolating the \( p \)-adic Galois representation attached to the classical newforms arising as specializations of \( \mathcal{F} \). Many of the hypotheses imposed in ?? are automatically satisfied by these representations.

1.3. An early investigation of how isogenies affect Iwasawa invariants was undertaken by Schneider [?], who gave a formula relating the \( \mu \)-invariants for Selmer groups of isogenous abelian varieties over \( \mathbb{Z}_p \)-extensions of number fields. This formula was generalized by Perrin-Riou [?] to more general \( p \)-adic representations. More recently, Ochiai [?] has given a similar formula for invariants of big Galois representations with coefficients in a power series ring \( \mathbb{Z}_p[T_1, \ldots, T_n] \). Our isogeny formula is a generalization of Ochiai’s and has a similar proof, which, in particular, depends on Euler-Poincaré characteristic formulas and Poitou-Tate duality.

1.4. In Theorem ??, we prove somewhat general Euler-Poincaré characteristic formulas for big Galois representations. For \( p > 2 \), the theorem can be deduced from the corresponding statements in Nekovář [?, 4.6.9 and 7.8.6] (which exclude the case of \( p = 2 \)). Our main result, the isogeny formula of Theorem ??, follows from a series of computations involving these. Fortunately, many of the needed computations are contained in Greenberg’s series of papers [?, ?, ?]. In a certain sense, therefore, this note may be viewed as an addition to that series. Some of the results contained here can also be found in the second author’s thesis [?, Ch. 1].

1.5. Under an additional “\( p \)-criticality” assumption on the representation \( T \) (cf. ??), we show in §?? that the corresponding normalized representation \( \tilde{T} \)
obtained by extending scalars to the normalization $\tilde{R}$ of $R$ gives Selmer groups which, when considered as $R$-modules, have the same divisor on $\text{Spec } R$. Using this fact and some elementary commutative algebra, we discuss how a main conjecture for the representation $\tilde{T}$ implies a corresponding main conjecture for $T$. Thus, under our admittedly somewhat strict hypotheses, main conjectures, roughly speaking, commute with normalization. This result should not be surprising to the experts; its study was suggested by Greenberg [?, §1].

1.6. We remark here on some of the hypotheses we impose, some of which could be considered rather strong. The conditions (?1)–(?4) and the $p$-criticality hypothesis imposed in §? are somewhat standard and are known to hold for many of the representations arising “in nature” from the study of Hida families as discussed briefly above, with the possible exception of (??), which has nonetheless been extensively studied. There are two additional, less standard, hypotheses we employ.

The first of these is that the Galois modules we consider are assumed to be free over the coefficient ring. There are two places where we make serious use of this hypothesis. The first is in the application of a result of Greenberg [?, Lemma 2.2.6] on vanishing of Galois invariants. We feel that this result is probably true for even torsion-free modules. The second is in the proof of Theorem ?, where we make use of the following property of free modules $M$ over a ring $R$ with module-finite normalization $\tilde{R}$: the divisor (on $\text{Spec } R$) associated to the torsion $R$-module $(M \otimes_R \tilde{R})/M$ is $\text{rank}_R M$ times the divisor associated to $\tilde{R}/R$. It is unclear to us whether there is a weaker hypothesis on $R$-modules which guarantees this to hold.

The second is the condition (??) on the rank of compact Selmer groups, which is necessary in order to conclude the surjectivity of a certain localization map. It is a difficult and interesting question whether this condition holds for representations arising from Hida theory and is not true in general (cf. [?, §4.9] or [?, §7(d)] for an example).

2 Notation

2.1. Fix a prime $p$. Let $R$ be a complete Noetherian local domain with maximal ideal $m$ and assume that $R$ is finite and free over $\mathbb{Z}_p[T_1, \ldots, T_n]$. If $V$ is a finite-dimensional vector space over the fraction field $\text{Frac } R$ of $R$, then we call an $R$-submodule $T \subseteq V$ an $R$-lattice in $V$ if $T$ is a finitely-generated $R$-module and $T \otimes_R \text{Frac } R = V$ (where, here and subsequently, “=” means “canonically isomorphic”).

Let $K/\mathbb{Q}$ be a finite extension. Fix a finite set $\Sigma$ of primes of $K$ containing the archimedean primes and the primes lying over $p$. Denote by $K_\Sigma$ the maximal extension of $K$ unramified outside $\Sigma$ and set $G_\Sigma = \text{Gal}(K_\Sigma/K)$. Our main objects of study in what follows are Galois representations

$$\rho : G_\Sigma \to \text{GL}_n(\text{Frac } R)$$
which are continuous in the sense that the representation space \( V = V_\rho \) of \( \rho \) admits a \( G_\Sigma \)-stable \( R \)-lattice \( T \) such that the induced representation, which by abuse of notation we still denote by \( \rho : G_\Sigma \to \text{Aut}_R(T) \), is continuous for the Krull topology on \( G_\Sigma \) and the topology on \( \text{Aut}_R(T) \) induced by the topology on \( R \).

In what follows, we shall be studying free lattices, i.e., \( R \)-submodules of \((\text{Frac} R)^{\oplus n}\) of rank \( n \) which are free \( R \)-modules. Without additional assumptions on \( R \), it may not be the case that any continuous \( R \)-linear representation of \( \text{Gal}(K_\Sigma/K) \) admits a \( \text{Gal}(K_\Sigma/K) \)-stable lattice which is free as an \( R \)-module.

2.2. If \( M \) is any \( \mathbb{Z}_p \)-module, denote by \( M^\vee \) the Pontryagin dual of \( M \), i.e., \( M^\vee = \text{Hom}_{\mathbb{Z}}(M, \mathbb{Q}/\mathbb{Z}_p) \). Note that \( R \) is a compact \( \mathbb{Z}_p \)-module, so its dual \( R^\vee \) is discrete; we endow \( R^\vee \) with the trivial Galois action. If \( M \) is a cofinitely-generated, discrete \( R \)-module, then, by, e.g., Greenberg [?, Prop. 3.2], the (continuous) Galois cohomology groups \( H^i(K_\Sigma/K, M) \) are likewise cofinitely-generated \( R \)-modules.

2.3 Ordinary data. For notational convenience, we now define a notion of ordinary datum over \( R \). Such a datum \( X \) consists of a pair \( (T, F) \), where \( T \) is a finitely-generated free \( R \)-module with continuous \( G_\Sigma \)-action and \( F \) consists of \( G_K \)-submodules \( F_v^+ T \subseteq T \), one for each prime \( v \) of \( K \) lying over \( p \), such that \( F_v^+ T \) and \( F_v^- T = T/F_v^+ T \) are free \( R \)-modules. We refer to the chain \( T \supseteq F_v^+ T \supseteq 0 \) as the local filtration on \( T \) at \( v \) given by \( F \).

Given ordinary data \( X_1 = (T_1, F_1) \) and \( X_2 = (T_2, F_2) \), we define a homomorphism \( \phi : X_1 \to X_2 \) to be an \( R[G_\Sigma] \)-linear homomorphism \( \phi : T_1 \to T_2 \) which is compatible with the filtrations in the sense that \( \phi(F_v^+ T_1) \subseteq F_v^+ T_2 \) for all \( v \mid p \).

2.4. We now define discrete modules associated to a datum \( X = (T, F) \).

Denote by \( W^* = W^*_X \) the discrete Galois module \( W^* = \text{Hom}_R(T, R^\vee(1)) \) dual to \( T \). Thus, \( W^* \cong T^\vee \) as an \( R \)-module. (Note that we do not define here a compact module \( T^\vee \) with Galois action the Tate dual of that on \( T \).) The filtrations \( F \) on \( T \) induce filtrations \( W^* \supseteq F_v^+ W^* \supseteq 0 \) for \( v \mid p \) via \( F_v^+ W^* = \text{Hom}_R(F_v^- T, R^\vee(1)) \).

2.5 Local conditions. A set \( \Delta \) of local conditions for an \( R[G_\Sigma] \)-module \( M \) is a choice of submodule \( H^1_i(K_v, M) \subseteq H^1(K_v, M) \) for each \( v \in \Sigma \). Given an ordinary datum \( X = (T, F) \), we define, following Greenberg [?, §4], the Greenberg local conditions for \( W^* \) as follows: if \( v \mid p \), then set

\[
H^1_i(K_v, W^*) = H^1_{\text{ur}}(K_v, W^*) = \ker(H^1(K_v, W^*) \xrightarrow{\text{res}} H^1(I_v, W^*)),
\]

where \( I_v \subseteq G_{K_v} \) is the inertia group. For \( v \mid p \), set

\[
H^1_i(K_v, W^*) = \ker(H^1(K_v, W^*) \longrightarrow H^1(I_v, F_v^- W^*)�)
\]

\( \text{ Documenta Mathematica } 16 \ (2011) \ 885–899 \)
where the homomorphism on the right is induced by the quotient $W^* \to F_v^* W^*$ and restriction to $I_v$.

Let $X = (T, F)$ be an ordinary datum. Recall that Tate local duality gives a perfect pairing
\[ H^1(K_v, T) \times H^1(K_v, W^*) \to R^\vee. \]

We define the Greenberg local conditions $H^1_f(K_v, T)$ for $T$ as the orthogonal complements of the Greenberg local conditions $H^1_f(K_v, W^*)$ for $W^*$ under this pairing.

2.6. Given a set of local conditions $\Delta$ on an $R[\Sigma]$-module $M$, set
\[ H^1_s(K_v, M) = H^1(K_v, M)/H^1_f(K_v, M) \]
and define the Selmer group over $K$ attached to $M$ and $\Delta$ by
\[ \text{Sel}_\Delta(M) = \ker \left( H^1(G_\Sigma, M) \to \bigoplus_{v \in \Sigma} H^1_s(K_v, M) \right), \]
where the homomorphism on the right-hand side is induced by the obvious local-to-global map. If $M = W^*$ or $T$ and $\Delta$ is the set of Greenberg local conditions for $M$, then we omit the $\Delta$ and denote the corresponding Selmer group by $\text{Sel}(M)$.

For a $G_\Sigma$-module $M$ and $i \geq 0$, we further define Shafarevich-Tate groups
\[ i(M) = \ker \left( H^i(K, M) \to \bigoplus_{v \in \Sigma} H^i(K_v, M) \right). \]
Thus, $^1(M) = \text{Sel}_\Delta(M)$ for $\Delta$ the set of local conditions defined by setting $H^1_f(K_v, M) = 0$ for all $v \in \Sigma$.

The representations arising in the case of Hida families come equipped with additional structure that allows other natural definitions of local conditions (e.g., the so-called Bloch-Kato local conditions) which in general give rise to Selmer groups different from those discussed above. Ochiai has studied the relationship between these Selmer groups, cf. [7, §3].

2.7. If $M$ is a finitely-generated $R$-module and $p \subseteq R$ is a prime ideal, then we denote the $p$-length of $M$ by
\[ \lgth_p M = \lgth_{R_p} M_p, \]
which is finite if $M$ is a torsion $R$-module. A simple argument shows that
\[ \lgth_p M = \sum_{j=0}^{\infty} \rank_{R_p} p^j M/p^{j+1} M. \quad (2.7.1) \]
A finitely-generated $R$-module $M$ is said to be pseudo-null if $\lgth_p M = 0$ for every height 1 prime $p \subseteq R$. Equivalently, $M$ is pseudo-null if the set $\text{Ass}_R(M)$
of associated primes of $M$ contains only primes of height 2 or greater. If $M$ is cofinitely-generated, we say $M$ is copseudo-null if $M^\vee$ is pseudo-null.

If $R$ has dimension 2 and finite residue field, then a finitely-generated, resp. cofinitely-generated, $R$-module is pseudo-null, resp. copseudo-null, if and only if it contains only finitely many elements.

2.8 Conditions on $X$. Fix an ordinary datum $X = (T, F)$. Below, we often subject $X$ to the following conditions.

(2.8.1) $T(-1)_{G_K}$ is a pseudo-null $R$-module.

(2.8.2) For each prime $v | p$ of $K$, $(F_v^+ T)(I_v) = 0$, $(F_v^+ T(-1))(K_v) = 0$, $(F_v^+ T)(K_v) = 0$, and $(F_v^- W^*)(K_v)$ is copseudo-null over $R$.

(2.8.3) $\operatorname{Sel}(W^*)$ is a cotorsion $R$-module.

(2.8.4) No subquotient of $W^*[m]$ is isomorphic to $\mu_p$ as a $G_K$-module.

(2.8.5) For all $v \in \Sigma$ with $v \not\mid p \infty$, $T(K_v) = 0$ and $W^*(K_v)$ is copseudo-null over $R$.

(2.8.6) $\operatorname{rank}_R \operatorname{Sel}(T) = 0$.

Note that $T(-1)_{G_K} \cong W^*(K)^\vee$, so (2.8.1) is equivalent to the statement that $W^*(K)$ is copseudo-null. A similar remark applies to (2.8.2) and the modules $(F_v^- W^*)(K_v)^\vee \cong F_v^+ T(-1)_{G_K_v}$. Condition (2.8.4) implies $T(K) = 0$. Indeed, $(T/mT) = \operatorname{Hom}_{Z_p}(W^*, \mu_{p\infty}) = 0$ under this assumption, so $T(K)/mT(K) = 0$ as well. As mentioned in the introduction, one cannot expect (2.8.6) to hold in general. As discussed, e.g., in [? , §4.9] or [? , §7(d)], there are interesting representations arising from Hida theory for which it should hold and for which it should not hold. In the context of those examples, though not generally, (2.8.1) and (2.8.4) should be equivalent.

3 Duality formulas

3.1. This section is devoted to the proof of various duality results for Selmer groups. The first several subsections (up to 3.3) are devoted to the proof of the following theorem, the global (3.1) and local (3.2) Euler-Poincaré characteristic formulas. This theorem can be deduced from Nekovář [?, 4.6.9 and 7.8.6], at least in the case $p > 2$.

3.2 Theorem. Suppose $K$ has $r_1$ real places and $r_2$ conjugate pairs of complex places. For any cofinitely-generated cotorsion $R$-module $D$ and height 1 prime $p \subseteq R$,

$$
\sum_{i=0}^2 (-1)^i \lgth_p \operatorname{H}^i(G_{\Sigma}, D)^\vee = \sum_{v \text{ real}} \lgth_p D(K_v)^\vee - (r_1 + r_2) \lgth_p D^\vee
$$

(3.2.1)
and, for every non-archimedean prime \( v \) of \( K \),
\[
\sum_{i=0}^{2} (-1)^i \text{lgth}_p H^i(K_v, D)^\vee = \begin{cases} 
0 & \text{if } v \nmid p \\
-[K_v : \mathbb{Q}_p] \text{lgth}_p D^\vee & \text{if } v \mid p
\end{cases}
\tag{3.2.2}
\]

3.3. Define, for any cofinitely-generated cotorsion \( R \)-module \( D \),
\[
\delta_*(-) = \sum_{i=0}^{2} (-1)^i \text{lgth}_p H^i(G, D)^\vee \]
and, for every non-archimedean prime \( v \) of \( K \),
\[
\sum_{i=0}^{2} (-1)^i \text{lgth}_p H^i(K_v, D)^\vee = \begin{cases} 
0 & \text{if } v \nmid p \\
-[K_v : \mathbb{Q}_p] \text{lgth}_p D^\vee & \text{if } v \mid p
\end{cases}
\]

3.4. The proof of Theorem 3.3 proceeds by induction on \( \text{lgth}_p D^\vee \) and dévissage.
The base case is the following.

LEMMA. If \( \text{lgth}_p p D^\vee = 0 \), then \( \delta_*(D) = 0 \) for \( * = \Sigma \) or a prime of \( K \).

Proof. Consider the short exact sequence
\[
0 \rightarrow D[p] \rightarrow D \rightarrow D/D[p] \rightarrow 0.
\]
By hypothesis, we have \( \text{lgth}_p D/D[p])^\vee = 0 \), so \( p \notin \text{Supp}(D/D[p])^\vee \). As \( \text{Supp} M^\vee \supset \text{Supp} H^i(G, M)^\vee \) for any \( i \geq 0 \), any cofinitely-generated \( R \)-module \( M \), and \( G = G_\Sigma \) or \( G_{K_v} \), we see from the definition of \( \delta_* \) that \( \delta_*(D) = \delta_*(D[p]) \).

3.5. Lemma (Dévissage). For any short exact sequence
\[
0 \rightarrow A \rightarrow B \rightarrow C \rightarrow 0
\]
of cofinitely-generated \( R \)-modules with \( G \)-action for \( G = G_{K_v} \), resp. \( G = G_\Sigma \),
we have \( \delta_*(A) - \delta_*(B) + \delta_*(C) = 0 \) for \( * = \Sigma \) or a prime of \( K \).
Proof. As \( \text{lgth}_p \) is additive in exact sequences, we may ignore the terms in the
definition of \( \delta_*(D) \) which are multiples of \( \text{lgth}_p D^\vee \). The lemma is slightly more
difficult when \( p = 2 \) and \( * = \Sigma \), so let us first assume either \( p > 2 \) or \( * \neq \Sigma \).
If \( v \) is an archimedean prime of \( K \), then, as \( p > 2 \), \( H^i(K_v, D) = 0 \) for \( i > 1 \)
and \( D = A, B, \) or \( C \). The result then follows from the long exact cohomology
sequence and the fact that \( G_\Sigma \) and \( G_{K_v} \) have \( p \)-cohomological dimension 2
under our assumptions.
Now suppose \( p = 2 \) and \( * = \Sigma \). By the long exact
\( G_\Sigma \)-cohomology sequence,
we have

\[
2 \sum_{i=0}^2 \text{lgth}_p H^i(G_\Sigma, A) - 2 \sum_{i=0}^2 \text{lgth}_p H^i(G_\Sigma, B) + 2 \sum_{i=0}^2 \text{lgth}_p H^i(G_\Sigma, C) =
\text{lgth}_p \ker[H^3(G_\Sigma, A) \rightarrow H^3(G_\Sigma, B)]. \tag{3.5.1}
\]
Recall (\cite{?, Thm. 4.10}, e.g.) that for any discrete ind-finite \( R[G_\Sigma] \)-module
\( D \), the natural map

\[
H^q(G_\Sigma, D) \rightarrow \prod_v \text{H}^q(K_v, D)
\]
given by the product of restrictions to decomposition groups at real places is
an isomorphism for \( q \geq 3 \), so the right-hand side of (3.5) is equal to

\[
\text{lgth}_p \prod_v \ker[H^3(K_v, A) \rightarrow H^3(K_v, B)].
\]

As \( G_{K_v} \) is cyclic of order 2 for \( v \) a real place, the cohomology groups \( H^i(K_v, D) \)
are periodic of period 2 for \( i > 0 \), and all have equal \( p \)-length (cf. \cite{?, Prop. 4.18}).
This implies that

\[
\text{lgth}_p \ker[H^3(K_v, A) \rightarrow H^3(K_v, B)] = \text{lgth}_p \ker[H^3(K_v, A) \rightarrow H^3(K_v, B)]
\]
for real \( v \), which, by the long exact \( G_{K_v} \)-cohomology sequences, shows that the
right hand side of (3.5) is equal to

\[
\sum_{v \text{ real}} (\text{lgth}_p A(K_v) - \text{lgth}_p B(K_v) + \text{lgth}_p C(K_v)),
\]
which proves the lemma for \( p = 2 \) and \( * = \Sigma \).

3.6 Proof of Theorem 3.5. The statement is true when \( \text{lgth}_p D^\vee = 0 \) by
Lemma 3.4, so assume \( \text{lgth}_p D^\vee > 0 \). Consider the short exact sequence

\[
0 \rightarrow p D^\vee \rightarrow D^\vee \rightarrow D^\vee/p D^\vee \rightarrow 0.
\]
Lemma 3.4 implies the result if \( \text{lgth}_p p D^\vee = 0 \). Similarly, if \( \text{lgth}_p D^\vee/p D^\vee = 0 \),
then \( \text{lgth}_p D^\vee = 0 \) by Nakayama’s Lemma, so a fortiori \( \text{lgth}_p \) \( D^\vee \) = 0 and we
are again done by Lemma 3.4. We may therefore assume that both \( \text{lgth}_p p D^\vee \)
and \( \text{lgth}_p D^\vee/p D^\vee \) are positive and thus less than \( \text{lgth}_p D^\vee \). The theorem then
follows from dèvissage (Lemma 3.4) and induction.
3.7 Theorem (Poitou-Tate global duality). There is a perfect pairing

\[ (W^*) \times 2(T) \to \mathbb{Q}_p/\mathbb{Z}_p \]

and a 9-term exact sequence

\[
0 \to H^0(G_\Sigma, W^*) \to \bigoplus_{v \in \Sigma} H^0(K_v, W^*) \to H^2(G_\Sigma, T)^\vee \to \\
H^1(G_\Sigma, W^*) \to \bigoplus_{v \in \Sigma} H^1(K_v, W^*) \to H^1(G_\Sigma, T)^\vee \to \\
H^2(G_\Sigma, W^*) \to \bigoplus_{v \in \Sigma} H^2(K_v, W^*) \to H^0(G_\Sigma, T)^\vee \to 0
\]

Proof. For all \( n \), \( R/m^n \) is finite. Note that \( W^* = \varprojlim W^*[m^n] \), and \( T = \varprojlim T/m^nT \). As \( W^*[m^n] \cong \text{Hom}_{\mathbb{Z}_p}(T/m^nT, \mathbb{Q}_p/\mathbb{Z}_p(1)) \), the theorem follows from the version for finite modules (see [?], Thm. I.4.10, for example) by taking limits. 

3.8. For a \( G_\Sigma \)-module \( M \) with local filtrations at each \( v \mid p \), e.g., for \( M \) arising from an ordinary datum, we define semi-local cohomology groups by

\[
H^i_{\text{loc}}(M) = \bigoplus_{v \mid p} H^i(K_v, F_{vM}) \oplus \bigoplus_{v \in \Sigma \setminus \{p\}} H^i(K_v, M).
\]

Additionally, let

\[
\text{loc}_M^i : H^i(G_\Sigma, M) \to H^i_{\text{loc}}(M)
\]

be the natural localization map.

3.9 Lemma. If \( X \) satisfies (??) and (??), then the natural homomorphism

\[
H^1(G_\Sigma, W^*) \to \bigoplus_{v \in \Sigma} H^1_{v}(K_v, W^*)
\]

is surjective.

Proof. Consider the exact sequence arising from local duality and the definitions of the various groups involved:

\[
H^1(G_\Sigma, W^*) \to \bigoplus_{v \in \Sigma} H^1_v(K_v, W^*) \to \text{Sel}(T)^\vee \to 1(T)^\vee \to 0.
\]

Under (??), [?], Prop. 2.2.1] states that \( H^1(G_\Sigma, T) \) is \( \Lambda \)-torsion-free, whence \( R \)-torsion-free, so the same is true of \( \text{Sel}(T) \). The lemma then follows from the assumption (??).
3.10 Lemma. If $X$ satisfies (??) and (??), then $\text{Sel}(W^*)/\ker\text{loc}_{1 W^*}$ is a copseudo-null $R$-module. In particular,

$$\text{lgth}_p \text{Sel}(W^*) = \text{lgth}_p \ker(\text{loc}_{1 W^*})$$

for every height 1 prime $p \subseteq R$. If $X$ further satisfies (??) and (??), then $\text{coker}\text{loc}_{1 W^*}$ is a copseudo-null $R$-module.

Proof. The inflation-restriction sequence for $I_v \subseteq G_K$ implies that the quotient $\text{Sel}(W^*)/\ker\text{loc}_{1 W^*}$ injects into

$$\bigoplus_{v | p} H^1(K_v/K_v, F_v^{- W^*(I_v)}) \oplus \bigoplus_{v \in \Sigma, v \nmid p_{\infty}} H^1(K_v/K_v, W^*(I_v)),$$

(3.10.1)

where $K_v^{ur}$ is the maximal unramified extension of $K_v$. The lemma thus follows from the assumptions (??) and (??), which state that $F_v^{- W^*(I_v)}$ and $W^*(I_v)$ are copseudo-null $R$-modules.

In case $X$ also satisfies (??) and (??), then Lemma ?? gives that the homomorphism

$$H^1(G_{\Sigma}, W^*) \longrightarrow \bigoplus_{v \in \Sigma} H^1(K_v, W^*)$$

defining $\text{Sel}(W^*)$ is surjective. The module (??) above is the kernel of the quotient map

$$H^1_{\text{loc}}(W^*) \longrightarrow \bigoplus_{v \in \Sigma} H^1_{\text{loc}}(K_v, W^*),$$

so the final statement in the lemma follows from the fact that (??) is copseudo-null.

3.11 Lemma. If $X$ satisfies (??), (??), and (??), then the $R$-modules $H^2_{\text{loc}}(W^*)$ and $H^2(G_{\Sigma}, W^*)$ are trivial.

Proof. By local Tate duality, we have

$$H^2(K_v, F_v^{- W^*}) \cong (F_v^{+ T})(K_v)$$

for $v \mid p$, and

$$H^2(K_v, W^*) \cong T(K_v)$$

for $v \in \Sigma, v \nmid p_{\infty}$. Both of these are trivial by (??) and (??), respectively, so that $H^2_{\text{loc}}(K_v, W^*) = 0$.

We first show $\hat{2}(W^*) = 0$. By [??, Prop. 6.6], $\hat{2}(W^*)$ is coreflexive, so it suffices to show that $\text{corank}_R \hat{2}(W^*) = 0$. By [??, Prop. 4.4], $\hat{2}(W^*)$ has the same $R$-corank as $\hat{1}(W^*)$. On the other hand, $\hat{1}(W^*) \subseteq \text{Sel}(W^*)$, which is assumed $R$-cotorsion by (??). By definition of , we have an exact sequence

$$0 \longrightarrow \hat{2}(W^*) \longrightarrow H^2(G_{\Sigma}, W^*) \longrightarrow \bigoplus_{v \in \Sigma} H^2(K_v, W^*).$$

We have just seen that $\hat{2}(W^*) = 0$, and $H^2(K_v, W^*) = 0$ for $v \in \Sigma$ by (??), (??), and local duality.
4 ISOGENIES

4.1. If $T_1$ and $T_2$ are finitely generated $R$-modules and $\phi: X_1 \to X_2$ is a homomorphism with torsion kernel and cokernel, then we say that $\phi$ is an \textit{isogeny} or that $T_1$ and $T_2$ are \textit{isogenous} if we do not wish to make the homomorphism explicit. Note that an isogeny of torsion-free $R$-modules is necessarily injective. Similarly, if $W_1$ and $W_2$ are cofinitely-generated $R$-modules, then we say that a homomorphism $\psi: W_1 \to W_2$ is an isogeny (and that $W_1$ and $W_2$ are isogenous) if its Pontryagin dual $\psi^\vee: W_2^\vee \to W_1^\vee$ is an isogeny. If $X_i = (T_i, F_i)$, $i = 1, 2$, are ordinary data, then we say that $\phi: X_1 \to X_2$ is an isogeny and hence that the $X_i$ are isogenous if $\phi: T_1 \to T_2$ is an isogeny. A homomorphism of ordinary data $\phi: X_1 \to X_2$ is an isogeny if and only if the induced homomorphism $W_1^\vee \to W_2^\vee$ is an isogeny. Isogeny is an equivalence relation on the categories of finitely-generated $R$-modules, cofinitely-generated $R$-modules, and ordinary data, cf. [?, §2].

4.2. For the remainder of the section, fix ordinary data $X_1 = (T_1, F_1)$ and $X_2 = (T_2, F_2)$ and an isogeny $\phi: X_1 \to X_2$. Our goal is to use Theorem ?? to prove a formula (Theorem ??) relating the $p$-lengths of Selmer groups for $X_1$ and $X_2$ in terms of various Galois invariants of the quotient module $T_2/\phi(T_1)$, or, more precisely, its dual $C = C_\phi = \ker[W_2^\vee \to W_1^\vee]$. The key tools we need are the global Euler-Poincaré characteristic formulas above and Poitou-Tate duality, Theorem ??'. The formula can be thought of as a reorganization of the information provided by Poitou-Tate duality under the assumptions (??)–(??).

4.3 Proposition. If $X$ satisfies (??)–(??), then for all height 1 primes $p \subseteq R$,

\[
\text{lgth}_p \text{Sel}(W_1)^\vee - \text{lgth}_p \text{Sel}(W_2)^\vee = \\
= \sum_{i=1}^{2} (-1)^i (\text{lgth}_p H^1(G_\Sigma, C)^\vee - \text{lgth}_p H^1_{\text{loc}}(C)^\vee)
\]

\textbf{Proof.} The commutative diagram

\[
\begin{array}{cccccc}
H^1(G_\Sigma, C) & \xrightarrow{\alpha} & H^1(G_\Sigma, W_1^\vee) & \xrightarrow{\gamma} & H^2(G_\Sigma, C) \\
| & | & | & | & | \\
H^1_{\text{loc}}(C) & \xrightarrow{\beta} & H^1_{\text{loc}}(W_1^\vee) & \xrightarrow{\delta} & H^2_{\text{loc}}(C) \\
\end{array}
\]

has exact rows. Assumptions (??), resp. (??) and (??), imply that $\ker \alpha$, resp. $\ker \beta$, is copseudo-null over $R$. Likewise, $\gamma$ and $\delta$ have trivial cokernel by Lemma ??'. By Lemma ??, $\text{loc}^{1}_{W_1^\vee}$ and $\text{loc}^{1}_{W_2^\vee}$ have copseudo-null cokernels. Examining the $p$-lengths in the above diagram for a height 1 prime $p \subseteq R$.
therefore gives
\[ \text{lgth}_p \ker(\text{loc}_{W_1})^\vee - \text{lgth}_p \ker(\text{loc}_{W_2})^\vee = \]
\[ = \sum_{i=1}^{2} (-1)^i (\text{lgth}_p \text{H}^0(G_{\Sigma}, C)^\vee - \text{lgth}_p \text{H}_{\text{loc}}^1(C)^\vee), \]
which implies the proposition by the first statement of Lemma ??.

4.4 Theorem. If \( X_1 \) and \( X_2 \) satisfy (??)–(??), then, for every height 1 prime \( p \subseteq R \),
\[ \text{lgth}_p \text{Sel}(W_1^\ast) = \text{lgth}_p \text{Sel}(W_2^\ast) \]
\[ = \sum_{v \text{ real}} \text{lgth}_p C(K_v)^\vee - (r_1 + r_2) \text{lgth}_p C^\vee + \sum_{v/p} [K_v : Q_p] \text{lgth}_p (F_v^\ast C)^\vee. \]

Proof. By (??), \( H^0(G_{\Sigma}, W_1^\ast) \) is copseudo-null, so \( H^0(G_{\Sigma}, C) \) and \( H^1_{\text{loc}}(W_1^\ast) \) are also copseudo-null. The theorem thus follows immediately from Proposition ?? and the global Euler-Poincaré characteristic formula, Theorem ??.

5 Application to normalization
5.1. We now apply the main result of §?? to study how Selmer groups behave with respect to normalization. Assume \( R \) is reduced and let \( \tilde{R} \) be the integral closure of \( R \) in its total ring of fractions. A well-known result of Nagata [?, Thm. 7] states that \( \tilde{R} \) is a finite \( R \)-module. If \( X \) is an ordinary datum over \( R \), then set \( \tilde{X} = (\tilde{T}, \tilde{F}) \), where \( \tilde{T} = T \otimes_R \tilde{R} \) and \( \tilde{F}_v^\ast T = (F_v^\ast T) \otimes_R \tilde{R} \). Since \( \tilde{R} \) is finite over \( R \), we may view \( \tilde{X} \) as an ordinary datum over \( R \) or over \( \tilde{R} \), and the natural inclusion \( T \rightarrow \tilde{T} \) is an isogeny of ordinary data over \( R \).

5.2. Fix an ordinary datum \( X \) over \( R \). For \( \Phi = \text{Frac} \ R \) the fraction field of \( R \), define \( V = T \otimes_R \Phi \), so \( V \) is a finite-dimensional \( \Phi \)-vector space with a \( \Phi \)-linear action of \( G_{\Sigma} \). The filtrations \( F \) induce filtrations \( V \supseteq F \supseteq 0 \) for \( v \mid p \). Define
\[ \alpha(X) = \text{dim}_\Phi (\text{res}_{K/Q} V)^+ = \sum_{v \mid \infty} \text{dim}_\Phi V(K_v). \]
For \( v \mid p \), define \( \varepsilon_v(X) = \text{dim}_K F_v^+ V \). We say that \( X \) is \( p \)-critical if \( \alpha(X) = \sum_{v/p} \varepsilon_v(x) \). For \( p \)-critical data, we have the following theorem regarding normalization.

5.3 Theorem. Let \( \mathfrak{p} \subseteq R \) be a height 1 prime. If \( p = 2, \) then assume that \( T(K_v) \) is a summand of \( T \) for each real place \( v \) of \( K \). If \( X \) and \( \tilde{X} \), both viewed as ordinary data over \( R \), are \( p \)-critical and satisfy (??)–(??), then
\[ \text{lgth}_p \text{Sel}(W^\ast)^\vee = \text{lgth}_p \text{Sel}(\tilde{W}^\ast)^\vee. \]
Proof. Let \( C = \ker[\tilde{W}^* \to W^*] \), the map being induced by the inclusion \( T \hookrightarrow \tilde{T} \). By (??), (??) and (??), we have that \( H^0(G_\Sigma, C) \) and \( H^0_{\text{loc}}(C) \) are copseudo-null over \( R \), so it suffices by Proposition ?? to show that
\[
\sum_{i=0}^{2} (-1)^i \text{lgth}_p H^i(G_\Sigma, C)^\vee = \sum_{i=0}^{2} (-1)^i H^i_{\text{loc}}(C)^\vee \tag{5.3.1}
\]
for all \( p \subseteq R \) of height 1. The global Euler-Poincaré formula (??) gives the left-hand side of (??) as
\[
\sum_{v \text{ real}} \text{lgth}_p C(K_v)^\vee - (r_1 + r_2) \text{lgth}_p C^\vee
\]
and the local formula (??) gives the right-hand side as
\[
\sum_{v \mid p} [K_v : Q_p] \text{lgth}_p (F_v^+ C)^\vee.
\]
Adding \([K : Q] \text{lgth}_p C^\vee\) to these yields
\[
\sum_{v \text{ real}} \text{lgth}_p C(K_v)^\vee + r_2 \text{lgth}_p C^\vee
\]
and
\[
\sum_{v \mid p} [K_v : Q_p] \text{lgth}_p (F_v^+ C)^\vee.
\]
By freeness of \( T \), \( \text{lgth}_p C^\vee = \text{rank}_R T \text{lgth}_p \tilde{R}/R \), and similarly for \((F_v^+ C)^\vee\) and \(C(K_v)\). The theorem thus follows from the \( p \)-criticality assumption on \( T \). 

5.4 Lemma. Let \( M \) and \( N \) be torsion \( \tilde{R} \)-modules and fix a height 1 prime \( q \subseteq R \). Then \( \text{lgth}_p M = \text{lgth}_p N \) for all height 1 primes \( p \subseteq \tilde{R} \) such that \( p \mid q \) if and only if \( \text{lgth}_q M = \text{lgth}_q N \).

Proof. The content of the lemma is that the \( q \)-length of a torsion \( \tilde{R} \)-module \( M \) (viewed as \( R \)-module) is determined by its \( p \)-lengths (viewed as \( \tilde{R} \)-module) for \( p \subseteq \tilde{R} \) lying over \( q \), and conversely. Let \( p_1, \ldots, p_n \) be the primes of \( \tilde{R} \) lying over \( q \subseteq R \) and set \( S = (R - q) \subseteq \tilde{R} \). First consider a chain
\[
M_q = M_{q,0} \supseteq M_{q,1} \supseteq \cdots \supseteq M_{q,n}
\]
computing the length of \( M_q \) as \( S^{-1} \tilde{R} \)-module. Each successive quotient in this chain is isomorphic to \( \tilde{R}/p_i \) for some \( i \). Localization of this chain to \( \tilde{R}_{p_i} \), therefore computes \( \text{lgth}_{p_i} M \) after removing repeated submodules and we see that \( \text{lgth}_q M \) determines and is determined by these lengths and \( \text{lgth}_q \tilde{R}/p_i \).
5.5. In the below corollary to Theorem 5.5., we say two finitely-generated \( R \)-modules \( M \) and \( N \) have the same divisor if \( \text{lgth}_p M = \text{lgth}_p N \) for all height 1 primes \( p \subseteq R \). Similarly, we say a finitely-generated \( R \)-module has the same divisor as an element \( L \in R \) if \( M \) and \( R/L \) have the same divisor.

**Corollary.** Let \( 0 \neq L \in R \) and let \( \tilde{L} \) be the image of \( L \) in \( \tilde{R} \). Using notation and assumptions as in Theorem 5.5., with the exception that we now view \( \tilde{X} \) as an ordinary datum over \( R \), \( \text{Sel}(\tilde{W}^*)^\vee \) has the same divisor as \( L \) if and only if \( \text{Sel}(\tilde{W}^*)^\vee \) has the same divisor as \( \tilde{L} \).

**Proof.** Viewing \( \tilde{R} \) as a rank 1 \( R \)-module, we use the formula [?, Lemma 11.7] to see that, for every height 1 prime \( q \subseteq R \),

\[
\text{lgth}_q \tilde{R}/(\tilde{L}) = \text{lgth}_q R/(L),
\]

so the result follows by combining Theorem 5.5. and Lemma ??.

5.6. Corollary 5.5. states, roughly speaking, that, under some assumptions, the formation of the divisor of the Selmer group of an ordinary datum commutes with normalization. In a situation where there is a \( p \)-adic \( L \)-function belonging to \( R \) associated with the ordinary datum \( X \), the corollary provides some flexibility in proving a main conjecture for \( X \), in that such a conjecture can be proved equivalently before or after normalization.

**References**

[1] T. Arnold, *Hida families, \( p \)-adic heights, and derivatives*, Ann. Inst. Fourier 61 (2011), to appear.

[2] R. Greenberg, *Iwasawa theory and \( p \)-adic deformations of motives*, Motives (Seattle, WA, 1991), Proc. Sympos. Pure Math., vol. 55, Amer. Math. Soc., Providence, RI, 1994, pp. 193–223.

[3] ______, *On the structure of certain Galois cohomology groups*, Doc. Math. (2006), Extra Vol., 335–391 (electronic).

[4] ______, *On the structure of Selmer groups*, preprint (2010).

[5] ______, *Surjectivity of the global-to-local map defining a Selmer group*, Kyoto J. Math. 50 (2011), 853–888.

[6] H. Hida, *Iwasawa modules attached to congruences of cusp forms*, Ann. Sci. École Norm. Sup. (4) 19 (1986), no. 2, 231–273.

[7] ______, *Modular forms and Galois cohomology*, Cambridge Studies in Advanced Mathematics, no. 69, Cambridge Univ. Press, Cambridge, 2000.
[8] K. Koo, *Change of Selmer group under isogeny, Iwasawa invariants of $\Lambda$-adic representation, and coefficient field of newforms*, Ph.D. thesis (2009), University of Washington.

[9] H. Matsumura, *Commutative ring theory*, second ed., Cambridge Studies in Advanced Mathematics, vol. 8, Cambridge University Press, 1989.

[10] J. Milne, *Arithmetic duality theorems*, Perspectives in Mathematics, vol. 1, Academic Press, 1986.

[11] M. Nagata, *Some remarks on local rings. II*, Mem. Coll. Sci. Univ. Kyoto. Ser. A. Math. 28 (1954), 109–120.

[12] J. Nekovář, *Selmer complexes*, Astérisque (2005), no. 310.

[13] T. Ochiai, *On the two-variable Iwasawa main conjecture*, Compositio Math. 142 (2006), 1157–1200.

[14] T. Ochiai, *The algebraic $p$-adic $L$-function and isogeny between families of Galois representations*, J. Pure Appl. Algebra 212 (2008), no. 6, 1381–1393.

[15] B. Perrin-Riou, *Variation de la fonction $L_p$-adique par isogénie*, Algebraic number theory, Adv. Stud. Pure Math., vol. 17, Academic Press, Boston, MA, 1989, pp. 347–358.

[16] P. Schneider, *The $\mu$-invariant of isogenies*, J. Indian Math. Soc. (N.S.) 52 (1987), 159–170 (1988).
Abstract. Suppose that \( G \) is a finite \( \pi \)-separable group. A classical result asserts that all irreducible characters of a Hall \( \pi \)-subgroup \( H \) of \( G \) extend to \( G \) if and only if \( H \) has a normal complement in \( G \).

Now, we fix a prime \( p \) and analyze when only the \( p' \)-degree irreducible characters of \( H \) extend to \( G \).

2010 Mathematics Subject Classification: 20C15 (primary), 20G40 (secondary)
Keywords and Phrases: extensions of characters, Hall subgroups, Sylow normalizers, \( \pi \)-separable groups

1. Introduction

We come back to an old result of C. Sah ([18]) that asserts that in a finite \( \pi \)-separable group, all irreducible complex characters of a Hall \( \pi \)-subgroup \( H \) of \( G \) extend to \( G \) if and only if \( G \) has a normal \( \pi \)-complement. Now, we fix a prime \( p \) and we wish to characterize when only the \( p' \)-degree characters of \( H \) extend to \( G \).

Theorem A. Let \( G \) be a finite \( \pi \)-separable group. Let \( H \) be a Hall \( \pi \)-subgroup of \( G \), let \( K \) be a \( \pi \)-complement of \( G \), and let \( p \) be a prime. Then every \( \alpha \in \text{Irr}(H) \) of \( p' \)-degree extends to \( G \) if and only if there is \( P \in \text{Syl}_p(H) \) such that \( \text{N}_G(P) \subseteq \text{N}_G(K) \).

Of course, Theorem A is far more general than Sah’s theorem, although we pay the price of using the Classification of Finite Simple Groups. This is not that surprising, however: in the case where \( H \) is normal in \( G \), Theorem A is equivalent to proving a well-known consequence of the (yet unproven) McKay conjecture (see [17, Thm. C]) which therefore now becomes established.

As an easy consequence of our main result we obtain:

1. The research of the second author is partially supported by the Spanish Ministerio de Educación y Ciencia proyecto MTM2010-15296, Programa de Movilidad, and Prometeo/Generalitat Valenciana.
Corollary B. Suppose that $A$ acts coprimely on a finite group $G$, and let $P \in \text{Syl}_p(G)$ be $A$-invariant. Then all $p'$-degree irreducible characters of $G$ are $A$-invariant if and only if $[N_G(P), A] = 1$.

The paper is split into two parts. In Section 2 we prove Theorem A and Corollary B modulo a statement (Theorem 2.2) on finite quasi-simple groups which is then shown in Section 3, using the classification, properties of algebraic groups and Deligne–Lusztig theory.

2. Proof of Theorem A

In our first result we use the Gajendragadkar’s $\pi$-special characters (whose main properties can be found in [3]).

Lemma 2.1. Suppose that $G$ is $\pi$-separable and let $H$ be a Hall $\pi$-subgroup of $G$. Let $L \triangleleft G$. Suppose that $\alpha \in \text{Irr}(H)$ extends to $G$ and is such that $H \cap L \subseteq \ker \alpha$. Then there is an extension $\beta \in \text{Irr}(G)$ such that $L \subseteq \ker \beta$. In particular, if all $p'$-degree irreducible characters of $H$ extend to $G$, then all $p'$-degree irreducible characters of $H/L$ extend to $G/L$.

Proof. Let $\tilde{\alpha} \in \text{Irr}(HL)$ be the unique irreducible character of $HL$ that extends $\alpha$ and has $L$ in its kernel. Notice that, by using the definition, $\tilde{\alpha}$ is $\pi$-special. By hypothesis, $\alpha$ extends to $G$. By [6, Thm. F], $\alpha$ has a $\pi$-special extension $\hat{\alpha}$ to $G$. Now by [3, Prop. (6.1)], we have that $\hat{\alpha}_H$ is a $\pi$-special extension of $\alpha_H$ to $HL$. But, again by the uniqueness part in loc.cit., we have that $\hat{\alpha}_H = \tilde{\alpha}$, and we are done. □

In order to prove Theorem A, we shall need the following non-trivial result whose proof (that uses the Classification of Finite Simple Groups) we defer until Section 3 below.

Theorem 2.2. Let $p$ be a prime. Suppose that $M \triangleleft G$ has $p$-power index, and has a normal Hall subgroup $S$ such that $M/S$ is not divisible by $p$. Assume that $S$ is a non-abelian quasi-simple group of order divisible by $p$ with $C_G(S) = Z(S)$ a $p'$-group. Let $P \in \text{Syl}_p(G)$. Then the following are equivalent:

(i) all $p'$-degree $P$-invariant irreducible characters of $S$ are $M$-invariant, and

(ii) there exists a complement $K$ of $S$ in $M$ normalized by $P$ and with $[N_S(P), K] = 1$.

We shall frequently write condition (ii) of the above theorem in the following more convenient form.

Lemma 2.3. Let $S \triangleleft M$. Suppose that $P, K$ are subgroups of $M$ with $S \cap K = 1$. Then $P$ normalizes $K$ and $[N_S(P), K] = 1$ if and only if $N_{SP}(P) \subseteq N_M(K)$.

Proof. Since $P \subseteq N_{SP}(P)$, by Dedekind’s lemma we have that $N_{SP}(P) = N_S(P)P$. Thus, if $P$ normalizes $K$ and $[N_S(P), K] = 1$ it is clear that $N_{SP}(P) \subseteq N_M(K)$. Conversely, we have that $P$ and $N_S(P)$ normalize $K$. Then, using that $S \triangleleft M$, we have that $[N_S(P), K] \subseteq S \cap K = 1$. □
Also, the group theoretical conclusion in Theorem A has another convenient form.

**Lemma 2.4.** Suppose that $G$ is a $\pi$-separable group with a Hall $\pi$-subgroup $H$ and a Hall $\pi$-complement $K$. Let $P \in \text{Syl}_p(H)$. If $P$ normalizes $K$, then $N_G(P) = N_H(P)N_K(P)$. In particular, $N_G(P) \subseteq N_G(K)$ if and only if $N_H(P) \subseteq N_G(K)$.

**Proof.** This is an application of Lemma (2.1) of [10]. □

**Lemma 2.5.** Suppose that $P$ is a $p$-group acting on a group $S$ that does not have a normal $p$-complement. Then there exists $1 \neq \chi \in \text{Irr}(S)$ which is $P$-invariant of $p'$-degree.

**Proof.** The semidirect product $SP$ cannot have a normal $p$-complement. By Thompson’s Theorem (Corollary (12.2) of [5]), there exists a non-linear $\gamma \in \text{Irr}(SP)$ of degree not divisible by $p$. Now, $\gamma_S = \chi$ is irreducible (by Corollary (11.29) of [5]) and $P$-invariant. It is clear that $\chi$ is not trivial since otherwise $\gamma$ would be a $p'$-degree character of a $p$-group, while $\gamma$ is not linear. □

The following is one direction of Theorem A.

**Theorem 2.6.** Let $G$ be a finite $\pi$-separable group. Let $H$ be a Hall $\pi$-subgroup of $G$, let $K$ be a $\pi$-complement of $G$, and let $p$ be a prime. Suppose that every $\alpha \in \text{Irr}(H)$ of $p'$-degree extends to $G$. Then there is $P \in \text{Syl}_p(H)$ such that $N_G(P) \subseteq N_G(K)$.

**Proof.** If $p$ does not divide $|H|$, then all irreducible characters of $H$ extend to $G$, and $G$ has a normal $\pi$-complement by Sah’s Theorem (see [18, Thm. 5]). So we may assume that $p \in \pi$.

We prove that there exists $P \in \text{Syl}_p(H)$ such that $N_G(P) \subseteq N_G(K)$ by induction on $|G|$. Since $H$ is a Hall subgroup of $G$, notice that every Sylow $p$-subgroup of $H$ is a Sylow $p$-subgroup of $G$.

Let $N$ be a fixed but arbitrary minimal normal subgroup of $G$.

**Step 1.** We can assume that there is $P \in \text{Syl}_p(H)$ such that $N_G(P) \subseteq N_G(K)N$. Also, $N \subseteq H$. In particular, $O_{\pi'}(G) = 1$.

By Lemma 2.1, we know that all $p'$-degree irreducible characters of $HN/N$ extend to $G/N$. Therefore, by induction and using that $N_G(KN) = N_G(K)N$, we conclude that there is $P \in \text{Syl}_p(H)$ such that $N_G(P) \subseteq N_G(K)N$. Now, since $G$ is $\pi$-separable, we have that either $N$ is a $\pi$-group or a $\pi'$-group. In the second case, $N \subseteq K$, and $N_G(P) \subseteq N_G(K)$. So we will assume in the following that $N$ is a $\pi$-group. Hence, $N \subseteq H$.

**Step 2.** It suffices to show that there is $m \in N$ such that $N_{NP}(P^m)$ normalizes $K$.

Indeed, suppose that $N_{NP}(P^m)$ normalizes $K$ for some $m \in N$. Hence $P^m$, which is contained in $NP$, normalizes $K$. Then we have by Step 1 that
$\mathcal{N}_G(P^m) \subseteq \mathcal{N}_G(K)N = G_1$. Now, $P^m \in \text{Syl}_p(G_1)$, and we may apply Lemma (2.1) of [10] to conclude that

$$\mathcal{N}_G(P^m) = \mathcal{N}_N(P^m)\mathcal{N}_{G_0}(K) \subseteq \mathcal{N}_G(K).$$

However $\mathcal{N}_G(P^m) \subseteq G_1$, so $\mathcal{N}_{G_0}(P^m) = \mathcal{N}_G(P^m)$. Now, $P^m \subseteq PN \subseteq H$ is a Sylow $p$-subgroup of $H$, and we are done in this case.

Step 3. If $\theta \in \text{Irr}(N)$ of $p'$-degree extends to $PN$, then $\theta$ is $K$-invariant.

If $\eta \in \text{Irr}(PN)$ is such an extension, then $\eta^H$ has $p'$-degree. Hence there exists $\psi \in \text{Irr}(H)$ over $\eta$ of $p'$-degree. By hypothesis, we have that $\psi$ extends to some $\chi \in \text{Irr}(G)$. Let $T$ be the stabilizer of $\chi$ in $G$, so that $PN \subseteq T$. If $\nu \in \text{Irr}(T|\theta)$ is the Clifford correspondent of $\chi$ over $\theta$, then $\nu^G = \chi$ has $\pi$-degree. It follows that $T$ contains some $\pi$-complement of $G$. Thus $K^g$ is contained in $T$ for some $g \in G$. Now, we know that $P$ normalizes $KN$ by Step 1, and hence $|G : \mathcal{N}_G(K)N|$ is not divisible by $p$. Thus $|G : \mathcal{N}_G(K^g)N|$ is not divisible by $p$. By Corollary (1.2) of [19] applied in the group $G/N$ with respect to the subgroup $T/N$, we have that $|T : \mathcal{N}_T(K^g)N|$ divides $|G : \mathcal{N}_G(K)N|$, and therefore $|T : \mathcal{N}_T(K^g)N|$ is not divisible by $p$ either. It follows that there is some $R \in \text{Syl}_p(T)$ normalizing $K^gN$. Hence $R \in \text{Syl}_p(G)$. Now, $R$ and $P^g$ are Sylow $p$-subgroups of $\mathcal{N}_G(K)N$. Thus $P^gm_0 = R$ for some $m_0 \in \mathcal{N}_G(K^g)$ and $m_0 \in N$. Also, $R = P^v$ for some $v \in T$ because $R$ and $P$ are Sylow $p$-subgroups of $T$. Now, write $m = x^g$ for some $x \in \mathcal{N}_G(K)$, so that $gm = xg$. We have that $P^{xgm_0} = R = P^v$ and thus $xgm_0v^{-1} \in \mathcal{N}_G(P) \subseteq \mathcal{N}_G(K)N$. Hence, $g_{m_0}v^{-1} \in \mathcal{N}_G(K)N$, and $g_{m_0}v^{-1} = wn$ for some $w \in \mathcal{N}_G(K)$ and $n \in N$. Finally, since $K^g \subseteq T$, we have that $K^{g_{m_0}v^{-1}} \subseteq T$ (because $m_0,v \in T$). Thus $K^{vn} \subseteq T$ and $K = K^w \subseteq T^{v^{-1}} = T$, as claimed.

Step 4. We can assume that $NKP = G$. Thus $H = NP$ and $M = NK \leq G$.

By Step 1, we have that $P \subseteq \mathcal{N}_G(KN) = \mathcal{N}_G(K)N$, so $G_0 = NKP$ is a subgroup of $G$ and $NK \leq G_0$. Write $H_0 = NP$ and $G_0 = NKP$, and notice that $H_0$ is a Hall $\pi$-subgroup of $G_0$. Also, $K$ is a $\pi$-complement of $G_0$. Let $\eta \in \text{Irr}(H_0)$ with $p'$-degree. Since $|H_0 : N|$ is a power of $p$, then, by Corollary (11.29) of [5], we have that $\theta = \eta_N \in \text{Irr}(N)$ has $p'$-degree. Now $\theta$ has an extension to $H_0$. By Step 3, we conclude that $\theta$ is $K$-invariant. Now, $(|KN : N|, |N|) = 1$ and therefore $\theta$ has a canonical extension $\hat{\theta}$ to $KN$ by Corollary (8.16) of [5], which is by uniqueness, therefore $P$-invariant. Hence, by Corollary (4.2) of [6], it follows that restriction defines a bijection

$$\text{Irr}(G_0) \hat{\theta} \rightarrow \text{Irr}(H_0|\theta).$$

We conclude that $\eta$ extends to $G_0$. Now suppose that $G_0 < G$. Then, by induction, we conclude that there is $P_0 \in \text{Syl}_p(H_0)$ such that $\mathcal{N}_{G_0}(P_0) \subseteq \mathcal{N}_{G_0}(K)$. Since $P \in \text{Syl}_p(H_0)$, we have that $P_0 = P^n$ for some $n \in N$. Then $\mathcal{N}_{NP}(P^n) \subseteq \mathcal{N}_{G_0}(P_0) \subseteq \mathcal{N}_{G_0}(K)$, and we apply Step 2 in this case. Hence, we are reduced to the case where $G_0 = G$, $H = NP$, and $M = NK \leq G$.

Step 5. We can assume that $N$ is a direct product of non-abelian simple groups of order divisible by $p$ which are transitively permuted by $G$. In particular,
\( \mathcal{O}_{p'}(G) = \mathcal{O}_p(G) = 1 \). Also, we can assume that every \( p' \)-degree \( P \)-invariant irreducible character of \( N \) is \( M \)-invariant.

Suppose first that \( N \) is a \( p' \)-group. Hence, \( P \) acts coprimely on \( NK \), and because \( NK \) is \( \pi \)-separable, it follows (using, for instance, Theorem (6.31) of [5]), we conclude that \( \mathcal{O}_p(K) = \mathcal{O}_p(G) = 1 \). Now every \( P \)-invariant character of \( N \) extends to \( PN \) (by Corollary (8.16) of [5]), and by Step 3 is \( K \)-invariant. Therefore, every \( P \)-invariant character of \( N \) is \( NK \)-invariant, and therefore \( K' \)-invariant. Thus every irreducible \( P \)-invariant character of \( N \) is \( PK' \)-invariant and by Lemma (2.2) of [16], we conclude that \( \mathcal{C}_N(PK') = \mathcal{C}_N(P) \). Hence, \( \mathcal{N}_N(P) = \mathcal{C}_N(P) \subseteq \mathcal{C}_N(K') \subseteq \mathcal{N}_G(K') \). We had that \( P^{n-1} \) normalizes \( K \) and now we have that \( \mathcal{N}_N(P^{n-1}) \subseteq \mathcal{N}_G(K) \).

Thus \( \mathcal{N}_NP(P^{n-1}) = \mathcal{N}_N(P^{n-1})P^{n-1} \subseteq \mathcal{N}_G(K) \), and this case is complete by Step 2.

Suppose now that \( N \) is a \( p \)-group. Hence \( N \subseteq P \) and \( H = P \). In this case, the hypotheses tell us that every linear character of \( P \) extends to \( G \). By Tate's Theorem (use, for instance, Theorem (6.31) of [5]), we conclude that \( G \) has a normal \( p \)-complement. Hence \( K < G \), and in this case the theorem is proved.

So we may assume that \( N \) is a direct product of non-abelian simple groups of order divisible by \( p \), which are transitively permuted by \( G \). In particular \( \mathcal{O}_p(N) = N \), and it follows that every \( P \)-invariant \( p' \)-degree irreducible character of \( N \) extends to \( PN \) by Corollary (8.16) of [5]. Therefore by Step 3, every \( P \)-invariant \( p' \)-degree character of \( N \) is \( K \)-invariant, and hence \( M \)-invariant.

**Step 6.** We can assume that \( N \) is a minimal normal subgroup of \( NP \). Hence \( N = S^{g_1} \times \cdots \times S^{g_t} \), where \( \{S^{g_1}, \ldots, S^{g_t}\} \) is the \( P \)-orbit of \( S \), a non-abelian simple group of order divisible by \( p \), \( g_i \in P \), and \( g_1 = 1 \). Also, we can assume that \( t > 1 \).

We can write \( N = U \times V \), where \( U > 1 \) and \( V \geq 1 \) are \( P \)-invariant, and \( U \) is the direct product of the \( P \)-orbit of a simple group \( S \). That is, \( U = S^{g_2} \times \cdots \times S^{g_t} \), where \( \{S^{g_2}, \ldots, S^{g_t}\} \) is the \( P \)-orbit of \( S \), \( g_i \in P \), and \( g_1 = 1 \). By Lemma 2.5, let \( 1 \neq \eta \in \text{Irr}(S) \) be \( \mathcal{N}_P(S) \)-invariant of \( p' \)-degree. Then it is straightforward to show that \( \nu = \eta^{g_2} \times \cdots \times \eta^{g_t} \) is \( P \)-invariant. Now, let \( \tau = \nu \times 1_V \in \text{Irr}(N) \), which is \( P \)-invariant of \( p' \)-degree. Then \( \tau \) is \( K \)-invariant by Step 5. Hence \( \ker \tau < N \) is \( K \)-invariant, and therefore \( G \)-invariant. Since \( V \subseteq \ker \tau \), we conclude that \( V = 1 \), because \( N \) is a minimal normal subgroup of \( G \).

Suppose now that \( N \) is simple. Since \( Z(N) = 1 \), we have that \( \mathcal{C}_M(N) \) is a \( p' \)-group. Since by Step 1, we know that \( \mathcal{O}_{p'}(G) = 1 \), then we have that \( \mathcal{C}_M(N) = 1 \). Since \( G/M \) is a \( p \)-group, then we conclude that \( \mathcal{C}_G(N) \) is a \( p \)-group. But we know that \( \mathcal{O}_p(G) = 1 \), and thus we conclude that \( \mathcal{C}_G(N) = 1 \). Then, and using Step 5, we are in the hypothesis of Theorem 2.2. We conclude by this theorem (and Lemma 2.3) that there is a complement \( K_1 \) of \( N \) in \( M \) such that \( \mathcal{N}_NP(P) \subseteq \mathcal{N}_G(K_1) \). Now, \( K_1 = K^n \) for some \( n \in N \), and we have that \( \mathcal{N}_NP(P^{n-1}) \subseteq \mathcal{N}_G(K) \). Then we are done by Step 2.

**Step 7.** Conclusion.
As before, write $N = S_1 \times \cdots \times S_t$, where $S_i = S_i^{g_i}, g_i \in P$ and $S = S_1$. Also recall that $t > 1$ and therefore $N_G(S) < G$. By using Steps 5 and 6, we have that $G = P N_G(S)$. Thus $|G : N_G(S)|$ is a power of $p$. Now, since $G/N$ has a normal $p$-complement $M/N$, and $N \subseteq N_G(S)$, we conclude that $M \subseteq N_G(S)$. In particular, $M \subseteq N_G(S^i)$ for all $i$.

Now, using that $G/M$ is a $p$-group, let $G_2 \triangleleft G$ be containing $N_G(S)$ with $|G_2| < |G|$. Notice that $N_P(S)$ is a Sylow $p$-subgroup of $N_G(S)$. Also $N_P(S) \subseteq P_2 = P \cap G_2 \in \text{Syl}_p(N_G(S))$. Now, let $\eta \in \text{Irr}(NP_2)$ of $p'$-degree. We have that $\eta_N \in \text{Irr}(N)$ is $P_2$-invariant of $p'$-degree. Write $\eta_N = \theta_1 \times \cdots \times \theta_t$, where $\theta_i \in \text{Irr}(S_i)$. Then $\theta_1$ is $N_P(S)$-invariant, because we can write $N = S \times S'$, where $S$ and $S'$ are $N_P(S)$-invariant. Now, let

$$\theta = \theta_1 \times (\theta_1)^{g_2} \times \cdots \times (\theta_1)^{g_t},$$

which is $P$-invariant. By Step 5, the character $\theta$ is $M$-invariant. In particular $\theta_1$ is $M$-invariant. Since $N_P(S^i) \subseteq P_2 \triangleleft P$, we can repeat the same argument with every $S^i$ and every $\theta_i$ to conclude that $\eta_N$ is $M$-invariant. By induction applied in the group $G_2$ with respect to the Hall $\pi$-subgroup $NP_2$ and Hall $\pi$-complement $K$, we conclude that there exists $P_3 \in \text{Syl}_p(N_3)$ such that $N_G(P_3) \subseteq N_G(K)$. Hence $P_3$ normalizes $K$ and is such that $|N_N(P_3), K| = 1$ by Lemma 2.3. Now, $P_3 \cap N \in \text{Syl}_p(N)$, and also $[P_3 \cap N, K] = 1$. In particular, $P_3 \cap N \subseteq N_N(K)$ and by the Frattini argument, we see that $|G : N_G(K)|$ is not divisible by $p$.

Now, $P_3 \subseteq P_1 \in \text{Syl}_p(N_G(K))$, and we have that $P_1 \in \text{Syl}_p(G)$. Recall that $G = MP = (KN)P$. Since $P \in \text{Syl}_p(G)$, we may write $(P_1)^{k_n - 1}x = P$ for some $k \in K$, $x \in P$ and $n \in N$. Hence $(P_1)^k = P^n$. Since $P_1$ normalizes $K$, we have that $(P_1)^k$ normalizes $K$ and thus $P^n$ normalizes $K$. Also, since $[N_N(P_3), K] = 1$, we have that $[N_N((P_3)^k), K] = 1$, where $(P_3)^k \subseteq P^n$.

Finally, let $Q = (P_3)^k \cap N = (P_3 \cap N)^k \in \text{Syl}_p(N)$. Now, notice that, by elementary group theory, if $R$ is any $p$-subgroup of $G$ such that $R \cap N = Q$, then $N_G(R) \subseteq N_G(Q)$ and $N_N(R)/Q = C_{N_N(Q)/Q}(R)$.

Since $(P_3)^k \cap N = Q \in \text{Syl}_p(N)$, and $(P_3)^k \subseteq (P_1)^k$, we also have that $(P_1)^k \cap N = Q$. In particular,

$$N_N((P_1)^k)/Q = C_{N_N(Q)/Q}(P_1)^k \subseteq C_{N_N(Q)/Q}((P_3)^k) = N_N((P_3)^k)/Q$$

and we conclude that $N_N((P_1)^k) \subseteq N_N((P_3)^k) \subseteq C_G(K)$. Hence, we have found $(P_1)^k \in \text{Syl}_p(G)$ such that $(P_1)^k$ normalizes $K$ and $[N_N((P_1)^k), K] = 1$. Hence

$$N_N(P_1)^k \subseteq N_G(K)$$

using Lemma 2.3. Since $(P_1)^k = P^n \in \text{Syl}_p(H)$, we use Step 2 to finish the proof of the theorem. $\square$

In order to prove the remaining direction of Theorem A, we need one more lemma.
**Theorem 2.8.** Let $G$ be a finite $\pi$-separable group. Let $H$ be a Hall $\pi$-subgroup of $G$, let $K$ be a $\pi$-complement of $G$, and let $p$ be a prime. Suppose that there is $P \in \text{Syl}_p(H)$ such that $N_G(P) \leq N_G(K)$. Then every $\alpha \in \text{Irr}(H)$ of $p'$-degree extends to $G$.

**Proof.** Again, we can assume that $p \in \pi$. We argue by double induction, first on $|G : O_{p'}(G)|$, and second on $|G|$. If $U \leq G$ and $N \triangleleft G$, then notice that $[U : O_{p'}(U)] \leq [G : O_{p'}(G)]$, and that $[G/N : O_{p'}(G/N)] \leq [G : O_{p'}(G)]$.

By hypothesis, we have that $K \triangleleft KP \leq G$. Also, if $KP \leq U < G$, by induction, we have that the theorem is valid for $U$ with respect to any Hall subgroup of $U$ containing $P$. Let $\alpha \in \text{Irr}(H)$ be of $p'$-degree. We want to show that $\alpha$ extends to $G$. If $N \triangleleft G$, then we have that $N_G(N) \leq N_G(KN/N)$. Therefore, if $1 < N$ is a $\pi'$-group, we easily see that $|G/N : O_{p'}(G/N)| < |G : O_{p'}(G)|$, and we deduce that $\tilde{\alpha} \in \text{Irr}(H/N)$ (the unique extension of $\alpha$ to $HN$ having $N$ in its kernel) extends to $G/N$, by induction. Hence, we may assume that $O_{p'}(G) = 1$.

Now, let $N = O_{p'}(G) \leq H$. Suppose that $Z \subseteq N$ is normal in $G$. Since $\alpha$ has $p'$-degree, there exists $\tilde{\eta} \in \text{Irr}(ZP)$ of $p'$-degree under $\alpha$. Hence $\eta = \eta_Z \in \text{Irr}(Z)$, because $ZP/Z$ is a $p$-group. Now, if $ZPK < G$, by induction, the group $ZPK$ with Hall subgroup $ZP$ and complement $K$ satisfies the hypothesis of the theorem. We conclude that $\tilde{\eta}$ extends to $ZPK$, and therefore that $\eta$ is $K$-invariant. Hence $PK \leq T = I_G(\eta)$, the stabilizer of $\eta$ in $G$. We have that $HT = G$ because $K \leq T$. Let $\nu \in \text{Irr}(T \cap H)$ be the Clifford correspondent of $\alpha$ over $\eta$. If $T < G$, by induction, we conclude that $\nu$ has an extension $\tilde{\nu} \in \text{Irr}(T)$. Then $(\tilde{\nu})_H = \nu^H = \alpha$ (using Mackey), and we are done in this case. Hence, we conclude that whenever $Z \subseteq N$ is normal in $G$ then either $ZKP = G$ or $\eta$ is $G$-invariant.

In the latter case, where $\eta$ is $G$-invariant, we use the theory of character triple isomorphisms, as developed in [7]. Since $\eta$ extends to $ZP$, then by using Theorem 5.2 of [7] and its proof, we can find a character triple $(G^*, Z^*, \eta^*)$ isomorphic to $(G, Z, \eta)$, where $Z^*$ is a $p'$-group, and also a $\pi$-group. Now, $(PZ)^* = P^* \times Z^*$ for a unique Sylow $p$-subgroup $P^*$ of $G^*$. Also, $H^*$ is a Hall $\pi'$-subgroup of $G^*$ and, using the Schur–Zassenhaus theorem, we have that $(KZ)^* = K^* \times Z^*$ for a unique subgroup $K^*$ of $G^*$, which turns out to be a Hall $\pi'$-complement of $G^*$. Also, using that $N_G(PZ)^* = N_G(P^* \times Z^*) = N_G(P^*)$ and that $N_{G^*}(K^* \times Z^*) = N_{G^*}(K^*)$, we see that the hypotheses of
the theorem are satisfied in \( G^* \). Furthermore, \( N^* = \text{O}_2(G^*) \), and therefore \( |G : \text{O}_2(G)| = |G^* : \text{O}_2(G^*)| \).

Now let \( \theta \in \text{Irr}(N) \) be an irreducible constituent of \( \alpha_N \). Suppose that \( \theta \) is \( G \)-invariant. Then, using the notation of the previous paragraph with \( Z = N \), we have that \( N^* = \text{O}_2(G^*) \) is central in \( G^* \). By the Hall–Higman’s 1.2.3 Lemma, it follows that \( V = \text{O}_{2'}(G^*) > 1 \). Then \( |G^* : \text{O}_{2'}(G^*/V)| < |G^* : N^*| = |G : \text{O}_{2'}(G)| \), and by induction, and arguing as in the first paragraph of the proof, we conclude that \( \alpha^* \) extends to \( G^* \), and therefore that \( \alpha \) extends to \( G \). Hence, by the previous paragraph, we may assume that \( NP(K) = G \). Thus \( H = NP \).

Suppose now that \( \theta \) is \( K \)-invariant. In this case, \( \theta \) has a canonical extension \( \rho \) to \( M = NK \triangleleft G \), using that \( (|M : N|, |N|) = 1 \). Also \( \rho \) is \( P \)-invariant by uniqueness. Also, in this case, we know by Corollary (4.2) of [6] that restriction defines a bijection \( \text{Irr}(G)[\rho] \rightarrow \text{Irr}(NP)[\theta] \), and we conclude that \( \alpha \) extends to \( G \). Hence, it is enough to show that \( \theta \) is \( K \)-invariant.

Now, let \( N/Z \) be a chief factor of \( G \). Since \( ZKP < G \), we conclude by the second paragraph of the proof that \( \alpha_Z \) has a \( G \)-invariant irreducible constituent. Hence, by using again character triple isomorphisms, it is no loss to assume that \( Z \) is a central \( p' \)-subgroup of \( G \).

In our present situation, and using Lemma 2.3, notice that our hypotheses now are that \( P \) normalizes \( K \) (that is, \( M \triangleleft G \)) and that \( |N(KP), K| = 1 \). If \( N/Z \) is a \( p' \)-group, then \( N\rho(G) = C_N(P) \), and thus \( \text{C}_N(KP) = C_N(P) \). In this case, and using that \( \theta \) is \( P \)-invariant, \( \theta \) is \( K \)-invariant by Lemma (2.2) of [16].

If \( N/Z \) is a \( p \)-group, since \( P \) normalizes \( K \), we have that \( |O_p(G), K| = 1 \), and in this case we have that \( \theta \) is \( K \)-invariant too (since \( N \subseteq O_p(G) \times Z \)).

Hence we may assume that \( N/Z \) is the direct product \( S_1/Z \times \ldots \times S_t/Z \) of nonabelian simple groups of order divisible by \( p \) which are transitive.permuted by \( G \). In particular \( O_p(G) = 1 \). Since \( O_{2'}(G) = 1 \), we easily deduce that \( C_N(Z) = Z \). Now, \( C_H(N) \) is a Hall \( \pi \)-subgroup of \( C_G(N) \), and \( C_K(N) \) is a Hall \( \pi \)-complement of \( C_G(N) \). Hence \( C_G(N) = C_H(N)C_K(N) \).

Since \( C_K(N) = Z \), we see that \( C_G(N) \subseteq H = NP \). In particular, \( C_G(N)/Z \) is a \( p \)-group. Since \( O_p(G) = 1 \), we see that \( C_G(C) = Z \). Now, since \( N/Z \) is a direct product of nonabelian simple groups, we have that \( N/Z = N \). Since \( Z \) is a \( p' \)-group, then \( N/N' \) is a \( p' \)-group, and \( N \cap P = N' \cap P \). Suppose that \( \alpha \) is not perfect. If \( N'P = NP \), then \( N = N \cap N'P = N'(N 
\text{Irr}(NP)\lambda \rightarrow \text{Irr}(NP)\lambda_{Z(N')} \), and also \( \text{Irr}(N)\lambda \rightarrow \text{Irr}(N')\lambda_{Z(N')} \).

By the inductive hypothesis applied in \( N'PK \), we conclude that \( \alpha_{NP'} \) extends to \( N'PK \). Thus \( \alpha_{N'} \) is \( K \)-invariant. By uniqueness in the restriction map, we deduce that \( \alpha_{N'} = \theta \) is also \( K \)-invariant, and we are done in this case too. Thus we may assume that \( N = \emptyset \) is perfect.

If \( t = 1 \), that is, if \( N/Z \) is simple, then we may apply Theorem 2.2 to conclude that \( \theta \) is \( K \)-invariant. So we may assume that \( t > 1 \). Hence \( N_G(S_i) < G \).

Let \( Q = P \cap N \in \text{Syl}_p(N) \), and let \( Q_i = Q \cap S_i = P \cap S_i \in \text{Syl}_p(S_i) \). Since \( |N_i(P), K| = 1 \), we have that \( [Q, K] = 1 \). Now, let \( 1 \neq x \in Q_i \), and let \( k \in K \).

Then \( x = x^k \in S_i \cap (S_i)^k \). If \( (S_i)^k \neq S_i \), then \( (S_i)^k \cap S_i = Z \), a \( p' \)-group, and
this is not possible. We conclude that $K \subseteq N_G(S_i)$ for all $i$. Since $S_i \triangleleft N$, then $M \subseteq N_G(S_i)$ and we conclude that all the $S_i$’s are $P$-conjugate, and $PN_G(S_i) = G$.

Now, using that $Z$ is a $p'$-group, we have that $Q = Q_1 \times \cdots \times Q_t$. Write $S_i = S_{g_i}$ for some $g_i \in P$, where $S = S_1$. Let $P_0 = N_P(S) \in Syl_p(N_G(S))$. Since $N \subseteq N_G(S)$, we have that $P_0 \cap N \in Syl_p(N)$ and $P_0 \cap S \in Syl_p(S)$. Necessarily, $P_0 \cap N = Q$ and $P_0 \cap S = Q_1$. Also, since $P$ normalizes $K$, then $P_0$ normalizes $KS$. Thus $KS \triangleleft KSP_0 = G_0 < G$.

We wish to apply the inductive hypothesis in $G_0$, where here $H_0 = SP_0$ is a Hall $\pi$-subgroup of $G_0$, and $K$ is a $\pi$-complement of $G_0$. Notice that $P_0 \in Syl_p(SP_0)$, since $P_0 \cap S \in Syl_p(S)$ and $|SP_0 : P_0| = |S : S \cap P_0|$.

By Lemma 2.4, we need to check that $N_{SP_0}(P_0)$ normalizes $K$. By Lemma 2.3, we need to check that $P_0$ normalizes $K$ and $[N_S(P_0), K] = 1$. Since $P$ normalizes $K$, we only need to check that $[N_S(P_0), K] = 1$. By hypothesis, we know that $[N_N(P), K] = 1$.

Now, $P_0 \cap N_S(Q_1) = Q_1$ and therefore $N_{G_0}(P_0) \subseteq N_{G_0}(Q_1)$. We easily conclude that

\[ N_S(P_0)/Q_1 = C_{N_S(Q_1)/Q_1}(P_0). \]

By the same argument,

\[ N_N(P)/Q = C_{N_N(Q)/Q}(P). \]

Since $[N_N(P), K] = 1$, then $C_{N_N(Q)/Q}(P) = C_{N_N(Q)/Q}(PK)$.

Now, we have that $N_N(Q)/Q$ is $KP$-isomorphic to the direct product of $N_{S_i}(Q_1)/Q_i$ and that these factors are transitively permuted by $P$. By Lemma 2.7, we know that

\[ C_{N_N(Q)/Q}(KP) = C_{N_N(Q)/Q}(P) \]

if and only if

\[ C_{N_N(Q_1)/Q_1}(KP_0) = C_{N_S(Q_1)/Q_1}(P_0) = N_S(P_0)/Q_1. \]

We conclude that $K$ acts trivially on $N_S(P_0)/Q_1$. Since $[Q_1, K] = 1$, by coprime action we have that $[K, N_S(P_0)] = 1$, as desired.

Hence, we can apply the inductive hypothesis in $G_0$. Now, by using the notation of central products used in Section 5.1 of [8], we can write $\theta = \theta_1 \cdot \cdots \cdot \theta_t$, where $\theta_i \in \text{Irr}(S_i)$. Since $\theta$ is $P$-invariant, then we conclude that $\theta_1$ is $P_0$-invariant. Since the determinantal order and the degree of $\theta_1$ are prime to $p$, we see that $\theta_1$ extends to a $p'$-degree character of $SP_0$. By induction, this character extends to $G_0$, and we conclude that $\theta_1$ is $K$-invariant. The same argument applies to every $\theta_i$, and we conclude that $\theta$ is $K$-invariant. This finishes the proof of the theorem.

The proof of Corollary B now is immediate:

Proof of Corollary B. Let $\Gamma = GA$ be the semidirect product. Then $\Gamma$ has a Hall $\pi$-subgroup $G$ and a $\pi$-complement $A$. By coprime action, $G$ has an $A$-invariant Sylow $p$-subgroup $P$, and all of them are $C_G(A)$-conjugate. Also $N_{\Gamma}(P) = N_G(P)A$ and $N_{\Gamma}(A) = C_G(A) \times A$. Suppose that $[N_G(P), A] = 1$. 

Documenta Mathematica 16 (2011) 901–919
Then $N_G(P) \subseteq N_G(A)$ and by Theorem A, we have that all irreducible $p'$-degree characters of $G$ extend to $\Gamma$ and are, therefore, $A$-invariant.

Conversely, if all $p'$-degree irreducible characters of $G$ are $A$-invariant, then all of them extend to $\Gamma$ since $(|\Gamma| : |G|) = 1$. Hence, by Theorem A, there is a Sylow $p$-subgroup $P_1$ of $G$ such that $N_G(P_1) \subseteq N_G(A)$. Hence $[N_G(P_1), A] = 1$. In particular, $P_1$ is $A$-invariant, and we conclude that $(P_1)^c = P$ for some $c \in C_G(A)$. Then $[N_G(P), A] = 1$, and the proof is complete. □

3. Proof of Theorem 2.2

The aim of this section is the proof of Theorem 2.2 which we restate as follows:

**Theorem 3.1.** Let $S$ be quasi-simple, normal in a group $X$ with $S/Z(S) \leq X/Z(S) \leq \text{Aut}(S/Z(S))$, and let $p$ be a prime dividing $|S|$, with $|Z(S)|$ prime to $p$. Let $M/S$ be normal in $X/S$ of order prime to $|S|$, such that $X/M$ is a $p$-group. Let $P \in \text{Syl}_p(X)$. Then the following are equivalent:

(i) all $P$-invariant characters in $\text{Irr}_p(S)$ are $M$-invariant;

(ii) there exists a complement $K$ of $S$ in $M$ normalized by $P$ with $[N_S(P), K] = 1$.

The statement is trivially true when $M = S$, so we may assume that the quasi-simple group $S$ has outer automorphisms of order prime to $|S|$, which by the classification forces $S$ to be of Lie type and $M/S$ to consist of field automorphisms:

**Proposition 3.2.** Let $S$ be finite quasi-simple and $\sigma \in \text{Aut}(S)$ with $\gcd(o(\sigma), |S|) = 1$. Then $S$ is of Lie type and $\sigma$ is a field automorphism. In particular, Out($S$) has a cyclic central $\pi'$-Hall subgroup, where $\pi = \pi(|S|)$.

**Proof.** According to [4, Cor. 5.1.4], any automorphism of a finite quasi-simple group $S$ is induced by an automorphism of the simple quotient $S/Z(S)$, so it suffices to deal with the case that $S$ is simple. Since the outer automorphism group of finite simple groups not of Lie type has order a power of 2, clearly $S$ must be of Lie type. In this case, Out($S$) is described in [4, §2.5]: it is an extension of the normal subgroup $D$ of diagonal automorphisms (which is cyclic or a Klein four group) by the commuting product of the cyclic group of field automorphisms with the group of graph automorphisms (the latter being a subgroup of $S_3$). Now the order of $D$ is only divisible by prime divisors of the order of the Weyl group of $S$. But by Remark 3.3 below, since $o(\sigma)$ is prime to $|S|$, it is only divisible by primes larger than those occurring in $|D|$, so $\sigma$ can only act trivially on $D$. This shows the claim. □

**Remark 3.3.** In the preceding proof, we used the following observation: if $W$ is a Weyl group, and $p$ a prime divisor of its order, then all primes dividing $p - 1$ also divide $|W|$. This can easily be checked by inspection. It also follows from the rationality of Weyl groups.

In fact, even more is true: all primes smaller than $p$ divide $|W|$; no a priori proof of this is known to the authors.
3.1. The general setup. We fix the following notation throughout this section. Let $G$ be a simple linear algebraic group of simply connected type over the algebraic closure of a finite field of characteristic $r$, and $T \leq G$ a maximal torus. Then for any graph automorphism of the Dynkin diagram of $G$ and any integral power $q = r^k$ there exists a Steinberg endomorphism $F_a : G \to G$ such that $F_a$ acts as $q\phi$ on the character group $X(T)$ of $T$, with $\phi$ of finite order inducing the given graph automorphism on the Weyl group. Similarly, if $G$ is of type $G_2$, $G_1$, or $F_4$, and $r = 2, 3$, respectively $2$, then for any odd power $q = \sqrt[r]{r}$ of $\sqrt[r]{r}$ there exists a Steinberg endomorphism $F_a : G \to G$ such that $F_a$ acts as $q\phi$ on $X(T) \otimes \mathbb{R}$ and $\phi$ induces the non-trivial graph automorphism of the Coxeter diagram of $G$ (see [15, Thm. 22.5], for example).

Now for a Steinberg endomorphism $F = F_a$ as above, let $G = GF$ be the corresponding finite group of fixed points. Then, with a finite number of exceptions, $G$ is a quasi-simple group of Lie type, and all finite simple groups of Lie type arise by this setup as $G/Z(G)$, except for $^2F_4(2)'$. Since the latter group has no coprime automorphisms, this exception is of no importance for our question. So we may and will now assume that $G = GF$ is perfect. Then, except for a finite number of cases, $G$ is the universal Schur cover of the simple group $G/Z(G)$. Since none of the simple groups of Lie type with exceptional Schur multiplier have coprime automorphisms (see [4, §6.1]), we may assume that $G/Z(G)$ is none of these. Thus, the quasi-simple group $S$ from the statement of Theorem 3.1 can be realized as a central quotient $S = G/Z$ of a group $G$ as above, for some $Z \leq Z(G)$, and in particular all irreducible characters of $S$ occur among the characters of $G$.

According to Proposition 3.2 we will have to consider field automorphisms of $S$. Now any coprime field automorphism $\gamma$ of $G$, and hence of $S$, is induced by a Steinberg endomorphism of $G$ as follows. For fixed $r$ and $G$, we may assume the $F_a$ to be chosen compatibly such that $F_{an} = F_a^n$ for any $n$ coprime to $\phi(\phi)$. In this setting, if $\gamma$ is a coprime field automorphism of $G$ of order $f$, then in particular, $f$ is prime to the order of $\phi$, and thus there exists a Steinberg endomorphism $F_c : G \to G$ such that $F = F_{cf} = F_c^f$ and $\langle \gamma \rangle = \langle F_c[G] \rangle$ in Out$(G)$. The centralizer of $\gamma$ in $G$ is then just the fixed point group $G^F$ under $F_c$, a finite group of Lie type of the same type as $G$.

3.2. Action on Irr$(G)$. In order to determine the action of coprime automorphisms on Irr$(G)$, we first recall Lusztig’s parametrization. For this, let $G^*$ be a group in duality with $G$, with compatible Steinberg endomorphisms $F_a^* : G^* \to G^*$, $a \geq 1$ (respectively $a$ odd, when $G$ is a Suzuki or Ree group). We’ll write $G^* := G^{*F^*}$ for the group of fixed points of $F^* = F_a^*$, and then by the results of Lusztig, there is a partition

$$\text{Irr}(G) = \bigsqcup_{s \in G^*/\sim} \mathcal{E}(G, s)$$

into Lusztig series $\mathcal{E}(G, s)$, indexed by semisimple elements $s$ of the dual group $G^*$ modulo conjugation. Now recall that $S = G/Z$ for some subgroup $Z \leq$
Let $T \leq G$ be a maximal torus, with dual torus $T^* \leq G^*$. So any $s \in T^*$ is a linear character of $T$, and $\chi \in \mathcal{E}(G, s)$ is trivial on $Z \leq T$ if and only if $Z \leq \ker(s)$ (see [13, Lemma 2.2]). This defines a subgroup of $T^*$ of index $|Z|$; we denote by $S^*$ the normal subgroup of $G^*$ generated by all these subgroups, so $G^*/S^* \cong Z$ (and hence $[G^*, G^*] \leq S^*$). Thus, for $s \in G^*$ semisimple, the characters in $\mathcal{E}(G, s)$ have $Z$ in their center, so descend to $S = G/Z$, if and only if $s \in S^*$.

Now any field automorphism of $G$ permutes the various Lusztig series $\mathcal{E}(G, s)$. More precisely, if $\gamma$ is induced by $F_c$ as described above, then by Lusztig $\mathcal{E}(G, s)$ is $\gamma$-stable if and only if the $G^*$-class of $s$ is $F_c^*$-stable. We need the following:

**Lemma 3.4.** In the above situation, assume that $\mathcal{E}(G, s)$ is $\gamma$-stable. Then $\gamma$ fixes each orbit in $\mathcal{E}(G, s)$ under diagonal automorphisms.

**Proof.** Let $H = C_G(s)$. Lusztig [12, Prop. 5.1] proves the existence of a surjection

$$\psi: \mathcal{E}(G, s) \rightarrow \mathcal{E}(H^{oF}, 1) \mod (H^F/H^{oF})$$

with fibres the orbits under the action of diagonal automorphisms, such that multiplicities of $\rho \in \mathcal{E}(G, s)$ in Deligne–Lusztig characters are determined by those of elements of $\psi(\rho)$. We claim that $\psi$ can be chosen $\gamma$-equivariant. The result then follows since field automorphisms act trivially on $\mathcal{E}(H^{oF}, 1)$ (see e.g. [2, Prop. 6.6]).

Now unipotent characters are uniquely determined by their multiplicities in Deligne–Lusztig characters, and thus the claim follows, unless $H$ has a component of exceptional type (see [2, Prop. 6.3]). But in the latter case, characters with same multiplicities in all Deligne–Lusztig characters have distinct eigenvalues of Frobenius attached, and since $\gamma$ commutes with $F$, these are respected by $\gamma$. \qed

In our situation, as remarked in the proof of Proposition 3.2, any prime divisor of $o(\gamma)$ is larger than the order of the group of diagonal automorphisms. Thus $\gamma$ must in fact fix all elements in $\mathcal{E}(G, s)$. We hence get the following characterization of $\gamma$-stable characters of $S$ (recall that all irreducible characters of $S$ occur among those of $G$):

**Proposition 3.5.** Let $\gamma$ be a coprime automorphism of $S$, where $S = G/Z$ with $G = G^F$, $Z \leq Z(G)$ as above. Then $\chi \in \text{Irr}(S)$ is $\gamma$-stable if and only if $\chi \in \mathcal{E}(G, s)$ for some $F_c^*$-stable semisimple element $s \in S^*$, where $\gamma$ is induced by $F_c$. 

---

[12, Prop. 5.1] proves the existence of a surjection

$$\psi: \mathcal{E}(G, s) \rightarrow \mathcal{E}(H^{oF}, 1) \mod (H^F/H^{oF})$$

with fibres the orbits under the action of diagonal automorphisms, such that multiplicities of $\rho \in \mathcal{E}(G, s)$ in Deligne–Lusztig characters are determined by those of elements of $\psi(\rho)$. We claim that $\psi$ can be chosen $\gamma$-equivariant. The result then follows since field automorphisms act trivially on $\mathcal{E}(H^{oF}, 1)$ (see e.g. [2, Prop. 6.6]).

Now unipotent characters are uniquely determined by their multiplicities in Deligne–Lusztig characters, and thus the claim follows, unless $H$ has a component of exceptional type (see [2, Prop. 6.3]). But in the latter case, characters with same multiplicities in all Deligne–Lusztig characters have distinct eigenvalues of Frobenius attached, and since $\gamma$ commutes with $F$, these are respected by $\gamma$. \qed

In our situation, as remarked in the proof of Proposition 3.2, any prime divisor of $o(\gamma)$ is larger than the order of the group of diagonal automorphisms. Thus $\gamma$ must in fact fix all elements in $\mathcal{E}(G, s)$. We hence get the following characterization of $\gamma$-stable characters of $S$ (recall that all irreducible characters of $S$ occur among those of $G$):
Proof. This follows from the above, using the fact that any $F_c^*$-stable conjugacy class of the connected group $G$ contains an $F_c^*$-stable (and hence $F^* = F_c^*F_p$-stable) element (see e.g., [15, Thm. 21.11]).

3.3. Defining characteristic. We now first dispose of an easy case; observe that in defining characteristic, all semisimple characters lie in $\text{Irr}_{p'}(G)$, by Lusztig’s Jordan decomposition of characters (but note that when $r$ is bad $\text{Irr}_{p'}(G)$ may contain non-semisimple characters).

Proposition 3.6. Let $S$ be quasi-simple of Lie type, $\gamma$ a coprime automorphism of $S$. Let $p = r$, the defining characteristic of $S$. Let $A$ be a $p$-group of automorphisms of $S$. Then $\gamma$ does not centralize a Sylow $p$-subgroup of $S$, and moreover not all $A$-invariant elements of $\text{Irr}_{p'}(G)$ are $\gamma$-invariant. In particular, Theorem 3.1 holds in this situation.

Proof. Let $f$ denote the order of $\gamma$. As above, we may choose $F, F_c : G \to G$ such that $S$ is a central quotient of $G = G^F$, $F = F_c^F$ and $\gamma$ is induced by $F_c|_{G_c}$. Then $G_0 := G^{F_c}$ is a group of the same type as $G$, but over a subfield $\mathbb{F}_q$ of $\mathbb{F}_p$ of index $[\mathbb{F}_q : \mathbb{F}_p] = f$. The order formula for groups of Lie type (see e.g., [15, Table 24.1]) then shows that the $p$-parts of $G$ and $G_0$ differ. Since $Z(G)$ has order prime to $p$, the same holds for the $p$-parts of $S$ and $C_S(\gamma)$. In particular, $\gamma$ does not centralize a Sylow $p$-subgroup of $S$, whence (ii) of Theorem 3.1 does not hold.

Since diagonal automorphisms have order prime to the defining characteristic, $A$ consists of graph and field automorphisms only. Thus, $A$ induces a group of automorphisms $A^*$ on the dual $G^*$ of $G$, and we write $H := C_{G^*}(A^*)$ for its fixed point group. Now let $s \in H \cap [G^*, G^*]$ be semisimple, not centralized by $\gamma^*$ (which exists by Zsigmondy’s theorem). Then by Proposition 3.5 the Lusztig series $E(G, s)$ consists of characters with $Z(G)$ in their kernel, invariant under $A$, but not invariant under $\gamma$. Moreover, the semisimple character in $E(G, s)$ is of $p'$-degree. This shows that (i) of Theorem 3.1 is not satisfied either, thus completing the proof.

3.4. Sylow subgroups. We now turn to the non-defining primes. There we require the following crucial result from [13, Prop. 7.3]:

Proposition 3.7. Let $G$ be as above, $p \neq r$ and $\chi \in \text{Irr}_{p'}(G)$. Then there is some semisimple $s \in G^*$ centralizing a Sylow $p$-subgroup of $G^*$ with $\chi \in E(G, s)$. Conversely, if $s \in G^*$ centralizes a Sylow $p$-subgroup, then there is some $\chi \in E(G, s)$ of $p$-height 0.

We next consider the following generic case, which occurs for all large enough primes $p$:

Proposition 3.8. Let $S = G/Z$ be quasi-simple of Lie type, with $G = G^F$, $Z \leq Z(G)$ as above. Let $p \neq r$ and $P$ a Sylow $p$-subgroup of $G$, $P^*$ a Sylow $p$-subgroup of $G^*$.

(a) Then $P$ is contained in a proper $F$-stable Levi subgroup of $G$ if and only if $P^*$ is contained in a proper $F'$-stable Levi subgroup of $G^*$.

Documenta Mathematica 16 (2011) 901–919
Let $L$ be a proper $F$-stable Levi subgroup of $G$ containing $P$. If $L^*$ denotes an $F^*$-stable Levi subgroup of $G^*$ dual to $L$, then $|G^*| = |G|$ and $|L^*F^*| = |L^*|$, so $L^*$ contains a Sylow $p$-subgroup $P^*$ of $G^*$. Since we may exchange the roles of $G, G^*$ in this argument, we obtain (a).

Now assume that $P$ is contained in the $F$-stable proper Levi subgroup $L$. Then $T := Z(L)$ is an $F$-stable torus of dimension at least 1 (see [15, Prop. 12.6]). Thus, for a coprime automorphism $\gamma$ of $S$, induced by $F_\gamma$, $T^F \setminus T^{F_\gamma}$ is non-empty. Let $\kappa : G \to G/Z = S$ denote the canonical epimorphism. Since $\kappa(T^F) \leq G_\gamma(P)$, this shows that $G_\gamma(P)$ is not centralized by $\gamma$.

On the other hand, the non-trivial torus $T^* := Z(L^*)$ lies in $C_{G^*}(P^*)$. It follows that there exists a semisimple $s \in T^*F_\gamma \setminus T^*F_{\gamma}^* \subseteq C_{G^*}(P^*)$ which hence is not $F_\gamma^*$-invariant. Then, by Proposition 3.5, the elements of $E(G, s)$ are not $\gamma$-invariant. But $E(G, s)$ contains characters of height 0 by Proposition 3.7 above.

Furthermore, by assumption the Sylow $p$-subgroup $B$ of $\Out(S)$ consists of field automorphisms, hence is cyclic. Thus there exists $F_\beta : G \to G$ such that $\beta := F_\beta|_G$ generates $B$ modulo inner automorphisms. Then we may argue as before with semisimple elements in $T^*F_{\gamma} \setminus T^*(F_{\gamma}F_{\epsilon})$, respectively $s \in T^*F_{\gamma} \setminus T^*(F_{\gamma}F_{\epsilon})$ to see that neither (i) nor (ii) of Theorem 3.1 are satisfied.

3.5. BAD PRIMES AND TORSION PRIMES. It remains to deal with the primes $p$ not satisfying the condition in Proposition 3.8(a). As announced before, these are small:

**Proposition 3.9.** Let $H$ be a simple algebraic group in characteristic $r$ with Steinberg endomorphism $F : H \to H$. Let $p \neq r$ be a prime and $P$ a Sylow $p$-subgroup of $H := H^F$. If $P$ is not contained in any proper $F$-stable Levi subgroup of $H$, then every semisimple element centralizing $P$ is quasi-isolated, and in particular, $p$ is a bad prime or torsion prime for $H$.

**Proof.** Clearly we have that $P \neq 1$. Let $g \in C_H(P)$ be semisimple, so $P \leq C^g_H(g)$. Now the connected group $C^g_H(g)$ acts by conjugation on the set $\Omega$ of proper Levi subgroups of $H$ containing $C^g_H(g)$, and any orbit therein is $F$-stable; hence by [15, Thm. 21.11], if $\Omega$ is non-empty then it contains an $F$-stable element $L$, and then $P \leq C^g_H(g) \leq L$, which is not the case. So $C^g_H(g)$ is not contained in any proper Levi subgroup of $H$, whence, by definition, $g$ is quasi-isolated in $H$ (see [15, Exmp. 14.4(2)]).

Apply this to $1 \neq g \in Z(P)$. If $p$ is not a torsion prime for $H$, the index $|C^g_H(g) : C^g_H(g)|$ is prime to $p$ (see [15, Prop. 14.20]), so $P \leq C^g_H(g)$. By assumption, this does not lie in any proper Levi subgroup of $H$, so by definition $g$ is isolated. Now by a result of Deriziotis (see [15, Rem. 14.5]) and the
algorithm of Borel and de Siebenthal (see [15, Thm. 13.12]), this implies that all prime divisors of \( o(g) \) are bad for \( H \). \( \square \)

We need the following two elementary results, where we write \( \Phi_m \) for the \( m \)th cyclotomic polynomial:

**Lemma 3.10.** Let \( q \geq 1 \), \( p \) a prime not dividing \( q \), and \( e \) the multiplicative order of \( q \) modulo \( p \). Then \( p|\Phi_d(q) \) if and only if \( d = ep^i \) for some \( i \geq 0 \).

See e.g. [13, Lemma 5.2] for a proof.

**Lemma 3.11.** Let \( m, f \geq 1 \). Then

\[
\Phi_m(X^f) = \prod_d \Phi_d(X)
\]

where the product runs over all divisors \( d \) of \( mf \) which do not divide \( m'f \) for any \( m' < m \). In particular, when \( \gcd(m, f) = 1 \) then \( \Phi_m(X^f) = \prod_{d\mid f} \Phi_{md}(X) \).

**Proof.** The roots of \( \Phi_m(X) \) are the primitive \( m \)th roots of unity. Thus the roots of \( \Phi_m(X^f) \) are \( m' \)th roots of unity whose \( f \)th power is a primitive \( m' \)th root of unity, hence whose order is not a divisor of \( m'f \), for any \( m' < m \). \( \square \)

**Proposition 3.12.** Let \( G \) be connected reductive with Steinberg endomorphisms \( F_c, F = F_c^1 \) such that \( \gcd(f, |G|) = 1 \), \( H \leq G \) a connected reductive \( F_c \)-stable subgroup, \( p \) a prime dividing the order of the Weyl group of \( G \), \( p \neq r \). Then \( H^{F_c} \) contains a Sylow \( p \)-subgroup of \( H^F \).

**Proof.** The group \( H_0 = H^{F_c} \) is a group of the same type as \( H = H^F \). Thus there are non-negative integers \( a(d) \) such that \( |H_0|_e = \prod_d \Phi_d(q_0)^{a(d)} \) and \( |H|_e = \prod_d \Phi_d(q)^{a(d)} \), with \( q_0^d = q \), and \( d \) divides the order of the Weyl group \( W \) of \( G \) whenever \( a(d) > 0 \) (see [15, Cor. 24.6 and 24.7]). Since \( f \) is coprime to \( |H| \), it’s coprime to \( p - 1 \) by Remark 3.3, so \( q, q_0 \) have the same order \( e \) modulo \( p \). Now, if \( p \) divides \( \Phi_d(q) = \Phi_d(q_0^d) \), then \( d = ep^i \) for some \( i \geq 0 \) by Lemma 3.10, and hence \( \gcd(f, d) = 1 \). Then the only factor of \( \Phi_d(q_0^d) \) (in the factorization given by Lemma 3.11), which is divisible by \( p \), is \( \Phi_d(q_0) \) (again using Lemma 3.10). Thus the \( p \)-parts of \( \Phi_d(q) \) and \( \Phi_d(q_0) \) coincide, and hence \( |H_0|_p = |H|_p \). \( \square \)

**Corollary 3.13.** Let \( S \) be quasi-simple of Lie type as above, \( \gamma \) a coprime automorphism and \( p \) a prime dividing the order of the Weyl group of \( S \). Then the fixed group \( C_S(\gamma) \) contains a Sylow \( p \)-subgroup of \( S \).

**Proof.** Write \( S = G/Z \) with \( G = G^{F_c}, Z \leq Z(G) \). Now \( \gamma \) is induced by some Steinberg endomorphism \( F_c \) with \( F_c^1 = F \), where \( f = o(\gamma) \) is coprime to \( |S| \). Then \( |G_0|_p = |G|_p \) by the preceding result, with \( G_0 = G^{F_c} \). Since the subgroup \( Z \leq Z(G) \) has order only divisible by prime divisors of the order of the Weyl group, it is fixed by \( F_c \) (again by Remark 3.3) and hence \( G_0/(Z \cap G_0) = G_0/Z \leq C_S(\gamma) \), which shows that \( |S|_p = |C_S(\gamma)|_p \). \( \square \)

After these preparations we can return to the proof of Theorem 3.1.
Proposition 3.14. Let $S = G/Z$ be quasi-simple of Lie type, with $G = G^F$ and $Z \leq Z(G)$ as above, $p \neq r$ a prime such that no proper $F$-stable Levi subgroup of $G$ contains a Sylow $p$-subgroup of $G$. Then (i) in Theorem 3.1 holds.

Proof. Let $\chi \in \text{Irr}_F(S)$, so $\chi \in \mathcal{E}(G,s)$ for some semisimple element $s \in G^*$ centralizing a Sylow $p$-subgroup $P^*$ of $S^*$, by Proposition 3.7. Now by Proposition 3.8, $P^*$ is not contained in a proper $F^*$-stable Levi subgroup of $G^*$, so by Proposition 3.9, applied to $H = G^*$, we conclude that all semisimple elements in $C_{G^*}(P^*)$ are quasi-isolated, their order is only divisible by torsion primes and bad primes, so in particular only by divisors of the order of the Weyl group of $G^*$. Let $T^*$ be an $F^*$-stable torus of $G^*$ containing $s$. Then Proposition 3.12 applied with $H = T^*$ shows that $s \in T^*F_*$. But then $\chi$ is $F_*$-stable by Proposition 3.5.

For $p$ odd let $e_p(q)$ denote the order of $q$ modulo $p$, respectively the order of $q$ modulo 4 when $p = 2$.

Proposition 3.15. Let $S = G/Z$ be quasi-simple of Lie type, with $G = G^F$ and $Z \leq Z(G)$ as above, $p \neq r$ a prime such that no proper $F$-stable Levi subgroup of $G$ contains a Sylow $p$-subgroup of $G$. Then (ii) in Theorem 3.1 holds.

Proof. By Proposition 3.9 our assumptions imply that $p$ is a torsion or bad prime for $S$. By [13, Thms. 5.14, 5.19, 8.4] the normalizer of a Sylow $p$-subgroup $S_e$ of $S$ is contained in the normalizer of a $\Phi_e$-torus $T_e$, where $e = e_p(q)$, unless $S = \text{SL}_3(q), \text{SU}_3(q), G_2(q), 2F_4(q^2)$ with $p = 3$ or $S = \text{Sp}_{2n}(q), 2G_2(q^2)$ with $p = 2$.

Now first assume that $S$ is of exceptional type. Since a Sylow 2-subgroup of $(2)E_6(q)$ (with $q$ odd) is contained in a Levi subgroup of type $(2)D_5(q)$, and a Sylow 3-subgroup of $E_7(q)$ (with $3 \not| q$) is contained in a Levi subgroup of type $E_6(q)$ or $2E_6(q)$, these situations do not arise here. In the remaining cases, which are collected in Table 1, the Sylow normalizers may easily be worked out explicitly inside $N_{S}(T_e)$, respectively they are already given in [14, Sect. 3 and 4] (see also [11, Sect. 4] for the case $p = 2$). In particular they depend only on the integer $a$, which is the same for $S$ and for the centralizer of any coprime automorphism by Corollary 3.13.

Now consider the case where $S$ is of classical type. If $S$ is not of type $A$, then $p = 2$ is the only torsion prime, and from the above mentioned result on $N_S(S_p)$ (or from [1, Thm. 4]) it follows that a Sylow 2-subgroup $S_2$ of $S$ is self-normalizing, or an extension of $S_2$ of degree $3^t$ where $t$ is the number of summands in the 2-adic expansion of $n$ for $S = \text{Sp}_{2n}(q)$ with $q \equiv \pm 3 \pmod{8}$. In particular, it’s the same in $S$ and in $C_S(\gamma)$.

Finally, when $S$ is linear the possible torsion primes are divisors of $\text{gcd}(n,q-1)$. For $\text{SL}_3(q)$ with $p = 3$ and $q \equiv 4, 7 \pmod{9}$ the Sylow $p$-normalizers is isomorphic to $3^{1+2}Q_8$, independently of $q$ (see [14, Sect. 3.1]), whence the
Extending Characters from Hall Subgroups

| $G$          | $p$   | $N(S_p)$                                    |
|--------------|-------|--------------------------------------------|
| $G_2(q)$     | $2$   | $(2^q)^2,2^2$                              |
|              |       | $\sum (3^q)^2, W(G_2)$ for $q \equiv 1, 8 \pmod{9}$ |
|              |       | $3^{1+2}, Q_8, 2$ for $q \equiv 2, 4, 5, 7 \pmod{9}$ |
| $D_4(q)$     | $2$   | $(2^q)^2, 2^2$                             |
|              |       | $\sum (3^q)^2, W(G_2)$ for $2^{2f+1} \equiv 8 \pmod{9}$ |
|              |       | $3^{1+2}, Q_8, 2$ for $2^{2f+1} \equiv 2, 5 \pmod{9}$ |
| $2F_4(2^{2f+1})$ | $2$ | $(2^q)^4, 3^2, 2^3$                      |
| $F_4(q)$     | $2$   | $(2^q)^4, S_2(W(F_4))$                     |
| $E_6(q)$     | $3$   | $(3^q)^6, S_3(W(E_6)), 2^2$ for $q \equiv 1 \pmod{3}$ |
|              |       | $(3^q)^4, 3^2, 2^3$ for $q \equiv 2 \pmod{3}$ |
| $2E_8(q)$    | $3$   | $(3^q)^6, S_3(W(E_8)), 2^2$ for $q \equiv 2 \pmod{3}$ |
|              |       | $(3^q)^4, 3^2, 2^3$ for $q \equiv 1 \pmod{3}$ |
| $E_7(q)$     | $2$   | $(2^q)^7, S_2(W(E_7))$                     |
| $E_8(q)$     | $2$   | $(2^q)^8, S_2(W(E_8))$                     |
|              |       | $(3^q)^6, S_3(W(E_8)), 2^4$ for $q \equiv 1 \pmod{5}$ |
|              |       | $\sum (5^q)^8, S_3(W(E_8)), [2^6]$ for $q \equiv 2 \pmod{5}$ |

Here, $p^a = |\Phi_s(q)|p$ with $e = e_p(q)$, respectively $3^a = |\Phi_4(2^{2f+1})|3$ for the Ree groups, and $S_p(H)$ denotes a Sylow $p$-subgroup of $H$. Moreover, $p^a$ stands for a cyclic group of that order, $[p^k]$ for an unspecified group of order $p^k$, $H^k$ for a direct product of $k$ copies of $H$.

Claim holds. Exclude this case. Let $n = \sum p^{k_i}$ be the $p$-adic expansion of $n$ (thus, any $p^k$ occurs at most $p - 1$ times). If this has at least two summands, then a proper Levi subgroup $GL_{n_1}(q) \circ GL_{n_2}(q)$, with $n_1 = p^{k_1}$, $n_2 = n - n_1$, contains a Sylow $p$-subgroup of $S$. Thus we may assume that $n = p^k$ is a $p$-power. For $p = 2$ the Sylow 2-subgroups are self-normalizing by [1, Thm. 4]. For odd $p$ an easy matrix calculation shows that the normalizer of $S_p$ (modulo the center) is an extension of a homocyclic group $(p^a)^{n-1}$ by the normalizer in $S_n$ of one of its Sylow $p$-subgroups, where $p^a = |q - 1|_p$. So again, it is the same in $S$ as in the centralizer of any coprime automorphism. The case of unitary groups is entirely similar.
The proof of Theorem 3.1 is now complete by Propositions 3.6, 3.8, 3.14 and 3.15. In the course of the proof we have established the following: if $S$ is of Lie type and $p$ not the defining characteristic, then (i) and (ii) in Theorem 3.1 are equivalent to

(iii) a Sylow $p$-subgroup of $S$ is not contained in any proper Levi subgroup of $S$.

Acknowledgement. We wish to thank the anonymous referee for a very careful reading of a previous version and for proposing several improvements.

References

[1] R. Carter, P. Fong, The Sylow 2-subgroups of the finite classical groups. J. Algebra 1 (1964), 139–151.
[2] F. Digne, J. Michel, On Lusztig’s parametrization of characters of finite groups of Lie type. Astérisque No. 181-182 (1990), 113–156.
[3] D. Gajendragadkar, A characteristic class of characters of finite \( \pi \)-separable groups. J. Algebra 59 (1979), 237–259.
[4] D. Gorenstein, R. Lyons, R. Solomon, The Classification of the Finite Simple Groups. Number 3. Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 1998.
[5] I. M. Isaacs, Character Theory of Finite Groups. AMS Chelsea Publishing, Providence, RI, 2006.
[6] I. M. Isaacs, Characters of \( \pi \)-separable groups. J. Algebra 86 (1984), 98–128.
[7] I. M. Isaacs, Partial characters of \( \pi \)-separable groups. Pp. 273–287 in: Representation Theory of Finite Groups and Finite-Dimensional Algebras (Bielefeld, 1991). Progr. Math., 95, Birkhäuser, Basel, 1991.
[8] I. M. Isaacs, G. Malle, G. Navarro, A reduction theorem for the McKay conjecture. Invent. Math. 170 (2007), 33–101.
[9] I. M. Isaacs, G. Navarro, Coprime actions, fixed-point subgroups and irreducible induced characters. J. Algebra 185 (1996), 125–143.
[10] I. M. Isaacs, G. Navarro, A characteristic subgroup and kernels of Brauer characters. Bull. Austral. Math. Soc. 72 (2005), no. 3, 381–384.
[11] A. S. Kondrat’ev, V. D. Mazurov, 2-signalizers of finite simple groups. (Russian) Algebra Logika 42 (2003), 594–623, 638; translation in Algebra Logic 42 (2003), 333–348.
[12] G. Lusztig, On the representations of reductive groups with disconnected centre. Astérisque 168 (1988), 157–166.
[13] G. Malle, Height 0 characters of finite groups of Lie type. Represent. Theory 11 (2007), 192–220.
[14] G. Malle, Extensions of unipotent characters and the inductive McKay condition. J. Algebra 320 (2008), 2963–2980.
[15] G. Malle, D. Testerman, Linear Algebraic Groups and Finite Groups of Lie Type. Cambridge University Press, Cambridge, 2011.
[16] G. NAVARRO, Actions and invariant character degrees. J. Algebra 160 (1993), 172–178.
[17] G. NAVARRO, Actions and characters in blocks. J. Algebra 275 (2004), 471–480.
[18] C. SAH, Existence of normal complements and extension of characters in finite groups. Illinois J. Math. 6 (1962), 282–291.
[19] A. TURULL, The number of Hall $\pi$-subgroups of a $\pi$-separable group. Proc. Amer. Math. Soc. 132 (2004), 2563–2565.

Gunter Malle
FB Mathematik
TU Kaiserslautern
Postfach 3049
67653 Kaiserslautern
Germany.
malle@mathematik.uni-kl.de

Gabriel Navarro
Departament d’Àlgebra
Universitat de València
Dr. Moliner 50
46100 Burjassot
Spain.
gabriel.navarro@uv.es
The Singularity Category of an Algebra with Radical Square Zero

Xiao-Wu Chen

Received: April 21, 2011
Communicated by Stefan Schwede

Abstract. To an artin algebra with radical square zero, a regular algebra in the sense of von Neumann and a family of invertible bimodules over the regular algebra are associated. These data describe completely, as a triangulated category, the singularity category of the artin algebra. A criterion on the Hom-finiteness of the singularity category is given in terms of the valued quiver of the artin algebra.

2010 Mathematics Subject Classification: 18E30, 13E10, 16E50
Keywords and Phrases: singularity category, von Neumann regular algebra, invertible bimodule, valued quiver

1. Introduction

Let $R$ be a commutative artinian ring. All algebras, categories and functors are $R$-linear. We recall that an $R$-linear category is Hom-finite provided that all the Hom sets are finitely generated $R$-modules.

Let $A$ be an artin $R$-algebra. Denote by $A$-mod the category of finitely generated left $A$-modules, and by $D^b(A$-mod) the bounded derived category. Following [14], the singularity category $D_{sg}(A)$ is the quotient triangulated category of $D^b(A$-mod) with respect to the full subcategory formed by perfect complexes; see also [3, 12, 10, 15, 2] and [13]. Here, we recall that a complex in $D^b(A$-mod) is perfect provided that it is isomorphic to a bounded complex consisting of finitely generated projective modules.

The singularity category measures the homological singularity of an algebra in the sense that an algebra $A$ has finite global dimension if and only if its singularity category $D_{sg}(A)$ vanishes. In the meantime, the singularity category captures the stable homological features of an algebra ([3]). A fundamental result of Buchweitz and Happel states that for a Gorenstein algebra $A$, the
singularity category $D_{sg}(A)$ is triangle equivalent to the stable category of
(maximal) Cohen-Macaulay $A$-modules ([3, 10]). This implies in particular
that the singularity category of a Gorenstein algebra is Hom-finite and has
Auslander-Reiten triangles. We point out that Buchweitz and Happel’s result
specializes to Rickard’s result ([15]) on self-injective algebras. However, for
non-Gorenstein algebras, not much is known about their singularity categories
([4]).

Our aim is to describe the singularity category of an algebra with radical square
zero. We point out that such algebras are usually non-Gorenstein ([5]). In what
follows, we describe the results in this paper.

We denote by $r$ the Jacobson radical of $A$. The algebra $A$ is said to be with
radical square zero provided that $r^2 = 0$. In this case, $r$ has a natural $A/r$-$A/r$
-bimodule structure. Set $r^{\otimes i} = A/r$ and $r^{\otimes i+1} = r \otimes_{A/r} (r^{\otimes i})$ for $i \geq 0$. Then
there are obvious algebra homomorphisms $\text{End}_{A/r}(r^{\otimes i}) \to \text{End}_{A/r}(r^{\otimes i+1})$ induced by $r \otimes_{A/r} -$. We denote by $\Gamma(A)$ the direct limit of this chain of algebra
homomorphisms. It is a regular algebra ([7, 8]) in the sense of von Neumann.

We call $\Gamma(A)$ the associated regular algebra of $A$. In most cases, the algebra
$\Gamma(A)$ is not semisimple.

For $n \in \mathbb{Z}$ and $i \geq \max\{0, n\}$, $\text{Hom}_{A/r}(r^{\otimes i}, r^{\otimes i-n})$ has a natural
$\text{End}_{A/r}(r^{\otimes i-n})$-$\text{End}_{A/r}(r^{\otimes i})$-bimodule structure. Set $K^n(A)$ to be the direct
limit of the chain of maps $\text{Hom}_{A/r}(r^{\otimes i}, r^{\otimes i-n}) \to \text{Hom}_{A/r}(r^{\otimes i+1}, r^{\otimes i+1-n})$, which are induced by $r \otimes_{A/r} -$. Then $K^n(A)$ is naturally a $\Gamma(A)$-$\Gamma(A)$-
bimodule for each $n \in \mathbb{Z}$. Observe that $K^0(A) = \Gamma(A)_\Gamma(A)$ as bimodules,
and that composition of maps induces $\Gamma(A)$-$\Gamma(A)$-bimodule morphisms $\phi^{n,m}: K^n(A) \otimes_{\Gamma(A)} K^m(A) \to K^{n+m}(A)$ for all $n, m \in \mathbb{Z}$. These bimodules
$K^n(A)$ are called the associated bimodules of $A$.

Recall that for an algebra $\Gamma$, a $\Gamma$-$\Gamma$-bimodule $K$ is invertible provided that the
functor $K \otimes -$ induces an auto-equivalence on the category of left $\Gamma$-modules.

**Theorem A.** Let $A$ be an artin algebra with radical square zero. Use the above
notation. Then the associated $\Gamma(A)$-$\Gamma(A)$-bimodules $K^n(A)$ are invertible and
the maps $\phi^{n,m}$ are isomorphisms of bimodules.

Since the algebra $\Gamma(A)$ is regular, the category $\text{proj} \Gamma(A)$ of finitely generated
right projective $\Gamma(A)$-modules is a semisimple abelian category. The invertible
bimodule $K^1(A)$ induces an auto-equivalence

$$\Sigma_A = - \otimes_{\Gamma(A)} K^1(A): \text{proj} \Gamma(A) \longrightarrow \text{proj} \Gamma(A).$$

We observe that the category $\text{proj} \Gamma(A)$ has a unique triangulated structure
with $\Sigma_A$ its shift functor; see Lemma 3.4. This unique triangulated category is
denoted by $(\text{proj} \Gamma(A), \Sigma_A)$.

The following result describes the singularity category of an artin algebra with
radical square zero, which is based on a result by Keller and Vossieck ([12]).
THEOREM B. Let $A$ be an artin algebra with radical square zero. Use the above notation. Then we have a triangle equivalence
\[ \mathcal{D}_{\text{sg}}(A) \simeq (\text{proj } \Gamma(A), \Sigma A). \]

We are interested in the Hom-finiteness of singularity categories. For this, we recall the notion of valued quiver of an artin algebra $A$. Choose a complete set of representatives of pairwise non-isomorphic simple $A$-modules $\{S_1, S_2, \ldots, S_n\}$. Set $\Delta_i = \text{End}_A(S_i)$; they are division algebras. Observe that $\text{Ext}^1_A(S_i, S_j)$ has a natural $\Delta_j$-$\Delta_i$-bimodule structure. The valued quiver $Q_A$ of $A$ is defined as follows: its vertex set is $\{S_1, S_2, \ldots, S_n\}$, here we identify each simple module $S_i$ with its isoclass; there is an arrow from $S_i$ to $S_j$ whenever $\text{Ext}^1_A(S_i, S_j) \neq 0$, in which case the arrow is endowed with a valuation $(\dim \Delta_j \text{ Ext}^1_A(S_i, S_j), \dim \Delta_i \text{ Ext}^1_A(S_i, S_j))$; here $\Delta_i^{\text{op}}$ denotes the opposite algebra of $\Delta_i$. We say that the valuation of $Q_A$ is trivial provided that all the valuations are $(1, 1)$. Recall that a vertex in a valued quiver is a source (resp. sink) provided that there is no arrows ending (resp. starting) at it. For a valued quiver, to adjoin a (new) source (resp. sink) is to add a vertex together with some valued arrows starting (resp. ending) at this vertex. For details, we refer to [1, III.1].

The following result characterizes when the singularity category is Hom-finite, using the valued quivers.

THEOREM C. Let $A$ be an artin algebra with radical square zero. Then the following statements are equivalent:

1. the singularity category $\mathcal{D}_{\text{sg}}(A)$ is Hom-finite;
2. the associated regular algebra $\Gamma(A)$ is semisimple;
3. the valued quiver $Q_A$ is obtained from a disjoint union of oriented cycles with the trivial valuation by repeatedly adjoining sources or sinks.

The paper is structured as follows. In Section 2, we collect some facts on singularity categories and recall a basic result due to Keller and Vossieck ([12]), which is applied to $\Omega^\infty$-finite algebras.

2. Preliminaries

In this section, we collect some facts on singularity categories of artin algebras. We recall a basic result due to Keller and Vossieck ([12]), which is applied to $\Omega^\infty$-finite algebras.

Let $A$ be an artin algebra over a commutative artinian ring $R$. Recall that $A$-mod denotes the category of finitely generated left $A$-modules. We denote
by $A$-proj the full subcategory formed by projective modules, and by $A$-mod the stable category of $A$-mod modulo projective modules ([1, p.104]). The morphism space $\text{Hom}_A(M, N)$ of two modules $M$ and $N$ in $A$-mod is defined to be $\text{Hom}_A(M, N)/p(M, N)$, where $p(M, N)$ denotes the $R$-submodule formed by morphisms that factor through projective modules.

Recall that for an $A$-module $M$, its syzygy $\Omega(M)$ is the kernel of its projective cover $P \to M$. This gives rise to the syzygy functor $\Omega: A$-mod $\to A$-mod ([1, p.124]). Set $\Omega^0(M) = M$ and $\Omega^{i+1}(M) = \Omega^i(\Omega(M))$ for $i \geq 0$. Denote by $\Omega^i(A$-mod) the full subcategory of $A$-mod formed by modules of the form $P \oplus \Omega^i(M)$ for some module $M$ and projective module $P$. Then an $A$-module $X$ belongs to $\Omega^i(A$-mod) if and only if there is an exact sequence $0 \to X \to P_{1-i} \to \cdots \to P_{-1} \to P_0$ with each $P_j$ projective.

Recall that $D^b(A$-mod) denotes the bounded derived category of $A$-mod, whose shift functor is denoted by $[1]$. For $n \in \mathbb{Z}$, $[n]$ denotes the $n$-th power of $[1]$. The module category $A$-mod is viewed as a full subcategory of $D^b(A$-mod) by identifying an $A$-module with the corresponding stalk complex concentrated at degree zero ([11, Proposition I.4.3]). Recall that a complex in $D^b(A$-mod) is perfect provided that it is isomorphic to a bounded complex consisting of projective modules; these complexes form a full triangulated subcategory $\text{perf}(A)$. Recall that, via an obvious functor, $\text{perf}(A)$ is triangle equivalent to the bounded homotopy category $K^b(A$-proj); compare [3, 1.1-1.2].

Following [14], we call the quotient triangulated category

$$D_{\text{sg}}(A) = D^b(A$-mod)/$\text{perf}(A)$$

the singularity category of $A$. Denote by $q: D^b(A$-mod) $\to D_{\text{sg}}(A)$ the quotient functor.

The following two results are known; compare [14, Lemma 1.11] and [3, Lemma 2.2.2].

**Lemma 2.1.** Let $X^\bullet$ be a complex in $D_{\text{sg}}(A)$ and $n_0 > 0$. Then for any $n$ large enough, there exists a module $M$ in $\Omega^n(A$-mod) such that $X^\bullet \simeq q(M)[n]$.

**Proof.** Take a quasi-isomorphism $P^\bullet \to X^\bullet$ with $P^\bullet$ a bounded above complex of projective modules ([11, Lemma I.4.6]). Take $n \geq n_0$ such that $H^i(X^\bullet) = 0$ for all $i < n_0 - n$, where $H^i(X^\bullet)$ denotes the $i$-th cohomology of $X^\bullet$. Consider the good truncation $\sigma_{\leq -n}P^\bullet = \cdots \to 0 \to M \to P_{1-n} \to P_{2-n} \to \cdots$ of $P^\bullet$, which is quasi-isomorphic to $P^\bullet$. Then the cone of the obvious chain map $\sigma_{\leq -n}P^\bullet \to M[n]$ is perfect, which becomes an isomorphism in $D_{\text{sg}}(A)$. This shows that $X^\bullet \simeq q(M)[n]$. We observe that $M$ lies in $\Omega^n(A$-mod). \hfill $\Box$

**Lemma 2.2.** Let $0 \to M \to P_{1-n} \to \cdots \to P_0 \to N \to 0$ be an exact sequence with each $P_j$ projective. Then we have an isomorphism $q(N) \simeq q(M)[n]$ in $D_{\text{sg}}(A)$. In particular, for an $A$-module $M$, we have a natural isomorphism $q(\Omega^n(M)) \simeq q(M)[-n]$.

**Proof.** The stalk complex $N$ is quasi-isomorphic to $\cdots \to 0 \to M \to P_{1-n} \to \cdots \to P_0 \to 0 \to \cdots$. This gives rise to a morphism $N \to M[n]$ in $D^b(A$-mod),
whose cone is \( \cdots \to 0 \to P^{1-n} \to \cdots \to P^0 \to 0 \to \cdots \) with \( P^0 \) at degree \(-1\); it is perfect. Then the morphism \( N \to M[n] \) becomes an isomorphism in \( D_{sg}(A) \). □

Consider the composite \( q' : A\text{-mod} \hookrightarrow D^b(A\text{-mod}) \xrightarrow{q} D_{sg}(A) \); it vanishes on projective modules. Then it induces uniquely a functor \( A\text{-mod} \to D_{sg}(A) \), which is still denoted by \( q' \). Then Lemma 2.2 yields, for each \( n \geq 0 \), the following commutative diagram

\[
\begin{array}{ccc}
A\text{-mod} & \xrightarrow{\Omega^n} & A\text{-mod} \\
\downarrow{q'} & & \downarrow{q'} \\
D_{sg}(A) & \xrightarrow{[-n]} & D_{sg}(A).
\end{array}
\]

We refer to [3, Lemma 2.2.2] for a similar statement. The functor \( q' \) induces a natural map

\[
\Phi^0 : \text{Hom}_A(M, N) \to \text{Hom}_{D_{sg}(A)}(q(M), q(N))
\]

for any modules \( M, N \). Let \( n \geq 1 \). Lemma 2.2 yields a natural isomorphism \( \theta_M : q(M) \xrightarrow{\sim} q(\Omega^n(M))[n] \). Then we have a map

\[
\Phi^n : \text{Hom}_A(\Omega^n(M), \Omega^n(N)) \to \text{Hom}_{D_{sg}(A)}(q(M), q(N))
\]

given by \( \Phi^n(f) = (\theta_M^n)^{-1} \circ (\Phi^0(f)[n]) \circ \theta_M^n \).

Consider the chain of maps

\[
\text{Hom}_A(\Omega^n(M), \Omega^n(N)) \to \text{Hom}_A(\Omega^{n+1}(M), \Omega^{n+1}(N))
\]

induced by the syzygy functor. It is routine to verify that \( \Phi^n \) are compatible with this chain of maps. Then we have an induced map

\[
\Phi : \text{lim}_{\rightarrow} \text{Hom}_A(\Omega^n(M), \Omega^n(N)) \to \text{Hom}_{D_{sg}(A)}(q(M), q(N)).
\]

We recall the following basic result.

**Proposition 2.3.** (Keller-Vossieck) Let \( M, N \) be \( A \)-modules as above. Then the map \( \Phi \) is an isomorphism.

**Proof.** The statement follows from [12, Exemple 2.3]. We refer to [2, Corollary 3.9(1)] for a detailed proof. □

Recall that an additive category \( \mathcal{A} \) is idempotent split provided that each idempotent \( e : X \to X \) splits, that is, it admits a factorization \( X \xrightarrow{u} Y \xrightarrow{v} X \) with \( u \circ v = \text{Id}_Y \). For example, a Krull-Schmidt category is idempotent split ([6, Appendix A]). In particular, for an artin algebra \( A \), the stable category \( A\text{-mod} \) is idempotent split.

**Corollary 2.4.** The singularity category \( D_{sg}(A) \) of an artin algebra \( A \) is idempotent split.
Proof. By Lemma 2.1 it suffices to show that for each module \( M \), an idempotent 
\( e : q(M) \to q(M) \) splits in \( D_{sg}(A) \). The above proposition implies that for a 
large \( n \), there is an idempotent \( e^n : \Omega^n(M) \to \Omega^n(M) \) in \( A\text{-mod} \) which is 
mapped by \( \Phi \) to \( e \). The idempotent \( e^n \) splits as \( \Omega^n(M) \to Y \xrightarrow{v} \Omega^n(M) \) with 
\( u \circ v = \text{Id}_Y \) in \( A\text{-mod} \). Then the idempotent \( e \) splits as \( q(M) \xrightarrow{(q(u)[n])e_{\Omega^n}} 
q(Y)[n] \xrightarrow{(q(v)[n])} q(M) \).

Let \( A \) be an additive category. For a subcategory \( C \), denote by \( \text{add} \ C \) the full 
subcategory of \( A \) formed by direct summands of finite direct sums of objects in \( C \). For any algebra \( \Gamma \), denote by \( \text{proj} \ \Gamma \) the category of finitely generated 
right projective \( \Gamma \)-modules. We observe that \( \text{proj} \ \Gamma = \text{add} \ \Gamma \)-finite algebra with an 
\( \Omega^n \)-generator \( A \).

Proposition 2.5. Let \( A \) be an \( \Omega^{\infty} \)-finite algebra with an \( \Omega^{\infty} \)-generator \( E \). 
Then we have \( D_{sg}(A) = \text{add} \ q(E) \). Consequently, we have an equivalence of 
categories 
\[ D_{sg}(A) \cong \text{proj} \text{End}_{D_{sg}(A)}(q(E)), \]
which sends \( q(E) \) to \( \text{End}_{D_{sg}(A)}(q(E)) \).

Proof. Observe that \( \Omega^{n+1}(A\text{-mod}) \subseteq \Omega^n(A\text{-mod}) \). Then we may assume that 
\( \text{add} \ (A \oplus E) \supseteq \text{add} \ \Omega^{n_0}(A\text{-mod}) = \text{add} \ (A \oplus E) \) for \( n_0 \) large 

For the first statement, it suffices to show that each object \( X^\bullet \) in \( D_{sg}(A) \) belongs to \( q(E) \). By Lemma 2.1, \( X^\bullet \cong q(M)[n_1] \) for a module \( M \in \Omega^{n_0}(A\text{-mod}) \) and \( n_1 > 0 \). Since add \( \Omega^{n_0}(A\text{-mod}) = \text{add} \ \Omega^{n_0+n_1}(A\text{-mod}) \), we may assume that \( M \otimes N \in \Omega^{n_0+n_1}(A\text{-mod}) \) for some module \( N \). Take an exact 
sequence \( 0 \to M \otimes N \to P_1 \to P_0 \to L \to 0 \) with each \( P_i \) projective 
and \( L \in \Omega^{n_0}(A\text{-mod}) \). By Lemma 2.2, \( q(L) \cong q(M \otimes N)[n_1] \) and then \( X^\bullet \) is a 
direct summand of \( q(L) \). Observing that \( L \in \text{add} \ (A \oplus E) \), we are done with the 
first statement.

The second statement follows from the projectivization; see [1, Proposition II.2.1]. The functor is given by \( \text{Hom}_{D_{sg}(A)}(q(E), -) \). We point out that Corollary 2.4 is needed here. \( \square \)

3. Algebras with radical square zero

In this section, we study the singularity category of an algebra with radical 
square zero, and prove Theorem A and B. An explicit example is given at the 
end.

Let \( A \) be an artin algebra. Denote by \( r \) the Jacobson radical of \( A \). The algebra 
\( A \) is said to be with \emph{radical square zero} provided that \( r^2 = 0 \). In this case, \( r \) 
has an \( A/r-A/r \)-bimodule structure, which is induced from the multiplication of \( A \).
Denote by $A\text{-ssmod}$ the full subcategory of $A\text{-mod}$ formed by semisimple modules. We observe that $r \otimes_{A/r} S = 0$ for a simple projective module $S$. Then the functor $r \otimes_{A/r} - : A\text{-ssmod} \to A\text{-ssmod}$ is well defined. We observe that the syzygy functor $\Omega$ sends semisimple modules to semisimple modules, and then we have the restricted functor $\Omega: A\text{-ssmod} \to A\text{-ssmod}$.

The following result is implicitly contained in the proof of [1, Lemma X.2.1].

**Lemma 3.1.** There is a natural isomorphism $\Omega \simeq r \otimes_{A/r} -$ of functors on $A\text{-ssmod}$.

**Proof.** Let $X$ be a semisimple $A$-module. Take a projective cover $P \to X$. Tensoring $P$ with the natural exact sequence of $A$-$A$-bimodules $0 \to r \to A \to A/r \to 0$ yields $\Omega(X) \simeq r \otimes_A P$. Using isomorphisms $r \otimes_A P \simeq r \otimes_{A/r} P/rP$ and $P/rP \simeq X$, we get an isomorphism $\Omega(X) \simeq r \otimes_{A/r} X$. It is routine to verify that this isomorphism is natural in $X$.

Recall that an algebra $\Gamma$ is regular in the sense of von Neumann provided that for each element $a$ there exists $a'$ such that $aa'a = a$. For example, a semisimple algebra is regular. Then a direct limit of semisimple algebras is regular. For details, we refer to [7, Theorem and Definition 11.24].

Recall that for an artin algebra $A$ with radical square zero, there is a chain of algebra homomorphisms $\text{End}_{A/r}(r^{\otimes i}) \to \text{End}_{A/r}(r^{\otimes i+1})$ induced by $r \otimes_{A/r} -$. Here, $r^{\otimes 0} = A/r$ and $r^{\otimes i+1} = r \otimes_{A/r} (r^{\otimes i})$. We set $\Gamma(A)$ to be the direct limit of this chain. Since each algebra $\text{End}_{A/r}(r^{\otimes i})$ is semisimple, the algebra $\Gamma(A)$ is regular. It is called the associated regular algebra of $A$. We refer to [8, 19.26B, Example] for a related construction.

We recall the associated $\Gamma(A)$-$\Gamma(A)$-bimodules $K^n(A)$ of $A$, $n \in \mathbb{Z}$. For $i \geq \max\{0, n\}$, $\text{Hom}_{A/r}(r^{\otimes i}, r^{\otimes i-n})$ has a natural $\text{End}_{A/r}(r^{\otimes i-n})$-$\text{End}_{A/r}(r^{\otimes i})$-bimodule structure. Consider a chain of maps $\text{Hom}_{A/r}(r^{\otimes i}, r^{\otimes i-n}) \to \text{Hom}_{A/r}(r^{\otimes i+1}, r^{\otimes i+1-n})$, which are induced by $r \otimes_{A/r} -$, and define $K^n(A)$ to be its direct limit. Then $K^n(A)$ is naturally a $\Gamma(A)$-$\Gamma(A)$-bimodule for each $n \in \mathbb{Z}$. Observe that $K^n(A) = \Gamma(A)\Gamma(A)\Gamma(A)$ as $\Gamma(A)$-$\Gamma(A)$-bimodules.

**Proposition 3.2.** Let $A$ be an artin algebra with radical square zero. Then there is a natural isomorphism

$$K^n(A) \simeq \text{Hom}_{D_{sc}(A)}(q(A/r), q(A/r)[n])$$

for each $n \in \mathbb{Z}$.

**Proof.** Consider the case $n \leq 0$ first. In this case, by Lemmas 2.2 and 3.1 we have $q(A/r)[n] \simeq q(\Omega^{-n}(A/r)) \simeq q(r^{\otimes -n})$. Then Proposition 2.3 yields an isomorphism $\text{Hom}_{D_{sc}(A)}(q(A/r), q(A/r)[n]) \simeq \lim \text{Hom}_{A/r}(\Omega^i(A/r), \Omega^i(r^{\otimes -n}))$. By Lemma 3.1 again we have $\Omega^i(A/r) \simeq r^{\otimes i}$ and $\Omega^i(r^{\otimes -n}) = r^{\otimes -i-n}$. Then we have a surjective map $\psi: K^n(A) \to \text{Hom}_{D_{sc}(A)}(q(A/r), q(A/r)[n])$. On the other hand, every morphism $f: r^{\otimes i} \to r^{\otimes -i-n}$ that is zero in $A\text{-mod}$ necessarily factors through a semisimple projective module. However, the functor $r \otimes_{A/r} -$ vanishes on semisimple projective modules. Then $r \otimes_{A/r} f$ is zero. This forces that $\psi$ is injective. We are done in this case.
Recall that an abelian category $\mathcal{A}$ is semi-simple.

**Proof.**

We observe that for any $\mathcal{A}$-equivalence on $\text{proj } \mathcal{A}$, the following observation is well known. The following isomorphism is an isomorphism $K^0(\mathcal{A}) = \Gamma(\mathcal{A}) \cong \text{End}_{\text{proj } \mathcal{A}}(q(\mathcal{A}/r))$ of algebras. Then for an arbitrary $n$, the above isomorphism becomes an isomorphism of $\Gamma(\mathcal{A})$-$\Gamma(\mathcal{A})$-bimodules.

Recall that an abelian category $\mathcal{A}$ is semi-simple provided that each short exact sequence splits. For example, for a regular algebra $\Gamma$, the category $\text{proj } \Gamma$ of finitely generated right projective $\Gamma$-modules is a semi-simple abelian category.

The following observation is well known.

**Lemma 3.4.** Let $\mathcal{A}$ be a semi-simple abelian category, and let $\Sigma$ be an auto-equivalence on $\mathcal{A}$. Then there is a unique triangulated structure on $\mathcal{A}$ with $\Sigma$ the shift functor.

The obtained triangulated category in this lemma will be denoted by $(\mathcal{A}, \Sigma)$.

**Proof.**

We use the fact that each morphism in $\mathcal{A}$ is isomorphic to a direct sum of morphisms of the forms $K \rightarrow 0$, $I \xrightarrow{\text{Id}} I$ and $0 \rightarrow C$. Then all possible triangles are a direct sum of the following trivial triangles $K \rightarrow 0 \rightarrow \Sigma(K) \xrightarrow{\text{Id}} \Sigma(K)$, $I \xrightarrow{\text{Id}} I \rightarrow 0 \rightarrow \Sigma(I)$ and $0 \rightarrow C \xrightarrow{\text{Id}} C \rightarrow \Sigma(0)$. 

**Proposition 3.5.** Let $\Lambda$ be an artin algebra with radical square zero, and let $\Gamma(\Lambda)$ be its associated regular algebra. Then there is a triangle equivalence

$$\Psi : \text{D}_{\text{sg}}(\Lambda) \cong (\text{proj } \Gamma(\Lambda), \Sigma)$$

for some auto-equivalence $\Sigma$ on proj $\Gamma(\Lambda)$, which sends $q(\Lambda/r)$ to $\Gamma(\Lambda)$.

**Proof.**

We observe that for any $\Lambda$-module $M$, its syzygy $\Omega(M)$ is semi-simple. Hence we have $\Omega(-) \cong \text{proj } \text{Hom}(\Lambda, -)$. We apply Proposition 2.5 to obtain an equivalence of categories $\text{D}_{\text{sg}}(\Lambda) \cong \text{proj } \text{End}_{\text{proj } \mathcal{A}}(q(\mathcal{A}/r))$. By Proposition 3.2 this yields an equivalence of categories $\text{D}_{\text{sg}}(\Lambda) \cong \text{proj } \Gamma(\Lambda)$. By transport of structures, the shift functor $[1]$ on $\text{D}_{\text{sg}}(\Lambda)$ corresponds to an auto-equivalence $\Sigma$ on proj $\Gamma(\Lambda)$, and then proj $\Gamma(\Lambda)$ becomes a triangulated category. However, by Lemma 3.3 the semi-simple abelian category proj $\Gamma(\Lambda)$ has a unique triangulated structure with $\Sigma$ the shift functor. Then this structure necessarily coincides with the transported one. Then we are done.

We are interested in the auto-equivalence $\Sigma$ above. The following result characterizes it using the bimodules $K^n(\Lambda)$.

**Lemma 3.6.** Use the notation as above. Then for each $n \in \mathbb{Z}$, the auto-equivalence $\Sigma^n$ is isomorphic to $- \otimes_{\Gamma(\Lambda)} K^n(\Lambda) : \text{proj } \Gamma(\Lambda) \rightarrow \text{proj } \Gamma(\Lambda)$. 
Proof. Recall that the above equivalence $\Psi$ is given by $\text{Hom}_{\mathcal{D}_{sg}(A)}(q(A/r), -)$, which sends $q(A/r)$ to $\Gamma(A)$. The auto-equivalence $\Sigma^n$ corresponds, via $\Psi$, to $[n]$ on $\mathcal{D}_{sg}(A)$. Then by Proposition 3.2 we have an isomorphism $\phi: K^n(A) \rightarrow \Sigma^n(\Gamma(A))$ of right $\Gamma(A)$-modules. Recall that $\Sigma^n(\Gamma(A))$ has a natural $\Gamma(A)$-$\Gamma(A)$-bimodule structure such that $\Sigma^n$ is isomorphic to $- \otimes_{\Gamma(A)} \Sigma^n(\Gamma(A))$. Thanks to Remark 3.3, the isomorphism $\phi$ is an isomorphism of bimodules. This proves the lemma. 

Recall that for an algebra $\Gamma$, a $\Gamma$-$\Gamma$-bimodule $K$ is invertible provided that the functor $- \otimes_{\Gamma} K$ induces an auto-equivalence on the category of right $\Gamma$-modules. For details, we refer to [7, Definition and Proposition 12.13]. We recall that for an artin algebra $A$ with radical square zero, the associated $\Gamma(A)$-$\Gamma(A)$-bimodules $K^n(A)$ are defined to be $\lim \text{Hom}_{A/r}(r^{\otimes i}, r^{\otimes -n})$, where $i \geq \max\{0, n\}$. Then composition of maps between the $A/r$-modules $r^{\otimes i}$ yields morphisms 

$$
\phi^{n,m}: K^n(A) \otimes_{\Gamma(A)} K^m(A) \rightarrow K^{n+m}(A)
$$

of $\Gamma(A)$-$\Gamma(A)$-bimodules, for all $n, m \in \mathbb{Z}$. More precisely, let $f \in K^n(A)$ and $g \in K^m(A)$ be represented by $f': r^{\otimes j-n} \rightarrow r^{\otimes j-m-n}$ and $g': r^{\otimes j} \rightarrow r^{\otimes j-m}$ for some large $j$, respectively. Then $\phi^{n,m}(f \otimes g)$ is represented by the composite $f' \circ g'$. 

The following result is Theorem A.

**Theorem 3.7.** Let $A$ be an artin algebra with radical square zero. Use the above notation. Then for all $n, m \in \mathbb{Z}$, the $\Gamma(A)$-$\Gamma(A)$-bimodules $K^n(A)$ are invertible and the morphisms $\phi^{n,m}$ are isomorphisms.

**Proof.** By Lemma 3.6 the functor $- \otimes_{\Gamma(A)} K^n(A): \text{proj} \Gamma(A) \rightarrow \text{proj} \Gamma(A)$ is an auto-equivalence for each $n \in \mathbb{Z}$. This functor extends naturally to an auto-equivalence on the category of all right $\Gamma(A)$-modules. Then $K^n(A)$ is an invertible bimodule. The second statement follows from Lemma 3.6 and the fact that $\Sigma^n \Sigma^n$ is isomorphic to $\Sigma^{n+m}$. Here, we use [7, Proposition 12.9] implicitly. 

We now have Theorem B. Denote the functor $- \otimes_{\Gamma(A)} K^1(A): \text{proj} \Gamma(A) \rightarrow \text{proj} \Gamma(A)$ by $\Sigma_A$.

**Theorem 3.8.** Let $A$ be an artin algebra with radical square zero. Use the above notation. Then we have a triangle equivalence

$$
\mathcal{D}_{sg}(A) \simeq (\text{proj} \Gamma(A), \Sigma_A),
$$

which sends $q(A/r)$ to $\Gamma(A)$.

**Proof.** It follows from Proposition 3.5 and Lemma 3.6. 

Let $A$ be an artin algebra with radical square zero. For each $n \geq 1$, we consider the artin algebra $G^n = A/r \oplus r^{\otimes n}$, which is the trivial extension of the $A/r$-$A/r$-bimodule $r^{\otimes n}$ ([1, p.78]). All these algebras $G^n$ have radical square zero. The following observation seems to be of independent interest.
Proposition 3.9. Use the above notation. Then for each \( n \geq 1 \), we have a triangle equivalence
\[
\text{D}_{\text{sg}}(G^n) \simeq \text{proj } \Gamma(A), \Sigma^n_A.
\]
In particular, we have a triangle equivalence \( \text{D}_{\text{sg}}(A) \simeq \text{D}_{\text{sg}}(G^1) \).

Proof. Write \( G^n = A' \). Then from the very definition, we have a natural identification \( \Gamma(A') = \Gamma(A) \). Moreover, the \( \Gamma(A')-\Gamma(A') \)-bimodule \( K^1(A') \) corresponds to the \( \Gamma(A)-\Gamma(A) \)-bimodule \( K^n(A) \). Then by Lemma 3.6 \( \Sigma_A \) corresponds to \( \Sigma^n_A \). Then the result follows from Theorem 3.8 immediately. \( \square \)

Remark 3.10. We point out that for \( n \geq 2 \), \( \text{D}_{\text{sg}}(G^n) \) might not be triangle equivalent to \( \text{D}_{\text{sg}}(A) \), although the underlying categories are equivalent.

We conclude this section with an example.

Example 3.11. Let \( k \) be a field and let \( n \geq 1 \). Consider the algebra \( A = k[x_1, x_2, \cdots, x_n]/(x_i x_j, 1 \leq i, j \leq n) \), which is with radical square zero. We identify \( A/\mathfrak{r} \) with \( k \), and \( \mathfrak{r} \) with the \( n \)-dimensional \( k \)-space \( V = kx_1 \oplus kx_2 \oplus \cdots \oplus kx_n \). Consequently, for each \( i \geq 0 \), the algebra \( \text{End}_{A/\mathfrak{r}}(\mathfrak{r}^{\oplus i}) \) is isomorphic to \( \text{End}_{k}(V^{\oplus i}) \), which is identified with the \( n^i \times n^i \) total matrix algebra \( M_{n^i}(k) \). Then the associated regular algebra \( \Gamma(A) \) is isomorphic to the direct limit of the following chain of algebra embeddings
\[
k \rightarrow M_n(k) \rightarrow M_{n^2}(k) \rightarrow M_{n^3}(k) \rightarrow \cdots
\]
Here, for each algebra \( B, B \rightarrow M_n(B) \) is the algebra embedding sending \( b \) to \( bI_n \), with \( I_n \) the \( n \times n \) identity matrix.

We observe that \( \Gamma(A) \) is a simple algebra. We point out that this construction is classical; see [8, 19.26 B, Example]. The algebra \( A \) is non-noetherian for \( n \geq 2 \), while for \( n = 1 \), it is isomorphic to \( k \).

Let \( 1 \leq r, s \leq n \). Define \( E_{rs} : V \rightarrow V \) to be the linear map such that \( E_{rs}(x_i) = \delta_{is} x_r \), where \( \delta \) is the Kronecker symbol. Consider, for all \( i \geq 0 \), the linear maps \( - \otimes k E_{rs} : \text{End}_{k}(V^{\otimes i}) \rightarrow \text{End}_{k}(V^{\otimes i+1}) \). Taking the limit, we have the induced linear map \( - \otimes k E_{rs} : \Gamma(A) \rightarrow \Gamma(A) \) for each pair of \( r, s \). Then we have an isomorphism \( \sigma : M_n(\Gamma(A)) \rightarrow \Gamma(A) \) of algebras, which sends an \( n \times n \) matrix \( (a_{ij}) \) to \( \sum_{1 \leq i, j \leq n} a_{ij} \otimes k E_{ij} \).

The associated \( \Gamma(A)-\Gamma(A) \)-bimodule \( K^1(A) \) is described as follows. As a \( k \)-space, \( K^1(A) = \Gamma(A) \oplus \Gamma(A) \oplus \cdots \oplus \Gamma(A) \) with \( n \) copies of \( \Gamma(A) \). The left action is given by \( a(a_1, a_2, \cdots, a_n) = (aa_1, aa_2, \cdots, aa_n) \), while the right action is given by \( (a_1, a_2, \cdots, a_n)a = (a_1, a_2, \cdots, a_n)\sigma^{-1}(a) \).

We remark that the regular algebra \( \Gamma(A) \) is related to a quotient abelian category studied in [16], which might relate to the singularity category \( \text{D}_{\text{sg}}(A) \) via a version of Koszul duality.

4. One-point (co)extensions and cyclicizations of algebras

In this section, we prove that one-point extensions and coextensions of algebras preserve their singularity categories. We then introduce the notion of cyclicization of an algebra, which is a repeated operation to remove sources and sinks.
Let \( A \) be an artin algebra. Let \( D \) be a simple artin algebra, and let \( _AM_D \) be an \( A-D \)-bimodule, on which \( R \) acts centrally. The one-point extension of \( A \) by \( M \) is the upper triangular matrix algebra \( A[M] = \begin{pmatrix} A & M \\ 0 & D \end{pmatrix} \). A left \( A[M] \)-module is denoted by a column vector \( \begin{pmatrix} X \\ V \end{pmatrix} \), where \( X \) and \( V \) are a left \( A \)-module and \( D \)-module, respectively, and that \( \phi : M \otimes_D V \to X \) is a morphism of \( A \)-modules.

We sometimes suppress the morphism \( \phi \), when it is clearly understood. For details, we refer to [1, III.2].

Recall from [1, III.1] the notion of valued quiver \( Q_A \) for an artin algebra \( A \).

We observe that for the unique simple \( D \)-module \( S \), the corresponding \( A[M] \)-module \( \begin{pmatrix} 0 \\ S \end{pmatrix} \) is simple injective, which corresponds to a source in the valued quiver \( Q_A[M] \) of the one-point extension \( A[M] \). Indeed, this valued quiver is obtained from \( Q_A \) by adding this source together with some valued arrows starting at it.

One-point extensions of algebras preserve singularity categories. Observe the natural exact embedding \( i : A\text{-mod} \to A[M]\text{-mod} \), which sends \( A \) \( X \) to \( i(X) = \begin{pmatrix} X \\ 0 \end{pmatrix} \).

**Proposition 4.1.** Let \( A[M] \) be the one-point extension of \( A \) as above. Then the exact embedding \( i : A\text{-mod} \to A[M]\text{-mod} \) induces a triangle equivalence

\[
\mathcal{D}_{sg}(A) \cong \mathcal{D}_{sg}(A[M]).
\]

**Proof.** The exact functor \( i \) extends naturally to a triangle functor \( \tilde{i} : D^b(A\text{-mod}) \to D^b(A[M]\text{-mod}) \). We observe that \( i(A) \) is projective, and then \( \tilde{i} \) sends perfect complexes to perfect complexes. Then it induces a triangle functor \( \tilde{\epsilon}_* : \mathcal{D}_{sg}(A) \to \mathcal{D}_{sg}(A[M]) \). We claim that \( \tilde{\epsilon}_* \) an equivalence.

For the claim, recall that the functor \( i \) admits a left adjoint \( j : A[M]\text{-mod} \to A\text{-mod} \) which sends \( \begin{pmatrix} X \\ V \end{pmatrix} \) to \( X/\text{Im} \phi \). Observe that the corresponding counit

\[
\begin{pmatrix} X \\ V \end{pmatrix} \sim \text{Id}_{A\text{-mod}}
\]

is an isomorphism. One checks that the cohomological dimension ([11, p.57]) of the functor \( j \) is at most one. In particular, the left derived functor \( L^b j : D^b(A[M]\text{-mod}) \to D^b(A\text{-mod}) \) is defined. Moreover, we have the adjoint pair \( (L^b j, i_*) \), and that the counit is an isomorphism. Since the functor \( j \) sends projective modules to projective modules, the functor \( L^b j \) preserves perfect complexes. Then it induces a triangle functor \( L^b j : \mathcal{D}_{sg}(A[M]) \to \mathcal{D}_{sg}(A) \). Moreover, we have the induced adjoint pair \( (L^b j, \tilde{\epsilon}_*) \), whose counit \( (L^b j)(\tilde{\epsilon}_*) \sim \text{Id}_{\mathcal{D}_{sg}(A)} \) is an isomorphism; see [14, Lemma 1.2]. In particular, the functor \( \tilde{\epsilon}_* \) is fully faithful.

It remains to show the denseness of \( \tilde{\epsilon}_* \). We now view the essential image \( \text{Im} \tilde{\epsilon}_* \) of \( \tilde{\epsilon}_* \) as a full triangulated subcategory of \( \mathcal{D}_{sg}(A[M]) \). It suffices to show that
for each $A[M]$-module $\begin{pmatrix} X \\ V \end{pmatrix}$, its image in $D_{sg}(A[M])$ lies in $\text{Im} \ i$; see Lemma 2.1. Observe that $\Omega(\begin{pmatrix} 0 \\ V \end{pmatrix})$ lies in $\text{Im} \ i$, and then by Lemma 2.2, $q(\begin{pmatrix} X \\ V \end{pmatrix})$ lies in $\text{Im} \ i$. The following natural exact sequence induces a triangle in $D_{sg}(A[M])$

$$0 \rightarrow i(X) \rightarrow \begin{pmatrix} X \\ V \end{pmatrix} \rightarrow \begin{pmatrix} 0 \\ V \end{pmatrix} \rightarrow 0.$$ 

This triangle implies that $q(\begin{pmatrix} X \\ V \end{pmatrix})$ lies in $\text{Im} \ i$. Then we are done. 

Let $D$ be a simple artin algebra, and let $D \cdot N$ be a $D$-$A$-bimodule, on which $R$ acts centrally. The one-point coextension of $A$ by $N$ is the upper triangular matrix algebra $[N]A = \begin{pmatrix} D & N \\ 0 & A \end{pmatrix}$. A left $[N]A$-module is written as $\begin{pmatrix} V \\ X \end{pmatrix}_A$, where $V$ and $X$ are a left $D$-module and $A$-module, respectively, and that $\phi: M \otimes_A X \rightarrow V$ is a morphism of $D$-modules. The valued quiver $Q_{[N]A}$ is obtained from $Q_A$ by adding a sink together with some valued arrows ending at it, where the sink corresponds to the simple projective $[N]A$-module $\begin{pmatrix} S \\ 0 \end{pmatrix}$ for a simple $D$-module $S$.

For the one-point coextension $[N]A$, we have an exact embedding $i: A\text{-mod} \rightarrow [N]A\text{-mod}$, which sends $A X$ to $i(X) = \begin{pmatrix} 0 \\ X \end{pmatrix}$.

The following result is similar to Proposition 4.1, while the proof is simpler. This result is closely related to [4, Theorem 4.1(1)].

**Proposition 4.2.** Let $[N]A$ be the one-point coextension as above. Then the embedding $i: A\text{-mod} \rightarrow [N]A\text{-mod}$ induces a triangle equivalence

$$D_{sg}(A) \simeq D_{sg}([N]A).$$

**Proof.** We observe that $i(A)$ has projective dimension at most one. Then the obviously induced functor $i_*: D^b(A\text{-mod}) \rightarrow D^b([N]A\text{-mod})$ preserves perfect complexes, and it induces the required functor $i_*: D_{sg}(A) \rightarrow D_{sg}([N]A)$. The functor $i$ admits an exact left adjoint $j: [N]A\text{-mod} \rightarrow A\text{-mod}$, which sends $\begin{pmatrix} V \\ X \end{pmatrix}$ to $X$; moreover, $j$ preserves projective modules. Then it induces a triangle functor $j_*: D_{sg}([N]A) \rightarrow D_{sg}(A)$, which is left adjoint to $i_*$. Then as in the proof of Proposition 4.1, we have that $i_*$ is fully faithful. The denseness of $i_*$ follows from the natural exact sequence

$$0 \rightarrow \begin{pmatrix} V \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} V \\ X \end{pmatrix} \rightarrow i(X) \rightarrow 0,$n

for each $[N]A$-module $\begin{pmatrix} V \\ X \end{pmatrix}$, and that the $[N]A$-module $\begin{pmatrix} V \\ 0 \end{pmatrix}$ is projective. We omit the details. 

\[\square\]
We use the above two propositions to reduce the study of the singularity category of arbitrary artin algebras to cyclic-like ones. Let $A$ be an artin algebra. Consider the valued quiver $Q_A$. A vertex $e$ is called cyclic provided that there is an oriented cycle containing it, and the corresponding simple $A$-module is called cyclic. More generally, a vertex $e$ is called cyclic-like provided that there is a path through $e$, which starts with a cyclic vertex and ends at a cyclic vertex, while the corresponding simple $A$-module is called cyclic-like. An artin algebra $A$ is called cyclic-like provided that its valued quiver $Q_A$ is cyclic-like. This is equivalent to that $A$ has neither simple projective nor simple injective modules.

For an artin algebra $A$, its cyclicization is an artin algebra $A_c$ which is either simple or cyclic-like, such that there is a sequence $A_c = A_0, A_1, \cdots, A_r = A$ satisfying that each $A_{i+1}$ is a one-point (co)extension of $A_i$. The following is an immediate consequence of the definition.

**Lemma 4.3.** Let $A$ be an artin algebra with its cyclicization $A_c$. Then we have a triangle equivalence 

$$D_{sg}(A_c) \simeq D_{sg}(A).$$

**Proof.** Apply Propositions 4.1 and 4.2, repeatedly. $\square$

The following result seems to be well known.

**Proposition 4.4.** The following statements hold.

1. Each artin algebra has a cyclicization.
2. Let $A_c$ and $A_{c'}$ be two cyclicizations of $A$. Then if $A_c$ is simple, so is $A_{c'}$. Otherwise, we have an isomorphism $A_c \simeq A_{c'}$ of algebras.

**Proof.** (1) It follows from the well-known fact that the existence of a simple injective (resp. projective) module of $A$ implies that $A$ is a one-point extension (resp. coextension) of $A'$. Moreover, the valued quiver $Q_{A'}$ of $A'$ is obtained from the one of $A$ by deleting the relevant source (resp. sink).

(2) The first statement follows from the observation that passing from $A$ to $A'$ in (1), the set of cyclic-like vertices stays the same. For the isomorphism of algebras, it suffices to observe that $A_c$-mod is equivalent to the smallest Serre subcategory ([7, Chapter 15]) of $A$-mod containing the cyclic-like simple $A$-modules $S$; moreover, the multiplicity of $P_{A_c}(S)$ in the indecomposable decomposition of $A_c$ equals the multiplicity of $P(S)$ in the one of $A$. Here, $P(S)$ and $P_{A_c}(S)$ denote the projective cover of $S$ as an $A$-module and $A_c$-module, respectively. $\square$

5. **Hom-finiteness of singularity categories**

In this section, we study the Hom-finiteness of the singularity category of an artin algebra with radical square zero, and prove Theorem C. Throughout, $A$ is an artin $R$-algebra such that its Jacobson radical $r$ satisfies $r^2 = 0$. Recall that in this case, the syzygy $\Omega(X)$ of any $A$-module $X$ is semisimple.
LEMA 5.1. Suppose that $A$ is cyclic-like. Then we have

1. each simple $A$-module has infinite projective dimension;
2. for each $i \geq 0$, the algebra homomorphism $\text{End}_{A/r}(r^{\otimes i}) \to \text{End}_{A/r}(r^{\otimes i+1})$ induced by $r \otimes \text{id}_{A/r}$ is injective.

Proof. (1) Recall that a cyclic-like algebra does not have simple projective modules. Then the statement follows from the observation that for a simple module $S$ with finite projective dimension, we have that $\text{proj.dim} \Omega(S) = \text{proj.dim} S - 1$.

(2) We recall that $A$-ssmod is the full subcategory of $A$-mod consisting of semisimple modules. Then by (1), $A$-ssmod is naturally equivalent to $A$-ssmod, and the syzygy functor $\Omega: A$-ssmod $\to A$-ssmod is faithful. Now the result follows from Lemma 3.1.

□

Recall that the singularity category $D_{sg}(A)$ is naturally $R$-linear. We are interested in the problem when it is Hom-finite, that is, all the Hom sets are finitely generated $R$-modules.

THEOREM 5.2. Let $A$ be an artin algebra with radical square zero. Then the following statements are equivalent:

1. the singularity category $D_{sg}(A)$ is Hom-finite;
2. the associated regular algebra $\Gamma(A)$ is semisimple;
3. the cyclicization $A_c$ of $A$ is either simple or isomorphic to a finite product of self-injective algebras.

We point out that the cyclicization $A_c$ of $A$ is necessarily with radical square zero. Recall that an indecomposable non-simple artin algebra with radical square zero is self-injective if and only if its valued quiver is an oriented cycle with the trivial valuation; see [1, Proposition IV.2.16] or the proof of [5, Corollary 1.3]. Then the statement (3) above is equivalent to the corresponding one in Theorem C.

Proof. Recall from Proposition 3.2 the isomorphism $\Gamma(A) \simeq \text{End}_{D_{sg}(A)}(q(A/r))$. Then we have the implication “(1)$\Rightarrow$(2)”, since an artin regular algebra is necessarily semisimple.

For “(2)$\Rightarrow$(1)”, consider the cyclicization $A_c$ of $A$, whose Jacobson radical is denoted by $r_c$. Then by Lemma 4.3 we have an equivalence $D_{sg}(A_c) \simeq D_{sg}(A)$. Applying Proposition 3.5 we have an equivalence $\text{proj.}\Gamma(A_c) \cong \text{proj.}\Gamma(A)$, that is, $\Gamma(A_c)$ and $\Gamma(A)$ are Morita equivalent. Then $\Gamma(A_c)$ is also semisimple. Recall that $\Gamma(A_c) = \lim_{\longrightarrow} \text{End}_{A_c/r_c}(r^{\otimes i}_c)$. By Lemma 5.1 all the canonical maps $\text{End}_{A_c/r_c}(r^{\otimes i}_c) \to \Gamma(A_c)$ are injective. Recall that for a semisimple algebra, the number of pairwise orthogonal idempotents is bounded. Then the $R$-lengths of the algebras $\text{End}_{A_c/r_c}(r^{\otimes i}_c)$ are uniformly bounded. Consequently, the algebra $\Gamma(A)$ is an artin $R$-algebra. By Proposition 3.5 the singularity category $D_{sg}(A_c)$ is Hom-finite. Then we are done by Lemma 4.3.
Recall from [15, Theorem 2.1] that the singularity category of a self-injective algebra is equivalent to its stable category. In particular, it is Hom-finite. Then the implication “(3)⇒(1)” follows from Lemma 4.3.

It remains to show “(1)⇒(3)”. Without loss of generality, we assume that the algebra $A$ is cyclic-like such that $D_{sg}(A)$ is Hom-finite. We will show that $A$ is self-injective.

We claim that the syzygy $\Omega(S)$ of any cyclic simple $A$-module $S$ is simple. Then there is only one arrow starting at $S$ in $Q_A$, which is valued by $(1, b)$ for some natural number $b$. Since $Q_A$ is cyclic-like, this forces that $Q_A$ is a disjoint union of oriented cycles. In each oriented cycle, every arrow has valuation $(1, b_i)$ for some $b_i$. Then the symmetrization condition implies that all these $b_i$’s are necessarily one; compare the proof of [1, Proposition VIII. 6.4]. As we point out above, this implies that $A$ is self-injective.

We prove the claim. Since by Corollary 2.4 $D_{sg}(A)$ is idempotent split, we have that $D_{sg}(A)$ is a Krull-Schmidt category ([6, Appendix A]). In particular, each object is uniquely decomposed as a direct sum of finitely many indecomposable objects. We observe that for each semisimple module $X$, $l_X \leq l_\Omega(X)$. Here, $l$ denotes the composition length. Consider a cyclic simple $A$-module $S$, and take a path $S = S_1 \rightarrow S_2 \rightarrow \cdots \rightarrow S_r \rightarrow S_{r+1} = S$ in $Q_A$. Assume on the contrary that $l_\Omega(S) \geq 2$. Then we have $\Omega(S) = S_2 \oplus X$ for some nonzero semisimple module $X$. Observe that $S$ is a direct summand of $\Omega^{r-1}(S_2)$, and then we have $\Omega^r(S) = S \oplus X'$ for a nonzero semisimple module $X'$. Consequently, we have $\Omega^{nr}(S) = S \oplus X' \oplus \Omega^r(X') \oplus \cdots \oplus \Omega^{(n-1)r}(X')$. Then the lengths of the semisimple modules $\Omega^{nr}(S)$ tend to the infinity, when $n$ goes to the infinity.

By Lemma 5.1(1), $q(T)$ is not zero for any simple $A$-module $T$. Recall from Lemma 2.2 that $q(S) \simeq q(\Omega^{nr}(S))[nr]$, and then $q(S) \simeq q(S)[nr] \oplus q(X')[nr] \oplus q(\Omega^r(X'))[nr] \oplus \cdots \oplus q(\Omega^{(n-1)r}(X'))[nr]$ for each $n \geq 1$. This contradicts to the Krull-Schmidt property of $D_{sg}(A)$, and we are done with the claim. □

ACKNOWLEDGEMENTS. The author thanks Professor Zhaoyong Huang, Professor Yu Ye and Dr. Longgang Sun for their helpful comments.

REFERENCES

[1] M. Auslander, I. Reiten and S.O. Smalø, Representation Theory of Artin Algebras, Cambridge Studies in Adv. Math. 36, Cambridge Univ. Press, Cambridge, 1995.

[2] A. Beligiannis, The homological theory of contravariantly finite subcategories: Auslander-Buchweitz contexts, Gorenstein categories and (co-)stabilization, Comm. Algebra 28(10) (2000), 4547–4596.

[3] R.O. Buchweitz, Maximal Cohen-Macaulay Modules and Tate Cohomology over Gorenstein Rings, Unpublished Manuscript, 1987. Available at: http://hdl.handle.net/1807/16682.

[4] X.W. Chen, Singularity categories, Schur functors and triangular matrix rings, Algebr. Represent. Theor. 12 (2009), 181–191.
[5] X.W. Chen, *Algebras with radical square zero are either self-injective or CM-free*, Proc. Amer. Math. Soc. 140 (1) (2012), 93–98.

[6] X.W. Chen, Y. Ye and P. Zhang, *Algebras of derived dimension zero*, Comm. Algebra 36 (2008), 1–10.

[7] C. Faith, *Algebra I*, Rings, Modules and Categories, Springer-Verlag, Berlin Heidelberg New York, 1973.

[8] C. Faith, *Algebra II*, Ring Theory, Springer-Verlag, Berlin Heidelberg New York, 1976.

[9] D. Happel, *Triangulated Categories in the Representation Theory of Finite Dimensional Algebras*, London Math. Soc., Lecture Notes Ser. 119, Cambridge Univ. Press, Cambridge, 1988.

[10] D. Happel, *On Gorenstein algebras*, In: Progress in Math. 95, Birkhäuser Verlag, Basel, 1991, 389–404.

[11] R. Hartshorne, *Duality and Residue*, Lecture Notes in Math. 20, Springer, Berlin, 1966.

[12] B. Keller and D. Vossieck, *Sous les catégories dérivées*, C.R. Acad. Sci. Paris, t. 305 Série I (1987) 225–228.

[13] H. Krause, *The stable derived category of a noetherian scheme*, Compositio Math. 141 (2005), 1128–1162.

[14] D. Orlov, *Triangulated categories of singularities and D-branes in Landau-Ginzburg models*, Trudy Steklov Math. Institute 204 (2004), 240–262.

[15] J. Rickard, *Derived categories and stable equivalence*, J. Pure Appl. Algebra 61 (1989), 303–317.

[16] S.P. Smith, *The non-commutative scheme having a free algebra as a homogeneous coordinate ring*, arXiv:1104.3822v1.

---

Xiao-Wu Chen
Department of Mathematics
University of Science and Technology of China
Wu Wen-Tsun Key Laboratory of Mathematics, USTC
Chinese Academy of Sciences
PR China
http://mail.ustc.edu.cn/~xwchen

Documenta Mathematica 16 (2011) 921–936
Jumps and Monodromy of Abelian Varieties

Lars Halvard Halle and Johannes Nicaise

Received: October 10, 2010
Revised: August 19, 2011
Communicated by Peter Schneider

Abstract. We prove a strong form of the motivic monodromy conjecture for abelian varieties, by showing that the order of the unique pole of the motivic zeta function is equal to the maximal rank of a Jordan block of the corresponding monodromy eigenvalue. Moreover, we give a Hodge-theoretic interpretation of the fundamental invariants appearing in the proof.

2010 Mathematics Subject Classification: MSC2000: 11G10, 14D05, 14D07

1. Introduction

Let $K$ be a henselian discretely valued field with algebraically closed residue field $k$, and let $A$ be a tamely ramified abelian $K$-variety of dimension $g$. In [14], we introduced the motivic zeta function $Z_A(T)$ of $A$. It is a formal power series over the localized Grothendieck ring of $k$-varieties $M_k$, and it measures the behaviour of the Néron model of $A$ under tame base change. We showed that $Z_A(L^{-s})$ has a unique pole, which coincides with Chai’s base change conductor $c(A)$ of $A$, and that the order of this pole equals $1 + t_{pot}(A)$, where $t_{pot}(A)$ denotes the potential toric rank of $A$. Moreover, we proved that for every embedding of $\mathbb{Q}_{\ell}$ in $\mathbb{C}$, the value $\exp(2\pi c(A)i)$ is an eigenvalue of the tame monodromy transformation on the $\ell$-adic cohomology of $A$ in degree $g$. The main ingredient of the proof is Edixhoven’s filtration on the special fiber of the Néron model of $A$ [12].

As we’ve explained in [14], this result is a global version of Denef and Loeser’s motivic monodromy conjecture for hypersurface singularities in characteristic zero. Denef and Loeser’s conjecture relates the poles of the motivic zeta function of the singularity to monodromy eigenvalues on the nearby

---

1The first author was partially supported by the Fund for Scientific Research-Flanders (G.0318.06) and by DFG under grant Hu 337/6-1. The second author was partially supported by the Fund for Scientific Research - Flanders (G.0415.10) and ANR-06-BLAN-0183.
cohomology. It is a motivic generalization of a conjecture of Igusa’s for the $p$-adic zeta function, which relates certain arithmetic properties of polynomials $f$ in $\mathbb{Z}[x_1, \ldots, x_n]$ (namely, the asymptotic behaviour of the number of solutions of the congruence $f \equiv 0$ modulo powers of a prime $p$) to the structure of the singularities of the complex hypersurface defined by the equation $f = 0$. The conjectures of Igusa and Denef-Loeser have been solved, for instance, in the case $n = 2$ [16][23], but the general case remains wide open. We refer to [20] for a survey.

There also exists a stronger form of Igusa’s conjecture, which says that the real parts of the poles of the $p$-adic zeta function of $f$ are roots of the Bernstein polynomial $b_f(s)$ of $f$, and that the order of each pole is at most the multiplicity of the corresponding root of $b_f(s)$. This stronger conjecture also has a motivic generalization, replacing the $p$-adic zeta function by the motivic zeta function, and taking for $f$ any polynomial with coefficients in a field of characteristic zero (or, more generally, any regular function on a smooth algebraic variety over a field of characteristic zero).

It is well-known that, for every complex polynomial $f$ and every root $\alpha$ of $b_f(s)$, the value $\alpha' = \exp(2\pi i \alpha)$ is a monodromy eigenvalue on the nearby cohomology $R\psi_f(\mathbb{C})_x$ of $f$ at some closed point $x$ of the zero locus $H_f$ of $f$ [15][17]. Moreover, if $H_f$ has an isolated singularity at $x$, then the multiplicity $m_\alpha$ of $\alpha$ as a root of the local Bernstein polynomial $b_{f,x}(s)$ of $f$ at $x$ is closely related to the maximal size $m_{\alpha'}$ of the Jordan blocks with eigenvalue $\alpha'$ of the monodromy transformation on $R^{n-1}\psi_f(\mathbb{C})_x$. In particular, $m_\alpha \leq m_{\alpha'}$ if $\alpha \notin \mathbb{Z}$ [17, 7.1].

The aim of the present paper is twofold. First, we prove a strong form of the motivic monodromy conjecture for abelian varieties. There is no good notion of Bernstein polynomial in our setting, but we can look at the size of the Jordan blocks. We show that the order $1 + t_{\text{pot}}(A)$ of the unique pole $s = c(A)$ of the motivic zeta function $Z_A(\mathbb{L}^{-s})$ is equal to the maximal rank of a Jordan block of the corresponding monodromy eigenvalue on the degree $g$ cohomology of $A$ (Theorem 5.9). Next, we use the theory of Néron models of variations of Hodge structures to give a Hodge-theoretic interpretation of the jumps in Edixhoven’s filtration. This is done in Theorems 6.2 and 6.3. In [14, 2.7], we speculated on a generalization of the monodromy conjecture to Calabi-Yau varieties over $\mathbb{C}((t))$ (i.e., smooth, proper, geometrically connected $\mathbb{C}((t))$-varieties with trivial canonical sheaf); we hope that the translation of Edixhoven’s invariants to Hodge theory will help to extend the proof of the monodromy conjecture to that setting.

In order to obtain these results, we divide Edixhoven’s jumps into three types: toric, abelian, and dual abelian. The basic properties of these types are discussed in Section 3. Not all of these results are used in the proofs of the main results of the paper. We include them for the sake of completeness and because we believe that they are of independent interest. The reader who is only interested in Theorems 4.4, 5.9, 6.2 and 6.3 may skip Lemma 3.2, Proposition 3.4 and all the results in Section 3 after Proposition 3.5.
The different types of jumps are related to the monodromy transformation on the Tate module of $A$ in Section 4. Toric jumps correspond to monodromy eigenvalues with Jordan block of size two, and the abelian and dual abelian jumps to monodromy eigenvalues with Jordan block of size one (Theorem 4.4). If $K = \mathbb{C}((t))$, then the abelian and dual abelian jumps can be distinguished by looking at the Hodge type in the limit mixed Hodge structure (Theorem 6.3).

Acknowledgements. We are grateful to the referee for carefully reading the manuscript and suggesting several improvements to the paper. The second author is indebted to Antoine Chambert-Loir for helpful suggestions concerning Section 5.1.

2. Preliminaries and notation

We denote by $R$ a Henselian discrete valuation ring, with quotient field $K$ and algebraically closed residue field $k$. We denote by $K^a$ an algebraic closure of $K$, by $K^s$ the separable closure of $K$ in $K^a$, and by $K^t$ the tame closure of $K$ in $K^s$. We fix a topological generator $\sigma$ of the tame monodromy group $G(K^t/K)$. We denote by $p$ the characteristic exponent of $k$, and by $N'\mathbb{Z}$ the set of integers $d > 0$ prime to $p$. We denote by $(\cdot)_s : (R - \text{Schemes}) \to (k - \text{Schemes}) : X \mapsto X_s = X \times_R k$

the special fiber functor.

For every abelian variety $B$ over a field $F$, we denote its dual abelian variety by $B^\vee$. For every abelian $K$-variety $A$ with Néron model $\mathcal{A}$, we denote by $t(A)$, $u(A)$ and $a(A)$ the reductive, resp. unipotent, resp. abelian rank of $\mathcal{A}_k$. We call these values the toric, resp. unipotent, resp. abelian rank of $A$. Obviously, their sum equals the dimension of $A$.

By Grothendieck’s semi-stable reduction theorem, there exists a finite extension $K'$ of $K$ in $K^s$ such that $A \times_K K'$ has semi-abelian reduction [2, IX.3.6]. This means that the identity component of the special fiber of its Néron model is a semi-abelian $k$-variety; equivalently, $u(A \times_K K') = 0$. The value $t_{\text{pot}}(A) = t(A \times_K K')$ is called the potential toric rank of $A$, and the value $a_{\text{pot}}(A) = a(A \times_K K')$ the potential abelian rank. It follows from [2, IX.3.9] that these values are independent of the choice of $K'$. We say that $A$ is tamely ramified if we can take for $K'$ a tame finite extension of $K$ (since $k$ is algebraically closed, this means that the degree $[K' : K]$ is prime to $p$).

For every scheme $S$, every $S$-group scheme $G$ and every integer $n > 0$, we denote by $n_G : G \to G$ the multiplication by $n$, and by $\pi_G$ its kernel. If $S$ is a set, and $g : S \to \mathbb{R}$ a function with finite support, we set

$$\|g\| = \sum_{s \in S} g(s).$$

We denote the support of $g$ by $\text{Supp}(g)$.
Definition 2.1. For every function
\[ f : \mathbb{Q}/\mathbb{Z} \to \mathbb{R} \]
we define its reflection
\[ f^* : \mathbb{Q}/\mathbb{Z} \to \mathbb{R} \]
by
\[ f^*(x) = f(-x). \]
We call \( f \) complete if for every \( x \in \mathbb{Q}/\mathbb{Z} \), the value \( f(x) \) only depends on the order of \( x \) in the group \( \mathbb{Q}/\mathbb{Z} \). We say that \( f \) is semi-complete if \( f + f^* \) is complete.

Consider a function
\[ f : \mathbb{Q}/\mathbb{Z} \to \mathbb{N} \]
and assume that there exists an element \( e \) of \( \mathbb{Z}_{>0} \) such that \( \text{Supp}(f) \) is contained in \( ((1/e)\mathbb{Z})/\mathbb{Z} \). Let \( F \) be any algebraically closed field such that \( e \) is prime to the characteristic exponent \( p' \) of \( F \). For each generator \( \zeta \) of \( \mu_e(F) \), we put
\[ P_{f,\zeta}(t) = \prod_{i \in ((1/e)\mathbb{Z})/\mathbb{Z}} (t - \zeta^i e)^{f(i)} \]
in \( F[t] \). For each integer \( d > 0 \), we denote by \( \Phi_d(t) \) the cyclotomic polynomial whose roots are the primitive \( d \)-th roots of unity. We say that \( \Phi_d(t) \) is \( F \)-tame if \( d \) is prime to \( p' \).

Lemma 2.2. The function \( f \) is complete if and only if for some generator \( \zeta \) of \( \mu_e(F) \), the polynomial \( P_{f,\zeta}(t) \) is the image of a product \( Q_f(t) \) of \( F \)-tame cyclotomic polynomials under the unique ring morphism
\[ \rho : \mathbb{Z}[t] \to F[t] \]
mapping \( t \) to \( t \). If \( f \) is complete, then \( P_{f,\zeta}(t) \) is independent of \( \zeta \) and \( e \), and \( Q_f(t) \) is unique. In that case, if we choose a primitive \( e \)-th root of unity \( \xi \) in an algebraic closure \( \mathbb{Q}^a \) of \( \mathbb{Q} \), then
\[ Q_f(t) = \prod_{i \in ((1/e)\mathbb{Z})/\mathbb{Z}} (t - \xi^i e)^{f(i)}. \]

Proof. First, assume that \( f \) is complete, and put
\[ Q_f(t) = \prod_{i \in ((1/e)\mathbb{Z})/\mathbb{Z}} (t - \xi^i e)^{f(i)} \]
for some primitive \( e \)-th root of unity \( \xi \) in \( \mathbb{Q}^a \). Then \( Q_f(t) \) is a product of \( F \)-tame cyclotomic polynomials, because \( f \) is complete and \( e \) is prime to \( p' \). There is a unique ring morphism
\[ \tilde{\rho} : \mathbb{Z}[\xi][t] \to F[t] \]
that maps \( \xi \) to \( \zeta \) and \( t \) to \( t \). We clearly have \( \tilde{\rho}(Q_f(t)) = P_{f,\zeta}(T) \). Since \( Q_f(t) \) belongs to \( \mathbb{Z}[t] \), it follows that \( \rho(Q_f(t)) = P_{f,\zeta}(T) \) so that \( P_{f,\zeta}(t) \) does not depend on \( \zeta \). Uniqueness of \( Q_f(t) \) follows from [14, 5.10].
Conversely, if $P_{f,\zeta}(t)$ is the image under $\rho$ of a product $Q(t)$ of $F$-tame cyclotomic polynomials, then it is easily seen that $f$ is complete. $\square$

3. Toric and abelian multiplicity

3.1. Galois action on Néron models. Let $A$ be a tamely ramified abelian $K$-variety of dimension $g$, and let $K'$ be a finite extension of $K$ in $K'$ such that $A' = A \times_K K'$ has semi-abelian reduction. We denote by $R'$ the integral closure of $R$ in $K'$, and by $\mathfrak{m}'$ the maximal ideal of $R'$. We put $d = [K' : K]$.

We denote by $\mu$ the Galois group $G(K'/K)$, and we let $\mu$ act on $K'$ from the left. The action of $\zeta \in \mu$ on $\mathfrak{m}'/(\mathfrak{m}')^2$ is multiplication by $\iota(\zeta)$, for some element $\iota(\zeta)$ in the group $\mu_d(k)$ of $d$-th roots of unity in $k$, and the map

$$\mu \to \mu_d(k) : \zeta \mapsto \iota(\zeta)$$

is an isomorphism.

We denote by $A$ and $A'$ the Néron models of $A$, resp. $A'$. By the universal property of the Néron model, there exists a unique morphism of $R'$-group schemes

$$h : A \times_R R' \to A'$$

that extends the canonical isomorphism between the generic fibers. It induces an injective morphism of free rank $g$ $R'$-modules

$$\text{Lie}(h) : \text{Lie}(A \times_R R') \to \text{Lie}(A').$$

**Definition 3.1 (Chai [5]).** The base change conductor $c(A)$ of $A$ is $[K' : K]^{-1}$ times the length of the cokernel of $\text{Lie}(h)$.

The definition does not require that $A$ is tamely ramified. The base change conductor is a positive rational number, independent of the choice of $K'$. It vanishes if and only if $A$ has semi-abelian reduction.

The right $\mu$-action on $A'$ extends uniquely to a right $\mu$-action on $A'$ such that the structural morphism

$$A' \to \text{Spec } R'$$

is $\mu$-equivariant. We denote by

$$(3.1) \quad 0 \to T \to (A'_o)^\mu \to B \to 0$$

the Chevalley decomposition of $(A'_o)^\mu$, with $T$ a $k$-torus and $B$ an abelian $k$-variety. There exist unique right $\mu$-actions on $T$, resp. $B$, such that (3.1) is $\mu$-equivariant. The right $\mu$-action on $B$ induces a left $\mu$-action on the dual abelian variety $B^\vee$.

**Lemma 3.2.** (1) The complex

$$(3.2) \quad 0 \to T^\mu \to ((A'_o)^\mu)^\mu \to B^\mu,$$

obtained from (3.1) by taking $\mu$-invariants, is an exact complex of smooth group schemes over $k$.

Taking identity components, we get a complex

$$(3.3) \quad 0 \to (T^\mu)^o \to ((A'_o)^\mu)^o \to (B^\mu)^o \to 0$$

Documenta Mathematica 16 (2011) 937–968
of smooth group schemes over $k$ that is exact at the left and at the right. The quotient

$$B' = (\left( A_s' \right)_{\mu}^o / \left( T_{\mu}^o \right)$$

is an abelian $k$-variety, and the natural morphism $f : B' \to (B^o)^{\mu}$ is a separable isogeny.

(2) If we denote by $h$ the unique morphism

$$h : A \times_R R' \to A'$$

extending the natural isomorphism between the generic fibers, then the $k$-morphism $h_s : A_s \to A'_s$ factors through a morphism

$$g : A_s \to (A'_s)^{\mu}.$$ 

The morphism $g$ is smooth and surjective, and its kernel is a connected unipotent smooth algebraic $k$-group. The identity component $((A'_s)^{\mu})^o$ is semi-abelian.

Proof. (1) It follows from [12, 3.4] that the group schemes in (3.2) are smooth over $k$. Exactness of (3.2) is clear. The morphism

$$\alpha : (A'_s)^o \to B$$

is smooth, since $T$ is smooth over $k$ [1, VIb.9.2]. It follows from [12, 3.5] that

$$\alpha^\mu : ((A'_s)^{\mu})^o \to B^\mu$$

is smooth, as well. Taking identity components in (3.2), we get a complex

$$(T^o) \xrightarrow{\beta} (A'_s)^o \xrightarrow{\gamma = (\alpha^\mu)^o} (B^o)^{\mu}$$

of smooth group schemes over $k$. It is obvious that $\beta$ is a monomorphism, and thus a closed immersion [1, VIb.1.4.2]. Surjectivity of $\gamma$ follows from [1, VIb.3.11], since $\gamma$ is smooth, and thus flat. We put

$$B' = ((A'_s)^{\mu})^o / (T^{\mu})^o.$$ 

This is a connected smooth algebraic $k$-group, by [1, VIb.3.2 and VIb.9.2]. Consider the natural morphism $f : B' \to (B^\mu)^{\mu}$. It is surjective, because $\gamma$ is surjective. The dimension of $B'$ is equal to

$$\dim (A'_s)^{\mu} - \dim T^{\mu},$$

which is at most $\dim B^\mu$ by exactness of (3.2). Surjectivity of $f$ then implies that $B'$ and $B^\mu$ must have the same dimension, so that $f$ has finite kernel. Thus $f$ is an isogeny and $B'$ is an abelian variety. The kernel of $f$ is canonically isomorphic to

$$\ker(\gamma) / (T^{\mu})^{\mu}.$$ 

Since $\gamma$ is smooth, we know that $\ker(\gamma)$ is smooth over $k$, so that $\ker(f)$ is smooth over $k$, by [1, VIb.9.2]. Hence, $f$ is a separable isogeny.

(2) Since $h$ is $\mu$-equivariant, and $\mu$ acts trivially on the special fiber $A_s$ of $A \times_R R'$, the morphism $h_s$ factors through a morphism $g : A_s \to (A'_s)^{\mu}$. By [12, 5.3], the morphism $g$ is smooth and surjective, and its kernel is a connected
unipotent smooth algebraic $k$-group. By [12, 3.4], $((A'_s)')^0$ is a connected smooth closed $k$-subgroup scheme of the semi-abelian $k$-group scheme $(A'_{s})^0$, so that $((A')^0)^0$ is semi-abelian by [14, 5.2].

3.2. Multiplicity functions. Fix an element $e \in \mathbb{N}$. For every finite dimensional $k$-vector space $V$ with a right $\mu(e)$-action $\ast : V \times \mu(e) \to V : (v, \zeta) \mapsto v \ast \zeta$ and for every integer $i$ in $\{0, \ldots, e-1\}$, we denote by $V[i]$ the maximal subspace of $V$ such that $v \ast \zeta = \zeta^i \cdot v$ for all $\zeta \in \mu(e)$ and all $v \in V[i]$. We define the multiplicity function $m_{V,\mu(e)} : \mathbb{Q}/\mathbb{Z} \to \mathbb{N}$ by

$$m_{V,\mu(e)}(i/e) = \dim(V[i]) \quad \text{for } i \in \{0, \ldots, e-1\}$$

$$m_{V,\mu(e)}(x) = 0 \quad \text{if } x \notin ((1/e)\mathbb{Z})/\mathbb{Z}$$

Note that $m_{V,\mu(e)}$ determines the $k[\mu(e)]$-module $V$ up to isomorphism, since the order of $\mu(e)$ is invertible in $k$.

In an analogous way, we define the multiplicity function $m_{\mu(e),W}$ for a finite dimensional $k$-vector space $W$ with left $\mu(e)$-action. The inverse of the right $\mu(e)$-action on $V$ is the left action

$$\mu(e)(k) \times V \to V : (\zeta, v) \mapsto v \ast \zeta^{-1}.$$ Its multiplicity function $m_{\mu(e),V}$ is equal to the reflection $(m_{V,\mu(e)})^*$ of the multiplicity function $m_{V,\mu(e)}$.

Let $A$ be a tamely ramified abelian $K$-variety. We adopt the notations of Section 3.1. In the set-up of (3.1), the group $\mu \cong \mu_d(k)$ acts on the $k$-vector spaces Lie($T$), Lie($A'_s$) and Lie($B$) from the right, and on Lie($B'$) from the left (via the dual action of $\mu$ on $B'$). Hence, we can state the following definitions.

**Definition 3.3.** We define the toric multiplicity function $m^t_A$ of $A$ by

$$m^t_A = m_{\text{Lie}(T),\mu}.$$ We define the abelian multiplicity function $m^{ab}_A$ of $A$ by

$$m^{ab}_A = m_{\text{Lie}(B),\mu}.$$ We define the dual abelian multiplicity function $\tilde{m}^{ab}_A$ of $A$ by

$$\tilde{m}^{ab}_A = m_{\mu,\text{Lie}(B')}.$$ Finally, we define the multiplicity function $m_A$ of $A$ by

$$m_A = m^t_A + m^{ab}_A = m_{\text{Lie}(A'),\mu}.$$ Using [2, IX.3.9], it is easily checked that these definitions only depend on $A$, and not on the choice of $K'$.
Proposition 3.4. For every tamely ramified abelian $K$-variety $A$, we have

$$\tilde{m}_{A}^{ab} = (m_{A}^{ab})^*.$$ 

Proof. We adopt the notations of Section 3.1. We set $(A^{\vee})' = A^{\vee} \times_{K} K'$ and we denote its Néron model by $(A^{\vee})'$ . The canonical divisorial correspondence on $A \times K A^{\vee}$ induces a divisorial correspondence on $B \times_{k} C$ that identifies $C$ with the dual abelian variety of $B$ [2, IX.5.4]. It suffices to show that the right $\mu$-action on $C$ is the inverse of the dual of the right $\mu$-action on $B$. To this end, we take a closer look at the construction of the divisorial correspondence on $B \times_{k} C$. Here we need the language of biextensions [2, VII and VIII]. We note that the following proof does not use the assumption that $A$ is tamely ramified and that $K'$ is a tame extension of $K$.

The canonical divisorial correspondence on $A \times K A^{\vee}$ can be interpreted as a Poincaré biextension $\mathcal{P}$ of $(A, A^{\vee})$ by $\mathbb{G}_{m,K}$ [2, VII.2.9.5], which is defined up to isomorphism. It induces a biextension $\mathcal{P}'$ of $(A', (A^{\vee})')$ by $\mathbb{G}_{m,K'}$ by base change. By [2, VIII.7.1], the biextension $\mathcal{P}'$ extends uniquely to a biextension of $(A'_{e}, ((A^{\vee})')_{e})$ by $\mathbb{G}_{m,K}$, which restricts to a biextension $\mathcal{P}'_{s}$ of $(A'_{s}, ((A^{\vee})')_{s})$ by $\mathbb{G}_{m,k}$. By [2, VIII.4.8], $\mathcal{P}'_{s}$ induces a biextension $\mathcal{Q}$ of $(B, C)$ by $\mathbb{G}_{m,k}$ that is characterized (up to isomorphism) by the fact that its pullback to $(A'_{e}, ((A^{\vee})')_{e})$ is isomorphic to $\mathcal{P}'_{s}$. The theorem in [2, IX.5.4] asserts that $\mathcal{Q}$ is a Poincaré biextension.

For every element $\zeta$ of $\mu$, we denote by $r_{\zeta}$ the right multiplication by $\zeta$ on $B$ and $C$. Since $\mathcal{P}'$ is obtained from the biextension $\mathcal{P}$ over $K$ by base change to $K'$, it follows easily from the construction that the pullback of the biextension $\mathcal{Q}$ through the $k$-morphisms

$$r_{\zeta} : B \rightarrow B$$

$$r_{\zeta} : C \rightarrow C$$

is isomorphic to $\mathcal{Q}$. Interpreting $\mathcal{Q}$ as an isomorphism

$$i : B \rightarrow C^{\vee}$$

in the way of [2, VIII.3.2.2], this means that the diagram

$$\begin{array}{ccc}
B & \xrightarrow{i} & C^{\vee} \\
\downarrow \quad r_{\zeta} & & \quad \uparrow (r_{\zeta})^{\vee} \\
B & \xrightarrow{i} & C^{\vee}
\end{array}$$

commutes, which is what we wanted to show. \qed

In the following proposition, we see how the multiplicity functions of a tamely ramified abelian $K$-variety $A$ are related to Edixhoven’s jumps and Chai’s elementary divisors of $A$. These jumps and elementary divisors are rational numbers in $[0, 1]$ that measure the behaviour of the Néron model of $A$ under
tame ramification of the base field $K$. For the definition of Edixhoven’s jumps, we refer to [12, 5.4.5]. The terminology we use is the one from [14, 4.12]; in particular, the multiplicity of a jump is defined there. For Chai’s elementary divisors, we refer to [5, 2.4]. By definition, the base change conductor $c(A)$ of $A$ is equal to the sum of the elementary divisors.

**Proposition 3.5.** Let $A$ be a tamely ramified abelian $K$-variety. The functions $m_A$, $m^\text{tor}_A$, $m^\text{ab}_A$ and $\tilde{m}^\text{ab}_A$ are supported on $((1/e)\mathbb{Z})/\mathbb{Z}$, with $e$ the degree of the minimal extension of $K$ where $A$ acquires semi-abelian reduction.

If we identify $[0,1] \cap \mathbb{Q}$ with $\mathbb{Q}/\mathbb{Z}$ via the bijection $[0,1] \cap \mathbb{Q} \to \mathbb{Q}/\mathbb{Z}: x \mapsto x \mod \mathbb{Z}$ then for every $x \in [0,1] \cap \mathbb{Q}$, the value $m_A(x)$ is equal to the multiplicity of $x$ as a jump in Edixhoven’s filtration for $A$. In particular, the support of $m_A$ is the set of jumps in Edixhoven’s filtration. The value $m_A(x)$ is also equal to the number of Chai’s elementary divisors of $A$ that are equal to $x$, and the base change conductor $c(A)$ of $A$ is given by

$$c(A) = \sum_{x \in [0,1] \cap \mathbb{Q}} (m_A(x) \cdot x).$$

**Proof.** See [12, 5.3 and 5.4.5] and [14, 4.8 and 4.13 and 4.18].

**Proposition 3.6.** We have the following equalities:

\[
\begin{align*}
\|m_A\| &= \dim(A), & \|m^\text{ab}_A\| &= \|\tilde{m}^\text{ab}_A\| &= a\text{pot}(A), \\
\|m^\text{tor}_A\| &= t\text{pot}(A), & m^\text{ab}_A(0) &= \tilde{m}^\text{ab}_A(0) &= a(A), \\
\end{align*}
\]

Moreover, we have

$$\sum_{x \in (\mathbb{Q}/\mathbb{Z}) \setminus \{0\}} m_A(x) = u(A).$$

**Proof.** We adopt the notations of Section 3.1. It follows immediately from the definitions that

\[
\begin{align*}
\|m_A\| &= \dim(\text{Lie}(A^1')), & \dim(\text{Lie}(A^1')) &= \dim(A), \\
\|m^\text{tor}_A\| &= \dim(\text{Lie}(T)), & \dim(\text{Lie}(T)) &= t\text{pot}(A), \\
\|\tilde{m}^\text{ab}_A\| &= \dim(\text{Lie}(B)), & \dim(\text{Lie}(B)) &= a\text{pot}(A).
\end{align*}
\]

By Lemma 3.2, the abelian, resp., reductive rank of $A^\circ$ is equal to the abelian, resp., reductive rank of the semi-abelian $k$-variety $((A^1')^\circ)^o$. In the notation of Lemma 3.2, the Chevalley decomposition of $((A^1')^\circ)^o$ is given by

$$0 \to (T^\nu)^o \to ((A^1')^\circ)^o \to B' \to 0$$
and there exists a separable isogeny \( f : B' \to (B')^\sigma \). By [12, 3.2], the natural morphisms

\[
\text{Lie}(T'^\sigma) \to \text{Lie}(T)^\sigma = \text{Lie}(T)[0] \\
\text{Lie}(B'^\sigma) \to \text{Lie}(B)^\sigma = \text{Lie}(B)[0]
\]

are isomorphisms. Since \( \text{Lie}(f) \) is also an isomorphism, we find

\[
m^\text{tor}_A(0) = t(A), \\
m^\text{ab}_A(0) = a(A).
\]

It follows that

\[
\sum_{x \in \mathbb{Q}/\mathbb{Z} \setminus \{0\}} m_A(x) = \|m_A\| - m^\text{tor}_A(0) - m^\text{ab}_A(0) \\
= \dim(A) - t(A) - a(A) \\
= u(A).
\]

By Proposition 3.4, we have

\[
\tilde{m}^\text{ab}_A(0) = m^\text{ab}_{A^\vee}(0)
\]

and we’ve just seen that this value is equal to the abelian rank \( a(A^\vee) \) of \( A^\vee \). But the abelian ranks of \( A \) and \( A^\vee \) coincide, by [2, 2.2.7], so that

\[
\tilde{m}^\text{ab}_A(0) = a(A).
\]

\[\square\]

**Lemma 3.7.** If \( A_1 \) and \( A_2 \) are tamely ramified abelian \( K \)-varieties, then

\[
m^\text{tor}_{A_1 \times_K A_2} = m^\text{tor}_{A_1} + m^\text{tor}_{A_2}, \\
m^\text{ab}_{A_1 \times_K A_2} = m^\text{ab}_{A_1} + m^\text{ab}_{A_2}, \\
\tilde{m}^\text{ab}_{A_1 \times_K A_2} = \tilde{m}^\text{ab}_{A_1} + \tilde{m}^\text{ab}_{A_2}.
\]

**Proof.** If we denote by \( A_1 \) and \( A_2 \) the Néron models of \( A_1 \), resp. \( A_2 \), then it follows immediately from the universal property of the Néron model that \( A_1 \times_K A_2 \) is a Néron model for \( A_1 \times_K A_2 \). Since the Chevalley decomposition of a connected smooth algebraic \( k \)-group commutes with finite fibered products over \( k \) and \( \text{Lie}(G_1 \times_K G_2) \) is canonically isomorphic to \( \text{Lie}(G_1) \oplus \text{Lie}(G_2) \) for any pair of algebraic \( k \)-groups \( G_1, G_2 \), the result follows. \[\square\]

**Proposition 3.8.** Let \( A \) be a tamely ramified abelian \( K \)-variety. Let \( L \) be a finite tame extension of \( K \) of degree \( e \), and put \( A_L = A \times_K L \). Then for each \( x \in \mathbb{Q}/\mathbb{Z} \), we have

\[
m^\text{tor}_{A_L}(x) = \sum_{y \in \mathbb{Q}/\mathbb{Z}, e \cdot y = x} m^\text{tor}_A(y) \\
m^\text{ab}_{A_L}(x) = \sum_{y \in \mathbb{Q}/\mathbb{Z}, e \cdot y = x} m^\text{ab}_A(y) \\
\tilde{m}^\text{ab}_{A_L}(x) = \sum_{y \in \mathbb{Q}/\mathbb{Z}, e \cdot y = x} \tilde{m}^\text{ab}_A(y)
\]
Proof. We adopt the notations of Section 3.1. Since the multiplicity functions do not depend on the choice of the field $K'$ where $A$ acquires semi-abelian reduction, we may assume that $L$ is contained in $K'$. If $\zeta$ is a generator of $\mu = G(K'/K)$, then the Galois group $G(K'/L)$ is generated by $\zeta^n$. Now the result easily follows from the definition of the multiplicity functions.

\begin{proposition}
If $f : A_1 \to A_2$ is an isogeny of tamely ramified abelian $K$-varieties and if the degree $\deg(f)$ of $f$ is prime to $p$, then
\[ m_{A_1}^{ab} = m_{A_2}^{ab} \quad \text{and} \quad m_{A_1}^{\text{tor}} = m_{A_2}^{\text{tor}}. \]
\end{proposition}

\begin{proof}
We put $n = \deg(f)$. Since $n$ is invertible in $K$, the morphism $f$ is separable, so that $\ker(f)$ is étale over $k$. Thus $\ker(f)$ is a finite étale $K$-group scheme of rank $n$. Every finite group scheme over a field is killed by its rank. (See [1, VII.18.5]; in our case, the result is elementary, because $\ker(f)$ is étale, so that we can reduce to the case of a constant group by base change to an algebraic closure of $K$. As an aside, we recall that Deligne has shown that every commutative finite group scheme over an arbitrary base scheme is killed by its rank [28, p.4].) It follows that $\ker(f)$ is contained in $n(A_1)$. Hence, there exists an isogeny $g : A_2 \to A_1$ such that $g \circ f = n_{A_1}$.

Let $K'$ be a tame finite extension of $K$ such that $A_1$ and $A_2$ acquire semi-abelian reduction over $K'$, and denote by $R'$ the integral closure of $R$ in $K'$. We denote the Néron model of $(A_i) \times_K K'$ by $A_i'$, for $i = 1, 2$. The morphisms $f : A_1 \to A_2$ and $g : A_2 \to A_1$ extend uniquely to morphisms of $R'$-group schemes
\[ f' : A_1' \to A_2' \]
\[ g' : A_2' \to A_1'. \]

For $i = 1, 2$, we denote by $B_i$ the abelian part of the semi-abelian $k$-variety $(A_i')^\text{tor}$. By functoriality of the Chevalley decomposition, $f_i'$ induces a morphism of $k$-group schemes $f_i' : B_1 \to B_2$. Likewise, $g_i'$ induces a morphism of $k$-group schemes $g_i' : B_2 \to B_1$. Since $g_i' \circ f_i'$ is multiplication by $n$, the same holds for $g_i' \circ f_i'$. In particular, $f_i'$ is an isogeny of degree prime to $p$. Thus $f_i'$ is a $\mu$-equivariant separable isogeny, so that $\Lie(f_i') : \Lie(B_1) \to \Lie(B_2)$ is a $\mu$-equivariant isomorphism, and $m_{A_1}^{ab} = m_{A_2}^{ab}$.

By [19, p.143], the dual morphism $(f_i')^\lor$ is again an isogeny, and its kernel is the Cartier dual of the kernel of $f_i'$. In particular, $f_i'$ and $(f_i')^\lor$ have the same degree, so that $(f_i')^\lor$ is separable. Since it is also equivariant for the left $\mu$-action on $B'$, we find that $\hat{m}_{A_1}^{ab} = \hat{m}_{A_2}^{ab}$.
\end{proof}

\begin{remark}
The same proof shows that $m^\text{tor}_A$ is invariant under isogenies of degree prime to $p$. We'll see in Corollary 4.5 that, more generally, the functions $m^\text{tor}_A$ and $m^{ab}_A + \hat{m}^{ab}_A$ are invariant under all isogenies.
\end{remark}

\begin{corollary}
Let $A$ be a tamely ramified abelian $K$-variety. If $k$ has characteristic zero, or $A$ is principally polarized, then
\[ m^{ab}_A = m^{ab}_{A'} \]
\end{corollary}
and
\[ \tilde{m}_{A}^{ab} = (m_{A}^{ab})^{*}. \]

**Proof.** The first equality follows from Proposition 3.9. Together with Proposition 3.4, it implies the second equality. \(\square\)

We will see in Theorem 6.3 that, when \(R\) is the ring of germs of holomorphic functions at the origin of \(\mathbb{C}\), the equality
\[ \tilde{m}_{A}^{ab} = (m_{A}^{ab})^{*} \]
expresses that the monodromy eigenvalues on the \((-1,0)\)-component of a certain limit mixed Hodge structure associated to \(A\) are the complex conjugates of the monodromy eigenvalues on the \((0,-1)\)-component. Corollary 3.11 generalizes this Hodge symmetry.

**Question 3.12.** Is it true that
\[ \tilde{m}_{A}^{ab} = (m_{A}^{ab})^{*} \]
for every tamely ramified abelian \(K\)-variety \(A\)?

### 4. Jumps and Monodromy

**Proposition 4.1.** Let \(B\) be an abelian \(k\)-variety, and let \(T\) be an algebraic \(k\)-torus. Fix an element \(e \in \mathbb{N}'\), and assume that \(\mu_{e}(k)\) acts on \(B\), resp. \(T\) from the right. We consider the dual left action of \(\mu_{e}(k)\) on \(B^{\vee}\). The functions
\[ m_{1} := m_{\text{Lie}(T),\mu_{e}(k)} \quad \text{and} \quad m_{2} := m_{\text{Lie}(B),\mu_{e}(k)} + m_{\mu_{e}(k),\text{Lie}(B^{\vee})} \]
are complete.

Moreover, for every prime \(\ell\) invertible in \(k\) and for each generator \(\zeta\) of \(\mu_{e}(k)\), the characteristic polynomial \(P_{1}(t)\) of \(\zeta\) on the \(\ell\)-adic Tate module
\[ \mathcal{G} T \cong \mathcal{H} T \otimes_{\mathbb{Z}_{\ell}} \mathbb{Q}_{\ell} \]
is equal to \(Q_{m_{1}}(t)\) (in the notation of Lemma 2.2). Likewise, the characteristic polynomial \(P_{2}(t)\) of \(\zeta\) on \(\mathcal{G} B\) is equal to \(Q_{m_{2}}(t)\).

**Proof.** We denote by
\[ \rho : \mathbb{Z}[t] \to k[t] \]
the unique ring morphism that maps \(t\) to \(t\). It is well-known that the characteristic polynomials \(P_{1}(t)\) and \(P_{2}(t)\) belong to \(\mathbb{Z}[t]\). For \(P_{1}(t)\), this follows from the canonical isomorphism
\[ \mathcal{G} T \cong \text{Hom}_{\mathbb{Z}}(X(T), \mathbb{Q}_{\ell}(1)), \]
where \(X(T)\) denotes the character module of \(T\). For \(P_{2}(t)\), it follows from [19, §19, Thm.4].

Since \(e\) is invertible in \(k\), \(P_{1}(t)\) and \(P_{2}(t)\) are products of \(k\)-tame cyclotomic polynomials. Thus, by Lemma 2.2 (and using the notation introduced there), we only have to show the following claims.
Claim 1: The image of $P_1(t)$ under $\rho$ equals $P_{m_1,\zeta}(t)$. Note that, by definition of the function $m_1$, the polynomial $P_{m_1,\zeta}(t)$ is the characteristic polynomial of the automorphism induced by $\zeta$ on $\text{Lie}(T)$. Thus Claim 1 is an immediate consequence of (4.1) and the canonical isomorphism

$$\text{Lie}(T) \cong \text{Hom}_Z(X(T), k).$$

Claim 2: The image of $P_2(t)$ under $\rho$ equals $P_{m_2,\zeta}(t)$. By definition of the function $m_2$, the polynomial $P_{m_2,\zeta}(t)$ is the product of the characteristic polynomials of the automorphism induced by $\zeta$ on $\text{Lie}(B)$ and the dual automorphism on $\text{Lie}(B^\vee)$. By [21, 5.1], the Hodge-de Rham spectral sequence of $B$ degenerates at $E_1$. This yields a natural short exact sequence

$$0 \to H^0(B, \Omega^1_B) \to H^1_{\text{dR}}(B) \to H^1(B, \mathcal{O}_B) \to 0$$

where $H^1_{\text{dR}}(B)$ is the degree one de Rham cohomology of $B$. We have natural isomorphisms

$$H^0(B, \Omega^1_B) \cong \text{Lie}(B)^\vee,$$

$$H^1(B, \mathcal{O}_B) \cong \text{Lie}(B^\vee).$$

(The first isomorphism follows from [4, 4.2.2]; the second isomorphism from [19, §13, Cor.3]). Thus it suffices to show that the image of $P_2(t)$ under $\rho$ is equal to the characteristic polynomial of $\zeta$ on $H^1_{\text{dR}}(B)$. As explained in the proof of [14, 5.12], this can be deduced from the fact that étale cohomology is a Weil cohomology, as well as de Rham cohomology (if $k$ has characteristic zero) and crystalline cohomology (if $k$ has characteristic $p > 0$), so that the characteristic polynomials of $\zeta$ on the respective cohomology spaces must coincide. □

For every $n \in \mathbb{Z}_{>0}$ and every $a \in \mathbb{C}$, we denote by $\text{Diag}_n(a)$ the rank $n$ diagonal matrix with diagonal $(a, \ldots, a)$, and by $\text{Jord}_n(a)$ the rank $n$ Jordan matrix with diagonal $(a, \ldots, a)$ and subdiagonal $(1, \ldots, 1)$. For any two complex square matrices $M$ and $N$, of rank $m$, resp. $n$, we denote by $M \oplus N$ the rank $m + n$ square matrix

$$M \oplus N = \begin{pmatrix} M & 0 \\ 0 & N \end{pmatrix}.$$ 

For every integer $q > 0$, we put

$$\oplus^q M = \underbrace{M \oplus \cdots \oplus M}_{q \text{ times}}.$$

Definition 4.2. For $i = 1, 2$, let

$$m_i : \mathbb{Q}/\mathbb{Z} \to \mathbb{N}$$

be a function with finite support. The Jordan matrix $\text{Jord}(m_1, m_2)$ associated to the couple $(m_1, m_2)$ is the complex square matrix of rank $\|m_1\| + 2 \cdot \|m_2\|$. 
given by
\[ \text{Jord}(m_1, m_2) = \bigoplus_{x \in \text{Supp}(m_1)} \left( \text{Diag}_{m_1(x)}(\exp(2\pi ix)) \right) \]
\[ \bigoplus_{y \in \text{Supp}(m_2)} \left( \bigoplus_{m^2(y)} \text{Jord}_2(\exp(2\pi iy)) \right) \]
where we ordered the set \( \mathbb{Q}/\mathbb{Z} \) using the bijection \( \mathbb{Q} \cap [0, 1[ \to \mathbb{Q}/\mathbb{Z} \) and the usual ordering on \( [0, 1[ \).

**Lemma 4.3.** Let \( V \) be a finite dimensional vector space over an algebraically closed field \( F \) of characteristic zero, and let \( M \) be an endomorphism of \( V \). Let \( d > 0 \) be an integer such that \( M^d \) is unipotent. Set
\[ W = \{ v \in V | M^d(v) = v \} \]
and assume that \( M^d \) acts trivially on \( V/W \) and that there exists an \( M \)-equivariant isomorphism between \( (V/W)^\vee \) and an \( M \)-stable subspace \( W' \) of \( W \).

Then the endomorphism \( M \) of \( V \) has the following Jordan form: for every eigenvalue \( \alpha \) of \( M \) on \( W' \) (counted with multiplicities), there is a Jordan block of size two with eigenvalue \( \alpha \), and for every eigenvalue \( \beta \) of \( M \) on \( W/W' \) (counted with multiplicities), there is a Jordan block of size one with eigenvalue \( \beta \).

**Proof.** Since \( M^d \) acts trivially on \( W \) and \( V/W \), we know that \((M^d - \text{Id})^2 = 0\) on \( V \), so that the minimal polynomial of \( M \) divides \((t^d - 1)^2\) and the Jordan blocks of \( M \) have size at most two. The subspace \( W \) of \( V \) is generated by all the eigenvectors of \( M \). Thus the dimension of \( V/W \) is equal to the number of Jordan blocks of \( M \) of size two, and the eigenvalues of these Jordan blocks are precisely the eigenvalues of \( M \) on \( V/W \), or, equivalently, \( W' \). It follows that the eigenvalues of \( M \) on \( V \) corresponding to a Jordan block of size one are the eigenvalues of \( M \) on \( W/W' \). \( \square \)

**Theorem 4.4.** We fix an embedding \( \mathbb{Q}_\ell \hookrightarrow \mathbb{C} \). If \( A \) is a tamely ramified abelian \( K \)-variety, then the monodromy action of \( \sigma \) on \( H^1(A \times_K K^t, \mathbb{Q}_\ell) \) has Jordan form
\[ \text{Jord}(m_{ab}^{\text{tor}} A + \tilde{m}_{ab}^{\text{tor}} A, m_{\text{tor}} A) \]
Moreover, the functions \( m_{\text{tor}}^{\text{tor}} A \) and \( m_{ab}^{\text{tor}} A + \tilde{m}_{ab}^{\text{tor}} A \) are complete.

**Proof.** We adopt the notations of Section 3.1. We denote by \( \mathcal{T}_A \) the \( \ell \)-adic Tate module of \( A \). We put \( I = G(K^*/K) \) and \( I' = G(K^*/K') \). Recall that there exists a canonical \( I \)-equivariant isomorphism
\[ H^1(A \times_K K^t, \mathbb{Q}_\ell) \cong \text{Hom}_{\mathcal{M}_2}(\mathcal{T}_A, \mathbb{Q}_\ell) \]
(see [18, 15.1]). Since \( A \) is tamely ramified, the wild inertia subgroup \( P \subset I \) acts trivially on \( H^1(A \times_K K^t, \mathbb{Q}_\ell) \) and \( \mathcal{T}_A \), so that the \( I \)-action on these modules factors through an action of \( I/P = G(K^t/K) \).
Since $P$ is a $p$-group and $p$ is prime to $\ell$, there exists for every $K$-variety $X$ and every integer $i \geq 0$ a canonical $G(K'/K)$-equivariant isomorphism

$$H^i(X \times_K K', \mathbb{Q}_\ell) \cong H^i(X \times_K K^s, \mathbb{Q}_\ell)^P$$

(see [2, I.2.7.1]). In our case, this yields a canonical $G(K'/K)$-equivariant isomorphism

$$H^1(A \times_K K^s, \mathbb{Q}_\ell) = H^1(A \times_K K^s, \mathbb{Q}_\ell)^P \cong H^1(A \times_K K^s, \mathbb{Q}_\ell).$$

By (4.2) and (4.3), it suffices to show that the action of $\sigma$ on

$$\mathcal{V}_A = \mathcal{G}_A \otimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell$$

has Jordan form

$$\text{Jord}(m_{A}^{ab} + \tilde{m}_{A}^{ab}, m_{A}^{\text{tor}})$$

and that $m_{A}^{ab} + \tilde{m}_{A}^{ab}$ and $m_{A}^{\text{tor}}$ are complete. Consider the filtration

$$\mathcal{X}(\mathcal{G}_A)^{\text{et}} \subset (\mathcal{G}_A)^{\text{et}} \subset \mathcal{G}_A$$

from [2, IX.4.1.1], with $(\mathcal{G}_A)^{\text{et}}$ the essentially fixed part of the Tate module $\mathcal{G}_A$, and $(\mathcal{G}_A)^{\text{et}}$ the essentially toric part. By definition,

$$(\mathcal{G}_A)^{\text{et}} = (\mathcal{G}_A)^{\ell T}$$

and $(\mathcal{G}_A)^{\text{et}}$ is stable under the action of $I$ on $\mathcal{G}_A$. We denote by

$$\mathcal{X}(\mathcal{G}_A)^{\text{et}} \subset (\mathcal{G}_A)^{\text{et}} = (\mathcal{G}_A)^{\ell T} \subset \mathcal{G}_A$$

the filtration obtained from (4.4) by tensoring with $\mathbb{Q}_\ell$. By [2, IX.4.1.2] there exists an $I$-equivariant isomorphism

$$\mathcal{X}(\mathcal{G}_A)/(\mathcal{G}_A)^{\text{et}} \cong ((\mathcal{G}_A)^{\text{et}})^{\ell T}.$$

In particular, $I'$ acts trivially on $\mathcal{X}(\mathcal{G}_A)/(\mathcal{G}_A)^{\text{et}}$. It follows that the $I'$-action on $\mathcal{X}(\mathcal{G}_A)$ is unipotent of level $\leq 2$ (this means that for every element $\gamma$ of $I'$, the endomorphism $(\gamma - 1)^2$ of $\mathcal{X}(\mathcal{G}_A)$ is zero), and that the $I$-action on $(\mathcal{G}_A)^{\text{et}}$ and $\mathcal{X}(\mathcal{G}_A)/(\mathcal{G}_A)^{\text{et}}$ factors through an action of $I/\ell I' \simeq \mu_d(k)$, where $d = [K'/K]$. We denote by $\bigtriangledown$ the image of $\sigma$ under the projection $G(K'/K) \to \mu_d$. The element $\sigma^d$ is a topological generator of $I'/\ell I'$. Combining (4.5) and (4.6) and applying Lemma 4.3 to the $\sigma$-action on $\mathcal{X}(\mathcal{G}_A)$, we see that it suffices to prove the following claims:

1. the $\bigtriangledown$-action on $(\mathcal{G}_A)^{\text{et}}$ has Jordan form $\text{Jord}(m_{A}^{\text{tor}}, 0)$, and $m_{A}^{\text{tor}}$ is complete,
2. the $\bigtriangledown$-action on $(\mathcal{G}_A)^{\text{et}}/(\mathcal{G}_A)^{\text{et}}$ has Jordan form $\text{Jord}(m_{A}^{ab} + \tilde{m}_{A}^{ab}, 0)$, and $m_{A}^{ab} + \tilde{m}_{A}^{ab}$ is complete.

Since $\bigtriangledown$ is the identity, the Jordan forms of the $\bigtriangledown$-actions in (1) and (2) are diagonal matrices. By [2, IX.4.2.7 and IX.4.2.9] there exist $\mu$-equivariant isomorphisms

$$(\mathcal{G}_A)^{\text{et}} \cong \mathcal{G}_T$$

and

$$(\mathcal{G}_A)^{\text{et}}/(\mathcal{G}_A)^{\text{et}} \cong \mathcal{G}_B.$$
so that claims (1) and (2) follow from Proposition 4.1. □

Corollary 4.5. The functions $m_A^{ab} + \tilde{m}_A^{ab}$ and $m_A^{tor}$ are invariant under isogeny. In particular, $m_A^{tor} = m_A^{AB}$, and

$$m_A^{ab} + \tilde{m}_A^{ab} = m_A^{ab} + \tilde{m}_A^{ab}.$$ 

The multiplicity function $m_A$, the jumps of $A$ (with their multiplicities), the elementary divisors of $A$ and the base change conductor $c(A)$ are invariant under isogenies of degree prime to $p$.

Proof. This follows from Propositions 3.5 and 3.9, and Theorem 4.4. □

Remark 4.6. Beware that the base change conductor, and thus the function $m_A^{ab}$, of a tamely ramified abelian $K$-variety $A$ can change under isogenies of degree $p$, if $p > 0$; see [5, 6.10.2] for an example.

Corollary 4.7. Using the notations of Proposition 3.5, the base change conductor of a tamely ramified abelian $K$-variety $A$ is given by

$$c(A) = \frac{1}{2}(t_{pot}(A) - t(A)) + \sum_{x \in [0,1] \cap \mathbb{Q}} m_A^{ab}(x)x.$$ 

In particular, if $A$ has potential purely multiplicative reduction (which means that $a_{pot}(A) = 0$), then

$$c(A) = \frac{u(A)}{2}.$$ 

Proof. By Proposition 3.5, we know that

$$c(A) = \sum_{x \in [0,1] \cap \mathbb{Q}} m_A^{tor}(x)x + \sum_{x \in [0,1] \cap \mathbb{Q}} m_A^{ab}(x)x.$$ 

Since $m_A^{tor}$ is complete, we have that

$$\sum_{x \in [0,1] \cap \mathbb{Q}} m_A^{tor}(x)x = \frac{1}{2} \left( \sum_{x \in [0,1] \cap \mathbb{Q}} m_A^{tor}(x)x + \sum_{x \in [0,1] \cap \mathbb{Q}} m_A^{tor}(x)(1-x) \right)$$

$$= \frac{1}{2} \left( \sum_{x \in [0,1] \cap \mathbb{Q}} m_A^{tor}(x) \right)$$

$$= \frac{1}{2}(\|m_A^{tor}\| - m_A^{tor}(0))$$

$$= \frac{1}{2}(t_{pot}(A) - t(A))$$

where the last equality follows from Proposition 3.6. If $a_{pot}(A) = 0$, then it follows from Proposition 3.6 that $m_A^{ab} = 0$ and $a(A) = 0$, so that

$$c(A) = \frac{1}{2}(t_{pot}(A) - t(A)) = \frac{1}{2} \left( \text{dim}(A) - t(A) \right) = \frac{u(A)}{2}.$$ 

Documenta Mathematica 16 (2011) 937–968
Remark 4.8. If $A$ has potential purely multiplicative reduction, then Corollary 4.7 can be viewed as a special case of Chai’s result that for every abelian $K$-variety $A$ (not necessarily tamely ramified) with potential purely multiplicative reduction, the base change conductor $c(A)$ equals one fourth of the Artin conductor of $\mathcal{V}_t A$ [5, 5.2]. If $A$ is tamely ramified, then this Artin conductor is simply the dimension of $\mathcal{V}_t A/((\mathcal{V}_t A)^s)^I$, where $I = G(K^*/K)$ and $(\mathcal{V}_t A)^s$ is the semi-simplification of the $I$-adic $I$-representation $\mathcal{V}_t A$. This value is precisely the number of eigenvalues of $\sigma$ (counted with multiplicities) that are different from one. Combining Proposition 3.6 with Theorem 4.4, we find that the Artin conductor of $\mathcal{V}_t A$ equals

$$2 \dim(A) - m_A^{ab}(0) - \tilde{m}_A^{ab}(0) - 2m_A^{tor}(0) = 2u(A).$$

Corollary 4.9. Let $A$ be a tamely ramified abelian $K$-variety, and let $e$ be the degree of the minimal extension of $K$ where $A$ acquires semi-abelian reduction. Fix a primitive $e$-th root of unity $\xi$ in an algebraic closure $\mathbb{Q}^b$ of $\mathbb{Q}$. The characteristic polynomial

$$P_\sigma(t) = \det(t \cdot \text{Id} - \sigma | H^1(A \times_K K^t, \mathbb{Q}_\ell))$$

of $\sigma$ on $H^1(A \times_K K^t, \mathbb{Q}_\ell)$ is given by

$$P_\sigma(t) = \prod_{i \in ((1/e)\mathbb{Z})/\mathbb{Z}} (t - \xi^{e^{-i}}m_A^{ab}(i) + \tilde{m}_A^{ab}(i) + 2m_A^{tor}(i) \in \mathbb{Z}[t].$$

Proof. This is an immediate consequence of Theorem 4.4. □

Corollary 4.10. Let $A$ be a tamely ramified abelian $K$-variety. Assume either that $k$ has characteristic zero, or that $A$ is principally polarized. Then $m_A^{ab}$ and $\tilde{m}_A^{ab}$ are semi-complete, and the monodromy action of $\sigma$ on $H^1(A \times_K K^t, \mathbb{Q}_\ell)$ has Jordan form

$$\text{Jord}(m_A^{ab} + (m_A^{ab})^*, m_A^{tor}).$$

In the notation of Corollary 4.9, we have

$$P_\sigma(t) = \prod_{i \in ((1/e)\mathbb{Z})/\mathbb{Z}} (t - \xi^{e^{-i}}m_A^{ab}(i) + m_A^{ab}(-i) + 2m_A^{tor}(i) \in \mathbb{Z}[t].$$

Proof. Semi-completeness of $m_A^{ab}$ and $\tilde{m}_A^{ab}$ follows from Corollary 3.11 and Theorem 4.4. The remainder of the statement is a combination of Corollaries 3.11 and 4.9. □

5. Potential toric rank and Jordan blocks

5.1. The weight filtration associated to a nilpotent operator.
Throughout this section, we fix a field $F$ of characteristic zero and a finite dimensional vector space $V$ over $F$. For every endomorphism $M$ on $V$, we consider its Jordan-Chevalley decomposition

$$M = M_s + M_n.$$
with $M_n$ the semi-simple part of $M$ and $N_n$ its nilpotent part. We recall the following well-known property.

**Proposition 5.1.** Let $N$ be a nilpotent endomorphism of $V$. Let $w$ be an integer. There exists a unique finite ascending filtration $W = (W_i, i \in \mathbb{Z})$ on $V$ such that

1. $NW_i \subset W_{i-2}$ for all $i$ in $\mathbb{Z}$,
2. the morphism of vector spaces $\text{Gr}^W_{w+\alpha} V \to \text{Gr}^W_{w-\alpha} V$

induced by $N^\alpha$ is an isomorphism for every integer $\alpha > 0$.

The filtration $W$ is called the weight filtration centered at $w$ associated to the nilpotent operator $N$.

**Proof.** See, for instance, [10, 1.6.1].

It is clear from the definition that the weight filtration centered at another integer $w'$ is the shifted filtration $W' = W_{w'-w}$. We define the amplitude of the filtration $W$ in Proposition 5.1 as the smallest integer $n \geq 0$ such that $W_n = V$. This value does not depend on the choice of the central weight $w$. The amplitude is related to sizes of Jordan blocks in the following way.

**Proposition 5.2.** Let $M$ be an endomorphism of $V$. We denote by $a$ the amplitude of the weight filtration $W$ associated to $M_n$ (centered at any weight $w \in \mathbb{Z}$). Then $a + 1$ is the maximum of the ranks of the Jordan blocks of $M$.

**Proof.** We may assume that $w = 0$ and that $M = M_n$. Then the proposition is an immediate consequence of the explicit description of the weight filtration in [10, 1.6.7].

**Proposition 5.3.** Let $N$ be a nilpotent endomorphism of $V$, and denote by $W$ the associated weight filtration centered at $w \in \mathbb{Z}$.

1. If $N'$ is another nilpotent operator on $V$ such that $(N - N')W_i \subset W_{i-3}$ for all $i \in \mathbb{Z}$, then the weight filtration associated to $N'$ centered at $w$ coincides with $W$.
2. The weight filtration $W$ does not change if we multiply $N$ with an automorphism $S$ of $V$ that commutes with $N$.

**Proof.** (1) This follows immediately from the definition of the weight filtration in Proposition 5.1.

(2) Clearly, the filtration $(S(W_i), i \in \mathbb{Z})$ on $V$ also satisfies properties (1) and (2) in Proposition 5.1, so that $S(W_i) = W_i$ for all $i$ in $\mathbb{Z}$ by uniqueness of the weight filtration. This implies at once that $W$ coincides with the weight filtration associated to $NS$ centered at $w$. 

---

*Documenta Mathematica 16 (2011) 937–968*
Definition 5.4. Let $W = (W_i, i \in \mathbb{Z})$ be an ascending filtration on $V$. The dual filtration $W^\vee$ on $V^\vee$ is the ascending filtration defined by

$$(W^\vee)_i = (W_{-i-1})^\perp$$

for all $i$ in $\mathbb{Z}$.

For every integer $j \geq 0$, the degree $j$ exterior power of $W$ is the ascending filtration $\wedge^j W$ on $\wedge^j V$ given by

$$(\wedge^j W)_i = \sum_{i_1 + \cdots + i_j = i} W_{i_1} \wedge \cdots \wedge W_{i_j}.$$}

If $V'$ is another finite dimensional vector space over $F$, endowed with an ascending filtration $W'$, then the tensor product of $W$ and $W'$ is the ascending filtration $W \otimes W'$ on $V \otimes V'$ given by

$$(W \otimes W')_i = \sum_{i_1 + i_2 = i} W_{i_1} \otimes W'_{i_2}.$$}

Proof. Point (1) and (2) are proven in [10, 1.6.9], using the theorem of Jacobson-Morosov. Point (3) can be proven in exactly the same way, since the morphism of linear groups

$$h : GL(V) \to GL(\wedge^j V) : M \mapsto \wedge^j M$$

induces a morphism of Lie algebras

$$\text{Lie}(h) : \text{End}(V) \to \text{End}(\wedge^j V)$$

that sends $N$ to $N^{(\wedge^j)}$.\hfill \Box

Corollary 5.6. Let $M$ be an automorphism of $V$, and consider an integer $w$ and an integer $j > 0$. We denote by $W$ the weight filtration centered at $w$ associated to the nilpotent operator $M_n$ on $V$.\hfill 
Let $V'$ be another finite dimensional vector space, endowed with an automorphism $M'$. If $W'$ is the weight filtration associated to $M'_n$ centered at $w' \in \mathbb{Z}$, then $W \otimes W'$ is the weight filtration centered at $w + w'$ associated to the nilpotent operator $(M \otimes M')_n$ on $V \otimes V'$.

(2) The exterior power filtration $\wedge^j W$ is the weight filtration centered at $w \cdot j$ associated to the nilpotent operator $(\wedge^j M)_n$ on $\wedge^j V$.

Proof. (1) By Proposition 5.3(2), we may assume that $M$ and $M'$ are unipotent, since multiplying these operators with $M^{-1}$ and $(M')^{-1}$, respectively, has no influence on the weight filtrations that we want to compare. Then

\[
(M \otimes M')_n - \text{Id} \otimes M'_n - M_n \otimes \text{Id} = M_n \otimes M'_n.
\]

It follows that

\[
((M \otimes M')_n - \text{Id} \otimes M'_n - M_n \otimes \text{Id})(W \otimes W')_i \subset (W \otimes W')_{i-4}
\]

for all $i \in \mathbb{Z}$. The result now follows from Propositions 5.3(1) and 5.5(2).

(2) The proof is completely similar to the proof of (1): one reduces to the case where $M$ is unipotent, and one shows by direct computation that

\[
(\wedge^j M)_n - (M_n)^{(\wedge^j)}
\]

shifts weights by at least $-4$. \[\square\]

The following lemma and proposition will allow us to compute the maximal size of certain Jordan blocks of monodromy on the cohomology of a tamely ramified abelian $K$-variety (Theorem 5.9).

**Lemma 5.7.** Let $F$ be an algebraically closed field of characteristic zero, and let $V \neq \{0\}$ be a finite dimensional vector space over $F$. Let $M$ be an automorphism of $V$, with Jordan form

\[
\text{Jord}_m(\xi)
\]

where $m \in \mathbb{Z}_{>0}$ and $\xi \in F^\times$. Then for every integer $j$ in $[1, m]$ and every integer $w$, the weight filtration centered at $w$ associated to the nilpotent operator $(\wedge^j M)_n$ on $\wedge^j V$ has amplitude $m(m - j)$.

**Proof.** We may assume that $w = 0$. Denote by $W$ the weight filtration associated to $M_n$ centered at 0. By Corollary 5.6, the weight filtration associated to $(\wedge^j M)_n$ centered at 0 coincides with the exterior power filtration $\wedge^j W$.

By the explicit description of the weight filtration in [10, 1.6.7], the dimension of $\text{Gr}^W_\alpha V$ is one if $\alpha$ is an integer in $[1 - m, m - 1]$ such that $\alpha - m$ is odd, and zero in all other cases. This easily implies that $\wedge^j W$ has amplitude

\[
(m - 1) + (m - 3) + \ldots + (m - 2j + 1) = m(m - j).
\]

\[\square\]
Proposition 5.8. Let $F$ be an algebraically closed field of characteristic zero, and let $V \neq \{0\}$ be a finite dimensional vector space over $F$. Let $M$ be an automorphism of $V$, with Jordan form

$$
\text{Jord}_{m_1}(\xi_1) \oplus \cdots \oplus \text{Jord}_{m_q}(\xi_q)
$$

where $q \in \mathbb{Z}_{>0}$, $m \in (\mathbb{Z}_{>0})^q$ and $\xi \in F^\times$ for $i = 1, \ldots, q$.

We fix an integer $j > 0$. For every element $\zeta$ of $F$, we denote by $\max_\zeta$ the maximum of the ranks of the Jordan blocks of $\wedge^j M$ on $\wedge^j V$ with eigenvalue $\zeta$.

If we denote by $S$ the set of tuples $s \in \mathbb{N}^q$ such that $\|s\| = j$ and $s_i \leq m_i$ for each $i \in \{1, \ldots, q\}$, then

$$
\max_\zeta = \max \{1 + \sum_{i=1}^{q} s_i (m_i - s_i) \mid s \in S, \prod_{i=1}^{q} (\xi_i)^{s_i} = \zeta \}
$$

for every $\zeta \in F$, with the convention that $\max_\emptyset = 0$.

Proof. We can write

$$
V = V_1 \oplus \cdots \oplus V_q
$$

such that $M(V_i) \subset V_i$ for each $i$ and such that the restriction $M_i$ of $M$ to $V_i$ has Jordan form $\text{Jord}_{m_i}(\xi_i)$. If we put

$$
V_s = (\wedge^{s_1} V_1) \otimes \cdots \otimes (\wedge^{s_q} V_q)
$$

for each $s \in S$, then we have a canonical isomorphism

$$
\wedge^j V \cong \bigoplus_{s \in S} V_s
$$

such that every summand $V_s$ is stable under $\wedge^j M$ and the restriction of $\wedge^j M$ to $V_s$ equals

$$
(\wedge^{s_1} M_1) \otimes \cdots \otimes (\wedge^{s_q} M_q).
$$

The automorphism $\wedge^j M$ has a unique eigenvalue on $V_s$, which is equal to

$$
\prod_{i=1}^{q} (\xi_i)^{s_i}.
$$

By Proposition 5.2, Corollary 5.6(1) and Lemma 5.7, the maximal rank of a Jordan block of $\wedge^j M$ on $V_s$ equals

$$
1 + \sum_{i=1}^{q} s_i (m_i - s_i).
$$

This yields the desired formula for $\max_\zeta$. \qed
5.2. The strong form of the monodromy conjecture.

Theorem 5.9. Let $A$ be a tamely ramified abelian $K$-variety of dimension $g$. For every embedding of $\mathbb{Q}_\ell$ in $\mathbb{C}$, the value $\alpha = \exp(2\pi i c(A))$ is an eigenvalue of $\sigma$ on $H^g(A \times_K K^t, \mathbb{Q}_\ell)$. Each Jordan block of $\sigma$ on $H^g(A \times_K K^t, \mathbb{Q}_\ell)$ has rank at most $t_{\text{pot}}(A) + 1$, and $\sigma$ has a Jordan block with eigenvalue $\alpha$ on $H^g(A \times_K K^t, \mathbb{Q}_\ell)$ with rank $t_{\text{pot}}(A) + 1$.

Proof. Since $A$ is tamely ramified, we have a canonical $G(K^t/K)$-equivariant isomorphism of $\mathbb{Q}_\ell$-vector spaces $H^g(A \times_K K^t, \mathbb{Q}_\ell) \cong \bigwedge^g H^1(A \times_K K^t, \mathbb{Q}_\ell)$. By Theorem 4.4, the monodromy operator $\sigma$ has precisely $\|m_{\text{tor}}A\|$ Jordan blocks of size 2 on $H^1(A \times_K K^t, \mathbb{Q}_\ell)$, and no larger Jordan blocks. It follows from Proposition 5.8 that the size of the Jordan blocks of $\sigma$ on $H^g(A \times_K K^t, \mathbb{Q}_\ell)$ is bounded by $1 + \|m_{\text{tor}}A\|$. By Proposition 3.6, we know that $\|m_{\text{tor}}A\| = t_{\text{pot}}(A)$. By Proposition 3.5, the image in $\mathbb{Q}/\mathbb{Z}$ of the base change conductor $c(A)$ equals

$$\sum_{x \in \mathbb{Q}/\mathbb{Z}} ((m_{\text{tor}}A(x) + m_{\text{ab}}A(x)) \cdot x)$$

and by Proposition 3.6, we have

$$\sum_{x \in \mathbb{Q}/\mathbb{Z}} (m_{\text{tor}}A(x) + m_{\text{ab}}A(x)) = g.$$ 

Hence, by Theorem 4.4 and Proposition 5.8, $\sigma$ has a Jordan block of size $1 + t_{\text{pot}}(A)$ with eigenvalue $\alpha$ on $H^g(A \times_K K^t, \mathbb{Q}_\ell)$. □

6. Limit Mixed Hodge structure

Let $A$ be a tamely ramified abelian $K$-variety of dimension $g$. Theorem 4.4 shows that the couple of functions $(m_{\text{tor}}^A, m_{\text{ab}}^A + \hat{m}_{\text{ab}}^A)$ and the Jordan form of $\sigma$ on the tame degree one cohomology of $A$ determine each other. It does not tell us how to recover $m_{\text{ab}}^A$ and $\hat{m}_{\text{ab}}^A$ individually from the cohomology of $A$.

In this section, we assume that $A$ is obtained by base change from a family of abelian varieties over a smooth complex curve. We will give an interpretation of the multiplicity functions $m_{\text{ab}}^A$, $\hat{m}_{\text{ab}}^A$ and $m_{\text{tor}}^A$ in terms of the limit mixed Hodge structure of the family.

6.1. Limit mixed Hodge structure of a family of abelian varieties.

Let $S$ be a connected smooth complex algebraic curve, let $s$ be a closed point on $S$, and choose a local parameter $t$ on $S$ at $s$. We put $K = \mathbb{C}((t))$, $R = \mathbb{C}[[t]]$ and $S = S \setminus \{s\}$. Let

$$f : X \to S$$

be a smooth projective family of abelian varieties over $S$, of relative dimension $g$, and put

$$A = X \times_S \text{Spec } K.$$
We choose an extension of $f$ to a flat projective morphism

$$\overline{f}: \overline{X} \to \overline{S},$$

and we denote by $\overline{X}_s$ the fiber of $\overline{f}$ over the point $s$.

We denote by $(\cdot)_{\text{an}}$ the complex analytic GAGA functor, but we will usually omit it from the notation if no confusion can occur. For instance, when we speak of the sheaf $R^i f_{\ast}(\mathbb{Z})$, it should be clear that we mean $R^i f^\text{an}_{\ast}(\mathbb{Z})$.

For every $i \in \mathbb{N}$, we consider the degree $i$ limit cohomology, resp. homology,

$$H^i(X_\infty, \mathbb{Z}) := H^i(\overline{X}_s(\mathbb{C}), R\overline{f}(\mathbb{Z}))$$

$$H_i(X_\infty, \mathbb{Z}) := H^{2g-i}(\overline{X}_s(\mathbb{C}), R\overline{f}(\mathbb{Z}))(g) = H^{2g-i}(X_\infty, \mathbb{Z})(g)$$

of $\overline{f}$ at $s$. Here

$$R\overline{f}(\mathbb{Z}) \in D^b_c(\overline{X}_s(\mathbb{C}), \mathbb{Z})$$

denotes the complex of nearby cycles associated to $\overline{f}_{\text{an}}$. For every $i \in \mathbb{N}$, the $\mathbb{Z}$-module $H^i(X_\infty, \mathbb{Z})$ is non-canonically isomorphic to $H^i(X_s(\mathbb{C}), \mathbb{Z})$, where $z$ is any point of $S(\mathbb{C})$ and $X_z$ is the fiber of $f$ over $z$ [3, XIV.1.3.2].

Likewise, by Poincaré duality, $H_i(X_{\infty, \mathbb{Z}})$ is non-canonically isomorphic to $H_i(X_{\infty, \mathbb{C}})$.

The limit cohomology and homology are independent of the chosen compactification $\overline{f}$, as can be deduced from [11, 4.2.11] by dominating two compactifications by a third one.

We set

$$H^i(X_\infty, \mathbb{Q}) := H^i(X_\infty, \mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{Q}$$

$$H^i(X_\infty, \mathbb{C}) := H^i(X_\infty, \mathbb{Z}) \otimes_{\mathbb{Z}} \mathbb{C}$$

and we use similar notations for the limit homology $H_i(X_\infty, \ast)$. For all $i \in \mathbb{N}$, we denote by $M$ the monodromy operators on $H_i(X_\infty, \mathbb{Z})$ and $H^i(X_\infty, \mathbb{Z})$, and by $N$ the logarithm of the nilpotent part $M_n$ of $M$.

The $\mathbb{Z}$-modules $H^i(X_{\infty, \mathbb{Z}})$ and $H_i(X_{\infty, \mathbb{Z}})$ carry natural mixed Hodge structures, which are the limits at $s$ of the variations of Hodge structures

$$R^i f_{\ast}(\mathbb{Z}), \text{ resp. } R^{2g-i} f_{\ast}(\mathbb{Z})(g),$$

on $S$. The existence of these limit mixed Hodge structures was conjectured by Deligne, and they were constructed by Schmid [25] and Steenbrink [26].

The weight filtrations on $H^i(X_{\infty, \mathbb{Q}})$ and $H_i(X_{\infty, \mathbb{Q}})$ are the weight filtrations centered at $i$, resp. $-i$, associated to the nilpotent operator $N$. We will briefly recall Schmid’s construction of the limit Hodge filtration below. The action of the semi-simple part $M_s$ of $M$ on $H_i(X_{\infty, \mathbb{Q}})$ and $H^i(X_{\infty, \mathbb{Q}})$ is a morphism of rational mixed Hodge structures, by [27, 2.13].

For every $i \in \mathbb{N}$, there exists an isomorphism of $\mathbb{Q}_\ell$-vector spaces

$$(6.1) \quad H^i(X_{\infty, \mathbb{Z}}) \otimes_{\mathbb{Z}} \mathbb{Q}_\ell \cong H^i(A \times_K K^a, \mathbb{Q}_\ell)$$

such that the monodromy action on the left hand side corresponds to the action of the canonical topological generator of $G(K^a/K) \cong \hat{\mathbb{Z}}(1)(\mathbb{C})$ on the right hand side. This follows from Deligne’s comparison theorem for $\ell$-adic versus complex analytic nearby cycles [3, XIV.2.8]. Thus, if $K^a$ is a finite extension of $K$ such
that $A \times_K K'$ has semi-abelian reduction and if we set $d = [K' : K]$, then $(M_s)^d$ is the identity on $H^i(X_\infty, \mathbb{Q})$ and $H_t(X_\infty, \mathbb{Q})$ for all $i \geq 0$. Identifying $M_s$ with the canonical generator $\exp(2\pi i/d)$ of $\mu_d(\mathbb{C})$, we obtain an action of $\mu_d(\mathbb{C})$ on the rational mixed Hodge structures $H^i(X_\infty, \mathbb{Q})$ and $H_t(X_\infty, \mathbb{Q})$, for all $i \geq 0$.

For the definition of the dual and the exterior powers of a mixed Hodge structure, we refer to [22, 3.2].

**Proposition 6.1.**

(1) For every $i \in \mathbb{N}$, there exists a natural isomorphism of mixed Hodge structures

$$\wedge^i \mathbb{H}^1(X_\infty, \mathbb{Z}) \to H^i(X_\infty, \mathbb{Z})$$

that is compatible with the action of $M$ on the underlying $\mathbb{Z}$-modules.

(2) For every $i \in \mathbb{N}$, there exists a natural isomorphism of mixed Hodge structures

$$H^i(X_\infty, \mathbb{Z})^\vee \to H_t(X_\infty, \mathbb{Z})$$

that is compatible with the action of $M$ on the underlying $\mathbb{Z}$-modules.

**Proof.** (1) The cup product defines a morphism

$$\wedge^2 R^1 f_* (\mathbb{Z}) \to R^i f_* (\mathbb{Z})$$

of sheaves on $S$. This is an isomorphism on every fiber, by the proper base change theorem and [19, p. 3], and thus an isomorphism of sheaves. Moreover, it is an isomorphism of variations of Hodge structures, because the cup product defines a morphism of pure Hodge structures on the cohomology of every fiber of $f$ [22, 5.45].

Looking at Schmid’s construction of the limit mixed Hodge structure in [25], one checks in a straightforward way that taking the limit of a variation of Hodge structures commutes with taking exterior powers. Compatibility of the Hodge filtrations is easy, since the exterior power defines a holomorphic map between the relevant classifying spaces. The compatibility of the weight filtrations follows from Corollary 5.6.

Thus, taking the limit at $s$ of the isomorphism (6.2), we obtain an isomorphism of mixed Hodge structures

$$\wedge^i \mathbb{H}^1(X_\infty, \mathbb{Z}) \to H^i(X_\infty, \mathbb{Z})$$

that is compatible with the action of $M$.

(2) For every $i \geq 0$, Poincaré duality yields a natural isomorphism of sheaves of $\mathbb{Z}$-modules

$$\alpha : R^i f_* (\mathbb{Z})^\vee \to R^{2g-i} f_* (\mathbb{Z})(g)$$

(note that Poincaré duality holds with coefficients in $\mathbb{Z}$ because $R^i f_* (\mathbb{Z})$ is a locally free sheaf of $\mathbb{Z}$-modules for all $i \geq 0$). By [22, 6.19], this isomorphism respects the Hodge structure on every fiber of $R^i f_* (\mathbb{Z})^\vee$ and $R^{2g-i} f_* (\mathbb{Z})(g)$. Thus (6.3) is an isomorphism of variations of Hodge structures on $S$. As in (1), one checks in a straightforward way that the limit of $R^i f_* (\mathbb{Z})^\vee$ at $s$ is the dual of the limit of $R^i f_* (\mathbb{Z})$, using Proposition 5.5 to verify the compatibility.
of the weight filtrations. Thus the isomorphism (6.3) induces an isomorphism
of mixed Hodge structures
\[ H^i(X_\infty, \mathbb{Z})^\vee \rightarrow H_i(X_\infty, \mathbb{Z}). \]

□

By Proposition 6.1, in order to describe the limit mixed Hodge structure on
\( H^i(X_\infty, \mathbb{Q}) \) and \( H^i(X_\infty, \mathbb{Z}) \) for all \( i \geq 0 \), it suffices to determine the limit mixed
Hodge structure on \( H_1(X_\infty, \mathbb{Z}) \).

6.2. Description of the mixed Hodge structure on \( H_1(X_\infty, \mathbb{Z}) \). We
denote by \( V \rightarrow S_{\text{an}} \) the polarized variation of Hodge structures
\( R^{2g-1} f_* (\mathbb{Z})(g) \) of type \( \{(0, -1), (-1, 0)\} \) \([8, 4.4.3]\). We denote by \( V^{\mathbb{Z}}, V^{\mathbb{Q}}, \) and \( V^{\mathbb{C}} \) the integer, resp. rational, resp. complex component of \( V \). The sheaf
\( V^{\mathbb{Z}} \) is a locally free sheaf of \( \mathbb{Z} \)-modules on \( S_{\text{an}} \) of rank \( 2g \). The fiber of \( V \) over a point \( z \) of \( S_{\text{an}} \) is
canonically isomorphic to the weight \(-1\) Hodge structure
\( H^{2g-1}(X_z(\mathbb{C}), \mathbb{Z})(g) \), where \( X_z \) denotes the fiber of \( f \) over \( z \). By Poincaré duality, there is a canonical
isomorphism of \( \mathbb{Z} \)-modules
\( H^{2g-1}(X_z(\mathbb{C}), \mathbb{Z})(g) \cong H_1(X_z(\mathbb{C}), \mathbb{Z}) \).

The limit of \( V \) at the point \( s \) is a mixed Hodge structure that was constructed
by Schmid \([25]\). In our notation, this limit is precisely the mixed Hodge
structure \( H_1(X_\infty, \mathbb{Z}) \). For a quick introduction to limit mixed Hodge structures,
we refer to \([13]\) and \([22, \S 10 \text{ and } \S 11]\). Here we will only briefly sketch the main
ingredients of the construction.

We consider a small disc \( \Delta \) around \( s \) in \( S_{\text{an}} \) and we denote by \( \Delta^* \) the punctured
disc \( \Delta \setminus \{ s \} \). It follows from the definition of the nearby cycles functor that
\( H_1(X_\infty, \mathbb{Z}) \) is the \( \mathbb{Z} \)-module of global sections of the pullback of \( V^{\mathbb{Z}} \) to a universal
cover of \( \Delta^* \) \([3, \text{ XIV.1.3.3}]\). By the comparison isomorphism (6.1) and \([2, \text{ IX.3.5}]\),
the action of the monodromy operator \( M^d \) on \( H_1(X_\infty, \mathbb{Q}) \) is unipotent of level
\( \leq 2 \) (meaning that \( (M^d - 1\text{Id})^2 = 0 \)). The level of unipotency can also be
deduced from the Monodromy Theorem \([25, \text{ 6.1}]\) and the fact that the fibers
of \( V \) are of type \( \{(-1, 0), (0, -1)\} \).

The weight filtration \( W \) on \( H_1(X_\infty, \mathbb{Q}) \) is the weight filtration with center \(-1\) associated to the nilpotent operator \( N \) (recall that \( N \) is the logarithm of the unipotent part \( M_u \) of the monodromy operator \( M \)). Since \( M^d_u = M^d \), we have
\[ dN = \log(M^d_u) = M^d - 1\text{Id} \]
so that \( N^2 = 0 \) and the weight filtration is of the form
\[ \{0\} \subset W_{-2} \subset W_{-1} \subset W_0 = H_1(X_\infty, \mathbb{Q}). \]
Explicitly, we have $W_{-1} = \ker(N)$ and $W_{-2} = \im(N)$. This filtration induces a weight filtration on the lattice $H_1(X_\infty, \mathbb{Z})$ in $H_1(X_\infty, \mathbb{Q})$. Note that

$$W_{-1}H_1(X_\infty, \mathbb{Z}) := W_{-1}H_1(X_\infty, \mathbb{Q}) \cap H_1(X_\infty, \mathbb{Z})$$

is stable under the action of $M_s$, since $H_1(X_\infty, \mathbb{Z})$ is stable under $M$ and $M$ is semi-simple on $W_{-1}H_1(X_\infty, \mathbb{Q}) = \ker(M^d - \Id)$.

Consider the finite covering

$$\tilde{\Delta} \to \Delta$$

obtained by taking a $d$-th root $t'$ of the coordinate $t$ on $\Delta$. This covering is totally ramified over the origin $s$ of $\Delta$. With a slight abuse of notation, we denote again by $s$ the unique point of the open disc $\tilde{\Delta}$ that lies over $s \in \Delta$. We denote by $\tilde{\Delta}^*$ the punctured disc $\tilde{\Delta} \setminus \{s\}$.

Pulling back $\mathcal{V}$ to a variation of Hodge structures $\mathcal{V}'$ on $\tilde{\Delta}^*$ has the effect of raising the monodromy operator $M$ to the power $d$. This has no influence on the associated weight filtration on $H_1(X_\infty, \mathbb{Q})$, since the logarithm of $(M_u)^d = M^d$ equals $dN$. Pulling back $\mathcal{V}$ to $\tilde{\Delta}^*$ is the first step in the construction of the limit Hodge filtration $F$ on $H_1(X_\infty, \mathbb{Q})$. The important point is that the monodromy $M^d$ of the variation $\mathcal{V}'$ is unipotent.

Schmid considers a complex manifold $\tilde{\mathcal{D}}$ that parameterizes descending filtrations $F_1 = H_1(X_\infty, \mathbb{C}) \supset F_0 \supset \{0\}$ on $H_1(X_\infty, \mathbb{C})$ that satisfy a certain compatibility relation with the polarization on $H_1(X_\infty, \mathbb{C})$ and such that $F_0$ has dimension $g$. “Untwisting” the monodromy action on the fibers of $\mathcal{V}'$, he constructs a map

$$\tilde{\Psi} : \tilde{\Delta}^* \to \tilde{\mathcal{D}}.$$

The Nilpotent Orbit Theorem (for one variable) in [25, 4.9] guarantees that $\tilde{\Psi}$ extends to a holomorphic map

$$\Psi : \tilde{\Delta} \to \tilde{\mathcal{D}}.$$

The point $\Psi(s)$ of $\tilde{\mathcal{D}}$ corresponds to a descending filtration

$$F^{-1} = H_1(X_\infty, \mathbb{C}) \supset F^0 \supset \{0\}$$

that is called the limit Hodge filtration on $H_1(X_\infty, \mathbb{C})$. Schmid’s fundamental result [25, 6.16] states that the weight filtration $W$ and the limit Hodge filtration $F$ define a polarized mixed Hodge structure on $H_1(X_\infty, \mathbb{Z})$, which is called the limit mixed Hodge structure of the variation of Hodge structures $\mathcal{V}'$.

We see from the shape of the weight filtration and the limit Hodge filtration that the limit mixed Hodge structure on $H_1(X_\infty, \mathbb{Z})$ is of type

$$\{(0, 0), (-1, 0), (0, -1), (-1, -1)\}.$$ 

Moreover, since

$$\text{Gr}^W_{-1}H_1(X_\infty, \mathbb{Z})$$

is polarizable, the mixed Hodge structure

$$(H_1(X_\infty, \mathbb{Z}), W, F)$$

is polarizable.
is a mixed Hodge 1-motive in the sense of [9, §10].

The construction of the limit Hodge filtration can be reformulated in terms of the Deligne extension or canonical extension. Consider the holomorphic vector bundle

$$ (V')^h = V'_Z \otimes \mathcal{O}_\Delta. $$

on the punctured disc $\Delta^*$. The locally constant subsheaf $V'_U$ of this vector bundle defines a connection $\nabla$ on $(V')^h$, called the Gauss-Manin connection. The vector bundle $(V')^h$ extends in a unique way to a vector bundle $\hat{V'}$ on $\tilde{\Delta}^*$. The locally constant subsheaf $V'_C$ of this vector bundle defines a connection $\nabla$ on $(V')^h$, called the Gauss-Manin connection. The vector bundle $(V')^h$ extends in a unique way to a vector bundle $\hat{V'}$ on $\tilde{\Delta}^*$ such that $\nabla$ extends to a logarithmic connection on $\hat{V'}$ whose residue at $s$ is nilpotent [7, 5.2]. We call $\hat{V'}$ the Deligne extension of the variation of Hodge structures $V'$. The fiber of the Deligne extension over the origin $s$ of $\tilde{\Delta}^*$ can be identified with $H_1(X_\infty, C)$, by [22, XI-8] (this identification depends on the choice of a local coordinate on $\tilde{\Delta}$; we take the coordinate $t'$).

The Hodge flags $F_0$ on the fibers of $V'$ glue to a holomorphic subbundle $F_0(V')^h$ of $(V')^h$, which extends uniquely to a holomorphic subbundle $F_0\hat{V'}$ of $\hat{V'}$ [22, 11.10]. Taking fibers at $s$, we obtain a descending filtration

$$ (\hat{V'}_s) = H_1(X_\infty, C) \supset (F_0\hat{V'}_s) \supset \{0\} $$

and this is precisely the limit Hodge filtration on $H_1(X_\infty, C)$ [22, 11.10].

Schmid’s construction of the limit mixed Hodge structure works for abstract variations of Hodge structures, which need not necessarily come from geometry. If the variation of Hodge structures comes from the cohomology of a proper and smooth family over $S$ (such as in the case that we are considering), the Deligne extension and the extension of the Hodge bundles can be constructed explicitly using a relative logarithmic de Rham complex associated to a suitable compactification $\overline{X}$. This is Steenbrink’s construction [26]. We will not need this approach in this paper.

**Theorem 6.2.** We apply the terminology of Section 3.1 to the abelian $K$-variety $A$ and define in this way the degree $d$ extension $K'$ of $K$, as well as the torus $T$ and the abelian variety $B$ over $C$, endowed with a right action of $\mu \cong \mu_d(C)$. There exist canonical $\mu$-equivariant isomorphisms of pure Hodge structures

$$ \text{Gr}_0^W(H_1(X, \mathbb{Q})) \cong H_1(T(C), \mathbb{Q})(-1) $$

$$ \text{Gr}_1^W(H_1(X, \mathbb{Z})) \cong H_1(B(C), \mathbb{Z}) $$

$$ \text{Gr}_2^W(H_1(X, \mathbb{Z})) \cong H_1(T(C), \mathbb{Z}). $$

**Proof.** We denote by $\mathbb{C}(S')$ the algebraic closure of the function field $\mathbb{C}(S)$ in $K'$, and we consider the normalization

$$ \overline{S} \to \overline{S} $$

of $\overline{S}$ in $\mathbb{C}(S')$. This is a ramified Galois covering, obtained by taking a $d$-th root of the local parameter $t$. Its Galois group is canonically isomorphic to $\mu$.
Lars Halvard Halle and Johannes Nicaise

With abuse of notation, we denote again by $s$ the unique point of the inverse image of $s$ in $S'$, and we put $S' = S' \setminus \{s\}$. Then

$$f' : X' = X \times_S S' \rightarrow S'$$

is a smooth projective family of abelian varieties, and we have a canonical isomorphism

$$A' = A \times_K K' \cong X' \times_S \text{Spec } K'.$$

As we’ve argued above, the fact that $A'$ has semi-abelian reduction implies that the variation of Hodge structures

$$V' = V \times_S S' \cong R^{2g-1} f'_*(\mathbb{Z}(g))$$

has unipotent monodromy around $s$.

We denote by $X'$ the Néron model of $X'$ over $S'$, and by $A'$ the Néron model of $A'$ over $R'$, where $R'$ denotes the integral closure of $R$ in $K'$. Note that there is a canonical isomorphism of $R'$-schemes

$$A' \cong X' \times_{S'} \text{Spec } R'.$$

The analytic family of abelian varieties

$$(f')^{\text{an}} : (X')^{\text{an}} \rightarrow (S')^{\text{an}}$$

is canonically isomorphic to the Jacobian

$$J(\mathcal{V}') \rightarrow (S')^{\text{an}}$$

of the variation of Hodge structures $\mathcal{V}'$ [24, 2.10.1]. We will now explain the relation between the complex semi-abelian variety $(A')_{\text{os}}$ and the limit mixed Hodge structure $H_1(X_{\infty}, \mathbb{Z})$ of $\mathcal{V}'$ at the point $s$. To simplify notation, we put $H_C = H_1(X_{\infty}, C)$ for $C = \mathbb{Z}, \mathbb{Q}, \mathbb{C}$, and we denote by $H$ the mixed Hodge structure

$$(H_\mathbb{Z}, W^\bullet H_\mathbb{Q}, F^\bullet H_\mathbb{C}).$$

By [24, 4.5(i)], $(X')^{\text{an}}$ is canonically isomorphic to Clemens’ Néron model of $X'$ over $S'$; see [6] and [24, 2.5] for a definition. In [24, 2.5], Clemens’ Néron model is constructed by gluing copies of the Zucker extension $J_Z^\mathbb{Z}(\mathcal{V}')$ of $\mathcal{V}'$, which is defined in [29] and [24, 2.1]. It follows immediately from this construction that the identity component

$$(A')^{\text{os}}$$

of Clemens’ Néron model is canonically isomorphic to the Zucker extension $J_Z^\mathbb{Z}(\mathcal{V}')$.

The Zucker extension is given explicitly by

$$J_Z^\mathbb{Z}(\mathcal{V}') = j_* \mathcal{V}'_Z / F^0 \mathcal{V}'$$

where $\mathcal{V}'$ is the Deligne extension of $\mathcal{V}'$ to $S'$, $j$ is the open immersion of $S'$ into $S$, and $F^0 \mathcal{V}'$ is the unique extension of the holomorphic vector bundle

$$F^0(\mathcal{V}'_Z \otimes_{\mathbb{Z}} \mathcal{O}(S')^{\text{an}})$$
to a holomorphic subbundle of $\tilde{V}'$. We can describe the fiber

$$J^\mathbb{Z}_S(\tilde{V}')_s \cong ((\mathcal{A}')^\alpha)_s^{\text{an}} \cong ((\mathcal{A}')_s^\alpha)^{\text{an}}$$

of $J^\mathbb{Z}_S(\tilde{V}')$ at $s$ in terms of the mixed Hodge structure $H$, as follows.

As we’ve explained above, the fiber of $\tilde{V}'$ over $s$ is isomorphic to $H_{\mathbb{C}}$, and $F^0(\tilde{V}')$ coincides with the degree zero part of the Hodge filtration on $H_{\mathbb{C}}$.

Moreover, by definition of the weight filtration on $H$, the fiber of $F^\alpha\mathcal{V}'_s$ at $s$ is the $\mathbb{Z}$-module of elements in $H_{\mathbb{Z}}$ that are invariant under the monodromy action of $M^d$. By definition, the weight filtration on $H_{\mathbb{Q}}$ is the filtration centered at $-1$ defined by the logarithm $N$ of the unipotent part $M_u$ of $M$, or, equivalently, the logarithm $N' = dN$ of $M^d$.

Since $(M^d - \text{Id})^2 = 0$, we have $N' = M^d - \text{Id}$ and $(N')^2 = 0$, and we see that

$$(j_*\mathcal{V}'_s)_s = \ker(N') = W_{-1}H_{\mathbb{Z}}.$$

Thus, we find canonical isomorphisms

$$((\mathcal{A}')_s^\alpha)^{\text{an}} \cong J^\mathbb{Z}_S(\tilde{V}')_s \cong W_{-1}H_{\mathbb{Z}}/H_{\mathbb{C}}/F^0H_{\mathbb{C}} \cong W_{-1}H_{\mathbb{Z}}/W_{-1}H_{\mathbb{C}}/(F^0H_{\mathbb{C}} \cap W_{-1}H_{\mathbb{C}}).$$

The last isomorphism is obtained as follows: since $\text{Gr}^W_0H$ is purely of type $(0,0)$, we have

$$F^0\text{Gr}^W_0H_{\mathbb{C}} = \text{Gr}^W_0H_{\mathbb{C}} = H_{\mathbb{C}}/W_{-1}H_{\mathbb{C}},$$

so that the morphism

$$W_{-1}H_{\mathbb{C}} \to H_{\mathbb{C}}/F^0H_{\mathbb{C}}$$

induced by the inclusion of $W_{-1}H_{\mathbb{C}}$ in $H_{\mathbb{C}}$ is surjective. Its kernel is precisely $F^0H_{\mathbb{C}} \cap W_{-1}H_{\mathbb{C}}$.

By [9, 10.1], we have an extension

$$(6.4) \quad 0 \to J(\text{Gr}^W_{-2}H) \to ((\mathcal{A}')_s^\alpha)^{\text{an}} \to J(\text{Gr}^W_{-1}H) \to 0$$

where

$$J(\text{Gr}^W_{-2}H) = \text{Gr}^W_{-2}H_{\mathbb{C}}/\text{Gr}^W_{-2}H_{\mathbb{Z}}$$

is a torus, and

$$J(\text{Gr}^W_{-1}H) = H_{\mathbb{Z}}/\text{Gr}^W_{-1}H_{\mathbb{C}}/F^0\text{Gr}^W_{-1}H_{\mathbb{C}}$$

an abelian variety, because the Hodge structure $\text{Gr}^W_{-1}H$ is polarizable. By [9, 10.1.3.3], the extension (6.4) must be the analytification of the Chevalley decomposition

$$0 \to T \to (\mathcal{A}')_s^\alpha \to B \to 0.$$

Hence, there exist canonical isomorphisms of pure Hodge structures

$$(6.5) \quad \text{Gr}^W_1(H) \cong H_1(B(\mathbb{C}), \mathbb{Z}),$$

$$(6.6) \quad \text{Gr}^W_2(H) \cong H_1(T(\mathbb{C}), \mathbb{Z}).$$

Moreover, by definition of the weight filtration on $H_{\mathbb{Q}}$, the operator $N$ defines a $\mu$-equivariant isomorphism of $\mathbb{Q}$-Hodge structures

$$\text{Gr}^W_0(H) \otimes_{\mathbb{Z}} \mathbb{Q} \to \text{Gr}^W_{-2}(H)(-1) \otimes_{\mathbb{Z}} \mathbb{Q}.$$
It remains to show that the isomorphisms (6.5) and (6.6) are \( \mu \)-equivariant. It is enough to prove that the Galois action of \( \mu \) on 

\[ \mathcal{V}' \rightarrow S' \]

extends analytically to the Zucker extension

\[ J_{\mathbb{Z}}^{\mathcal{V}}(\mathcal{V}') \rightarrow \overline{\mathfrak{s}} \]

in such a way that the action of the canonical generator of \( \mu = \mu_\mathbb{d}(\mathbb{C}) \) on 

\[ J_{\mathbb{Z}}^{\mathcal{V}}(\mathcal{V}')_s = W^{-1}_{-1}H_\mathbb{Z}\backslash H_\mathbb{C}/F^0H_\mathbb{C} \]

coinsides with the semi-simple part \( M_s \) of the monodromy action. This follows easily from the constructions. \( \square \)

6.3. Multiplicity functions and limit mixed Hodge structure.

**Theorem 6.3.** We keep the notations of Section 6.1.

1. The potential toric rank \( t_{\text{pot}}(A) \) is equal to the largest integer \( \alpha \) such that \( \exp(2\pi i c(A)i) \) is an eigenvalue of \( M_s \) on \( \text{Gr}^{W}_{g+\alpha}H^g(X_\infty, \mathbb{Q}) \).

2. The Jordan form of \( M_s \) on 

\[ \text{Gr}^{W}_{-1}H_1(X_\infty, \mathbb{Q})^{1,0} \]

is \( \text{Jord}(m^{ab}_{A}, 0) \), 

\[ \text{Gr}^{W}_{-1}H_1(X_\infty, \mathbb{Q})^{0,1} \]

is \( \text{Jord}(\tilde{m}^{ab}_{A}, 0) \), 

\[ \text{Gr}^{W}_{2}H_1(X_\infty, \mathbb{Q}) \]

is \( \text{Jord}(m^{ws}_{A}, 0) \), 

\[ \text{Gr}^{W}_{0}H_1(X_\infty, \mathbb{Q}) \]

is \( \text{Jord}(m^{ws}_{A}, 0) \).

**Proof.** We denote by \( M_u \) the unipotent part of the monodromy operator \( M \), and by \( N \) its logarithm. By definition, the weight filtration on \( H^g(X_\infty, \mathbb{Q}) \) is the filtration with center \( g \) associated to the nilpotent operator \( N \). Thus (1) follows from Proposition 5.2, Theorem 5.9 and the isomorphism (6.1).

Point (2) follows from Theorem 6.2 and the canonical \( \mu \)-equivariant isomorphisms

\[ H_1(B(\mathbb{C}), \mathbb{C})^{1,0} \cong \text{Lie}(B) \]

\[ H_1(B(\mathbb{C}), \mathbb{C})^{0,1} \cong \text{Lie}(B^\vee) \]

\[ H_1(T(\mathbb{C}), \mathbb{C}) \cong \text{Lie}(T) \]

(see [19, pp. 4 and 86] for the dual isomorphisms on the level of cohomology). \( \square \)

**References**

[1] Schémas en groupes. I: Propriétés générales des schémas en groupes. Séminaire de Géométrie Algébrique du Bois Marie 1962/64 (SGA 3). Dirigé par M. Demazure et A. Grothendieck. Lecture Notes in Mathematics, Vol. 151. Springer-Verlag, Berlin, 1970.
Jumps and Monodromy of Abelian Varieties

[2] Groupes de monodromie en géométrie algébrique. I. Séminaire de Géométrie Algébrique du Bois-Marie 1967–1969 (SGA 7 I), Dirigé par A. Grothendieck. Avec la collaboration de M. Raynaud et D. S. Rim. Lecture Notes in Mathematics, Vol. 288. Springer-Verlag, Berlin, 1972.
[3] Groupes de monodromie en géométrie algébrique. II. Séminaire de Géométrie Algébrique du Bois-Marie 1967–1969 (SGA 7 II), Dirigé par P. Deligne et N. Katz. Lecture Notes in Mathematics, Vol. 340. Springer-Verlag, Berlin, 1973.
[4] S. Bosch, W. Lütkebohmert, and M. Raynaud. Néron models, volume 21 of Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer-Verlag, 1990.
[5] C.-L. Chai. Néron models for semialbelian varieties: congruence and change of base field. Asian J. Math., 4(4):715–736, 2000.
[6] H. Clemens. The Néron model for families of intermediate Jacobians acquiring “algebraic” singularities. Inst. Hautes Études Sci. Publ. Math. 58:5–18, 1983.
[7] P. Deligne. Équations différentielles à points singuliers réguliers, volume 163 of Lecture Notes in Mathematics. Berlin: Springer-Verlag, 1970.
[8] P. Deligne. Théorie de Hodge: II. Inst. Hautes Études Sci. Publ. Math., 40:5–57, 1971.
[9] P. Deligne. Théorie de Hodge: III. Inst. Hautes Études Sci. Publ. Math., 44:5–77, 1974.
[10] P. Deligne. La conjecture de Weil: II. Inst. Hautes Études Sci. Publ. Math., 52:137-252, 1980.
[11] A. Dimca. Sheaves in topology. Universitext. Springer-Verlag, Berlin, 2004.
[12] B. Edixhoven. Néron models and tame ramification. Compos. Math., 81:291–306, 1992.
[13] R. Hain. Periods of limit mixed Hodge structures. In: Current developments in mathematics, pages 113-133, Int. Press, Somerville, MA, 2003.
[14] I.H. Halle and J. Nicaise. Motivic zeta functions of abelian varieties, and the monodromy conjecture. Adv. Math., 227:610–653, 2011.
[15] M. Kashiwara. Holonomic systems of linear differential equations with regular singularities and related topics in topology. In: Algebraic varieties and analytic varieties, Proc. Symp., Tokyo 1981, volume 1 of Adv. Stud. Pure Math., pages 49–54, North-Holland, Amsterdam-New York, 1983.
[16] F. Loeser. Fonctions d’Igusa p-adiques et polynômes de Bernstein. Am. J. of Math., 110:1–22, 1988.
[17] B. Malgrange. Polynômes de Bernstein-Sato et cohomologie évanescente. Astérisque, 101/102:243–267, 1983.
[18] J.S. Milne. Abelian varieties. In: G. Cornell and J.H. Silverman (eds.), Arithmetic geometry. Springer Verlag, 1986.
[19] D. Mumford. Abelian varieties. With appendices by C. P. Ramanujam and Yuri Manin, 2nd ed., volume 5 of Tata Institute of Fundamental Research Studies in Mathematics. London: Oxford University Press, 1974.
[20] J. Nicaise. An introduction to p-adic and motivic zeta functions and the monodromy conjecture. In: G. Bhowmik, K. Matsumoto and H. Tsumura (eds.), *Algebraic and analytic aspects of zeta functions and L-functions*, volume 21 of *MSJ Memoirs*, Mathematical Society of Japan, pages 115-140, 2010.

[21] T. Oda. The first de Rham cohomology group and Dieudonné modules. *Ann. Sci. École Norm. Sup. (4)*, 2(1):63-135, 1969.

[22] C. Peters and J.H.M. Steenbrink. *Mixed Hodge structures*, volume 52 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge*. Berlin: Springer, 2008.

[23] B. Rodrigues. On the monodromy conjecture for curves on normal surfaces. *Math. Proc. Camb. Philos. Soc.*, 136(2):313-324, 2004.

[24] M. Saito. Admissible normal functions. *J. Algebraic Geom.* 5, 235-276, 1996.

[25] W. Schmid. Variation of Hodge structure: the singularities of the period mapping. *Invent. Math.* 22:211-319, 1973.

[26] J.H.M. Steenbrink. Limits of Hodge structures. *Invent. Math.*, 31:229–257, 1976.

[27] J.H.M. Steenbrink. Mixed Hodge structure on the vanishing cohomology. In: *Real and complex singularities (Proc. Ninth Nordic Summer School/NAVF Sympos. Math., Oslo, 1976)*, pages 525–563. Alphen aan den Rijn: Sijthoff and Noordhoff, 1977.

[28] J. Tate and F. Oort. Group schemes of prime order. *Ann. Sci. École Norm. Sup. (4)* 3:1-21, 1970.

[29] S. Zucker. Generalized intermediate Jacobians and the theorem on normal functions. *Invent. Math.* 33(3):185–222, 1976.

Lars Halvard Halle  
Institut für Algebraische Geometrie  
Gottfried-Wilhelm-Leibniz-Universität  
30167 Hannover  
Germany  
larshhal@math.uio.no

Johannes Nicaise  
KULeuven  
Department of Mathematics  
Celestijnenlaan 200B  
3001 Heverlee  
Belgium  
johannes.nicaise@wis.kuleuven.be

Mathematisk Institutt  
Universitetet i Oslo  
Postboks 1053  
Blindern  
0316 Oslo  
Norway

**Documenta Mathematica 16 (2011) 937–968**
ON FAMILIES OF WEAKLY ADMISSIBLE FILTERED ϕ-MODULES AND THE ADJOINT QUOTIENT OF GL_d

EUGEN HELLMANN

Received: April 8, 2011
Revised: November 20, 2011
Communicated by Peter Schneider

Abstract. We study the relation of the notion of weak admissibility in families of filtered ϕ-modules, as considered in [He], with the adjoint quotient. We show that the weakly admissible subset is an open subvariety in the fibers over the adjoint quotient. Further we determine the image of the weakly admissible set in the adjoint quotient generalizing earlier work of Breuil and Schneider.

2010 Mathematics Subject Classification: 11F80, 11F85, 14G20, 14G22
Keywords and Phrases: Filtered ϕ-modules, p-adic Hodge theory, rigid geometry

1. Introduction

Filtered ϕ-modules appear in p-adic Hodge-theory as a category of linear algebra data describing crystalline representations of the absolute Galois group of a local p-adic field. More precisely, there is an equivalence of categories between crystalline representations and weakly admissible filtered ϕ-modules, see [CF]. Here weak admissibility is a semi-stability condition relating the slopes of the ϕ-linear endomorphism Φ with the filtration.

In our companion paper [He] we define and study arithmetic families of filtered ϕ-modules and crystalline representations. Our families are parameterized by rigid analytic spaces or adic spaces in the sense of Huber, see [Hu2] for example. We show that the condition of being weakly admissible is an open condition [He, Theorem 1.1] and that there is an open subset of the weakly admissible locus over which there exists a family of crystalline representations giving rise to the family of filtered ϕ-modules [He, Theorem 1.3].

In this paper we study the weakly admissible locus in more detail. In the setting of period domains in the sense of Rapoport and Zink [RZ], the weakly admissible locus is an admissible open subset of a flag variety. Contrarily, the
weakly admissible locus in our set up has an algebraic nature as soon as we fix the Frobenius $\Phi$, or even the conjugacy class of its semi-simplification. Further we analyze the image of the weakly admissible locus in the adjoint quotient. The question whether there exists a weakly admissible filtration for a fixed conjugacy class of the semisimplification of the Frobenius already appears in work of Breuil and Schneider [BS] on the $p$-adic Langlands correspondence. Unlike the characterization in [BS], our characterization of the set of automorphisms $\Phi$ for which there exists a weakly admissible filtration is purely in terms of the adjoint quotient of $GL_d$.

Our main results are as follows: Fix a finite extension $K$ of $\mathbb{Q}_p$ and write $K_0$ for the maximal unramified extension of $\mathbb{Q}_p$ inside $K$. Let $d > 0$ be an integer and denote by $A \subset GL_d$ the diagonal torus. For a dominant cocharacter $\nu : \mathbb{G}_m, \overline{\mathbb{Q}}_p \to (\text{Res}_K/\mathbb{Q}_p A)_{\overline{\mathbb{Q}}_p}$ we write $\text{Gr}_\nu$ for the partial flag variety of $\text{Res}_K/\mathbb{Q}_p GL_d$ parametrizing flags of "type $\nu$", see section 3.1 for the precise definition. This variety is defined over the reflex field $E$ of $\nu$. As in [He, 4.1] we denote by $D_\nu = ((\text{Res}_{K_0}/\mathbb{Q}_p GL_d)_E \times \text{Gr}_\nu)/(\text{Res}_{K_0}/\mathbb{Q}_p GL_d)_E$ the stack of filtered $\phi$-modules with filtration of "type $\nu$" on the category of adic spaces locally of finite type. The action of $(\text{Res}_{K_0}/\mathbb{Q}_p GL_d)_E$ is explicitly described in (3.4). Let $W$ denote the Weyl group of $GL_d$. We will define a morphism

$$\alpha : D_\nu \to (A/W)^{\text{ad}}$$

to the adification (see [Hu2, Remark 4.6 (i)]) of the adjoint quotient $A/W$ and prove the following theorem.

**Theorem 1.1.** Let $x \in (A/W)^{\text{ad}}_E$ and form the 2-fiber product

$$\begin{array}{ccc}
\alpha^{-1}(x)^{\text{wa}} & \to & D_\nu^{\text{wa}} \\
\downarrow & & \downarrow \\
x & \to & (A/W)^{\text{ad}}_E.
\end{array}$$

Then there Artin stack in schemes $\mathfrak{A}$ over the field $k(x)$ such that

$$\alpha^{-1}(x)^{\text{wa}} = \mathfrak{A}^{\text{ad}}.$$

The stack $\mathfrak{A}$ is the stack quotient of a quasi-projective $k(x)$-variety.

Further we determine the image of the weakly admissible locus $D_\nu^{\text{wa}}$ under the morphism $\alpha$. The description of this image works in the category of analytic spaces in the sense of Berkovich.

**Theorem 1.2.** Let $\nu$ be a dominant coweight as above. There is a dominant coweight $\mu(\nu)$ of $GL_d$ associated to $\nu$ such that

$$\alpha^{-1}(x)^{\text{wa}} \neq \emptyset \iff x \in (A/W)^{\leq \mu(\nu)}.$$

Here $(A/W)^{\leq \mu(\nu)}$ is a Newton-stratum in the sense of Kottwitz [Ko].
The coweight $\mu(\nu)$ which appears in the theorem is explicit and defined in Definition 5.4.

ACKNOWLEDGEMENTS: I thank my advisor M. Rapoport for his interest and advice. Further I thank the referee for his careful reading and his comments. This work was supported by the SFB/TR 45 "Periods, Moduli Spaces and Arithmetic of Algebraic Varieties" of the DFG (German Research Foundation).

2. Filtered $\varphi$-modules

Throughout this section we denote by $F$ a topological field containing $\mathbb{Q}_p$ with a continuous valuation $v_F : F \rightarrow \Gamma_F \cup \{0\}$ in the sense of [Hu1, 2, Definition] that is $\Gamma_F$ is a totally ordered abelian group (written multiplicative) and

\[
\begin{align*}
v_F(0) &= 0, \\
v_F(1) &= 1, \\
v_F(ab) &= v_F(a)v_F(b), \\
v_F(a+b) &\leq \max\{v_F(a), v_F(b)\},
\end{align*}
\]

where the order on $\Gamma_F$ is extended to $\Gamma_F \cup \{0\}$ by $0 < \gamma$ for all $\gamma \in \Gamma_F$. We will introduce the notion of a filtered $\varphi$-module with coefficients in $F$ and define weak admissibility for these objects.

Recall that $K_0$ is an unramified extension of $\mathbb{Q}_p$ with residue field $k$ and write $f = [K_0 : \mathbb{Q}_p]$. We write $\varphi$ for the lift of the absolute Frobenius to $K_0$.

2.1. $\varphi$-modules with coefficients.

In this subsection we define and study what we call isocrystals over $k$ with coefficients in $F$.

Definition 2.1. An isocrystal over $k$ with coefficients in $F$ is a free $F \otimes \mathbb{Q}_p K_0$-module $D$ of finite rank together with an automorphism $\Phi : D \rightarrow D$ that is semi-linear with respect to $\id \otimes \varphi : F \otimes \mathbb{Q}_p K_0 \rightarrow F \otimes \mathbb{Q}_p K_0$.

A morphism $f : (D, \Phi) \rightarrow (D', \Phi')$ is an $F \otimes \mathbb{Q}_p K_0$-linear map $f : D \rightarrow D'$ such that

\[f \circ \Phi = \Phi' \circ f.\]

The category of isocrystals over $k$ with coefficients in $F$ is denoted by $\text{Isoc}(k)_F$.

It is easy to see that $\text{Isoc}(k)_F$ is an $F$-linear abelian $\otimes$-category with the obvious notions of direct sums and tensor products.

Remark 2.2. (i) Given an $F \otimes \mathbb{Q}_p K_0$-module $D$ of finite type, the existence of a semi-linear automorphism $\Phi : D \rightarrow D$ implies that $D$ is free over $F \otimes \mathbb{Q}_p K_0$.

This fact will be used in the sequel.

(ii) In the classical setting an isocrystal over $k$ is a finite-dimensional $K_0$-vector space with $\varphi$-linear automorphism $\Phi$, i.e. an object in $\text{Isoc}(k)_{\mathbb{Q}_p}$.

If $F$ is finite over $\mathbb{Q}_p$, then an isocrystal over $k$ with coefficients in $F$ is the same as an object $(D, \Phi) \in \text{Isoc}(k)_{\mathbb{Q}_p}$ together with a map $F \rightarrow \text{End}_k(D)$.
where the subscript \( \Phi \) on the right hand side indicates that the endomorphisms commute with \( \Phi \) (compare [DOR, VIII, 5] for example). This is clearly equivalent to our definition.

Let \( F' \) be an extension of \( F \) with valuation \( v_{F'} : F' \to \Gamma_{F'} \cup \{0\} \) extending the valuation \( v_F \). The extension of scalars from \( F \) to \( F' \) is the functor
\[
(2.1) \quad - \otimes_F F' : \text{Isoc}(k)_F \to \text{Isoc}(k)_{F'},
\]
that maps \((D, \Phi) \in \text{Isoc}(k)_F\) to the object \((D \otimes_F F', \Phi \otimes \text{id})\).

If \( F' \) is a finite extension of \( F \), then we also define the restriction of scalars
\[
(2.2) \quad \epsilon_{F'/F} : \text{Isoc}(k)_{F'} \to \text{Isoc}(k)_{F}.
\]
This functor maps \((D', \Phi') \in \text{Isoc}(k)_{F'}\) to itself, forgetting the \( F' \)-action but keeping the \( F \)-action.

We write \( \Gamma_F \otimes \mathbb{Q} \) for the localisation of the abelian group \( \Gamma_F \). Then every element \( \gamma' \in \Gamma_F \otimes \mathbb{Q} \) can be written as a single tensor \( \gamma \otimes r \) and we extend the total order of \( \Gamma_F \) to \( \Gamma_F \otimes \mathbb{Q} \) by
\[
a \otimes \frac{1}{m} < b \otimes \frac{1}{n} \iff a^n < b^m.
\]

**Definition 2.3.** Let \((D, \Phi) \in \text{Isoc}(k)_F\) and \( d = \text{rk}_{F \otimes \mathbb{Q}_p, K_0} D \). The map \( \Phi^f : D \to D \) is an \( F \)-linear map on the \( f \)-dimensional \( F \)-vector space \( D \).

(i) Define the Newton slope of \((D, \Phi)\) as
\[
\lambda_N^{(F)}(D, \Phi) = v_F(\det_F \Phi^f) \otimes \frac{1}{d} \in \Gamma_F \otimes \mathbb{Q}.
\]
Here \( \det_F \) means that we take the determinant of an \( F \)-linear map on an \( F \)-vector space.

(ii) Let \( \lambda \in \Gamma_F \otimes \mathbb{Q} \). An object \((D, \Phi) \in \text{Isoc}(k)_F\) is called purely of Newton-slope \( \lambda \) if for all \( \Phi \)-stable \( F \otimes \mathbb{Q}_p, K_0 \)-submodules \( D' \subset D \) we have \( \lambda_N^{(F)}(D', \Phi|_{D'}) = \lambda \).

**Lemma 2.4.** Let \( F' \) be an extension of \( F \) with valuation \( v_{F'} \) extending \( v_F \) and \((D, \Phi) \in \text{Isoc}(k)_F\). Then
\[
\lambda_N^{(F')} \circ (D \otimes_F F', \Phi \otimes \text{id}) = \lambda_N^{(F)}(D, \Phi).
\]
If in addition \( F' \) is finite over \( F \) and \((D', \Phi') \in \text{Isoc}(k)_{F'}\), then
\[
\lambda_N^{(F)} \circ (\epsilon_{F'/F}(D', \Phi')) = \lambda_N^{(F')}(D', \Phi').
\]

**Proof.** These are straightforward computations. \( \Box \)

As the Newton slope is preserved under extension and restriction of scalars we will just write \( \lambda_N \) in the sequel.

**Remark 2.5.** Let \((D, \Phi) \in \text{Isoc}(k)_{\mathbb{Q}_p}\) be an object of rank \( d \) and denote for the moment by \( v_p \) the usual \( p \)-adic valuation on \( \mathbb{Q}_p \). Write \( | - | = p^{-v_p(\cdot)} \). Then the value group of the absolute value is \( \Gamma_{\mathbb{Q}_p} = p\mathbb{Z} \) and we identify \( \Gamma_{\mathbb{Q}_p} \otimes \mathbb{Q} \) with the subgroup \( p\mathbb{Z} \) of \( \mathbb{R} \setminus \{0\} \). Our definitions then imply
\[
(2.3) \quad \lambda_N(D, \Phi) = p^{-\frac{1}{d} v_p(\det_{K_0} \Phi)}.
\]
Filtered $\varphi$-modules and the adjoint quotient

Here $v_p(\det_{K_0} \Phi)$ is the $p$-adic valuation of the determinant over $K_0$ of any matrix representing the semi-linear map $\Phi$ in some chosen basis. This matrix is well defined up to $\varphi$-conjugation and hence the valuation of the determinant is independent of choices. Note that (the negative of) the exponent in (2.3) is the usual Newton slope of the isocrystal $(D, \Phi)$ over $k$, compare [Zi] for example.

**Proposition** 2.6. Let $(D, \Phi) \in \text{Isoc}(k)_F$, then there exist unique elements $\lambda_1 < \lambda_2 < \cdots < \lambda_r \in \Gamma_F \otimes \mathbb{Q}$ and a unique decomposition

$$D = D_1 \oplus D_2 \oplus \cdots \oplus D_r$$

of $D$ into $\Phi$-stable $F \otimes_{\mathbb{Q}_p} K_0$-submodules such that $(D_i, \Phi|_{D_i})$ is purely of Newton slope $\lambda_i$.

**Proof.** First we show the existence of such a decomposition. The uniqueness will then follow from Lemma 2.9 below.

**Step 1:** Assume first that there exists an embedding $\psi_0 : K_0 \hookrightarrow F$. We obtain an isomorphism

$$F \otimes_{\mathbb{Q}_p} K_0 \cong \prod_{\psi : K_0 \hookrightarrow F} F.$$ 

The endomorphism $\text{id} \otimes \varphi$ on the left hand side translates to the shift of the factors on the right hand side. Further we obtain the corresponding decomposition

$$D = \prod_{\psi} V_\psi$$

into $F$-vector spaces $V_\psi$ and $F$-linear isomorphisms

$$\Phi_\psi = \Phi|_{V_\psi} : V_\psi \xrightarrow{\cong} V_{\psi \circ \varphi}.$$

There is a bijection between the $\Phi$-stable subspaces $D'$ of $D$ and the $\Phi^f|_{V_{\psi_0}}$-stable subspaces of $V_{\psi_0}$ given by $D' \mapsto D' \cap V_{\psi_0}$.

Given $D' \subset D$ and $U = D' \cap V_{\psi_0}$ we have

$$\lambda_N(D', \Phi|_{D'}) = v_F(\det_F \Phi^f|_U) \otimes \frac{1}{\dim_F U} \in \Gamma_F \otimes \mathbb{Q}.$$ 

Hence the desired decomposition of $D$ is induced by the decomposition of $V_{\psi_0}$ into the maximal $\Phi^f|_{V_{\psi_0}}$-stable subspaces $U \subset V_{\psi_0}$ such that

$$v_F(\det_F \Phi^f|_U) \otimes \frac{1}{\dim_F U} = \lambda_i$$

for all $\Phi^f$-stable subspaces $U' \subset U$.

**Step 2:** If there is no embedding $\psi$ of $K_0$ into $F$, then we find a finite extension $F' = FK_0$ of $F$ such that $K_0$ embeds into $F'$. We want to deduce the result from Step 1 by Galois descent. The extension $F'$ is Galois over $F$, as $K_0$ is Galois over $\mathbb{Q}_p$. Further we extend the valuation from $F$ to $F'$ by setting $v_F(\mathcal{O}_{K_0}) = \{1\}$, where $\mathcal{O}_{K_0} \subset K_0$ is the ring of integers.
Write \((D', \Phi')\) for the extension of scalars of \((D, \Phi) \in \text{Isoc}(k)_F\) to \(\text{Isoc}(k)_{F'}\). Then there exists \(\lambda_1 < \lambda_2 < \cdots < \lambda_r \in \Gamma_F \otimes \mathbb{Q}\) and a decomposition
\[
D' = D'_1 \oplus D'_2 \oplus \cdots \oplus D'_r
\]
such that the \(D'_i\) are \(\Phi'\)-stable and \((D'_i, \Phi'|_{D'_i})\) is purely of slope \(\lambda_i\). Now the action of the Galois group \(\text{Gal}(F'/F)\) preserves the valuation on \(F'\) and hence also the Newton slope of a \(\Phi'\)-stable subobject of \(D'\). It follows that \(\text{Gal}(F'/F)\) preserves the decomposition \((2.4)\) and hence this decomposition descends to \(D\).

\(\square\)

Remark 2.7. Proposition 2.6 replaces the slope decomposition in the classical context (c.f. [Zi, VI, 3] for example).

Definition 2.8. Let \((D, \Phi) \in \text{Isoc}(k)_F\) and denote by \(D = \bigoplus D_i\) a decomposition of \(D\) into \(\Phi\)-stable submodules purely of slope \(\lambda_i \in \Gamma_F \otimes \mathbb{Q}\) as in Proposition 2.6. We will refer to this as the slope decomposition. Further, for \(\lambda \in \Gamma_F \otimes \mathbb{Q}\) we define
\[
D_{\lambda} = \begin{cases} 
D_i, & \lambda = \lambda_i \\
0, & \text{otherwise.}
\end{cases}
\]

Lemma 2.9. Let \(f : (D, \Phi) \to (D', \Phi')\) be a morphism in \(\text{Isoc}(k)_{F}\). Consider slope decompositions \(D = \bigoplus D_i\) and \(D' = \bigoplus D'_j\) as in Proposition 2.6. Then for all \(\lambda \in \Gamma_F \otimes \mathbb{Q}\)
\[
f(D_{\lambda}) \subset D'_{\lambda}.
\]

Proof. This is an immediate consequence of \(f \circ \Phi = \Phi' \circ f\).

\(\square\)

2.2. Filtered Isocrystals with coefficients. Recall that \(K\) is a totally ramified extension of \(K_0\). We denote by \(e = [K : K_0]\) the ramification index of \(K\). In this section we define the basic object of our study.

Definition 2.10. A \(K\)-filtered isocrystal over \(k\) with coefficients in \(F\) is a triple \((D, \Phi, F^*)\), where \((D, \Phi) \in \text{Isoc}(k)_F\) and \(F^*\) is a descending, separated and exhaustive \(\mathbb{Z}\)-filtration on \(D_K = D \otimes_{K_0} K\) by (not necessarily free) \(F \otimes_{\mathbb{Q}_p} K\)-submodules.

A morphism
\[
f : (D, \Phi, F^*) \to (D', \Phi', F'^*)
\]
is a morphism \(f : (D, \Phi) \to (D', \Phi')\) in \(\text{Isoc}(k)_F\) such that \(f \otimes \text{id} : D_K \to D'_{K}\) respects the filtrations.

The category of \(K\)-filtered isocrystals over \(k\) with coefficients in \(F\) is denoted by \(\text{Fil Isoc}(k)_F^K\).

It is easy to see that \(\text{Fil Isoc}(k)_F^K\) is an \(F\)-linear \(\otimes\)-category. Further it has obvious notions of kernels, cokernels and exact sequences. For an extension \(F'\) of \(F\) we again have an extension of scalars like in \((2.1)\),
\[
\otimes_{\mathbb{F}} F' : \text{Fil Isoc}(k)_F^K \to \text{Fil Isoc}(k)_{F'}^K.
\]
If $F'$ is finite over $F$, we also have a restriction of scalars like in (2.2),
\[
\epsilon_{F'/F} : \Fil \Isoc(k)_{F'}^K \to \Fil \Isoc(k)_F.
\]

In the following we will often shorten our notation and just write $D$ for an object $(D, \Phi, \Fil)$ in $\Fil \Isoc(k)_F^K$.

We now want to develop a slope theory for filtered isocrystals and define weakly admissible objects.

**Definition 2.11.** Let $(D, \Phi, \Fil) \in \Fil \Isoc(k)_F^K$. We define
\[
\deg(D) = \left( v_F(\det_F(\Phi^f)) \otimes \frac{1}{F'} \right)^{-1} \deg_F(D) \in \Gamma_F \otimes Q.
\]

We call $\mu(D)$ the slope of $D$.

**Remark 2.12.** As in Lemma 2.4, one easily sees that the slope $\mu_D$ is preserved under extension and restriction of scalars. Hence we will just write $\mu$ in the sequel.

Now we have a Harder-Narasimhan formalism as in [DOR, Chapter 1]. The only difference is that our valuations are written multiplicatively, while in the usual theory they are written additively. We will only sketch the proofs and refer to [DOR] for the details.

**Lemma 2.13.** Let $0 \to D' \to D \to D'' \to 0$ be a short exact sequence in $\Fil \Isoc(k)_F^K$. Then
\[
\deg(D) = \deg(D') \deg(D'').
\]

Further
\[
\max\{\mu(D'), \mu(D'')\} \geq \mu(D) \geq \min\{\mu(D'), \mu(D'')\}.
\]
The sequence $\mu(D), \mu(D), \mu(D')$ is either strictly increasing or strictly decreasing or stationary.

**Proof.** The first assertion is obvious from the definitions and the second is a direct consequence.

**Lemma 2.14.** Let $f : D \to D'$ be a morphism in $\Fil \Isoc(k)_F^K$. Then
\[
\deg_f(\text{coim } f) \geq \deg_f(\text{im } f).
\]

**Proof.** Replacing $D$ by $\text{coim } f$ and $D'$ by $\text{im } f$, we may assume that $f$ is an isomorphism in $\Isoc(k)_F$. Now the assertion follows easily from
\[
(f \otimes \text{id})(\Fil) \subset \Fil.
\]
Definition 2.15. An object \((D, \Phi, \mathcal{F}^\bullet) \in \text{Fil Isoc}(k)^K_F\) is called semi-stable if, for all \(\Phi\)-stable subobjects \(D' \subset D\), we have \(\mu(D') \geq \mu(D)\). It is called stable if the inequality is strict for all proper subobjects. Finally, \(D\) is called weakly admissible if it is semi-stable of slope 1.

Note that semi-stability is defined using "\(\geq\)" instead of "\(\leq\)" (as in [DOR]), since our valuations are written multiplicatively.

Remark 2.16. Let \((D, \Phi, \mathcal{F}^\bullet) \in \text{Fil Isoc}(k)^K_{Q_p}\). Using the notations of Remark 2.5, we find

\[
\mu(D) = p^\frac{1}{d} \left( v_p(\det_{K_0} \Phi) - \sum_i \dim_{K_0} (\mathcal{F}^i/\mathcal{F}^{i+1}) \right).
\]

Hence we see that \(D\) is weakly admissible if and only if it is weakly admissible in the sense of [CF, 3.4].

Proposition 2.17. Let \(D, D' \in \text{Fil Isoc}(k)^K_F\) be semi-stable objects.

(i) If \(\mu(D) < \mu(D')\), then \(\text{Hom}(D, D') = 0\).

(ii) If \(\mu(D) = \mu(D') = \mu\), then for all \(f \in \text{Hom}(D, D')\) we have \(\text{im} f \cong \text{coim} f\) and the objects \(\ker f\), \(\text{coker} f\) and \(\text{im} f\) are semi-stable of slope \(\mu\).

Proof. The proof is the same as in [DOR, Proposition 1.1.20] \(\square\)

Corollary 2.18. Let \(\mu \in \Gamma \otimes \mathbb{Q}\), then the full subcategory of \(\text{Fil Isoc}(k)^K_F\) consisting of semi-stable objects of slope \(\mu\) is an abelian, artinian and noetherian category which is stable under extensions. The simple objects are exactly the stable ones.

Proof. The proof is the same as the proof of [DOR, Corollary 1.2.21]. \(\square\)

The main result of this section is the existence of a Harder-Narasimhan filtration for the objects in \(\text{Fil Isoc}(k)^K_F\). The existence of this filtration will also imply that semi-stability (and hence weak admissibility) is preserved under extension and restriction of scalars.

Proposition 2.19. Let \(D \in \text{Fil Isoc}(k)^K_F\), then there exist unique elements \(\mu_1 < \mu_2 < \cdots < \mu_r \in \Gamma_F \otimes \mathbb{Q}\) and a unique filtration

\[
0 = D_0 \subset D_1 \subset D_2 \subset \cdots \subset D_r = D
\]

of \(D\) in \(\text{Fil Isoc}(k)^K_F\) such that \(D_i/D_{i-1}\) is semi-stable of slope \(\mu_i\).

Proof. The proof is similar to the proof of [DOR, Proposition 1.3.1 (a)].

First we prove the existence of the filtration. The uniqueness will then follow from Lemma 2.21 below.

By the existence of the slope decomposition in Proposition 2.6, the set

\[
\{ \mu(D') \mid D' \subset D \text{ stable under } \Phi \}
\]

is finite. Hence there is a unique minimal element \(\mu_1\) and we claim that there is a maximal subobject \(D_1 \subset D\) of slope \(\mu_1\) which then must be semi-stable. This follows, as the sum of two subobjects of slope \(\mu_1\) has again slope \(\mu_1\), by Lemma 2.13 and the minimality of \(\mu_1\).

Proceeding with \(D/D_1\) the claim follows by induction. \(\square\)
Definition 2.20. Let $D \in \text{Fil} \text{Isoc}(k)^F$ and denote by

$$0 = D_0 \subset D_1 \subset D_2 \subset \cdots \subset D_r = D$$

a filtration as in Proposition 2.19. This filtration is called the Harder-Narasimhan filtration of $D$. For $\mu \in \Gamma_F \otimes \mathbb{Q}$ we define

$$D(\mu) = \begin{cases} 
0 & \text{if } \mu < \mu_1 \\
D_i & \text{if } \mu_i \leq \mu < \mu_{i+1} \\
D & \text{if } \mu \geq \mu_r.
\end{cases}$$

Lemma 2.21. Let $f : D \rightarrow D'$ be a morphism in $\text{Fil} \text{Isoc}(k)^F$ and fix filtrations of $D$ and $D'$ as in Proposition 2.19. Let $\mu \in \Gamma_F \otimes \mathbb{Q}$, then (with the notation of Definition 2.20):

$$f(D(\mu)) \subset D'(\mu).$$

Proof. The proof is the same as in [DOR, Proposition 1.3.1 (b)]. □

Corollary 2.22. Let $F'$ be an extension of $F$ with valuation $v_{F'}$, extending $v_F$ and $D \in \text{Fil} \text{Isoc}(k)^F$. Then $D$ is semi-stable of slope $\mu$, if and only if $D' = D \otimes_F F'$ is semi-stable of slope $\mu$. If in addition $F'$ is finite over $F$, then $D' \in \text{Fil} \text{Isoc}(k)^F$ is semi-stable of slope $\mu$, if and only if $\epsilon_{F'/F}(D') \in \text{Fil} \text{Isoc}(k)^F$ is semi-stable of slope $\mu$.

Proof. First it is clear that semi-stability of $D'$ implies semi-stability of $D$, as every $\Phi$-stable $F$-subspace of $D$ defines a $\Phi$-stable $F'$-subspace of $D'$ which has the same slope.

Now assume that $D$ is semi-stable. We may assume that $F'$ is finitely generated over $F$, as every counterexample for the semi-stability condition is defined over a finitely generated extension. Then $F'$ is an algebraic extension of a purely transcendental extension and we can treat both cases separately.

Assume first that $F'$ is an algebraic extension of $F$. We may replace it by its Galois hull and denote by $G = \text{Gal}(F'/F)$ the Galois group of $F'$ over $F$. Then the action of $G$ preserves the valuation on $F'$. We denote by

$$0 = D'_0 \subset D'_1 \subset D'_2 \subset \cdots \subset D'_r = D'$$

the Harder-Narasimhan filtration of $D'$. The action of $G$ commutes with $\Phi$ and preserves the filtration $F^* \otimes_F F'$ of $D' \otimes_{K, F} K$. It follows that it preserves the slope of a $\Phi$-stable subobject and hence preserves the Harder-Narasimhan filtration. It follows that the filtration descends to $F$ and hence it can only have one step, as $D$ is semi-stable.

Assume now that $F'$ is purely transcendental over $F$. By the above discussion of algebraic extensions we may also assume that $F$ is algebraically closed. Again we write $G = \text{Aut}(F'/F)$ for the group of $F$-automorphisms of $F'$. As above we only need to check that $G$ preserves the slope of a $\Phi$-stable subobject of $D'$. Let $U \subset D'$ be such a $\Phi$-stable subspace. Then $U$ is a direct sum of
indecomposable $\Phi$-modules $U_i$ such that the isomorphism class of $U_i$ is defined over $F$. This can be seen as follows: We decompose $D$ as a product

$$D = V_1 \times \cdots \times V_f$$

of $F$-vector spaces such that $\Phi$ induces a linear map $\Phi_i : V_i \to V_{i+1}$, where $V_{f+1} := V_1$. We can choose bases of these vector spaces such that the matrix of $\Phi$ is represented by the tuple $(\text{id}, \ldots, \text{id}, A)$ for some $A \in \text{GL}_d(F)$ in Jordan canonical form. If we choose a similar canonical form for $\Phi|_U$, then it is clear that every Jordan-block for $\Phi|_U$ is a sub-Jordan-block of $A$.

It follows that $\det_F(\Phi_f|_U) \in F$ and hence the action of $G$ preserves the slope of $U$, which yields the claim by the same descent argument as above.

Now assume that $F'$ is finite over $F$ and $D'$ is a semi-stable object of $\text{Fil}\text{Isoc}(k)_{F'}$. Consider the Harder-Narasimhan filtration of $\epsilon_{F'/F}(D')$. By Lemma 2.21 the filtration steps are stable under the operation of $F'$. Hence the filtration can have only one step.

If conversely $\epsilon_{F'/F}(D')$ is known to be semi-stable, then every $\Phi$-stable $F'$-subspace of $D'$ is a $\Phi$-stable $F$-subspace of $\epsilon_{F'/F}(D')$, and hence semi-stability of $D'$ follows. \qed

3. Families of filtered $\varphi$-modules

It is shown in [He, 4] that the stack of weakly admissible filtered $\varphi$-modules is an open substack of the stack of filtered $\varphi$-modules. We briefly recall this result before we study the weakly admissible locus in the fibers over the adjoint quotient. We write $\text{Rig}_E$ for the category of rigid analytic spaces over a finite extension $E$ of $\mathbb{Q}_p$ (see [BGR] and also [Bo] for an introduction to rigid geometry) and $\text{Ad}^\text{lt}_E$ for the category of adic spaces locally of finite type over $E$, see [Hu2].

3.1. Stacks of filtered $\varphi$-modules. Let $d$ be a positive integer and $\nu$ an algebraic cocharacter

$$\nu : \hat{\mathbb{Q}}_p^\times \to (\text{Res}_{K/Q_p} A_K)(\hat{\mathbb{Q}}_p),$$

where $A \subset \text{GL}_d$ is the diagonal torus. We assume that this cocharacter is dominant with respect to the restriction $B$ of the Borel subgroup of upper triangular matrices in $(\text{GL}_d)_K$. We write $E \subset \hat{\mathbb{Q}}_p$ for the reflex field of $\nu$, i.e. the field of definition of the cocharacter $\nu$, see below for a precise characterization. Let $\Delta$ denote the set of simple roots (defined over $\hat{\mathbb{Q}}_p$) of $\text{Res}_{K/Q_p} \text{GL}_d$ with respect to $B$ and denote by $\Delta_\nu \subset \Delta$ the set of all simple roots $\alpha$ such that $\langle \alpha, \nu \rangle = 0$. Here $\langle - , - \rangle$ is the canonical pairing between characters and cocharacters. We write $P_\nu$ for the parabolic subgroup of $(\text{Res}_{K/Q_p} \text{GL}_d)$ containing $B$ and corresponding to $\Delta_\nu \subset \Delta$. This parabolic subgroup is defined over $E$, and the quotient by this parabolic is a projective $E$-variety

$$\text{Gr}_{K,\nu} = (\text{Res}_{K/Q_p} \text{GL}_d)_E/P_\nu$$
representing the functor
\[ S \mapsto \{ \text{filtrations } F^\bullet \text{ of } \mathcal{O}_S \otimes_{\mathbb{Q}_p} K^d \text{ of type } \nu \} \]
on the category of \( E \)-schemes. Here the filtrations are locally on \( S \) direct summands. Being of type \( \nu \) means the following. Assume that the cocharacter \( \nu : \hat{\mathbb{Q}}_p^\times \rightarrow \prod_{\psi : K \rightarrow \hat{\mathbb{Q}}_p} \text{GL}_d(\hat{\mathbb{Q}}_p) \)
is given by cocharacters
\[ \nu_\psi : \lambda \mapsto \text{diag}(\lambda^{i_1}(\psi)^{m_1(\psi)}, \ldots, \lambda^{i_r}(\psi)^{m_r(\psi)}) \]
for some integers \( i_j(\psi) \in \mathbb{Z} \), with \( i_j(\psi) \neq i_{j'}(\psi) \) for \( j \neq j' \), and multiplicities \( m_j(\psi) > 0 \). Then any point \( F^\bullet \in \text{Gr}_{K,\nu}(\mathbb{Q}_p) \) is a filtration \( \prod_\nu F^\bullet_\psi \) of \( \prod_\nu \mathbb{Q}_p^d \)
such that
\[ \dim_{\mathbb{Q}_p} \text{gr}_i(F^\bullet_\psi) = \begin{cases} 0 & \text{if } i \notin \{ i_1(\psi), \ldots, i_r(\psi) \} \\ m_j(\psi) & \text{if } i = i_j(\psi). \end{cases} \]
In terms of the integers \( i_j(\psi) \) and \( m_j(\psi) \) the fact that \( \nu \) is dominant means that \( i_j(\psi) \geq i_{j+1}(\psi) \) for all \( j \) and \( \psi \). The reflex field \( E \) of the character \( \nu \) is characterized by the requirement
\[ \text{Gal}(\hat{\mathbb{Q}}_p/E) = \{ \sigma \in \text{Gal}(\hat{\mathbb{Q}}_p/\mathbb{Q}_p) \mid i_j(\psi) = i_j(\sigma\psi), \ m_j(\psi) = m_j(\sigma\psi) \}. \]
We denote by \( \text{Gr}^\text{rig}_{K,\nu} \) resp. \( \text{Gr}^\text{ad}_{K,\nu} \) the associated rigid space, resp. the associated adic space (cf. [BGR, 9.3.4] and [Hu2, Remark 4.6 (i)]).

Given \( \nu \) as in (3.1) and denoting as before by \( E \) the reflex field of \( \nu \), we consider the following fpqc-stack \( \mathcal{D}_\nu \) on the category \( \text{Rig}_E \) (resp. on the category \( \text{Ad}^\text{rig}_E \)). For \( X \in \text{Rig}_E \) (resp. \( \text{Ad}^\text{rig}_E \)) the groupoid \( \mathcal{D}_\nu(X) \) consists of triples \((D, \Phi, F^\bullet)\), where \( D \) is a coherent \( \mathcal{O}_X \otimes_{\mathbb{Q}_p} K_0 \)-module which is locally on \( X \) free over \( \mathcal{O}_X \otimes_{\mathbb{Q}_p} K_0 \) and \( \Phi : D \rightarrow D \) is an id \( \otimes \varphi \)-linear automorphism. Finally \( F^\bullet \) is a filtration of \( D = D \otimes_{K_0} K \) of type \( \nu \), i.e. after choosing fpqc-locally on \( X \) a basis of \( D \), the filtration \( F^\bullet \) induces a map to \( \text{Gr}^\text{rig}_{K,\nu} \) (resp. \( \text{Gr}^\text{ad}_{K,\nu} \)), compare also [PR, 5.1].

One easily sees that the stack \( \mathcal{D}_\nu \) is the stack quotient of the rigid space
\[ X_\nu = (\text{Res}_{K_0/\mathbb{Q}_p} \text{GL}_d)_E^{\text{rig}} \times \text{Gr}^\text{rig}_{K,\nu} \]
by the \( \varphi \)-conjugation action of \((\text{Res}_{K_0/\mathbb{Q}_p} \text{GL}_d)_E^{\text{rig}} \), given by
\[ (A, F^\bullet) \cdot g = (g^{-1}A\varphi(g), g^{-1}F^\bullet). \]
Here the canonical map \( X_\nu \rightarrow \mathcal{D}_\nu \) is given by
\[ (A, F^\bullet) \mapsto (\mathcal{O}_{X_\nu} \otimes_{\mathbb{Q}_p} K_0^d, A(\text{id} \otimes \varphi), F^\bullet). \]
3.2. The weakly admissible locus. Fix a cocharacter \( \nu \) with reflex field \( E \) as in the previous section. If \( X \in \text{Ad}^\text{left}_E \) and \( x \in X \), then our definitions imply that, given \((D, \Phi, \mathcal{F}^\bullet) \in \mathcal{D}_\nu(X)\), we have

\[
(D \otimes k(x), \Phi \otimes \text{id}, \mathcal{F}^\bullet \otimes k(x)) \in \text{Fil Isoc}\left(k\right)
\]

One of the main results of [He] is concerned with the structure of the weakly admissible locus in the stacks \( \mathcal{D}_\nu \) defined above.

**Theorem 3.1.** Let \( \nu \) be a cocharacter as in (3.1) and \( X \) be an adic space locally of finite type over the reflex field of \( \nu \). If \((D, \Phi, \mathcal{F}^\bullet) \in \mathcal{D}_\nu(X)\), then the weakly admissible locus

\[X^{wa} = \{x \in X \mid (D \otimes k(x), \Phi \otimes \text{id}, \mathcal{F}^\bullet \otimes k(x)) \text{ is weakly admissible}\}\]

is an open subset. Especially it has a canonical structure of an adic space.

**Proof.** This is [He, Theorem 4.1]. \( \square \)

We can define a substack \( \mathcal{D}_\nu^{wa} \subset \mathcal{D}_\nu \) consisting of the weakly admissible filtered isocrystals. More precisely, for an adic space \( X \) the groupoid \( \mathcal{D}_\nu^{wa}(X) \) consists of those triples \((D, \Phi, \mathcal{F}^\bullet)\) such that \((D \otimes k(x), \Phi \otimes \text{id}, \mathcal{F}^\bullet \otimes k(x))\) is weakly admissible for all \( x \in X \). Thanks to Corollary 2.22 it is clear that this is again an fpqc-stack. The following result is now an obvious consequence of Theorem 3.1.

**Corollary 3.2.** The stack \( \mathcal{D}_\nu^{wa} \) on the category of adic spaces locally of finite type over the reflex field of \( \nu \) is an open substack of \( \mathcal{D}_\nu \).

4. The fibers over the adjoint quotient

We now come to the main results of this paper. We want to link the weakly admissible locus in

\[(\text{Res}_{K_0/\mathbb{Q}_p} \text{GL}_d \times \text{Gr}_{K,\nu})^{\text{ad}}\]

as considered in the previous section to the adjoint quotient of the group \( \text{GL}_d \). This relation was studied by Breuil and Schneider in [BS]. In this section we show that the fibers over the adjoint quotient are (base changes of) analytifications of schemes over \( \mathbb{Q}_p \) and hence the period stacks considered here have a much more algebraic nature than the period spaces considered by Rapoport and Zink in [RZ]. In the next section we determine the image of the weakly admissible locus in the adjoint quotient and identify it with a closed Newton-stratum in the sense of Kottwitz [Ko].

First we need to recall some notations and facts about the adjoint quotient from [Ko]. We write \( \text{GL}_d = \text{GL}(V) \) for the general linear group over \( \mathbb{Q}_p \), where \( V = \mathbb{Q}_p^d \), and \( B \subset \text{GL}_d \) for the Borel subgroup of upper triangular matrices. Further we denote by \( A \subset B \) the diagonal torus and identify \( X_*(A) \) and \( X^*(A) \) with \( \mathbb{Z}^d \) in the usual way, i.e. \((m_1, \ldots, m_d) \in \mathbb{Z}^d \) defines

\[ (t \mapsto \text{diag}(t^{m_1}, \ldots, t^{m_d})) \in X^*(A) \]

resp. \((\text{diag}(t_1, \ldots, t_d) \mapsto t_1^{m_1} \cdots t_d^{m_d}) \in X_*(A)\).
Let $\Delta = \{\alpha_1, \ldots, \alpha_{d-1}\}$ be the simple roots defined by $B$, i.e. $\langle \alpha_i, \nu \rangle = \nu_i - \nu_{i+1}$ for all $\nu \in X_*(A)$. We also choose lifts

$$\omega_i = (1^{(i)}, 0^{(d-i)}) \in \mathbb{Z}^d = X^*(A)$$

denotes the Weyl group of $(\text{GL}_d, A)$. There is a map

$$c : A \rightarrow \mathbb{A}^{d-1} \times \mathbb{G}_m$$

which maps an element of $A$ to the coefficients of its characteristic polynomial. This morphism identifies $A/W$ with $\mathbb{A}^{d-1} \times \mathbb{G}_m$. Now we will define a map

$$G = \text{Res}_{\mathbb{K}_0/\mathbb{Q}_p} (\text{GL}_d)_{\mathbb{K}_0} \rightarrow A/W \quad (4.1)$$

that is invariant under $\varphi$-conjugation on the left side. Recall that we have identifications $\text{GL}_d = \text{GL}(V)$ and $(\text{GL}_d)_{\mathbb{K}_0} = \text{GL}(V_0)$, where $V_0 = V \otimes_{\mathbb{Q}_p} \mathbb{K}_0$. For an $\mathbb{Q}_p$-algebra $R$ and $g \in G(R)$ we have the $R \otimes_{\mathbb{Q}_p} \mathbb{K}_0$-linear automorphism $\Phi_f = (g(\text{id} \otimes \varphi))^f$ of $R \otimes_{\mathbb{Q}_p} V_0$. Its characteristic polynomial is an element of $(R \otimes_{\mathbb{Q}_p} \mathbb{K}_0)[T]$. Now this polynomial is invariant under $\text{id} \otimes \varphi$ and hence it already lies in $R[T]$ which can be seen as follows: First we may assume that $\mathbb{K}_0$ embeds into $R$, as $R = R' \cap (R \otimes_{\mathbb{Q}_p} \mathbb{K}_0) \subseteq R' \otimes_{\mathbb{Q}_p} \mathbb{K}_0$ for any extension $R'$ of $R$. As in the proof of Proposition 2.6, we choose decompositions $R \otimes_{\mathbb{Q}_p} \mathbb{K}_0 = \prod_{i=1}^{f} R_i$ and $D = \prod_{i=1}^{f} V_i$ such that $\Phi_f$ maps $V_i$ to $V_{i+1}$, where $V_{f+1} = V_1$. Then $\Phi_f$ induces automorphisms $\Phi_i$ on $V_i$. It follows that

$$\text{charpoly}(\Phi_f) = (\text{charpoly}(\Phi_1), \ldots, \text{charpoly}(\Phi_f)).$$

However, $\Phi_{i+1} = \Phi_{i/V_i} \circ \Phi_i \circ (\Phi_i|_{V_i})^{-1}$ and hence $\text{charpoly}(\Phi_1) = \text{charpoly}(\Phi_{i+1})$. We define the morphism in (4.1) by mapping $g \in G(R)$ to the coefficients of this polynomial. It is easy to check that this morphism is invariant in $\varphi$-conjugation on $G$ and hence we get morphisms

$$G \times \text{Gr}_{K, \nu} \overset{\alpha}{\longrightarrow} (A/W)_E \quad (4.2)$$

where $D_\nu$ is the stack-quotient

$$D_\nu = (G_E \times \text{Gr}_{K, \nu})/G$$

on the category of $E$-schemes, where the action of $G$ on $G_E \times \text{Gr}_{K, \nu}$ is the same as in (3.4). Here $\nu$ is a cocharacter as in (3.1) and $E$ is the reflex field of $\nu$. We also write $\alpha$ and $\tilde{\alpha}$ for the analytification of these morphisms.

**Theorem 4.1.** Let $x \in (A/W)_E^{\text{ad}}$ and $\nu$ be a cocharacter as in (3.1). Then there exists a quasi-projective $k(x)$-scheme $X$ which is an open subscheme of $\tilde{\alpha}^{-1}(x)$ such that the weakly admissible locus in the fiber over $x$ is given by

$$\tilde{\alpha}^{-1}(x)^{\text{wa}} = X^{\text{ad}}.$$
Further we denote by \( f \) (recall projection for this action and consider the subscheme \( Z \)). We also write linear automorphism of \( \mathcal{O} \).

Let \( x = (c_1, \ldots, c_d) \in k(x)^{d-1} \times k(x)^x \) and let \( v_x \) denote the (multiplicative) valuation on \( k(x) \). First note that

\[
c_d = \det_{k(x) \otimes \mathbb{Q}_p} K_0(\Phi^f) = \det_{k(x)} (\Phi^f)^{1/f}
\]

and hence \( \tilde{\alpha}^{-1}(x) = 0 \) unless

\[
v_x(c_d)^{1/f} = v_x(p)^{\sum_{j \geq 0} \frac{1}{f_j} \dim_k(x^i) g_{ij}^f},
\]

where \( \mathcal{F}^* \) is the universal filtration on \( \text{Gr}_{K_0} \). In the following we will assume that this condition is satisfied. For \( i \in \{0, \ldots, d\} \), consider the following functor on the category of \( \mathbb{Q}_p \)-schemes,

\[
S \mapsto \begin{cases} \mathcal{E} \subset \mathcal{O}_S \otimes \mathbb{Q}_p V_0 \text{ locally free } \mathcal{O}_S \otimes \mathbb{Q}_p \text{-submodule} \\ i \text{ that is locally on } S \text{ a direct summand} \end{cases}.
\]

Using the theory of Quot-schemes (see [FGA, Thm 3.1] for example) this functor is easily seen to be representable by a projective \( \mathbb{Q}_p \)-scheme \( \text{Gr}_{K_0,i} \).

We let \( G = \text{Res}_{K_0/K} \mathbb{Q}_p \text{GL}_d \) act on \( \text{Gr}_{K_0,i} \) in the following way: for a \( \mathbb{Q}_p \)-scheme \( S \), let \( A \in G(S) \) and \( \mathcal{E} \in \text{Gr}_{K_0,i}(S) \). We get a linear endomorphism \( A \) of \( \mathcal{O}_S \otimes \mathbb{Q}_p V_0 \) and define the action of \( A \) on \( \mathcal{E} \) by

\[
A \cdot \mathcal{E} = A((id \otimes \varphi)(\mathcal{E})�).
\]

Write

\[
a : G \times \text{Gr}_{K_0,i} \longrightarrow \text{Gr}_{K_0,i}
\]

for this action and consider the subscheme \( Z_i \subset G \times \text{Gr}_{K_0,i} \) defined by the following fiber product:

\[
\begin{array}{ccc}
Z_i & \longrightarrow & G \times \text{Gr}_{K_0,i} \\
\downarrow & & \downarrow a \times \text{id} \\
\text{Gr}_{K_0,i} & \longrightarrow & \text{Gr}_{K_0,i} \times \text{Gr}_{K_0,i}.
\end{array}
\]

An \( S \)-valued point \( x \) of the scheme \( Z_i \) is a pair \((g_x, U_x)\), where \( g_x \in G(S) \) is a linear automorphism of \( \mathcal{O}_S \otimes \mathbb{Q}_p V_0 \) and \( U_x \) is an \( \mathcal{O}_S \otimes \mathbb{Q}_p K_0 \)-submodule of rank \( i \) stable under \( \Phi_x = g_x(id \otimes \varphi) \). The scheme \( Z_i \) is projective over \( G \) via the first projection

\[
\text{pr}_1 : Z_i \longrightarrow G.
\]

Further we denote by \( f_i \in \Gamma(Z_i, \mathcal{O}_{Z_i}) \) the global section defined by

\[
f_i(g_x, U_x) = \det((g_x(id \otimes \varphi))^{f}|_{U_x})
\]

(recall \( f = [K_0 : \mathbb{Q}_p] \)), where the determinant is the determinant as \( \mathcal{O}_{Z_i} \)-modules. We also write \( f_i \) for the global section on the associated adic space \( Z_i^{ad} \).

We write \( \mathcal{E} \) for the pullback of the universal bundle on \( Z_i \) to \( Z_i \times \text{Gr}_{K_0} \) and \( \mathcal{F}^* \) for the pullback of the universal filtration on \( \text{Gr}_{K_0} \). Then the fiber product

\[
\mathcal{G}^* = (\mathcal{E} \otimes K_0) \cap \mathcal{F}^*
\]

Documenta Mathematica 16 (2011) 969–991

\[\text{Proof.}\] The proof will be similar to the proof of [He, Theorem 4.1].
is a filtration of $E \otimes_{K_0} K$ by coherent sheaves. By the semi-continuity theorem the function

$$h_i : x \mapsto \sum_{j \in \mathbb{Z}} \frac{1}{e_j} \dim_{k(x)} \text{gr}_j G^\bullet$$

is upper semi-continuous on $Z_i \times \text{Gr}_{K,\nu}$ and hence so is

$$h^\text{ad}_i : x \mapsto \sum_{j \in \mathbb{Z}} \frac{1}{e_j} \dim_{k(x)} \text{gr}_j (G^\bullet)^{\text{ad}}.$$

For $m \in \mathbb{Z}$ we write $Y_{i,m} \subset Z_i \times \text{Gr}_{K,\nu}$ (resp. $Y_{i,m}^{\text{ad}} \subset Z_i^{\text{ad}} \times \text{Gr}_{K,\nu}^{\text{ad}}$) for the closed subscheme (resp. the closed adic subspace)

$$Y_{i,m} = \{ y \in Z_i \times \text{Gr}_{K,\nu} \mid h_i(y) \geq m \},$$

$$Y_{i,m}^{\text{ad}} = \{ y \in Z_i^{\text{ad}} \times \text{Gr}_{K,\nu}^{\text{ad}} \mid h^\text{ad}_i(y) \geq m \}.$$

Then the definitions imply that

$$\text{pr}_{i,m} : Y_{i,m} \rightarrow G \times \text{Gr}_{K,\nu}$$

$$\text{pr}_{i,m} : Y_{i,m}^{\text{ad}} \rightarrow (G \times \text{Gr}_{K,\nu})^{\text{ad}}$$

are proper morphism. Now

$$S_{i,m} = \{ y = (g_y, U_y, F^\bullet) \in Y_{i,m} \times (G \times \text{Gr}_{K,\nu})^{\text{ad}} \mid v_y(f_i(g_y, U_y)) > v_y(p)^{f^m} \}$$

is a union of connected components of

$$Y_{i,m} \times (G \times \text{Gr}_{K,\nu})^{\text{ad}},$$

which can be seen as follows: Let $\lambda_1, \ldots, \lambda_d$ denote the zeros of the polynomial

$$X^d + c_1 X^{d-1} + \cdots + c_{d-1} X + c_d.$$

Then every possible value of the $f_i$ is a product of some of the $\lambda_i$ and hence $f_i$ can take only finitely many values.

We conclude that the subset $\bigcup_{i,m} \text{pr}_{i,m}(S_{i,m})$ is closed and claim that

$$\tilde{\alpha}^{-1}(x)^{\text{wa}} = (\tilde{\alpha}^{-1}(x) \bigcup_{i,m} \text{pr}_{i,m}(S_{i,m}))^{\text{ad}}.$$

Indeed, let $z = (g_z, F^\bullet) \in \tilde{\alpha}^{-1}(x) \subset G^{\text{ad}} \times \text{Gr}_{K,\nu}^{\text{ad}}$. Then the object

$$(k(z) \otimes V_0, g_z(\text{id} \otimes \varphi), F^\bullet)$$

is not weakly admissible if and only if there exists a $g_z(\text{id} \otimes \varphi)$-stable subspace $U_z \subset k(z) \otimes V_0$ of some rank, violating the weak admissibility condition. This means $z \in \bigcup_{i,m} \text{pr}_{i,m}(S_{i,m})^{\text{ad}}$. Here we implicitly use that fact that weak admissibility is stable under extension of scalars (see Corollary 2.22).

Remark 4.2. In view of the period domains considered in [RZ] it can be surprising that this weakly admissible locus is indeed the adification of a scheme, not just an analytic space. The main reason is the following: In [RZ] the isocrystal is fixed and the counter examples one has to exclude for the weak admissibility condition are parametrized by the $\mathbb{Q}_p$-valued points of an algebraic variety. In
our setting the isocrystal is not fixed and the Frobenius $\Phi$ may vary. Hence the set of counter examples is the algebraic variety itself rather than its $\mathbb{Q}_p$-valued points.

**Example 4.3.** This example illustrates the difference with period spaces in the sense of Rapoport-Zink. Let $K = \mathbb{Q}_p$ and $d = 2$. We consider the Frobenius $\Phi = \text{diag}(p,p)$ and fix the filtration $\mathcal{F}^*$ such that

$$\dim \mathcal{F}^i = \begin{cases} 2 & i \leq 0 \\ 1 & i = 1, 2 \\ 0 & i \geq 3. \end{cases}$$

In this case the flag variety $\text{Gr}_{K,\nu}$ is the projective line $\mathbb{P}^1$ and the period space of [RZ] (or rather [DOR]) is the Drinfeld upper halfplane $\mathbb{P}^1 \setminus \mathbb{P}^1(\mathbb{Q}_p)$ which is not a scheme. On the other hand the weakly admissible set in the sense discussed above is obviously empty.

**Corollary 4.4.** Let $x \in (A/W)^{\text{ad}}_E$ and consider the $2$-fiber product

$$\xymatrix{ \alpha^{-1}(x)^{\text{wa}} \ar[r] \ar[d] & \mathcal{D}^{\text{wa}}_E \ar[d] \\ x \ar[r] & (A/W)^{\text{ad}}_E. }$$

Then there exists an Artin stack in schemes $\mathcal{X}$ over the field $k(x)$ such that

$$\alpha^{-1}(x)^{\text{wa}} \cong \mathcal{X}^{\text{ad}}_E.$$

**Proof.** This is an immediate consequence of Theorem 4.1.

We end the discussion of the fibers over the adjoint quotient by discussing three examples.

**Example 4.5.** Let $K = \mathbb{Q}_p$ and $d = 3$. We take $\Phi = \text{diag}(1,p,p^2)$ and fix the type of the filtration $\mathcal{F}^*$ such that

$$\dim \mathcal{F}^i = \begin{cases} 3 & i \leq 0 \\ 2 & i = 1 \\ 1 & i = 2 \\ 0 & i \geq 3. \end{cases}$$

We write $G = \text{GL}_3$ and $B \subset G$ for the Borel subgroup of upper triangular matrices. Further $X = G/B$ is the full flag variety, and we are interested in the weakly admissible locus in $X$. One easily checks that

$$X^{\text{wa}} = \{ \mathcal{F}^* \in X \mid \mathcal{F}^1 \cap V_1 = 0, \text{ and } \mathcal{F}^2 \not\subset V_{12} \},$$

where $0 \subset V_1 \subset V_{12} \subset \mathbb{Q}_p^3$ is the standard flag fixed by the Borel $B$. The subset $X^{\text{wa}}$ is obviously stable under $B$ and, in fact,

$$X^{\text{wa}} = Bw_0B/B,$$
where \( w_0 \) is the longest Weyl group element. If \( x \) denotes the image of \( \Phi \) in the adjoint quotient, then \( \Phi \) is a representative of the unique \( \varphi \)-conjugacy class in \( G \) mapping to \( x \) and we further have \( \alpha^{-1}(x) = A \backslash G/B \) and

\[
\alpha^{-1}(x)_{wa} = A \backslash B w_0 B / B.
\]

**Example 4.6.** We use the same notations as in the example above, but this time \( \Phi = \text{diag}(1, 1, p^3) \). Then

\[
X_{wa} = \{ \mathcal{F}^\bullet \in X \mid \mathcal{F}^1 \cap V_{12} = 0 \}.
\]

As \( \dim V_{12} = \dim \mathcal{F}^1 = 2 \) it follows that \( X_{wa} = \emptyset \). In this case there is a second \( \varphi \)-conjugacy class in \( G \) mapping to the same point in the adjoint quotient as \( \Phi \). A representative of this second \( \varphi \)-conjugacy class is given by

\[
\Phi' = \begin{pmatrix}
1 & 1 & 0 \\
0 & 1 & 0 \\
0 & 0 & p^3
\end{pmatrix}.
\]

**Example 4.7.** In this example let \( d = 2 \) and \( \Phi = \text{diag}(1, p) \). Let \( K \) be a ramified extension of \( \mathbb{Q}_p \) of degree \( e \) and consider flags of the type \((1, \ldots, 1)\), i.e. the cocharacter is defined over \( \mathbb{Q}_p \), the only non-trivial filtration step is \( \mathcal{F}^1 = (\mathcal{F}^1_i)_{i=1,\ldots,e} \) and the base change of the flag variety \( X = \text{Gr}_{K,\nu} \) to \( K \) is

\[
X_K = \mathbb{P}_{K}^1 \times \cdots \times \mathbb{P}_{K}^1.
\]

The weakly admissible locus in \( X_K \) is given by

\[
X_{wa}^K = \{ \mathcal{F}^\bullet = (\mathcal{F}^\bullet_i)_{i=1,\ldots,e} \mid \mathcal{F}^i_i \neq \infty \text{ for all } i \in \{1,\ldots,e\} \}.
\]

Let \( G = \text{Res}_{K/\mathbb{Q}_p} \text{GL}_2 \) and \( B \subset G \) the Weil-restriction of the Borel subgroup of upper triangular matrices. Again we write \( w_0 \) for the longest Weyl group element of \( G \). Then

\[
X_{wa} = Bu_0 B / B \subset X = G / B.
\]

If again \( x \) denotes the image of \( \Phi \) in the adjoint quotient, then \( \alpha^{-1}(x) = A \backslash X \) and

\[
\alpha^{-1}(x)_{wa} = A \backslash B u_0 B / B.
\]

5. **Newton strata and weak admissibility**

The proof of Theorem 4.1 suggests that the weakly admissible locus in the fibers over a point in \( A/W \) does only depend on the valuation of the zeros of the characteristic polynomial associated to the points of the adjoint quotient. Hence we want to extend the result that the fibers over the adjoint quotient are nice spaces to the pre-image of the Newton strata in the adjoint quotient. Here we work in the category of analytic spaces in the sense of Berkovich (see [Be]), as it is not obvious how to generalize the notion of Newton strata (as defined in [Kol]) to adic spaces. Though the weakly admissible locus is not a Berkovich space in general [He, Example 4.4], we show that it becomes a Berkovich space, if we restrict ourselves to the pre-images of the Newton strata. Further we want to identify the image of the weakly admissible locus in the adjoint quotient with
a (closed) Newton-stratum. As usual we will write \( \mathcal{H}(x) \) for the residue field at a point \( x \) in an analytic space and \( X^{an} \) for the analytic space associated to a scheme \( X \).

5.1. NEWTON STRATA. We first need to recall more notations from [Ko]. We write \( a = X_*(A) \otimes \mathbb{Z} \mathbb{R} \), and \( a_{\text{dom}} \subset a \) for the subset of dominant elements, i.e the elements \( \mu \in a \) such that \( \langle \alpha_i, \mu \rangle \geq 0 \) for all \( i \in \{1, \ldots, d-1 \} \). For \( c = (c_1, \ldots, c_d) \in (A/W)^{an} \) we write
\[
\langle \lambda, \nu \rangle = -v_a(\lambda(a))
\]
for all \( \lambda \in X^*(A) \), where we write \( v_a \) for the (additive) valuation on \( \mathcal{H}(a) \). By [Ko, Proposition 1.4.1] there is a continuous map \( r : a \to a_{\text{dom}} \) mapping \( x \in a \) to the dominant element with the smallest distance to \( x \), and this map extends in a continuous way to \( \tilde{\mathbb{R}}^{d-1} \times \mathbb{R} \). Here \( a \subset \tilde{\mathbb{R}}^{d-1} \times \mathbb{R} \) via the chosen identification \( X_*(A) = \mathbb{Z}^d \). Then we find that \( r(d_{e(a)}) \) is the unique dominant element in the \( W \)-orbit of \( \nu_a \). This follows from [Ko, Theorem 1.5.1] for all \( a \in A(\mathbb{Q}_p) \) and, for an arbitrary point, from the fact that \( A(\mathbb{Q}_p) \) is dense in \( A^{an} \) and the continuity of the construction. Note that \( r(d_{e}) = (-v_a(\lambda_1) \geq \cdots \geq -v_a(\lambda_d)) \) if the \( \lambda_i \) are the roots of the characteristic polynomial associated to \( c \) and Kottwitz shows that this does only depend on \( d_e \) and not on \( c \) itself.

**Definition 5.1.** For \( \mu \in a_{\text{dom}} \) we define
\[
(A/W)_\mu = \{ c \in A/W \mid r(d_c) = \mu \}
\]
\[
(A/W)_{\leq \mu} = \{ c \in A/W \mid r(d_c) \leq \mu \}.
\]
Here "\( \leq \)" is the usual dominance order on dominant coweights. We will call the first of these subsets the **Newton stratum** defined by \( \mu \) and the second the **closed Newton stratum** defined by \( \mu \).

We need another description of these sets to identify them as analytic subspaces of the adjoint quotient.

**Proposition 5.2.** Let \( \mu \in a_{\text{dom}} \) and \( I_\mu = \{ i \in \{1, \ldots, d-1 \} \mid \langle \alpha_i, \mu \rangle = 0 \} \).

Then
\[
(A/W)_\mu = \left\{ c = (c_1, \ldots, c_d) \in (\mathbb{A}^{d-1} \times G_m)^{an} \mid \begin{array}{l}
v_c(c_i) \geq -\langle \omega_i, \mu \rangle, i \in I_\mu \\
v_c(c_i) = -\langle \omega_i, \mu \rangle, i \notin I_\mu
\end{array} \right\}
\]
\[
(A/W)_{\leq \mu} = \left\{ c = (c_1, \ldots, c_d) \in (\mathbb{A}^{d-1} \times G_m)^{an} \mid \begin{array}{l}
v_c(c_i) \geq -\langle \omega_i, \mu \rangle, i \neq d \\
v_c(c_d) = -\langle \omega_d, \mu \rangle
\end{array} \right\}
\]

**Proof.** For all points in \((A/W)^{an}(\mathbb{Q}_p)\) this follows from [Ko, Theorem 1.5.2]. Again the proposition follows from continuity, and the fact that the points in \((A/W)^{an}(\mathbb{Q}_p)\) are dense in \((A/W)^{an}\). \(\square\)
The category of (strict) analytic spaces is a full subcategory of the category of adic spaces locally of finite type, see [Hu3, 8.3] and the references cited there. Hence we can restrict the stacks $\mathcal{D}_\nu$ and $\mathcal{D}_\nu^{wa}$ to the category of analytic spaces. We write again $\mathcal{D}_\nu$ and $\mathcal{D}_\nu^{wa}$ for these restrictions. Further we write $\tilde{\alpha}^{an}$ (resp. $\alpha^{an}$) for the analytifications of the morphisms defined in (4.2).

**Theorem 5.3.** Let $\nu$ be a cocharacter as in (3.1) and $\mu \in \mathfrak{a}_{\text{dom}}$. Then the weakly admissible locus in the inverse image $((\tilde{\alpha}^{an})^{-1}(A/W)^{an}_\mu)$ is an analytic space.

**Proof.** The proof is almost identical with the proof of Theorem 4.1. If we replace $Z_i$ by

$$Z_i = (Z_i)^{an} \times (G \times_{\text{Gr}_{K,\nu}})^{an} (\tilde{\alpha}^{an})^{-1}(A/W)^{an}_\mu$$

and $Y_{i,m}$ by

$$Y_{i,m} = Y_{i,m} \times (G \times_{\text{Gr}_{K,\nu}})^{an} (\tilde{\alpha}^{an})^{-1}(A/W)^{an}_\mu$$

Here the functions $f_i$ are not locally constant on $Z_i$ and $\tilde{Y}_{i,m}$ but their valuations (or absolute values) are. As in the proof of Theorem 4.1\(^1\) it follows that

$$\tilde{S}_{i,m} = \{ y = (g_y, U_y, F^*_y) \in Y_{i,m} \mid f_i(g_y, U_y)|_y > |p|^j \}$$

is a union of connected components of $\tilde{Y}_{i,m}$ and we use the properness of the projection $\text{pr}_{i,m} : \tilde{Y}_{i,m} \to (\tilde{\alpha}^{an})^{-1}(A/W)^{an}_\mu)$ to conclude that

$$(\tilde{\alpha}^{an})^{-1}(A/W)^{an}_\mu \cong (\tilde{\alpha}^{an})^{-1}(A/W)^{an}_\mu) \setminus \bigcup_{\text{pr}_{i,m}(\tilde{S}_{i,m})}$$

is an open subspace. \(\square\)

### 5.2. The image of the weakly admissible locus

In this section we determine the image of the weakly admissible locus under the map defined in (4.1). In the case of a regular cocharacter $\nu$ it was shown by Breuil and Schneider that the set of points $a \in A$ such that there exists a weakly admissible filtered $\varphi$-module $(D, \Phi, F^*)$ with $(\Phi^J)^{an} = a$ is an affinoid domain, see [BS, Proposition 3.2]. Here we extend this result to the general case and give a description of this image purely in terms of the adjoint quotient $A/W$. The difference with the description in [BS] is that we do not need to fix an order of the eigenvalues corresponding to the order of their valuations. We fix a coweight $\nu$ as in (3.1). This coweight determines the jumps of the filtration $F^*$ on $\text{Gr}_{K,\nu}$. After passing to $\bar{Q}_p$ the filtration is given by

$$F^* = \prod_{\psi} F^*_\psi$$

where the product runs over all embeddings $\psi : K \hookrightarrow \bar{Q}_p$.

We write $\{x_{\psi,1} > x_{\psi,2} > \cdots > x_{\psi,r}\}$ for the jumps of the filtration $F^*_\psi$, i.e.

$$\text{gr}_i F^*_\psi \neq 0 \iff i \in \{x_{\psi,1}, \ldots, x_{\psi,r}\}.$$  

\(^1\)Note that in this case this does not show that the weakly admissible locus is a scheme, as our base is not a point but the Newton-stratum $(A/W)^{an}_\mu$, which is not a scheme.
Further denote by \( n_{\psi,i} \) the rank of \( F^\psi_{\psi,i} \) and write
\[
m_{\psi,j}(i) = \max(0, n_{\psi,j} + i - d).
\]
This is the minimal dimension of the intersection of \( F^\psi_{\psi,i} \) with a subspace of dimension \( i \). For \( i \in \{0, \ldots, d\} \) define
\[
(5.2) \quad l_i = \sum_{\psi} \frac{1}{\psi} \left( \sum_{j=1}^{r-1} (x_{\psi,j} - x_{\psi,j+1})m_{\psi,j}(i) + x_{\psi,r}m_{\psi,r}(i) \right).
\]

**Definition 5.4.** For a cocharacter \( \nu \) and \( i \in \{1, \ldots, d\} \) define \( l_i \) as in (5.2).
Define a rational dominant coweight \( \mu(\nu) \in \mathfrak{d}_{\text{dom}} \) by requiring that
\[
\frac{1}{\nu}(\omega_i, \mu(\nu)) = -l_i \quad \text{for all } i \in \{1, \ldots, d\}.
\]
The following result generalizes [BS, Proposition 3.2].

**Theorem 5.5.** Let \( \nu \) be a cocharacter as in (3.1) and define \( \mu(\nu) \) as in Definition 5.4. Let \( x \in (A/W)^{\text{an}} \), then \( \tilde{\psi}^{-1}(x)^{\text{wa}} \neq \emptyset \) if and only if \( x \in (A/W)^{\text{an}}_{\leq \mu(\nu)} \).

**Proof.** Let \( e = (c_1, \ldots, c_d) \in (A/W)^{\text{an}}_{\leq \mu(\nu)} \) and denote by \( \lambda_1, \ldots, \lambda_t \) the roots of
\[
X^d + c_1X^{d-1} + \cdots + c_{d-1}X + c_d
\]
with multiplicities \( m_i \) in some finite extension \( L \) of \( \mathcal{H}(e) \) containing \( K_0 \). Let \( D = L \otimes_{\mathbb{Q}_p} V_0 \cong \prod_{i=1}^d L^d \) and
\[
g = (\text{id}, \ldots, \text{id}, A) \in \prod_{i=1}^t \mathrm{GL}_d(L) \cong G(L),
\]
where \( A \) is a matrix consisting of \( t \) Jordan blocks of size \( m_i \) with diagonal entries \( \lambda_i \). Now the pair \( (D, \Phi) = (D, g(\text{id} \otimes \varphi)) \in \text{Isoc}(k)_L \) has the property that there are only finitely many \( \Phi \)-stable subobjects \( D' \subset D \). If \( D' \subset D \) is a rank \( i \) subobject then
\[
t_N(D') := \frac{1}{t} v_p(\det \Phi|_{D'}) = \frac{1}{t} \sum_{j=1}^t m_j' v_p(\lambda_j)
\]
for some multiplicities \( m_j' \), where we write \( v_p \) for the additive valuation. Write \( a = (\lambda_1^{(m_1)}, \ldots, \lambda_t^{(m_t)}) \in A^{\text{an}} \), then \( c(a) = c \) and \( r(d_{c(a)}) \leq \mu(\nu) \) by assumption. It follows that
\[
t_N(D') = \frac{1}{t} v_p(w\omega_i(a)) = \frac{1}{t} v_p(w\omega_i, \nu_a)
\]
\[
= -\frac{1}{t} \langle w'\omega_i, d_{c(a)} \rangle \geq -\frac{1}{t} \langle \omega_i, r(d_{c(a)}) \rangle
\]
\[
\geq -\frac{1}{t} \langle \omega_i, \mu(\nu) \rangle = l_i.
\]
for some \( w, w' \in W \). Now for all \( \Phi \)-stable \( D' \subset D \) consider the open subset
\[
U_{D'} \subset \text{Gr}_{K,\nu} \otimes_{\mathbb{Q}_p} K_0
\]
of all filtrations $\mathcal{F}^*$ such that $\dim(F^*_\psi \cap D'_K) = \max(0, n_{\psi, j} + i - d)$ for all embeddings $\psi$. This is open as the right hand side is the minimal possible dimension of such an intersection. Since $Gr_{K, \nu}$ is geometrically irreducible we find that the intersection $\bigcap_{D' \subset D} U_{D'}$ is non-empty and hence there exists an $F$-valued point $F^*$ in this intersection, where $F$ is some extension of $L$. Now we have $(D \otimes_{K_0} F, \Phi \otimes \text{id}, F^*) \in \text{Fil Isoc}(k)_{\Phi}$ and this object is weakly admissible since for all $\Phi$-stable $D' \subset D$ we have

$$\deg(D') = l_i - t_N(D') \leq 0$$

where $i$ is the rank of the subobject $D'$ (and here we write the degree additively). Further, by the definition of $g$, we find that $g$ maps to $c$ under the map $\tilde{\alpha}$.

Conversely assume that $c \in (A/W)^{an}$ such that $\emptyset \neq \tilde{\alpha}^{-1}(c)$. Let $(D, \Phi, F^*)$ be an $F$-valued point of this fiber for some field $F$ containing $K_0$. Then $D$ decomposes into $D_1 \times \cdots \times D_I$ and we denote by $\mu_1, \ldots, \mu_t$ the distinct eigenvalues of $\Phi|_{D_1}$ and by $d_i$ their multiplicities (as zeros of the characteristic polynomial).

We write $(\lambda_1, \ldots, \lambda_d) = (\mu_1^{(d_1)}, \ldots, \mu_t^{(d_t)})$. Then $c = c(\lambda_1, \ldots, \lambda_d)$ and we claim that

$$\frac{1}{\ell} \sum_{j \in I} v_p(\lambda_j) \geq l_i$$

for all $I \subset \{1, \ldots, d\}$ with $\sharp I = i$. This claim clearly implies $c \in (A/W)^{an}_{\leq \mu(c)}$.

Let $I \subset \{1, \ldots, d\}$ and write $(\lambda_j)_{j \in I} = (\lambda_1^{(m_1)}, \ldots, \lambda_t^{(m_t)})$, where we assume that the $\lambda_j$ are pairwise distinct. Then $\sum_{j=1}^I m_j = i = \sharp I$. Using the Jordan canonical form on easily sees that there exists a subobject $D' \subset D$ such that

$$(\Phi|_{D' \cap D_i})^{an} = \text{diag}(\lambda_1^{(m_1)}, \ldots, \lambda_t^{(m_t)})$$

and hence

$$\sum_{j \geq 1} m_j \frac{1}{\ell} v_p(\lambda_j) \geq \sum_{j \in I} \frac{1}{\ell} \dim gr_{\lambda_j} D'_K \geq l_i,$$

which yields the claim. □

We end by giving two examples of closed Newton strata in the adjoint quotient.

**Example 5.6.** Let $K = \mathbb{Q}_p$ and $d = 3$. We fix the cocharacter $\nu$ as in Example 4.5 and Example 4.6, i.e.

$$\dim \mathcal{F}^i = \begin{cases} 3 & i \leq 0 \\ 2 & i = 1 \\ 1 & i = 2 \\ 0 & i \geq 3. \end{cases}$$

One easily checks that $l_1 = 0$, $l_2 = 1$ and $l_3 = 3$, i.e.

$$\mu(\nu) : t \mapsto \text{diag}(1, t^{-1}, t^{-2}).$$
The image of the weakly admissible locus in the adjoint quotient is given by

\[(A/W)^{an}_{\leq \mu(\nu)} = \left\{ c = (c_1, c_2, c_3) \in A^2 \times \mathbb{G}_m \mid \begin{array}{l}
v_c(c_1) \geq 0, \\
v_c(c_2) \geq 1, \\
v_c(c_3) = 3.
\end{array} \right\}\]

If \(a = (a_1, a_2, a_3) \in A\) with \(v_a(a_1) \leq v_a(a_2) \leq v_a(a_3)\), then [BS, Proposition 3.2] says that there exists a weakly admissible filtered \(\phi\)-module \((D, \Phi, F^*\)) with filtration of type \(\nu\) such that \(\Phi^{ss} = a\) if and only if

\[
\begin{cases}
0 \leq v_a(a_1) \\
1 \leq v_a(a_1) + v_a(a_2) \\
3 = v_a(a_1) + v_a(a_2) + v_a(a_3).
\end{cases}
\]

This is clearly equivalent to our condition in the adjoint quotient. This result also explains Example 4.5 and Example 4.6.

**Example 5.7.** Again we let \(K = \mathbb{Q}_p\) and \(d = 3\). Fix a cocharacter \(\nu\) such that

\[
\dim F^i = \begin{cases}
3 & i \leq 0 \\
2 & i = 1 \\
0 & i \geq 2.
\end{cases}
\]

One easily checks that \(l_1 = 0\), \(l_2 = 1\) and \(l_3 = 2\), and the image of the weakly admissible locus is

\[(A/W)^{an}_{\leq \mu(\nu)} = \left\{ c = (c_1, c_2, c_3) \in A^2 \times \mathbb{G}_m \mid \begin{array}{l}
v_c(c_1) \geq 0, \\
v_c(c_2) \geq 1, \\
v_c(c_3) = 2.
\end{array} \right\}\]

**References**

[Be] V. Berkovich, *Spectral theory and analytic geometry over non-archimedian fields*, Math. surveys and monographs 33, American Math. Soc. 1990.

[BGR] S. Bosch, U. Güntzer, R. Remmert, *Non-archimedian analysis*, Springer-Verlag, Berlin 1984.

[Bo] S. Bosch, *Lectures on formal and rigid geometry*, Preprint 378, University of Münster, SFB 478-Preprint Series, Münster 2005. Available at http://www.math.uni-muenster.de/sfb/about/publ/heft378.pdf.

[BS] C. Breuil, P. Schneider, *First steps towards p-adic Langlands functoriality*, J. reine angew. Math. 610 (2007), 149 - 180.

[CF] P. Colmez, J.-M. Fontaine, *Constructions des représentations p-adiques semi-stables*, Invent. Math. 140 (2000), 1-43.

[DOR] J-F. Dat, S. Orlik, M. Rapoport, *Period domains over finite and p-adic fields*, Cambridge Tracts in Math. 183, Cambridge University Press 2010.

[FGA] A. Grothendieck, *Fondements de la Géométrie Algébrique IV. Les schémas de Hilbert*, Séminaire Bourbaki, t. 13, 1960/61, no. 221.

[He] E. Hellmann, *On arithmetic families of filtered \(\phi\)-modules and crystalline representations*, preprint 2011, arXiv:1010.4577v2.
Filterd $\varphi$-modules and the adjoint quotient

[Hu1] R. Huber, Continuous valuations, Math. Z. 212 (1993), 445-447.
[Hu2] R. Huber, A generalization of formal schemes and rigid analytic varieties, Math. Z. 217 (1994), 513-551.
[Hu3] R. Huber, Étale cohomology of rigid analytic varieties and adic spaces, Aspects of Math., E30, Friedr. Vieweg & Sohn, Braunschweig, 1996.
[Ko] R. E. Kottwitz, Dimensions of Newton strata in the adjoint quotient of reductive groups, Pure Appl. Math. Q. 2 (2006), 817-836.
[PR] G. Pappas, M. Rapoport, $\Phi$-modules and coefficient spaces, Moscow Math. J. 9 no. 3 (2009), 625-664.
[RZ] M. Rapoport, T. Zink, Period spaces for p-divisible groups, Ann. of Math. Studies, vol 141, Princeton University Press, Princeton, NJ, 1996.
[Zi] T. Zink, Cartiertheorie kommutativer formaler Gruppen, Teubner Texte zur Math. 68, Leipzig, 1984.

Eugen Hellmann
Mathematisches Institut
der Universität Bonn
Endenicher Allee 60
53115 Bonn
Germany
hellmann@math.uni-bonn.de
