$p$-ADIC PROPERTIES OF PICARD MODULAR SCHEMES AND MODULAR FORMS

EHUD DE SHALIT AND EYAL Z. GOREN

Table of contents

0. Introduction
1. Background
1.1 The unitary group and its symmetric space
1.2 Picard modular surfaces over $\mathbb{C}$
1.3 The Picard moduli scheme
1.4 Smooth compactifications
1.5 The universal semi-abelian scheme $\mathcal{A}$
1.6 Degeneration of $\mathcal{A}$ along a geodesic connecting to a cusp

2. The basic automorphic vector bundles
2.1 The vector bundles $\mathcal{P}$ and $\mathcal{L}$
2.2 The Kodaira-Spencer isomorphism
2.3 Extensions to the boundary of $S$
2.4 Fourier-Jacobi expansions
2.5 The Gauss-Manin connection in a neighborhood of a cusp
2.6 Fields of rationality

3. Picard moduli schemes modulo an inert prime $p$
3.1 The stratification
3.2 New relations between automorphic vector bundles in characteristic $p$
3.3 The Hasse invariant $h_{\Sigma}$
3.4 A secondary Hasse invariant on the supersingular locus
3.5 The open Igusa surfaces
3.6 Compactification of the Igusa surface along the supersingular locus

4. Modular forms modulo $p$
4.1 Modular forms mod $p$ as functions on $Ig(p)$
4.2 The theta operator
4.3 A study of the theta operator along the supersingular locus
4.4 Relation to the filtration and theta cycles
4.5 Compatibility between theta operators for elliptic and Picard modular forms

The $p$-adic theory of modular curves and modular forms, as developed in the early 70’s by Serre, Katz, Mazur, Manin and many others, inspired over the last thirty years an extensive study of higher dimensional Shimura varieties, with spectacular applications to almost every area in number theory and representation theory.

Date: August 1, 2014.
In the course of these vast generalizations, some of the more special phenomena pertaining to modular curves did not find appropriate analogues, and are only beginning to get addressed in recent years. In this work we study what we believe to be the simplest Shimura varieties beyond modular curves (and Shimura curves), namely Picard modular surfaces. For arithmetic questions over quadratic imaginary fields they play a role similar to the role played by modular curves over $\mathbb{Q}$. Our goal is to explore the arithmetic and geometry of these surfaces, and of modular forms over them, concentrating on aspects that have been only recently discovered, or appear to be new. Our hope is that staying within the bounds of $U(2,1)$ will keep the notation simple, and our tool-box relatively light, yet at the same time will allow for the exposition of new principles.

Our work relies on important work of Larsen, Bellaiche, Vollaard, Büttel, and Wedhorn. The modulo-$p$ geometry of Picard modular surfaces, and, more generally, Shimura varieties associated to unitary groups, has seen important contributions recently also by Goldring and Nicole, Koskivirta, Kudla, Rapoport, Terstiege, Howard and Pappas, and maybe others of which we are unaware. For completeness, we have included more than usual background material. We hope that we have not neglected to give appropriate credit when citing others’ results.

We have chosen to work with a prime $p$ which is inert in the underlying quadratic imaginary field, as this is the more interesting case, both from the algebro-geometric point of view, where the geometry modulo $p$ is more challenging, and from the representation theoretic point of view, where the unitary group is non-split. We do not touch upon questions of automorphic $L$-functions or representation theory. For these see the comprehensive volume [L-R] and the references therein. We do not touch upon Galois representations either, although eventually we hope that our results will be relevant to both these areas.

This paper concerns the algebraic geometry of Picard modular schemes and modular forms modulo $p$. In a subsequent work we hope to treat $p$-adic modular forms in the style of Serre and Katz, the canonical subgroup, overconvergence, and further topics. We shall now explain the main results and the structure of the paper.

Chapters 1-2 summarize known results, and serve as a systematic introduction on which chapters 3-4 and future work will be based. In particular, sections 1.4 and 2.4 rely on the theses of Larsen [La1] and Bellaiche [Bel]. We recommend the latter for its very readable exposition. Sections 1.2.4, 1.6, 2.1, 2.5 and 2.6 contain results or computations for which we did not find adequate references in the literature. Chapter 2 introduces the two basic automorphic vector bundles $\mathcal{P}$ (a plane bundle) and $\mathcal{L}$ (a line bundle). Modular forms are sections of vector bundles from the tensor algebra generated by these two, but we contend ourselves with scalar valued modular forms, which are sections of $\mathcal{L}^k$ ($k$ being the weight).

Chapter 3 contains new results. There are three strata to the Picard surface $S$ modulo $p$: the $\mu$-ordinary (open and dense) locus $S_\mu$, the general supersingular (one dimensional) locus $S_{gss}$, and the superspecial points $S_{ssp}$ (see [Bu-We],[V]). After recalling this, we find various relations between $\mathcal{P}$ and $\mathcal{L}$ that are peculiar to characteristic $p$. We study the Verschiebung homomorphism $V_\mathcal{L}$ from $\mathcal{L}$ to $\mathcal{P}^{(p)}$ and $V_\mathcal{P}$ from $\mathcal{P}$ to $\mathcal{L}^{(p)}$. We find that outside the superspecial points, both maps have rank one, but that $\text{Im}(V_\mathcal{L}) = \ker(V_\mathcal{P}^{(p)})$ is the defining equation of the general supersingular stratum $S_{gss}$. We prove that the line bundle $\mathcal{P}_0 = \ker(V_\mathcal{P})$, defined
over $S_\mu \cup S_{gss}$, does not extend across the superspecial points. Using these relations we define a Hasse invariant $h_{ss}$, a section of $L^{p^2 - 1}$, whose zero divisor is precisely the supersingular locus $S_{ss} = S_{gss} \cup S_{ssp}$. Although the same definition has been already given by Goldring and Nicole in [Go-Ni] for all unitary Shimura varieties, and recently generalized to all Shimura varieties of Hodge type by Koskivirta and Wedhorn [Ko-We], in the special case of Picard surfaces our analysis goes deeper.

Following the definition of $h_{ss}$ we define on $S_{ss}$ a secondary Hasse invariant $h_{ssp}$, a section of $L^{p^2 + 1}|_{ss}$, whose zeroes are the superspecial points $S_{ss}$ (albeit with high multiplicity). This secondary Hasse invariant is closely related to recent work of Boxer [Bo], and points at a general phenomenon that deserves further exploration. It seems reasonable that Boxer’s method will generalize to Shimura varieties of PEL type and will produce invariants which agree with the partial Hasse invariants appearing in [Gor]. As an application of our analysis of the secondary Hasse invariant we derive a striking formula, expressing the number of irreducible components of the supersingular stratum in terms of the Euler number of $\bar{S}_C$. By well-known results, due in this case to Holzapfel, this number is given essentially by an $L$-value.

As for modular curves, one can introduce an Igusa scheme $Ig(p^n)$ of any level $p^n$. We focus on $Ig = Ig(p)$, but note in passing that the higher $Ig(p^n)$ will become instrumental in the theory of $p$-adic modular forms, deferred to future work. Although initially defined only over the $\mu$-ordinary stratum, $Ig$ can be compactified over the whole of $S$ (and over the cuspidal divisors in $\bar{S}$ as well) by “extracting a $p^2 - 1$ root of $h_{ss}$.” The line bundle $L$ is trivialized over the $\mu$-ordinary locus $Ig_\mu$, and this allows us to regard modular forms as functions on $Ig$, with prescribed poles along $Ig_{gss}$, the supersingular locus. This is the geometric basis for developing an analogue of Serre’s theory of “modular forms modulo $p$” in our context. In particular, we study their Fourier-Jacobi expansions (complex FJ expansions are $q$-expansions with theta functions as coefficients, but here we employ the arithmetic analysis of Bellaiche, explained in Section 2.4). This leads to the notion of a “filtration” of a modular form mod $p$, as in the case of elliptic modular forms.

Chapter 4 seems to be entirely new. The theory becomes richer once we bring in the Gauss-Manin connection and the Kodaira-Spencer isomorphism. We define a theta operator on modular forms mod $p$ (in weight $k$) as follows. We first divide our modular form $f$ by the $k$-th power of the canonical section $a(1)$ of $L$ on $Ig$ to get a function $g = r(f)$ on $Ig$, with a pole of order $k$ along $Ig_{gss}$. We apply the (inverse of the) Kodaira-Spencer isomorphism to $dg$ to get a section of $P \otimes L$, which we map via $V_P \otimes 1$ to get a meromorphic section of $L^{p+1}$. Multiplying by $a(1)^k$ allows us to descend back to $S$, so we obtain a meromorphic modular form $\Theta(f)$ of weight $k + p + 1$. This construction is motivated by [An-Go], although substantial new phenomena appear in the present case. We study this operator. On the one hand, near the cusps we relate the retraction of a formal neighborhood of the cuspidal divisor, which was introduced by Bellaiche in [Bel], to the complex computations of section 1.6 and 2.6. This allows us to show that the effect of $\Theta$ on FJ expansions, is, as in the classical case, a “Tate twist”. On the other hand, we study $\Theta(f)$ along the supersingular locus, and show that, thanks to the fact that we have divided out by $P_0 = \ker V_P$, $\Theta(f)$ is in fact everywhere holomorphic! We derive some interesting consequences for “theta cycles”, where there are similarities, but also surprising deviations from the classical theory. We end the paper with a comparison of our
theta operator with the Serre-Katz theta operator on modular curves, embedded in $\bar{S}$. Theta operators on classical modular forms have been instrumental in studying congruences between them, with applications to Galois representations. We expect the same to be true for Picard modular forms.

1. Background

1.1. The unitary group and its symmetric space.

1.1.1. Notation. Let $K$ be an imaginary quadratic field, contained in $\mathbb{C}$. We denote by $\Sigma : K \hookrightarrow \mathbb{C}$ the inclusion and by $\bar{\Sigma} : K \hookrightarrow \mathbb{C}$ its complex conjugate. We use the following notation:

- $d_K$ - the square free integer such that $K = \mathbb{Q}(\sqrt{d_K})$.
- $D_K$ - the discriminant of $K$, equal to $d_K$ if $d_K \equiv 1 \mod 4$ and $4d_K$ if $d_K \equiv 2, 3 \mod 4$.
- $\delta_K = \sqrt{D_K}$ - the square root with positive imaginary part, a generator of the different of $K$, sometimes simply denoted $\delta$.
- $\omega_K = (1 + \sqrt{d_K})/2$ if $d_K \equiv 1 \mod 4$, otherwise $\omega_K = \sqrt{d_K}$, so that $O_K = \mathbb{Z} + \mathbb{Z}\omega_K$.
- $\bar{a}$ - the complex conjugate of $a \in K$.
- $\text{Im}(\delta(a)) = (a - \bar{a})/\delta$, for $a \in K$.

We fix an integer $N \geq 3$ (the “tame level”) and let $R_0 = O_K[1/2d_KN]$. This is our base ring. If $R$ is any $R_0$-algebra and $M$ is any $R$-module with $O_K$-action, then $M$ becomes an $O_K \otimes R$-module and we have a canonical type decomposition

$$M = M(\Sigma) \oplus M(\bar{\Sigma})$$

where $M(\Sigma) = e_{\Sigma}M$ and $M(\bar{\Sigma}) = e_{\bar{\Sigma}}M$, and where the idempotents $e_{\Sigma}$ and $e_{\bar{\Sigma}}$ are defined by

$$e_{\Sigma} = \frac{1 \otimes 1}{2} + \frac{\delta \otimes \delta^{-1}}{2}, \quad e_{\bar{\Sigma}} = \frac{1 \otimes 1}{2} - \frac{\delta \otimes \delta^{-1}}{2}.$$ 

Then $M(\Sigma)$ (resp. $M(\bar{\Sigma})$) is the part of $M$ on which $O_K$ acts via $\Sigma$ (resp. $\bar{\Sigma}$). The same notation will be used for sheaves of modules on $R$-schemes, endowed with an $O_K$ action. If $M$ is locally free, we say that it has type $(p, q)$ if $M(\Sigma)$ is of rank $p$ and $M(\bar{\Sigma})$ is of rank $q$.

We denote by

$$T = re_{\Sigma}G_m$$

the non-split torus whose $\mathbb{Q}$-points are $K^\times$, and by $\rho$ the non-trivial automorphism of $T$, which on $\mathbb{Q}$-points induces $\rho(a) = \bar{a}$. The group $G_m$ embeds in $T$ and the homomorphism $a \mapsto a \cdot \rho(a)$ from $T$ to itself factors through a homomorphism

$$N : T \to G_m,$$

the norm homomorphism. Its kernel $\ker(N)$ is denoted $T^1$.

1.1.2. The unitary group. Let $V = K^3$ and endow it with the hermitian pairing

$$\langle u, v \rangle = t \bar{u} \begin{pmatrix} \delta^{-1} \\ -\delta^{-1} \end{pmatrix} v.$$
We identify $V_\mathbb{R}$ with $\mathbb{C}^3$ ($\mathcal{K}$ acting via the natural inclusion $\Sigma$). It then becomes a hermitian space of signature $(2,1)$. Conversely, any 3-dimensional hermitian space over $\mathcal{K}$ whose signature at the infinite place is $(2,1)$ is isomorphic to $V$ after rescaling the hermitian form by a positive rational number.

Let
(1.6)  \[ G = GU(V, (\cdot, \cdot)) \]
be the general unitary group of $V$, regarded as an algebraic group over $\mathbb{Q}$. For any $\mathbb{Q}$-algebra $A$ we have
(1.7)  \[ G(A) = \{(g, \mu) \in GL_3(A \otimes \mathcal{K}) \otimes A^\times \mid (gu, gv) = \mu \cdot (u, v) \ \forall u, v \in V_A\}. \]

We write $G = G(\mathbb{Q})$, $G_\infty = G(\mathbb{R})$ and $G_p = G(\mathbb{Q}_p)$. A similar notational convention will apply to any algebraic group over $\mathbb{Q}$ without further ado. If $p$ splits in $\mathcal{K}$, $\mathbb{Q}_p \otimes \mathcal{K} \simeq \mathbb{Q}_p^2$ and $G_p$ becomes isomorphic to $GL_3(\mathbb{Q}_p) \times \mathbb{Q}_p^\times$. The isomorphism depends on the embedding of $\mathcal{K}$ in $\mathbb{Q}_p$, i.e. on the choice of a prime above $p$ in $\mathcal{K}$. For a non-split prime $p$ the group $G_p$, like $G_\infty$, is of (semisimple) rank $1$.

As $\mu$ is determined by $g$ we often abuse notation and write $g$ for the pair $(g, \mu)$ and $\mu(g)$ for the multiplier $\mu$. It is a character of algebraic groups over $\mathbb{Q}$, $\mu : G \to \mathbb{G}_m$. Another character is $\det : G \to \mathbb{T}$, defined by $\det(g, \mu) = \det(g)$. If we let
(1.8)  \[ \nu = \mu^{-1} \det : G \to \mathbb{T} \]
then both $\mu$ and $\det$ are expressible in terms of $\nu$, namely $\mu = \nu \cdot (\rho \circ \nu)$ and $\det = \nu^2 \cdot (\rho \circ \nu)$. The first relation is a consequence of the relation $\det \circ (\rho \circ \det) = \mu^3$, and the second is a consequence of the first and the definition of $\nu$.

The groups
(1.9)  \[ U = \ker \mu, \quad SU = \ker \nu = \ker \mu \cap \ker(\det) \]
are the unitary and the special unitary group, respectively.

We also introduce an alternating $\mathbb{Q}$-linear pairing $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{Q}$ defined by $\langle u, v \rangle = \text{Im}_A(u, v)$. We then have the formulae
(1.10)  \[ \langle au, v \rangle = \langle u, \bar{a}v \rangle, \quad 2(u, v) = \langle u, \bar{u}v \rangle + \delta(u, v). \]
We call $\langle \cdot, \cdot \rangle$ the polarization form, for a reason that will become clear soon.

1.1.3. The hermitian symmetric domain. The group $G_\infty = G(\mathbb{R})$ acts on $\mathbb{P}_2^\mathbb{C} = \mathbb{P}(V_\mathbb{R})$ by projective linear transformations and preserves the open subdomain $X$ of negative definite lines (in the metric $\langle \cdot, \cdot \rangle$). If we switch to coordinates in which the hermitian quadratic form $(u, u)$ assumes the standard shape $x\bar{x} + y\bar{y} - z\bar{z}$, it becomes evident that $X$ is biholomorphic to the open unit ball in $\mathbb{C}^2$, hence is connected, and $G_\infty$ acts on it transitively. We shall nevertheless stick to the coordinates introduced above. Every negative definite line is represented by a unique vector $^t(z, u, 1)$ and such a vector represents a negative definite line if and only if
(1.11)  \[ \lambda(z, u) \overset{\text{def}}{=} \text{Im}_A(z) - u\bar{u} > 0. \]
One refers to the realization of $X$ as the set of points $(z, u) \in \mathbb{C}^2$ satisfying this inequality as a Siegel domain of the second kind. It is convenient to think of the point $x_0 = (\delta/2, 0)$ as the “center” of $X$.

If we let $K_\infty$ be the stabilizer of $x_0$ in $G_\infty$, then $K_\infty$ is compact modulo center $(K_\infty \cap U(\mathbb{R}))$ is compact and isomorphic to $U(2) \times U(1))$. Since $G_\infty$ acts transitively on $X$, we may identify $X$ with $G_\infty/K_\infty$. 

The space $\mathcal{X}$ carries a Riemannian metric which is invariant under the action of $G_\infty$, the Bergmann metric. It may be described as follows. Switching once again to coordinates on $\mathbb{C}^3$ in which the hermitian quadratic form $(u, v)$ becomes $x\bar{x} + y\bar{y} - z\bar{z}$, the real symmetric bilinear form $\text{Re}(u, v)$ has signature $(4, 2)$, and the manifold given by the equation $x\bar{x} + y\bar{y} - z\bar{z} = -1$ becomes a circle bundle over $\mathcal{X}$. The restriction of $\text{Re}(u, v)$ to the tangent bundle of this manifold has signature $(4, 1)$ and when we take the quotient by the circle action we get an invariant Riemannian metric on $\mathcal{X}$.

The usual upper half plane embeds in $\mathcal{X}$ (holomorphically and isometrically) as the set of points where $u = 0$.

1.1.4. The cusps of $\mathcal{X}$. The boundary $\partial \mathcal{X}$ of $\mathcal{X}$ is the set of points $(z, u)$ where $\text{Im}_d(z) = u\bar{u}$, together with a unique point “at infinity” $c_\infty$ represented by the line $t(1 : 0 : 0)$. The lines represented by $\partial \mathcal{X}$ are the isotropic lines in $V_\mathbb{R}$. The set of cusps $C\mathcal{X}$ is the set of $K$-rational isotropic lines, or, equivalently, the set of points on $\partial \mathcal{X}$ with coordinates in $K$. If $s \in K$ and $r \in \mathbb{Q}$ we write

$$c_s^r = (r + \delta s\bar{s}/2, s).$$

Then $C\mathcal{X} = \{c_s^r | r \in \mathbb{Q}, s \in K\} \cup \{c_\infty\}$. The group $G = G(\mathbb{Q})$ acts transitively on the cusps.

The stabilizer of a cusp is a Borel subgroup in $G_\infty$. Since $G$ acts transitively on the cusps, we may assume that our cusp is $c_\infty$. It is then easy to check that its stabilizer $P_\infty$ has the form $P_\infty = M_\infty N_\infty$, where

$$M_\infty = \left\{ tm(\alpha, \beta) = t \begin{pmatrix} \alpha & \beta \\ \bar{\alpha}^{-1} & \bar{\beta} \end{pmatrix} | t \in \mathbb{R}_+^\times, \alpha \in \mathbb{C}^\times, \beta \in \mathbb{C}^1 \right\},$$

$$N_\infty = \left\{ n(u, r) = \begin{pmatrix} 1 & \delta \bar{u} & r + \delta u\bar{u}/2 \\ \delta u & 1 & u \\ 0 & 0 & 1 \end{pmatrix} | u \in \mathbb{C}, r \in \mathbb{R} \right\}.$$

The matrix $tm(\alpha, \beta)$ belongs to $U_\infty$ if and only if $t = 1$, and to $SU_\infty$ if furthermore $\beta = \bar{\alpha}/\alpha$. The group $N_\infty$ is contained in $SU_\infty$. The group $P = P_\infty \cap G$ consists of the same matrices with $K$-rational entries. Since $N = N_\infty \cap G$ still acts transitively on the set of finite cusps $c_s^r$, we conclude that $G$ acts doubly transitively on $C\mathcal{X}$.

Of particular interest to us will be the geodesics connecting an interior point $(z, u)$ to a cusp $c \in C\mathcal{X}$. If $(z, u) = n(u, r)m(d, 1)x_0$ (recall $x_0 = t(\delta/2 : 0 : 1)$) where $d$ is real and positive (i.e. $r = \text{Re} z$ and $d = \sqrt{\lambda(z, u)}$) then the geodesic connecting $(z, u)$ to $c_\infty$ can be described by the formula

$$\gamma_r^u(t) = n(u, r)m(t, 1)x_0 = (r + \delta(u\bar{u} + t^2)/2, u) \ (d \leq t < \infty).$$

The same geodesic extends in the opposite direction for $0 < t \leq d$, and if $u$ and $r$ lie in $K$, it ends there in the cusp $c_s^r$. We shall call $\gamma_r^u(t)$ the geodesic retraction of $\mathcal{X}$ to the cusp $c_\infty$. As $0 < t < \infty$ these parallel geodesics exhaust $\mathcal{X}$, they converge to $c_\infty$ as $t \to \infty$, and they pass through $(z, u)$ precisely when $r = \text{Re} z$ and $t = \sqrt{\lambda(z, u)}$.

\footnote{Note that any proper $\mathbb{R}$-parabolic subgroup of $G_\infty$ is conjugate to $P_\infty$, as $SU(2, 1)$ has $\mathbb{R}$-rank 1.}
The points \((z, u)\) and \((z', u')\) lie on the same geodesic if and only if \(u = u'\) and \(\text{Re}(z) = \text{Re}(z')\).

1.2. Picard modular surfaces over \(\mathbb{C}\).

1.2.1. Lattices and their arithmetic groups. Fix an \(\mathcal{O}_K\)-invariant lattice \(L \subset V\) which is self-dual in the sense that

\[
L = \{u \in V | \langle u, v \rangle \in \mathbb{Z} \forall v \in L\}.
\]

Equivalently, \(L\) is its own \(\mathcal{O}_K\)-dual with respect to the hermitian pairing \(\langle , \rangle\). We assume also that the Steinitz class \(\mathcal{B}\) of \(L\) as an \(\mathcal{O}_K\)-module is \([\mathcal{O}_K]\), or, what amounts to the same, that \(L\) is a free \(\mathcal{O}_K\)-module. When we introduce the Shimura variety later on, we shall relax this last assumption, but the resulting scheme will be disconnected (over \(\mathbb{C}\)).

Fix an integer \(N \geq 1\) and let

\[
\Gamma = \{g \in G | gL = L \text{ and } g(u) \equiv u \mod NL \forall u \in L\}.
\]

This \(\Gamma\) is a discrete subgroup of \(G_\infty\), contained in \(U_\infty\). It is easy to see that if \(N \geq 3\) then \(\Gamma\) is torsion free, acts freely and faithfully on \(X\), and is contained in \(SU_\infty\). From now on we assume that this is the case.

If \(g \in G\) and \(\mu(g) = 1\) (i.e. \(g \in U\)) the lattice \(gL\) is another lattice of the same sort and the discrete group corresponding to it is \(g \Gamma g^{-1}\). Since \(U\) acts transitively on the cusps, this reduces the study of \(\Gamma \backslash X\) near a cusp to the study of a neighborhood of the standard cusp \(c_\infty\) (at the price of changing \(L\) and \(\Gamma\)).

It is important to know the classification of lattices \(L\) as above (self-dual and \(\mathcal{O}_K\)-free). Let \(e_1, e_2, e_3\) be the standard basis of \(K^3\). Let

\[
L_0 = \text{Span}_{\mathcal{O}_K}\{\delta e_1, e_2, e_3\}
\]

and

\[
L_1 = \text{Span}_{\mathcal{O}_K}\{\frac{\delta}{2} e_1 + e_3, e_2, \frac{\delta}{2} e_1 - e_3\}.
\]

These two lattices are self-dual and of course, \(\mathcal{O}_K\)-free. The following theorem is based on the local-global principle and a classification of lattices over \(\mathbb{Q}_p\) by Shimura [Sh1].

**Lemma 1.1.** ([La1], p.25). For any lattice \(L\) as above there exists a \(g \in U\) such that \(gL = L_0\) or \(gL = L_1\). If \(D_K\) is odd, \(L_0\) and \(L_1\) are equivalent. If \(D_K\) is even, they are inequivalent.

Indeed, if \(D_K\) is even, \(L_0 \otimes \mathbb{Q}_p\) and \(L_1 \otimes \mathbb{Q}_p\) are \(U_p\)-equivalent for every \(p \neq 2\), but not for \(p = 2\).

1.2.2. Picard modular surfaces and the Baily-Borel compactification. We denote by \(X_\Gamma\) the complex surface \(\Gamma \backslash X\). Since the action of \(\Gamma\) is free, \(X_\Gamma\) is smooth. We describe a topological compactification of \(X_\Gamma\). A standard neighborhood of the cusp \(c_\infty\) is an open set of the form

\[
\Omega_R = \{(z, u) | \lambda(z, u) > R\}.
\]

The set \(C_\Gamma = \Gamma \backslash \mathcal{C}X\) is finite, and we write \(C_\Gamma = C_c\). We let \(X_\Gamma^*\) be the disjoint union of \(X_\Gamma\) and \(C_\Gamma\). We topologize it by taking \(\Gamma \backslash \Omega_R \cup \{c_\infty, t\}\) as a basis of neighborhoods.\(^2\)

\(^2\)The Steinitz class of a finite projective \(\mathcal{O}_K\)-module is the class of its top exterior power as an invertible module.
at $c_{\infty,\Gamma}$. If $c = g(c_{\infty})$ where $g \in U$, we take $g(g^{-1} \Gamma g \backslash \Omega_R) \cup \{c_{\Gamma}\}$ instead. The following theorem is well-known.

**Theorem 1.2.** (Satake, Baily-Borel) $X_{\Gamma}^*$ is projective and the singularities at the cusps are normal. In other words, there exists a normal complex projective surface $S_{\Gamma}^*$ and a homeomorphism $\iota : S_{\Gamma}^*(\mathbb{C}) \simeq X_{\Gamma}^*$, which on $S_{\Gamma}(\mathbb{C}) = \iota^{-1}(X_{\Gamma})$ is an isomorphism of complex manifolds. $S_{\Gamma}^*$ is uniquely determined up to isomorphism.

1.2.3. The universal abelian variety over $X_{\Gamma}$. With $x \in \mathfrak{X}$ and with our choice of $L$ we shall now associate a PEL-structure $\mathbf{A}_x = (A_x, \lambda_x, \iota_x, \alpha_x)$ where

1. $A_x$ is a 3-dimensional complex abelian variety,
2. $\lambda_x$ is a principal polarization on $A_x$ (i.e. an isomorphism $A_x \simeq A_x^\vee$ with its dual abelian variety induced by an ample line bundle),
3. $\iota_x : \mathcal{O}_K \hookrightarrow End(A_x)$ is an embedding of CM type (2, 1) (i.e. the action of $\iota(a)$ on the tangent space of $A_x$ at the origin induces the representation $2\Sigma + \overline{\Sigma}$ such that the Rosati involution induced by $\lambda_x$ preserves $\iota(\mathcal{O}_K)$ and is given by $\iota(a) \mapsto \iota(\overline{a})$),
4. $\alpha_x : N^{-1}L/L \simeq A_x[N]$ is a full level $N$ structure, compatible with the $\mathcal{O}_K$-action and the polarization. The latter condition means that if we denote by $\langle, \rangle_\lambda$ the Weil “$e_N$-pairing” on $A_x[N]$ induced by $\lambda_x$, then for $l, l' \in N^{-1}L$

\[
\langle \alpha_x(l), \alpha_x(l') \rangle_\lambda = e^{2\pi i \langle l, l' \rangle}.
\]

Let $W_x$ be the negative definite complex line in $V_{\mathbb{R}} = \mathbb{C}^3$ defined by $x$, and $W_x^\perp$ its orthogonal complement, a positive definite plane. Let $J_x$ be the complex structure which is multiplication by $i$ on $W_x^\perp$ and by $-i$ on $W_x$. Let $A_x = (V_{\mathbb{R}}, J_x)/L$. Then the polarization form $\langle, \rangle$ is a Riemann form on $L$. One only has to verify that $\langle u, J_x v \rangle + i \langle u, v \rangle$ is a positive definite hermitian form. But up to the factor $|\delta|/2$ this is the same as $\langle u, v \rangle$ on $W_x^\perp$ and $-\langle u, v \rangle$ on $W_x$. The Riemann form determines a principal polarization on $A_x$ as usual. The action of $\mathcal{O}_K$ is derived from the underlying $\mathcal{K}$ structure of $V$. As we have changed the complex structure on $W_x$, the CM type is now (2, 1). Finally the level $N$ structure $\alpha_x$ is the identity map.

If $\gamma \in \Gamma$ then $\gamma$ induces an isomorphism between $\mathbf{A}_x$ and $\mathbf{A}_{x'}$ (x). Conversely, if $\mathbf{A}_x$ and $\mathbf{A}_{x'}$ are isomorphic structures, it is easy to see that $x'$ and $x$ must belong to the same $\Gamma$-orbit. It follows that points of $X_{\Gamma}$ are in a bijection with PEL structures of the abelian variety for which the triple

\[
(H_1(A_x, \mathbb{Z}), \iota_x, \langle, \rangle_\lambda)
\]

is isomorphic to $(L, \iota, \langle, \rangle)$ (here $\iota$ refers to the $\mathcal{O}_K$ action on $L$), with the further condition that $\alpha_x$ is compatible with the isomorphism between $L$ and $H_1(A_x, \mathbb{Z})$ in the sense that we have a commutative diagram

\[
\begin{array}{ccc}
0 & \to & L \\
\downarrow & & \downarrow \\
0 & \to & H_1(A_x, \mathbb{Z}) \\
\end{array}
\begin{array}{ccc}
& & \alpha_x \\
\downarrow & & \downarrow \\
& & A_x[N] \\
\end{array}
\rightarrow \begin{array}{ccc}
N^{-1}L & \rightarrow & N^{-1}L/L \\
\downarrow & & \downarrow \\
N^{-1}H_1(A_x, \mathbb{Z}) & \rightarrow & A_x[N] \\
\end{array} \rightarrow 0
\]

1.2.4. A “moving lattice” model for the universal abelian variety. We want to assemble the individual $A_x$ into an abelian variety $A$ over $\mathfrak{X}$. In other words, we want to construct a 5-dimensional complex manifold $A$, together with a holomorphic map
A \to \mathcal{X} whose fiber over x is identified with A_x. For that, as well as for the computation of the Gauss-Manin connection below, it is convenient to introduce another model, in which the complex structure on \mathbb{C}^3 is fixed, but the lattice varies.

For simplicity we assume from now on that \(L = L_0\) is spanned over \(\mathcal{O}_K\) by \(\delta e_1, e_2\) and \(e_3\). The case of \(L_1\) can be handled similarly.

Let \(\mathbb{C}^3\) be given the usual complex structure, and let \(a \in \mathcal{O}_K\) act on it via the matrix

\[
(1.24) \quad \iota'(a) = \begin{pmatrix} a & \bar{a} \\ a & \bar{a} \end{pmatrix}.
\]

Given \(x = (z, u) \in \mathcal{X}\) consider the lattice

\[
(1.25) \quad L'_x = \text{Span}_{\iota'(\mathcal{O}_K)} \left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ -1 \\ -u \end{pmatrix}, \begin{pmatrix} u \\ -z/\delta \\ z/\delta \end{pmatrix} \right\} \subset \mathbb{C}^3.
\]

The map \(T_x : \mathbb{C}^3 \to \mathbb{C}^3\) which sends \(\zeta = (\zeta_1, \zeta_2, \zeta_3)\) to

\[
(1.26) \quad T(\zeta) = \lambda(z, u)^{-1} \left\{ -\zeta_1 \begin{pmatrix} \bar{u}z \\ \bar{u} (z - \bar{z})/\delta \end{pmatrix} - \zeta_2 \begin{pmatrix} \bar{z} + \bar{\delta} u \bar{u} \\ u \\ 1 \end{pmatrix} + \zeta_3 \begin{pmatrix} z \\ u \\ 1 \end{pmatrix} \right\}
\]

is a complex linear isomorphism between \(\mathbb{C}^3\) and \((V_\mathbb{R}, J_x)\). In fact, it sends \(\mathbb{C}e_1 + \mathbb{C}e_2\) linearly to \(W_\mathbb{R}^+\) and \(\mathbb{C}e_3\) conjugate-linearly to \(W_x\). It intertwines the \(\iota'\) action of \(\mathcal{O}_K\) on \(\mathbb{C}^3\) with its \(\iota\) action on \((V_\mathbb{R}, J)\). It furthermore sends \(L'_x\) to \(L\). In fact, an easy computation shows that it sends the three generating vectors of \(L'_x\) to \(\delta e_1, e_2\) and \(e_3\), respectively. We conclude that \(T_x\) induces an isomorphism

\[
(1.27) \quad T_x : A'_x = \mathbb{C}^3/L'_x \simeq A_x.
\]

Consider the differential forms \(d\zeta_1, d\zeta_2\) and \(d\zeta_3\). As their periods along any \(l \in L'_x\) vary holomorphically in \(z\) and \(u\), the five coordinates \(\zeta_1, \zeta_2, \zeta_3, z, u\) form a local system of coordinates on the family \(A' \to \mathcal{X}\). Identifying \(A'\) with \(A\) allows us to put the desired complex structure on the family \(A\). Alternatively, we may define \(A'\) as the quotient of \(\mathbb{C}^3 \times \mathcal{X}\) by \(\zeta \mapsto \zeta + l(z, u)\) where \(l(z, u)\) varies over the holomorphic lattice-sections.

The model \(A'\) has another advantage, that will become clear when we examine the degeneration of the universal abelian variety at the cusp \(c_\infty\). It suffices to note at this point that the first two of the three generating vectors of \(L'_x\) depend only on \(u\).

1.3. The Picard moduli scheme.

1.3.1. The moduli problem. Fix \(N \geq 3\) as before. Fix the lattice \(L = L_0 \subset V = K^3\).

Let \(R\) be an \(R_0\)-algebra (recall \(R_0 = \mathcal{O}_K[1/2d_KN]\)). Let \(\mathcal{M}(R)\) be the collection of (isomorphism classes of) PEL structures \((A, \lambda, \iota, \alpha)\) where

\begin{enumerate}
    \item \(A/R\) is an abelian scheme of relative dimension 3
    \item \(\lambda : A \simeq A^t\) is a principal polarization
    \item \(\iota : \mathcal{O}_K \to \text{End}(A/R)\) is a homomorphism such that
        \begin{enumerate}
            \item \(\iota\) makes \(\text{Lie}(A/R)\) a locally free \(R\)-module of type \((2, 1)\),
            \item the Rosati involution induced on \(\iota(\mathcal{O}_K)\) by \(\lambda\) is \(\iota(a) \mapsto \iota(\bar{a})\).
        \end{enumerate}
\end{enumerate}
(4) $α : N^{-1}L/L \simeq A[N]$ is an isomorphism of $\mathcal{O}_K$-group schemes over $R$ which is compatible with the polarization in the sense that there exists an isomorphism $ν_N : \mathbb{Z}/NZ \simeq μ_N$ of group schemes over $R$ such that

$$\left(α \left(\frac{l}{N}\right), α \left(\frac{l'}{N}\right)\right) = ν_N(⟨l, l'⟩ \mod N).$$

In addition we require that for every multiple $N'$ of $N$, locally étale over $\text{Spec}(R)$, there exists a similar level $N'$-structure $α'$, restricting to $α$ on $N^{-1}L/L$. One says that $α$ is locally étale symplectic liftable ([Lan], 1.3.6.2).

In view of Lemma [1.1] the last condition of symplectic liftable is void if $D_K$ is odd, while if $D_K$ is even it is equivalent to the following condition ([Bel], 1.3.1):

- For any geometric point $η : R → k$ ($k$ algebraically closed field, necessarily of characteristic different from 2), the $\mathcal{O}_K \otimes \mathbb{Z}_2$ polarized module $(T_2A_η; ⟨⟩; 〈, 〉)$ is isomorphic to $(L \otimes \mathbb{Z}_2; ⟨⟩)$ under a suitable identification of $\text{lim}_H μ_2^n(k)$ with $\mathbb{Z}_2$.

The choice of $L_0$ was arbitrary. If we took $L_1$ as our basic lattice we would get a similar moduli problem.

A level $N$ structure $α$ can exist only if the group schemes $\mathbb{Z}/NZ$ and $μ_N$ become isomorphic over $R$, but the isomorphism $ν_N$ is then determined by $α$.

$\mathcal{M}$ becomes a functor on the category of $R_0$-algebras (and more generally, on the category of $R_0$-schemes) in the obvious way. The following theorem is of fundamental importance ([Lan], 1.4.1.11).

**Theorem 1.3.** The functor $R → \mathcal{M}(R)$ is represented by a smooth quasi-projective scheme $S$ over $\text{Spec}(R_0)$, of relative dimension 2.

We call $S$ the (open) Picard modular surface of level $N$. It comes equipped with a universal structure $(A, λ, τ, α)$ of the above type over $S$. We call $A$ the universal abelian scheme over $S$. For every $R_0$-algebra $R$ and PEL structure in $\mathcal{M}(R)$, there exists a unique $R$-point of $S$ such that the given PEL structure is obtained from the universal one by base-change.

1.3.2. **The Shimura variety $Sh_K$.** We briefly recall the interpretation of the Picard modular surface as a canonical model of a Shimura variety. The symmetric domain $X$ can be interpreted as a $G_∞$-conjugacy class of homomorphisms

$$h : S = res^G_{\mathbb{C}} G_m → G$$

turning $(G, X)$ into a Shimura datum in the sense of Deligne [De]. The reflex field associated to this datum turns out to be $K$. Let $K_∞$ be the stabilizer of $x_0$ in $G_∞$ and $K_f^0 \subset G(\mathbb{A}_f)$ the subgroup stabilizing $\tilde{L} = L \otimes \hat{\mathbb{Z}}$. Let $K_f$ be the subgroup of $K_f^0$ inducing the identity on $L/NL$. Let $K = K_∞K_f \subset G(\mathbb{A})$. Then the Shimura variety $Sh_K$ is a complex quasi-projective variety whose complex points are isomorphic, as a complex manifold, to the double coset space

$$Sh_K(\mathbb{C}) \simeq G(\mathbb{Q}) \backslash G(\mathbb{A}) / K \simeq G(\mathbb{Q}) \backslash (X \times G(\mathbb{A}_f) / K_f).$$

The theory of Shimura varieties provides a canonical model for $Sh_K$ over $K$. The following important theorem complements the one on the representability of the functor $\mathcal{M}$.

**Theorem 1.4.** The canonical model of $Sh_K$ is the generic fiber $S_K$ of $S$. 
Let us explain only how to associate to a point of $Sh_K(\mathbb{C})$ a point in $S(\mathbb{C})$. For that we have to associate an element of $\mathcal{M}(\mathbb{C})$ to $g \in G(\mathbb{A})$, and show that the structures associated to $g$ and to $\gamma g k$ ($\gamma \in G, k \in K$) are isomorphic. Let $x = x_g = g_? (x_0) \in \mathbb{C}$. Let $L_g = g_f (\hat{L}) \cap V$ (the intersection taking place in $V_{\mathbb{A}} = \hat{L} \otimes \mathbb{Q}$) and
\begin{equation}
A_g = (V_{\mathbb{R}}, J_x)/L_g.
\end{equation}
Note that $J_x$ depends only on $g_? K_\infty$ and $L_g$ only on $g_f K_f^0$, so $A_g$ depends only on $g K^0$.

Let $\hat{\mu}(g)$ be the unique positive rational number such that for every prime $p$,
\begin{equation}
\text{ord}_p \hat{\mu}(g) = \text{ord}_p \mu(g_p).
\end{equation}
Such a rational number exists since $\mu(g_p)$ is a $p$-adic unit for almost all $p$ and $\mathbb{Q}$ has class number 1. We claim that
\begin{equation}
\langle \cdot, \cdot \rangle_g = \hat{\mu}(g)^{-1} \langle \cdot, \cdot \rangle : L_g \times L_g \to \mathbb{Q}
\end{equation}
induces a principal polarization $\lambda_g$ on $A_g$. That this is a (rational) Riemann form follows from the fact that $\langle u, v \rangle_{J_x} = \langle u, J_x v \rangle + i \langle u, v \rangle$ is hermitian positive definite. That $\langle \cdot, \cdot \rangle_g$ is indeed $\mathbb{Z}$-valued and $L_g$ is self-dual follows from the choice of $\hat{\mu}(g)$ since locally at $p$ the dual of $g_p L_p$ under $\langle \cdot, \cdot \rangle : V_p \times V_p \to \mathbb{Q}_p$ is $\mu(g_p)^{-1} g_p L_p$. We conclude that there exists a unique polarization $\lambda_g : A_g \to A_g^*$ such that
\begin{equation}
\langle u, v \rangle_{\lambda_g} = \exp(2\pi i \langle u, v \rangle_g)
\end{equation}
for every $u, v \in A_g[l] = l^{-1} L_g/L_g$ and every $l \geq 1$. This polarization is principal.

Since $g_f$ commutes with the $K$-structure on $V_{\mathbb{A}}$, $L_g$ is still an $O_K$-lattice, hence $\dot{\iota}_g$ is defined.

Finally $\alpha_g$ is derived from
\begin{equation}
N^{-1} L/L = N^{-1} \hat{L}/\hat{L} \xrightarrow{g_f} N^{-1} \hat{L}_g/\hat{L}_g = N^{-1} L_g/L_g = A_g[N].
\end{equation}
We note that $\alpha_g$ depends only on $g K$ because $K_f \subset K_f^0$ is the principal level-$N$ subgroup, and that it lifts to level $N'$ structure for any multiple $N'$ of $N$, by the same formula. The isomorphism $\nu_{N,g}$ between $\mathbb{Z}/N\mathbb{Z}$ and $\mu_N(\mathbb{C})$ that makes (1.28) work is self-evident (see (1.49)). Let $A_g \in \mathcal{M}(\mathbb{C})$ be the structure just constructed.

Let now $\gamma \in G(\mathbb{Q})$. Then the action of $\gamma$ on $V$ induces an isomorphism between the tuples $A_g$ and $A_{\gamma g}$. Indeed, $\gamma : V_{\mathbb{R}} \to V_{\mathbb{R}}$ intertwines the complex structures $x_g$ and $x_{\gamma g}$, and carries $L_g$ to $L_{\gamma g}$, so induces an isomorphism of the abelian varieties, which clearly commutes with the PEL structures.

This shows that $A_g$ depends solely on the double coset of $g$ in $G(\mathbb{Q}) \backslash G(\mathbb{A})/K$. One is left now with two tasks which we do not do in this survey: (i) Proving that if $A_g \simeq A_{g'}$, then $g$ and $g'$ belong to the same double coset, and that every $A \in \mathcal{M}(\mathbb{C})$ is obtained in this way, (ii) Identifying the canonical model of $Sh_K$ over $\mathbb{C}$ with $S_K$. 

1.3.3. The connected components of $Sh_K$. Recall that $G' = SU = \ker(\nu : G \to T)$. Since $G'$ is simple and simply connected, strong approximation holds and
\begin{equation}
G'(\mathbb{A}) = G'(\mathbb{Q}) G'_\infty K_f^0.
\end{equation}
Here $K' = K \cap G'(\mathbb{A}), K'_f = K \cap G'(\mathbb{A}_f)$. From the connectedness of $G'_\infty$ we deduce that
\begin{equation}
G'(\mathbb{Q})\backslash G'(\mathbb{A})/K'
\end{equation}
is connected.

As $N \geq 3$, $\nu(K) \cap K^\infty = \{1\}$. Here $K^\infty = \nu(G(\mathbb{Q}))$, and it follows that
\begin{equation}
G'(\mathbb{Q})\backslash G'(\mathbb{A})/K' \rightarrow G(\mathbb{Q})\backslash G(\mathbb{A})/K
\end{equation}
is injective. We now claim (see also Theorem 2.4 and 2.5 of [De]) that
\begin{equation}
\nu : \pi_0(G(\mathbb{Q})\backslash G(\mathbb{A})/K) \approx \pi_0(T(\mathbb{Q})\backslash T(\mathbb{A})/\nu(K))
\end{equation}
is a bijection. For $\nu$ is surjective ([De] (0.2)) and continuous (on double coset spaces) so clearly induces a surjective map between the sets of connected components. On the other hand if $[g_1]$ and $[g_2]$ (double cosets of $g_i \in G(\mathbb{A})$) are mapped by $\nu$ to the same connected component in $T(\mathbb{Q})\backslash T(\mathbb{A})/\nu(K)$, then since $G_\infty$ is mapped onto the connected component of the identity in $T(\mathbb{Q})\backslash T(\mathbb{A})/\nu(K)$, modifying $g_1$ by an element of $G_\infty$ we may assume that
\begin{equation}
\nu([g_1]) = \nu([g_2]) \in T(\mathbb{Q})\backslash T(\mathbb{A})/\nu(K),
\end{equation}
without changing the connected component in which $[g_1]$ lies. Once this has been established, for appropriate representatives $g_i$ of the double cosets, $g_i^{-1}g_2 \in G'(\mathbb{A})$, so by the connectedness of $G'(\mathbb{Q})\backslash G'(\mathbb{A})/K'$, $[g_1]$ and $[g_2]$ lie in the same connected component of $G(\mathbb{Q})\backslash G(\mathbb{A})/K$.

The group $\pi_0(T(\mathbb{Q})\backslash T(\mathbb{A})/\nu(K))$ is the group
\begin{equation}
K^\infty\backslash K_\mathbb{A}^\infty/\mathbb{C}^\times \nu(K_f) = K^\infty\backslash K_f^\infty/\nu(K_f).
\end{equation}
It sits in a short exact sequence
\begin{equation}
0 \rightarrow \mu_K \backslash U_K/\nu(K_f) \rightarrow K^\infty\backslash K_f^\infty/\nu(K_f) \overset{\rho_f}{\rightarrow} Cl_K \rightarrow 0,
\end{equation}
where $U_K$ is the product of local units at all the finite primes and $Cl_K$ is the class group.

### 1.3.4. The $cl$ and $\nu_N$ invariants of a connected component.

The norm $N : K^\times \rightarrow \mathbb{Q}^\times$ satisfies $N \circ \nu = \nu \circ \rho = \mu$, hence induces a map
\begin{equation}
K^\times\backslash K_f^\times/\nu(K_f) \rightarrow Q_+^\times\backslash Q_f^\times/\mu(K_f).
\end{equation}

Using the lattice $L$ as an integral structure in $V$, we see that $G$ comes from a group scheme $G_Z$ over $\mathbb{Z}$, whose points in any ring $A$ are
\begin{equation}
G_Z(A) = \{(g, \mu) \in GL_{O_{\mathbb{Z}} \otimes \mathbb{A}}(L_A) \times A^\times | \langle gu, gv \rangle = \mu \langle u, v \rangle \}.
\end{equation}
We likewise get that $\mu$ is a homomorphism from $G_Z$ to $G_m$. The diagram
\begin{equation}
\begin{array}{c}
G_Z(\mathbb{Z}_p) \downarrow \mu \rightarrow Z_p^\times \\
\rightarrow G_Z(\mathbb{Z}_p/\mathbb{N}\mathbb{Z}_p) \downarrow \mu \\
\end{array}
\end{equation}
commutes, $G_Z(\mathbb{Z}_p) = K_p^0$ and the kernel of $G_Z(\mathbb{Z}_p) \rightarrow G_Z(\mathbb{Z}_p/\mathbb{N}\mathbb{Z}_p)$ is $K_p$. This shows that $\mu(K_f) \subset \mathbb{Z}^\times(N)$, the product of local units congruent to 1 mod $N$. But
\begin{equation}
Q_+^\times\backslash Q_f^\times/\mathbb{Z}^\times(N) = (\mathbb{Z}/\mathbb{N}\mathbb{Z})^\times.
\end{equation}
To conclude, we have shown the existence of two maps from the set of connected components:

\[(1.46) \quad cl : \pi_0(G(\mathbb{Q}) \setminus G(\mathbb{A})/K) \to Cl_K \]

\[(1.47) \quad \nu_N : \pi_0(G(\mathbb{Q}) \setminus G(\mathbb{A})/K) \to (\mathbb{Z}/N\mathbb{Z})^\times.\]

These two maps are independent: together they map \(\pi_0(G(\mathbb{Q}) \setminus G(\mathbb{A})/K)\) onto \(Cl_K \times (\mathbb{Z}/N\mathbb{Z})^\times\). On the other hand, they have a non-trivial common kernel, which grows with \(N\), as is evident from the interpretation of \(K^\times \setminus \mathbb{C}^\times /\nu(K_f)\) as the Galois group of a certain class field extension of \(K\). The map \(cl\) gives the restriction to the Hilbert class field, while the map \(\nu_N\) gives the restriction to the cyclotomic field \(\mathbb{Q}(\mu_N)\). We have singled out \(cl\) and \(\nu_N\), because when \(N \geq 3\), they have an interpretation in terms of the complex points of \(Sh_K\).

**Proposition 1.5.** Let \([g] \in G(\mathbb{Q}) \setminus G(\mathbb{A})/K = Sh_K(\mathbb{C})\). Then

(i) \(cl([g])\) is the Steinitz class of the lattice \(L_g = g_f(L) \cap V\) in \(Cl_K\).

(ii) \(\nu_N([g])\) is (essentially) the \(\nu_N, g\) that appears in the definition of \(\alpha_g\) (see \(1.3.2\)).

**Proof.** (i) \(cl([g])\) is the class of the ideal \((\nu(g_f))\) associated to the idele \(\nu(g_f) \in K_f^\times\).

This ideal is in the same class as \((\det(g_f))\), because \(\mu(g_f) \in \mathbb{Q}_f^\times\), so \((\mu(g_f))\) is principal. But the class of \((\det(g_f))\) is the Steinitz class of \(L_g\), since the Steinitz class of \(L\) is trivial.

(ii) To find \(\nu_N([g])\) we first project the idele \(\mu(g_f)\) to \(\hat{\mathbb{Z}}^\times\) using \(\mathbb{Q}_f^\times \to \hat{\mathbb{Z}}^\times\). But this is just \(\tilde{\mu}(g_f)^{-1}\mu(g_f)\). We then take the result modulo \(N\), so

\[(1.48) \quad \nu_N([g]) = \tilde{\mu}(g_f)^{-1}\mu(g_f) \mod N.\]

Now the definition of the tuple \((A_g, \lambda_g, \iota_g, \alpha_g)\) is such that if \(u, v \in N^{-1}L/L\) then

\[
\left\langle \alpha_g(u), \alpha_g(v) \right\rangle_{\lambda_g} = \exp \left(2\pi i N \left(g_fu, g_f v \right)_{g_f} \right)
\]

\[= \exp \left(2\pi i \tilde{\mu}(g_f)^{-1}\mu(g_f)N \left(g_fu, g_f v \right) \right)
\]

\[= \exp \left(2\pi i \tilde{\nu}_N([g])N \left(u, v \right) \right)
\]

\[(1.49) \quad \left\langle \alpha_g(u), \alpha_g(v) \right\rangle_{\lambda_g} = \exp \left(2\pi i \tilde{\nu}_N([g])N \left(u, v \right) \right)
\]

Part (ii) follows if we identify \(\nu_{N,g} \in Isom_{\mathbb{C}}(N^{-1}\mathbb{Z}/\mathbb{Z}, \mu_N)\) with \(\nu_N([g]) \in (\mathbb{Z}/N\mathbb{Z})^\times\) using \(\exp(2\pi i \cdot)\). \(\square\)

1.3.5. The complex uniformization. Recall that \(X = G_\infty/K_\infty\) and that it was equipped with a base point \(x_0\) (corresponding to \((z, u) = (\delta_K/2, 0)\) in the Siegel domain of the second kind). Let \(1 = g_1, \ldots, g_m \in G(\mathbb{A}_f)\) \((m = \#(K^\times \setminus \mathbb{C}^\times /\nu(K_f)))\) be representatives of the connected components of \(G(\mathbb{Q}) \setminus G(\mathbb{A})/K\), and define congruence groups

\[(1.50) \quad \Gamma_j = G(\mathbb{Q}) \cap g_jK_fg_j^{-1}.\]

We write \([x, g_j]\) for \(G(\mathbb{Q})(x, g_jK_f) \in G(\mathbb{Q}) \setminus (X \times G(\mathbb{A}_f)/K_f) = G(\mathbb{Q}) \setminus G(\mathbb{A})/K\).

Then \([x', g_j] = [x, g_j]\) if and only if \(x' = \gamma x\) for \(\gamma \in \Gamma_j\). The map

\[(1.51) \quad \prod_{j=1}^{m} X_{\Gamma_j} = \prod_{j=1}^{m} \Gamma_j \backslash X \simeq Sh_K(\mathbb{C})\]

sending \(\Gamma_j x\) to \([x, g_j]\) is an isomorphism.
Note that $\Gamma_1 = \Gamma$ is the principal level-$N$ congruence subgroup in $G_2(Z)$, the stabilizer of $L$. Similarly, $\Gamma_j$ is the principal level-$N$ congruence subgroup in the stabilizer of $L_{O_j}$, and is thus a group of the type considered in [1.2.1] except that we have dropped the assumption on the Steinitz class of $L_{O_j}$. As $N \geq 3$, $\det(\gamma) = 1$ and $\mu(\gamma) = 1$ for all $\gamma \in \Gamma_j$, for every $j$. Indeed, on the one hand these are in $K^\times$ and $Q_+^\times$ respectively. On the other hand, they are local units which are congruent to $1 \mod N$ everywhere. It follows that $\Gamma_j$ are subgroups of $G^\prime(Q) = SU(Q)$.

We get a similar decomposition to connected components (as an algebraic surface)

$$S_C = \prod_{j=1}^m S_{\Gamma_j}$$

and we write $S_C = \prod_{j=1}^m S_{\Gamma_j}^*$ for the Baily-Borel compactification.

1.4. Smooth compactifications.

1.4.1. The smooth compactification of $X_\Gamma$. We begin by working in the complex analytic category and follow the exposition of [Cog]. The Baily-Borel compactification $X_\Gamma^*$ is singular at the cusps, and does not admit a modular interpretation. For general unitary Shimura varieties, the theory of toroidal compactifications provides smooth compactifications that depend, in general, on extra data. It is a unique feature of Picard modular surfaces, stemming from the finiteness of $O_K^\times$, that this smooth compactification is canonical. As all cusps are equivalent (if we vary the lattice $L$ or $\Gamma$), it is enough, as usual, to study the smooth compactification at $c_\infty$. In [Cog] this is described for an arbitrary $L$ (not even $O_K$-free), but for simplicity we write it down only for $L = L_0$.

As $N \geq 3$, elements of $\Gamma$ stabilizing $c_\infty$ lie in $N_\infty$. The computations, which we omit, are somewhat simpler if $N$ is even, an assumption made for the rest of this section. Let

$$\Gamma_{\text{cusp}} = \Gamma \cap N_\infty.$$

**Lemma 1.6.** Let $N \geq 3$ be even. The matrix $n(s,r) \in \Gamma_{\text{cusp}}$ if and only if: (i) $(d_K \equiv 1 \mod 4)$ $s \in NO_K$, $r \in ND_KZ$, (ii) $(d_K \equiv 2, 3 \mod 4)$ $s \in NO_K$ and $r \in 2^{-1}ND_KZ$.

Let $M = N|D_K|$ in case (i) and $M = 2^{-1}N|D_K|$ in case (ii). This is the width of the cusp $c_\infty$. Let

$$q = q(z) = e^{2\pi iz/M}.\tag{1.54}$$

For $R > 0$, the domain $\Omega_R = \{(z,u) \in \mathfrak{X} | \lambda(z,u) > R\}$ is invariant under $\Gamma_{\text{cusp}}$ and if $R$ is large enough, two points of it are $\Gamma$-equivalent if and only if they are $\Gamma_{\text{cusp}}$-equivalent. A sufficiently small punctured neighborhood of $c_\infty$ in $X_\Gamma^*$ therefore looks like $\Gamma_{\text{cusp}} \setminus \Omega_R$. As

$$n(s,r)(z,u) = (z + \delta s(u + s/2) + r, u + s)\tag{1.55}$$

we obtain the following description of $\Gamma_{\text{cusp}} \setminus \Omega_R$. Let $\Lambda = NO_K$ and $E = C/\Lambda$, an elliptic curve with complex multiplication by $O_K$. Let $T$ be the quotient

$$T = (C \times C)/\Lambda\tag{1.56}$$

$\text{No confusion should arise from the use of the letter } N \text{ to denote both the level and the unipotent radical of } P.$
where the action of \( s \in \Lambda \) is via
\[
(1.57) \quad [s] : (t, u) \mapsto (e^{2\pi i \delta \bar{u}} s_t, u + s).
\]
It is a line bundle over \( E \) via the second projection. We denote the class of \((t, u)\)
modulo the action of \( \Lambda \) by \([t, u]\).

**Proposition 1.7.** Let \( T_R \subset T \) be the disk bundle consisting of all the points \([t, u]\)
where
\[
(1.58) \quad |t| < e^{-\pi |\delta|(R + u\bar{u})}/M.
\]
(This condition is invariant under the action of \( \Lambda \).) Let \( T'_{R} \) be the punctured disk
bundle obtained by removing the zero section from \( T_R \). Then the map \((z, u) \mapsto (q(z), u)\)
duces an analytic isomorphism between \( \Gamma_{\text{cusp}} \setminus \Omega_R \) and \( T'_{R} \).

**Proof.** This follows from the discussion so far and the fact that \( \lambda(z, u) > R \) is
equivalent to the above condition on \( t = q(z) \) ([Cog], Prop. 2.1). \( \square \)

To get a smooth compactification \( \bar{X}_\Gamma \) of \( X_\Gamma \) (as a complex surface), we glue the
disk bundle \( T_R \) to \( X_\Gamma \) along \( T'_{R} \). In other words, we complete \( T'_{R} \) by adding the
zero section, which is isomorphic to \( E \). The same procedure should be carried out
at any other cusp of \( C_\Gamma \).

Note that the geodesic (1.15) connecting \((z, u) \in X\) to the cusp \( c_\infty \) projects in
\( X_\Gamma \) to a geodesic which meets \( E \) transversally at the point \( u \mod \Lambda \).
We caution that this geodesic in \( X_\Gamma \) depends on \((z, u)\) and \( c_\infty \) and not only on their images
modulo \( \Gamma \).

The line bundle \( T \) is the inverse of an ample line bundle on \( E \). In fact, \( T^\vee \)
is the \( N \)-th (resp. 2\( N \)-th) power of one of the four basic theta line bundles if \( d_K \equiv 1 \mod 4 \) (resp. \( d_K \equiv 2, 3 \mod 4 \)). A basic theta function of the lattice \( \Lambda \) satisfies,
for \( u \in \mathbb{C} \) and \( s \in \Lambda \),
\[
(1.59) \quad \theta(u + s) = \alpha(s) e^{2\pi i (u + s)/|\delta|N^2} \theta(u)
\]
where \( \alpha : \Lambda \to \pm 1 \) is a quasi-character (see [Mu], p.25). Recalling the relation
between \( M \) and \( N \), and the assumption that \( N \) was even, we easily get the relation
between \( T \) and the theta line bundles.

1.4.2. *The smooth compactification of \( S \).* The arithmetic compactification \( \tilde{S} \) of
the Picard surface \( S \) (over \( R_0 \)) is due to Larsen [La1] (see also [Bel] and [Lan]). We
summarize the results in the following theorem. We mention first that as \( S_C \) has a
canonical model \( S \) over \( R_0 \), its Baily-Borel compactification \( S^{*} \) over \( R_0 \), and \( S \)
embeds in \( S^{*} \) as an open dense subscheme.

**Theorem 1.8.** (i) There exists a projective scheme \( \bar{S} \), smooth over \( R_0 \), of relative
dimension 2, together with an open dense immersion of \( S \) in \( \bar{S} \), and a proper
morphism \( p : \bar{S} \to S^{*} \), making the following diagram commutative
\[
(1.60) \quad \begin{array}{ccc}
S & \to & \bar{S} \\
\downarrow & & \downarrow \\
S^{*} & \to & \bar{S}^{*}
\end{array}
\]

(ii) As a complex manifold, there is an isomorphism
\[
(1.61) \quad \bar{S}_C \simeq \prod_{j=1}^{m} \bar{X}_{\Gamma_j},
\]
extending the isomorphism of \( S_C \) with \( \prod_{i=1}^m X_{\Gamma_i} \).

(iii) Let \( C = p^{-1}(S^* - S) \). Let \( R_N \) be the integral closure of \( R_0 \) in the ray class field \( K_N \) of conductor \( N \) over \( K \). Then the connected components of \( C_{R_N} \) are geometrically irreducible, and are indexed by the cusps of \( S^*_R \) over which they sit. Furthermore, each component \( E \subset C_{R_N} \) is an elliptic curve with complex multiplication by \( \mathcal{O}_K \).

We call \( C \) the cuspidal divisor. If \( c \in S^*_C - S_C \) is a cusp, we denote the complex elliptic curve \( p^{-1}(c) \) by \( E_c \). Bear in mind that while \( E_c \) is in principle defnable over the Hilbert class field \( K_1 \), no canonical model of it over that field is provided by \( \tilde{S} \). On the other hand, \( E_c \) does come with a canonical model over \( K_N \), and even over \( R_N \).

We refer to [La1] and [Bel] for a moduli-theoretic interpretation of \( C \) as a moduli space for semi-abelian schemes with a suitable action of \( \mathcal{O}_K \) and a “level-\( N \) structure”. Unfortunately these references do not give such a moduli interpretation to \( S \). While they do construct a universal semi-abelian scheme over \( \tilde{S} \) (see the next section), the level-\( N \) structure over \( S \) does not extend to a flat level-\( N \) structure over \( S \) in the ordinary sense, and the notion has to be modified over the boundary. What is evidently missing is the construction of a “Tate-Picard” object similar to the “generalized elliptic curve” which was constructed over the complete modular curve by Deligne and Rapoport in [De-Ra].

1.4.3. Change of level. Assume that \( N \geq 3 \) is even, and \( N' = QN \). We then obtain a covering map \( X_{\Gamma(N')} \to X_{\Gamma(N)} \) where by \( \Gamma(N) \) we denote the group previously given above \( \Gamma \). Near any of the cusps, the analytic model allows us to analyze this map locally. Let \( E' \) be an irreducible cuspidal component of \( X_{\Gamma(N')} \) mapping to the irreducible component \( E \) of \( X_{\Gamma(N)} \). The following is a consequence of the discussion in the previous sections.

Proposition 1.9. The map \( E' \to E \) is a multiplication-by-\( Q \) isogeny, hence étale of degree \( Q^2 \). When restricted to a neighborhood of \( E' \), the covering \( X_{\Gamma(N')} \to X_{\Gamma(N)} \) is of degree \( Q^3 \), and has ramification index \( Q \) along \( E \), in the normal direction to \( E \).

Corollary 1.10. The pull-back to \( E' \) of the normal bundle \( T(N) \) of \( E \) is the \( Q \)th power of the normal bundle \( T(N') \) of \( E' \).

1.5. The universal semi-abelian scheme \( \mathcal{A} \).

1.5.1. The universal semi-abelian scheme over \( \tilde{S} \). As Larsen and Bellaiche explain, the universal abelian scheme \( \pi : \mathcal{A} \to S \) extends canonically to a semi-abelian scheme \( \pi : \mathcal{A} \to \tilde{S} \). The polarization \( \lambda \) extends over the boundary \( C = \tilde{S} - S \) to a principal polarization \( \lambda \) of the abelian part of \( \mathcal{A} \). The action \( \iota \) of \( \mathcal{O}_K \) extends to an action on the semi-abelian variety, which necessarily induces separate actions on the toric part and on the abelian part.

Let \( E \) be a connected component of \( C_{R_N} \), mapping (over \( \mathbb{C} \) and under the projection \( p \)) to the cusp \( c \in S^*_C \). Then there exist (1) a principally polarized elliptic curve \( B \) defined over \( R_N \), with complex multiplication by \( \mathcal{O}_K \) and CM type \( \Sigma \), and (2) an ideal \( a \) of \( \mathcal{O}_K \), such that every fiber \( A_x \) of \( \mathcal{A} \) over \( E \) is an \( \mathcal{O}_K \)-group extension of \( B \) by the \( \mathcal{O}_K \)-torus \( a \otimes \mathbb{G}_m \). Both \( B \) (with its polarization) and the ideal class \( [a] \in Cl_K \) are uniquely determined by the cusp \( c \). Only the extension class in the
category of $\mathcal{O}_K$-groups varies as we move along $E$. Note that since the Lie algebra of the torus is of type $(1, 1)$, the Lie algebra of such an extension $A_\mathfrak{a}$ is of type $(2, 1)$, as is the case at an interior point $x \in S$. If we extend scalars to $\mathbb{C}$, the isomorphism type of $B$ is given by another ideal class $[b]$ (i.e. $B(\mathbb{C}) \simeq \mathbb{C}/b$). In this case we say that the cusp $c$ is of type $(a, b)$.

The above discussion defines a homomorphism (of fpf sheaves over $\text{Spec}(R_N)$)

\begin{equation}
E \to \text{Ext}_{\mathcal{O}_K}^1(B, \mathfrak{a} \otimes \mathbb{G}_m).
\end{equation}

As we shall see soon, the $\text{Ext}$ group is represented by an elliptic curve with CM by $\mathcal{O}_K$, defined over $R_N$, and this map is an isogeny.

1.5.2. $\mathcal{O}_K$-semi-abelian schemes of type $(a, b)$. We digress to discuss the moduli space for semi-abelian schemes of the type found above points of $E$. Let $R$ be an $R_0$-algebra, $B$ an elliptic curve over $R$ with complex multiplication by $\mathcal{O}_K$ and CM type $\Sigma$, and $a$ an ideal of $\mathcal{O}_K$. Consider a semi-abelian scheme $\mathcal{G}$ over $R$, endowed with an $\mathcal{O}_K$ action $\iota : \mathcal{O}_K \to \text{End}(\mathcal{G})$, and a short exact sequence

\begin{equation}
0 \to \mathfrak{a} \otimes \mathbb{G}_m \to \mathcal{G} \to B \to 0
\end{equation}

of $\mathcal{O}_K$-group schemes over $R$. We call all this data a semi-abelian scheme of type $(a, B)$ (over $R$). The group classifying such structures is $\text{Ext}_{\mathcal{O}_K}^1(B, \mathfrak{a} \otimes \mathbb{G}_m)$. Any $\chi \in \mathfrak{a}^* = \text{Hom}(\mathfrak{a}, \mathbb{Z})$ defines, by push-out, an extension $\mathcal{G}_\chi$ of $B$ by $\mathbb{G}_m$, hence a point of $B^t = \text{Ext}^1(B, \mathbb{G}_m)$. We therefore get a homomorphism from $\text{Ext}_{\mathcal{O}_K}^1(B, \mathfrak{a} \otimes \mathbb{G}_m)$ to $\text{Hom}(\mathfrak{a}^*, B^t)$. A simple check shows that its image is in $\text{Hom}_{\mathcal{O}_K}(\mathfrak{a}^*, B^t) = \delta_K \mathfrak{a} \otimes_{\mathcal{O}_K} B^t$, and that this construction yields an isomorphism

\begin{equation}
\text{Ext}_{\mathcal{O}_K}^1(B, \mathfrak{a} \otimes \mathbb{G}_m) \simeq \delta_K \mathfrak{a} \otimes_{\mathcal{O}_K} B^t.
\end{equation}

Here we have used the canonical identification $\mathfrak{a}^* = \delta^{-1} \mathfrak{a}^{-1}$ (via the trace pairing). Although $(\delta_K)$ is a principal ideal, so can be ignored, it is better to keep track of its presence. We emphasize that the CM type of $B^t$, with the natural action of $\mathcal{O}_K$ derived from its action on $B$, is $\Sigma$ rather than $\Sigma$.

Thus over $\delta_K \mathfrak{a} \otimes_{\mathcal{O}_K} B^t$ there is a universal semi-abelian scheme $\mathcal{G}(a, B)$ of type $(a, B)$, and any $\mathcal{G}$ as above, over any base $R'/R$, is obtained from $\mathcal{G}(a, B)$ by pullback (specialization) with respect to a unique map $\text{Spec}(R') \to \delta_K \mathfrak{a} \otimes_{\mathcal{O}_K} B^t$.

When $R = \mathbb{C}$, $B \simeq \mathbb{C}/b$ for a unique ideal class $[b]$ (with $\mathcal{O}_K$ acting via $\Sigma$). Then, canonically, $B^t = \mathbb{C}/\delta_K^{-1}b^{-1}$ (with $\mathcal{O}_K$ acting via $\Sigma$). The pairing between the lattices, $b \times \delta_K^{-1}b^{-1} \to \mathbb{Z}$ is $(x, y) \mapsto Tr_{K/Q}(xy)$. Since the $\mathcal{O}_K$ action on $B^t$ is via $\Sigma$,

\begin{equation}
\text{Ext}_{\mathcal{O}_K}^1(\mathbb{C}/b, \mathfrak{a} \otimes \mathbb{G}_m) \simeq \delta_K \mathfrak{a} \otimes_{\mathcal{O}_K} \mathbb{C}/\delta_K^{-1}b^{-1} = \mathbb{C}/\delta_K^{-1}b^{-1}.
\end{equation}

The universal semi-abelian variety $\mathcal{G}(a, B)$ will now be denoted $\mathcal{G}(a, b)$. In 1.6.2 below we give a complex analytic model of this $\mathcal{G}(a, b)$.

1.6. Degeneration of $\mathcal{A}$ along a geodesic connecting to a cusp.

1.6.1. The degeneration to a semi-abelian variety. It is instructive to use the “moving lattice model” to compute the degeneration of the universal abelian scheme along a geodesic, as we approach a cusp. To simplify the computations, assume for the rest of this section, as before, that $N \geq 3$ is even, and that the cusp is the standard cusp at infinity $c = c_\infty$. In this case we have shown that $E_c = \mathbb{C}/\Lambda$, where
\(\Lambda = N\mathcal{O}_K\), and we have given a neighborhood of \(E_c\) in \(\bar{X}_\Gamma\) the structure of a disk bundle in a line bundle \(\mathcal{J}\). See Proposition 1.7.

Consider the geodesic (1.15) connecting \((z, u)\) to \(c_\infty\). Consider the universal abelian scheme in the moving lattice model (cf. (1.27)). Of the three vectors used to span \(L'_z\) over \(\mathcal{O}_K\) in (1.25) the first two do not depend on \(z\). As \(u\) is fixed along the geodesic, they are not changed. The third vector represents a cycle that vanishes at the cusp (together with all its \(\mathcal{O}_K\)-multiples). We conclude that \(A'_z\) degenerates to

\[
\mathbb{C}^3 / \text{Span}_{\mathcal{O}_K}(\mathcal{O}_K) = \left\{ \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & u \end{pmatrix} \right\}.
\]

Making the change of variables \((\zeta'_1, \zeta'_2, \zeta'_3) = (\zeta_1, \zeta_2 + \bar{u}\zeta_1, \zeta_3)\) does not alter the \(\mathcal{O}_K\) action and gives the more symmetric model

\[
\mathcal{G}_u = \mathbb{C}^3 / \text{Span}_{\mathcal{O}_K}(\mathcal{O}_K) = \left\{ \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & \bar{u} \\ 1 & u & u \end{pmatrix} \right\}
\]

(by note that \(\zeta'_3\), unlike \(\zeta_2\), does not vary holomorphically in the family \(\{\mathcal{G}_u\}\), only in each fiber individually).

Let \(e(x) = e^{2\pi i x} : \mathbb{C} \to \mathbb{C}^\times\) be the exponential map, with kernel \(\mathbb{Z}\). For any ideal \(a\) of \(\mathcal{O}_K\) it induces a map

\[
e_a : a \otimes \mathbb{C} \to a \otimes \mathbb{C}^\times
\]

with kernel \(a \otimes 1\). As usual we identify \(a \otimes \mathbb{C}\) with \(\mathbb{C}(\mathcal{A}) \otimes \mathbb{C}(\mathcal{O}_K)\), sending \(a \otimes \zeta \mapsto (a\zeta, a\zeta)\). We now note that if we use this identification to identify \(\mathbb{C}^3\) with \(\mathbb{C} \oplus (\mathcal{O}_K \otimes \mathbb{C})\) (an identification which is compatible with the \(\mathcal{O}_K\) action) then the \(\iota'\mathcal{(O)_K}\)-span of the vector \(\iota'(0, 1, 1)\) is just the kernel of \(e_{\mathcal{O}_K}\). We conclude that

\[
\mathcal{G}_u \simeq \{ \mathbb{C} \oplus (\mathcal{O}_K \otimes \mathbb{C}^\times) \} / L_u
\]

where \(L_u\) is the sub-\(\mathcal{O}_K\)-module

\[
L_u = \{ (s, e_{\mathcal{O}_K}(s\bar{u}, s\bar{u})) | s \in \mathcal{O}_K \}.
\]

This clearly gives \(\mathcal{G}_u\) the structure of an \(\mathcal{O}_K\)-semi-abelian variety of type \((\mathcal{O}_K, \mathcal{O}_K)\), i.e. an extension

\[
0 \to \mathcal{O}_K \otimes \mathbb{C}^\times \to \mathcal{G}_u \to \mathbb{C} / \mathcal{O}_K \to 0.
\]

1.6.2. The analytic uniformization of the universal semi-abelian variety of type \((a, b)\). We now compare the description that we have found for the degeneration of \(\mathcal{A}\) along the geodesic connecting \((z, u)\) to \(c_\infty\) with the analytic description of the universal semi-abelian variety of type \((a, b)\).

**Proposition 1.11.** Let \(a\) and \(b\) be two ideals of \(\mathcal{O}_K\). For \(u \in \mathcal{C}\) consider

\[
\mathcal{G}_u \simeq \{ \mathbb{C} \oplus (a \otimes \mathbb{C}^\times) \} / L_u
\]

where

\[
L_u = \{ (s, e_{a}(s\bar{u}, s\bar{u})) | s \in b \}.
\]

Then \(\mathcal{G}_u\) is a semi-abelian variety of type \((a, b)\), any complex semi-abelian variety of this type is a \(\mathcal{G}_u\), and \(\mathcal{G}_u \simeq \mathcal{G}_v\) if and only if \(u - v \in \mathfrak{m}^{-1}\).
Proof. That $G_u$ is a semi-abelian variety of type $(a, b)$ is obvious. That any abelian variety of this type is a $G_u$ follows by passing to the universal cover $\mathbb{C}^2(\Sigma) \oplus \mathbb{C}(\Sigma)$, and noting that by a change of variables in the $\Sigma$- and $\overline{\Sigma}$-isotypical parts, we may assume that the lattice by which we divide is of the form

$$
\begin{pmatrix}
  a & 0 \\
  1 & 1 \\
  1 & u
\end{pmatrix}
\oplus
\begin{pmatrix}
  1 \\
  \bar{u} \\
  u
\end{pmatrix}.
$$

Finally, the map $u \mapsto [G_u]$ is a homomorphism $\mathbb{C} \to \text{Ext}_1^{\mathcal{O}_K}(\mathbb{C}/\mathcal{O}_K, \mathcal{O}_K \otimes \mathbb{C}^\times)$, so we only have to prove that $G_u$ is split if and only if $u \in a b^{-1}$. But one can check easily that $G_u$ is trivial if and only if $(s \bar{u}, \bar{s}u) \in \ker e_a = a \otimes 1 = \{(a, \bar{a})| a \in a\}$ for every $s \in b$, and this holds if and only if $u \in a b^{-1}$. □

**Corollary 1.12.** Let $N \geq 3$ be even. Let $c = c_\infty$ be the cusp at infinity. Then the map

$$
E_c \to \text{Ext}_1^{\mathcal{O}_K}(\mathbb{C}/\mathcal{O}_K, \mathcal{O}_K \otimes \mathbb{C}^\times)
$$

sending $u$ to the isomorphism class of the semi-abelian variety above $u \mod \Lambda$ is the isogeny of multiplication by $N$.

**Proof.** In view of the computations above, and the description of a neighborhood of $E_c$ in $\bar{X}_\Gamma$ given in Proposition 1.7 this map is identified with the canonical map

$$
\mathbb{C}/N\mathcal{O}_K \to \mathbb{C}/\mathcal{O}_K.
$$

□

The extra data carried by $u \in E_c$, which is forgotten by the map of the corollary, comes from the level $N$ structure. As mentioned before, according to [La1] and [Bel] the cuspidal divisor $C$ has a modular interpretation as the moduli space for semi-abelian schemes of the type considered above, together with level-$N$ structure ($\mathcal{M}_{\infty,N}$ structures in the language of [Bel]). A level-$N$ structure on a semi-abelian variety $G$ of type $(a, b)$ consists of (i) a level-$N$ structures $\alpha: N^{-1}\mathcal{O}_K/\mathcal{O}_K \simeq a \otimes \mu_N$ on the toric part (ii) a level-$N$ structure $\beta: N^{-1}\mathcal{O}_K/\mathcal{O}_K \simeq N^{-1}b/b = B[N]$ on the abelian part (iii) an $\mathcal{O}_K$-splitting $\gamma$ of the map $G[N] \to B[N]$.

Over $c = c_\infty$, when $a = b = \mathcal{O}_K$, there are obvious natural choices for $\alpha$ and $\beta$ (independent of $u$) but the splittings $\gamma$ in (iii) form a torsor under $\mathcal{O}_K/N\mathcal{O}_K$. If we consider the splitting

$$
\gamma_u: N^{-1}\mathcal{O}_K/\mathcal{O}_K \ni s \mapsto (s, e\mathcal{O}_K(s\bar{u}, \bar{s}u)) \mod L_u
$$

then the tuples $(G_u, \alpha, \beta, \gamma_u)$ and $(G_v, \alpha, \beta, \gamma_v)$ are isomorphic if and only if $u \equiv v \mod N\mathcal{O}_K$, i.e. if and only if $u$ and $v$ represent the same point of $E_c$.

2. **The basic automorphic vector bundles**

2.1. **The vector bundles $\mathcal{P}$ and $\mathcal{L}$**.
2.1.1. **Definition and first properties.** Recall our running assumptions and notation. The tame level $N \geq 3$, $S$ is the Picard modular scheme over the base ring $R_0$, $\overline{S}$ is its smooth compactification, and $\mathcal{A}$ is the universal semi-abelian scheme over $\overline{S}$ (an abelian scheme over $S$) constructed by Larsen and Bellaiche.

Let $\omega_\mathcal{A}$ be the relative cotangent space at the origin of $\mathcal{A}$. If $e : \overline{S} \rightarrow \mathcal{A}$ is the zero section,
\begin{equation}
\omega_\mathcal{A} = e^*(\Omega^1_{\mathcal{A}/\overline{S}}).
\end{equation}
This is a rank 3 vector bundle over $\overline{S}$ and the action of $\mathcal{O}_\mathcal{K}$ allows to decompose it according to types. We let
\begin{equation}
\mathcal{P} = \omega_\mathcal{A}(\Sigma), \quad \mathcal{L} = \omega_\mathcal{A}(\overline{\Sigma}).
\end{equation}
Then $\mathcal{P}$ is a plane bundle, and $\mathcal{L}$ a line bundle.

Over $S$ (but not over the cuspidal divisor $C = \overline{S} - S$) we have the usual identification $\omega_\mathcal{A} = \pi_1*\Omega^1_{\mathcal{A}/S}$. The relative de Rham cohomology of $\mathcal{A}/S$ is a rank 6 vector bundle sitting in an exact sequence (the Hodge filtration)
\begin{equation}
0 \rightarrow \omega_\mathcal{A} \rightarrow H^1_{dR}(\mathcal{A}/S) \rightarrow R^1\pi_*\mathcal{O}_\mathcal{A} \rightarrow 0.
\end{equation}
Since, for any abelian scheme, $R^1\pi_*\mathcal{O}_\mathcal{A} = \omega^{\vee}_{\mathcal{A}}$ (canonical isomorphism, see [Mu]), and $\lambda : \mathcal{A} \rightarrow \mathcal{A}^\vee$ is an isomorphism which reverses CM types, we obtain an exact sequence
\begin{equation}
0 \rightarrow \omega_\mathcal{A} \rightarrow H^1_{dR}(\mathcal{A}/S) \rightarrow \omega^{\vee}_{\mathcal{A}}(\rho) \rightarrow 0.
\end{equation}
The notation $\mathcal{M}(\rho)$ means that $\mathcal{M}$ is a vector bundle with an $\mathcal{O}_\mathcal{K}$ action and in $\mathcal{M}(\rho)$ the vector bundle structure is that of $\mathcal{M}$ but the $\mathcal{O}_\mathcal{K}$ action is conjugated. Decomposing according to types, we have two short exact sequences
\begin{equation}
0 \rightarrow \mathcal{P} \rightarrow H^1_{dR}(\mathcal{A}/S)(\Sigma) \rightarrow \mathcal{L}^{\vee}(\rho) \rightarrow 0
\end{equation}
\begin{equation}
0 \rightarrow \mathcal{L} \rightarrow H^1_{dR}(\mathcal{A}/S)(\overline{\Sigma}) \rightarrow \mathcal{P}^{\vee}(\rho) \rightarrow 0.
\end{equation}
The pairing $\langle \cdot, \cdot \rangle_\lambda$ on $H^1_{dR}(\mathcal{A}/S)$ induced by the polarization is $\mathcal{O}_S$-linear, alternating, perfect, and satisfies $\langle \lambda(a)x, y \rangle = \langle x, \lambda(a)y \rangle_\lambda$. It follows that $H^1_{dR}(\mathcal{A}/S)(\Sigma)$ and $H^1_{dR}(\mathcal{A}/S)(\overline{\Sigma})$ are maximal isotropic subspaces, and are set in duality. As $\omega_\mathcal{A}$ is also isotropic, this pairing induces pairings
\begin{equation}
\mathcal{P} \times \mathcal{P}^{\vee}(\rho) \rightarrow \mathcal{O}_S, \quad \mathcal{L} \times \mathcal{L}^{\vee}(\rho) \rightarrow \mathcal{O}_S.
\end{equation}
These two pairings (in this order) are the tautological pairings between a vector bundle and its dual.

Another consequence of this discussion that we wish to record is the canonical isomorphism over $S$
\begin{equation}
\det \mathcal{P} = \mathcal{L}(\rho) \otimes \det \left( H^1_{dR}(\mathcal{A}/S)(\Sigma) \right).
\end{equation}

2.1.2. **The factors of automorphy corresponding to $\mathcal{L}$ and $\mathcal{P}$.** The formulae below can be deduced also from the matrix calculations in the first few pages of [Sh2]. Let $\Gamma = \Gamma_j$ be one of the groups used in the complex uniformization of $\mathcal{S}_\mathcal{C}$, cf Section 13.3.5. Via the analytic isomorphism $X_\Gamma \simeq \mathcal{S}_\mathcal{C}$ with the $j$th connected component, the vector bundles $\mathcal{P}$ and $\mathcal{L}$ are pulled back to $X_\Gamma$ and then to the symmetric space $\mathfrak{X}$, where they can be trivialized, hence described by means of factors of automorphy. Let us denote by $\mathcal{P}_{an}$ and $\mathcal{L}_{an}$ the two vector bundles on $X_\Gamma$, in the complex analytic category, or their pull-backs to $\mathfrak{X}$. 
To trivialize $\mathcal{L}_\text{an}$ we must choose a nowhere vanishing global section over $\mathfrak{X}$. As usual, we describe it only on the connected component containing the standard cusp, corresponding to $j = 1$ (where $L = L_{a_1} = L_0$). Recalling the “moving lattice model” (1.27) and the coordinates $\zeta_1, \zeta_2, \zeta_3$ introduced there, we note that $d\zeta_3$ is a generator of $\mathcal{L}_\text{an} = \omega_{\mathcal{A}(\Sigma)}$. For reasons that will become clear later (cf Section 2.6) we use $2\pi i \cdot d\zeta_3$ to trivialize $\mathcal{L}_\text{an}$ over $\mathfrak{X}$. Suppose

$$\gamma = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \in \Gamma \subset SU_\infty.$$  

If $\gamma(z, u) = (z', u')$ then

$$z' = \frac{a_1 z + b_1 u + c_1}{a_3 z + b_3 u + c_3}, \quad u' = \frac{a_2 z + b_2 u + c_2}{a_3 z + b_3 u + c_3}$$

and

$$\gamma \begin{pmatrix} z \\ u \\ 1 \end{pmatrix} = j(\gamma; z, u) \begin{pmatrix} z' \\ u' \\ 1 \end{pmatrix}, \quad j(\gamma; z, u) = a_3 z + b_3 u + c_3.$$

**Lemma 2.1.** The following relation holds for every $\gamma \in U_\infty$

$$\lambda(z, u) = \lambda(\gamma(z, u)) \cdot |j(\gamma; z, u)|^2.$$ 

**Proof.** Let $v = v(z, u) = i(z, u, 1)$. Then

$$\lambda(z, u) = -(v, v).$$

As $\varepsilon(\gamma(z, u)) = j(\gamma; z, u)^{-1} \cdot \varepsilon(v(z, u))$ the lemma follows from $(\gamma v, \gamma v) = (v, v)$. □

Let $\mathcal{V} = \text{Lie}(\mathcal{A}/\mathfrak{X}) = \omega_{\mathcal{A}/\mathfrak{X}}^\vee$ and $\mathcal{W} = \mathcal{V}(\Sigma) = \mathcal{L}_\text{an}^\vee$ (a line bundle). At a point $x = (z, u) \in \mathfrak{X}$ the fiber $\mathcal{V}_x$ is identified canonically with $(\mathcal{W}_x, J_x)$ and then $\mathcal{W}_x = \mathcal{W}_x = \mathbb{C} \cdot i(z, u, 1)$.

**Proposition 2.2.** For $x = (z, u) \in \mathfrak{X}$ let

$$v_3(z, u) = \lambda(z, u)^{-1} \begin{pmatrix} z \\ u \\ 1 \end{pmatrix} \in \mathcal{W}_x.$$ 

Then (i) $v_3(z, u)$ is a nowhere vanishing holomorphic section of $\mathcal{W}$, (ii) $(d\zeta_3, v_3) \equiv 1$, (iii) the automorphy factor corresponding to $d\zeta_3$ is the function $j(\gamma; z, u)$.

**Proof.** Since, by construction, $d\zeta_3$ is a nowhere vanishing holomorphic section of $\mathcal{L}$ (over $\mathfrak{X}$), (i) follows from (ii). To prove (ii) we transfer $v_3(z, u)$ to the moving lattice model and get $i(0, 0, 1)$, which is the dual vector to $d\zeta_3$. To prove (iii) we compute in $V_\mathbb{R}$ (with the original complex structure!)

$$\frac{\gamma_* v_3(z, u)}{v_3(\gamma(z, u))} = \frac{\lambda(\gamma(z, u))}{\lambda(z, u)} j(\gamma; z, u) = j(\gamma; z, u)^{-1},$$

and recall that since $W_{\gamma(z, u)}$ is precisely the line where the complex structure in $(V_\mathbb{R}, J_{\gamma(z, u)})$ has been reversed, in $(V_\mathbb{R}, J_{\gamma(z, u)})$ we have

$$\frac{\gamma_* v_3(z, u)}{v_3(\gamma(z, u))} = j(\gamma; z, u)^{-1}.$$
Dualizing, we get \((x = (z, u))\)

\[
\frac{(\gamma^{-1})^* d\zeta_3|_{\gamma(x)}}{d\zeta_3|_{\gamma(x)}} = j(\gamma, x).
\]

This concludes the proof. \(\square\)

Consider next the plane bundle \(P_{an}\). As we will only be interested in scalar-valued modular forms, we do not compute its matrix-valued factor of automorphy (but see [Sh2]). It is important to know, however, that the line bundle \(\det P_{an}\) gives nothing new.

**Proposition 2.3.** There is an isomorphism of analytic line bundles over \(X_{\Gamma}\), \(\det P_{an} \simeq L_{an}\).

Moreover, \(d\zeta_1 \wedge d\zeta_2\) is a nowhere vanishing holomorphic section of \(\det P_{an}\) over \(X\), and the factor of automorphy corresponding to it is \(j(\gamma; z, u)\).

**Proof.** Since a holomorphic line bundle on \(X_{\Gamma} = \Gamma \backslash X\) is determined, up to an isomorphism, by its factor of automorphy, and \(j(\gamma; z, u)\) is the factor of automorphy of \(L_{an}\) corresponding to \(d\zeta_3\), it is enough to prove the second statement. Let \(U = V(\Sigma)\) be the plane bundle dual to \(P_{an}\). Let

\[
v_1(z, u) = -\lambda(z, u)^{-1} \begin{pmatrix} \bar{u}z \\ (z - \bar{z})/\delta \\ \bar{u} \end{pmatrix}
\]

and

\[
v_2(z, u) = -\lambda(z, u)^{-1} \begin{pmatrix} \bar{z} + \delta u\bar{u} \\ u \\ 1 \end{pmatrix}
\]

(considered as vectors in \((V_\mathbb{R}, J_x) = V_x\)). As we have seen in (1.27), these two vector fields are sections of \(U\) and at each point \(x \in X\) form a basis dual to \(d\zeta_1 \wedge d\zeta_2\). It follows that they are holomorphic sections, and that \(v_1 \wedge v_2\) is the basis dual to \(d\zeta_1 \wedge d\zeta_2\). We must show that the factor of automorphy corresponding to \(v_1 \wedge v_2\) is \(j(\gamma; z, u)^{-1}\), i.e. that

\[
\frac{\gamma_*(v_1 \wedge v_2(z, u))}{v_1 \wedge v_2(\gamma(z, u))} = j(\gamma; z, u)^{-1}.
\]

Working in \(V_\mathbb{R} = \mathbb{C}^3\) (with the original complex structure)

\[
\frac{\gamma_*(v_1 \wedge v_2(z, u))}{v_1 \wedge v_2(\gamma(z, u))} \cdot \frac{1}{J(\gamma; z, u)} = \frac{\gamma_*(v_1 \wedge v_2 \wedge v_3(z, u))}{v_1 \wedge v_2 \wedge v_3(\gamma(z, u))}.
\]

But

\[
v_1 \wedge v_2 \wedge v_3(z, u) = \delta \lambda(z, u)^{-1} e_1 \wedge e_2 \wedge e_3,
\]

because

\[
det \begin{pmatrix} \bar{u}z \\ (z - \bar{z})/\delta \\ \bar{u} \end{pmatrix} \begin{pmatrix} \bar{z} + \delta u\bar{u} \\ u \\ 1 \end{pmatrix} = \delta \lambda(z, u)^2.
\]
As $\det(\gamma) = 1$, this gives
\[
\frac{\gamma_s(\nu_1 \wedge \nu_2(z, u))}{\nu_1 \wedge \nu_2(\gamma(z, u))} \cdot \frac{1}{j(\gamma; z, u)} = \frac{\lambda(\gamma(z, u))}{\lambda(z, u)} \cdot \frac{1}{j(\gamma; z, u)},
\]
and the proof is complete. \hfill \Box

2.1.3. The relation $\det P \simeq L$ over $\tilde{S}_K$. The isomorphism between $\det P$ and $L$ is in fact algebraic, and even extends to the generic fiber $\tilde{S}_K$ of the smooth compactification.

**Proposition 2.4.** One has $\det P \simeq L$ over $\tilde{S}_K$.

**Proof.** Since $\text{Pic}(\tilde{S}_K) \subset \text{Pic}(\tilde{S}_C)$ it is enough to prove the proposition over $\mathbb{C}$. By GAGA, it is enough to establish the triviality of $\det P \otimes L^{-1}$ in the analytic category. For each connected component $X_\Gamma$ of $S_C$, the section $(d\zeta_1 \wedge d\zeta_2) \otimes d\zeta_3^{-1}$ descends from $X$ to $X_\Gamma$, because $d\zeta_1 \wedge d\zeta_2$ and $d\zeta_3$ have the same factor of automorphy $j(\gamma, x)$ ($\gamma \in \Gamma, x \in X$). This section is nowhere vanishing on $X_\Gamma$, and extends to a nowhere vanishing section on $\tilde{X}_\Gamma$, trivializing $\det P \otimes L^{-1}$. In fact, if $c$ is the standard cuspidal, $d\zeta_1 \wedge d\zeta_2$ and $d\zeta_3$ are already well-defined and nowhere vanishing sections of $\det P$ and $L$ in the neighborhood
\[
(\Gamma_{\text{cusp}} \setminus \Omega_R) = (\Gamma_{\text{cusp}} \setminus \Omega_R) \cup E_c
\]
of $E_c$ (see [1.4.1]). This is a consequence of the fact that $j(\gamma, x) = 1$ for $\gamma \in \Gamma_{\text{cusp}}$.

An alternative proof is to use Theorem 4.8 of [Ha]. In our case it gives a functor $V \mapsto [V]$ from the category of $G(\mathbb{C})$-equivariant vector bundles on the compact dual $\mathbb{P}^2$ of $Sh_K$ to the category of vector bundles with $G(A_f)$-action on the inverse system of Shimura varieties $Sh_K$. Here $\mathbb{P}^2 = G(\mathbb{C})/H(\mathbb{C})$, where $H(\mathbb{C})$ is the parabolic group stabilizing the line $\mathbb{C}$. \textbf{[δ/2, 0, 1]} in $G(\mathbb{C}) = GL_3(\mathbb{C}) \times \mathbb{C}^\times$, and the irreducible $V$ are associated with highest weight representations of the Levi factor $L(\mathbb{C})$ of $H(\mathbb{C})$. It is straightforward to check that $\det P$ and $L$ are associated with the same character of $L(\mathbb{C})$, up to a twist by a character of $G(\mathbb{C})$, which affects the $G(A_f)$-action (hence the normalization of Hecke operators), but not the structure of the line bundles themselves. The functoriality of Harris’ construction implies that $\det P$ and $L$ are isomorphic also algebraically. \hfill \Box

We do not know if $\det P$ and $L$ are isomorphic as algebraic line bundles over $S$. This would be equivalent, by (2.7), to the statement that for every PEL structure $(A, \lambda, \iota, c) \in M(R)$, for any $R_0$-algebra $R$, $\det(H_{\text{dR}}^1(A/R)(\Sigma))$ is the trivial line bundle on $\text{Spec}(R)$. To our regret, we have not been able to establish this, although a similar statement in the “Siegel case”, namely that for any principally polarized abelian scheme $(A, \lambda)$ over $R$, $\det H_{\text{dR}}^1(A/R)$ is trivial, follows at once from the Hodge filtration (2.4). Our result, however, suffices to guarantee the following corollary, which is all that we will be using in the sequel.

**Corollary 2.5.** For any characteristic $p$ geometric point $\text{Spec}(k) \to \text{Spec}(R_0)$, we have $\det P \simeq L$ on $\tilde{S}_k$. A similar statement holds for morphisms $\text{Spec}W(k) \to \text{Spec}(R_0)$.

**Proof.** Since $\tilde{S}$ is a regular scheme, $\det P \otimes L^{-1} \simeq \mathcal{O}(D)$ for a Weil divisor $D$ supported on vertical fibers over $R_0$. Since any connected component $Z$ of $\tilde{S}_k$ is irreducible, we can modify $D$ so that $D$ and $Z$ are disjoint, showing that $\det P \otimes L^{-1}|_{Z}$ is trivial. The second claim is proved similarly. \hfill \Box
2.1.4. Modular forms. Let $R$ be an $R_0$-algebra. A modular form of weight $k \geq 0$ and level $N \geq 3$ defined over $R$ is an element of the finite $R$-module

$$M_k(N, R) = H^0(\bar{S}_R, \mathcal{L}^k).$$

(2.26)

We usually omit the subscript $R$, remembering that $\bar{S}$ is now to be considered over $R$. The well-known Koecher principle says that $H^0(\bar{S}, \mathcal{L}^k) = H^0(S, \mathcal{L}^k)$. See [Bel], Section 2.2, for an arithmetic proof that works integrally over any $R_0$-algebra $R$.

A cusp form is an element of the space

$$M^0_k(N, R) = H^0(\bar{S}, \mathcal{L}^k \otimes \mathcal{O}(C)^\vee).$$

(2.27)

As we shall see below (cf Corollary 2.10), if $k \geq 3$, there is an isomorphism $\mathcal{L}^k \otimes \mathcal{O}(C)^\vee \simeq \Omega^k_S \otimes \mathcal{L}^{k-3}$. In particular, cusp forms of weight 3 are “the same” as holomorphic 2-forms on $\bar{S}$.

An alternative definition (à la Katz) of a modular form of weight $k$ and level $N$ defined over $R$, is as a “rule” $f$ which assigns to every $R$-scheme $T$, and every $A = (A, \lambda, t, \alpha) \in \mathcal{M}(T)$, together with a nowhere vanishing section $\omega \in H^0(T, \omega_{A/T}(\bar{S}))$, an element $f(A, \omega) \in H^0(T, \mathcal{O}_T)$ satisfying

- $f(A, \lambda \omega) = \lambda^{-k} f(A, \omega)$ for every $\lambda \in H^0(T, \mathcal{O}_T)^\times$.
- The “rule” $f$ is compatible with base change $T' \to T$.

Indeed, if $f$ is an element of $M_k(N, R)$, then given such an $A$ and $\omega$, the universal property of $S$ produces a unique morphism $\varphi : T \to S$ over $R$, $\varphi^* A = A$, and we may let $f(A, \omega) = \varphi^* f/\omega$. Conversely, given such a rule $f$ we may cover $S$ by Zariski open sets $T$ where $\mathcal{L}$ is trivialized, and then the sections $f(A_T, \omega_T)\omega_T^k$ (a trivializing section over $T$) glue to give $f \in M_k(N, R)$. While viewing $f$ as a “rule” rather than a section is a matter of language, it is sometimes more convenient to use this language.

Let $R \to R'$ be a homomorphism of $R_0$-algebras. Then Bellaiche proved the following theorem ([Bel], 1.1.5).

**Theorem 2.6.** If $k \geq 3$ (resp. $k \geq 6$) then $M^0_k(N, R)$ (resp. $M_k(N, R)$) is a locally free finite $R$-module, and the base-change homomorphism

$$R' \otimes M^0_k(N, R) \simeq M^0_k(N, R')$$

(2.28)

is an isomorphism (resp. base change for $M_k(N, R)$).

Bellaiche considers only weights divisible by 3, but his proofs generalize to all $k$ (cf remark on the bottom of p.43 in [Bel]).

Over $\mathbb{C}$, pulling back to $\mathfrak{X}$ and using the trivialization of $\mathcal{L}$ given by the nowhere vanishing section $2\pi i \cdot d\zeta_3$, a modular form of weight $k$ is a collection $(f_j)_{1 \leq j \leq m}$ of holomorphic functions on $\mathfrak{X}$ satisfying

$$f_j(\gamma(z, u)) = j(\gamma; z, u)^k f_j(z, u) \quad \forall \gamma \in \Gamma_j$$

(2.29)

(the Koecher principle means that no condition has to be imposed at the cusps).

2.2. The Kodaira Spencer isomorphism. Let $\pi : A \to S$ be an abelian scheme of relative dimension 3, as in the Picard moduli problem. The Gauss-Manin connection

$$\nabla : H^1_{dR}(A/S) \to H^1_{dR}(A/S) \otimes \mathcal{O}_S \Omega^1_S$$

(2.30)

defines the Kodaira-Spencer map

$$KS \in Hom_{\mathcal{O}_S}(\omega_A \otimes \mathcal{O}_S \omega_{A^1}, \Omega^1_S)$$

(2.31)
as the composition of the maps
\[
\omega_A = H^0(A, \Omega^1_{A/S}) \to H^1_{dR}(A/S) \xrightarrow{\nabla} H^1_{dR}(A/S) \otimes_{\mathcal{O}_S} \Omega^1_S
\]
(2.32) 
\[\to \mathcal{O}^*_A \otimes_{\mathcal{O}_S} \Omega^1_S,\]
and finally using \(\text{Hom}(L, M^\vee \otimes N) = \text{Hom}(L \otimes M, N)\). Recall that if \(A\) is endowed with an \(\mathcal{O}_K\) action via \(\iota\) then the induced action of \(a \in \mathcal{O}_K\) on \(A^t\) is induced from the action on \(\text{Pic}(A)\), taking a line bundle \(M\) to \(\iota(a)^* M\). As the polarization \(\lambda: A \to A^t\) is \(\mathcal{O}_S\)-linear but satisfies \(\lambda \circ \iota(a) = \iota(a^\rho) \circ \lambda\), it follows that the induced \(\mathcal{O}_K\) action on \(A^t\) is of type \((1, 2)\), hence \(\omega^\vee_A\) is of type \((1, 2)\).

**Lemma 2.7.** The map \(KS\) induces maps
\[
KS(\Sigma) : \omega_A(\Sigma) \to \omega^\vee_A(\Sigma) \otimes_{\mathcal{O}_S} \Omega^1_S
\]
(2.33)
\[
KS(\bar{\Sigma}) : \omega_A(\bar{\Sigma}) \to \omega^\vee_A(\bar{\Sigma}) \otimes_{\mathcal{O}_S} \Omega^1_S,
\]
hence maps, denoted by the same symbols,
\[
KS(\Sigma) : \omega_A(\Sigma) \otimes_{\mathcal{O}_S} \omega_A^t(\Sigma) \to \Omega^1_S
\]
(2.34)
\[
KS(\bar{\Sigma}) : \omega_A(\bar{\Sigma}) \otimes_{\mathcal{O}_S} \omega_A^t(\bar{\Sigma}) \to \Omega^1_S.
\]
The CM-type-reversing isomorphism \(\lambda^* : \omega_A^t \to \omega_A\) induced by the principal polarization satisfies
\[
KS(\Sigma)(\lambda^* x \otimes y) = KS(\bar{\Sigma})(\lambda^* y \otimes x)
\]
for all \(x \in \omega_A^t(\bar{\Sigma})\) and \(y \in \omega_A^t(\Sigma)\).

**Proof.** The first claim follows from the fact that the Gauss-Manin connection commutes with the endomorphisms, hence preserves CM types. The second claim is a consequence of the symmetry of the polarization, see [Fa-Ch], Prop. 9.1 on p.81 (in the Siegel modular case). \(\square\)

Observe that \(\omega_A(\Sigma) \otimes_{\mathcal{O}_S} \omega_A^t(\Sigma)\), as well as \(\omega_A(\bar{\Sigma}) \otimes_{\mathcal{O}_S} \omega_A^t(\bar{\Sigma})\), are vector bundles of rank 2.

**Lemma 2.8.** If \(S\) is the Picard modular surface and \(A = A^t\) is the universal abelian variety, then
\[
KS(\Sigma) : \omega_A(\Sigma) \otimes_{\mathcal{O}_S} \omega_A^t(\Sigma) \to \Omega^1_S
\]
is an isomorphism, and so is \(KS(\bar{\Sigma})\).

**Proof.** This is well-known and follows from deformation theory. For a self-contained proof, see [Bel], Prop. II.2.1.5. \(\square\)

**Proposition 2.9.** The Kodaira-Spencer map induces a canonical isomorphism of vector bundles over \(S\)
\[
\mathcal{P} \otimes \mathcal{L} \simeq \Omega^1_S.
\]
(2.37)

**Proof.** We need only use \(\lambda^*\) to identify \(\omega_A^t(\Sigma)\) with \(\omega_A(\Sigma)\). \(\square\)

We refer to Corollary 2.14 for an extension of this result to \(\bar{\Sigma}\).

**Corollary 2.10.** There is an isomorphism of line bundles \(\mathcal{L}^3 \simeq \Omega^2_S\).

**Proof.** Take determinants and use \(\det \mathcal{P} \simeq \mathcal{L}\). We emphasize that while \(KS(\Sigma)\) is canonical, the identification of \(\det \mathcal{P}\) with \(\mathcal{L}\) depends on a choice, which we shall fix later on once and for all. \(\square\)
The last corollary should be compared to the case of the open modular curve $Y(N)$, where the square of the Hodge bundle $\omega_{E}$ of the universal elliptic curve, becomes isomorphic to $\Omega_{Y(N)}^{1}$. Over $\mathbb{C}$, as the isomorphism between $\mathcal{L}^{3}$ and $\Omega_{\tilde{S}}^{2}$ takes $d\zeta_{3}$ to a constant multiple of $dz \wedge du$ (see Corollary 2.18), the differential form corresponding to a modular form $(f_{j})_{1 \leq j \leq m}$ of weight 3, is (up to a constant) $(f_{j}(z, u)dz \wedge du)_{1 \leq j \leq m}$.

### 2.3. Extensions to the boundary of $S$.

#### 2.3.1. The vector bundles $\mathcal{P}$ and $\mathcal{L}$ over $C$.

Let $E \subset C_{R_{N}}$ be a connected component of the cuspidal divisor (over the integral closure $R_{N}$ of $R_{0}$ in the ray class field $K_{N}$). As we have seen, $E$ is an elliptic curve with CM by $O_{K}$. If the cusp at which $E$ sits is of type $(a, B)$ $(a$ an ideal of $O_{K}$, $B$ an elliptic curve with CM by $O_{K}$ defined over $R_{N}$) then $E$ maps via an isogeny to $\delta_{K}a \otimes O_{K} B^{t} = \text{Ext}_{O_{K}}^{1}(B, a \otimes \mathbb{G}_{m})$. In particular, $E$ and $B$ are isogenous over $K_{N}$.

Consider $\mathcal{G}$, the universal semi-abelian $O_{K}$-abelian three-fold of type $(a, B)$, over $\delta_{K}a \otimes O_{K} B^{t}$. The semi-abelian scheme $A$ over $E$ is the pull-back of this $\mathcal{G}$. Clearly, $\omega_{A/E} = \mathcal{P} \oplus \mathcal{L}$ and $\mathcal{P} = \omega_{A/E}(\Sigma)$ admits over $E$ a canonical rank 1 sub-bundle $\mathcal{P}_{0} = \omega_{B}$. Since the toric part and the abelian part of $\mathcal{G}$ are constant, $\mathcal{L}, \mathcal{P}_{0}$ and $\mathcal{P}_{\mu} = \mathcal{P}/\mathcal{P}_{0}$ are all trivial line bundles when restricted to $E$. It can be shown that $\mathcal{P}$ itself is not trivial over $E$.

#### 2.3.2. More identities over $\tilde{S}$.

We have seen that $\Omega_{\tilde{S}}^{2} \simeq \mathcal{L}^{3}$. For the following proposition, compare [Bel], Lemme II.2.1.7.

**Proposition 2.11.** Working over $K_{N}$, let $E_{j}$ ($1 \leq j \leq h$) be the connected components of $C$. Then

\begin{equation}
\Omega_{\tilde{S}}^{2} \simeq \mathcal{L}^{3} \otimes \bigotimes_{j=1}^{h} O(E_{j})^{\vee}.
\end{equation}

**Proof.** By [Hart] II.6.5, $\Omega_{\tilde{S}}^{2} \simeq \mathcal{L}^{3} \otimes \bigotimes_{j=1}^{h} O(E_{j})^{n_{j}}$ for some integers $n_{j}$ and we want to show that $n_{j} = -1$ for all $j$. By the adjunction formula on the smooth surface $\tilde{S}$, if we denote by $K_{\tilde{S}}$ a canonical divisor, $O(K_{\tilde{S}}) = \Omega_{\tilde{S}}^{2}$, then

\begin{equation}
0 = 2g_{E_{j}} - 2 = E_{j}(E_{j} + K_{\tilde{S}}).
\end{equation}

We conclude that

\begin{equation}
\deg(\Omega_{\tilde{S}}^{2}|_{E_{j}}) = E_{j}.K_{\tilde{S}} = -E_{j}.E_{j} > 0.
\end{equation}

Here $E_{j}.E_{j} < 0$ because $E_{j}$ can be contracted to a point (Grauert’s theorem). As $\mathcal{L}|_{E_{i}}$ and $O(E_{i})|_{E_{j}}$ ($i \neq j$) are trivial we get

\begin{equation}
-E_{j}.E_{j} = n_{j}E_{j}.E_{j},
\end{equation}

hence $n_{j} = -1$ as desired. □

### 2.4. Fourier-Jacobi expansions.
2.4.1. The infinitesimal retraction. We follow the arithmetic theory of Fourier-Jacobi expansions as developed in [Bel]. Let $\hat{S}$ be the formal completion of $\bar{S}$ along the cuspidal divisor $C = \bar{S} - S$. We work over $R_0$, and denote by $C^{(n)}$ the $n$-th infinitesimal neighborhood of $C$ in $\bar{S}$. The closed immersion $i : C \hookrightarrow \hat{S}$ admits a canonical left inverse $r : \hat{S} \to C$, a retraction satisfying $r \circ i = \text{Id}_C$. This is not automatic, but rather a consequence of the rigidity of tori, as explained in [Bel], Proposition II.2.4.2. As a corollary, the universal semi-abelian scheme $A_{/C^{(n)}}$ is the pull-back of $A_C$ via $r$. The same therefore holds for $\mathcal{P}$ and $\mathcal{L}$, namely there are natural isomorphisms $r^*(\mathcal{P}|_C) \simeq \mathcal{P}|_{C^{(n)}}$ and $r^*(\mathcal{L}|_C) \simeq \mathcal{L}|_{C^{(n)}}$. As a consequence, the filtration

$$0 \to \mathcal{P}_0 \to \mathcal{P} \to \mathcal{P}_\mu \to 0$$

extends canonically to $C^{(n)}$. Since $\mathcal{L}, \mathcal{P}_0$ and $\mathcal{P}_\mu$ are trivial on $C$, they are trivial over $C^{(n)}$ as well.

2.4.2. Arithmetic Fourier-Jacobi expansions. We fix an arbitrary noetherian $R_0$-algebra $R$ and consider all our schemes over $R$, without a change in notation. As usual, we let $\mathcal{O}_{\hat{S}} = \lim_{\to} \mathcal{O}_{C^{(n)}}$ (a sheaf in the Zariski topology on $C$). Via $r^*$, this is a sheaf of $\mathcal{O}_C$-modules. Choose a global nowhere vanishing section $s \in H^0(C, \mathcal{L})$ trivializing $\mathcal{L}$. Such a section is unique up to a unit of $R$ on each connected component of $C$. This $s$ determines an isomorphism

$$\mathcal{L}^k|_{\hat{S}} \simeq \mathcal{O}_{\hat{S}}, \quad f \mapsto f/s^k$$

for each $k$, hence a ring homomorphism

$$FJ : \bigoplus_{k=0}^{\infty} M_k(N, R) \to H^0(C, \mathcal{O}_{\hat{S}}).$$

We call $FJ(f)$ the (arithmetic) Fourier-Jacobi expansion of $f$. It depends on $s$ in an obvious way.

To understand the structure of $H^0(C, \mathcal{O}_{\hat{S}})$ let $\mathcal{I} \subset \mathcal{O}_{\hat{S}}$ be the sheaf of ideals defining $C$, so that $C^{(n)}$ is defined by $\mathcal{I}^n$. The conormal sheaf $\mathcal{N} = \mathcal{I}/\mathcal{I}^2$ is the restriction $i^*\mathcal{O}_{\hat{S}}(-C)$ of $\mathcal{O}_{\hat{S}}(-C)$ to $C$. It is an ample invertible sheaf on $C$, since (over $R_N$) its degree on each component $E_j$ is $-E_j^2 > 0$.

Now $r^*$ supplies, for every $n \geq 2$, a canonical splitting of

$$0 \to \mathcal{I}/\mathcal{I}^n \to \mathcal{O}_{\hat{S}}/\mathcal{I}^n \to \mathcal{O}_{\hat{S}}/\mathcal{I} \to 0.$$  

Inductively, we get a direct sum decomposition

$$\mathcal{O}_{\hat{S}}/\mathcal{I}^n \simeq \bigoplus_{m=0}^{n-1} \mathcal{I}^m/\mathcal{I}^{m+1}$$

as $\mathcal{O}_C$-modules, hence, since $\mathcal{I}^m/\mathcal{I}^{m+1} \simeq \mathcal{N}^m$, an isomorphism

$$H^0(C, \mathcal{O}_{C^{(n)}}) \simeq \bigoplus_{m=0}^{n-1} H^0(C, \mathcal{N}^m), \quad f \mapsto \sum_{m=0}^{n} c_m(f).$$

This isomorphism respects the multiplicative structure, so is a ring isomorphism. Going to the projective limit, and noting that the $c_m(f)$ are independent of $n$, we get

$$FJ(f) = \sum_{m=0}^{\infty} c_m(f) \in \prod_{m=0}^{\infty} H^0(C, \mathcal{N}^m).$$
2.4.3. Fourier-Jacobi expansions over C. Working over C, we shall now relate
the infinitesimal retraction r to the geodesic retraction, and the powers of the conormal
bundle N to theta functions. Recall the analytic compactification of Xr described in
Proposition 2.17. Let E be the connected component of Xr − Xr corresponding to the
standard cusp c∞. As before, we denote by E(n) its nth infinitesimal neighborhood.
The line bundle T|E is just the analytic normal bundle to E, hence we have an
isomorphism
\begin{equation}
N_{an} \cong T^\vee
\end{equation}
between the analytification of N = I/I^2 and the dual of T.

**Lemma 2.12.** The infinitesimal retraction r : E(n) → E coincides with the map
induced by the geodesic retraction (1.13).

**Proof.** The meaning of the lemma is this. The infinitesimal retraction induces a
map of ringed spaces
\begin{equation}
r_{an} : E_{an}^{(n)} \rightarrow E_{an}
\end{equation}
where E_{an} is the analytic space associated to E with its sheaf of analytic functions
O^an_E, and E_{an}^{(n)} is the same topological space with the sheaf O^an_E/I^an_{an}. The geodesic
retraction (sending (z, u) to u mod Λ) is an analytic map r_{geo} : E_{an}(ε) → E_{an},
where E_{an}(ε) is our notation for some tubular neighborhood of E_{an} in S_{an}. On the
other hand, there is a canonical map can of ringed spaces from E_{an}^{(n)} to E_{an}(ε). We
claim that these three maps satisfy r_{geo} ∘ can = r_{an}.

To prove the lemma, note that the infinitesimal retraction r : E(n) → E is
uniquely characterized by the fact that the O_k-semi-abelian variety A_x at a point
x of E(n), obtained as a specialization of the universal family A, is the pull-back of the
universal semi-abelian variety at r(x). See [Bel], II.2.4.2. The computations of
section 1.1 show that the same is true for the infinitesimal retraction obtained from
the geodesic retraction. We conclude that the two retractions agree on the level of
“truncated Taylor expansions”.

Consider now a modular form of weight k and level N over C, f ∈ M_k(N, C).
Using the trivialization of L_{an} over the symmetric space X given by 2πi · dζ, as discussed in Section 2.1.2 we identify f with a collection of functions f_j on X,
transforming under Γ_j according to the automorphy factor j(γ; z, u)^k. As usual we
look at Γ = Γ_1 only, and at the expansion of f = f_1 at the standard cusp c∞,
the other cusps being in principle similar. On the arithmetic FJ expansion side
this means that we concentrate on one connected component E of C, which lies
on the connected component of S_C corresponding to g_1 = 1. It also means that as
the section s used to trivialize L along E, we must use a section that, analytically,
coincides with 2πi · dζ.

Pulling back the sheaf N_{an} from E = C/Λ to C, it is clear that q = q(z) =
e^{2πiz/M} maps, at each u ∈ C, to a generator of T^\vee = N_{an} = I_{an}/I_{an}^2, and we
denote by q^m the corresponding generator of N_{an}^m = I_{an}^m/I_{an}^{m+1}. If
\begin{equation}
f(z, u) = \sum_{m=0}^{∞} θ_m(u)e^{2πimz/M} = \sum_{m=0}^{∞} θ_m(u)q^m
\end{equation}
is the complex analytic Fourier expansion of f at a neighborhood of c∞, then
c_m(z, u) = θ_m(u)q^m ∈ H^0(E, N_{an}^m) is just the restriction of the section denoted
above by $c_m(f)$ to $E$. The functions $\theta_m$ are classical elliptic theta functions (for the lattice $\Lambda$).

2.5. **The Gauss-Manin connection in a neighborhood of a cusp.**

2.5.1. A computation of $\nabla$ in the complex model. We shall now compute the Gauss-Manin connection in the complex ball model near the standard cusp $c_{\infty}$. Recall that we use the coordinates $(z, u, \zeta_1, \zeta_2, \zeta_3)$ as in Section 1.2.3. Here $d\zeta_1$ and $d\zeta_2$ form a basis for $\mathcal{P}$ and $d\zeta_3$ for $\mathcal{L}$. The same coordinates served to define also the semi-abelian variety $\mathcal{G}_n$ (denoted also $\mathcal{A}_n$) over the cuspidal component $E$ at $c_{\infty}$, cf Section 1.6. As explained there (1.69), the projection to the abelian part is given by the coordinate $\zeta_1$ (modulo $\mathcal{O}_K$), so $d\zeta_1$ is a basis for the sub-line-bundle of $\omega_{\mathcal{A}/E}$ coming from the abelian part, which was denoted $\mathcal{P}_0$. In section 2.4.1 above it was explained how to extend the filtration $\mathcal{P}_0 \subset \mathcal{P}$ canonically to the formal neighborhood $\hat{\mathcal{S}}$ of $E$ using the retraction $r$, by pulling back from the boundary. It was also noted that complex analytically, the retraction $r$ is the germ of the geodesic retraction introduced earlier. From the analytic description of the degeneration of $\mathcal{A}_{(z,u)}$ along a geodesic, it becomes clear that $\mathcal{P}_0 = r^*(\mathcal{P}_0|_E)$ is just the line bundle $\mathcal{O}_{\hat{\mathcal{S}}} \cdot d\zeta_1 \subset \omega_{\mathcal{A}/\hat{\mathcal{S}}}$. It follows that $\mathcal{P}_\mu = \mathcal{O}_{\hat{\mathcal{S}}} \cdot d\zeta_2 \mod \mathcal{P}_0$.

We shall now pull back these vector bundles to the ball $\mathfrak{X}$, and compute the Gauss Manin connection $\nabla$ complex analytically on $\omega_{\mathcal{A}/\mathfrak{X}}$. We write $\mathcal{P}_0 = \mathcal{O}_\mathfrak{X} \cdot d\zeta_1$ for $\mathcal{P}_{0,an}$ etc. dropping the decoration an. Let

$$\alpha_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \alpha_2 = \begin{pmatrix} 1 \\ 0 \\ u \end{pmatrix}, \alpha_3 = \begin{pmatrix} u \\ -z/\delta \\ z/\delta \end{pmatrix}$$

and

$$\alpha_1' = {\iota}'(\omega_\mathfrak{X})\alpha_1 = \begin{pmatrix} 0 \\ \omega_\mathfrak{K} \\ \bar{\omega}_\mathfrak{K} \end{pmatrix}, \alpha_2' = {\iota}'(\omega_\mathfrak{X})\alpha_2 = \begin{pmatrix} \omega_\mathfrak{K} \\ 0 \\ \bar{\omega}_\mathfrak{K}u \end{pmatrix},$$

$$\alpha_3' = {\iota}'(\omega_\mathfrak{X})\alpha_3 = \begin{pmatrix} \omega_\mathfrak{K}u \\ -\omega_\mathfrak{K}z/\delta \\ \bar{\omega}_\mathfrak{K}z/\delta \end{pmatrix}.$$ 

These 6 vectors span $L_{(z,u)}$ over $\mathbb{Z}$. Let $\beta_1, \ldots, \beta_3'$ be the dual basis to $\{\alpha_1, \ldots, \alpha_3\}$ in $H_{dR}^1(\mathcal{A}/\mathcal{O}_\mathfrak{X})$, i.e. $\int_{\alpha_i} \beta_1 = 1$ etc. As the periods of the $\beta_i$’s along the integral homology are constant, the $\beta$-basis is horizontal for the Gauss-Manin connection. The first coordinate of the $\alpha_i$ and $\alpha_i'$ gives us

$$d\zeta_1 = 0 \cdot \beta_1 + 1 \cdot \beta_2 + u \cdot \beta_3 + 0 \cdot \beta_1' + \omega_\mathfrak{K} \cdot \beta_2' + \omega_\mathfrak{K}u \cdot \beta_3',$$

and we find that

$$\nabla(d\zeta_1) = (\beta_3 + \omega_\mathfrak{K}\beta_3') \otimes du.$$ 

Similarly, we find

$$\nabla(d\zeta_2) = -\delta^{-1}(\beta_3 + \omega_\mathfrak{K}\beta_3') \otimes dz,$$

$$\nabla(d\zeta_3) = (\beta_2 + \bar{\omega}_\mathfrak{K}\beta_2') \otimes du + \delta^{-1}(\beta_3 + \bar{\omega}_\mathfrak{K}\beta_3') \otimes dz.$$
2.5.2. A computation of $KS$ in the complex model. We go on to compute the Kodaira-Spencer map on $\mathcal{P}$, i.e., the map denoted $KS(\Sigma)$. For that we have to take $\nabla(d\zeta_1)$ and $\nabla(d\zeta_2)$ and project them to $R^1\pi_*\mathcal{O}_A(\Sigma) \otimes \Omega^1_X$. We then pair the result, using the polarization form $\langle \cdot, \cdot \rangle_\lambda$ on $H^1_{dR}(A)$ (reflecting the isomorphism (2.57) $R^1\pi_*\mathcal{O}_A(\Sigma) = \text{Lie}(A^\vee)(\Sigma) = \omega_{A^\vee}(\Sigma) \simeq L^\vee(\rho)$ coming from $\lambda$), with $d\zeta_3$.

To perform the computation we need two lemmas

**Lemma 2.13.** The Riemann form on $L'_\Sigma$, associated to the polarization $\lambda$, is given in the basis $\alpha_1, \alpha_2, \alpha_3, \alpha'_1, \alpha'_2, \alpha'_3$ by the matrix

$$J = \begin{pmatrix} 0 & 1 \\ -I & 0 \\ 1 & 0 \end{pmatrix}.$$

**Proof.** This is an easy computation using the transition map $T$ between $L$ and $L'_\Sigma$ and the fact that on $L$ the Riemann form is the alternating form $\langle \cdot, \cdot \rangle_\lambda$ for $H_1^{dR}(A)$ as well as a Riemann form on the integral homology $H_1(A, \mathbb{Z})$. We compare the two.

**Lemma 2.14.** Let $(A, \lambda)$ be a principally polarized complex abelian variety. If $\alpha_1, \ldots, \alpha_{2g}$ is a symplectic basis for $H_1(A, \mathbb{Z})$ in which the associated Riemann form is given by a matrix $J$, and $\beta_1, \ldots, \beta_{2g}$ is the dual basis of $H^1_{dR}(A)$, then the matrix of the bilinear form $\langle \cdot, \cdot \rangle_\lambda$ on $H^1_{dR}(A)$ in the basis $\beta_1, \ldots, \beta_{2g}$ is $(2\pi i)^{-1}J$.

**Proof.** These are essentially Riemann’s bilinear relations. For example, if $A$ is the Jacobian of a curve $C$ and the basis $\alpha_1, \ldots, \alpha_{2g}$ has the standard intersection matrix

$$J = \begin{pmatrix} 0 & 1 \\ -I & 0 \end{pmatrix}$$

then the lemma follows from the well-known formula for the cup product ($\xi, \eta$ being differentials of the second kind on $C$)

$$\xi \cup \eta = \frac{1}{2\pi i} \sum_{i=1}^g \left( \int_{\alpha_i} \xi \int_{\alpha_i+\eta} \eta - \int_{\alpha_i} \eta \int_{\alpha_i+\eta} \xi \right).$$

Using the two lemmas we get

$$KS(d\zeta_1 \otimes d\zeta_3) = \langle \beta_3 + \omega_\Sigma \beta'_3, d\zeta_3 \rangle_\lambda \cdot du = -\delta(2\pi i)^{-1} du.$$

Similarly,

$$KS(d\zeta_2 \otimes d\zeta_3) = \langle -\delta^{-1}(\beta_3 + \omega_\Sigma \beta'_3), d\zeta_3 \rangle_\lambda \cdot dz = (2\pi i)^{-1} dz.$$

We summarize.
Proposition 2.15. Let \( z, u, \zeta_1, \zeta_2, \zeta_3 \) be the standard coordinates in a neighborhood of the cusp \( c_\infty \). Then, complex analytically, the Kodaira-Spencer isomorphism

\[
KS(\Sigma) : \mathcal{P} \otimes \mathcal{L} \simeq \Omega^1_X
\]

is given by the formulae

\[
KS(d\zeta_1 \otimes d\zeta_3) = -\delta(2\pi i)^{-1} du, \quad KS(d\zeta_2 \otimes d\zeta_3) = (2\pi i)^{-1} dz.
\]

Corollary 2.16. The Kodaira-Spencer isomorphism \( \mathcal{P} \otimes \mathcal{L} \simeq \Omega^1_S \) extends meromorphically over \( \hat{S} \). Moreover, in a formal neighborhood \( \hat{S} \) of \( C \), its restriction to the line sub-bundle \( \mathcal{P}_0 \otimes \mathcal{L} \) is holomorphic, and on any direct complement of \( \mathcal{P}_0 \otimes \mathcal{L} \) in \( \mathcal{P} \otimes \mathcal{L} \) it has a simple pole along \( C \).

Proof. As we have seen, \( d\zeta_1 \otimes d\zeta_3 \) and \( d\zeta_2 \otimes d\zeta_3 \) define a basis of \( \mathcal{P} \otimes \mathcal{L} \) at the boundary, with \( d\zeta_1 \otimes d\zeta_3 \) spanning the line sub-bundle \( \mathcal{P}_0 \otimes \mathcal{L} \). On the other hand \( du \) is holomorphic there, while \( dz \) has a simple pole along the boundary. \( \square \)

Corollary 2.17. The induced map

\[
(2.65) \quad \kappa : \Omega^1_X \to \mathcal{P}_\mu \otimes \mathcal{L}
\]

(\( \mathcal{P}_\mu = \mathcal{P}/\mathcal{P}_0 \)) obtained by inverting the isomorphism \( KS(\Sigma) \) and dividing \( \mathcal{P} \) by \( \mathcal{P}_0 \) kills \( du \) and maps \( dz \) to \( 2\pi i \cdot d\zeta_2 \otimes d\zeta_3 \).

Proof. As we have seen, \( d\zeta_1 \) is a basis for \( \mathcal{P}_0 \). \( \square \)

Corollary 2.18. The isomorphism \( \mathcal{L}^3 \simeq \Omega^2_S \) maps \( d\zeta_3 \otimes d\zeta_3 \) to a constant multiple of \( dz \wedge du \).

Proof. The isomorphism \( \det \mathcal{P} \simeq \mathcal{L} \) carries \( d\zeta_1 \wedge d\zeta_2 \) to a constant multiple of \( d\zeta_3 \), so the corollary follows from \( (2.64) \). \( \square \)

2.5.3. Transferring the results to the algebraic category. The computations in the analytic category over \( X \) of course descend (still in the analytic category) to \( S_C \), because they are local in nature. They then hold a fortiori in the formal completion \( \hat{S}_C \) along the cuspidal component \( E \). But the Gauss-Manin and Kodaira-Spencer maps are defined algebraically on \( S \), and both \( \Omega^1_S \) and \( \omega_{A/\hat{S}} \) are flat over \( R_0 \), so from the validity of the formulae over \( C \) we deduce their validity in \( \hat{S} \) over \( R_0 \), provided we identify the differential forms figuring in them (suitably normalized) with elements of \( \Omega^1_S \) and \( \omega_{A/\hat{S}} \) defined over \( R_0 \). In particular, they hold in the characteristic \( p \) fiber as well.

From the relation

\[
(2.66) \quad \frac{dq}{q} = \frac{2\pi i}{M} dz
\]

we deduce that the map \( \kappa \) has a simple zero along the cuspidal divisor.

Finally, although we have done all the computations at one specific cusp, it is clear that similar computations hold at any other cusp.

2.6. Fields of rationality.
2.6.1. Rationality of local sections of $\mathcal{P}$ and $\mathcal{L}$. We have compared the arithmetic surface $S$ with the complex analytic surfaces $\Gamma_j \backslash \mathfrak{X}$ (1 $\leq j \leq m$), and the compactifications of these two models. We have also compared the universal semi-abelian scheme $\mathcal{A}$ and the automorphic vector bundles $\mathcal{P}$ and $\mathcal{L}$ in both models. In this section we want to compare the local parameters obtained from the two presentations, and settle the question of rationality. To avoid issues of class numbers, we shall work rationally and not integrally.

We shall need to look at local parameters at the cusps, and as the cusps are defined only over $K_N$, we shall work with $S_{K_N}$ instead of $S_K$. With a little more care, working with Galois orbits of cusps, we could probably prove rationality over $K$, but for our purpose $K_N$ is good enough.

If $\xi$ and $\eta$ belong to a $K_N$-module, we write $\xi \sim \eta$ to mean that $\eta = c\xi$ for some $c \in K_N^\times$.

We begin with the vector bundles $\mathcal{P}$ and $\mathcal{L}$. Over $\mathbb{C}$ they yield analytic vector bundles $\mathcal{P}_{an}$ and $\mathcal{L}_{an}$ on each $X_{\Gamma_j}$ (1 $\leq j \leq m$). Assume for the rest of this section that $j = 1$ and write $\Gamma = \Gamma_1$. Similar results will hold for every $j$. The vector bundles $\mathcal{P}$ and $\mathcal{L}$ are trivialized over the unit ball $\mathfrak{X}$ by means of the nowhere vanishing sections $d\zeta_3 \in H^0(\mathfrak{X}, \mathcal{L}_{an})$ and $d\zeta_1, d\zeta_2 \in H^0(\mathfrak{X}, \mathcal{P}_{an})$. These sections do not descend to $X_{\Gamma}$, but

\begin{equation}
\sigma_{an} = (d\zeta_1 \otimes d\zeta_2) \otimes d\zeta_3^{-1} \in H^0(X_{\Gamma}, \det \mathcal{P} \otimes \mathcal{L}^{-1})
\end{equation}

does, as the factors of automorphy of $d\zeta_1 \otimes d\zeta_2$ and $d\zeta_3$ are the same (cf Section 2.1.2). Furthermore, this factor of automorphy (i.e. $j(\gamma; z, u)$) is trivial on $\Gamma_{cusps}$, the stabilizer of $c_\infty$ in $\Gamma$, so $d\zeta_1 \otimes d\zeta_2$ and $d\zeta_3$ define sections of $\det \mathcal{P}$ and $\mathcal{L}$ on $\hat{S}_C$, the formal completion of $S_C$ along the cuspidal divisor $E_c = p^{-1}(c_\infty) \subset S_C$. We have noted already that along $E_c$, $\mathcal{P}$ has a canonical filtration

\begin{equation}
0 \to \mathcal{P}_0 \to \mathcal{P} \to \mathcal{P}_\mu \to 0
\end{equation}

and that $d\zeta_1$ is a generator of $\mathcal{P}_0$. (Compare (1.63) and (1.71) and note that the projection to $\mathbb{C}/\mathcal{O}_K = B(\mathbb{C})$ is via the coordinate $\zeta_1$, so $d\zeta_1$ is a generator of $\mathcal{P}_{0|E_c} = \omega_B$.) As we have shown in Section 2.4.1 this filtration extends to the formal neighborhood $\hat{S}_C$ of $E_c$. The vector bundles $\mathcal{P}$ and $\mathcal{L}$, as well as the filtration on $\mathcal{P}$ over $\hat{S}_C$, are defined over $K_N$. It makes sense therefore to ask if certain sections are $K_N$-rational. Recall that the cusp $c_\infty$ is of type $(\mathcal{O}_K, \mathcal{O}_K)$.

**Proposition 2.19.** (i) $2\pi i \cdot d\zeta_3 \in H^0(\hat{S}_{K_N}, \mathcal{L})$. In other words, this section is $K_N$-rational.

(ii) Similarly $2\pi i \cdot d\zeta_2$ projects (modulo $\mathcal{P}_0$) to a $K_N$-rational section of $\mathcal{P}_\mu$.

(iii) Let $B$ be the elliptic curve over $K_N$ associated with the cusp $c_\infty$ as in Section 1.5.1. Let $\Omega_B \in \mathbb{C}^\times$ be a period of a basis $\omega$ of $\omega_B = H^0(B, \Omega_{B/K_N}^1)$ (i.e. the lattice of periods of $\omega$ is $\Omega_B \cdot \mathcal{O}_K$). This $\Omega_B$ is well-defined up to an element of $K_N^\times$. Then $\Omega_B \cdot d\zeta_1 \in H^0(\hat{S}_{K_N}, \mathcal{P}_0)$ is $K_N$-rational.

**Proof.** Let $E$ be the component of $C_{K_N}$ which over $\mathbb{C}$ becomes $E_c$. Let $\mathcal{G}$ be the universal semi-abelian scheme over $E$. Then $\mathcal{G}$ is a semi-abelian scheme which is an extension of $B \times_{K_N} E$ by the torus $(\mathcal{O}_K \otimes \mathbb{G}_{m, K_N}) \times_{K_N} E$. At any point $u \in E(\mathbb{C})$ we have the analytic model $\mathcal{G}_u$ (1.69) for the fiber of $\mathcal{G}$ at $u$, but the abelian part and the toric part are constant. Over $E$ the line bundle $\mathcal{P}_0$ is (by definition) $\omega_{B \times E/E}$. As the lattice of periods of a suitable $K_N$-rational differential is $\Omega_B \cdot \mathcal{O}_K$, while the
lattice of periods of $d\zeta_1$ is $\mathcal{O}_K$, part (iii) follows. For parts (i) and (ii) observe that
the toric part of $\mathcal{G}$ is in fact defined over $K$, and that $e^c_{\mathcal{O}_K}$ maps the cotangent space
of $\mathcal{O}_K \otimes \mathbb{G}_{m,K}$ isomorphically to the $K$-span of $2\pi i \zeta_2$ and $2\pi i \zeta_3$. $\square$

**Corollary 2.20.** $\Omega_B \cdot \sigma_{an}$ is a nowhere vanishing global section of $\det \mathcal{P} \otimes \mathcal{L}^{-1}$
over $S_F$, rational over $K_N$.

**Proof.** Recall that we denote by $S_F$ the connected component of $S_{K,N}$ whose associated analytic space is
the complex manifold $X_F$. We have seen that as an analytic section $\Omega_B \cdot \sigma_{an}$ descends to $X_F$ and
extends to the smooth compactification $\hat{X}_F$. By GAGA, it is algebraic. Since $\hat{X}_F$ is connected, to check its field of definition, it
is enough to consider it at one of the cusps. By the Proposition, its restriction to the formal neighborhood of $E_c$ ($c = c_{\infty}$) is defined over $K_N$. $\square$

The complex periods $\Omega_B$ (and their powers) appear as the transcendental parts of special values of $L$-functions
associated with Grossencharacters of $K$. They are therefore instrumental in the construction of $p$-adic $L$ functions on $K$. We expect them to appear in the $p$-adic interpolation of holomorphic Eisenstein series on the

2.6.2. *Rationality of local parameters at the cusps.* We keep the assumptions and the notation of the previous section. Analytically, neighborhoods of $E_{c_{\infty}}$ were described in Section 2.4.1 with the aid of the parameters $(z, u)$. Let $\hat{S}$ denote the formal completion of $\hat{S}_c$ along $E$. Let $r : \hat{S} \to E$ be the infinitesimal retraction discussed in Section 2.4.1. If $i : E \hookrightarrow \hat{S}$ is the closed embedding then $r \circ i = 1_{E}$. If $\mathcal{I}$ is the sheaf of definition of $E$, then $\mathcal{N} = \mathcal{I}/\mathcal{I}^2$ is the conormal bundle to $E$, hence its analytification is the dual of the line bundle $\mathcal{I}$,

\begin{equation}
(2.69)
\mathcal{N}_{an} = \mathcal{I}'.
\end{equation}

Consider $r^*\mathcal{N}$ on $\hat{S}$. The retraction allows us to split the exact sequence
\begin{equation}
(2.70)
0 \to \mathcal{N} \to r^*\Omega^1_{\hat{S}} \to \Omega^1_{\hat{E}} \to 0
\end{equation}

using $\Omega^1_{\hat{E}} = i^*r^*\Omega^1_{\hat{E}} \subset i^*\Omega^1_{\hat{S}}$. Thus $i^*\Omega^1_{\hat{S}} = \mathcal{N} \times \Omega^1_{\hat{E}}$. The map $i \circ r : \hat{S} \to \hat{S}$ induces a sheaf homomorphism $r^*i^*\Omega^1_{\hat{S}} \to \Omega^1_{\hat{S}}$ which becomes the identity if we restrict it to $E$ (i.e. follow it with $i^*$). By Nakayama’s lemma, it is an isomorphism. It follows that

\begin{equation}
(2.71)
\Omega^1_{\hat{S}} = r^*i^*\Omega^1_{\hat{S}} = r^*\mathcal{N} \times r^*\Omega^1_{\hat{E}}.
\end{equation}

Let $x \in E$ and represent it by $u \in \mathbb{C}$ (modulo $\Lambda$). Then $q = e^{2\pi iz/M}$, where $M$ is the width of the cusp (1.54), is a local analytic parameter on a classical neighborhood $U_x$ of $x$ which vanishes to first order along $E$. Note that $q$ depends on the choice of $u$ (see Remark below). It follows that $dq$, the image of $q$ in $\mathcal{I}_{an}/\mathcal{I}_{an}^2$, is a basis of $\mathcal{N}_{an}$ (on $U_x \cap E$). But

\begin{equation}
(2.72)
2\pi i \cdot dz = M \frac{dq}{q}
\end{equation}
is independent of $u$ (see (1.55)), so represents a global meromorphic section of $r^*\mathcal{N}_{an}$, with a simple pole along $E \subset \hat{S}_c$. By GAGA, this section is (meromorphic) algebraic.
Proposition 2.21. (i) The section $2\pi i \cdot dz$ is $K_N$-rational, i.e. it is the analytification of a section of $r^* L$. (ii) The section $\Omega_B \cdot du$ is $K_N$-rational, i.e. belongs to $H^0(E, \Omega^1_{E/K_N})$.

Proof. The proof relies on the Kodaira-Spencer isomorphism $KS(\Sigma)$ (2.64), which is a $K_N$-rational (even $K$-rational) algebraic isomorphism between $P \otimes L$ and $\Omega^1_{\tilde{S}}$. As we have shown, it extends to a meromorphic homomorphism from $P \otimes L$ to $\Omega^1_{\tilde{S}}$ over $\tilde{S}$. Over $\tilde{S}$ it induces an isomorphism of $P_0 \otimes L$ onto $r^* \Omega^1_{E} \subset \Omega^1_{\tilde{S}}$ carrying the $K_N$-rational section $\Omega_B d\zeta_1 \otimes 2\pi i d\zeta_3$ to $-\Omega_B \delta \cdot du$, proving part (ii) of the proposition. It also carries $2\pi i d\zeta_2 \otimes 2\pi i d\zeta_3$ to $2\pi i dz$, but the latter is only meromorphic. We may summarize the situation over $\tilde{S}$ by the following commutative diagram with exact rows:

$$
\begin{array}{cccccc}
0 & \rightarrow & \tilde{I} \otimes P_0 \otimes L & \rightarrow & \tilde{I} \otimes P \otimes L & \rightarrow & \tilde{I} \otimes P_0 \otimes L & \rightarrow & 0 \\
\downarrow & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \rightarrow & r^* \Omega^1_{E} & \rightarrow & \Omega^1_{\tilde{S}} & \rightarrow & r^* L & \rightarrow & 0.
\end{array}
$$

(2.73)

Let $h$ be a $K_N$-rational local equation of $E$, i.e. a $K_N$-rational section of $\tilde{I}$ in some Zariski open $U$ intersecting $E$ non-trivially, vanishing to first order along $E \cap U$. The differential $\eta = h \cdot (2\pi i dz)$ is regular on $U$, and to prove that it is $K_N$-rational we may restrict it to $\tilde{S}$ and check rationality there. But in $\tilde{S}$ we have a $K_N$-rational product decomposition $\Omega^1_{\tilde{S}} = r^* L \times r^* \Omega^1_{E}$ and the projection of $\eta$ to the second factor is 0, so it is enough to prove rationality of its projection to $r^* L$. This projection is the image, under $KS(\Sigma)$, of $h \cdot (2\pi i d\zeta_2 \otimes 2\pi i d\zeta_3 \mod P_0 \otimes L)$, so our assertion follows from parts (i) and (ii) of the previous proposition. This proves that $\eta$, hence $h^{-1} \eta = 2\pi i dz$ is a $K_N$-rational differential. An alternative proof of part (ii) is to note that $E$ is isogenous over $K_N$ to $B$, so up to a $K_N$-multiple has the same period. \qed

Remark 2.1. Unlike $2\pi i dz = M dq/q$, the parameter $q$ is not a well-defined parameter at $x$, and depends not only on $x$, but also on the point $v$ used to uniformize it. If we change $u$ to $u + s$ ($s \in \Lambda$) then $q$ is multiplied by the factor $e^{2\pi i \delta (u + s/2)} M$, so although $O^{hol}_{S_{c,x}} \subset O_{S_{c,x}}$ and analytic parameters may be considered as formal parameters, the question whether $q$ itself is $K_N$-rational is not well-defined (in sharp contrast to the case of modular curves!).

2.6.3. Normalizing the isomorphism $\det P \simeq L$. Let us fix a nowhere vanishing section

$$
\sigma \in H^0(S_K, \det P \otimes L^{-1}).
$$

(2.74)

This section is determined up to $K^\times$. From now on we shall use this section to identify $\det P$ with $L$ whenever such an identification is needed. From Corollary 2.20 we deduce that when we base change to $\mathbb{C}$, on each connected component $X_F$

$$
\sigma \sim \Omega_B \cdot \sigma_{an}.
$$

(2.75)

3. Picard moduli schemes modulo an inert prime $p$

3.1. The stratification.
3.1.1. The three strata. Fix a prime $p > 2$ which is inert in $\mathcal{K}$ and relatively prime to $N$. Then $R_0/pR_0 = \mathbb{F}_p$ and we consider the characteristic $p$ fiber of the smooth compactification $\tilde{S}$ of $S$.

\begin{equation}
\tilde{S} \times_{\text{Spec}(R_0)} \text{Spec}(R_0/pR_0).
\end{equation}

From now on we shall write $\tilde{S}$ to denote this scheme, rather than the original one over $R_0$. We let $\mathcal{A}_x$ as before, stand for the universal semi-abelian variety over $\tilde{S}$. The structure of $\tilde{S}$ has been worked out by Vollaard [V]. We record her main results.

Recall that an abelian variety over a field of characteristic $p$ is called supersingular if the Newton polygon of its $p$-divisible group is of constant slope $1/2$. It is called superspecial if it is isomorphic to a product of supersingular elliptic curves.

**Theorem 3.1.** (i) There exists a closed reduced 1-dimensional subscheme $S_{ss} \subset \tilde{S}$ (the supersingular locus), disjoint from the cuspidal divisor (i.e. contained in $S$), which is uniquely characterized by the fact that for every geometric point $x$ of $\tilde{S}$, the abelian variety $\mathcal{A}_x$ is supersingular if and only if $x \in S_{ss}$.

(ii) Let $S_{ss}$ be the singular locus on $S_{ss}$. Then $x$ lies in $S_{ss}$ if and only if $\mathcal{A}_x$ is superspecial. If $x_0 \in S_{ss}$ then $x_0$ is rational over $\mathbb{F}_p^2$ and

\begin{equation}
\hat{\mathcal{O}}_{S, x_0} \simeq \mathbb{F}_p^2[[u,v]]/(u^{p+1} + v^{p+1}).
\end{equation}

(iii) Assume that $N$ is large enough (depending on $p$). Then the irreducible components of $S_{ss}$ are rational over $\mathbb{F}_p^2$, nonsingular, and in fact are all isomorphic to the Fermat curve

\begin{equation}
x^{p+1} + y^{p+1} + z^{p+1} = 0.
\end{equation}

There are $p^3 + 1$ points of $S_{ss}$ on each irreducible component, and through each such point pass $p + 1$ irreducible components. Any two irreducible components are either disjoint or intersect transversally at a unique point.

We call $\tilde{S}_{\mu} = \tilde{S} - S_{ss}$ (or $\mathcal{S}_{\mu} = \tilde{S}_{\mu} \cap S$) the $\mu$-ordinary or generic locus, $S_{gss} = S_{ss} - S_{ss}$ the general supersingular locus, and $S_{ss}$ the superspecial locus. Then $\tilde{S} = \mathcal{S}_{\mu} \cup S_{gss} \cup S_{ss}$ is a stratification. The three strata are of dimensions 2, 1, and 0 respectively, the closure of each stratum contains the lower dimensional ones, and each of the three is open in its closure.

3.1.2. The $p$-divisible groups. Bültel and Wedhorn [Bu-We] and Vollaard describe the $p$-divisible group $\mathcal{A}_x(p)$ of the abelian variety $\mathcal{A}_x$ for $x$ in the various strata. Let $x : \text{Spec}(k) \to \tilde{S}_{\mu}$ be a geometric point ($k$ an algebraically closed field of characteristic $p$) and $A = \mathcal{A}_x$ (an abelian variety over $k$). Then the $p$-divisible group $A(p)$ has a three-step $\mathcal{O}_K$-invariant filtration

\begin{equation}
0 \subset \text{Fil}^2 A(p) \subset \text{Fil}^3 A(p) \subset \text{Fil}^0 A(p) = A(p)
\end{equation}

such that $gr^2 = \text{Fil}^2$ is a group of multiplicative type (connected, with étale dual), $gr^1 = \text{Fil}^1 / \text{Fil}^2$ is connected with a connected dual, and $gr^0 = \text{Fil}^0 / \text{Fil}^1$ is étale. Each of the three graded pieces is a $p$-divisible group of height 2 ($\mathcal{O}_K$-height 1). As every $p$-divisible group over an algebraically closed field of characteristic $p$ splits canonically into a product of a group of multiplicative type, a group which is connected with a connected dual, and an étale group, the above filtration splits.
We have
\[
\begin{align*}
gr^2 A(p) & \simeq a \otimes \mu_{p^\infty} \\
gr^1 A(p) & \simeq \mathfrak{G} \\
gr^0 A(p) & \simeq c \otimes \mathbb{Q}_p / \mathbb{Z}_p
\end{align*}
\]
(3.5)
where $a$ and $c$ are ideals of $\mathcal{O}_K$ and $\mathfrak{G}$ is the $p$-divisible group of a supersingular elliptic curve over $\mathbb{F}_p$ (the group denoted by $G_{1,1}$ in the Manin-Dieudonné classification [Dem]). Although, up to isomorphism, it is possible to substitute $\mathcal{O}_K$ for $a$ or $c$, it is sometimes more natural to allow this extra freedom in notation.

Geometric points of $S_{ss}$ classify abelian varieties $A = A_x$ for which $A(p)$ is isogenous to $\mathfrak{G}^4$, and such a point $x$ is superspecial if and only if $A(p)$ is isomorphic to $\mathfrak{G}^3$.

The $a$-number of $A$ is the dimension of the $k$-linear space $\text{Hom}(a_p, A[p])$. It is 1 if $x \in S_1 \cup S_{gss}$ and 3 if $x \in S_{ssp}$. The stratification which we have described coincides, in our simple case, with the Ekedahl-Oort stratification [Oo],[Mo],[We].

3.2. New relations between automorphic vector bundles in characteristic $p$.

3.2.1. The line bundles $\mathcal{P}_0$ and $\mathcal{P}_\mu$ over $\bar{S}_\mu$. Consider the universal semi-abelian variety $\mathcal{A}$ over the Zariski open set $\bar{S}_\mu$. As we have seen in Section 2.3.1 over the cuspidal divisor $C$, $\mathcal{P} = \omega_{\mathcal{A}}(\Sigma)$ admits a canonical filtration
\[
0 \to \mathcal{P}_0 \to \mathcal{P} \to \mathcal{P}_\mu \to 0
\]
(3.6)
where $\mathcal{P}_0$ is the cotangent space to the abelian part of $\mathcal{A}$, and $\mathcal{P}_\mu$ the $\Sigma$-component of the cotangent space to the toric part of $\mathcal{A}$. This filtration exists already in characteristic 0, but when we reduce modulo $p$ it extends, as we now show, to the whole of $\bar{S}_\mu$.

Let $\mathcal{A}[p]^0$ be the connected part of the subgroup scheme $\mathcal{A}[p]$. Then $\mathcal{A}[p]^0$ is finite flat over $\bar{S}_\mu$ of rank $p^4$. (It is clearly flat and quasi-finite, and the fiber rank can be computed separately on $C$ and on $S_\mu$. Since the rank is constant, the morphism to $\bar{S}_\mu$ is actually finite, cf. [De-Ra], Lemme 1.19.) Let
\[
0 \subset \mathcal{A}[p]^\mu \subset \mathcal{A}[p]^0
\]
(3.7)
be the maximal subgroup-scheme of multiplicative type. Since at every geometric point of $\bar{S}_\mu$, $\mathcal{A}[p]^\mu$ is of rank $p^2$, this subgroup is also finite flat over $\bar{S}_\mu$. It is also $\mathcal{O}_K$-invariant. Over the cuspidal divisor $C$, $\mathcal{A}[p]^\mu$ is the $p$-torsion in the toric part of $\mathcal{A}$, and over $S_\mu$
\[
\mathcal{A}[p]^\mu = \mathcal{A}[p] \cap \text{Fil}^2 A(p).
\]
(3.8)
As $\omega_{\mathcal{A}}$ is killed by $p$, we have $\omega_{\mathcal{A}} = \omega_{\mathcal{A}[p]} = \omega_{\mathcal{A}[p]^0}$. Let $\omega_{\mathcal{A}}^\mu = \omega_{\mathcal{A}[p]^\mu}$, a rank-2 $\mathcal{O}_K$-vector bundle of type $(1,1)$. The kernel of $\omega_{\mathcal{A}[p]^0} \to \omega_{\mathcal{A}[p]^\mu}$ is then a line bundle $\mathcal{P}_0$ of type $(1,0)$ and we get the short exact sequence
\[
0 \to \mathcal{P}_0 \to \omega_{\mathcal{A}} \to \omega_{\mathcal{A}}^\mu \to 0
\]
(3.9)
over the whole of $\bar{S}_\mu$. Decomposing according to types and setting $\mathcal{P}_\mu = \omega_{\mathcal{A}}^\mu(\Sigma)$, we get the desired filtration.
3.2.2. Frobenius and Verschiebung. Write $G$ for $A[p]$ (a finite flat $\mathcal{O}_K$-group scheme over $S$) and $G^{(p)}$ for the base-change of $G$ with respect to the absolute Frobenius morphism of degree $p$. In other words, if we denote by $\phi$ the homomorphism $x \mapsto x^p$ (of any $\mathbb{F}_p$-algebra), and by $\Phi : S \to S$ the corresponding map of schemes, then

\begin{equation}
G^{(p)} = S \times_{\Phi, S} G.
\end{equation}

The relative Frobenius is an $\mathcal{O}_S$-linear homomorphism $\text{Frob}_G : G \to G^{(p)}$, characterized by the fact that $pr_2 \circ \text{Frob}_G$ is the absolute Frobenius morphism of $G$. The relative Verschiebung is an $\mathcal{O}_S$-linear homomorphism $\text{Ver}_G : G^{(p)} \to G$. Under Cartier duality $\text{Ver}_G$ is dual to $\text{Frob}_G^D$, where we denote by $G^D$ the Cartier dual of $G$.

Recall that $\omega_G = \omega_A$, and similarly $\omega_{G^{(p)}} = \omega_A^{(p)} = \omega_A^{(p)} = \mathcal{O}_S \otimes_{\Phi, \mathcal{O}_S} \omega_A$. The morphism $\text{Ver}_G : G^{(p)} \to G$ induces a homomorphism of vector bundles $V : \omega_A \to \omega_A^{(p)}$, and taking $\Sigma$-components we get

\begin{equation}
V : \mathcal{P} = \omega_A(\Sigma) \to \omega_A(\Sigma)^{(p)} = \omega_A(\Sigma)^{(p)} = \mathcal{L}^{(p)}.
\end{equation}

Over $\bar{S}_\mu$ this map fits in a commutative diagram

\begin{equation}
\begin{array}{cccc}
0 & \leftarrow & \omega_A(\Sigma) & \leftarrow & \mathcal{P}_\mu & \leftarrow & \mathcal{P} & \leftarrow & 0 \\
\downarrow & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \leftarrow & (\omega_A(\Sigma))^{(p)} & \leftarrow & \mathcal{L}^{(p)} & \leftarrow & 0 & \leftarrow & 0
\end{array}
\end{equation}

The right vertical arrow is 0 since $V$ kills $\mathcal{P}_0$, as $\mathfrak{G}$ is of local-local type. The left vertical map is an isomorphism, since $Ver$ is an isomorphism on $p$-divisible groups of multiplicative type. We conclude that

\begin{equation}
\mathcal{P}_0 = \ker(V : \mathcal{P} \to \mathcal{L}^{(p)}).
\end{equation}

3.2.3. Relations between $\mathcal{P}_0, \mathcal{P}_\mu$ and $\mathcal{L}$ over $\bar{S}_\mu$. We first recall a general lemma.

**Lemma 3.2.** Let $\mathcal{M}$ be a line bundle over a scheme $S$ in characteristic $p$. Let $\Phi : S \to S$ be the absolute Frobenius and $\mathcal{M}^{(p)} = \Phi^* \mathcal{M}$. Then the map $\mathcal{M}^{(p)} \to \mathcal{M}$

\begin{equation}
a \otimes m \mapsto a \cdot m \otimes \cdots \otimes m
\end{equation}

is an isomorphism of line bundles over $S$.

Since $\mathcal{L}^{(p)} \simeq \mathcal{L}^p$ by the lemma, we have

\begin{equation}
\mathcal{L}^p \simeq \mathcal{P}/\mathcal{P}_0 = \mathcal{P}_\mu.
\end{equation}

Finally, from $\mathcal{P}_0 \otimes \mathcal{P}_\mu \simeq \det \mathcal{P} \simeq \mathcal{L}$ we get

\begin{equation}
\mathcal{P}_0 \simeq \mathcal{L}^{1-p}.
\end{equation}

We have proved:

**Proposition 3.3.** Over $\bar{S}_\mu$, $\mathcal{P}_\mu \simeq \mathcal{L}^p$ and $\mathcal{P}_0 \simeq \mathcal{L}^{1-p}$.

In the same vein we get a commutative diagram for the $\bar{\Sigma}$ parts

\begin{equation}
\begin{array}{cccc}
0 & \leftarrow & \omega_A^{(p)}(\bar{\Sigma}) & \leftarrow & \mathcal{L} & \leftarrow & 0 & \leftarrow & 0 \\
\downarrow & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \leftarrow & (\omega_A^{(p)}(\bar{\Sigma}))^{(p)} & \leftarrow & \mathcal{P}^{(p)} & \leftarrow & \mathcal{P}_0^{(p)} & \leftarrow & 0
\end{array}
\end{equation}

and deduce that $V$ is injective on $\mathcal{L}$ and

\begin{equation}
\mathcal{P}^{(p)} = \mathcal{P}_0^{(p)} \oplus V(\mathcal{L}).
\end{equation}
Thus over $\tilde{S}_\mu$, $\mathcal{P}$ has a canonical filtration by $\mathcal{P}_0$, but the induced filtration on $\mathcal{P}^{(p)}$ already splits as a direct sum.

**Remark 3.1.** Working over $\bar{\mathbb{F}}_p$, and restricting attention to a connected component $E$ of $C$, $\mathcal{P}|_E$ is a non-split extension of $\mathcal{P}_\mu$ by $\mathcal{P}_0$. However, both $\mathcal{P}_\mu$ and $\mathcal{P}_0$ are trivial on $E$, so the extension is described by a non-zero class $\xi \in H^1(E, \mathcal{O}_E)$.

The extension $\mathcal{P}^{(p)}|_E$ is then described by $\xi^{(p)}$. The semilinear map $\xi \mapsto \xi^{(p)}$ is the Cartier-Manin operator, and since $E$ is a supersingular elliptic curve, $\xi^{(p)} = 0$ and $\mathcal{P}^{(p)}|_E$ splits. Thus at least over $C$, the splitting of $\mathcal{P}^{(p)}$ is consistent with what we know so far.

Since $V$ induces an isomorphism of $\mathcal{L}$ onto $\mathcal{P}^{(p)}/\mathcal{P}_0^{(p)} \simeq (\mathcal{P}/\mathcal{P}_0)^p$ and $\mathcal{P}/\mathcal{P}_0 \simeq \mathcal{L}^p$ we conclude that over $\tilde{S}_\mu$, $\mathcal{L} \simeq \mathcal{L}^{e^2}$. In the next section we realize this isomorphism via the Hasse invariant. Combining what was proved so far we easily get the following.

**Proposition 3.4.** Over $\tilde{S}_\mu$, $\mathcal{L}^{e^2} \simeq \mathcal{L}$. For $k \geq 1$ odd, $\mathcal{P}^{(p)} \simeq \mathcal{L}^{e-1} \oplus \mathcal{L}$. For $k \geq 2$ even, $\mathcal{P}^{(p)} \simeq \mathcal{L}^{1-p} \oplus \mathcal{L}^p$, but for $k = 0$ we only have an exact sequence

$$0 \to \mathcal{L}^{1-p} \to \mathcal{P} \to \mathcal{L}^p \to 0.$$  \hfill (3.19)

**Corollary 3.5.** Over $\tilde{S}_\mu$, $\mathcal{L}^{p^2-1}$, $\mathcal{P}_0^{p^2-1}$ and $\mathcal{P}_0^{p+1}$ are trivial line bundles.

3.2.4. **Extending the filtration on $\mathcal{P}$ over $S_{gss}$.** In order to determine to what extent the filtration on $\mathcal{P}$ and the relation between $\mathcal{L}$ and the two graded pieces of the filtration extend into the supersingular locus, we have to employ Dieudonné theory.  

**Proposition 3.6.** Let $\mathcal{P}_0 = \ker(V : \mathcal{P} \to \mathcal{L}^{(p)})$. Then over the whole of $\tilde{S} - S_{gss}$, $V(\mathcal{P}) = \mathcal{L}^{(p)}$ and $\mathcal{P}_0$ is a rank 1 submodule. Let $\mathcal{P}_\mu = \mathcal{P}/\mathcal{P}_0$. Then $\mathcal{P}_\mu \simeq \mathcal{L}^p$, $\mathcal{P}_0 \simeq \mathcal{L}^{1-p}$, and the filtration (3.14) is valid there.

**Proof.** Everything is a formal consequence of the fact that $V$ maps $\mathcal{P}$ onto $\mathcal{L}^{(p)}$, and the relation $\det \mathcal{P} \simeq \mathcal{L}$. Over $\tilde{S}_\mu$, the proposition was verified in the previous subsection, so it is enough to prove that $V(\mathcal{P}) = \mathcal{L}^{(p)}$ in the fiber of any geometric point $x \in S_{gss}(k)$ ($k$ algebraically closed). We use the description of $H^1_{DR}(\mathbb{A}_x/k)$ given in Lemma 2.9 below, due to Büttel and Wedhorn. In the notation of that lemma, $\mathcal{P}_x$ is spanned over $k$ by $e_1$ and $e_2$ and $\mathcal{L}_x$ by $f_3$, while $V(e_1) = 0$, $V(e_2) = f_3^{(p)}$. This concludes the proof. \hfill \square

**Proposition 3.7.** Over the whole of $\tilde{S} - S_{gss}$, $V$ maps $\mathcal{L}$ injectively onto a sub-line-bundle of $\mathcal{P}^{(p)}$.

**Proof.** Once again, we know it already over $\tilde{S}_\mu$, and it remains to check the assertion fiber-wise on $S_{gss}$. We refer again to Lemma 2.9 and find that $V(f_3) = e_1^{(p)}$, which proves our claim. \hfill \square

The emerging picture is this: Outside the superspecial points, $V$ maps $\mathcal{L}$ injectively onto a sub-line-bundle of $\mathcal{P}^{(p)}$, and $V^{(p)}$ maps $\mathcal{P}^{(p)}$ surjectively onto $\mathcal{L}^{(p^2)}$. However, the line $V(\mathcal{L})$ coincides with the line $\mathcal{P}_0^{(p)} = \ker(V^{(p)})$ only on the general supersingular locus, while on its complement $\tilde{S}_\mu$ the two lines make up a frame for $\mathcal{P}^{(p)}$ (3.18). One can be a little more precise. The equation

$$V(\mathcal{L}) = \mathcal{P}_0^{(p)}$$  \hfill (3.20)
uniquely, but the extension does not come from $S_p$ we study the deformation of the action of $V$ distinct. In other words, on the normalization of \((3.22)\)

$$H(V) = (3.21) \]$$

we cast it in the language of de Rham cohomology. At a point and we end up using Grothendieck’s crystalline deformation theory, even though \(\text{Proof. To simplify the computations we change the base field to } \bar{\mathbb{F}}_p\). At any superspecial point there are \(2 \) branches of \(S_{ss}\) meeting transversally. We shall prove the proposition by showing that along any one of these branches (labelled by \(\zeta\), a \(p+1\)st root of \(-1\)) \(P_0\) approaches a line \(P_x[\zeta] \subset P_x\), but these \(p+1\) lines are distinct. In other words, on the normalization of \(S_{ss}\) we can extend the filtration uniquely, but the extension does not come from \(S_{ss}\).

Before we go into the proof a word of explanation is needed. In Section \(4.3\) below we study the deformation of the action of \(V\) on \(\omega_A\) near a general supersingular point \(x \in S_{gss}\). For that purpose the first infinitesimal neighborhood of \(x\) suffices, and we end up using Grothendieck’s crystalline deformation theory, even though we cast it in the language of de Rham cohomology. At a point \(x \in S_{gss}\), in contrast, we need to work in the full formal neighborhood of \(x\) in \(S\), or at least in an Artinian neighborhood which no longer admits a divided power structure. The reason is that the singularity of \(S_{ss}\) at \(x\) is formally of the type \(\text{Spec}(\bar{\mathbb{F}}_p[[u,v]]/(u^{p+1} + v^{p+1}))\), as we shall see in \(3.27\). Crystalline deformation theory is inadequate, and we need to use Zink’s “displays”. As the theory of displays is covariant, we start with the covariant Cartier module of \(A = \mathcal{A}_x\) rather than the contravariant Dieudonné module, and look for its universal deformation.

Let us review the (confusing) functoriality of these two modules. For the moment, let \(A\) be an abelian variety over \(\bar{\mathbb{F}}_p\). If \(D\) is the (contravariant) Dieudonné module of \(A\) and \(M\) is its (covariant) Cartier module, then \(D/pD = H^1_{dR}(A)\) and \(M/pM = H^1_{dR}(A')\) are set in duality. The dual of \(V = \text{Ver}_A^*\), \(\text{Ver}_A\) being the Verschiebung isogeny from \(A'^{(p)}\) to \(A\), is the map \(F : (M/pM)^{(p)} \to M/pM\) (\(F = \text{Frob}_A^\ell\), \(\text{Frob}_A\) being the Frobenius isogeny from \(A\) to \(A'^{(p)}\)). As usual, since \(\bar{\mathbb{F}}_p\) is perfect, we may view \(V\) as a \(\phi^{-1}\)-linear map of \(D/pD\), and \(F\) as a \(\phi\)-linear map of \(M/pM\). Replacing \(A\) by \(A'\) we then also have a map \(F\) on \(D/pD\) and \(V\) on \(M/pM\). The Hodge filtration \(\omega_A \subset H^1_{dR}(A)\) is \((D/pD)[F]\). Its dual is the quotient \(\text{Lie}(A) = H^1(A', \mathcal{O})\) of \(H^1_{dR}(A')\), identified with \(M/VM\). Compare [Oda], Corollary 5.11.

This reminder tells us that when we pass from the contravariant theory to the covariant one, instead of looking for the deformation of \(V\) on \(\omega_A\) we should look for the deformation of \(F\) on \(\text{Lie}(A) = M/VM\). At a superspecial point \(F\) annihilates \(\text{Lie}(A)\), but at nearby points in \(S\) it need not annihilate it anymore.

Now let \(x \in S_{gss}\) and \(A = \mathcal{A}_x\). The Cartier module (modulo \(p\)) \(M/pM\) of \(A\) admits a symplectic basis \(f_3, e_1, e_2, e_3, f_1, f_2\) where \(\mathcal{O}_K\) acts on the \(e_i\) via \(\Sigma\) and
on the $f_i$ via $\tilde{\Sigma}$, where the polarization pairing is $(e_i, f_j) = -(f_j, e_i) = \delta_{ij}$ and $(e_i, e_j) = (f_i, f_j) = 0$, and where $f_3, e_1, e_2$ project to a basis of $\text{Lie}(A) = M/V M$.

With an appropriate choice of the basis, the Frobenius $F$ on $M/pM$ is the $\phi$-linear map whose matrix with respect to the basis $f_3, e_1, e_2, f_1, f_2$ is

$$
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
\end{pmatrix}.
$$

(3.23)

All this can be deduced from [Bu-We], 3.2.

To construct the universal display we follow the method of [Go-O]. See also [An-Go]. With local coordinates $u$ and $v$ we write $\tilde{S} = Spf \tilde{\mathbb{F}}_p[[u, v]]$ for the formal completion of $S$ at $x$. We study the deformation of $F$ to

$$
F : H^1_{dR}(A^t/\tilde{S})^{(p)} \rightarrow H^1_{dR}(A^t/\tilde{S}).
$$

(3.24)

We use a basis $f_3, e_1, \ldots, f_2$ satisfying the same assumptions as above with respect to the $O_K$-type and the polarization pairing. Then one can choose $u$ and $v$ and the basis of $H^1_{dR}(A^t/\tilde{S})$ so that the universal Frobenius is given by the matrix

$$
F = \begin{pmatrix}
u^{p+1} & u^{p+1} \\
u & u^p \\
v & v^p \\
1 & 0 \\
1 & 0 \\
\end{pmatrix}.
$$

(3.25)

Since the first three vectors project onto a basis of $\text{Lie}(A)$, the matrix of $F : \text{Lie}(A)^{(p)} \rightarrow \text{Lie}(A)$ is the $3 \times 3$ upper left corner, and the matrix of $F^2 (= F \circ F^{(p)})$ is (note the semilinearity)

$$
\begin{pmatrix}
u^{p+1} & u^{p+1} & u^p \\
u & u^p & u^p \\
v & v^p & v^p \\
1 & 0 & 0 \\
1 & 0 & 0 \\
\end{pmatrix}.
$$

(3.26)

This matrix is sometimes called the Hasse-Witt matrix. Thus on $\text{Lie}(A)(\tilde{\Sigma})^{(p^2)} = \mathcal{L}^{(p^2)}$ the action of $F^2$ is given by multiplication by $u^{p+1} + v^{p+1}$. As the supersingular locus is the locus where the action of $F$ on the Lie algebra is nilpotent, it follows that the local (formal) equation of $S_{ss}$ at $x$ is

$$
u^{p+1} + v^{p+1} = 0.
$$

(3.27)

Note that this equation guarantees also that the lower $2 \times 2$ block, representing the action of $F^2$ on the $\Sigma$-part of the Lie algebras is (semi-linearly) nilpotent, i.e.

$$
\begin{pmatrix}
u^{p+1} & u^p \\
u & u^p \\
v^{p+1} & v^p \\
v^p & v^p \\
\end{pmatrix}
\begin{pmatrix}
u^{p+1} & u^p \\
v^{p+1} & v^p \\
v^{p+1} & v^p \\
v^{p+1} & v^p \\
\end{pmatrix} = 0.
$$

(3.28)

We write $\tilde{S}_{ss} = Spf(\tilde{\mathbb{F}}_p[[u, v]]/(u^{p+1} + v^{p+1}))$ for the formal completion of $S_{ss}$ at $x$. Letting $\zeta$ run over the $p + 1$ roots of $-1$ we recover the $p + 1$ formal branches through $x$ as the “lines”

$$
u = \zeta v.
$$

(3.29)
We write \( \hat{S}_{ss}[\zeta] = Spec(\mathbb{F}_p[[u, v]]/(u - \zeta v)) \) for this branch. When we restrict (pull back) the vector bundle \( \text{Lie}(A) \) to \( \hat{S}_{ss}[\zeta] \), \( \ker(F) \cap \text{Lie}(A)^{(p)} \) (the dual of \( P_\mu \)) becomes
\[
\ker \begin{pmatrix}
0 & \zeta v & v \\
\zeta v & 0 & 0 \\
v & 0 & 0
\end{pmatrix}.
\]
(3.30)

When \( v \neq 0 \) (i.e. outside the point \( x \)) this is the line
\[
\mathbb{F}_p \begin{pmatrix}
0 \\
1 \\
-\zeta
\end{pmatrix}
\]
(3.31)

As these lines are distinct, the filtration of \( P \) can not be extended across \( x \). \( \square \)

3.3. The Hasse invariant \( h_\Sigma \).

3.3.1. Definition of \( h_\Sigma \). The construction and main properties of the Hasse invariant that we are about to describe, have been given (for any unitary Shimura variety) by Goldring and Nicole in [Go-Ni]. Let \( R \) be an \( R_0/pR_0 = \mathbb{F}_p^2 \)-algebra. Let \( A \) be an abelian scheme over \( R \) and
\[
A^{(p)} = A \times_{\text{Spec}R, \phi} \text{Spec}R
\]
(3.32)
where \( \phi \) is the \( p \)-power map. If \( \alpha \) is an endomorphism of \( A \) then \( \alpha^{(p)} = \alpha \times 1 \) is an endomorphism of \( A^{(p)} \) and
\[
\alpha^{(p)} \circ \text{Frob}_{A/R} = \text{Frob}_{A/R} \circ \alpha.
\]
(3.33)

If \( \iota : \mathcal{O}_K \to \text{End}_R(A) \) is a ring homomorphism we define \( \iota^{(p)} : \mathcal{O}_K \to \text{End}_R(A^{(p)}) \) by
\[
\iota^{(p)}(a) = (a)^{(p)}.
\]
(3.34)

If, via \( \iota \), \( A \) has CM by \( \mathcal{O}_K \) and type \((2, 1)\), then \( A^{(p)} \) will have, via \( \iota^{(p)} \), type \((1, 2)\). The isogenies \( \text{Frob}_{A/R} \) and \( \text{Ver}_{A/R} \) induce \( R \)-linear maps \( (F_{A/R})_* \) and \( (V_{A/R})_* \) on Lie algebras. Since
\[
(\iota^{(p)}(a))_* = (F_{A/R})_* \circ (V_{A/R})_* = (F_{A/R})_* \circ (V_{A/R})_*
\]
(3.35)
(and similarly for \( (V_{A/R})_* \), \( (F_{A/R})_* \) and \( (V_{A/R})_* \) preserve types. As \( \text{Lie}(A^{(p)})(\Sigma) \) is 2-dimensional but \( \text{Lie}(A)(\Sigma) \) is 1-dimensional,
\[
(V_{A/R})_* : \text{Lie}(A^{(p)}) \to \text{Lie}(A)
\]
(3.36)
must have a kernel in \( \text{Lie}(A^{(p)})(\Sigma) \) which is at least one-dimensional. Consider
\[
V_*^2 := (V_{A/R})_* \circ (V_{A^{(p)}/R})_* : \text{Lie}(A)^{(p^2)} \to \text{Lie}(A)
\]
(3.37)
which again preserves types. We remark that under the identification of \( \text{Lie}(A) \) with \( H^1(A', \mathcal{O}) \), \( (V_{A/R})_* = (F_{A'/R})^* \).

Definition 3.1. The Hasse invariant is
\[
h_\Sigma = V_*^2(\Sigma) : \text{Lie}(A)(\Sigma)^{(p^2)} \to \text{Lie}(A)(\Sigma).
\]
(3.38)
Recall (Lemma 3.2) that if \( \mathcal{M} \) is a line bundle on a scheme \( S \) in characteristic \( p \), then \( \mathcal{M}^{(p)} = \mathcal{O}_S \otimes_{\mathcal{O}_S} \mathcal{M} \) is another line bundle over \( S \) and the map
\[
\mathcal{M}^{(p)} \ni 1 \otimes m \mapsto m \otimes \cdots \otimes m \in \mathcal{M}^{p}
\]
is an isomorphism of line bundles over \( S \). This applies in particular to the line bundle \( L \) on the Picard modular surface \( \overline{S} \) (over \( \mathbb{R}_0/p\mathbb{R}_0 \)).

We conclude that
\[
\hat{h}_{\Sigma} \in \text{Hom}_S(Lie(A)(\overline{\Sigma})^p, Lie(A)(\overline{\Sigma}))
\]
(3.40)
This means that the Hasse invariant is a modular form of weight \( p^2 - 1 \) in characteristic \( p \).

**Theorem 3.9.** The Hasse invariant is invertible on \( S_\mu \) and vanishes on \( S_{ss} = S - S_\mu \) to order one. More precisely, \( S_\mu \) is an open subscheme whose complement \( S_{ss} \) is a divisor, and when we endow this divisor with its induced reduced subscheme structure, it becomes the Cartier divisor \( \text{div}(h_{\Sigma}) \).

**Proof.** Dualizing the definition, the Hasse invariant vanishes precisely where \( V_A^{(p)}(\mathcal{L}) \) is contained in \( \ker(V_{A_{(p)}}^* : \mathcal{P}_{(p)} \to \mathcal{L}^{(p^2)}) \). We have already seen that over \( S_\mu \) the latter is the line bundle \( \mathcal{P}_{(p)} \) and that \( V^* \) sends \( \mathcal{L} \) isomorphically onto a direct complement of \( \mathcal{P}_{(p)}^{(p)} \), cf [3.18]. Since the Hasse invariant does not vanish on \( S_\mu \), then the restriction of \( h_{\Sigma} \) vanishes on \( S_{ss} \) to order 1 we must study the Dieudonné module at an infinitesimal neighborhood of a point \( x \in S_{ss} \) and compute \( V^*(p) \circ V^* \) using local coordinates there. This can be extracted from [Bu-We], but since later in our work we explain it in detail (for the purpose of studying the theta operator), we shall now refer to Section 3.3. In Lemma 4.3 we describe the (contravariant) Dieudonné module at \( x \). In subsection 3.3.3 we describe its infinitesimal deformation. Using the local coordinates \( u \) and \( v \), and the notation used there, \( f_3 - uf_1 - vf_2 \) becomes a basis for \( \mathcal{L} \) over the first infinitesimal neighborhood of \( x \). We then compute
\[
V^*(f_3 - uf_1 - vf_2) = e_1^{(p)} - ue_2^{(p)}
\]
(3.41)
\[
V^*(p)(e_1^{(p)} - ue_2^{(p)}) = -uf_3^{(p^2)} = -u \cdot (f_3 - uf_1 - vf_2)^{(p^2)}.
\]
Apart from working in the cotangent space instead of the tangent space, \( V^*(p) \circ V^* \) is the Hasse invariant (use \( \text{Hom}(M, N) = \text{Hom}(N^\vee, M^\vee) \)). It follows that after \( \mathcal{L} \) has been locally trivialized, the equation \( h_{\Sigma} = 0 \) becomes \( u = 0 \), which is the local equation for \( S_{ss} \).

\[ \square \]

3.3.2. Nonvanishing of \( h_{\Sigma} \) at the cusps.

**Proposition 3.10.** The Hasse invariant extends to a holomorphic section of \( \mathcal{L}^{p^2 - 1} \) over \( S \), which is nowhere vanishing on \( S_\mu \). If we trivialize \( \mathcal{L}|_C \) then the restriction of \( h_{\Sigma} \) to the cuspidal divisor \( C \) becomes a nowhere vanishing locally constant function.

**Proof.** Extendibility holds by the Koecher principle for any modular form. One can even deduce non-vanishing at the cusps, but here we may argue directly. The same definition as the one given over \( S \), with the abelian variety \( A \) replaced by \( \mathcal{G} = \mathcal{A}_{[p]} \) (\( A \) signifying now the semi-abelian variety over \( S \)), defines \( h_{\Sigma} \) over the complete Picard surface:
\[
h_{\Sigma} = (V_{\mathcal{G}})_* \circ (V_{\mathcal{G}(p)})_* : \text{Lie}(\mathcal{G})(\overline{\Sigma})^{(p^2)} \to \text{Lie}(\mathcal{G})(\overline{\Sigma}).
\]
(3.42)
The same argument as above shows that $h_{\Sigma}$ is nowhere vanishing on the whole of $S_{\mu}$. Since $L|_C$ is trivial, the last statement is obvious. □

**Corollary 3.11.** The scheme $S^*_\mu = S^* - S_{ss}$ is affine and over $\bar{\mathbb{F}}_p$, the intersection of $S_{ss}$ with every connected component of $S$ is connected.

**Proof.** The line bundle $L$ is ample on $S$, even over $\mathbb{F}_p$. Hence for large enough $m$, which we can take to be a multiple of $p^2 - 1$, $L^m$ is very ample, and by [La1] the Baily-Borel compactification $S^*$ is the closure of $S$ in the projective embedding supplied by the linear system $H^0(S, L^m)$. It follows that $L^m$ has an extension to a line bundle on $S^*$ which we denote $O_{S^*}(1)$, since it comes from the restriction of the $O(1)$ of the projective space to $S^*$. Moreover, Larsen proves that on the smooth compactification $S$, $L^m = \pi^* O_{S^*}(1)$ where $\pi : S \to S^*$.

Replacing $h_{\Sigma}$ by its power $h_{\Sigma}^{m/(p^2-1)}$, this power becomes a global section of $L^m$, hence its zero locus $S_{ss}$ a hyperplane section of $S^*$ in the projective embedding supplied by $H^0(S, L^m)$. Its complement is therefore affine. The last claim follows from the fact [Hart], III, 7.9, that a positive dimensional hyperplane section of a smooth (or more generally, normal) projective variety is connected. □

The scheme $S_{\mu}$ is of course far from affine, as it contains the complete curves $E$ as cuspidal divisors.

### 3.4. A secondary Hasse invariant on the supersingular locus.

In his forthcoming Ph.D. thesis [Bo], Boxer develops a general theory of secondary Hasse invariants defined on lower strata of Shimura varieties of Hodge type. See also [Kos]. In this section we provide an independent approach, in the case of Picard modular surfaces, affording a detailed study of its properties. As an application we relate the number of irreducible components of $S_{ss}$ to the Euler number of $S_\Sigma$, and through it to the value of the function $L(s, \frac{\omega}{2\pi})$ at $s = 3$.

#### 3.4.1. Definition of $h_{ssp}$.  
As we have seen, along the general supersingular locus $S_{gss}$, Verschiebung induces isomorphisms

$$V_{\Sigma} : L \simeq \mathcal{P}^{(p)}_D, \ V_P : \mathcal{P}_D \simeq \mathcal{L}^{(p)}.$$ (3.43)

(The first is unique to $S_{gss}$, the second holds also on the $\mu$-ordinary stratum.) Consider the isomorphism

$$V_p^{(p)} \otimes V_{\Sigma}^{-1} : \mathcal{P}_D^{(p)} \otimes \mathcal{P}_D^{(p)} \simeq \mathcal{L}^{(p^2)} \otimes \mathcal{L} \simeq L^{p^2+1}.$$ (3.44)

Its source is the line bundle $\det \mathcal{P}^{(p)}$ which is identified with $\mathcal{L}^{(p)} \simeq L^p$. We therefore get a nowhere vanishing section

$$\tilde{h}_{ssp} \in H^0(S_{gss}, \mathcal{L}^{p^2+1}).$$ (3.45)

Our “secondary” Hasse invariant is the nowhere vanishing section

$$h_{ssp} = \tilde{h}_{ssp}^{p+1} \in H^0(S_{gss}, \mathcal{L}^{p^2+1}).$$ (3.46)

---

4One way to see it is to use the ampleness of the Hodge bundle $\det \omega_A \simeq \mathcal{L}^2$ (pull back from Siegel space, where it is known to be ample by [Fa-Ch]).

5It is not clear that $\mathcal{L}$ itself has an extension to a line bundle on $S^*$, or that $\pi_* \mathcal{L}$, which is a coherent sheaf extending $L|_S$, is a line bundle (the problem lying of course only at the cusps). In other words, it is not clear that we can extract an $m$th root of $O_{S^*}(1)$ as a line bundle.
We shall show that $h_{\text{ssp}}$ extends to a holomorphic section on $S_{ss}$, and vanishes at the superspecial points (to a high order).

3.4.2. Computations at the superspecial points. We refer again to the computations of Proposition 3.3 and work over $\overline{\mathbb{F}}_p$. Dualizing the display described there, and using the letters $e_i, f_j$ to denote the dual basis to the basis used there we get the following.

**Lemma 3.12.** Let $x \in S_{ss}$ be a superspecial point. There exist formal coordinates $u$ and $v$ so that the formal completion of $S$ at $x$ is $\hat{S} = \text{Spf}(\overline{\mathbb{F}}_p[[u,v]])$, and $D = \text{Hom}(\mathcal{A}/\overline{\mathbb{S}})$ has a basis $f_3, e_1, e_2, e_3, f_1, f_2$ over $\overline{\mathbb{F}}_p[[u,v]]$ with the following properties:

(i) $f_3, e_1, e_2$ is a basis for $\omega_A$
(ii) The basis is symplectic, i.e. the polarization form is $\langle e_i, f_j \rangle = -\langle f_j, e_i \rangle = \delta_{ij}$, $\langle e_i, e_j \rangle = \langle f_i, f_j \rangle = 0$.
(iii) $\mathcal{O}_K$ acts on the $e_i$ via $\Sigma$ and on the $f_j$ via $\Sigma$.
(iv) $V : D \to D^{(p)}$ is given by $V f_3 = u e_1^{(p)} + ve_2^{(p)}$, $V e_1 = u f_3^{(p)}$, $V e_2 = v f_3^{(p)}$, $V e_3 = f_3^{(p)}$, $V f_1 = e_1^{(p)}$, $V f_2 = e_2^{(p)}$.

Using the lemma, we compute along $\hat{S}_{ss}[\zeta]$, where $u = \zeta v$ ($\zeta^{p+1} = -1$). Denote by $\mathcal{P}^{(p)}[\zeta]$, $\mathcal{P}_0^{(p)}[\zeta]$ and $\mathcal{L}[\zeta]$ the pull-backs of the corresponding vector bundles to the branch $\hat{S}_{ss}[\zeta]$. The map $V_{\zeta}$ is given by

$$f_3 \mapsto u e_1^{(p)} + ve_2^{(p)} = v \cdot (\zeta e_1^{(p)} + e_2^{(p)}) = v \cdot (\zeta^p e_1 + e_2)^{(p)} \in \mathcal{P}_0^{(p)}[\zeta].$$

Use $e_1 \wedge e_2 = e_1 \wedge (\zeta^p e_1 + e_2)$ as a basis for $\det \mathcal{P}[\zeta]$. Since $V_{\zeta}^{(p)}$ maps $e_1^{(p)}$ to $(\zeta^p)^{p} f_3^{(p)}$, $\tilde{h}_{ss}$ maps $e_1^{(p)} \wedge e_2^{(p)} = e_1^{(p)} \wedge (\zeta^p e_1 + e_2)^{(p)}$ to

$$\tilde{h}_{ss}(e_1^{(p)} \wedge e_2^{(p)}) = \zeta^p v^{p-1} f_3^{(p^2)} \otimes f_3 = \zeta^p v^{p-1} f_3^{p^2+1} = \zeta u^{p-1} f_3^{p^2+1}.$$

**Lemma 3.13.** There does not exist a function $g \in \overline{\mathbb{F}}_p[[u,v]]/(u^{p+1} + v^{p+1})$ on $\hat{S}_{ss}$ whose restriction to the branch $\hat{S}_{ss}[\zeta]$ is $\zeta u^{p-1}$.

**Proof.** Had there been such a function, represented by a power series $G \in \overline{\mathbb{F}}_p[[u,v]]$, then we would get $vG = u^p$ on $\hat{S}_{ss}[\zeta]$ for every $\zeta$, hence

$$vG - u^p \in (u^{p+1} + v^{p+1}) \subset \overline{\mathbb{F}}_p[[u,v]].$$

But any power series in the ideal $(u^{p+1} + v^{p+1})$ contains only terms of degree $\geq p+1$, while in $vG - u^p$ we can not cancel the term $u^p$. \hfill $\Box$

The lemma means that $\tilde{h}_{ss}$ can not be extended over $S_{ss}$ to a section of $\text{Hom}(\det \mathcal{P}^{(p)}, \mathcal{L}^{p^{2}+1}) \simeq \mathcal{L}^{p^{2}+1}$. However, when we raise it to a $p+1$ power the dependence on $\zeta$ disappears. It then extends to a section $h_{ss}$ of $\mathcal{L}^{p^3+1}$ over $S_{ss}$, given over $\hat{S}$ by the equation

$$h_{ss} = e u^{p^2-1} f_3^{p^2+1},$$

where $e \in \overline{\mathbb{F}}_p[[u,v]]^\times$ depends on the isomorphism between $\det \mathcal{P}$ and $\mathcal{L}$. 

**Theorem 3.14.** The secondary Hasse invariant $h_{ssp}$ belongs to $H^0(S_{ss}, \mathcal{L}^{p^3+1})$. It vanishes precisely at the points of $S_{ss}$. The subscheme “$h_{ssp} = 0$” of $S_{ss}$ is not reduced. At $x \in S_{ss}$, with $u$ and $v$ as above, it is the spectrum of

$$\bar{\mathbb{F}}_p[[u, v]]/(u^{p+1} + v^{p+1}, u^{p^2-1}, v^{p^2-1}).$$

From now on we assume that $N$ is large enough (depending on $p$) so that Theorem 3.1(iii) holds. Each irreducible component of $S_{ss}$ is non-singular, and $h_{ssp}$ has a zero of order $p^2 - 1$ at each superspecial point on such a component. Each component contains $p^3 + 1$ superspecial points. It follows that if $Z$ is such a component,

$$\text{deg}(\mathcal{L}^{p^3+1}|_Z) = \text{deg}(\text{div}_Z(h_{ssp})) = (p^3 + 1)(p^2 - 1).$$

We get the following corollary.

**Corollary 3.15.** Let $Z$ be an irreducible component of $S_{ss}$, and assume that $N$ is large enough. Then $\text{deg}(\mathcal{L}|_Z) = p^2 - 1$.

3.4.3. Application to the number of irreducible components of $S_{ss}$. We want to give an interesting application of the last corollary to the computation of the number of irreducible components of $S_{ss}$. Let

$$Z = \bigcup_{i=1}^{n} Z_i$$

be the decomposition into irreducible components of a single connected component $Z$ of $S_{ss}$ (recall that $Z$ is the intersection of $S_{ss}$ with a connected component of $\bar{S}$). The $Z_i$ are smooth, and their genus is $p(p-1)/2$.

**Theorem 3.16.** Let $c_2$ be the Euler characteristic of the connected component of $\bar{S}$ containing $Z$, i.e. if over $\mathbb{C}$ this connected component is $\bar{X}_\Gamma$ then

$$c_2 = \sum_{i=0}^{4} (-1)^i \dim \mathbb{C} H^i(\bar{X}_\Gamma, \mathbb{C}).$$

Then

$$3n = c_2.$$ 

**Proof.** Computing intersection numbers,

$$\text{(Z.Z)} = n(p^3 + 1)p + \sum_{i=1}^{n} (Z_i.Z_i).$$

Denote by $K_{\bar{S}}$ a canonical divisor on the given connected component of $\bar{S}$. From the adjunction formula,

$$p(p-1) - 2 = 2g(Z_i) - 2 = Z_i.(Z_i + K_{\bar{S}}).$$

As we have seen in Proposition 2.11 $\mathcal{O}(K_{\bar{S}} + C) \simeq \mathcal{L}^3$ where $C$ is the cuspidal divisor. Hence

$$(Z_i.K_{\bar{S}}) = Z_i.(K_{\bar{S}} + C) = \text{deg}(\mathcal{L}^3|_{Z_i}) = 3(p^2 - 1)$$

by the last Corollary. We get

$$(Z_i.Z_i) = -2p^2 - p + 1.$$ 

Plugging this into the expression for $(Z.Z)$ we get

$$(Z.Z) = n(p^2 - 1)^2.$$
On the other hand, since $Z$ is the divisor of the Hasse invariant on $\bar{S}$, $\text{div}(h_{\bar{Z}}) = Z$, we get $O(Z) = L^{p^2-1}$ so

$$n = c_1(L)^2.$$  

(3.61)

From this we get $9n = (K_{\bar{S}} + C)(K_{\bar{S}} + C)$. Holzapfel [Ho] (4.3.11') on p.184, implies

$$9n = 3c_2(X_\Gamma) = 3c_2(\bar{X}_\Gamma)$$

proving the theorem. \qed

In [Ho] (5A.4.3), p.325 Holzapfel proves that for $\Gamma$ the stabilizer of $L_0$

$$c_2(X_\Gamma) = \frac{3|D_K|^{5/2}}{32\pi^3} L \left(3, \left(\frac{D_K}{.} \right) \right) - \frac{3\varepsilon}{16} L \left(-2, \left(\frac{D_K}{.} \right) \right),$$

(3.63)

where $\varepsilon = 1$, unless $D_K = -3$, in which case $\varepsilon = 3$. For a congruence subgroup $\Gamma'$ we have $c_2(X_{\Gamma'}) = [\Gamma : \Gamma']c_2(X_\Gamma)$. Formulae of this form have been previously obtained, for split reductive groups, by Harder.

3.4.4. The classes of Ekedahl-Oort strata in the Chow ring of $\bar{S}$. Let us denote by $\text{CH} = \text{CH}(\bar{S})$ the Chow ring with $\mathbb{Q}$ coefficients of $\bar{S}$ over $\overline{\mathbb{F}}_p$, and let $\text{CH}_L$ the $\mathbb{Q}$-subalgebra generated by $c_1(L)$ (note that we now use $c_i$ to denote Chern classes of vector bundles in $\text{CH}$ and not in cohomology). Recall that the Ekedahl-Oort strata of $\bar{S}$ consist of $\bar{S}$ itself, $Z = S_{ss}$ in codimension 1, and $W = S_{ssp}$ in codimension 2.

**Proposition 3.17.** The classes $[\bar{S}], [Z]$ and $[W]$ in $\text{CH}$ belong to $\text{CH}_L$.

A similar result for Siegel modular varieties has been proved by van der Geer in [vdG] and for Hilbert modular varieties in [Go-O].

**Proof.** Without loss of generality we may assume that $N$ is large enough, so that Theorem [3.1(iii)] applies. As the divisor of the Hasse invariant is $Z$, $[Z] = (p^2 - 1)c_1(L)$. It remains to deal with $[W]$. Let $f : \bar{Z} \to \bar{S}$ be the map from the normalization of $\bar{Z}$ to $\bar{S}$ with image $Z$. Then the projection formula implies

$$[Z]c_1(L) = f_*(f^*(c_1(L))) = f_*(c_1(f^*L)).$$

(3.64)

Write $\bar{Z} = \coprod_{i=1}^n \bar{Z}_i$ and note that $(p^3 + 1)c_1(f^*L)$ is the class of $\text{div}(h_{ssp})$ on $\bar{Z}$, which is represented by the 0-cycle consisting of the superspecial points on each $\bar{Z}_i$ with multiplicities $p^2 - 1$. It follows that

$$(p^3 + 1)[Z]c_1(L) = (p^2 - 1)(p + 1)[W].$$

(3.65)

Substituting $[Z] = (p^2 - 1)c_1(L)$ we get

$$[W] = (p^2 - p + 1)c_1(L)^2.$$  

(3.66)

\qed

3.5. The open Igusa surfaces.
3.5.1. The Igusa scheme. Let $N \geq 3$ as always, and let $\mathcal{M}$ be the moduli problem of Section 3.3.1. Let $n \geq 1$ and consider the following moduli problem on $R_0/pR_0$-algebras:

- $\mathcal{M}_{Ig(p^n)}(R)$ is the set of isomorphism classes of triples $(A, A[p^n]^\mu, \varepsilon)$ where $A \in \mathcal{M}(R)$, $A[p^n]^\mu \subset A[p^n]$ is a finite flat $O_K$-subgroup scheme of rank $p^{2n}$ of multiplicative type, and

$$(3.67) \quad \varepsilon : \delta_K^{-1}O_K \otimes \mu_{p^n} \simeq A[p^n]^\mu$$

is an isomorphism of $O_K$-group schemes over $R$.

It is clear that if $(A, A[p^n]^\mu, \varepsilon) \in \mathcal{M}_{Ig(p^n)}(R)$ then $A$ is fiber-by-fiber $\mu$-ordinary and therefore $A \in \mathcal{M}(R)$ defines an $R$-point of $S_\mu$. It is also clear that the functor $R \mapsto \mathcal{M}_{Ig(p^n)}(R)$ is relatively representable over $\mathcal{M}$, and therefore as $N \geq 3$ and $\mathcal{M}$ is representable, this functor is also representable by a scheme $Ig_\mu(p^n)$ which maps to $S_\mu$. See [Ka-Ma] for the notion of relative representability. We call $Ig_\mu(p^n)$ the Igusa scheme of level $p^n$.

**Proposition 3.18.** The morphism $\tau : Ig_\mu(p^n) \to S_\mu$ is finite and étale, with the Galois group $\Delta(p^n) = (O_K/p^nO_K)^\times$ acting as a group of deck transformations.

**Proof.** Every $\mu$-ordinary abelian variety has a unique finite flat $O_K$-subgroup scheme of multiplicative type $A[p^n]^\mu$ of rank $p^{2n}$. Such a subgroup scheme is, locally in the étale topology, isomorphic to $\delta_K^{-1}O_K \otimes \mu_{p^n}$, and any two isomorphisms differ by a unique automorphism of $\delta_K^{-1}O_K \otimes \mu_{p^n}$. But $\Delta(p^n) = Aut_{O_K}(\delta_K^{-1}O_K \otimes \mu_{p^n})$. If we let $\gamma \in \Delta(p^n)$ act on the triple $(A, A[p^n]^\mu, \varepsilon)$ via

$$(3.68) \quad \gamma((A, A[p^n]^\mu, \varepsilon)) = (A, A[p^n]^\mu, \varepsilon \circ \gamma^{-1})$$

$\Delta(p^n)$ becomes a group of deck transformation and the proof is complete. \qed

3.5.2. A compactification over the cusps. The proof of the following proposition mimics the construction of $\overline{S}$. We omit it.

**Proposition 3.19.** Let $\overline{Ig_\mu}(p^n)$ be the normalization of $\overline{S}_\mu = \overline{S} - S_{ss}$ in $Ig_\mu(p^n)$. Then $\overline{Ig_\mu}(p^n) \to S_\mu$ is finite étale and the action of $\Delta(p^n)$ extends to it. The boundary $\overline{Ig_\mu}(p^n) - Ig_\mu(p^n)$ is non-canonically identified with $\Delta(p^n) \times C$.

We define similarly $\overline{Ig_\mu}$, and note that it is finite étale over $S_\mu^*$.

**Proposition 3.20.** Let $\mathcal{A}$ denote the pull-back of the universal semi-abelian variety from $\overline{S}_\mu$ to $\overline{Ig_\mu}(p^n)$. Then $\mathcal{A}$ is equipped with a canonical Igusa level structure

$$(3.69) \quad \varepsilon : \delta_K^{-1}O_K \otimes \mu_{p^n} \simeq A[p^n]^\mu.$$ 

Over $C$ and after base change to $R_N/pR_N$ the toric part of $\mathcal{A}$ is locally Zariski of the form $a \otimes \mathbb{G}_m$ and $\varepsilon$ is then an $O_K$-linear isomorphism between $\delta_K^{-1}O_K \otimes \mu_{p^n}$ and $a \otimes \mu_{p^n}$.

3.5.3. A trivialization of $\mathcal{L}$ over the Igusa surface. From now on we focus on $\overline{Ig_\mu} = \overline{Ig_\mu}(p)$ although similar results hold when $n > 1$, and would be instrumental in the study of $p$-adic modular forms. The vector bundle $\omega_{\mathcal{A}}$ pulls back to a similar vector bundle over $\overline{Ig_\mu}$. But there

$$(3.70) \quad \omega^\mu_{\mathcal{A}} := \omega_{\mathcal{A}[p]^n}$$
is a rank 2 quotient bundle stable under \(O_K\) (of type \((1,1)\)), and the isomorphism \(\varepsilon\) induces an isomorphism

\[
\varepsilon^*: \omega_{\mathcal{A}}^* \simeq \omega_{\mathcal{K}}^{-1} \otimes \mu_p^1.
\]

Now \(\operatorname{Lie}(\delta_{\mathcal{K}}^{-1} O_K \otimes \mu_p) = \delta_{\mathcal{K}}^{-1} O_K \otimes \operatorname{Lie}(\mu_p) = \delta_{\mathcal{K}}^{-1} O_K \otimes \operatorname{Lie}(\mathcal{G}_m)\) and by duality

\[
\omega_{\delta_{\mathcal{K}}^{-1} O_K \otimes \mu_p} = O_K \otimes \omega_{\mathcal{G}_m}.
\]

with \(1 \otimes dT/T\) as a generator (if \(T\) is the parameter of \(\mathcal{G}_m\)). Here we have used the fact that the \(\mathbb{Z}\)-dual of \(\delta_{\mathcal{K}}^{-1} O_K\) is \(O_K\) via the trace pairing. This is the constant vector bundle \(O_K \otimes R = R(\Sigma) \oplus R(\bar{\Sigma})\).

**Proposition 3.21.** The line bundles \(\mathcal{L}, \mathcal{P}_0\) and \(\mathcal{P}_\mu\) are trivial over \(\overline{\text{Ig}_\mu}\).

**Proof.** Use \(\varepsilon^*\) as an isomorphism between vector bundles and note that \(\mathcal{L} = \omega_{\mathcal{A}}^*(\bar{\Sigma})\) and \(\mathcal{P}_\mu = \omega_{\mathcal{A}}^*(\Sigma)\). The relation \(\mathcal{P}_0 \otimes \mathcal{P}_\mu = \det \mathcal{P} \simeq \mathcal{L}\) implies the triviality of \(\mathcal{P}_0\) as well.

Note that the trivialization of \(\mathcal{L}\) and \(\mathcal{P}_\mu\) is canonical, because it uses only the tautological map \(\varepsilon\), which exists over the Igusa scheme. The trivialization of \(\mathcal{P}_0\) on the other hand depends on how we realize the isomorphism \(\det \mathcal{P} \simeq \mathcal{L}\).

We can now give an alternative proof to the fact that \(\mathcal{L}^{p^2-1}\) and \(\mathcal{P}_\mu^{p^2-1}\) are trivial on \(\overline{\mathcal{S}}\). Since \(\overline{\text{Ig}_\mu}\) is an étale cover of \(\overline{\mathcal{S}}\) of order \(p^2 - 1\), \(\operatorname{det} \tau^*(\mathcal{L}) \simeq \mathcal{L}^{p^2-1}\). As \(\tau^* \mathcal{L}\) is already trivial, so is \(\mathcal{L}^{p^2-1}\) on the base. The same argument works for \(\mathcal{P}_\mu\) and \(\mathcal{P}_0\). The fact that \(\mathcal{P}_0^{p^2+1}\) is already trivial could be deduced by a similar argument had we worked out an analogue of \(\text{Ig}(p)\) classifying symplectic isomorphisms of \(\mathfrak{g}[p]\) with \(\operatorname{gr}^1 A[p]\). The role of \(\Delta(p)\) for such a moduli space would be assumed by

\[
\Delta^1(p) = \ker(N: (O_K/pO_K) \times \rightarrow \mathbb{F}_p^\times),
\]

which is a group of order \(p + 1\). We do not go any further in this direction here.

### 3.6. Compactification of the Igusa surface along the supersingular locus.

#### 3.6.1. Extracting a \(p^2 - 1\) root from \(h_\Sigma\) over \(\overline{\text{Ig}_\mu}\). Let \(a(1)\) be the canonical nowhere vanishing section of \(\mathcal{L}\) over \(\overline{\text{Ig}_\mu}\), which is sent to \(\varepsilon_\Sigma \cdot (1 \otimes dT/T)\) under the trivialization

\[
\varepsilon^*: \mathcal{L} = \omega_{\mathcal{A}}^*(\bar{\Sigma}) \simeq (O_K \otimes \omega_{\mathcal{G}_m})(\bar{\Sigma}) = R(\bar{\Sigma}).
\]

Here \(R\) is any \(R_0/pR_0\)-algebra over which we choose to work. In other words, \(a(1) = (\varepsilon^*)^{-1}(\varepsilon_\Sigma \cdot 1 \otimes dT/T)\). Dually, \(a(1)\) is the homomorphism from \(\operatorname{Lie}(\mathcal{A})(\bar{\Sigma})\) to \(\delta_{\mathcal{K}}^{-1} \otimes \operatorname{Lie}(\mathcal{G}_m)(\bar{\Sigma})\) arising from \(\varepsilon^{-1}\). Let \(a(k) = a(1)^{\otimes k} \in H^0(\overline{\text{Ig}_\mu}, \mathcal{L}^k)\).

**Proposition 3.22.** (i) Let \(\gamma \in \Delta(p) = (O_K/pO_K) \times\). Then \(\Delta(p)\) acts on \(H^0(\overline{\text{Ig}_\mu}, \mathcal{L})\) and

\[
\gamma^* a(1) = \bar{\Sigma}(\gamma)^{-1} \cdot a(1).
\]

(ii) The section \(a(1)\) is a \(p^2 - 1\) root of the Hasse invariant over \(\overline{\text{Ig}_\mu}\), i.e.

\[
a(p^2 - 1) = h_\Sigma.
\]
Proof. (i) This part is a restatement of the action of $\Delta(p)$. At two points of $Ig_{\mu}(R)$ lying over the same point of $S_{\mu}(R)$ and differing by the action of $\gamma \in \Delta(p)$, the canonical embeddings

$$\delta^{-1}_K \otimes \mu_p \hookrightarrow A[p]$$

(3.77)

differ by $\iota(\gamma)$ (ii) The induced trivializations of $Lie(A)(\Sigma)$ differ by $\Sigma(\gamma)$ and by duality we get (i).

(ii) Since over any $\mathbb{F}_p$-base, $Ver_{G_m} = 1$, we have a commutative diagram

$$\begin{array}{ccc}
Lie(A)(\Sigma)(p^2) & \overset{\iota}{\leftarrow} & Lie(A)(\Sigma) \\
\downarrow a(1)(p^2) & & \downarrow a(1) \\
\delta^{-1}_K \otimes Lie(G_m)(\Sigma) & = & \delta^{-1}_K \otimes Lie(G_m)(\Sigma)
\end{array}$$

(3.78)

Using the isomorphism $Lie(A)(\Sigma)(p^2) \simeq Lie(A)(\Sigma)^{p^2}$ to which we alluded before, we get the commutative diagram

$$\begin{array}{ccc}
Lie(A)(\Sigma)^{p^2} & \overset{h_\Sigma}{\leftarrow} & Lie(A)(\Sigma) \\
\downarrow a(p^2) & & \downarrow a(1) \\
\delta^{-1}_K \otimes Lie(G_m)(\Sigma) & = & \delta^{-1}_K \otimes Lie(G_m)(\Sigma)
\end{array}$$

(3.79)

from which we deduce that $h_\Sigma = a(p^2 - 1)$. Note that we can interprete $a(1)$ as an arrow as above rather than as a section of $\mathcal{L} = Lie(A)(\Sigma)^{p^2}$ because of the identity $Hom(M, N) = Hom(N^\vee, M^\vee)$. \hfill \Box

3.6.2. The compactification $\overline{\text{Th}}$ of $\overline{Ig}_{\mu}$. Quite generally, let $L \to X$ be a line bundle associated with an invertible sheaf $\mathcal{L}$ on a scheme $X$. Write $L^n$ for the line bundle $L^{\otimes n}$ over $X$. Let $s : X \to L^n$ be a section. Consider the fiber product

$$Y = L \times_{L^n} X$$

(3.80)

where the two maps to $L^n$ are $\lambda \mapsto \lambda^n$ and $s$. Let $p : Y \to L^n$ be the projection which can also factor as $Y \overset{\iota}{\rightarrow} L \to L^n \to X$ (since $X \overset{\iota}{\rightarrow} L^n \to X$ is the identity). Consider

$$p^*L = L \times_X (L \times_{L^n} X).$$

(3.81)

This line bundle on $Y$ has a tautological section $t : Y \to p^*L,$

$$t : y = (\lambda, x) \mapsto (\lambda, y) = (\lambda, (\lambda, x))$$

(3.82)

Here $s(x) = \lambda^n$ and

$$t^n(y) = (\lambda^n, y) = (s(x), y) = p^*s(y)$$

(3.83)

so $t$ is an $n$th root of $p^*$s. Moreover, $Y$ has the universal property with respect to extracting $n$th roots from $s$: If $p_1 : Y_1 \to X$, and $t_1 \in \Gamma(Y_1, p_1^*L)$ is such that $t_1^n = p_1^*s$, then there exists a unique morphism $h : Y_1 \to Y$ covering the two maps to $X$ such that $t_1 = h^*t$.

The map $L \to L^n$ is finite flat of degree $n$ and if $n$ is invertible on the base, finite étale away from the zero section. Indeed, locally on $X$ it is the map $\mathbb{A}^1 \times X \to \mathbb{A}^1 \times X$ which is just raising to $n$th power in the first coordinate. By base-change, it follows that the same is true for the map $p : Y \to X$ : this map is finite flat of degree $n$ and étale away from the vanishing locus of the section $s$ (assuming $n$ is invertible). We remark that if $L$ is the trivial line bundle, we recover usual Kummer theory.
Applying this in our example with \( n = p^2 - 1 \) we define the complete Igusa surface of level \( p \), \( \overline{T_\ell} = T_\ell(p) \) as
\[
\overline{T_\ell} = \mathcal{L} \times_{\mathcal{L}^{p^2-1}} \bar{S}
\]
where the map \( \mathcal{S} \to \mathcal{L} \) is \( h_\eta \). From the universal property and part (ii) of Proposition \( \text{[7.22]} \) we get a map of \( \bar{S} \)-schemes
\[
T_{g, \mu} \to \bar{T}_g.
\]
This map is an isomorphism over \( \bar{S}_\mu \) because both schemes are étale torsors for \( \Delta(p) = (\mathcal{O}_K/p\mathcal{O}_K)^\times \) and the map respects the action of this group. We summarize the discussion in the following theorem (for the last point, consult [Mu2], Proposition 2, p.198).

**Theorem 3.23.** The morphism \( \tau : \overline{T_g} \to \bar{S} \) satisfies the following properties:

(i) It is finite flat of degree \( p^2 - 1 \), étale over \( \bar{S}_\mu \), totally ramified over \( S_{ss} \).

(ii) \( \Delta \) acts on \( \overline{T_g} \) as a group of deck transformations and the quotient is \( \bar{S} \).

(iii) \( \Delta \) acts on \( \overline{T_g} \) as a group of deck transformations and the quotient is \( \bar{S} \).

(iv) \( \Delta \) acts on \( \overline{T_g} \) as a group of deck transformations and the quotient is \( \bar{S} \).

In particular, \( \bar{S}_\mu \) is regular in codimension 1.

3.6.3. Irreducibility of \( I_g \). So far we have avoided the delicate question of whether \( \overline{T_g} \) is "relatively irreducible", i.e. whether \( \tau^{-1}(T) \) is irreducible if \( T \subset \bar{S} \) is an irreducible (equivalently, connected) component. Using an idea of Katz, and following the approach taken by Ribet in [Ri], the irreducibility of \( \tau^{-1}(T) \) could be proven for any level \( p^n \) if we could prove the following:

- Let \( q = p^2 \). For any \( r \) sufficiently large and for any \( \gamma \in (\mathcal{O}_K/p^n\mathcal{O}_K)^\times \) there exists a \( \mu \)-ordinary abelian variety with PEL structure \( \mathcal{A} \in S_{ss}(\mathbb{F}_q) \) such that the image of \( \text{Gal}(\mathbb{F}_q/\mathbb{F}_q') \) in
\[
\text{Aut} \left( \text{Isom}_{\mathbb{F}_q}(\delta_{K}^{-1} \otimes \mu_{p^{n}}, \mathcal{A}[p^n]^{\mu}) \right) = (\mathcal{O}_K/p^n\mathcal{O}_K)^\times
\]
contains \( \gamma \).

Instead, we shall give a different argument valid for the case \( n = 1 \).

**Proposition 3.24.** The morphism \( \tau : \overline{T_g} \to \bar{S} \) induces a bijection on irreducible components.

**Proof.** Since \( \overline{T_g} \) is a normal surface, connected components and irreducible components are the same. Let \( T \) be a connected component of \( \bar{S} \) and \( T_{ss} = T \cap S_{ss} \). Let \( \tau^{-1}(T) = \bigsqcup Y_i \) be the decomposition into connected components. As \( \tau \) is finite and flat, each \( \tau(Y_i) = T \). Since \( \tau \) is totally ramified over \( T_{ss} \), there is only one \( Y_i \). \( \square \)

4. Modular forms modulo \( p \)

4.1. Modular forms \( \mod p \) as functions on \( I_g \).
4.1.1. Representing modular forms by functions on $Ig$. The Galois group $\Delta(p) = (\mathcal{O}_k/p\mathcal{O}_k)^\times$ acts on the coordinate ring $H^0(Ig_\mu, \mathcal{O})$ and we let $H^0(Ig_\mu, \mathcal{O})^{(k)}$ be the subspace where it acts via the character $\bar{\Sigma}^k$. Then

\begin{equation}
H^0(Ig_\mu, \mathcal{O}) = \bigoplus_{k=0}^{p^2-2} H^0(Ig_\mu, \mathcal{O})^{(k)}
\end{equation}

and each $H^0(Ig_\mu, \mathcal{O})^{(k)}$ is free of rank 1 over $H^0(S_\mu, \mathcal{O}) = H^0(Ig_\mu, \mathcal{O})^{(0)}$.

For any $0 \leq k$ the map $f \mapsto f/a(k)$ is an embedding

\begin{equation}
M_k(N, R_0/pR_0) \to H^0(Ig_\mu, \mathcal{O})^{(k)}.
\end{equation}

**Lemma 4.1.** Fix $0 \leq k < p^2 - 1$. Then we have a surjective homomorphism

\begin{equation}
\bigoplus_{n \geq 0} M_{k+n(p^2-1)}(N, R_0/pR_0) \to H^0(Ig_\mu, \mathcal{O})^{(k)}.
\end{equation}

**Proof.** Take $f \in H^0(Ig_\mu, \mathcal{O})^{(k)}$, so that $f \cdot a(k) \in H^0(Ig_\mu, \mathcal{O})^{(0)}$, hence descends to $g \in H^0(S_\mu, \mathcal{L}^k)$. This $g$ may have poles along $S_{ss}$, but some $h_{\Sigma}^n g$ will extend holomorphically to $S$, hence represents a modular form of weight $k + n(p^2 - 1)$, which will map to $f$ because $a(k + n(p^2 - 1)) = h_{\Sigma}^n a(k)$. \hfill $\square$

**Proposition 4.2.** The resulting ring homomorphism

\begin{equation}
r : \bigoplus_{k \geq 0} M_k(N, R_0/pR_0) \to H^0(Ig_\mu, \mathcal{O})
\end{equation}

obtained by dividing a modular form of weight $k$ by $a(k)$ is surjective, respects the $\mathbb{Z}/(p^2-1)\mathbb{Z}$-grading on both sides, and its kernel is the ideal generated by $h_{\Sigma} - 1$.

**Proof.** We only have to prove that anything in $\ker(r)$ is a multiple of $h_{\Sigma} - 1$, the rest being clear. Since $r$ respects the grading, we may assume that for some $k \geq 0$ we have $f_j \in M_{k+j(p^2-1)}(S, R_0/pR_0)$ and

\begin{equation}
\sum_{j=0}^m a(k)^{-1} h_{\Sigma}^{-j} f_j = 0.
\end{equation}

But then $f_m = -h_{\Sigma}^m \left( \sum_{j=0}^{m-1} h_{\Sigma}^{-j} f_j \right)$, so $\sum_{j=0}^m f_j = \sum_{j=0}^{m-1} (1 - h_{\Sigma}^{m-j}) f_j$ belongs to $(1 - h_{\Sigma})$. \hfill $\square$

As a result we get that

\begin{equation}
Ig_\mu^* = \text{Spec} \left( \bigoplus_{k \geq 0} M_k(N, R_0/pR_0)/(h_{\Sigma} - 1) \right)
\end{equation}

and

\begin{equation}
S_\mu^* = \text{Spec} \left( \bigoplus_{k \geq 0} M_{k(p^2-1)}(N, R_0/pR_0)/(h_{\Sigma} - 1) \right).
\end{equation}
4.1.2. Fourier-Jacobi expansions modulo $p$. The arithmetic Fourier-Jacobi expansion (2.48) depended on a choice of a nowhere vanishing section $s$ of $L$ along the boundary $C = \tilde{S} - S$ of $\tilde{S}$. As the boundary $\tilde{C} = \overline{Tg_{\mu} - Ig_{\mu}}$ is (non-canonically) identified with $\Delta(p) \times C$, we may “compute” the Fourier-Jacobi expansion on the Igusa surface rather than on $S$. But on the Igusa surface, $a(1)$ is a canonical choice for such an $s$. We may therefore associate a canonical Fourier-Jacobi expansion

\begin{equation}
F J(f) = \sum_{m=0}^{\infty} c_m(f) \in \prod_{m=0}^{\infty} H^0(\tilde{C}, N^m)
\end{equation}

along the boundary of $Ig$, for every

\begin{equation}
f \in M_k(N, R) = \bigoplus_{k=0}^{\infty} M_k(N, R)
\end{equation}

($R$ an $R_0/pR_0$-algebra). The following proposition becomes almost a tautology.

**Proposition 4.3.** The Fourier-Jacobi expansion $F J(h_{ss})$ of the Hasse invariant is 1. Moreover, for $f_1$ and $f_2$ in the graded ring $M_s(N, R)$, $r(f_1) = r(f_2)$ if and only if $F J(f_1) = F J(f_2)$.

**Proof.** The first statement is tautologically true. For the second, note that for $f \in M_k(N, R)$, $F J(f)$ is the (expansion of the) image of $f/a(k)$ in $H^0(\tilde{C}, O_{\tilde{I}g})$ where $\tilde{I}g$ is the formal completion of $Ig$ along $\tilde{C}$, while $r(f)$ is the image of $f/a(k)$ in $H^0(\overline{Tg_{\mu}}, O)$. The proposition follows from the fact that by Proposition 3.24 the irreducible components of $\overline{Tg_{\mu}}$ are in bijection with the connected components of $\tilde{S}$, so every irreducible component of $\overline{Tg_{\mu}}$ contains at least one cuspidal component (“$q$-expansion principle”). A function on $\overline{Tg_{\mu}}$ that vanishes in the formal neighborhood of any cuspidal component must therefore vanish on any irreducible component, so is identically 0. \hfill \Box

4.1.3. The filtration of a modular form modulo $p$. Let $f \in M_k(N, R)$, where $R$ is an $R_0/pR_0$-algebra as before. Define the filtration $\omega(f)$ to be the minimal $j$ such that $r(f) = r(f')$ (equivalently $F J(f) = F J(f')$) for some $f' \in M_j(N, R)$. The following proposition follows immediately from previous results.

**Proposition 4.4.** Let $f \in M_k(N, R)$. Then $\omega(f) \leq k$ and

\begin{equation}
\omega(f) \equiv k \mod (p^2 - 1).
\end{equation}

Let $\omega(f) = k - (p^2 - 1)n$. Then $n$ is the order of vanishing of $f$ along $S_{ss}$. Equivalently, $k - \omega(f)$ is the order of vanishing of the pull-back of $f$ to $Ig$ along $Ig_{ss}$. In addition, $\omega(f^m) = m\omega(f)$.

4.2. The theta operator.

4.2.1. Definition of $\Theta(f)$. From now on we work over $\overline{F}_p$. We first recall some notation. Let $S$ be the (open) Picard surface over $\overline{F}_p$ and $Ig = Ig(p)$ the Igusa surface of level $p$ over $S$ (completed along the supersingular locus as explained above). We denote by $Z = S_{ss} = S - S_p$ the supersingular locus of $S$, by $\tilde{Z} = Ig_{ss} = Ig - Ig_{\mu}$ its pre-image under the covering map $\tau : Ig \to S$, by $Z' = S_{ss} = S_{ss} - S_{ssp}$ the smooth part of $Z$, and by $\tilde{Z}' = Ig_{gss} = Ig_{ss} - Ig_{ssp}$ the pre-image of $Z'$ under $\tau$. When we need to compactify these schemes at the cusps, we let $\tilde{S}$ and $\overline{Tg_{\mu}}$ stand for the smooth compactifications. By the Koecher principle, the space of $\mod p$...
modular forms of weight \( k \geq 0 \) and level \( N \) can be regarded as sections of \( \mathcal{L}^k \) over either \( S \) or \( \tilde{S} \):

\[
M_k(N, \tilde{\mathbb{F}}_p) = H^0(S, \mathcal{L}^k) = H^0(\tilde{S}, \mathcal{L}^k).
\]

We recall (Theorem 3.23) that \( \tau : \tilde{\mathbb{F}}_p \rightarrow \tilde{S} \) is finite Galois of degree \( p^2 - 1 \), that it is étale outside \( Z \), and fully ramified over \( Z \). The Galois group of the covering may be canonically identified with

\[
\Delta = (\mathcal{O}_K/p\mathcal{O}_K)^\times.
\]

If \( M \) is an \( \tilde{\mathbb{F}}_p[\Delta] \) module, and \( \chi : \Delta \rightarrow \tilde{\mathbb{F}}_p^\times \) a character, we let \( M^\chi \) be the submodule on which \( \Delta \) acts via \( \chi \). We continue to denote by \( \mathcal{A} \) the universal abelian scheme over \( Ig \) (or its extension to a semi-abelian scheme over \( \tilde{Tg} \)) and by \( \mathcal{P} \) and \( \mathcal{L} \) the \( \Sigma \)- and \( \bar{\Sigma} \)-parts of \( \omega_{\mathcal{A}/Ig} \). These vector bundles are just the base-change by \( \tau^* \) of their counterparts over \( S \). We let \( a(1) \) be the canonical section of \( \mathcal{L} \) over \( \tilde{Tg} \), trivializing \( \mathcal{L} \) over \( \tilde{Tg}_\mu = \tilde{Tg} - \tilde{Z} \), and vanishing to first order along \( \tilde{Z} \). The Galois group \( \Delta \) acts on \( a(1) \) via \( \Sigma^{-1} \).

Let \( f \in H^0(S, \mathcal{L}^k) \). Let \( g = r(f) = \tau^* f/a(1)^k \in H^0(Ig_\mu, \mathcal{O}) \). This function has a pole of order \( k \) along \( \tilde{Z} \), and the Galois group acts on it via \( \Sigma^k \). Let \( dg \in H^0(Ig_\mu, \Omega_{Ig}^{1}) = H^0(Ig_\mu, \tau^* \Omega_S^{1}) \). The Kodaira-Spencer isomorphism \( KS(\Sigma) \) is an isomorphism

\[
KS(\Sigma) : \mathcal{P} \otimes \mathcal{L} \cong \Omega_{S}^{1}.
\]

Write \( KS(\Sigma)^{-1} \) for its inverse: \( \Omega_{S}^{1} \cong \mathcal{P} \otimes \mathcal{L} \). Let

\[
\kappa = (V \otimes 1) \circ KS(\Sigma)^{-1} : \Omega_{S}^{1} \rightarrow \mathcal{L}^{(p)} \otimes \mathcal{L} \cong \mathcal{L}^{p+1}
\]

be the map obtained from \( KS(\Sigma)^{-1} \) when we apply \( V \) to \( \mathcal{P} \). Over \( S - S_{ssp} \) this is the same as dividing out by the line sub-bundle \( \mathcal{P}_0 = \ker V_\mathcal{P} \), since

\[
\mathcal{P}/\mathcal{P}_0 = \mathcal{P}_\mu \cong V(\mathcal{P}) = \mathcal{L}^{(p)}.
\]

We denote by \( \kappa \) also the map induced on the base-change of these vector bundles by \( \tau^* \) to \( Ig - Ig_{ssp} \), and consider \( \kappa(dg) \). As \( \Delta \) still acts on \( \kappa(dg) \) via \( \Sigma^k \), its action on \( a(1)^k \kappa(dg) \) is trivial, so it descends to \( S_{ssp} \). We define

\[
\Theta(f) = a(1)^k \kappa(dg) \in H^0(S_{ssp}, \mathcal{L}^{k+p+1}).
\]

A priori, this extends only to a meromorphic modular form of weight \( k + p + 1 \), as it may have poles along \( Z \).

4.2.2. The main theorem. For the formulation of the next theorem we need to define what we mean by the standard cuspidal component of \( S \) or \( \tilde{Tg} \). Recall that according to [Bel] and [La1] the cuspidal scheme \( C = \tilde{S} - S \) classifies \( \mathcal{O}_K \)-semi-abelian varieties with level \( N \) structure. The standard component of \( C \) is the component (over \( \mathbb{C} \)) which classifies extensions of the elliptic curve \( \mathbb{C}/\mathcal{O}_K \) by the \( \mathcal{O}_K \)-torus \( \mathcal{O}_K \otimes \mathbb{C}^\times \) (thus sits over a cusp of type \( (\mathcal{O}_K, \mathcal{O}_K) \) in \( S^2 \)), together with a level-\( N \)-structure \( (\alpha, \beta, \gamma) \) (see [Bel], I.4.2 and Section 1.6.2), where

\[
\alpha : \mathcal{O}_K/N\mathcal{O}_K = \mathcal{O}_K \otimes \mathbb{Z}/N\mathbb{Z} \rightarrow \mathcal{O}_K \otimes \mathbb{C}^\times
\]

is given by \( 1 \otimes (a \mapsto \exp(2\pi ia/N)) \) and

\[
\beta : \mathcal{O}_K/N\mathcal{O}_K = N^{-1}\mathcal{O}_K/\mathcal{O}_K \rightarrow \mathbb{C}/\mathcal{O}_K
\]

is the canonical embedding. (The splitting \( \gamma \) varies along the component.) The standard component of \( C \) over \( R_N \) is the one which becomes this component after
base change to \( \mathbb{C} \) (our convention is that all number fields are contained in \( \mathbb{C} \)). Let \( P \) be a prime of \( R_N \) above \( p \) at which we reduce the Picard surface. The standard component of \( C \) over \( R_N/P R_N \) is the reduction modulo \( P \) of the standard component of \( C \) over \( R_N \). Finally, \( \overline{Tg} \) maps to \( \overline{S} \) (over \( R_N/P R_N \)) and the cuspidal components mapping to a given component \( E \) of \( C \) are classified by the embedding of \( \delta_K^{-1} O_K \otimes \mu_p \) in the toric part of \( A \). Since the toric part of the universal semi-abelian variety over the standard component is \( O_K \otimes \mathbb{G}_m \), we may define the standard cuspidal component of \( \overline{Tg} \) to be the component where the map

\[
\varepsilon : \delta_K^{-1} O_K \otimes \mu_p \rightarrow O_K \otimes \mathbb{G}_m
\]

is the natural embedding. Here we use the fact that \( \delta_K \) is invertible in \( O_K/pO_K \). Let \( \tilde{E} \subset \tilde{C} = \overline{Tg} - Ig \) be this standard component.

**Theorem 4.5.** (i) The operator \( \Theta \) maps \( H^0(S, \mathcal{L}^k) \) to \( H^0(S, \mathcal{L}^{k+p+1}) \).

(ii) The effect of \( \Theta \) on Fourier-Jacobi expansions is a “Tate twist”. More precisely, let

\[
FJ(f) = \sum_{m=0}^{\infty} c_m(f)
\]

be the Fourier-Jacobi expansion of \( f \) along \( \tilde{E} \) (thus \( c_m(f) \in H^0(\tilde{E}, N^m) \)). Then

\[
FJ(\Theta(f)) = M^{-1} \sum_{m=0}^{\infty} mc_m(f).
\]

Here \( M \) (equal to \( N \) or \( 2^{-1} N \)) is the width of the cusp.

(iii) If \( f \in H^0(S, \mathcal{L}^k) \) and \( g \in H^0(S, \mathcal{L}^l) \) then

\[
\Theta(fg) = f\Theta(g) + \Theta(f)g.
\]

(iv) \( \Theta(h_{\Sigma} f) = h_{\Sigma} \Theta(f) \) (equivalently, \( \Theta(h_{\Sigma}) = 0 \)).

**Corollary 4.6.** The operator \( \Theta \) extends to a derivation of the graded ring of modular forms, and for any \( f \), \( \Theta(f) \) is a cusp form.

Parts (iii) and (iv) of the theorem are clear from the construction. The proof of (i), that \( \Theta(f) \) is in fact holomorphic along \( S_{ss} \), will be given in the next section. We shall now study its effect on Fourier-Jacobi expansions, i.e. part (ii). That a factor like \( M^{-1} \) is necessary in (ii) becomes evident if we consider what happens to FJ expansions under level change. If \( N \) is replaced by \( N' = NQ \) then the conormal bundle becomes the \( Q \)-th power of the conormal bundle of level \( N' \), i.e. \( N = N'^Q \) (see Section 1.4.3). It follows that what was the \( m \)-th FJ coefficient at level \( N \) becomes the \( Qm \)-th coefficient at level \( N' \). The operator \( \Theta \) commutes with level-change, but the factor \( M^{-1} \), which changes to \( (QM)^{-1} \), takes care of this.

4.2.3. The effect of \( \Theta \) on FJ expansions. Let \( E \) be the standard cuspidal component of \( \overline{S} \) (over the ring \( R_N \)). As the reader probably noticed, we have trivialized the line bundle \( \mathcal{L} \) along \( E \) on two different occasions in two seemingly different ways, that we now have to compare. On the one hand, after reducing modulo \( P \) and pulling \( \mathcal{L} \) back to the Igusa surface, we got a canonical nowhere vanishing section \( a(1) \) trivializing \( \mathcal{L} \) over \( \overline{mg} \), and in particular along any of the \( p^2 - 1 \) cuspidal components lying over \( E \) in \( \overline{mg} \). On the other hand, extending scalars from \( R_N \) to
Lemma 4.7. The sections $a(1)$ and $2\pi i d\zeta_3$ “coincide” in the sense that they come from the same section in $H^0(E,L)$.

Proof. Let $A$ be the universal semi-abelian variety over $E$. Its toric part is $O_K \otimes \mathbb{G}_m$, hence, taking $\Sigma$-component of the cotangent space at the origin

\begin{equation}
L|_E = \omega_{A/E}(\Sigma) = (\delta_\Sigma^{-1} O_K \otimes \omega_{\mathbb{G}_m})(\Sigma)
\end{equation}

admits the canonical section $e_\Sigma \cdot (1 \otimes dT/T)$. Tracing back the definitions and using \[1.09\], this section becomes, under the base change $R_N \hookrightarrow \mathbb{C}$, just $2\pi i d\zeta_3$. On the other hand, when we reduce it modulo $P$ and use the Igusa level structure $\varepsilon$ at the standard cusp, it pulls back to the section “with the same name” $e_\Sigma \cdot (1 \otimes dT/T)$, because along $E$ \[4.19\] induces the identity on cotangent spaces. The lemma follows from the fact that, by definition, $\varepsilon^* a(1) = e_\Sigma \cdot (1 \otimes dT/T)$ too. \hfill $\Box$

Lemma 4.8. The sections $a(1)^{p+1}$ and $2\pi i d\zeta_2 \otimes 2\pi i d\zeta_3$ “coincide” in the sense that they come from the same section in $H^0(E,P_\mu \otimes L)$.

Proof. Let $\sigma_2$ (resp. $\sigma_3$) be the $K_N$-rational section of $P_\mu$ (resp. $L$) along $E$, which over $\mathbb{C}$ becomes the section $2\pi i d\zeta_2$ (resp. $2\pi i d\zeta_3$). We have just seen that modulo $P$, when we identify $E$ with $E$ (via the covering map $\tau : \overline{Tg} \to \hat{S}$), $\sigma_3$ reduces to $a(1)$. To conclude, we must show that the map

\begin{equation}
V : P/P_0 = P_\mu \simeq L^{(p)}
\end{equation}

carries $\sigma_2$ to $\sigma_3^{(p)}$. This will map, under $L^{(p)} \simeq L^p$, to $a(1)^p$. Along $E$ the line bundles $P_\mu$ and $L$ are just the $\Sigma$- and $\hat{\Sigma}$-parts of the cotangent space at the origin of the torus $O_K \otimes \mathbb{G}_m$, and $\sigma_2$ and $\sigma_3$ are the sections

\begin{equation}
\sigma_2 = e_\Sigma \cdot (1 \otimes dT/T), \quad \sigma_3 = e_\Sigma \cdot (1 \otimes dT/T).
\end{equation}

Since in characteristic $p$, $V = Ver^* : \omega_{\mathbb{G}_m} \to \omega_{\mathbb{G}_m}^{(p)}$ maps $dT/T$ to $(dT/T)^{(p)}$, for the $O_K$-torus, $V(\sigma_2) = \sigma_3^{(p)}$, and we are done. \hfill $\Box$

To prove part (ii) we argue as follows. Let $g = f/a(1)^k$ be the function on $\overline{Tg_\mu}$ obtained by trivializing the line bundle $L$. We have to study the FJ expansion along $E$ of $\kappa(dg)/a(1)^{p+1}$, where $\kappa$ is the map defined in \[4.11\]. For that purpose we may restrict to a formal neighborhood of $E$. This formal neighborhood is isomorphic, under the covering map $\tau : \overline{Tg_\mu} \to \hat{S}_\mu$, to the formal neighborhood $\overline{S}$ of $E$ in $S$. We may therefore regard $dg$ as an element of $\Omega^1_S$. Now

\begin{equation}
\kappa : \Omega^1_S \to P_\mu \otimes L
\end{equation}

is a homomorphism of $O_S$-modules defined over $R_N$ so, having restricted to $\overline{S}$, we may study the effect of $\kappa$ on FJ expansions by embedding $\overline{S}$ in a tubular neighborhood $\overline{S}(\varepsilon)$ of $E$ and using complex analytic Fourier-Jacobi expansions. We are thus reduced to a complex-analytic computation, near the standard cusp at infinity.
Let
\begin{equation}
(4.28) \quad g(z, u) = \sum_{m=0}^{\infty} \theta_m(u)q^m
\end{equation}
where \( q = e^{2\pi iz/M} \) and \( \theta_m \) is a theta function, so that \( \theta_m(u)q^m \) is a section of \( \mathcal{N}^m \) along \( E \) (now over \( \mathbb{C} \)). Then
\begin{equation}
(4.29) \quad dg = 2\pi i M^{-1} \sum_{m=0}^{\infty} m\theta_m(u)q^m dz + \sum_{m=0}^{\infty} \theta_m'(u)q^m du.
\end{equation}
According to Corollary 2.17 \( \kappa(du) = 0 \), and \( \kappa(dz) = 2\pi i d\zeta_2 \otimes d\zeta_3 \). It follows that
\begin{equation}
(4.30) \quad \kappa(dg) = M^{-1} \sum_{m=0}^{\infty} m\theta_m(u)q^m \cdot 2\pi i d\zeta_2 \otimes 2\pi i d\zeta_3.
\end{equation}
Recalling that in characteristic \( p \), \( 2\pi i d\zeta_2 \otimes 2\pi i d\zeta_3 \) reduced to \( (1)^{p+1} \), the proof of part (ii) of the theorem is now complete. For the convenience of the reader we summarize the transitions between complex and \( p \)-adic maps in the following diagram:
\begin{equation}
(4.31)
\begin{array}{ccc}
/\mathbb{F}_p & \Omega^1_S/\mathbb{F}_p & \overset{K^S(\Sigma)-1}{\rightarrow} \mathcal{P} \otimes \mathcal{L} \\
\cap & & \downarrow \text{mod } \mathcal{P}_0 \\
/\mathbb{F}_p & \Omega^1_S/\mathbb{F}_p & \overset{K^S(\Sigma)-1}{\rightarrow} \mathcal{P}_\mu \otimes \mathcal{L} \\
\uparrow \text{mod } \mathcal{P} & & \uparrow \\
/\mathbb{R} & \Omega^1_S/\mathbb{R} & \overset{K^S(\Sigma)-1}{\rightarrow} \mathcal{P}_\mu \otimes \mathcal{L} \\
\downarrow & & \downarrow \\
/\mathbb{C} & \Omega^1_S/\mathbb{C} & \overset{K^S(\Sigma)-1}{\rightarrow} \mathcal{P} \otimes \mathcal{L} \\
\cup & & \\
/\mathbb{C} & \Omega^1_S/\mathbb{C} & \overset{K^S(\Sigma)-1}{\rightarrow} \mathcal{P} \otimes \mathcal{L} \\
\end{array}
\end{equation}
We next turn to part (i).

4.3. A study of the theta operator along the supersingular locus.

4.3.1. De Rham cohomology in characteristic \( p \). We continue to consider thePicard surface \( S \) over \( \overline{\mathbb{F}}_p \). Let \( U = \text{Spec}(R) \rightarrow S \) be a closed point \( s_0 = (R = \overline{\mathbb{F}}_p, \mathcal{O}_{S, s_0}/\mathfrak{m}_{S, s_0}) \), a nilpotent thickening of a closed point, or an affine open subset of \( S \). We consider the restriction of the universal abelian scheme to \( R \) and denote it by \( A/R \). Let \( A^{(p)} = R \otimes_{\phi, R} A \) be its base change with respect to the map \( \phi(x) = x^p \). Let
\begin{equation}
(4.32) \quad D = H^4_{dR}(A/R),
\end{equation}
a locally free \( R \)-module of rank 6. The de Rham cohomology of \( A^{(p)} \) is
\begin{equation}
(4.33) \quad D^{(p)} = R \otimes_{\phi, R} D.
\end{equation}
The \( R \)-linear Frobenius and Verschiebung morphisms \( \text{Frob} : A \rightarrow A^{(p)}, \text{Ver} : A^{(p)} \rightarrow A \) induce (by pull-back) linear maps
\begin{equation}
(4.34) \quad F : D^{(p)} \rightarrow D, \ V : D \rightarrow D^{(p)}.
\end{equation}
Both $F$ and $V$ are everywhere of rank 3, which implies that their kernel and image are locally free direct summands. Moreover, $\text{Im} F = \ker V$ and $\text{Im} V = \ker F = \omega_{A/R}$. The maps $F$ and $V$ preserve the types $\Sigma, \bar{\Sigma}$, but note that $D^{(p)}(\Sigma) = D(\Sigma)^{(p)}$ etc.

The principal polarization on $A$ induces one on $A^{(p)}$, and these polarizations induce symplectic forms

\[ \langle \cdot, \cdot \rangle : D \times D \to R, \quad \langle \cdot, \cdot \rangle^{(p)} : D^{(p)} \times D^{(p)} \to R \]

where the second form is just the base-change of the first. For $x \in D^{(p)}, y \in D$ we have

\[ \langle FX, y \rangle = \langle x, V y \rangle^{(p)}. \]

In addition, for $a \in O_K$

\[ \langle i(a)x, y \rangle = \langle x, i(\bar{a})y \rangle. \]

As $VF = FV = 0$, the first relation implies that $\text{Im} F$ and $\text{Im} V$ are isotropic subspaces. So is $\omega_{A/R}$.

The Gauss-Manin connection is an integrable connection

\[ \nabla : D \to \Omega^1_R \otimes D. \]

It is a priori defined (e.g. in [Ka-O]) when $R$ is smooth over $\bar{\mathbb{F}}_p$, but we can define it by base change also when $R$ is a nilpotent thickening of a point of $S$ (see [Kob], where $R$ is a local Artinian ring). Note however that if $R = O_{S,s_0}$ and $R_m = O_{S,s_0}/m_{S,s_0}^{m+1}$ ($m \geq 0$), and if we extend scalars from $R$ to $R_m$ we get by base change the connection

\[ \nabla : D_{R_m} \to (\Omega^1_R \otimes_R R_m) \otimes_{R_m} D_{R_m} \to \Omega^1_{R_m} \otimes_{R_m} D_{R_m} \]

but $\Omega^1_R \otimes_R R_m \to \Omega^1_{R_m}$ is not an isomorphism. In fact, $\Omega^1_{R_m}$ is not $R_m$-free. As a result the Kodaira-Spencer map

\[ KS(\Sigma) : P_{R_m} \otimes_{R_m} L_{R_m} \to \Omega^1_{R_m} \]

will not be an isomorphism over $R_m$, as it is over $R$. It is nevertheless true that $KS(\Sigma)$ induces an isomorphism

\[ KS(\Sigma) : P_{R_{m-1}} \otimes_{R_{m-1}} L_{R_{m-1}} \cong \Omega^1_{R_m} \otimes_{R_m} R_{m-1}. \]

This follows from the fact that $\Omega^1_{R_m} \otimes_{R_m} R_{m-1} = \Omega^1_R \otimes_R R_{m-1}$ (if $u, v$ are local parameters at $s_0$ then $\{u^i v^j du, u^i v^j dv\}_{i+j \leq m-2}$ are linearly independent over $\bar{\mathbb{F}}_p$ in $\Omega^1_{R_m}$; the only dependencies between these differentials occurring for $i + j = m - 1$).

We shall need to deal only with the first infinitesimal neighborhood of a point, $R = O_{S,s_0}/m_{S,s_0}^2$. In this case, $D$ has a basis of horizontal sections. Indeed, $R = \bar{\mathbb{F}}_p[u,v]/(u^2, uv, v^2)$ where $u$ and $v$ are local parameters at $s_0$, and

\[ \Omega^1_R = (Rdu + Rdv)/\langle udu, vdv, udv + vdu \rangle \]

($p$ is odd). If $x \in D$ and

\[ \nabla x = du \otimes x_1 + dv \otimes x_2 \]

then $\partial x = x - u x_1 - v x_2$ satisfies

\[ \nabla \partial x = -u \nabla x_1 - v \nabla x_2. \]
But if \( \nabla x_1 = du \otimes x_{11} + dv \otimes x_{12} \) and \( \nabla x_2 = du \otimes x_{21} + dv \otimes x_{22} \) then
\[
0 = \nabla^2 x = du \wedge dv \otimes (x_{21} - x_{12})
\]
hence \( x_{21} - x_{12} \in m_{S',s_0} D \). It follows that
\[
\nabla \tilde{x} = -udv \otimes x_{12} - vdu \otimes x_{21} = du \otimes (x_{12} - x_{21}) = 0.
\]
This means that \( \tilde{x} \) is a horizontal section having the same specialization as \( x \) in the special fiber, so the horizontal sections span \( D \) over \( R \) by Nakayama’s lemma.

Let \( e_1, e_2, f_3, f_1, f_2, e_3 \) be any six horizontal sections over \( R \), specializing to a basis of \( H^1_{dR}(A_{s_0}/\mathbb{F}_p) \). Let \( D_0 \) be their \( \mathbb{F}_p \)-span. As we have just seen,
\[
R \otimes_{\mathbb{F}_p} D_0 = D
\]
and \( \nabla = d \otimes 1 \). Since \( R^{d=0} = \mathbb{F}_p \), it follows that \( D_0 = D^\nabla \), i.e. there are no more horizontal sections besides \( D_0 \). Thus every \( x \in H^1_{dR}(A_{s_0}/\mathbb{F}_p) \) has a unique extension to a horizontal section \( x \in H^1_{dR}(A/R) \).

There is a similar connection on \( D^{(p)} \). The isogenies \( \text{Frob} \) and \( \text{Ver} \), like any isogeny, take horizontal sections with respect to the Gauss-Manin connection to horizontal sections, e.g. if \( x \in D \) and \( \nabla x = 0 \) then \( Vx \in D^{(p)} \) satisfies \( \nabla(Vx) = 0 \).

The pairing \( \langle \cdot, \cdot \rangle \) is horizontal for \( \nabla \), i.e.
\[
d \langle x, y \rangle = \langle \nabla x, y \rangle + \langle x, \nabla y \rangle.
\]

**Remark 4.1.** In the theory of Dieudonné modules one works over a perfect base. It is then customary to identify \( D \) with \( D^{(p)} \) via \( x \leftrightarrow 1 \otimes x \). This identification is only \( \sigma \)-linear where \( \sigma = \phi \), now viewed as an automorphism of \( R \). The operator \( F \) becomes \( \sigma \)-linear, \( V \) becomes \( \sigma^{-1} \)-linear and \( \{4.30\} \) reads \( \langle Fx, y \rangle = \langle x, Vy \rangle^\sigma \). With this convention \( F \) and \( V \) switch types, rather than preserve them.

**4.3.2. The Dieudonné module at a gss point.** Assume from now on that \( s_0 \in Z' = S_{\text{gss}} \) is a closed point of the general supersingular locus. We write \( D_0 \) for \( H^1_{dR}(A_{s_0}/\mathbb{F}_p) \).

**Lemma 4.9.** There exists a basis \( e_1, e_2, f_3, f_1, f_2, e_3 \) of \( D_0 \) with the following properties. Denote by \( e_1^{(p)} = 1 \otimes e_1 \in D_0^{(p)} \) etc.

(i) \( \mathcal{O}_K \) acts on the \( e_i \) via \( \Sigma \) and on the \( f_i \) via \( \bar{\Sigma} \) (hence it acts on the \( e_i^{(p)} \) via \( \Sigma \) and on the \( f_i^{(p)} \) via \( \bar{\Sigma} \)).

(ii) The symplectic pairing satisfies
\[
\langle e_i, f_j \rangle = -\langle f_j, e_i \rangle = \delta_{ij}, \quad \langle e_i, e_j \rangle = \langle f_i, f_j \rangle = 0.
\]

(iii) The vectors \( e_1, e_2, f_3 \) form a basis for the cotangent space \( \omega_{A_0/\mathbb{F}_p} \). Hence \( e_1 \)
and \( e_2 \) span \( \mathcal{P} \) and \( f_3 \) spans \( \mathcal{L} \).

(iv) \( \ker(V) \) is spanned by \( e_1, f_2, e_3 \). Hence \( \mathcal{P}_0 = \mathcal{P} \cap \ker(V) \) is spanned by \( e_1 \).

(v) \( V f_2 = f_3^{(p)}, V f_3 = e_1^{(p)}, V f_1 = e_2^{(p)} \).

(vi) \( F f_1^{(p)} = -e_3, F f_2^{(p)} = -e_1, F e_3^{(p)} = -f_2 \).

**Proof.** Up to a slight change of notation, this is the unitary Dieudonné module which Büttel and Wedhorn call a “braid of length 3” and denote by \( \tilde{B}(3) \), cf [Bu-We] (3.2). The classification in loc. cit. Proposition 3.6 shows that the Dieudonné module of a \( \mu \)-ordinary abelian variety is isomorphic to \( \tilde{B}(2) \oplus \tilde{S} \), that of a gss abelian variety is isomorphic to \( \tilde{B}(3) \) and in the superspecial case we get \( \tilde{B}(1) \oplus \tilde{S}^2 \). \( \square \)
4.3.3. **Infinitesimal deformations.** Let $\mathcal{O}_{S,s_0}$ be the local ring of $S$ at $s_0$, $\mathfrak{m}$ its maximal ideal, and $R = \mathcal{O}_{S,s_0}/\mathfrak{m}^2$. This $R$ is a truncated polynomial ring in two variables, isomorphic to $\overline{\mathbb{F}}_p[u, v]/(u^2, uv, v^2)$.

As remarked above, the de Rham cohomology $D = H^1_{dR}(A/R)$ has a basis of horizontal sections,

\begin{equation}
D = R \otimes_{\overline{\mathbb{F}}_p} D^\nabla, \quad \nabla = d \otimes 1
\end{equation}

and since $D_0 = \overline{\mathbb{F}}_p \otimes_R D$, we may identify $D^\nabla$ canonically with $D_0$.

Grothendieck tells us that $A/R$ is completely determined by $A_0$ and by the Hodge filtration $\omega_{A/R} \subset D = R \otimes_{\overline{\mathbb{F}}_p} D_0$. Since $A$ is the universal infinitesimal deformation of $A_0$, we may choose the coordinates $u$ and $v$ so that

\begin{equation}
\mathcal{P} = \text{Span}_R \{e_1 + ue_3, e_2 + ve_3\}.
\end{equation}

The fact that $\omega_{A/R}$ is isotropic implies then that

\begin{equation}
\mathcal{L} = \text{Span}_R \{f_3 - uf_1 - vf_2\}.
\end{equation}

Consider the abelian scheme $A^{(p)}$. It is not the universal deformation of $A_0^{(p)}$ over $R$. In fact, the map $\phi : R \to R$ factors as

\begin{equation}
R \xrightarrow{\pi} \overline{\mathbb{F}}_p \xrightarrow{d_3} \overline{\mathbb{F}}_p \xrightarrow{i} R,
\end{equation}

and therefore $A^{(p)}$, unlike $A$, is constant: $A^{(p)} = \text{Spec}(R) \times_{\text{Spec} \overline{\mathbb{F}}_p} A_0^{(p)}$. As with $D$, $D^{(p)} = R \otimes_{\overline{\mathbb{F}}_p} D_0^{(p)}$, $\nabla = d \otimes 1$, but this time the basis of horizontal sections can be obtained also from the trivialization of $A^{(p)}$, and $\omega_{A^{(p)}/R} = \text{Span}_R \{e_1^{(p)}, e_2^{(p)}, f_3^{(p)}\}$.

Since $V$ and $F$ preserve horizontality, $e_1, f_2, e_3$ span $\ker(V)$ over $R$ in $D$, and the relations in $(v)$ and $(vi)$ of Lemma 4.9 continue to hold. Indeed, the matrix of $V$ in the basis at $s_0$ prescribed by that lemma, continues to represent $V$ over $\text{Spec}(R)$ by “horizontal continuation”. The matrix of $F$ is then derived from the relation $(vi)$.

The Hodge filtration nevertheless varies, so we conclude that

\begin{equation}
\mathcal{P}_0 = \mathcal{P} \cap \ker(V) = \text{Span}_R \{e_1 + ue_3\}.
\end{equation}

The condition $V(\mathcal{L}) = \mathcal{P}_0^{(p)}$, which is the “equation” of the closed subscheme $Z' \cap \text{Spec}(R)$ (see Proposition 3.9) means

\begin{equation}
V(f_3 - uf_1 - vf_2) = e_1^{(p)} - ue_2^{(p)} \in R \cdot e_1^{(p)}
\end{equation}

and this holds if and only if $u = 0$. We have proved the following lemma.

**Lemma 4.10.** Let $s_0 \in S_{gg}$ as above. Then the closed subscheme $S_{gg} \cap \text{Spec}(R)$ is given by the equation $u = 0$.

4.3.4. **The Kodaira-Spencer isomorphism along the general supersingular locus.** We keep the assumptions of the previous subsections, and compute what the Gauss-Manin connection does to $\mathcal{P}_0$. A typical element of $\mathcal{P}_0$ is $g(e_1 + ue_3)$ for some $g \in R$.

Then

\begin{equation}
\nabla(g(e_1 + ue_3)) = dg \otimes (e_1 + ue_3) + d(ue_3) = dg \otimes e_3 + gdu \otimes e_3.
\end{equation}

Note that when we divide by $\omega_{A/R}$ and project $H^1_{dR}(A/R)$ to $H^1(A, \mathcal{O})$, $e_1 + ue_3$ dies, and the image $\overline{e}_3$ of $e_3$ becomes a basis for the line bundle that we called $\mathcal{L}^{\vee}(\rho) = H^1(A, \mathcal{O})(\Sigma)$. Recall the definition of $\kappa$ given in (4.14), but note that this
definition only makes sense over $\text{Spec}(O_{S,s_0})$ or its completion, where $KS(\Sigma)$ is an isomorphism, and can be inverted.

**Proposition 4.11.** Let $s_0 \in Z' = S_{gss}$. Choose local parameters $u$ and $v$ at $s_0$ so that in $O_{S,s_0}$ the local equation of $Z'$ becomes $u = 0$. Then at $s_0$, $\kappa(du)$ has a zero along $Z'$.

**Proof.** Let $i : Z' \hookrightarrow S$ be the locally closed embedding. We must show that in a suitable Zariski neighborhood of $s_0$, where $u = 0$ is the local equation of $Z'$, $i^*\kappa(du) = 0$. It is enough to show that the image of $\kappa(du)$ in the fiber at every point $s$ of $Z'$ near $s_0$, vanishes. All points being alike, it is enough to do it at $s_0$. In other words, we denote by $\kappa_0$ the map

$$
\kappa_0 : \Omega^1_{S,s_0} \rightarrow \mathcal{P}_\mu \otimes L|_{s_0} \simeq L^{p+1}|_{s_0}.
$$

and show that $\kappa_0(du) = 0$. We may now work over $\text{Spec}(R)$, where $R = O_{S,s_0}/m^2$.

It is enough to show that in the diagram

$$
\begin{array}{ccc}
\mathcal{P}_R \otimes L_R & \xrightarrow{KS(\Sigma)} & \Omega^1_R \\
\downarrow & & \downarrow \\
\mathcal{P}_{s_0} \otimes L_{s_0} & \simeq & \Omega^1_{S,s_0}
\end{array}
$$

$KS(\Sigma)$ maps the line sub-bundle $\mathcal{P}_{0,R} \otimes L_R$ onto $Rdu$. Once we have passed to the infinitesimal neighborhood $\text{Spec}(R)$ we can replace the local parameters $u, v$ by any two formal parameters for which $u = 0$ defines $Z' \cap \text{Spec}(R)$. We may therefore assume, in view of Lemma 4.10 that $u$ and $v$ have been chosen as in section 4.3.3. But then equation (4.55) shows that the restriction of $KS(\Sigma)$ to $Z'$, i.e. the homomorphism $i^*KS(\Sigma)$, maps $i^*(\mathcal{P})$ onto $i^*R \cdot du \otimes \tau_\Sigma$. This concludes the proof. □

**4.3.5. A computation of poles along the supersingular locus.** We are now ready to prove the following.

**Proposition 4.12.** Let $k \geq 0$, and let $f \in H^0(S, L^k)$ be a modular form of weight $k$ in characteristic $p$. Then $\Theta(f) \in H^0(S, L^{k+p+1})$.

**Proof.** A priori, the definition that we have given for $\Theta(f)$ produces a meromorphic section of $L^{k+p+1}$ which is holomorphic on the $\mu$-ordinary part $S_\mu$ but may have a pole along $Z = S_{ss}$. Since $S$ is a non-singular surface, it is enough to show that $\Theta(f)$ does not have a pole along $Z' = S_{gss}$, the non-singular part of the divisor $Z$. Consider the degree $p^2 - 1$ covering $\tau : Ig \rightarrow S$, which is finite, étale over $S_\mu$ and totally ramified along $Z$. Let $s_0 \in Z'$ and let $\tilde{s}_0 \in Ig$ be the closed point above it. Let $u, v$ be formal parameters at $s_0$ for which $Z'$ is given by $u = 0$, as in Theorem 3.23. As explained there we may choose formal parameters $w, v$ at $\tilde{s}_0$ where $w^{p^2 - 1} = u$ (and $v$ is the same function $v$ pulled back from $S$ to $Ig$). It follows that in $\Omega^1_{Ig}$ we have

$$
du = -w^{p^2 - 2}dw.
$$

We now follow the steps of our construction. Dividing $f$ by $a(1)^k$ we get a function $g = f/a(1)^k$ on $Ig$ with a pole of order $k$ along $\tilde{Z}$, the supersingular divisor on $Ig$, whose local equation is $w = 0$. In $\tilde{O}_{Ig,\tilde{s}_0}$ we may write

$$
g = \sum_{l=-k}^{\infty} g_l(v)w^l.
$$
Then
\begin{equation}
dg = \sum_{l=-k}^{\infty} \log(v) w^{l-1} \, dw + \sum_{l=-k}^{\infty} w^l g_l(v) \, dv.
\end{equation}
(4.60)

Applying the map \( \kappa \) (extended \( \mathcal{O}_{\mathcal{I}_g} \)-linearly from \( S \) to \( \mathcal{I}_g \)), and noting that \( \kappa(du) \) has a zero along \( Z' \), hence a zero of order \( p^2 - 1 \) along \( Z' \), we conclude that \( \kappa(dg) \) has a pole of order \( k \) (at most) along \( Z' \). Finally \( \Theta(f) = a(1)^k \cdot \kappa(dg) \) becomes holomorphic along \( Z' \), and also descends to \( S \). It is therefore a holomorphic section of \( \mathcal{P}_\mu \otimes \mathcal{L}^{k+1} \simeq \mathcal{L}^{k+p+1} \).

It is amusing to compare the reasons for the increase by \( p+1 \) in the weight in \( \Theta(f) \) for modular curves and for Picard modular surfaces. In the case of modular curves the Kodaira-Spencer isomorphism is responsible for a shift by 2 in the weight, but the section acquires simple poles at the supersingular points. One has to multiply by the Hasse invariant, which has weight \( p - 1 \), to make the section holomorphic, hence a total increase by \( p + 1 = 2 + (p - 1) \) in the weight. In our case, the Kodaira-Spencer isomorphism is responsible for a shift by \( p+1 \) (the \( p \) coming from \( \mathcal{P}_\mu \simeq \mathcal{L}^p \)), but the section turns out to be holomorphic along the supersingular locus. See Section 4.5.

4.4. Relation to the filtration and theta cycles. In part (ii) of the main theorem we have described the way \( \Theta \) acts on Fourier-Jacobi expansions at the standard cusp. A similar formula inevitably exists at any other cusp. We may deduce from it that modular forms in the image of \( \Theta \) have vanishing FJ coefficients in degrees divisible by \( p \). Moreover, for such a form \( f \in \text{Im}(\Theta) \), \( \Theta^{p-1}(f) \) and \( f \) have the same FJ expansions, and hence the same filtration. Note also that if \( r(f_1) = r(f_2) \) then \( r(\Theta(f_1)) = r(\Theta(f_2)) \). We may therefore define unambiguously
\begin{equation}
\Theta(r(f)) = r(\Theta(f)).
\end{equation}
(4.61)

As we clearly have
\begin{equation}
\omega(\Theta(f)) = \omega(f) + p + 1 - a(p^2 - 1)
\end{equation}
for some \( a \geq 0 \) we deduce the following result.

**Proposition 4.13.** Let \( f \in M_k(N, \overline{\mathbb{F}}_p) \) be a modular form modulo \( p \), and assume that \( r(f) \in \text{Im}(\Theta) \). Then
\begin{equation}
r(f) = r(\Theta^{p-1}(f)).
\end{equation}
(4.63)

There exists a unique index \( 0 \leq i \leq p - 2 \) such that
\begin{equation}
\omega(\Theta^{i+1}(f)) = \omega(\Theta^i(f)) + p + 1 - (p^2 - 1).
\end{equation}
(4.64)

For any other \( i \) in this range
\begin{equation}
\omega(\Theta^{i+1}(f)) = \omega(\Theta^i(f)) + p + 1.
\end{equation}
(4.65)

This is reminiscent of the “theta cycles” for classical (i.e. elliptic) modular forms modulo \( p \), see [Se], [Ka2] and [Joc]. Recall that if \( f \) is a mod \( p \) modular form of weight \( k \) on \( \Gamma_0(N) \) with \( q \)-expansion \( \sum a_n q^n \) \( (a_n \in \overline{\mathbb{F}}_p) \), then \( \Theta(f) \) is a mod \( p \) modular form of weight \( k + p + 1 \) with \( q \) expansion \( \sum na_n q^n \) (Katz denotes \( \Theta(f) \)
by $A\theta(f)$. One has $\omega(\Theta(f)) < \omega(f) + p + 1$ if and only if $\omega(f) \equiv 0 \mod p$. In such a case we say that the filtration “drops” and we have

$$\omega(\Theta(f)) = \omega(f) + p + 1 - a(p - 1)$$

(4.66)

for some $a > 0$. As a corollary, $\omega(f)$ can never equal 1 mod $p$ for an $f \in \text{Im}(\Theta)$. Assume now that $f \in \text{Im}(\Theta)$ is a “low point” in its “theta cycle”, namely, $\omega(f)$ is minimal among all $\omega(\Theta^{i}(f))$. Then $\omega(\Theta^{i+1}(f)) < \omega(\Theta^{i}(f)) + 1$ for one or two values of $i \in \{0, p - 2\}$, which are completely determined by $\omega(f) \mod p$ [Joc].

This is not true anymore for Picard modular forms. Not only the drop in the theta cycle is unique, but the question of when exactly it occurs is mysterious and deserves further study. We make the following elementary observation showing that whether a drop in the filtration occurs in passing from $f$ to $\Theta(f)$ can not be determined by $\omega(f)$ modulo $p$ alone. Let $f$ and $k$ be as in the Proposition.

1. If $k \leq p^2 - 1$ then $\omega(f) = k$.
2. If $k < p + 1$ then $\omega(\Theta^{i}(f)) = k + i(p + 1)$ for $0 \leq i \leq p - 2$, so the drop occurs at the last step of the theta cycle, i.e. at weight $k + (p - 2)(p + 1)$, which is congruent to $k - 2$ modulo $p$.
3. If $k < p + 1$ but $f \notin \text{Im}(\Theta)$ then starting with $\Theta(f)$ instead of $f$, one sees that the drop in the theta cycle of $\Theta(f)$ occurs either in passing from $\Theta^{p-2}(f)$ to $\Theta^{p-1}(f)$, or in passing from $\Theta^{p-1}(f)$ to $\Theta^{p}(f)$.

4.5. Compatibility between theta operators for elliptic and Picard modular forms.

4.5.1. The theta operator for elliptic modular forms. The theta operator for elliptic modular forms modulo $p$ was introduced by Serre and Swinnerton-Dyer in terms of $q$-expansions, cf. [Se], but its geometric construction was given by Katz in [Ka1] and [Ka2], where it is denoted $A\theta$. Katz’ construction relied on a canonical splitting of the Hodge filtration over the ordinary locus, but it coincides with a slightly modified construction, similar to the one we have been using over the Picard modular surface. This construction was suggested in [Gr], Proposition 5.8, see also [An-Go] in the Hilbert modular case.

Let $X$ be the modular curve $X(N)$ over $\overline{\mathbb{F}}_p$ $(N \geq 3, p \nmid N)$ and $I_{ord}$ the Igusa curve of level $p$ lying over $X_{ord} = X - X_{ss}$, the ordinary part of $X$. Let $X$ and $I_{ord}$ be the curves obtained by adjoining the cusps to $X$ and $I_{ord}$ respectively. Let $L = \omega_{E/X}$ be the cotangent bundle of the universal elliptic curve, extended over the cusps as usual. Classical modular forms of weight $k$ and level $N$ are sections of $L^k$ over $X$. Let $a(1)$ be the tautological nowhere vanishing section of $L$ over $I_{ord}$. Given a modular form $f$ of weight $k$, we consider $r(f) = \tau^{*}f/a(1)^{k}$ where $\tau : I_{ord} \to X$ is the covering map, and apply the inverse of the Kodaira-Spencer isomorphism $KS : L^2 \to \Omega_{I_{ord}}^{1}$ to get a section $\kappa(dr(f))$ of $L^2$ over $I_{ord}$. When multiplied by $a(1)^{k}$ it descends to $X_{ord}$, and when this is multiplied further by $h = a(1)^{p-1}$, the Hasse invariant for elliptic modular forms, it extends holomorphically over $X_{ss}$ to an element

$$\theta(f) = a(1)^{k+p-1}\kappa(dr(f)) \in \Gamma(\tilde{X}, L^{k+p+1}).$$

(4.67)

4.5.2. An embedding of a modular curve in $\bar{S}$. To illustrate our idea, and to simplify the computations, we assume that $N = 1$ and $d_{K} \equiv 1 \mod 4$, so that $D = D_{K} =$
We shall treat only one special embedding of the modular curve $\bar{X} = X_0(D)$ into $\bar{S}$ (there are many more).

Embed $SL_2(\mathbb{R})$ in $G'_\infty$ via

$$
\left( \begin{array}{cc} a & b \\ c & d \end{array} \right) \mapsto \left( \begin{array}{cc} a & b \\ c & d \end{array} \right).
$$

This embedding induces an embedding of symmetric spaces $\bar{\mathfrak{H}} \hookrightarrow \mathfrak{X}$, $z \mapsto \iota(z,0)$.

One can easily compute that the intersection of $\Gamma$, the stabilizer of the lattice $L_0$ in $G'_\infty$, with $SL_2(\mathbb{R})$, is the subgroup of $SL_2(\mathbb{Z})$ given by

$$
\Gamma^0(D) = \left\{ \left( \begin{array}{cc} a & b \\ c & d \end{array} \right) : D | b \right\}.
$$

Let $E_0 = \mathbb{C}/\mathcal{O}_K$, endowed with the canonical principal polarization and $CM$ type $\Sigma$. For $z \in \bar{\mathfrak{H}}$ let $\Lambda_z = \mathbb{Z} + \mathbb{Z}z$ and $E_z = \mathbb{C}/\Lambda_z$. Let $M_z$ be the cyclic subgroup of order $D$ of $E_z$ generated by $D^{-1}z \mod \Lambda_z$. Using the model (1.27) of the abelian variety $A_z$ associated to the point $\iota(z,0) \in \mathfrak{X}$, we compute that

$$
A_z \simeq E_0 \times (\mathcal{O}_K \otimes E_z/\delta_K \otimes M_z)
$$

with the obvious $\mathcal{O}_K$-structure. The principal polarization on $A_z$ provided by the complex uniformization is the product of the canonical polarization of $E_0$ and the principal polarization of $\mathcal{O}_K \otimes E_z/\delta_K \otimes M_z$ obtained by descending the polarization

$$
\lambda_{can} : \mathcal{O}_K \otimes E_z \rightarrow \delta_K^{-1} \otimes E_z = (\mathcal{O}_K \otimes E_z)^t
$$

of degree $D^2$, modulo the maximal isotropic subgroup $\delta_K \otimes M_z$ of $\ker(\lambda_{can})$.

It is now clear that over any $R_0$-algebra $R$ we have the same moduli theoretic construction, sending a pair $(E,M)$ where $M$ is a cyclic subgroup of degree $D$ to $A(E,M)$, with $\mathcal{O}_K$ structure and polarization given by the same formulae. This gives a modular embedding $X \rightarrow S$ which is generically injective. To make this precise at the level of schemes (rather than stacks) one would have to add a level $N$ structure and replace the base ring $R_0$ by $R_N$.

4.5.3. Comparison of the two theta operators. From now on we work over $\bar{\mathbb{F}}_p$. The modular interpretation of the embedding $j : \bar{X} \rightarrow \bar{S}$ allows us to complete it to a diagram

$$
\begin{array}{ccc}
\bar{I}_{\text{ord}} & \stackrel{j}{\rightarrow} & \bar{T}\mathfrak{g}_{\mu} \\
\tau \downarrow & & \downarrow \tau \\
\bar{X}_{\text{ord}} & \stackrel{j}{\rightarrow} & \bar{S}_{\mu}
\end{array}
$$

Lemma 4.14. The pull-back $j^* \omega_{A/S}$ decomposes as a product $\omega_{E_0} \times (\mathcal{O}_K \otimes \omega_{E/X})$. Under this isomorphism

$$
\begin{align*}
\quad & j^* \mathcal{L} = (\mathcal{O}_K \otimes \omega_{E/X})(\Sigma) \\
\quad & j^* \mathcal{P}_0 = \omega_{E_0} \\
\quad & j^* \mathcal{P}_{\mu} \simeq (\mathcal{O}_K \otimes \omega_{E/X})(\Sigma).
\end{align*}
$$

The line bundle $j^* \mathcal{P}_0$ is constant, and $\mathcal{P}_{\mu}$, originally a quotient bundle of $\mathcal{P}$, becomes a direct summand when restricted to $\bar{X}$.
Proof. This is straightforward from the construction of \( j \), and the fact that \( E_0 \) is supersingular, while \( E \) is ordinary over \( \bar{X}_{\text{ord}} \). Note that \( \mathcal{O}_K \otimes E/\delta_K \otimes M \) and \( \mathcal{O}_K \otimes E \) have the same cotangent space. \( \square \)

Proposition 4.15. Identify \( j^* \mathcal{L} \) with \( \omega_{E/X} \) (\( \mathcal{O}_K \) acting via \( \bar{\Sigma} \)). Then for \( f \in \Gamma(\bar{S}, \mathcal{L}^k) = M_k(N, \overline{\mathbb{F}}_p) \)

\[
\theta(j^*(f)) = j^*(\Theta(f)).
\]

Proof. We abbreviate \( I_{\text{ord}} \) by \( I \) and \( I_g \) by \( I_g \). The pull-back via \( j \) of the tautological section \( a(1) \) of \( \mathcal{L} \) over \( I_g \) is the tautological section \( a(1) \) of \( j^* \mathcal{L} = \omega_{E/X} \). We therefore have

\[
j^*(\text{dr}(f)) = \text{dr}(j^*(f))
\]

(4.75)

\((r(f) = \tau^* f/a(1)^k \) is the function on \( I_g \) denoted earlier also by \( g \). It remains to check the commutativity of the following diagram

\[
\begin{array}{cccc}
\Omega^1_{I_g} & \xrightarrow{K \Sigma^{-1}} & \mathcal{P} \otimes \mathcal{L} & \xrightarrow{\mathcal{V} \otimes 1} & \mathcal{L}^{p+1} \\
j^*_0 & \downarrow & j^* & \downarrow & . \\
\end{array}
\]

Here \( j^*_0 \) is the map \( j^* \Omega^1_{I_g} \rightarrow \Omega^1_I \) on differentials whose kernel is the conormal bundle of \( I \) in \( I_g \). For that we have to compare the Kodaira-Spencer maps on \( S \) and on \( X \). As we have seen in the lemma, \( \mathcal{P}/\mathcal{P}_0 = \mathcal{P}_\mu \) pulls back under \( j \) to \( \mathcal{L}(\rho) \) (the line bundle \( \mathcal{L} \) with the \( \mathcal{O}_K \) action conjugated). But, \( KS(\Sigma)(\mathcal{P}_0 \otimes \mathcal{L}) \) maps under \( j^* \) to the conormal bundle, so we obtain a commutative diagram

\[
\begin{array}{cccc}
\Omega^1_{I_g} & \xrightarrow{K \Sigma^{-1}} & \mathcal{P} \otimes \mathcal{L} & \xrightarrow{\mathcal{V} \otimes 1} & \mathcal{L}^{p+1} \\
j^*_0 & \downarrow & j^* & \downarrow & . \\
\Omega^1_I & \leftarrow & j^* \mathcal{L}(\rho) \otimes j^* \mathcal{L} \\
\end{array}
\]

The commutativity of the diagram

\[
\begin{array}{cccc}
\mathcal{P}_\mu & \xrightarrow{\mathcal{V}} & \mathcal{L}(p) \\
\downarrow & \downarrow & . \\
\end{array}
\]

(4.78)

follows from the definition of the Hasse invariant \( h \) on \( X \). Identifying \( \mathcal{L}(p) \) with \( \mathcal{L}^p \) as usual and tensoring the last diagram with \( \mathcal{L} \) provides the last piece of the puzzle. \( \square \)

Bibliography

[An-Go] F. Andreatta, E. Z. Goren: Hilbert modular forms: mod \( p \) and \( p \)-adic aspects, Memoirs A.M.S. 819, 2005.

[Bel] J. Bellaïche: Congruences endoscopiques et représentations Galoisiennes, Thèse, Paris XI (Orsay), 2002.

[Bo] G. Boxer, Ph.D. Thesis, Harvard, to appear.

[Bu-We] O.Büttel, T.Wedhorn: Congruence relations for Shimura varieties associated to some unitary groups, J. Inst. Math. Jussieu 5, 229-261 (2006).
[Cog] J. Cogdell: Arithmetic cycles on Picard modular surfaces and modular forms of nebentypus, J. Reine Angew. Math. 357 (1985), 115-137.

[De] P. Deligne: Travaux de Shimura, Sém. Bourbaki 389 (1971).

[De-Ra] P. Deligne, M. Rapoport: Les schémas de modules de courbes elliptiques, in: Modular Functions of One Variable II, LNM 349 (1973), 143-316.

[Den] M. Demazure: Lectures on $p$-divisible groups, LNM 302, Springer, 1972.

[Fa-Ch] G. Faltings, C.-L. Chai: Degeneration of abelian varieties, Springer, 1990.

[Go-Ni] W. Goldring, M.-H. Nicole: A $\mu$-ordinary Hasse invariant, preprint, arXiv:1302.1614.

[Gor] E. Z. Goren: Hasse invariants for Hilbert modular varieties, Israel J. Math. 122 (2001), 157-174.

[Go-O] E. Z. Goren, F. Oort: Stratification of Hilbert modular varieties, J. Alg. Geom. 9 (2000), 111-154.

[Gr] B. Gross: A tameness criterion for Galois representations associated to modular forms (mod $p$), Duke Math. J. 61 (1990), 445-517.

[Ha] M. Harris: Arithmetic vector bundles and automorphic forms on Shimura varieties. I, Invent. Math. 82 (1985), 151-189.

[Hart] R. Hartshorne: Algebraic Geometry, Graduate Texts in Mathematics, No. 52, Springer, 1977.

[Ho] R.-P. Holzapfel: Ball and surface arithmetics, Aspects of Mathematics, E29, Vieweg & Sohn, 1998.

[Joc] N. Jochnowitz: A study of the local components of the Hecke algebra mod $l$, Trans. Amer. Math. Soc. 270 (1982), 253-267.

[Ka1] N. Katz: $p$-adic properties of modular schemes and modular forms, in: Modular functions of one variable III, 69-190. LNM 350, Springer, 1973.

[Ka2] N. Katz: A result on modular forms in characteristic $p$, in: Modular functions of one variable V, 53-61, LNM 601, Springer, 1977.

[Ka-Ma] N. Katz, B. Mazur: Arithmetic moduli of elliptic curves. Annals of Mathematics Studies 108, Princeton, 1985.

[Ka-O] N. Katz, T. Oda: On the differentiation of de Rham cohomology classes with respect to parameters, J. Math. Kyoto Univ. 8 (1968), 199-213.

[Kob] N. Koblitz: $p$-adic variation of the zeta-function over families of varieties defined over finite fields, Compositio Math. 31 (1975), 119-218.

[Kos] J.-S. Koskivirta: Sections of the Hodge bundle over Ekedahl Oort strata of Shimura varieties of Hodge type, preprint, arXiv:1410.1317.

[Ko-We] J.-S. Koskivirta, T. Wedhorn: Hasse invariants for Shimura varieties of Hodge type, preprint, arXiv:1406.2178.

[Lau] K.-W. Lan: Arithmetic compactifications of PEL-type Shimura varieties, London Mathematical Society Monographs 36, Princeton, 2013.
[L-R] R. Langlands, D. Ramakrishnan: *The zeta functions of Picard modular surfaces*, Univ. Montréal, Montréal, 1992.

[La1] M. Larsen: Unitary groups and $l$-adic representations, *Ph.D. Thesis*, Princeton University, 1988.

[La2] M. Larsen: Arithmetic compactification of some Shimura surfaces *in:* [L-R], 31-45.

[Mo] B. Moonen: Group schemes with additional structures and Weyl group cosets *in:* *Moduli of abelian varieties*, Progr. Math. 195, 255-298, Birkhäuser, 2001.

[Mu] D. Mumford: *Abelian varieties*, Oxford University Press, London 1970.

[Mu2] D. Mumford: *The red book of varieties and schemes*, second expanded edition, LNM 1358, Springer (1999).

[Oda] T. Oda: On the first de Rham cohomology group and Dieudonné modules, Ann. Sci. ENS 2 (1969), 63-135.

[Oo] F. Oort: A stratification of a moduli space of abelian varieties, *in:* *Moduli of abelian varieties*, Progr. Math. 195, 417-440, Birkhäuser, 2001.

[Ri] K. Ribet: A. $p$-adic interpolation via Hilbert modular forms, *in:* *Algebraic geometry*, Proc. Sympos. Pure Math. 29, 581–592, AMS, 1975.

[Se] J.-P. Serre: Formes modulaires et fonctions zêta $p$-adiques. *in:* *Modular functions of one variable III*, LNM 350, 191-268, Springer, 1973.

[Sh1] G. Shimura: Arithmetic of unitary groups, Ann. of Math. 79 (1964), 369-409.

[Sh2] G. Shimura: The arithmetic of automorphic forms with respect to a unitary group, Ann. of Math. 107 (1978), 569-605.

[V] I. Vollaard: The supersingular locus of the Shimura variety for $GU(1,s)$, Canad. J. Math. 62 (2010), 668-720.

[We] T. Wedhorn: The dimension of Oort strata of Shimura varieties of PEL-type, *in:* *Moduli of abelian varieties*, Progr. Math. 195, 441-471, Birkhäuser, 2001.

Hebrew University, Jerusalem, Israel
E-mail address: deshalit@math.huji.ac.il.

McGill University, Montreal, Canada
E-mail address: eyal.goren@mcgill.ca