WICK ROTATION AND THE POSITIVITY OF ENERGY IN QUANTUM FIELD THEORY

by MAXIM KONTSEVICH

(I. H. E. S., Bures-sur-Yvette, France)

and GRAEME SEGAL†

(All Souls College, University of Oxford, Oxford, UK)

[Received 23 December 2020]

Dedicated to the memory of Sir Michael Atiyah

Abstract

We propose a new axiom system for unitary quantum field theories on curved space-time backgrounds, by postulating that the partition function and the correlators extend analytically to a certain domain of complex-valued metrics. Ordinary Riemannian metrics are contained in the allowable domain, while Lorentzian metrics lie on its boundary.

1. Introduction

In conventional quantum theory, the states of a system are represented by the rays in a complex Hilbert space \( \mathcal{H} \), and the time evolution is given by a one-parameter group of unitary operators

\[
U_t = e^{iHt} : \mathcal{H} \to \mathcal{H}
\]

(for \( t \in \mathbb{R} \)), generated by an unbounded self-adjoint operator \( H \) called the Hamiltonian. Positivity of the energy corresponds to the fact that \( H \) is positive-semidefinite, i.e., the spectrum of \( H \) is contained in \( \mathbb{R}_+ \). This is clearly equivalent to saying that the operator-valued function \( t \mapsto U_t \) is the boundary value of a holomorphic function \( t \mapsto U_t \), which is defined in the upper half-plane

\[
\{ t \in \mathbb{C} : \text{Im}(t) > 0 \}
\]

and is bounded in the operator norm. (The physically relevant condition is actually that the energy is bounded below, so rather than asking for \( U_t \) to be bounded for \( \text{Im}(t) > 0 \) we could require \( \| U_t \| \leq e^{c\text{Im}(t)} \) for some \( c \).)

The holomorphic formulation helps us see what a strong constraint the positivity of energy is. The boundary value of a bounded holomorphic function in the upper half-plane must vanish identically if it vanishes on an open interval of the real axis, and so, if taken literally, positive energy implies that a state \( \xi \in \mathcal{H} \) for which \( U_t(\xi) \) belongs to a closed subspace \( \mathcal{H}_0 \) of \( \mathcal{H} \) for all \( t < 0 \) must remain

† Corresponding author. E-mail: graeme.segal@all-souls.ox.ac.uk

© The Author(s) 2021. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
in $\mathcal{H}_0$ for all $t \geq 0$, i.e., ‘nothing can happen for the first time’—a paradox pointed out by Fermi as early as 1932 [3].

How can this notion be adapted to the context of quantum field theory? The essential feature of quantum field theory is that the observables of the theory are organized by their positions in a given space-time $M$, which we shall take to be a smooth $d$-dimensional manifold with a Lorentzian metric $g = (g_{ij})$. We expect that energy, and its positivity, should also have a local aspect, which is encoded, for any quantum field theory, in the energy-momentum tensor, its most basic local observable.

In the usual formulations of quantum field theory such as [13], for each space-time point $x \in M$ there is a topological vector space $\mathcal{O}_x$ of observables at $x$, and the $\mathcal{O}_x$ fit together to form a vector bundle on $M$. The content of the theory is completely encoded in multilinear ‘functions’, called vacuum expectation values,

$$\mathcal{O}_{x_1} \times \cdots \times \mathcal{O}_{x_k} \rightarrow \mathbb{C}$$

$$\langle f_1, \ldots, f_k \rangle \mapsto \langle f_1, \ldots, f_k \rangle_{(M,g)}$$

for all sequences $\{x_1, \ldots, x_k\}$ of points in $M$. (We are oversimplifying here just for this introduction. First, the expectation values are usually required to be distributions on $M^k$, which excludes examples we can include, such as the $d = 3$ sigma model with a circle as target, for which the expectation values are hyperfunctions rather than distributions. Secondly, we shall not exclude observables such as ‘Wilson loops’, which are localized at submanifolds of space-time of dimension greater than zero.)

To come from a field theory, the vacuum expectation values must satisfy a long list of conditions such as the Wightman axioms given in [13]. These include a causality axiom, which asserts that if the points $x_1, \ldots, x_k$ are spatially separated (that is, no two can be joined by a path whose speed is never faster than light), then the expectation value is independent of the ordering of the points.

One motivation for this formulation is the ‘path-integral’ picture (cf. [4]), according to which the theory arises from a mythological superstructure consisting of a space $\Phi_M$ of ‘fields’ of some kind that are locally defined on the Lorentzian manifold $(M,g)$. In this picture, the vector space $\mathcal{O}_x$ of observables at $x$ is the space of smooth functions $f : \Phi_M \rightarrow \mathbb{C}$ such that $f(\phi)$ depends only on the restriction of $\phi$ to an arbitrarily small neighbourhood of $\phi$.

The physics of the theory is determined by an action functional $S_g : \Phi_M \rightarrow \mathbb{R}$, which notionally defines a complex-valued measure on the space $\Phi_M$, symbolically denoted by $e^{-iS_g(\phi)/\hbar} D\phi$. The parameter $\hbar$ here—the unit of action—is Planck’s constant. The vacuum expectation values are given in terms of the measure by

$$\langle f_1, \ldots, f_k \rangle_{(M,g)} = \int_{\Phi_M} f_1(\phi) \ldots f_k(\phi) e^{-iS_g(\phi)/\hbar} D\phi.$$ 

The smallness of the unit $\hbar$ of action means that the notional integral is very highly oscillatory, and so the measure on $\Phi_M$ is effectively concentrated near the critical points of the action. These points are the solutions of the classical equations of motion, and they form the classical state space of the system.

There are two ways to introduce the idea of positive energy into this picture. Both involve holomorphicity, and we shall refer to both—rather vaguely—as ‘Wick rotation’. They derive from two different ways of viewing the time $t$ in the evolution operator $U_t$ of quantum mechanics. The traditional way is to regard the possibility of extending the map $t \mapsto U_t$ to the upper half-plane as ‘creating’ a complex time-manifold with the physical time-axis at its boundary. In field theory, this leads to viewing space-time $M$ as part of the boundary of a complex manifold $M_C$, and then the positivity of
energy is expressed by the property that the vacuum expectation values (1) are the boundary values of holomorphic functions on a domain in \((M_C)^k\). This makes good sense when \(M\) is the standard Minkowski space \(\mathbb{M} \cong \mathbb{R}^{3,1}\). It is less natural in the case of a curved space-time, if only because a smooth manifold does not have a complexification (or even a way of putting it on the boundary of a complex manifold) until one chooses—non-canonically—a real-analytic structure on it. Even then, \(M\) may have only a small thickening as a complex manifold, while the holomorphic characterization of positive energy makes use of the whole upper half of the \(t\)-plane.

An alternative approach—the one we present in this paper—is to treat the time parameter \(t\) as the length of an oriented time interval, thinking of it as a one-dimensional manifold equipped with a Riemannian (or pseudo-Riemannian) metric. Then we do not need to complexify the time-manifold: we simply allow the metric on it to be complex-valued. (The authors have independently spoken about this idea from time to time since the late 1980s—see Section 4—but as far as we know no one else has pursued it systematically.) There are two reasons why the approach fits well with the path-integral picture when the time interval of quantum mechanics is replaced by the space-time \(M\) of quantum field theory. First, the usual action functionals \(S_g\) depend explicitly on the Lorentzian metric \(g\) of \(M\) in a way that makes sense when \(g\) is complex. Secondly and more importantly, the path integral is an oscillatory integral which does not converge even schematically. Its archetype is an improper Gaussian integral of the form

\[
F(A) = \int_{\mathbb{R}^n} \exp \left( \frac{i}{2} x^T A x \right) \, dx_1 \ldots dx_n,
\]

where \(A\) is a real symmetric \(n \times n\) matrix. The standard way to treat such an integral is to begin with a complex symmetric matrix \(A\) whose imaginary part is positive definite—that is, a point \(A\) of the Siegel ‘generalized upper half-plane’. For such matrices, the integral converges and defines a holomorphic function of \(A\) in the Siegel domain. The value of the original improper integral is defined as the limit as \(A\) moves to the boundary of the domain.

The main point of the present paper is to introduce an interesting domain \(\text{Met}_C(M)\) of complex-valued metrics on a smooth manifold \(M\). The positivity of energy of a quantum field theory is expressed by the property that it is defined for space-time manifolds with metrics belonging to this domain. The domain is a complexification of the manifold \(\text{Met}(M)\) of ordinary Riemannian metrics on \(M\), and the real Lorentzian metrics (but not real metrics of other signatures) are a subset of its boundary. The special role of Lorentzian signature is perhaps the most notable feature of our work. In Section 5, we shall explain how a theory defined on space-times with complex metrics gives rise, under appropriate conditions, to a theory defined for Lorentzian space-times which automatically satisfies the expected causality axiom when the Lorentzian metric is globally hyperbolic. Finally, although we avoid complexifying space-time, our approach leads us to a conjecture about a question arising in the traditional treatment of quantum field theories defined in Minkowski space \(\mathbb{M}\): how to characterize the largest domain in \((M_C)^k\) in which the vacuum expectation values are holomorphic.

The Shilov boundary

The relevant meaning of ‘boundary’ for the complex domains we are interested in is the Shilov boundary. An account of this concept can be found in [5], but our use of it will be little more than heuristic. It is the analogue for a complex domain of the ‘extremal points’ of a bounded open subset \(U\) of \(\mathbb{R}^n\). In the Euclidean situation, the extremal points are the smallest subset \(K\) of the closure of
$U$ such that every affine-linear function on $U$ attains its maximum on $K$, and the convex hull of $U$ is the set of points $x \in \mathbb{R}^n$ where

$$\inf_K f < f(x) < \sup_K f$$

for every affine function $f$. For a complex domain, we replace affine-linear functions by holomorphic functions. Thus, if $U$ is an open subset of a finite-dimensional complex Stein manifold $U^+$ (for example, an affine space or an affine algebraic variety) and the closure of $U$ in $U^+$ is a compact manifold $X$ with a piecewise-smooth boundary, then the Shilov boundary of $U$ is the smallest compact subset $K$ of $X$ with the property that for every holomorphic function $f$ defined in a neighbourhood of $X$ we have

$$\sup_U |f| = \sup_K |f|.$$  

(For a simple proof that there is a smallest such subset, see p. 67 of [5].) The Shilov boundary of a manifold is part of its topological boundary, but can be much smaller, just as, for example, the only extremal points of a Euclidean simplex are its vertices. In our examples, its real dimension will always be equal to the complex dimension of the domain. Thus, for the polydisc

$$U = \{(z_1, \ldots, z_n) \in \mathbb{C}^n : |z_i| < 1\},$$

the Shilov boundary is the torus $|z_1| = \ldots = |z_n| = 1$.

The most relevant example for us is the Siegel ‘generalized half-plane’ $\mathbb{U}_n$ of complex-valued quadratic forms on $\mathbb{R}^n$ with positive-definite real part, i.e., the $n \times n$ complex symmetric matrices $A$ with Re$(A)$ positive-definite. As so presented, $\mathbb{U}_n$ is not bounded in the vector space of matrices, but it has an alternative ‘unit disc’ description as the complex symmetric matrices $A$ such that $||A|| < 1$, or, equivalently, such that $1 - AA$ is positive definite. (The second description is obtained from the first by the Cayley transform $A \mapsto (A - 1)(A + 1)^{-1}$.) In the first description, the purely imaginary symmetric matrices lie on the boundary of the domain as the ‘generalized imaginary axis’. They form a dense open subset of the Shilov boundary, just as the imaginary axis is a dense open subset of the boundary of the usual right half-plane. But to understand the Shilov boundary in this case it is better to pass to yet another description of $\mathbb{U}_n$, as an open subset of the compact complex manifold $\text{Lag}(\mathbb{C}^{2n})$ of complex Lagrangian subspaces of a symplectic vector space $\mathbb{C}^{2n}$.

To obtain this description, let us start from a real vector space $V$ with complexification $V_\mathbb{C}$. Complex-valued quadratic forms on $V$ are the same as symmetric maps $A : V_\mathbb{C} \to V_\mathbb{C}^*$, and the graph of such a map is a Lagrangian subspace of the complex symplectic vector space $V_\mathbb{C} \oplus V_\mathbb{C}^*$. Now any Lagrangian subspace $W$ of $V_\mathbb{C} \oplus V_\mathbb{C}^*$ acquires a Hermitian inner product by the formula $\langle w_1, w_2 \rangle = iS(\bar{w}_1, w_2)$, where $S$ is the $\mathbb{C}$-bilinear symplectic form of $V_\mathbb{C} \oplus V_\mathbb{C}^*$. The Siegel domain $\mathbb{U}(V)$ consists precisely of those $W$ for which the Hermitian form of $W$ is positive-definite. The topological boundary of the domain consists of all $W$ whose Hermitian form is positive-semidefinite but not positive-definite. It is a piecewise-smooth manifold stratified by the rank of the Hermitian form. The lowest-dimensional stratum, where the Hermitian form vanishes, is the smooth compact manifold $\text{Lag}(V \oplus V^*)$ of real Lagrangian subspaces of $V \oplus V^*$. It has a dense open subset consisting of subspaces that do not intersect $V^*$: these are the graphs of the real symmetric maps $A : V \to V^*$.

In this example, and all the others we shall encounter, there is no difficulty in identifying the Shilov boundary inside the topological boundary, for by the maximum-modulus principle a point
does not belong to it if it can be written as \( f(0) \) for some non-constant holomorphic map \( f \) from a neighbourhood of \( 0 \) in \( \mathbb{C} \) to the closure of the domain. In particular, we shall meet tube domains of the form \( \mathbb{R}^N \times iC \subset \mathbb{C}^N \), where \( C \) is a convex open subset of \( \mathbb{R}^N \); for them the Shilov boundary is \( \mathbb{R}^N \times iK \), where \( K \) is the set of extremal points of \( C \).

2. The domain of complex metrics

A Riemannian metric on a manifold \( M \) is a positive-definite symmetric bilinear form \( g : T_x \times T_x \to \mathbb{R} \) on the tangent space \( T_x \) at each point \( x \in M \). The metrics we shall consider will be defined by symmetric \( \mathbb{R} \)-bilinear maps \( g : T_x \times T_x \to \mathbb{C} \) at each point, with an appropriate generalization of the positivity condition.

To see what condition we should require, let us consider the simplest example of a field theory: a free real scalar field of mass \( m \). Then, the space of ‘fields’ \( \Phi_M \) is the vector space \( C^\infty(M; \mathbb{R}) \) of smooth functions, and in the exponent of the path-integral we have the quadratic form

\[
iS_g(\phi) = \frac{1}{2} \int_M (d\phi \wedge *d\phi + m^2 \phi \wedge *)
= \frac{1}{2} \int_M \left\{ \sum g^{ij} \frac{\partial \phi}{\partial x^i} \frac{\partial \phi}{\partial x^j} + m^2 \phi^2 \right\} (\det g)^{1/2} |dx^1 \ldots dx^d|.
\]

Here, \((g^{ij})\) denotes the inverse of the matrix \( g = (g_{ij}) \), and \(*\) is the Hodge star-operator defined by the metric, which takes differential forms of degree \( p \) to forms of degree \( d - p \) twisted by the orientation bundle. (We shall not assume the space-time \( M \) is orientable.) In particular, the star-operator takes the constant function \( 1 \) to the volume element

\[
*1 = \text{vol}_g = (\det g)^{1/2}|dx^1 \ldots dx^d| \quad (3)
\]

Notice that for a Lorentzian metric \( g \) the volume element \(*1\) is pure imaginary. This agrees with the fact that the ‘action’ \( S_g \) should be real for a Lorentzian manifold. We want the real part of the quadratic form \( iS_g \) to be positive-definite for all the complex metrics we allow. This imposes two conditions. First, we need the real part of the twisted \( d \)-form \( \text{vol}_g \) defined by the formula (3) to be a positive volume-form on \( M \). We therefore require that \( \det g \), which is invariantly defined up to multiplication by a positive real number, is not real and negative, and we choose \((\det g)^{1/2}\) to have positive real part.

The second condition we need is that the real part of the matrix \((\det g)^{1/2}g^{-1}\), or equivalently of the inverse matrix \((\det g)^{-1/2}g\), is positive-definite. The two conditions together would give us a domain whose Shilov boundary (like that of the Siegel generalized half-plane) contains indefinite real quadratic forms of all signatures, and not only the Lorentzian ones. But we shall impose further conditions. A clue to what more is needed comes from the theory of the electromagnetic field on \( M \), with its field-strength given by a real 2-form \( F \) on \( M \), and with the action-functional

\[
iS_g(F) = \frac{1}{2} \int_M F \wedge *F.
\]

The Hodge \(*\)-operator makes sense for a complex metric: for a \( p \)-form \( \alpha \) we define a twisted \((d - p)\)-form \(*\alpha\) by taking the inner-product of \( \alpha \) with \( \text{vol}_g = *1 \), using the complex inner-product \( g \). We
regard $\text{vol}_x$ as an element of the complex line $|\wedge^d(T^*_x)|_\mathbb{C}$, where $|\wedge^d(T^*_x)|$ is the tensor product of $\wedge^d(T^*_x)$ with the real orientation line of $T_x$, and $|\wedge^d(T^*_x)|_\mathbb{C}$ is its complexification, but with the convention that the orientation-reversing automorphisms of $T_x$ act antilinearly. We say that an element of the real part of the line is positive if it is a positive volume-element.

For the electromagnetic field, we need the real part of the quadratic form $^d(T^*_x)!j^d(T^*_x)j_\mathbb{C}$ given by $F\times F \to F$ to be positive-definite.

This makes it natural, if we are going to consider space-time manifolds $M$ of all dimensions, to propose

**Definition 2.1** On a $d$-dimensional real vector space $V$ a quadratic form $g: V \to \mathbb{C}$ is called an allowable complex metric if, for all degrees $p \geq 0$, the real part of the quadratic form $\wedge^p(V^*) \to |\wedge^d(V^*)|_\mathbb{C}$ given by $\alpha \mapsto \alpha \wedge \ast \alpha$ is positive-definite.

Fortunately, this definition has an equivalent formulation which is much more explicit and illuminating.

**Theorem 2.2** Definition 2.1 is equivalent to: there is a basis of the real vector space $V$ in which the quadratic form $g$ can be written

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_d y_d^2,$$

where $y_i$ are coordinates with respect to the basis and $\lambda_i$ are non-zero complex numbers, not on the negative real axis, such that

$$|\arg(\lambda_1)| + |\arg(\lambda_2)| + \ldots + |\arg(\lambda_d)| < \pi. \quad (4)$$

The complex-valued quadratic forms $g: V \to \mathbb{C}$ on a real vector space $V$ which satisfy the conditions of Definition 2.1 or Theorem 2.2 form an open subset $Q_\mathbb{C}(V)$ of the complex vector space $S^2(V^*_\mathbb{C})$. It follows from Theorem 2.2 that the real inner products with signature $(d-1,1)$—but not those with other signatures—lie on the boundary of the domain $Q_\mathbb{C}(V)$. For if the metric is real then each $|\arg(\lambda_i)|$ is either 0 or $\pi$, and the inequality (4) shows that at most one of the $|\arg(\lambda_i)|$ can become $\pi$ on the boundary.

Another consequence of (4) is that

$$\max \arg \lambda_i - \min \arg \lambda_i < \pi,$$

which shows that when $\nu$ runs through $V$ the complex numbers $g(\nu)$ form a closed convex cone in $\mathbb{C}$ disjoint from the open negative real axis. In particular, $g(\nu)$ can never be real and negative.

Using the criterion mentioned at the end of Section 1, we see that the real Lorentzian metrics—and no other non-degenerate metrics—belong to the Shilov boundary of $Q_\mathbb{C}(V)$, when it is regarded as a
bounded domain in an affine variety (cf. the proof of 2.7). Indeed, if \( g = \sum \lambda_j \gamma_j^2 \) is a complex metric for which the inequality (4) becomes an equality, and at least two of the eigenvalues \( \lambda_j \) and \( \lambda_k \) are not on the negative real axis, then (after rescaling the basis vectors \( e_j \) and \( e_k \) so that \( |\lambda_j| = |\lambda_k| = 1 \)) we get a holomorphic curve through \( g \), in the closure of \( Q_C(V) \), by changing \( \lambda_j \) to \( (\lambda_j)^{1+\varepsilon} \) and \( \lambda_k \) to \( (\lambda_k)^{1-\varepsilon} \), where \( \varepsilon \) is +1 or −1 according as the arguments of \( \lambda_j \) and \( \lambda_k \) have the same or opposite signs.

In fact, the Shilov boundary of \( Q_C(V) \) contains two disjoint copies of the space of Lorentzian metrics on \( V \), for an eigenvalue \( \lambda \) can approach the negative real axis either from above or from below. The two copies are interchanged by the complex-conjugation map on \( Q_C(V) \). Because of our choice to make the orientation-reversing elements of \( GL(V) \) act antilinearly on the orientation-line of \( V \), we can say that the nondegenerate points of the Shilov boundary of \( Q_C(V) \) are the time-oriented Lorentzian metrics.

We define the space \( Met_C(M) \) of allowable complex metrics on a smooth manifold \( M \) as the space of smooth sections of the bundle on \( M \) whose fibre at \( x \) is \( Q_C(T_x) \).

Before giving the surprisingly simple proof of Theorem 2.2, let us say what motivated the two different-looking conditions. The desire to make the real parts of natural quadratic action functionals positive-definite hardly needs further comment, but choosing to focus on the ‘higher abelian gauge field’ actions \( \alpha \wedge * \alpha – the ‘Ramond-Ramond’ fields of superstring theory—may well seem arbitrary. Why not allow other kinds of tensor fields? Our conditions do not imply that they will be positive-definite. Witten has kindly suggested to us a justification for our focus, based on properties of the classical energy-momentum tensor explained in [15]. Including the higher gauge theories does, in any case, impose an upper bound on the class of complex metrics we can allow, for the partition functions of these theories on a \( d \)-dimensional torus \( M \) with a flat Riemannian metric \( g \) are explicitly known (cf. [8], [14](4.4)), and we can see to which complex metrics they can be analytically continued. The gauge-equivalence classes of fields form an infinite-dimensional Lie group which is a product of a torus, a lattice, and an infinite-dimensional real vector space, and the partition function is the product of three corresponding factors. More precisely, an abelian gauge \((p−1)\)-field \( A \) has a field-strength \( F_A \), a closed \( p \)-form on \( M \) with integral periods, which determines \( A \) up to the finite-dimensional torus \( H^{p−1}(M; \mathbb{T}) \) of flat gauge fields with \( F_A = 0 \). The space of fields is therefore a product

\[
H^{p−1}(M; \mathbb{T}) \times \Phi_p \times \Gamma_p,
\]

where \( \Phi_p \) is the vector space of exact \( p \)-forms on \( M \), and \( \Gamma_p \cong \text{Harm}_p^C(M) \cong H^p(M; \mathbb{Z}) \) is the finite-dimensional lattice of harmonic (and hence constant) \( p \)-forms with integral periods. The partition function is a Gaussian integral on this product: the torus of flat fields contributes its volume (for an appropriate metric determined by the geometry of \( M \)), the lattice \( \Gamma_p \) of harmonic \( p \)-forms contributes its theta-function

\[
\sum_{\alpha \in \Gamma_p} \exp \left( -\frac{1}{2} \int_M \alpha \wedge * \alpha \right),
\]

while the vector space \( \Phi_p \) contributes an ‘analytic torsion’ which is a power of the determinant of the Laplace operator acting on smooth functions on \( M \) (with the zero-eigenvalue omitted)—an analogue of the Dedekind eta-function, but with the lattice of characters of the torus \( M \) replacing the lattice \( \mathbb{Z} + \tau \mathbb{Z} \subset \mathbb{C} \). Of these three factors, the first clearly extends holomorphically to the space of all flat complex metrics on \( M \), and the analytic torsion can be continued to a non-vanishing holomorphic
function in the open set of complex metrics $g$ for which $(\det g)^{-1/2}$ belongs to the Siegel domain $\mathbb{U}(V)$; however, the theta-function cannot be continued beyond those metrics for which the real part of $\int \alpha \wedge \ast \alpha$ is positive.

Approaching from the opposite direction, the inequality (4) is motivated by the traditional analytical continuation of vacuum expectation values to an open subset of the $k$-fold product of complexified Minkowski space $\mathbb{M}_C$. The Wightman axioms imply that the expectation values extend holomorphically to a domain $\mathcal{U}_k$ called the ‘permutated extended tube’, which is functorially associated to the complex vector space with its $\mathbb{C}$-bilinear metric. (A set of points $x_1, \ldots, x_k$ belongs to $\mathcal{U}_k$ if, after ordering them suitably, there is an element $\gamma$ of the complexified Lorentz group such that the imaginary part of $\gamma(x_i - x_{i+1})$ belongs to the forward light-cone for each $i$.) It is a basic result in the Wightman theory (cf. [13] or [7](2.1)) that $\mathcal{U}_k$ contains the configuration space $\text{Conf}_k(\mathbb{E})$ of all $k$-tuples of distinct points of the standard Euclidean subspace $\mathbb{E} \subset \mathbb{M}_C$. For a $d$-dimensional real vector space $V$ with a complex metric, the complexification $V_C$ is isomorphic to $\mathbb{M}_C$, uniquely up to a complex Lorentz transformation, and so the domain $\mathcal{U}_k(V)$ is well-defined in $(V_C)^k$. In the next section, we shall give a definition of a quantum field theory on space-times $M$ with complex metrics: it implies that the expectation values are smooth functions on the configuration spaces $\text{Conf}_k(M)$ of distinct $k$-tuples in $M$. This makes it natural to ask which (constant) complex metrics on $V$ have the property that the configuration space $\text{Conf}_k(V)$ is contained in the holomorphic envelope of $\mathcal{U}_k(V)$, i.e., the largest Stein manifold to which all holomorphic functions on $\mathcal{U}_k(V)$ automatically extend. The original motivation of condition (4) was

**Proposition 2.3** If a complex metric on a $d$-dimensional real vector space $V$ satisfies condition (4), then $\text{Conf}_k(V)$ is contained in the holomorphic envelope of $\mathcal{U}_k(V)$.

We shall postpone the proof of this result to an appendix at the end of this section.

**Proof of Theorem 2.2** The first point is to show that a quadratic form that satisfies the conditions of Definition 2.1 can be written in the diagonal form $\sum \lambda_i y_j^2$ with respect to real coordinates $y_j$ on $V$. To diagonalize a complex form $g = A + iB$ with respect to a real basis is to diagonalize its real and imaginary parts simultaneously, which is possible if either $A$ or $B$—or, more generally, a real linear combination of them such as the real part of $(\det g)^{-1/2}g$—is positive-definite. However, 2.1, applied when $p = 1$, implies that the real part of $(\det g)^{-1/2}g$ is positive.

Suppose now that $g$ is diagonalized with respect to a basis $\{e_i\}$ of $V$. Then the form $\alpha \mapsto \alpha \wedge \ast \alpha$ on $\wedge^p(V^\ast)$ is diagonal with respect to the basis $\{e^*_S = e^*_{i_1} \wedge \ldots \wedge e^*_{i_p}\}$, where $\{e^*_i\}$ is the dual basis to $\{e_i\}$, and $S$ runs through $p$-tuples $S = (i_1, \ldots, i_p)$. The value of the form $\alpha \wedge \ast \alpha$ on the basis element $e^*_S$ is

$$\left(\lambda_1 \ldots \lambda_d\right)^{1/2} \prod_{i \in S} \lambda_i^{-1},$$

which has positive real part if its argument

$$\frac{1}{2} \left\{ \sum_{i \in S} \arg(\lambda_i) - \sum_{i \notin S} \arg(\lambda_i) \right\}.$$
lies in the open interval \((-\pi/2, \pi/2)\). But to say that this is true for every subset \(S\) of \(\{1, \ldots, d\}\) is precisely condition (4).

The proof of Theorem 2.2 shows that to give an element \(g\) of \(Q_C(V)\) is the same as to give a finite sequence \(\theta_1 \geq \theta_2 \geq \ldots \geq \theta_m\) in the interval \((-\pi, \pi)\) together with a decomposition

\[ V = V_1 \oplus \ldots \oplus V_m \]

such that

\[ \sum_k \dim V_k \cdot |\theta_k| < \pi. \]

Thus on \(V_k\) the bilinear form \(g\) is \(e^{i\theta_k}\) times a real positive-definite form. The only ambiguity in this description is that if, say, \(\theta_k = \theta_{k+1}\), we can replace \(V_k\) by \(V_k \oplus V_{k+1}\) and omit \(\theta_{k+1}\) and \(V_{k+1}\). This means that the subspace \(P = \bigoplus e^{-i\theta_k/2} V_k\) of the complexification \(V_C\) of \(V\) is canonically associated with the form \(g\). On the real subspace \(P\), the complex bilinear form \(g\) is real and positive-definite. Our argument gives us canonical isomorphisms

\[ V = \exp(i\pi \Theta/2)(P) \subset P_C = V_C, \]

where \(\Theta : P \to P\) is the self-adjoint operator which is multiplication by \(\theta_k\) on \(P_k = e^{-i\theta_k/2}V_k\). Condition (4) becomes the assertion that \(||\Theta||_1 < 1\), where \(||\Theta||_1\) is the trace-norm of \(\Theta\), i.e., the sum of the absolute values of its eigenvalues.

This shows that the space \(Q_C(V)\) is parameterized by the pairs \((g_0, \Theta)\), where \(g_0\) is a positive-definite inner-product on \(V\) and \(\Theta\) belongs to the convex open set \(\Pi(V, g_0)\) of operators in \(V\) which are self-adjoint with respect to \(g_0\) and satisfy \(||\Theta||_1 < 1\), i.e., the interior of the convex hull of the rank 1 orthogonal projections in \(V\). In fact we have proved

**Proposition 2.4** \(Q_C(V)\) is a fibre-bundle over the space of positive-definite inner-products on \(V\) whose fibre at a point \(g_0\) is \(\Pi(V, g_0)\). Equivalently, choosing a reference inner-product on \(V\), we have

\[ Q_C(V) \cong GL(V) \times_{O(V)} \Pi(V). \]

In particular, \(Q_C(V)\) is contractible.

It is an important fact that an allowable complex metric on \(V\) remains allowable when restricted to any subspace \(W\) of \(V\). This follows from an analogous property of the trace-norm, but we shall give a direct proof, as its point of view on the angles \(\theta_i\) as critical values helps give a feeling for allowable complex metrics.

**Proposition 2.5** If \(g \in Q_C(V)\) and \(W\) is any vector subspace of \(V\), then \(g|W\) belongs to \(Q_C(W)\).

**Proof.** For any \(g \in Q_C(V)\), the function \(v \mapsto \arg(g(v))\) is a smooth map from the real projective space \(\mathbb{P}(V)\) to the open interval \((-\pi, \pi) \subset \mathbb{R}\). By rescaling the basis elements \(\{e_k\}\) we can write \(g\) as \(\sum e^{i\theta_k} y_k^2\). The numbers \(\theta_k\) are precisely the critical values of \(\arg(g)\). We shall order the basis elements so that
\[ \pi > \theta_1 \geq \theta_2 \geq \ldots \geq \theta_d > -\pi. \]

For each vector subspace \( A \) of \( V \), let us write \( \theta^A \) and \( \theta_A \) for the supremum and infimum of \( \arg(g) \) on \( \mathbb{P}(A) \). Then, we have

\[ \theta_k = \sup \{ \theta_A : \dim(A) = k \} = \inf \{ \theta^A : \dim(A) = d - k + 1 \}. \]

It is enough to prove Proposition 2.5 when \( W \) is a subspace of \( V \) of codimension 1. In that case, the preceding characterization of the critical values shows that if \( \theta'_1 \geq \ldots \geq \theta'_{d-1} \) are the critical values of \( \arg(g|W) \), we have \( \theta_k \geq \theta'_k \geq \theta_{k+1} \). The critical values for \( g|W \) therefore interleave those for \( g \):

\[ \theta_1 \geq \theta'_1 \geq \theta_2 \geq \theta'_2 \geq \ldots \geq \theta_{d-1} \geq \theta'_{d-1} \geq \theta_d. \]

This implies that \( \sum |\theta'_k| \leq \sum |\theta_k| < \pi \), as we want. \( \square \)

In Section 5, we shall need the following variant of the preceding formulation. Suppose that \( Z \) is a \( d \)-dimensional complex vector space with a nondegenerate quadratic form \( g \). (Any such pair \((Z, g)\) is isomorphic to \( \mathbb{C}^d \) with the standard form \( \sum z_k^2 \).) Let \( \mathcal{R}(Z) \) denote the space of all \( d \)-dimensional real subspaces \( A \) of \( Z \) such that \( g|A \) belongs to \( Q_\mathbb{C}(A) \). This is an open subset of the Grassmannian of all real subspaces of \( Z \). If \( Z_\mathbb{R} \) is any \( d \)-dimensional real vector subspace of \( Z \) for which \( g|Z_\mathbb{R} \) is real and positive-definite, then the projection \( A \subset Z \to Z_\mathbb{R} \) is an isomorphism, for any non-zero element of its kernel would have the form \( iv \) with \( v \in Z_\mathbb{R} \), and so \( g(iv) \) would be real and negative, which cannot happen if \( g|A \) is allowable.

**Proposition 2.6** The space \( \mathcal{R}(Z) \) is contractible and is isomorphic to

\[ O_\mathbb{C}(Z) \times_{O(Z_\mathbb{R})} \Pi(Z_\mathbb{R}). \]

**Proof.** This is essentially a reformulation of what has been said, but it may be helpful to relate the spaces \( Q_\mathbb{C}(V) \) and \( \mathcal{R}(Z) \) by considering, for a complex quadratic vector space \((Z, g)\) as above, the intermediate space \( \mathcal{R}(V; Z) \) of \( \mathbb{R} \)-linear embeddings \( f : V \to Z \) of the real vector space \( V \) such that \( f^*(g) \) is allowable. This space has two connected components, corresponding to the orientation of the projection \( V \to Z_\mathbb{R} \).

The groups \( \text{GL}(V) \) and \( O_\mathbb{C}(Z) \) act by right- and left-composition on \( \mathcal{R}(V; Z) \), and each action is free. Thus, \( \mathcal{R}(V; Z) \) is at the same time a principal \( \text{GL}(V) \)-bundle with base \( \mathcal{R}(Z) \) and a principal \( O_\mathbb{C}(Z) \)-bundle with base \( Q_\mathbb{C}(V) \). However, the Lie groups \( \text{GL}(V) \) and \( O_\mathbb{C}(Z) \) are homotopy equivalent to their maximal compact subgroups, that is, in both cases to the compact orthogonal group \( O_d \). More precisely, the contractibility of \( Q_\mathbb{C}(V) \) implies that \( \mathcal{R}(V; Z) \) is homotopy-equivalent to the fibre \( O_\mathbb{C}(Z)f \) for any \( f \in \mathcal{R}(V; Z) \). If we choose \( f \) so that \( f^*(g) \) is a positive-definite real form on \( V \), this gives us a homotopy-equivalence \( O(V) \to O_\mathbb{C}(Z)f \to \mathcal{R}(V; Z) \). However, \( O(V) \) is also contained in and equivalent to the fibre \( f\text{GL}(V) \) of the other fibration \( \mathcal{R}(V; Z) \to \mathcal{R}(Z) \), which implies the contractibility of its base \( \mathcal{R}(Z) \). \( \square \)

The last property of \( Q_\mathbb{C}(V) \) that we shall record briefly, for the sake of experts, is

**Proposition 2.7** The domain \( Q_\mathbb{C}(V) \) is holomorphically convex, that is, a ‘domain of holomorphy’.
\textbf{Proof.} The Siegel domain $\mathcal{U}(V)$ of complex-valued inner-products with positive-definite real part on a real vector space $V$ is known to be a domain of holomorphy in $S^2(V^\ast_C)$. So, therefore, is the product

$$\prod_{0 \leq p \leq d/2} \mathcal{U}(\wedge^p(V))$$

inside its ambient complex vector space. The space $Q_C(V)$ is the intersection of this product domain with the affine variety which is the natural embedding of $S^2(V^\ast_C)$ in this ambient vector space, and so it too is a domain of holomorphy. \qed

\textbf{The two-dimensional case}

The case $d = 2$ is especially simple because then the matrix $(\det g)^{-1/2}g$ depends only on the conformal structure and decouples from the volume element.

A non-degenerate complex inner product $g$ on a two-dimensional real vector space $V$ is determined up to a scalar multiple by its two distinct null-directions in the complexified space $V_C$. We can think of these as two points of the Riemann sphere $\mathbb{P}(V_C)$. Then, $(\det g)^{-1/2}g$ has positive real part precisely when the two points lie one in each of the open hemispheres of the sphere $\mathbb{P}(V_C)$ separated by the real equatorial circle $\mathbb{P}(V)$. When the two points move to distinct points of the equator, we get a Lorentzian inner-product, with its two light-directions in $\mathbb{P}(V)$.

A point of the sphere $\mathbb{P}(V_C)$ not on the equator can be regarded as a complex structure on the real vector space $V$, and the two hemispheres correspond to the two possibilities for the orientation which a complex structure defines. On a smooth surface $\Sigma$, any almost-complex structure is integrable, so a point of $\text{Met}_C(\Sigma)$ is a pair of complex structures on $\Sigma$ of opposite orientations, together with a complex volume element. The Riemannian metrics are those for which the two complex structures are complex-conjugate to each other, and the volume element is real.

When $d = 2$ the domain $Q_C(V)$ is, thus, a three-dimensional polydisc, one disc for each of the complex structures, and the third for the volume-element.

\textbf{The one-dimensional case: electric circuits}

Our concept of an allowable complex metric does not at first look interesting in the one-dimensional case, but if we allow singular one-manifolds—identified with finite graphs $M$—we find that complex metrics arise naturally in electrical circuit theory. A Riemannian metric on $M$ is determined (up to isometry) by the assignment of a positive real number to each edge of the graph, and can be interpreted as its \textit{resistance} when the edge is regarded as a wire in an electrical circuit. A state of the system (perhaps with current entering or leaving at each node) is determined by a continuous potential function $\phi : M \to \mathbb{R}$ which is smooth on each closed edge, and whose gradient is the current flowing in the circuit. Because $\phi$ is determined only up to adding a constant, we shall normalize it by $\int_M \phi = 0$. The energy of a state is

$$\frac{1}{2} \int_M ||\nabla \phi||^2 ds,$$

and so the system can be regarded as a free massless field theory on the graph: in particular, the vacuum expectation value $\langle \phi(x)\phi(y) \rangle$, when $x$ and $y$ are two nodes of the graph, is the ratio of the
potential-difference \(\phi(x) - \phi(y)\) to the current flowing in at \(x\) and out at \(y\) when no current is allowed to enter or leave at other nodes.

We encounter complex metrics when we consider a circuit in which an alternating current with frequency \(\omega\) is flowing, and in which each branch has not only a resistance \(R\) but also a positive inductance \(L\) and a positive capacitance \(C\). In that situation, the volume element \(\sqrt{g} = R\) is replaced by the impedance

\[
\sqrt{g} = R + i\omega L + 1/i\omega C,
\]

a complex number which defines an allowable metric because \(\text{Re}\sqrt{g} > 0\).

Quite apart from electric circuitry, however, singular one-dimensional manifolds with allowable complex metrics can arise in quantum field theory as the Gromov–Hausdorff limits of non-singular space-times of higher dimension. For example, if we embed a smooth graph \(M\) in \(\mathbb{R}^3\), then for almost all sufficiently small \(\varepsilon > 0\) the boundary of the \(\varepsilon\)-neighbourhood of \(M\) is a smooth surface \(M_\varepsilon\) whose limit is \(M\) as \(\varepsilon \to 0\): this is one way of viewing the passage from closed string theory to quantum field theory.

### Appendix to Section 2: proof of 2.3

If \(V\) is a real vector space with an allowable complex metric, then the preceding discussion shows that it can be identified with the subspace

\[
V = \exp(i\Theta/2)(E)
\]

of \(\mathbb{M}_C\). Here \(E = \mathbb{R}^d\) is the standard Euclidean subspace of \(\mathbb{M}_C\), and \(\Theta\) is a real diagonal matrix whose entries \(\theta_1, \ldots, \theta_d\) belong to the ‘generalized octahedron’ \(\Pi_0 \subset \mathbb{R}^d\) consisting of those \(\Theta\) whose diagonal entries \(\theta_1, \ldots, \theta_d\) satisfy the inequality (4). We want to prove that \(\exp(i\Theta/2)\) maps each \(k\)-tuple \(x = \{x_1, \ldots, x_k\}\) of distinct points of \(E\) to a point of the holomorphic envelope \(\mathcal{U}_k\) of the Wightman permuted extended tube \(\mathcal{U}_k\). In fact, we shall prove the stronger statement that \(\exp(i\Theta/2)(x) \in \mathcal{U}_k\) when \(\Theta\) is a complex diagonal matrix with \(\text{Re}(\Theta) \in \Pi_0\).

The crucial fact is that \(\Pi_0\) is the convex hull of its intersection \(\Pi_{00}\) with the coordinate axes in \(\mathbb{R}^d\), (i.e., \(\Pi_{00}\) consists of the diagonal matrices with only one entry \(\theta\), non-zero, and \(-\pi < \theta < \pi\)). Our strategy is to show that \(\exp(i\Theta/2)(x) \in \mathcal{U}_k\) when \(\text{Re}(\Theta) \in \Pi_{00}\), and to deduce that the same is true when \(\text{Re}(\Theta)\) belongs to the convex hull \(\Pi_0\). The essential tool is Bochner’s ‘tube theorem’ ([5] Thm 2.5), which asserts that if \(P\) is a connected open subset of \(\mathbb{R}^d\) then a holomorphic function defined in the tube domain \(P \times i\mathbb{R}^d\) extends holomorphically to the tube domain \(P' \times i\mathbb{R}^d\), where \(P'\) is the convex hull of \(P\).

Having fixed a \(k\)-tuple \(x\) in \(\mathbb{M}_C\), let us first show that if \(\text{Re}(\Theta) \in \Pi_{00}\) then \(\exp(i\Theta/2)(x)\) is contained in \(\mathcal{U}_k\). Suppose that the non-zero diagonal element of \(\Theta\) is in the \(r\)th place. Because \(\mathcal{U}_k\) is invariant under the orthogonal group \(O(E)\) we can assume that the \(r\)th basis vector \(e_r\) of \(E\) is the Wick-rotated time-axis of \(\mathbb{M}\), so that \(e_r\) belongs to \(iC\), where \(C\) is the forward light-cone in \(\mathbb{M}\). With respect to the real structure \(\mathbb{M}_C = \mathbb{M} \oplus i\mathbb{M}\), the imaginary part of the \(k\)-tuple

\[
y = \exp(i\Theta/2)(x)
\]

lies on the line \(\mathbb{R}e_r\), and so, after ordering the points appropriately, \(y\) will belong to the forward tube in \(\mathbb{M}_C\) providing the points of \(x\) have distinct \(r\)th coordinates. However, if the \(r\)th coordinates of \(\text{Im}(y)\)
are not distinct, we can make them so by choosing a unit vector \( e \in \mathbb{E} \) perpendicular to \( e_r \) such that the coordinates \((x, e)\) are distinct, and rotating the \( k \)-tuple \( y \) by a small amount in the \( \{e, e_r\} \)-plane, again using the \( O(\mathbb{E}) \)-invariance of \( \mathcal{U}_k \).

We now know that \( \mathcal{U}_k \) contains an open neighbourhood of \( \Pi_0 \times \mathbb{R}^d \) in \( \mathbb{C}^d \). To apply Bochner's theorem, we need to know that the envelope \( \mathcal{U}_k \) contains a tube \( P \times \mathbb{R}^d \), where \( P \) is an open neighbourhood of \( \Pi_0 \) in \( \mathbb{R}^d \). In fact it is enough, by induction, to treat the case \( d = 2 \), for that case, together with Bochner's theorem, implies that a function holomorphic in a neighbourhood of \( (\Pi_0(\mathbb{R}^d) \cup \Pi_0(\mathbb{R}^{d-r}) \times \mathbb{R}^d) \) is holomorphic in a neighbourhood of \( (\Pi_0(\mathbb{R}^r+1) \cup \Pi_0(\mathbb{R}^{d-r-1})) \times \mathbb{R}^d \).

To reduce the \( d = 2 \) case to the standard Bochner theorem, it is enough to prove the following

**Lemma 2.8** Let \( L \) be the \( L \)-shaped subset \((\{0\} \times [0, 1]) \cup ([0, 1] \times \{0\})\) of the quadrant \((\mathbb{R}_+)^2\). Then, any holomorphic function \( F \) defined in a neighbourhood of \( L \times \mathbb{R}^2 \subset \mathbb{C}^2 \) can be extended holomorphically to \( P \times \mathbb{R}^2 \), where \( P \) is the intersection of \((\mathbb{R}_+)^2\) with a neighbourhood of \( L \) in \( \mathbb{R}^2 \).

**Proof.** For any \( t \in (0, 1/2) \), we define \( D_t \) as the intersection of the two unit discs \( \{z \in \mathbb{C} : |z - (1 - t)| \leq 1\} \) and \( \{z \in \mathbb{C} : |z + (1 - t)| \leq 1\} \). Then, we define \( f : D_t \rightarrow \mathbb{C}^2 \) by

\[
    f(z) = (-\log((1 - t) - z), -\log((1 - t) + z)).
\]

The map \( f \) is a holomorphic embedding in a neighbourhood of \( D_t \) in \( \mathbb{C} \), and \( \text{Re } f(\partial D_t) \) is contained in the coordinate axes of \( \mathbb{R}^2 \). If we choose \( T = (1 - e^{-1})/2 \) then \( \text{Re } f(\partial D_T) \) is precisely the closure of \( L \).

For any \( \eta \in \mathbb{R}^2 \), define \( f_\eta : D_T \rightarrow \mathbb{C}^2 \) by \( f_\eta(z) = f(z) + i\eta \). Then, the holomorphic map \( F \) is defined in a neighbourhood of the curve \( f_\eta(\partial D_T) \), and if we can show that \( F \circ f_\eta \) extends holomorphically over \( D_T \) then we shall have continued \( F \) analytically to the tube domain \( f(D_T) + i\mathbb{R}^2 \), and the proof will be complete.

When a function \( F \) is holomorphic in an open domain containing the boundary of a holomorphically embedded disc—in this case \( f_\eta(D_T) \)—then to show that \( F \) can be extended over the whole disc the standard method is to show that the disc can be moved holomorphically, keeping its boundary within the domain of \( F \), until the whole disc is contained in the domain of \( F \); the Cauchy integral formula then defines the desired extension. In our case, we can deform \( f_\eta(D_T) \) through the family \( f_\eta(D_t) \) as \( t \) decreases from \( T \) towards 0. As \( t \downarrow 0 \) the domain \( D_t \) shrinks to the origin in \( \mathbb{C} \), and \( f_\eta(D_t) \rightarrow i\eta \), which is contained in the domain of \( F \). \( \square \)

### 3. Quantum field theories as functors

The traditional Wightman approach to quantum field theory is not well-adapted to important examples such as gauge theories, especially when the space-time is not flat. Another formulation—potentially more general—views a \( d \)-dimensional field theory as something more like a group representation, except that the group is replaced by a category \( \mathcal{C}_d^\mathbb{R} \) of space-time manifolds. The guiding principle of this approach is to preserve as much as possible of the path-integral intuition. We shall present it very briefly here, with minimal motivation.

Roughly, the objects of the category \( \mathcal{C}_d^\mathbb{R} \) are compact smooth \((d - 1)\)-dimensional manifolds \( \Sigma \) equipped with complex metrics \( g \in \text{Met}_c(\Sigma) \). A morphism from \( \Sigma_0 \) to \( \Sigma_1 \) is a cobordism \( M \) from \( \Sigma_0 \) to \( \Sigma_1 \), also with a complex metric. We shall write \( M : \Sigma_0 \sim \Sigma_1 \) to indicate a cobordism. Composition
of morphisms is by concatenation of the cobordisms. The reason for the word ‘roughly’ is that, because there is no canonical way to give a smooth structure to the concatenation of two smooth cobordisms, we must modify the definition slightly so that an object of $C_d^C \Sigma$ is not a $(d - 1)$-manifold but rather is a germ of a $d$-manifold along a given $(d - 1)$-manifold $\Sigma$—that is, $\Sigma$ is given as a closed submanifold of a $d$-manifold $U$, but any two open neighbourhoods of $\Sigma$ in $U$ define the same object of $C_d^C \Sigma$. We require $\Sigma$ to be two-sided in $U$ and equipped with a co-orientation, which tells us which side is incoming and which is outgoing. (Nevertheless, we shall usually suppress the thickening cobordism category.)

To say that the functor is ‘holomorphic’ means that, for a given smooth manifold-germ $\Sigma \subset U$, the topological vector spaces $E_\Sigma$ form a locally trivial holomorphic vector bundle on the complex manifold $\text{Met}_C(U)$ of complex metrics on $U$, and that the maps $Z_M : E_{\Sigma_0} \to E_{\Sigma_1}$ define a morphism of holomorphic vector bundles on the manifold $\text{Met}_C(M)$ (to which the bundles $\{E_{\Sigma_0}\}$ and $\{E_{\Sigma_1}\}$ are pulled back).

In practice, theories are usually defined on cobordism categories where the manifolds are required to have additional structure such as an orientation or a spin-structure. These can easily be included, but are not relevant to our account. For the same reason, we do not mention that, for a theory including fermions, the vector spaces $E_\Sigma$ will have a mod 2 grading, and the usual sign-conventions must be applied when we speak of their tensor products.

Because our objects $\Sigma \subset U$ are really germs of $d$-manifolds, we automatically have a family of cobordisms $\Sigma' \sim \sim \Sigma$ embedded in $U$, each diffeomorphic to the trivial cobordism $\Sigma \times [0,1]$ with the outgoing boundary $\Sigma \times \{1\}$ corresponding to $\Sigma \subset U$. These cobordisms can be ordered by inclusion, giving us a direct system of objects $\Sigma'$ with cobordisms to $\Sigma$. Similarly, looking downstream rather than upstream, we have a family of cobordisms $\Sigma \sim \sim \Sigma''$ contained in $U$, giving us an inverse system of objects $\Sigma''$ to which $\Sigma$ maps. For any field theory, therefore, there are natural maps

$$\lim E_{\Sigma'} \to E_\Sigma \to \lim E_{\Sigma''},$$

which we shall write

$$\hat{E}_\Sigma \to E_\Sigma \to \hat{E}_\Sigma,$$

introducing the notations $\hat{E}_\Sigma = \lim E_{\Sigma'}$ and $\hat{E}_\Sigma = \lim E_{\Sigma''}$ for the upstream and downstream limits.

We shall assume the functor has the continuity property that each of these maps is injective with dense image. The space $\hat{E}_\Sigma$, being the inverse-limit of a countable sequence of nuclear maps of Fréchet spaces, is a nuclear Fréchet space. (A very useful concise account of nuclear spaces can be
found in [2].) The other space $\tilde{E}_\Sigma$ is also nuclear, but usually not metrizable; it is the dual of the nuclear Fréchet space $\hat{E}_\Sigma^*$, where $\Sigma^*$ denotes the germ $\Sigma$ with its co-orientation reversed. As this is such a basic point, we have included a proof as an Appendix at the end of this section.

When we have a cobordism $M: \Sigma_0 \rightsquigarrow \Sigma_1$ we automatically get maps $E_{\Sigma_0} \to \tilde{E}_\Sigma$ and $\tilde{E}_{\Sigma_0} \to \tilde{E}_{\Sigma_1}$. The space $\tilde{E}_\Sigma$ with which we began plays only a provisional role in the theory, serving to construct the fundamental nuclear spaces between which it is sandwiched.

The essential requirement we place on the functor is that it takes disjoint unions to tensor products, that is, we are given an isomorphism of functors

$$\tilde{E}_\Sigma \otimes \tilde{E}_{\Sigma'} \to \tilde{E}_{\Sigma \sqcup \Sigma'},$$

which is associative and commutative in terms of the usual isomorphisms for the disjoint union and tensor product. There is a unique natural concept of tensor product here, because all the vector spaces are nuclear, and $\tilde{E}_\Sigma \otimes \tilde{E}_{\Sigma'} \cong \tilde{E}_{\Sigma \sqcup \Sigma'}$ is equivalent to $\tilde{E}_\Sigma \otimes \tilde{E}_{\Sigma'} \cong \tilde{E}_{\Sigma \sqcup \Sigma'}$. The functoriality means that we are assuming

$$Z_M \otimes Z_{M'} = Z_{M \sqcup M'},$$

for two cobordisms $M$ and $M'$.

The tensoring assumption implies that $E_\emptyset = \mathbb{C}$, where $\emptyset$ denotes the empty $(d-1)$-manifold. Thus for a closed $d$-manifold $M$ we have a partition function $Z_M \in \text{End}(E_\emptyset) = \mathbb{C}$. The whole structure of the theory is a way of expressing the sense in which the number $Z_M$ depends ‘locally’ on $M$.

In this discussion we have still committed an abuse of language: the ‘category’ $C_d^\Sigma$ is not really a category because it does not have identity maps. We could deal with this by agreeing that an isomorphism $\Sigma_0 \to \Sigma_1$ is a cobordism of zero length, but then these degenerate cobordisms are represented by operators which are not nuclear. The true replacement for the missing identity operators is our assumption that the maps $\tilde{E}_\Sigma \to \tilde{E}_\Sigma'$ are injective with dense image. To avoid the abuse of language we can say that a field theory is a functor $\Sigma \mapsto E_\Sigma$ from $(d-1)$-manifolds and isomorphisms to vector spaces, together with a transformation $Z_M : E_{\Sigma_0} \to E_{\Sigma_1}$ for each cobordism. Whichever line we take, we must assume that an isomorphism $f : \Sigma_0 \to \Sigma_1$ of germs of $d$-manifolds induces an isomorphism $f_* : E_{\Sigma_0} \to E_{\Sigma_1}$ which depends smoothly on $f$, in the sense that for any family $P \times \Sigma_0 \to \Sigma_1$ parameterized by a finite-dimensional manifold $P$ the induced map $P \times E_{\Sigma_0} \to E_{\Sigma_1}$ is smooth.

Let us explain briefly how to get from this functorial picture to the traditional language of local observables and vacuum expectation values. For a point $x$ of a $d$-manifold $M$, we define the vector space $O_x$ of observables at $x$ as follows. We consider the family of all closed discs $D$ smoothly embedded in $M$ which contain $x$ in the interior $\mathring{D}$. If $D' \subset \mathring{D}$ then $D \setminus D'$ is a cobordism $\partial D' \rightsquigarrow \partial D$ and gives us a trace-class map $E_{\partial D'} \to E_{\partial D}$. We therefore have an inverse system $\{E_{\partial D}\}$ indexed by the discs $D$, and we define $O_x$ as its inverse-limit.

Now suppose that $M$ is closed and that $x_1, \ldots, x_k$ are distinct points of $M$. Let $D_1, \ldots, D_k$ be disjoint discs in $M$ with $x_i \in D_i$. Then, $M' = M \setminus \bigcup D_i$ is a cobordism from $\bigcup \partial D_i$ to the empty $(d-1)$-manifold $\emptyset$, and defines $Z_{M'} : E_{\cup \partial D_i} \to E_\emptyset = \mathbb{C}$. Using the tensoring property, we can write this

$$Z_{M'} : \bigotimes E_{\partial D_i} \to \mathbb{C},$$
and then we can pass to the inverse-limits to get the expectation-value map

$$\bigotimes O_x \rightarrow \mathbb{C}.$$ 

We might prefer the language of ‘field operators’ to that of vacuum expectation values. If the space-time $M$ is a cobordism $\Sigma_0 \rightarrow \Sigma_1$, then for any $x$ in the interior of $M$—say $x \in D \subset M$—the cobordisms $M \setminus D$ from $\partial D \sqcup \Sigma_0$ to $\Sigma_1$ define maps

$$O_x \rightarrow \text{Hom}_{\text{nucl}}(E_{\Sigma_0}; E_{\Sigma_1}),$$

while if $x$ lies on a hypersurface $\Sigma$ an observable at $x$ defines a map $\hat{E}_{\Sigma} \rightarrow \hat{E}_{\Sigma}$, that is, it acts on $E_{\Sigma}$ as an unbounded operator. However, on a Lorentzian space-time $M$ we sometimes want to make the observables at all points $x \in M$ act on a single vector space, and to ask whether they commute when space-like separated. We shall postpone that discussion to Section 5.

One observable which we should mention is the energy-momentum tensor. If we think of a field theory as analogous to a group representation, then the energy-momentum tensor is the analogue of the induced representation of Lie algebras: for every cobordism $M : \Sigma_0 \rightarrow \Sigma_1$ it is the derivative of the operator $Z_M$ with respect to the metric of $M$. This makes it a distributional symmetric tensor-density $T_{ij}$ on $\hat{\mathcal{M}}$ with values in $\text{Hom}_{\text{nucl}}(E_{\Sigma_0}; E_{\Sigma_1})$. If we cover $M$ with small balls $D_i$, then by using a partition of unity we can write an infinitesimal change in the metric as the sum of contributions supported in the interiors of the $D_i$, and so the change in $Z_M$ is the sum of contributions coming from the spaces $E_{\partial D_i}$, and hence from a field operators placed at the centres of the balls $D_i$. However, developing this picture properly needs much more discussion, which we shall not embark on here; it probably requires the assumption that the theory is asymptotically conformal at short distances. The case of a two-dimensional conformal theory is treated fully in Section 9 of [12].

**Lorentzian manifolds**

There is a category $\mathcal{C}_d^{\text{Lor}}$ which at first sight looks more relevant to quantum field theory than $\mathcal{C}_d^C$. Its objects are compact Riemannian manifolds of dimension $(d - 1)$ and its morphisms are $d$-dimensional cobordisms equipped with real Lorentzian metrics. Fredenhagen and his coworkers (cf. [1]) have developed the theory of quantum fields in curved space-time using a version of this category. The category $\mathcal{C}_d^{\text{Lor}}$ lies ‘on the boundary’ of the category $\mathcal{C}_d^C$. In Section 5, we shall discuss the sense in which a representation of $\mathcal{C}_d^C$ has a ‘boundary value’ on $\mathcal{C}_d^{\text{Lor}}$, at least if it is unitary.

**Unitarity**

So far, we have not asked for an inner-product on the topological vector space $E_{\Sigma}$ associated with a $(d - 1)$-manifold $\Sigma$. Our main concern in this work is with unitary theories, even though not all interesting quantum field theories are unitary.

To define unitarity in our context, recall that, if $\Sigma^*$ denotes the manifold germ $\Sigma$ with its co-orientation reversed, then $\hat{E}_{\Sigma^*}$ is the dual topological vector space to $\hat{E}_{\Sigma}$. Furthermore, a cobordism $M : \Sigma_0 \rightarrow \Sigma_1$ can also be regarded as a cobordism from $\Sigma_1^*$ to $\Sigma_0^*$, and the two maps $E_{\Sigma_0} \rightarrow \hat{E}_{\Sigma_1}$ and $E_{\Sigma_1} \rightarrow \hat{E}_{\Sigma_0}$ are automatically algebraic transposes of each other. Thus $\Sigma \mapsto \Sigma^*$ is a contravariant functor.
In a unitary theory we shall not expect the vector space $E_{\Sigma}$ to have an inner product for every $(d-1)$-manifold $\Sigma$. A complex metric $g \in \text{Met}_{C}(\Sigma)$ has a complex conjugate $\bar{g}$. If we write $\bar{\Sigma}$ for $\Sigma$ with the metric $\bar{g}$ but with its co-orientation unchanged, then $\Sigma \mapsto \bar{\Sigma}$ is a covariant functor. (If our theory is defined on a category of oriented space-time manifolds, we must give $\bar{\Sigma}$ the opposite orientation in $U$.) It is natural to require that there is an antilinear involution

$$E_{\Sigma} \cong \bar{E}_{\Sigma}. \quad (5)$$

For a theory satisfying condition (5), the conjugate dual of the vector space $\bar{E}_{\Sigma}$ is $\hat{E}_{\bar{\Sigma}}$. We expect $\bar{E}_{\Sigma}$ to have an inner-product only when $\Sigma \cong \bar{\Sigma}^*$, that is, when the $d$-manifold germ $\Sigma \subset U$ admits a reflection with fixed-point set $\Sigma$, which reverses the co-orientation and changes the metric to its complex conjugate. Such a hypersurface-germ $\Sigma$ will be called time-symmetric. Its metric is real and Riemannian when restricted to the $(d-1)$-dimensional hypersurface $\Sigma$ itself.

We can now define a unitary theory as one which satisfies two conditions:

(i) the reality condition (5) and
(ii) reflection-positivity, in the sense that when we have a time-symmetric hypersurface $\Sigma \cong \bar{\Sigma}^*$ the hermitian duality between $\bar{E}_{\Sigma}$ and $\hat{E}_{\bar{\Sigma}}$ is positive-definite.

For a unitary theory, when we have a time-symmetric germ $\Sigma$, we can complete the pre-Hilbert space $\bar{E}_{\Sigma}$ to obtain a Hilbert space $E_{\Sigma}^{\text{Hilb}}$ with

$$\bar{E}_{\Sigma} \to E_{\Sigma}^{\text{Hilb}} \to \hat{E}_{\Sigma}. \quad$$

The theory on flat tori

The partition function of a theory on oriented flat Riemannian tori already gives us a lot of information about the theory. The moduli space of such tori is the double-coset space

$$O_d \backslash \text{GL}_d(\mathbb{R}) / \text{SL}_d(\mathbb{Z}) \cong Q(\mathbb{R}^d) / \text{SL}_d(\mathbb{Z}),$$

where $Q(\mathbb{R}^d) = O_d \backslash \text{GL}_d(\mathbb{R})$ is the space of positive-definite real $d \times d$ matrices. This space is an orbifold, so the partition function is best described as a smooth function $Z : Q(\mathbb{R}^d) \to \mathbb{C}$ which is invariant under $\text{SL}_d(\mathbb{Z})$. Our axioms imply that $Z$ extends to a holomorphic function

$$Q_{\mathbb{C}}(\mathbb{R}^d) \to \mathbb{C},$$

but they also imply very strong constraints beyond that. Notably, each choice of a surjection $\mathbb{Z}^d = \pi_1(M) \to \mathbb{Z}$ gives us a way of writing the torus $M$ as a cobordism $M : \Sigma \rightsquigarrow \Sigma$ from a $(d-1)$-dimensional torus $\Sigma$ to itself, and then we have $Z(M) = \text{trace}(Z_M)$, where $Z_M : E_{\Sigma} \to E_{\Sigma}$ is a nuclear operator in the vector space $E_{\Sigma}$, which is graded by the characters $\chi$ of the translation-group $T_{\Sigma}$ of
Appendix to Section 3: The duality $(\hat{E}_\Sigma)^* \cong \hat{E}_{\Sigma^*}$

To keep things as general as possible, we suppose that $\Sigma \mapsto E_\Sigma$ is a functor from the $d$-dimensional cobordism category to a category of metrizable topological vector spaces and nuclear maps. We suppose also that the category of vector spaces is equipped with a tensor product functor, which is coherently associative and commutative, and that we are given natural isomorphisms $E_{\Sigma_1} \otimes E_{\Sigma_2} \to E_{\Sigma_1 \sqcup \Sigma_2}$. (For example, we could work with the category of Hilbert spaces with the natural Hilbert space tensor product.)

Composable cobordisms $\Sigma_1 \rightsquigarrow \Sigma_2 \rightsquigarrow \Sigma_3$ give us maps

$$E_{\Sigma_1} \to E_{\Sigma_2} \to E_{\Sigma_3}. \quad (6)$$

By reinterpreting $\Sigma_1 \rightsquigarrow \Sigma_2$ as a cobordism $\Sigma_1 \sqcup \Sigma_2^* \rightsquigarrow \emptyset$ we get a map $E_{\Sigma_1} \otimes E_{\Sigma_2^*} \to \mathbb{C}$, and hence $E_{\Sigma_1} \to (E_{\Sigma_2^*})^*$. Similarly, we can reinterpret $\Sigma_2 \rightsquigarrow \Sigma_3$ as $\emptyset \rightsquigarrow \Sigma_2^* \sqcup \Sigma_3$, which gives $(E_{\Sigma_2^*})^* \to E_{\Sigma_1}$. It is easy to see that the composite $E_{\Sigma_1} \to (E_{\Sigma_2^*})^* \to E_{\Sigma_3}$ coincides with $E_{\Sigma_1} \to E_{\Sigma_2} \to E_{\Sigma_3}$. Yet again, performing the reinterpretations in the reverse order, we get maps

$$(E_{\Sigma_1^*})^* \to E_{\Sigma_2} \to (E_{\Sigma_3^*})^*$$

whose composite is the transpose of the map induced by the composite cobordism $\Sigma_2^* \rightsquigarrow \Sigma_1^*$. Now suppose that we have an infinite sequence of cobordisms

$$\ldots \rightsquigarrow \Sigma_{i+1} \rightsquigarrow \Sigma_i \rightsquigarrow \Sigma_{i-1} \rightsquigarrow \ldots, \quad (7)$$

indexed by $i \geq 0$, which form the downstream tail of a manifold-germ $\Sigma$, that is, the sequence which we used above to define the space $\hat{E}_\Sigma = \lim_{\rightarrow} E_{\Sigma_i}$. Let us perform the two manipulations that we performed on (6) alternately on the sequence (7), thereby obtaining a sequence whose even terms are $E_{\Sigma_{2i}}$ and whose odd terms are $(E_{\Sigma_{2i+1}})^*$. The inverse-limit of the whole sequence is the same as that of any cofinal subsequence. Considering the cofinal subsequence of even terms shows that the inverse-limit is $E_\Sigma$ but the inverse-limit of the cofinal sequence of odd terms is

$$\lim_{\rightarrow} (E_{\Sigma_{2i+1}})^* = (\lim_{\rightarrow} E_{\Sigma_{2i+1}})^*.$$

This shows that $\hat{E}_\Sigma \cong (\hat{E}_{\Sigma^*})^*$. However, because $\hat{E}_\Sigma$ is automatically a nuclear Fréchet space, we can dualize again and conclude that $(\hat{E}_\Sigma)^* \cong \hat{E}_{\Sigma^*}$ also.
4. Some analogies from representation theory

The relation between representations of the category $\mathcal{C}_d^\mathbb{C}$ and of the Lorentzian category $\mathcal{C}_d^\mathbb{L}$ which lies ‘on its boundary’ follows a pattern familiar in the representation theory of many Lie groups that have a special class of unitary representations characterized as the boundary values of holomorphic representations of a complex semigroup by contraction operators. The essential features can all be seen in the simplest example.

The group $G = \text{PSL}_2(\mathbb{R})$ is the group of Möbius transformations of the Riemann sphere $\Sigma = \mathbb{C} \cup \infty$ which maps the open upper half-plane $\mathbb{U}$ to itself. It lies on the boundary of the complex sub-semigroup of $\mathbb{G}_\mathbb{C} = \text{PSL}_2(\mathbb{C})$ consisting of Möbius transformations which map the closure of $\mathbb{U}$ into its own interior. It is natural, however, to consider a slightly larger semigroup $\mathbb{G}_\mathbb{C}^\subset$ by including the degenerate Möbius transformations which collapse $\mathbb{U}$ to a single point in $\mathbb{U}$—these correspond to complex $2 \times 2$ matrices of rank one. The topological boundary of $\mathbb{G}_\mathbb{C}^\subset$ consists of the Möbius transformations which take $\mathbb{U}$ to a disc or point in the upper half-plane that touches the real axis, and the Shilov boundary consists of the group $G$ of real Möbius transformations—an open solid torus—compactified by its two-torus boundary, which is the hyperboloid $\text{det}(A) = 0$ in $\mathbb{P}^3_\mathbb{R}$ consisting of the degenerate real Möbius transformations. (Thus, the complete Shilov boundary is the part of $\mathbb{P}^3_\mathbb{R}$ where $\text{det}(A) \geq 0$.)

The irreducible unitary representations of the group $G = \text{PSL}_2(\mathbb{R})$ are essentially of two kinds, the principal series and the discrete series (we shall ignore the ‘supplementary series’, which is of measure zero in the space of unitary representations). The best-known principal series representation is the action of $G$ on the Hilbert space of $1/2$-densities on the circle $\mathbb{P}^1_\mathbb{R}$ which is the boundary of $\mathbb{U}$—the general member of the series is the action on densities of complex degree $s$ with $\text{Re}(s) = 1/2$. The best-known discrete-series representation is the action of $G$ on the square-summable holomorphic one-forms on $\mathbb{U}$, with the natural norm

$$\| \alpha \|^2 = i \int_{\mathbb{U}} \alpha \wedge \bar{\alpha}$$

—more generally, for each positive integer $p$, we have the action on holomorphic $p$-forms $\alpha = f(z)(dz)^{\otimes p}$, when one must divide $\alpha \wedge \bar{\alpha}$ by the $(p - 1)^{\text{st}}$ power of the $G$-invariant area form. The discrete-series representations obviously extend to bounded holomorphic representations of the semigroup $\mathbb{G}_\mathbb{C}^\subset$ by contraction operators. They are singled out by this ‘positive energy’ property: the principal series representations cannot extend to $\mathbb{G}_\mathbb{C}^\subset$, because when $|a| < 1$ the element $w \mapsto aw$ (here $w = (z-i)/(z+i)$ is the coordinate in the unit-disc model $|w| < 1$ of $\mathbb{U}$) of the semigroup $\mathbb{G}_\mathbb{C}^\subset$ would be represented by an operator whose eigenvalues are $a^n$ for all $n \in \mathbb{Z}$. However, let us notice that, though the discrete-series representations are unitary on the boundary group $G = \text{PSL}_2(\mathbb{R})$, the degenerate elements of $\mathbb{G}_\mathbb{C}^\subset$, which collapse $\mathbb{U}$ to a point $p \in \mathbb{U}$, are represented by bounded operators of rank 1. Therefore, these unitary representations of $\text{PSL}_2(\mathbb{R})$ do not extend unitarily to the whole Shilov boundary: the degenerate elements correspond to unbounded rank 1 operators $\xi \mapsto \langle \xi, \eta \rangle \eta$, where $\eta$ and $\xi$ are ‘non-normalizable elements’ of the Hilbert space—that is, they belong to an appropriate completion of it.

The group $G$ is a subgroup of the group $\text{Diff}^+(S^1)$ of orientation-preserving diffeomorphisms of the circle. This infinite-dimensional Lie group does not possess a complexification $[10]$, although its Lie algebra, the space of smooth vector fields on the circle, can of course be complexified. The
beginning of the present work was the observation, made in the 1980s quite independently by the two authors and also by Yu Neretin [9, 11], that there is an infinite-dimensional complex semigroup \( A \) which has exactly the same relation to \( \text{Diff}^+(S^1) \) as \( G^\prec_C \) has to \( G = \text{PSL}_2(\mathbb{R}) \). Its elements are complex annuli with parametrized boundary circles: one can think of them as ‘exponentiations’ of outward-pointing complex vector fields defined on a circle in the complex plane. The annuli form a complex semigroup when concatenated as cobordisms, and the lowest-weight or ‘positive-energy’ representations of \( \text{Diff}^+(S^1) \)—and of loop groups—which arise in two-dimensional conformal field theory are precisely those which are boundary values of holomorphic representations of the semigroup \( A \) by trace-class operators.

The discussion of \( \text{PSL}_2(\mathbb{R}) \) generalizes to the symplectic group \( G = \text{Sp}(V) \cong \text{Sp}_{2n}(\mathbb{R}) \) of a real symplectic vector space \( V \) of dimension \( 2n \). The role of the upper half-plane \( \mathbb{U} \) is played by the Siegel ‘generalized upper half-plane’—the domain \( \mathbb{U}(V) \) of positive Lagrangian subspaces of the complexification \( V_\mathbb{C} \) described in Section 1. The group \( G \) lies on the boundary of a semigroup \( G^\prec_C \) which is the Siegel domain \( \mathbb{U} (\hat{V} \oplus V) \), where \( \hat{V} \) denotes \( V \) with sign of its symplectic form reversed. A generic element of this domain is the graph of a complex symplectic transformation of \( V_\mathbb{C} \) which maps the closure of \( \mathbb{U}(V) \) into its own interior, but, just as was the case with \( \text{PSL}_2(\mathbb{C}) \), there are degenerate elements which map \( \mathbb{U}(V) \) non-injectively into themselves. The complex semigroup \( G^\prec_C \) has been carefully studied by Roger Howe [6], who called it the ‘oscillator semigroup’.

The Shilov boundary of \( G^\prec_C \) is the Grassmannian of real Lagrangian subspaces of \( \hat{V} \oplus V \): generically, these are the graphs of elements of the real group \( G = \text{Sp}(V) \), but this group is compactified by the addition of Lagrangian subspaces which intersect the axes of \( \hat{V} \oplus V \) nontrivially, and thus correspond to Lagrangian correspondences from \( V \) to \( V \), which are not actual maps \( V \to V \). Once again, whereas \( \text{Sp}_C^\prec(V_\mathbb{C}) \) is a genuine semigroup, the composition-law of the real group \( \text{Sp}(V) \) does not extend to the compactification.

The group \( G = \text{Sp}(V) \) has a discrete series of unitary representations generalizing those of \( \text{PSL}_2(\mathbb{R}) \). The most important is the metaplectic representation—actually a representation of a double covering \( \hat{G} \) of \( \text{Sp}(V) \)—which is the action on the quantization \( \mathcal{H}_V \) of the symplectic space \( V \). The Hilbert space \( \mathcal{H}_V \) is characterized by the property that it contains a copy of the ray \( \langle \wedge^n(W) \rangle^{(1/2)} \) for each point \( W \) of the domain \( \mathbb{U}(V) \)—the square-root of the natural hermitian holomorphic line bundle \( \{ \wedge^n(W) \} \) on \( \mathbb{U}(V) \) is canonical up to multiplication by \( \pm 1 \), and is holomorphically embedded in \( \mathcal{H}_V \). It is acted on by \( \hat{G} \) rather than \( G \).

The action of \( \hat{G} \) on \( \mathcal{H}_V \) is the boundary-value of a holomorphic projective representation of the oscillator semigroup \( G^\prec_C \). For \( G^\prec_C \) is just the domain \( \mathbb{U} (\hat{V} \oplus V) \), each point of which defines a ray in

\[
\mathcal{H}_{\hat{V} \oplus V} \cong \mathcal{H}_V^* \otimes \mathcal{H}_V \cong \text{End}_{HS}(\mathcal{H}_V),
\]

where \( \text{End}_{HS} \) denotes the Hilbert–Schmidt endomorphisms. (A more careful discussion shows that \( G^\prec_C \) is represented by operators of trace class.)

When \( n = 1 \), the group \( \text{Sp}(V) \) is \( \text{SL}_2(\mathbb{R}) \), a double covering of the group \( \text{PSL}_2(\mathbb{R}) \) of Möbius transformations we considered before. To relate the cases of \( \text{PSL}_2(\mathbb{R}) \) and \( \text{Sp}(V) \), recall that \( \text{PSL}_2(\mathbb{C}) \) is an open subspace of the complex projective space \( \mathbb{P}^3_\mathbb{C} \) formed from the vector space of \( 2 \times 2 \) matrices: in fact, it is the complement of the quadric \( Q_2^\mathbb{C} \cong \mathbb{P}^1_\mathbb{C} \times \mathbb{P}^1_\mathbb{C} \) defined by the vanishing of the determinant, that is, by the matrices of rank 1. The double covering group \( \text{SL}_2(\mathbb{C}) \) sits inside the Grassmannian of complex Lagrangian subspaces of \( \mathbb{C}^4 \), which is a quadric three-fold \( Q_3^\mathbb{C} \) in \( \mathbb{P}^4_\mathbb{C} \); it is a non-singular hyperplane section (corresponding to the Lagrangian condition) of the Klein quadric formed by all the lines in \( \mathbb{P}^3(\mathbb{C}) \). The quadric \( Q_3^\mathbb{C} \) is the branched double-covering of the projective
space $\mathbb{P}^3_{\mathbb{C}}$ of $2 \times 2$ matrices, branched along the quadric $Q_2^1$ of rank 1 matrices. The contractible semigroup $\text{SL}_2^<(\mathbb{C})$ is the open subset of the Lagrangian Grassmannian of $\mathbb{C}^4$ consisting of the positive Lagrangian subspaces, and it is a double covering of $\text{PSL}_2^<(\mathbb{C})$.

5. Unitarity and global hyperbolicity

In the previous section, we saw how a holomorphic representation of a complex semigroup by contraction operators on a Hilbert space can give rise—on passing to the boundary—to a unitary representation of a group which is a dense open subset of the Shilov boundary of the semigroup. The remaining points of the Shilov boundary are not represented by unitary operators; the representation extends to them only in some ‘weak’ sense. We now come to the analogue of this phenomenon in quantum field theory, where the Lorentzian cobordism category $C^\text{Lor}_{d,\text{diff}}$ lies on the boundary of $C^\text{C}_d$, and the role of the open dense subgroup of the Shilov boundary is played by the subcategory of \textit{globally hyperbolic} cobordisms which we shall define below. We should mention, however, that although the category of globally hyperbolic cobordisms is very natural, the category $C^\text{Lor}_{d,\text{diff}}$ may be smaller than the optimal category we could put on the boundary of $C^\text{C}_d$. For example, the Lorentzian cobordisms could possibly be allowed to contain ‘black holes’ surrounded by horizons, rather analogous to the ‘cobordisms-with-boundaries’ used to describe two-dimensional theories with both open and closed strings. We shall not pursue such speculations here.

When we have a theory defined on $C^\text{C}_d$ let us first consider how to extend the assignment $\Sigma \mapsto E_\Sigma$ to a Lorentzian germ $\Sigma \subset U$, with $\Sigma$ co-oriented in $U$. We can identify $U$ with $\Sigma \times (-\varepsilon, \varepsilon)$ by exponentiating the geodesic curves emanating perpendicularly from $\Sigma$. The metric then takes the form $h_t - dt^2$, where $t \mapsto h_t$ is a smooth map from $(-\varepsilon, \varepsilon)$ to the manifold of Riemannian metrics on $\Sigma$. If the germ is time-symmetric, then we can define $E_\Sigma$ by replacing the Lorentzian metric by the ‘Wick rotated’ Riemannian metric $h_t + dt^2$, which makes sense because if $h_t = h_{-t}$ then $h_t$ is a function of $t^2$, so that $h_t$ is defined and real. However, this does not help for a general hypersurface, and in any case seems rather arbitrary: we shall return to this point in Remark 5.3.

It is less easy to assign an operator $Z_M : E_{\Sigma_0} \to E_{\Sigma_1}$ to a Lorentzian cobordism $M : \Sigma_0 \rightleftharpoons \Sigma_1$. Even if $M$ is a cylinder topologically, it can be complicated in its ‘causal’ structure. Consider, for example, a two-dimensional cylindrical space-time. We saw in Section 2 that, up to a conformal multiplier, a complex metric on a surface is a pair of complex structures with opposite orientations. At the Shilov boundary, the complex structures degenerate to the foliations by the left- and right-moving light-lines of a Lorentzian surface. If each light-line that sets out from the incoming boundary circle of the cylinder eventually reaches the outgoing boundary circle, then each family of light-lines gives us a diffeomorphism from the incoming to the outgoing boundary. In fact (cf. [12] p.8 and p.16), the isomorphism classes of Lorentzian cylinders of this kind are determined up to conformal equivalence by the pair of diffeomorphisms together with a positive integer which counts the number of times that the left- and right-moving lines emanating from a given point of the incoming circle cross before hitting the outgoing circle. This agrees with the well-known fact that the Hilbert space associated to a circle in two-dimensional conformal field theory comes with a projective unitary representation of the group $\text{Diff}^+(S^1) \times \text{Diff}^+(S^1)$.

However the light-lines from the incoming circle can behave in a more complicated way. For example, one set of light-lines may spiral closer and closer to a closed limit-cycle of the foliation, a light-line which is a circle parallel to the incoming boundary circle of the annulus. That set of lines will then never reach the outgoing circle. One might think of this phenomenon as akin to a black hole in the space-time, although, unlike a black hole, the Lorentzian metric here has no singularity. The
In fact, the space Emb \( \Sigma \to \Sigma \) is isomorphic to the space of all \( C^1 \) maps \( \Sigma \to \Sigma \) that are fibered over \( \Sigma \) with fibers of the form \( \Sigma \times \Sigma \). This is usually regarded as an ‘annulus of infinite length’, and it acts on an irreducible positive-energy representation of \( \text{Diff}^+(S^1) \) by a projection operator of rank one, like the action of a degenerate complex Möbius transformation in a discrete-series representation of \( \text{PSL}_2(\mathbb{R}) \).

In works on general relativity, a Lorentzian cobordism \( M: \Sigma_0 \to \Sigma_1 \) between Riemannian manifolds is called \textit{globally hyperbolic} if every maximally extended time-like geodesic in \( M \) travels from \( \Sigma_0 \) to \( \Sigma_1 \). Such an \( M \) must be diffeomorphic to \( \Sigma_0 \times [0, 1] \). It is only for globally hyperbolic manifolds that, for example, the Cauchy problem for the wave-equation on \( M \) is solvable.

Of course here we are only considering \textit{compact} cobordisms, which are not the usual focus in relativity theory. In the compact situation, we can take the definition of global hyperbolicity to be the existence of a smooth time-function \( t: M \to [0, 1] \) whose gradient is everywhere in the positive light-cone, and which is, therefore, a fibration with Riemannian fibres. From \( t \) we obtain a diffeomorphism \( M \to \Sigma_0 \times [0, 1] \) by following the orthogonal trajectories to the time-slices.

The existence of a time-function on a compact Lorentzian cobordism is clearly an open condition, and so the globally hyperbolic cobordisms form an open subcategory \( \text{C}^{d}_{\text{gh}} \) of \( \text{C}^{d}_{\text{Lor}} \) which should play the role of the real Lie group to which the holomorphic contraction representations of Section 4 can be extended (however, the result (5.2) we prove below is unfortunately weaker).

For a globally hyperbolic cobordism equipped with a time-function, the metric, in terms of the diffeomorphism \( M \to \Sigma_0 \times [0, 1] \), takes the form \( h_t + c^2 dt^2 \) for some function \( c: \Sigma_0 \times [0, 1] \to i \mathbb{R} \). A small deformation \( \delta c \) of \( c \) into the right half-plane changes the Lorentzian metric into an allowable complex metric, and we could hope to define \( Z_M \) in the Lorentzian case as the limit of the operators associated to such deformations. That, however, encounters the problem that the deformed metric depends not only on the choice of the deformation \( \delta c \), but, more importantly, on the choice of the time-function, which should be irrelevant to the identity operator—quite unlike what happens with Riemannian cobordisms. (A possible analogy is the process of making a based loop-space into a topological group by collapsing paths which retrace their steps.)

The passage from \( \text{C}^d \) to \( \text{C}^d_{\text{Lor}} \) is already interesting when \( d = 1 \), that is, for quantum mechanics rather than quantum field theory—the case when the Euclidean path-integral can be treated by traditional measure-theory. It is worthwhile to spell out the argument in this case, before passing to higher dimensions.

We began this work with the relation of positive energy to one-parameter contraction semigroups. Our first task now is to understand why a holomorphic representation of the category \( \text{C}^d_1 \) is just such a one-parameter semigroup, where the parameter runs through the open half-plane \( \mathbb{C}_+ = \{ z \in \mathbb{C} : \text{Re}(z) > 0 \} \). Whereas a Riemannian structure on a closed interval is completely determined by its length, the allowable complex metrics on the interval have an infinite-dimensional moduli-space.

Any complex metric on \( I = [0, 1] \) can be pulled back from the holomorphic quadratic differential \( dz^2 \) on \( \mathbb{C} \) by means of a smooth embedding \( f: I \to \mathbb{C} \) such that \( f(0) = 0 \) and \( \text{Re} f'(t) > 0 \) for all \( t \in I \). In fact, the space \( \text{Emb}(I; \mathbb{C}) \) of such embeddings is isomorphic to \( \text{Met}_\mathbb{C}(I) \) as a complex manifold. If \( f'(t) = 1 \) when \( t \) is sufficiently close to the ends of the interval \( I \), then the pulled-back metric defines a morphism \( I_f: P \to P \) in the category \( \text{C}^d_1 \), where \( P \) denotes the object defined by the germ of the standard metric on the line \( \mathbb{R} \) at the origin.
The crucial observation is that the operator \( Z_f : E_P \to E_P \) defined by \( I_f \) depends only on the point \( f(1) \in \mathbb{C}_+ \). It is as if \( Z_f \) were the ‘contour integral’ of a holomorphic differential on \( \mathbb{C} \) along the path \( f \). The argument is as follows. First, \( Z_f \) does not change if \( f \) is replaced by \( \tilde{f} = f \circ \phi \) where \( \phi \) is any diffeomorphism \( I \to I \), which is the identity near the ends of the interval. This means that \( Z_f \) does not change if \( f \) moves along a curve in \( \text{Emb}(I; \mathbb{C}) \) whose tangent vector at each point is the action of an element of the Lie algebra \( \text{Vect}(I) \) of compactly supported vector fields on the interior of \( I \). But then—because \( Z_f \) depends holomorphically on \( f \)—it does not change if each tangent vector is the action of an element of the complexified Lie algebra \( \text{Vect}_C(I) \). Finally, if \( f, \tilde{f} \in \text{Emb}(I; \mathbb{C}) \) define two morphisms \( P \to P \) and have \( f(1) = \tilde{f}(1) \), the tangent vectors to the obvious linear path from \( f \) to \( \tilde{f} \) are given by the action of elements of \( \text{Vect}_C(I) \).

We can therefore write \( Z_f = u(z) \), where \( z = f(1) \). Obviously, we have \( u(z_1)u(z_2) = u(z_1 + z_2) \) for any \( z_1, z_2 \in \mathbb{C}_+ \). Furthermore, because the object \( P \) of \( \mathcal{C}_1^{\text{gh}} \) is time-symmetric, the vector space \( \hat{E}_P \) is a pre-Hilbert space, and the unitarity condition tells us that \( u(z) \) is the hermitian transpose of \( u(z) \).

The desired unitary semigroup \( \{u(iT)\}_{T \in \mathbb{R}} \), which will act on the triple \( \hat{E}_P \to \hat{E}_P \to \hat{E}_P \), can now be defined as follows. As explained in Section 3, any vector \( \xi \in \hat{E}_P \) can be written \( \xi = u(\varepsilon)\eta \) for some \( \varepsilon > 0 \) and some \( \eta \in \hat{E}_P \). We define \( u(iT)\xi = u(\varepsilon + iT)\eta \), which is plainly independent of \( \varepsilon \). Finally, \( u(iT) \) is unitary because

\[
u(-iT)u(iT)\xi = u(-iT)u(\varepsilon + iT)\eta
\]

\[
= u(-iT)u(\varepsilon/2)u(\varepsilon/2 + iT)\eta
\]

\[
= u(\varepsilon/2 - iT)u(\varepsilon/2 + iT)\eta
\]

\[
= u(\varepsilon)\eta = \xi
\]

To pass from \( d = 1 \) to higher-dimensional cobordisms, we observe that the essential step in our argument was the first case of the following

**Principle 5.1** If a \( d \)-dimensional cobordism \( M \) is a real submanifold of a complex \( d \)-manifold \( M_C \), and \( M \) has an allowable complex metric induced from a holomorphic symmetric form \( g \) on the tangent bundle \( TM_C \), then the linear map \( Z_M \) does not change when \( M \) is moved around smoothly inside \( M_C \) (leaving its ends fixed), providing the restriction of \( g \) to \( M \) remains an allowable complex metric.

As in the \( d = 1 \) case, this principle holds because any infinitesimal movement of \( M \) inside \( M_C \) is given by a complex vector field on \( M \), while \( Z_M \) depends holomorphically on \( M \) and, being invariant under the action of \( \text{Diff}(M \text{ rel } \partial M) \), does not change when \( M \) moves in a direction given by the action of a complexified tangent vector to this group.

Unfortunately, to use the principle we need the cobordism \( M \) to be embedded in a complexification \( M_C \), and the only natural way to ensure this is to pass from the smooth Lorentzian category \( \mathcal{C}_d^{\text{Lor}} \) to the corresponding real-analytic cobordism category \( \mathcal{C}_d^{\text{Lor}, \omega} \), where both the manifolds and their metrics are assumed real-analytic. Inside this category, there is the subcategory \( \mathcal{C}_d^{\text{gh}, \omega} \) of globally hyperbolic cobordisms: we shall also assume that the time-function \( \tau : M \to [0, 1] \) is real-analytic, although that could be avoided, because any smooth function can be approximated real-analytically.

There are two ways of thinking about restricting to real-analytic cobordisms. One might think that the smooth cobordism category is the natural object and try to eliminate the analyticity hypothesis. However, one could also think that the natural allowable space-times really do come surrounded
by a thin holomorphic thickening, within which the choice of a smooth totally-real representative is essentially arbitrary. In any case, we can prove the following theorem.

**Theorem 5.2** A unitary quantum field theory as defined in Section 3 on the category $C^C_d$ induces a functor from $C^d_{gh, \omega}$ to topological vector spaces. The functor takes time-symmetric objects to Hilbert spaces and takes cobordisms between them to unitary operators.

To be quite precise, the theorem asserts that if $\Sigma$ is a time-symmetric $(d-1)$-manifold germ then there is a Hilbert space $E^{H, \text{fib}}_{\Sigma}$ with

$$\hat{E}_\Sigma \subset E^{H, \text{fib}}_{\Sigma} \subset \hat{E}_\Sigma,$$

and a real-analytic globally hyperbolic cobordism $\Sigma_0 \sim \Sigma_1$ between time-symmetric hypersurfaces induces a unitary isomorphism $E^{H, \text{fib}}_{\Sigma_0} \to E^{H, \text{fib}}_{\Sigma_1}$ which also maps $\hat{E}_{\Sigma_0}$ to $\hat{E}_{\Sigma_1}$ and $\hat{E}_{\Sigma_0}$ to $\hat{E}_{\Sigma_1}$.

**Proof of 5.2** Given a real-analytic globally hyperbolic cobordism $M : \Sigma_0 \sim \Sigma_1$, we choose a time function $t : M \to [0, 1]$ whose level surfaces foliate $M$ by Riemannian manifolds, and, following the orthogonal trajectories to the foliation, we identify $M$ with $\Sigma_0 \times [0, 1]$ as before.

Using the real-analyticity assumptions, we can find a complexification $M_C$ of $M$ to which both $t$ and $g$ can be extended holomorphically, and we can assume that $\tau = it : M_C \to U \subset \mathbb{C}$ is a holomorphic fibre bundle over a neighbourhood $U$ of the interval $i[0, 1]$. Furthermore, the isomorphism $\Sigma_0 \times [0, 1] \to M$ extends to a holomorphic trivialization of the bundle $M_C \to U$. For any smooth curve $f : [0, 1] \to U$ such that $f(0) = 0$ and $\text{Re} f'(s) > 0$ for $s \in [0, 1]$, this gives us a totally real submanifold $M_f$ of $M_C$ sitting over the curve. We can use the morphism associated with the cobordism $M_f$ in exactly the way we used $Z_f$ in discussing the one-dimensional case, to obtain a unitary operator $Z_M$ associated with the Lorentzian cobordism.

It is important that $Z_M$ does not depend on the choice of the time-function $t$ defining the foliation. For two choices of $t$ are linearly homotopic, and changing from one to the other amounts to deforming the totally real embedding $\Sigma_0 \times [0, 1] \to M_C$ by a real-analytic diffeomorphism of $\Sigma_0 \times [0, 1]$. $\square$

**Remark 5.3** We can apply the Principle 5.1 to understand better how a theory defined on $C^C_d$ assigns a vector space $E_\Sigma$ to a Lorentzian germ $\Sigma \subset U$.

If the Lorentzian metric on $U$ is real-analytic then the complex theory gives us a holomorphic bundle $\{\hat{E}_f\}$ on the space $\mathcal{J}$ of germs of embeddings $f : (-\varepsilon, \varepsilon) \to \mathbb{C}$ such that $f(0) = 0$ and $\text{Re} f'(t) > 0$ for all $t$. In particular, for $\lambda \in \mathbb{C}_+$ we have the radial paths $f_\lambda, f_{\lambda'} \in \mathcal{J}$ for which $f_\lambda(t) = \lambda t$. But recall that $\hat{E}_f$ is the inverse-limit of a sequence of spaces associated with the germs of $f$ at the points $f(t_k)$, for any sequence $\{t_k \downarrow 0\}$.

Now consider two neighbouring rays $f_\lambda, f_{\lambda'}$ with $|\lambda| = |\lambda'|$, and choose a sequence $\{t'_k \downarrow 0\}$ which interleaves $\{t_k\}$, that is, $t_k > t'_k > t_{k+1}$. We can choose a path $f \in \mathcal{J}$ which lies in the sector bounded by the rays $f_\lambda$ and $f_{\lambda'}$ and coincides with them alternately in the neighbourhoods of the points $\lambda t_k$ and $\lambda' t'_k$. This $f$ gives us a family of cobordisms from the germ at $\lambda t_k$ to the germ at $\lambda t_k$ and from the germ at $\lambda t_{k+1}$ to the germ at $\lambda' t'_k$. Putting these together, we obtain inverse canonical isomorphisms between $\hat{E}_{f_\lambda}$ and $\hat{E}_{f_{\lambda'}}$. The coherence of these isomorphisms when we consider three nearby rays also follows from the Principle 5.1.

By this means, we see that we could have chosen any smooth path $f$ to define $\hat{E}_\Sigma$. However the family $\hat{E}_f$ has the property that $\hat{E}_f$ is the complex-conjugate space to $E_f$, so that reversing the complex
time-direction conjugates the identification of $\hat{E}_\Sigma$ with the Euclidean choice $\hat{E}_t$. If the Lorentzian germ $\Sigma \subset U$ is time-symmetric—but not otherwise—the arguments we have already used will give us a hermitian inner product on $\hat{E}_\Sigma$.

**Field operators**

Finally, we come to the Wick rotation of field operators, although our account will be sketchy. The first step is to understand how the vector space $O_x$ of observables at a point $x$ of a space-time $M$ behaves as the metric of $M$ passes from complex to Lorentzian. We shall continue to assume that $M$ and its Lorentzian metric are real-analytic.

In Section 3 we associated a space $O_x$ with a germ at $x$ of a complex metric on a manifold containing $x$: it is the fibre of a bundle on the space $\text{Met}_C(\hat{x})$ of such germs. If we embed a Lorentzian $M$ in a complexification $M_C$, there will be a holomorphic exponential map from a neighbourhood of 0 in the complexified tangent space $T^C_x = T_xM \otimes \mathbb{C}$ to $M_C$. Inside $T^C_x$, we can consider the $d$-dimensional real vector subspaces $V$ on which the metric induced from the complex bilinear form of $T^C_x$ is allowable. We saw in Proposition 2.6 that these $V$ form a contractible open subset $\mathcal{U}$ of the real Grassmannian $\text{Gr}_d(T^C_x)$. Exponentiating $V$ will give us a germ of a $d$-manifold with a complex metric, and hence a map $\mathcal{U} \to \text{Met}_C(\hat{x})$. Pulling back the bundle of observables by this map gives us a bundle on $\mathcal{U}$, which, using the Principle 5.1 as we did in Remark 5.3, we see to be trivial. Identifying its fibres gives us our definition of $O_x$ for Lorentzian $M$.

We need no new ideas to see that for any Lorentzian cobordism $M : \Sigma_0 \rightarrow \Sigma_1$ and any $x \in \hat{M}$ an element $\psi \in O_x$ acts as an operator $E_{\Sigma_0} \to E_{\Sigma_1}$. Furthermore, if $x$ lies on a time-slice $\Sigma$ we get an operator $\psi \in \text{Hom}(\hat{E}_\Sigma; \hat{E}_\Sigma)$, that is, an unbounded operator in $E_{\Sigma}$, simply by considering the cobordisms corresponding to a sequence of successively thinner collars of $\Sigma$. Indeed, the same argument shows that if $x_1, \ldots, x_k$ are distinct points on $\Sigma$, we have a map

$$O_{x_1} \otimes \ldots \otimes O_{x_k} \rightarrow \text{Hom}(\hat{E}_\Sigma; \hat{E}_\Sigma)$$

which does not depend on choosing an ordering of the points.

In the introduction, we mentioned the Wightman axiom that field operators at space-like separated points must commute. We can now see how this follows from our framework, at least in a globally hyperbolic space-time. This is because the spaces $\hat{E}_{\Sigma_t} \subset E^{\text{Hilb}}_{\Sigma_t} \subset E_{\Sigma}$ for all times $t_0 \leq t \leq t_1$ can be identified with those at time $t_0$ by the unitary propagation $Z_{t,t'}$ from time $t$ to a later time $t'$ to get a single rigged Hilbert space $\hat{E} \subset E^{\text{Hilb}} \subset \hat{E}$, and we can define an unbounded operator

$$\tilde{\psi} = Z_{t_0,t}^{-1} \circ \psi \circ Z_{t_0,t} : \hat{E} \rightarrow \hat{E}$$

for any $\psi \in O_x$ with $x \in \Sigma_t$. Furthermore, if we change the choice of time-function on the cobordism, so that $x$ lies on a different time-slice, then $\tilde{\psi}$ will not change.

The fact that two observables $\psi, \psi'$ situated at space-like separated points $x, x'$ give rise to operators $\tilde{\psi}, \tilde{\psi}'$ which are composable and commute is now clear. For if $x$ and $x'$ are space-like separated we can choose a single time-slice $\Sigma_t$ that contains them both, and we see that the composed operator, in either order, is $Z_{t_0,t}^{-1} \circ (\psi \otimes \psi') \circ Z_{t_0,t}$.
The domain of holomorphicity of the vacuum expectation values

We end with a conjecture about a question arising in the traditional treatment of field theories defined in the standard Minkowski space \( \mathbb{M} = \mathbb{R}^{d-1,1} \). There, the Wightman axioms imply that the vacuum expectation values, initially defined as distributions or other generalized functions on the \( k \)-fold products \( \mathbb{M} \times \ldots \times \mathbb{M} \), are boundary values of holomorphic functions defined in an open domain \( \mathcal{U}_k \) in the complexified space \( \mathbb{M}_\mathbb{C} \times \ldots \times \mathbb{M}_\mathbb{C} \). The definition of \( \mathcal{U}_k \), known as the ‘permuted extended tube’, was given in Section 2. Recall that \( \mathcal{U}_2 \) consists of all pairs of points \( x, y \) such that \( ||x-y||^2 \) is not real and \( \leq 0 \).

If \( k > 2 \), however, \( \mathcal{U}_k \) is known not to be holomorphically convex, so it cannot be the largest complex manifold to which the expectation values can be analytically continued. It is an old problem to describe this largest manifold \( \mathcal{V}_k \), or even the holomorphic envelope of \( \mathcal{U}_k \).

The ideas of this paper suggest a candidate for \( \mathcal{V}_k \). It sits over the open subset \( \check{\mathcal{V}}_k \) of all \( k \)-tuples \( x = \{x_1, \ldots, x_k\} \) of distinct points in \( \mathbb{M}_\mathbb{C} \) which lie on some totally real submanifold \( M \) with two properties:

(i) The metric on \( M \) induced from \( \mathbb{M}_\mathbb{C} \) is allowable and
(ii) \( M \) projects surjectively onto the usual real Euclidean subspace \( \mathbb{E} = \mathbb{R}^d \) of \( \mathbb{M}_\mathbb{C} = \mathbb{E} \oplus i\mathbb{E} \).

Notice that, by the remark before Proposition 2.6, the projection \( M \to \mathbb{E} \) is a local diffeomorphism if the metric of \( M \) is allowable, so (ii) implies that \( M \) is the graph of a smooth map \( \mathbb{E} \to i\mathbb{E} \).

Let \( \mathcal{F}_k \) denote the space of all pairs \( (M, x) \) satisfying the above conditions. It is an infinite-dimensional complex manifold projecting to the open subset \( \check{\mathcal{V}}_k \) of \( \text{Conf}_k(\mathbb{M}_\mathbb{C}) \), and it is easy to see that the map \( \pi : \mathcal{F}_k \to \check{\mathcal{V}}_k \) is open. We define \( \mathcal{V}_k \) as the largest Hausdorff quotient manifold of \( \mathcal{F}_k \) through which \( \pi \) factorizes and which maps to \( \check{\mathcal{V}}_k \) by a local diffeomorphism. Thus, two points \( (M, x), (M', x) \) of the fibre \( \mathcal{F}_{k,x} \) of \( \pi \) at \( x \) have the same image in \( \mathcal{V}_k \) if they are in the same connected component of the fibre, but—as that equivalence relation need not give a Hausdorff quotient—also if there are paths \( \gamma, \gamma' \) from \( (M, x) \) and \( (M', x) \) to a third point \( (M'', x'') \) of \( \mathcal{F}_k \) which cover the same path from \( x \) to \( x'' \) in \( \check{\mathcal{V}}_k \).

To motivate the definition of \( \mathcal{V}_k \), we must enlarge our framework to allow Lorentzian space-times whose time-slices are not compact. The simplest way to do this is to introduce the cobordism category in which a morphism is the part of a \( d \)-dimensional allowable submanifold \( M \) of \( \mathbb{M}_\mathbb{C} \) cut off between two time-symmetric hypersurfaces.

A field theory said to be defined and holomorphic on this category, if it has a Lorentz-invariant vacuum state in a natural sense, will have vacuum expectation values, which are holomorphic functions \( \mathcal{E}_k \) on the spaces \( \mathcal{F}_k \) of pairs \( (M, x) \). Strictly speaking, \( \mathcal{E}_k \) is a holomorphic section of a bundle on \( \mathcal{F}_k \), but we can use the local diffeomorphism \( M \to \mathbb{E} \) to trivialize the bundle, giving us a holomorphic function

\[
\mathcal{E}_k : \mathcal{F}_k \to \text{Hom}(\mathcal{O}^\otimes k, \mathbb{C}),
\]

where \( \mathcal{O} \) is the space of observables at a point of \( \mathbb{E} \).

Our much-used Principle 5.1 tells us that the value of the function \( \mathcal{E}_k \) does not change if, while holding the marked points \( x \) fixed in \( \mathbb{M}_\mathbb{C} \), we move \( M \) smoothly in the allowable class. Therefore, in fact we have a holomorphic function on \( \mathcal{F}_k \) which is constant on the connected components of the fibres of \( \mathcal{F}_{k,x} \) of the map \( \pi \), that is, to the isotopy classes of allowable manifolds containing \( x \).

Unfortunately, we have no proof that \( \mathcal{V}_k \) is a domain of holomorphy, but at least we can assert
**Proposition 5.4** \( \mathcal{V}_k \) contains the Wightman domain \( \mathcal{U}_k \).

Furthermore, we saw in Proposition 2.3 that the holomorphic envelope of \( \mathcal{U}_k \) contains \( \mathcal{V}_k^{\text{flat}} \), the part of \( \mathcal{V}_k \) represented by flat affine submanifolds \( M \subset \mathbb{M}^\mathbb{C} \).

**Proof of 5.4** First, \( \mathcal{V}_k \) is invariant under the complex orthogonal group of \( \mathbb{M}^\mathbb{C} \) and under reorderings of the points \( x = \{x_1, \ldots, x_k\} \). Therefore, it is enough to consider \( x \) such that the imaginary part of \( x_{i+1} - x_i \) belongs to the forward light-cone \( C \subset \mathbb{M} \) for each \( i \).

Smoothing the obvious polygonal path joining the points, we can thus assume that the \( x_i \) lie on a curve \( x : \mathbb{R} \rightarrow \mathbb{M}^\mathbb{C} \) whose derivative \( \text{Im}(x'(t)) \) belongs to \( C \) for all \( t \). But then we can choose, smoothly in \( t \), a set of \( d-1 \) orthonormal vectors \( e_j(t) \) in \( \mathbb{M}^\mathbb{C} \) which are all orthogonal to \( x'(t) \). Let \( M_t \) be the real vector subspace of \( \mathbb{M}^\mathbb{C} \) spanned by the vectors \( e_j(t) \). The points \( x \) lie on the \( d \)-dimensional real ruled manifold \( M \) swept out by the affine \( d-1 \)-planes \( x(t)+M_t \), and the metric of \( M \) is clearly allowable. \( \square \)

**References**

1. R. Brunetti and K. Fredenhagen, *Quantum Field Theory on Curved Backgrounds*, Lecture Notes in Physics, Vol. 786, Springer, 2009, 129–155.
2. K. Costello, *Renormalization and Effective Field Theory*, Mathematical Surveys and Monographs, Vol. 170, Amer. Math. Soc, Providence, R.I, 2011.
3. E. Fermi, Quantum theory of radiation, *Rev. Mod. Phys.* 4 (1932), 87–132.
4. R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York, 1965.
5. L. Hörmander, *An Introduction to Complex Analysis in Several Variables*, North-Holland, Amsterdam, 1990.
6. R. Howe, *The Oscillator Semigroup*, Proc. Symp. Pure Math., Vol. 48, Amer. Math. Soc, Providence, R.I, 1988, 1196–1200.
7. D. Kazhdan, Introduction to QFT. *Quantum Fields and Strings: a course for mathematicians*, Princeton 1996/1997, Amer. Math. Soc, Providence, R.I, 1999, 377–418.
8. G. Kelnhofer, Functional integration and gauge ambiguities in generalized abelian gauge theories, *J. Geom. Phys.* 59 no. 7 (2009), 1017–1035, arXiv:0711.4085 [hep-th].
9. Y. A. Neretin, Holomorphic continuations of representations of the group of diffeomorphisms of the circle, *Mat. Sb.* 180 (1989), 635–57, English translation Math. USSR-Sb. 67 (1990).
10. A. Pressley and G. Segal, *Loop Groups*, Oxford University Press, Oxford, 1986.
11. G. Segal, The definition of conformal field theory. *Differential Geometrical Methods in Theoretical Physics (Como 1987)*, Vol. 250, NATO Adv. Sci. Inst. Ser. C, Math. Phys. Sci., Kluwer, 1988, 165–171.
12. G. Segal, The definition of conformal field theory. *Topology, Geometry, and Quantum Field Theory*, (Ed. U. Tillmann), Vol. 308, London Math. Soc. Lecture Notes, London, 2004, 421–577.
13. R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics, and All that*, Princeton University Press, Princeton, 2000 (1st edn, Benjamin, 1964).
14. R. Szabo, Quantization of higher abelian gauge theory in generalized differential cohomology, *Proc. 7th Internat. Conf. on Math. Methods in Physics (ICMP 2012)*, 2012, arXiv:1209.2530 [hep-th].
15. S. Weinberg and E. Witten, Limits on massless particles, *Phys. Lett. B*, 96 (1980), 59–62.