Wick rotation and the positivity of energy in quantum field theory

Maxim Kontsevich, I.H.E.S. Paris
and
Graeme Segal, All Souls College, Oxford

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

$$U_t \;=\; e^{\mathrm{i}Ht} : \mathcal{H} \to \mathcal{H}$$

(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. that 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

$$\{t \in \mathbb{C} : \mathrm{Im}(t) > 0\}$$

and is bounded in the operator norm.[1]

The holomorphic formulation helps us see what a strong constraint the positivity of energy is. It implies, for example, that if, for some $\xi \in \mathcal{H}$, the state $U_t(\xi)$ belongs to a closed subspace $\mathcal{H}_0$ of $\mathcal{H}$ for all $t < 0$ then it belongs to $\mathcal{H}_0$ for all $t \geq 0$, i.e. "nothing can happen for the first time".

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

[1]The physically relevant condition is actually that the energy is bounded below: replacing the Hamiltonian H by $H - c$ makes no observable difference. Rather than asking for U_t to be bounded for $\mathrm{Im}(t) > 0$ it is better to require $||U_t|| \leq \mathrm{e}^{c\mathrm{Im}(t)}$ for some c.

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.

In the usual formulation of quantum field theory, 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[2] in multilinear 'maps'

$$\begin{array}{ccc} \mathcal{O}_{x_1} \times \ldots \times \mathcal{O}_{x_k} & \longrightarrow & \mathbb{C} \\ (f_1, \ldots, f_k) & \mapsto & \langle f_1, \ldots, f_k \rangle_{(M,g)} \end{array} \tag{1}$$

for all sequences $\{x_1, \ldots, x_k\}$ of points in M, defining generalized functions[3] on the products M^k. The functions (1) are called *vacuum expectation values*. To come from a field theory they must satisfy a long list of conditions such as the Wightman axioms [SW]. These include a *causality* axiom which asserts that if the points $x_1, \ldots, x_k$ are spatially separated (i.e. no two can be joined by a time-like curve) then the expectation value is independent of the ordering of the points.

One motivation for this formulation is the "path-integral" picture, according to which the theory arises from a mythological superstructure consisting of a space Φ_M of "fields" of some kind which 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 $f : \Phi_M \to \mathbb{C}$ such that $f(\phi)$ depends only on the restriction of ϕ to an arbitrarily small neighbourhood of x. All of the physics of the theory is determined by an *action functional* $S_g : \Phi_M \to \mathbb{R}$ which notionally defines a complex-valued measure on the space Φ_M, symbolically denoted by $\mathrm{e}^{-\mathrm{i}S_g(\phi)/\hbar}\mathcal{D}\phi$. The parameter $\hbar$ here — the unit of action — is Planck's constant. The vacuum expectation values

[2]This is an oversimplification just for this introduction. In a gauge theory, for example, an observable such as a "Wilson loop" — the holonomy of the gauge field around a closed loop in space-time — is localized not at a point but at a loop, and we shall not exclude such features.

[3]The Wightman axioms ask for the vacuum expectation values to be distributions on M^k (which morally means that the theory has a logarithmic-conformal limit at short distances), but when the space-time dimension is three or more this is too strong to include natural examples such as the sigma-model with a circle as target, for which the vacuum expectation values are hyperfunctions but not distributions.

are given in terms of the measure by

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

The smallness of the unit $\hbar$ of action means that the notional integral is very highly oscillatory, and so the measure on Φ_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 more 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 $M_{\mathbb{C}}$, and asking for the the vacuum expectation values (1) to be the boundary-values of holomorphic functions on a domain in $(M_{\mathbb{C}})^k$. 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[4] 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.

The alternative approach — the one we shall pursue in this paper — is to treat the time-parameter t as the *length* of an oriented time-interval M equipped with a varying metric. Then we do not need to complexify M: we simply allow the *metric* on M to be complex-valued. There are two reasons why this 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

[4]In fact the motivation suggests not a complexification of M but a way of putting it on the boundary of a complex manifold. We shall come back to this in Section 5.

which does not converge even schematically. Its archetype is an improper Gaussian integral of the form

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

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 — i.e. 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 $\mathrm{Met}_{\mathbb{C}}(M)$ of complex-valued metrics on a smooth manifold M. The domain is a complexification of the manifold $\mathrm{Met}(M)$ of ordinary Riemannian metrics on M, and it has the real Lorentzian metrics (but *not* real metrics of other signatures) as a dense open 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, when the Lorentzian metric is globally hyperbolic, automatically satisfies the expected causality axiom. Finally, although we avoid complexifying space-time, our approach leads us to a conjecture about a question arising in the rival approach: for a theory defined in Minkowski space M, how can one characterize the largest domain in $(M_{\mathbb{C}})^k$ to which the vacuum expectation values extend holomorphically?

The relevant meaning of 'boundary' for the complex domains we are interested in is the *Shilov boundary*, which is usually defined only for finite-dimensional domains. If U is an open subset of a finite-dimensional complex manifold U^+, 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 on U which extends continuously to X we have

$$\sup_{x\in U} |f(x)| \;=\; \sup_{x\in K} |f(x)|.$$

The prime example is the polydisc

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

whose Shilov boundary is the torus $|z_1| = \ldots = |z_n| = 1$. The Shilov boundary of a manifold is part of its topological boundary, but can be much smaller. In our examples its real dimension is equal to the complex dimension of the domain. In the case of the Siegel domain of complex symmetric matrices the real symmetric matrices form the 'generalized real axis', which is a dense open subset of the Shilov boundary of the domain, as we shall describe in detail in Section 4.

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 Φ_M is the vector space $\mathrm{C}^\infty(M;\mathbb{R})$ of smooth functions, and the action is given by the quadratic form

$$\begin{aligned} \mathrm{i}S_g(\phi) &= \frac{1}{2}\int_M (d\phi \wedge *d\phi + m^2\phi \wedge *\phi) \\ &= \frac{1}{2}\int_M \left\{ \sum g^{ij} \frac{\partial\phi}{\partial x^i}\frac{\partial\phi}{\partial x^j} + m^2\phi^2 \right\} (\det g)^{1/2} |dx^1 \ldots dx^d|. \end{aligned}$$

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

$$*1 = \mathrm{vol}_g = (\det g)^{1/2} |dx^1 \ldots dx^d| \qquad (3)$$

Notice that for a Lorentzian metric g the volume element $*1$ is pure imaginary, corresponding to the $\sqrt{-1} = \mathrm{i}$ in front of the action S_g. The first condition we require of our complex metrics is that the real part of the twisted d-form vol_g defined by the formula (3) is 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.

To ensure that the real part of the quadratic form $\mathrm{i}S_g$ is positive-definite we also need the real part of the matrix $(\det g)^{1/2}g^{-1}$ — or equivalently of the inverse matrix $(\det g)^{-1/2}g$ — to be positive-definite.

We need more than this, however. The conditions so far would give us a domain whose boundary (like that of the Siegel generalized upper half-plane) contains indefinite real quadratic forms of all signatures, and not only the Lorentzian ones. 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

$$\mathrm{i}S_g(F) \;=\; \frac{1}{2}\int_M F \wedge *F.$$

The Hodge $*$-operator makes sense for a complex metric: for a p-form α we define a twisted $(d-p)$-form $*\alpha$ by taking the inner-product of α with $\mathrm{vol}_g \;=\; *1$, using the complex inner-product g.

For the electromagnetic field we need the real part of the quadratic form

$$\wedge^2(T_x^*) \;\longrightarrow\; |\wedge^d(T_x^*)| \otimes \mathbb{C}$$

given by $F \mapsto F \wedge *F$ to be positive-definite. (Here $|\wedge^d(T_x^*)|$ denotes the real line of volume elements on T_x obtained by tensoring $\wedge^d(T_x^*)$ by the line of orientation of T_x, and an element of it is positive if it is a positive volume-element.) This makes it natural, if we are going to consider space-time manifolds M of all dimensions, to require

Condition 2.1

for all *degrees* $p \geq 0$ *the real part of the quadratic form*

$$\wedge^p(T_x^*) \;\longrightarrow\; |\wedge^d(T_x^*)| \otimes \mathbb{C}$$

given by $\alpha \mapsto \alpha \wedge *\alpha$ *is positive-definite.*

Theorem 2.2 *Condition 2.1 is equivalent to: there is a basis of the real vector space* T_x *in which the metric* g *takes the form*

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

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

$$|\arg(\lambda_1)| + |\arg(\lambda_2)| + \ldots + |\arg(\lambda_d)| < \pi. \tag{4}$$

The complex-valued quadratic forms $g : V \to \mathbb{C}$ on a real vector space V which satisfy the conditions of the theorem form an open subset $Q_{\mathbb{C}}(V)$ of the complex vector space $S^2(V^*_{\mathbb{C}})$. We shall refer to them as the *allowable* complex metrics. 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 π, and the inequality (4) of Theorem 2.2 shows that at most *one* of the $|\arg(\lambda_i)|$ can become π on the boundary. Another consequence of the inequality (4) is that

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

which shows that when v runs through V the complex numbers $g(v)$ form a closed convex cone in $\mathbb{C}$ disjoint from the open negative real axis.

We define the space $\mathrm{Met}_{\mathbb{C}}(M)$ 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 $Q_{\mathbb{C}}(T_x)$.

Before giving the surprisingly simple proof of Theorem 2.2 let us say a little more to motivate its conditions. The desire to make the path integral look more convergent hardly needs further comment, but choosing to focus on the quadratic '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? Including the higher gauge theories, however, does at least 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. [Ke], [Sz](4.4)). The gauge-equivalence classes of fields form an infinite-dimensional Lie group, and 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

$$H^{p-1}(M;\mathbb{T}) \;\times\; \Phi_0 \;\times\; \mathrm{Harm}^p_{\mathbb{Z}}(M),$$

where Φ_0 is the vector space of exact p-forms on M, and $\mathrm{Harm}^p_{\mathbb{Z}}(M)$ is the finite-dimensional lattice of harmonic (and hence constant) p-forms with integral periods. The partition function is likewise a product of three terms: the torus of flat fields contributes its volume (for an appropriate metric determined by the geometry of M), the lattice Γ of harmonic p-forms contributes its theta-function

$$\sum_{\alpha\in\Gamma} \exp\left(-\frac{1}{2}\int_M \alpha\wedge *\alpha\right) ,$$

while the vector space Φ_0 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 'Siegel half-plane' of complex metrics with positive real part; but the theta-function can not be continued beyond those metrics for which the real part of the form $\int \alpha\wedge *\alpha$ is positive.

In the opposite direction, the inequality (4) is motivated by a *lower* bound on the class of complex metrics we would like to include, coming from the traditional analytical continuation of vacuum expection values to an open subset of the k-fold product of complexified Minkowski space $\mathbb{M}_{\mathbb{C}}$. The Wightman axioms imply that the expectation values extend holomorphically to a domain $\mathcal{U}_k$ called the 'permuted extended tube'[5]. It is a very basic result in the Wightman approach to quantum field theory (cf. [SW], or [Ka](2.1)) that $\mathcal{U}_k$ contains the configuration space $\mathrm{Conf}_k(\mathbb{E})$ of all k-tuples of *distinct* points of the standard Euclidean subspace $\mathbb{E}\subset\mathbb{M}_{\mathbb{C}}$. That makes it very natural to include among the allowable metrics those of all the d-dimensional real subspaces V of $\mathbb{M}_{\mathbb{C}}$ such that the configuration space $\mathrm{Conf}_k(V)$ of distinct k-tuples in V is contained in the holomorphic hull of $\mathcal{U}_k$. But we have:

Proposition 2.3 *If the induced metric on a d-dimensional real subspace V of $\mathbb{M}_{\mathbb{C}}$ satisfies condition (4) above then $\mathrm{Conf}_k(V)$ is contained in the holomorphic hull of $\mathcal{U}_k$.*

[5] A set of points $x_1,\ldots,x_k$ belongs to $\mathcal{U}_k$ if, after ordering them suitably, there is an element γ 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.

We shall sketch a proof of this result after Proposition 2.7 below.

Proof of Theorem 2.2 To diagonalize a complex symmetric matrix $g = A + \mathrm{i}B$ 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. If g is diagonalized as in the theorem with respect to a basis $\{e_i\}$ of T_x, then the form $\alpha \mapsto \alpha \wedge *\alpha$ on $\wedge^p(T_x^*)$ 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 $\alpha \wedge *\alpha$ on e_S^* is

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

whose argument is

$$\frac{1}{2} \left\{ \sum_{i \in S} \arg(\lambda_i) - \sum_{i \notin S} \arg(\lambda_i) \right\}.$$

The result follows by taking S to be the the set with $\arg(\lambda_i)$ negative. ♠

The proof of Theorem 2.2 shows that to give an element g of $Q_{\mathbb{C}}(V)$ 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

$$\sum_k \dim V_k \cdot |\theta_k| < \pi,$$

and on V_k the bilinear form g is $\mathrm{e}^{\mathrm{i}\theta_k}$ times a real positive-definite form. The only ambiguity in this is that if, say, $\theta_k = \theta_{k+1}$ we can replace V_k by $V_k \oplus V_{k+1}$ and omit θ_{k+1} and V_{k+1}. This means that the subspace $P = \bigoplus \mathrm{e}^{-\mathrm{i}\theta_k/2} V_k$ of the complexification $V_{\mathbb{C}}$ of V is *canonically* associated to the form g. On the real subspace P the complex bilinear form g is real and positive-definite. Our argument shows that

$$V \;\cong\; \exp(\mathrm{i}\pi\Theta/2)(P) \;\subset\; P_{\mathbb{C}} \;=\; V_{\mathbb{C}},$$

where $\Theta : P \to P$ is a self-adjoint operator with trace-norm[6] $||\Theta||_1 < 1$. This shows that the space $Q_{\mathbb{C}}(V)$ is parametrized by the pairs (g_0, Θ), where g_0 is a positive-definite inner-product on V and Θ 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 *$Q_{\mathbb{C}}(V)$ 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*

$$Q_{\mathbb{C}}(V) \;\cong\; \mathrm{GL}(V) \times_{\mathrm{O}(V)} \Pi(V).$$

In particular, $Q_{\mathbb{C}}(V)$ 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 because its point of view on the angles θ_i as critical values helps give a feeling for allowable complex metrics.

Proposition 2.5 *If $g \in Q_{\mathbb{C}}(V)$ and W is any vector subspace of V then $g|W$ belongs to $Q_{\mathbb{C}}(W)$.*

Proof For any $g \in Q_{\mathbb{C}}(V)$ 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 $\sum \mathrm{e}^{\mathrm{i}\theta_k} y_k^2$. The numbers θ_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 θ^A and θ_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

[6]The trace-norm is the sum of the absolute values of the eigenvalues.

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. ♠

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 bilinear 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 A 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 is the complexification of a real vector space $Z_{\mathbb{R}}$ with a positive-definite metric then (by (2.5)) the projection $V \subset Z \to Z_{\mathbb{R}}$ is an isomorphism.

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

$$\mathrm{O}_{\mathbb{C}}(Z) \times_{\mathrm{O}(Z_{\mathbb{R}})} \Pi(Z_{\mathbb{R}}).$$

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 $\mathrm{GL}(V)$ and $\mathrm{O}_{\mathbb{C}}(Z)$ 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 $\mathrm{GL}(V)$-bundle with base $\mathcal{R}(Z)$ and a principal $\mathrm{O}_{\mathbb{C}}(Z)$-bundle with base $Q_{\mathbb{C}}(V)$. But the Lie groups $\mathrm{GL}(V)$ and $\mathrm{O}_{\mathbb{C}}(Z)$ are homotopy equivalent to their maximal compact subgroups, i.e. 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 $\mathrm{O}_{\mathbb{C}}(Z)f$ 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 $\mathrm{O}(V) \to \mathrm{O}_{\mathbb{C}}(Z)f \to \mathcal{R}(V; Z)$. But $\mathrm{O}(V)$ is also contained in and equivalent to the fibre $f\mathrm{GL}(V)$ of the other fibration $\mathcal{R}(V; Z) \to \mathcal{R}(Z)$, which implies the contractibility of its base $\mathcal{R}(Z)$. ♠

The last property of $Q_{\mathbb{C}}(V)$ which we shall record is

Proposition 2.7 *The domain $Q_{\mathbb{C}}(V)$ is holomorphically convex.*

Proof The Siegel domain $\mathcal{U}(V)$ of complex-valued inner products with positive-definite imaginary part on a real vector space V is known to be a domain of holomorphy in $S^2(V^*_{\mathbb{C}})$. 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 $Q_{\mathbb{C}}(V)$ is the intersection of this product domain with the affine variety which is the natural embedding of $S^2(V^*_{\mathbb{C}})$ in this ambient vector space, and so it too is a domain of holomorphy.

Sketch proof of 2.3

If V is a totally-real subspace of $\mathbb{M}_{\mathbb{C}}$ whose induced metric is allowable then the preceding discussion shows that, up to a complex orthogonal transformation of $\mathbb{M}_{\mathbb{C}}$,

$$V \;=\; \exp(\mathrm{i}\Theta/2)(\mathbb{E}).$$

Here $\mathbb{E}$ is the standard Euclidean subspace of $\mathbb{M}_{\mathbb{C}}$, identified with $\mathbb{R}^d$ so that the first coordinate is the Wick-rotated time, and Θ is a real diagonal matrix whose entries $\theta_1, \ldots, \theta_d$ belong to the 'generalized octagon' $\Pi_0 \subset \mathbb{R}^d$ consisting of those Θ whose diagonal entries $\theta_1, \ldots, \theta_d$ satisfy the inequality (4). What we want to prove is that when $\Theta \in \Pi_0$ the diagonal matrix $\exp(\mathrm{i}\Theta/2)$ maps each k-tuple $\mathbf{x} = \{x_1, \ldots, x_k\}$ of distinct points of $\mathbb{E}$ to a point of the Wightman 'permuted extended tube' $\mathcal{U}_k$. It is enough to do this when the points of $\mathbf{x}$ have their r^{th}-coordinates all distinct for $1 \leq r \leq d$, for $\mathcal{U}_k$ is invariant under the complex orthogonal group $\mathrm{O}(\mathbb{M}_{\mathbb{C}})$, and the coordinates can be made distinct by applying an element of $\mathrm{O}(\mathbb{E}) \subset \mathrm{O}(\mathbb{M}_{\mathbb{C}})$. (Notice that the map $\exp(\mathrm{i}\Theta/2)$ commutes with the action of $\mathrm{O}(\mathbb{E})$.)

Having fixed a set of points $\mathbf{x}$, let us consider the open subset U of $\mathbb{C}^d$ consisting of all *complex* diagonal matrices Θ such that $\exp(\mathrm{i}\Theta/2)(\mathbf{x})$ is contained in the holomorphic hull of $\mathcal{U}_k$. We propose to show that U contains the open 'tube domain' $\Pi_0 \times \mathrm{i}\mathbb{R}$.

The crucial fact is that Π_0 is the convex hull of its subspace Π_{00} consisting of the matrices with rank 1, i.e. those for which only one of the θ_r is non-zero. If the real part of Θ belongs to Π_{00} then $\exp(\mathrm{i}\Theta/2)(\mathbf{x})$ is contained in $\mathcal{U}_k$, for

a point of $\mathbb{C}^d$ with only one non-zero coordinate λ in the rth place belongs to the extended forward tube if $r = 1$ and $\mathrm{Re}(\lambda) > 0$ or $r > 1$ and $\mathrm{Im}(\lambda) > 0$, and for each r we can re-order the points $x_1, \ldots, x_k$ so that $x_i - x_{i+1}$ has its rth coordinate positive for $1 \leq i \leq k-1$, and then $\exp(\mathrm{i}\Theta/2)$ will take $x_i - x_{i+1}$ into the extended forward tube.

We should now like to apply Bochner's 'tube theorem', which asserts that if P is a connected open subset of $\mathbb{R}^d$ then a holomorphic function defined in the tube domain $P \times \mathrm{i}\mathbb{R}$ extends holomorphically to the tube domain whose base is the convex hull of P. Unfortunately our domain U is not an open tube domain, though it contains a neighbourhood of the non-open tube domain $\Pi_{00} \times \mathrm{i}\mathbb{R}^d$. Experts seem to believe that Bochner's theorem applies in our situation, but among the many generalizations in the literature we have not found one that exactly covers it. As Proposition 2.3 plays only a motivational role in this paper we have not pursued the question further.

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 2-dimensional real vector space V is determined up to a scalar multiple by its two distinct null-directions in the complexified space $V_{\mathbb{C}}$. We can think of these as two points of the Riemann sphere $\mathbb{P}(V_{\mathbb{C}})$. 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 $\mathbb{P}(V_{\mathbb{C}})$ 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 $\mathbb{P}(V_{\mathbb{C}})$ 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 Σ any almost-complex structure is integrable, so a point of $\mathrm{Met}_{\mathbb{C}}(\Sigma)$ is a pair of complex structures of opposite orientations on Σ, 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 $Q_{\mathbb{C}}(V)$ is thus a 3-dimensional polydisc, one disc for each of the complex structures, and the third for the volume-element.

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* 1-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 *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. The energy of a state is

$$\frac{1}{2}\int_M ||\nabla\phi||^2 \mathrm{d}s,$$

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 ω 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*

$$\sqrt{g} = R + \mathrm{i}\omega L + 1/\mathrm{i}\omega C,$$

a complex number which defines an allowable metric because $\mathrm{Re}\sqrt{g} > 0$.

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.

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* $\mathcal{C}_d^{\mathbb{C}}$ 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 $\mathcal{C}_d^{\mathbb{C}}$ are compact smooth $(d-1)$-dimensional manifolds Σ equipped with complex metrics $g \in \mathrm{Met}_{\mathbb{C}}(\Sigma)$. A morphism from Σ_0 to Σ_1 is a cobordism M from Σ_0 to Σ_1, also with a complex metric. We shall write $M : \Sigma_0 \rightsquigarrow \Sigma_1$ 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 $\mathcal{C}_d^{\mathbb{C}}$ is not a $(d-1)$-manifold but rather is a *germ* of a d-manifold along a given $(d-1)$-manifold Σ — i.e. Σ is given as a closed submanifold of a d-manifold U, but any two open neighbourhoods of Σ in U define the same object of $\mathcal{C}_d^{\mathbb{C}}$. We require Σ 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 U, the co-orientation, and the complex metric g from the notation.) Furthermore, two morphisms M and M' from Σ_0 to Σ_1 are identified if there is an isometry $M \to M'$ which is the identity on the germs Σ_0 and Σ_1.

In terms of the category $\mathcal{C}_d^{\mathbb{C}}$ we define a d-dimensional field theory as a *holomorphic* functor from $\mathcal{C}_d^{\mathbb{C}}$ to the category of Fréchet topological vector spaces and *nuclear* (i.e. trace-class) linear maps. We shall write E_Σ for the vector space associated to an object Σ, and $Z_M : E_{\Sigma_0} \to E_{\Sigma_1}$ for the linear map associated to a cobordism $M : \Sigma_0 \rightsquigarrow \Sigma_1$. To say that the functor is 'holomorphic' means that, for a given smooth manifold-germ $\Sigma \subset U$, the topological vector spaces E_Σ form a locally trivial holomorphic vector bundle on the complex manifold $\mathrm{Met}_{\mathbb{C}}(U)$ 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 $\mathrm{Met}_{\mathbb{C}}(M)$ (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_Σ will have a mod 2 grading, and the usual sign-conventions will 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 $\Sigma' \rightsquigarrow \Sigma$ 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 Σ' with cobordisms to Σ. Similarly, looking downstream rather than upstream, we have a family of cobordisms $\Sigma \rightsquigarrow \Sigma''$ contained in U, giving us an inverse system of objects Σ'' to which Σ maps. For any field theory, therefore, there are natural maps

$$\lim_{\rightarrow} E_{\Sigma'} \quad \rightarrow \quad E_\Sigma \quad \rightarrow \quad \lim_{\leftarrow} E_{\Sigma''}.$$

We shall assume the functor has the *continuity* property that each of these maps is injective with dense image. We shall write $\check{E}_\Sigma$ for the upstream limit $\lim_\rightarrow E_{\Sigma'}$, and $\hat{E}_\Sigma$ for the downstream limit. Then $\hat{E}_\Sigma$, as the inverse-limit of a countable sequence of nuclear maps of Fréchet spaces, is a *nuclear* Fréchet space[7]. The other space $\check{E}_\Sigma$ is also nuclear, but not usually metrizable: it is the dual of the nuclear Fréchet space $\hat{E}_{\Sigma^*}$, where Σ^* denotes the germ Σ 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 $\check{E}_{\Sigma_0} \rightarrow \check{E}_{\Sigma_1}$ and $\hat{E}_{\Sigma_0} \rightarrow \hat{E}_{\Sigma_1}$, and both of them factorize though a map $\hat{E}_{\Sigma_0} \rightarrow \check{E}_{\Sigma_1}$. In fact this is equivalent to the original assumption that U_M is nuclear, because any map from the dual of a nuclear Fréchet space to a nuclear Fréchet space is nuclear.

The space E_Σ 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:

$$\check{E}_{\Sigma \sqcup \Sigma'} \cong \check{E}_\Sigma \otimes \check{E}_{\Sigma'}.$$

[7] A very useful concise account of nuclear spaces can be found in [C].

This implies — and is equivalent to — $\hat{E}_{\Sigma\sqcup\Sigma'} \cong \hat{E}_\Sigma \otimes \hat{E}_{\Sigma'}$. Of course for two cobordisms M and M' we also assume

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

There is a unique natural concept of tensor product here, because all the vector spaces are nuclear.

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* $Z_M \in \mathrm{End}(E_\emptyset) = \mathbb{C}$. The whole structure of the theory is a way of expressing the sense in which the number Z_M deppends *locally* on M.

In this discussion we have still committed some abuses of language. The "category" $\mathcal{C}_d^{\mathbb{C}}$ 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 $\check{E}_\Sigma \to \hat{E}_\Sigma$ 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. But whatever 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 $P \times \Sigma_0 \to \Sigma_1$ parametrized 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 $\mathcal{O}_x$ 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 $D \setminus \mathring{D}'$ 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 $\mathcal{O}_x$ 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 $x_i \in \mathring{D}_i$. Then $M' = M \setminus \bigcup \mathring{D}_i$

is a cobordism from $\bigsqcup \partial D_i$ to the empty $(d-1)$-manifold $\emptyset$, and defines $Z_{M'} : E_{\sqcup \partial D_i} \to E_\emptyset = \mathbb{C}$. Using the tensoring property we can write this

$$Z_{M'} : \bigotimes E_{\partial D_i} \longrightarrow \mathbb{C},$$

and then we can pass to the inverse-limits to get the expectation-value map

$$\bigotimes \mathcal{O}_{x_i} \longrightarrow \mathbb{C}.$$

We might prefer the language of "field operators" to that of vacuum expectation values. If the space-time M is a cobordism $\Sigma_0 \rightsquigarrow \Sigma_1$, then for any x in the interior of M — say $x \in \mathring{D} \subset M$ — the cobordisms $M \setminus \mathring{D}$ define maps

$$\mathcal{O}_x \to \mathrm{Hom}(E_{\Sigma_0}; E_{\Sigma_1}),$$

while if x lies on a hypersurface Σ an observable at x defines a map $\check{E}_\Sigma \to \hat{E}_\Sigma$, i.e. it acts on E_Σ as an unbounded operator. But 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.

Lorentzian manifolds

There is a category $\mathcal{C}_d^{\mathrm{Lor}}$ which at first sight looks more relevant to quantum field theory than $\mathcal{C}_d^{\mathbb{C}}$. 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. [BF]) have developed the theory of quantum fields in curved space-time using a version of this category. The category $\mathcal{C}_d^{\mathrm{Lor}}$ lies "on the boundary" of the category $\mathcal{C}_d^{\mathbb{C}}$. In section 5 we shall discuss the sense in which a representation of $\mathcal{C}_d^{\mathbb{C}}$ has a "boundary value" on $\mathcal{C}_d^{\mathrm{Lor}}$, at least if it is *unitary*.

Unitarity

So far we have not asked for an inner product on the topological vector space E_Σ associated to a $(d-1)$-manifold Σ. 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 Σ^* denotes the manifold germ Σ with its co-orientation reversed, then $\check{E}_{\Sigma^*}$ is the dual topological

vector space to $\hat{E}_\Sigma$. Furthermore, a cobordism $M : \Sigma_0 \rightsquigarrow \Sigma_1$ can also be regarded as a cobordism from Σ_1^* to Σ_0^*, and the two maps $E_{\Sigma_0} \to E_{\Sigma_1}$ and $E_{\Sigma_1^*} \to E_{\Sigma_0^*}$ 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_Σ to have an inner product for every $(d-1)$-manifold Σ. A complex metric $g \in \mathrm{Met}_{\mathbb{C}}(\Sigma)$ has a complex conjugate $\bar{g}$. If we write $\bar{\Sigma}$ for Σ with the metric $\bar{g}$ but with its co-orientation unchanged[8] then $\Sigma \mapsto \bar{\Sigma}$ is a *covariant* functor. It is natural to require that

$$E_{\bar{\Sigma}} \cong \bar{E}_\Sigma. \tag{5}$$

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

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

$$\check{E}_\Sigma \;\to\; E_\Sigma^{Hilb} \;\to\; \hat{E}_\Sigma.$$

The theory on flat tori

[8] If our theory is defined on a category of *oriented* space-time manifolds, we must give $\bar{\Sigma}$ the opposite orientation to Σ, although the same co-orientation in U.

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

$$\mathrm{O}_d\backslash\mathrm{GL}_d(\mathbb{R})/\mathrm{SL}(\mathbb{Z}) \;\cong\; Q(\mathbb{R}^d)/\mathrm{SL}_d(\mathbb{Z}),$$

where $Q(\mathbb{R}^d) = \mathrm{O}_d\backslash\mathrm{GL}_d(\mathbb{R})$ 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 $\mathrm{SL}_d(\mathbb{Z})$. The axioms we have proposed 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, for each way of writing a torus M as a cobordism $\tilde{M} : \Sigma \rightsquigarrow \Sigma$ from a $(d-1)$-dimensional torus Σ to itself we have $Z(M) = \mathrm{trace}(Z_{\tilde{M}})$, where $Z_{\tilde{M}} : E_\Sigma \to E_\Sigma$ is a nuclear operator in the vector space E_Σ, which is graded by the characters of the translation-group of Σ. More explicitly, we have

$$Q(\mathbb{R}^d) \;=\; Q(\mathbb{R}^{d-1}) \times \mathbb{R}_+^\times \times \mathbb{R}^{d-1},$$

and with repect to this decomposition we must have

$$Z(A, \mathrm{e}^t, \xi) \;=\; \sum_{i,\alpha} \chi_{i,\alpha}(\xi)\; \mathrm{e}^{-\lambda_i t},$$

where $\{\lambda_i = \lambda_i(A)\}$ is the sequence (tending to $+\infty$) of eigenvalues of the Hamiltonian operator on E_Σ, which is graded by the character-group $\mathbb{Z}^{d-1}$ of the torus $\Sigma = \mathbb{R}^{d-1}/\mathbb{Z}^{d-1}$.

Appendix: The duality $(\check{E}_\Sigma)^* \;\cong\; \hat{E}_{\Sigma^*}$

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[9] which is coherently associative and commutative, and that we are given natural isomorphisms $E_{\Sigma_1} \otimes E_{\Sigma_2} \;\to\; E_{\Sigma_1 \sqcup \Sigma_2}$.

[9]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

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

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

Now suppose that we have an infinite sequence of cobordisms

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

indexed by $i \geq 0$, which form the downstream tail of a manifold-germ Σ, i.e. the sequence which we used above to define the space $\hat{E}_\Sigma = \lim_{\leftarrow} E_{\Sigma_i}$. 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 $\hat{E}_\Sigma$. But the inverse-limit of the cofinal sequence of odd terms is

$$\lim_{\leftarrow} (E_{\Sigma^*_{21+1}})^* \ = \ (\lim_{\rightarrow} E_{\Sigma^*_{21+1}})^*.$$

This shows that $\hat{E}_\Sigma \cong (\check{E}_{\Sigma^*})^*$. But, because $\hat{E}_\Sigma$ is automatically a nuclear Fréchet space, we can dualize again and conclude that $(\hat{E}_\Sigma)^* \cong \check{E}_{\Sigma^*}$ also.

4 Some analogies from representation theory

To understand the relation between representations of the category $\mathcal{C}_d^{\mathbb{C}}$ and of the Lorentzian category $\mathcal{C}_d^{\mathrm{Lor}}$ which lies "on its boundary" it is helpful to consider the representation theory of some finite-dimensional groups which lie on the boundary of complex semigroups.

The group $G = \mathrm{PSL}_2(\mathbb{R})$ is the group of Möbius transformations of the Riemann sphere $\Sigma = \mathbb{C} \cup \infty$ which map the open upper half-plane $\mathbb{U}$ to itself. It lies on the boundary of the complex sub-semigroup of $G_{\mathbb{C}} = \mathrm{PSL}_2(\mathbb{C})$ 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 $G_{\mathbb{C}}^{<}$ by including the degenerate Möbius transformations which collapse $\mathbb{U}$ to a single point in $\mathbb{U}$ — these correspond to complex 2×2 matrices of rank one. The resulting semigroup is then a contractible open subset of the 3-dimensional complex projective space formed from the 2×2 matrices. The topological boundary of $G_{\mathbb{C}}^{<}$ consists of the Möbius transformations which take $\mathbb{U}$ to a disc or point in the upper half-plane which touches the real axis, and the Shilov boundary consists of the group G of real Möbius transformations — an open solid torus — compactified by its 2-torus boundary, which is the hyperboloid $\det(A) = 0$ 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 $\det(A) \geq 0$.)

The irreducible unitary representations of the group $G = \mathrm{PSL}_2(\mathbb{R})$ are essentially[10] of two kinds, the *principal series* and the *discrete series*. 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 $\mathrm{Re}(s) = 1/2$. The best-known discrete series representation is the action of G on the square-summable holomorphic 1-forms on $\mathbb{U}$, with the natural norm

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

— 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 $(p-1)^{\mathrm{st}}$ power of the G-invariant area form on the Poincaré plane $\mathbb{U}$ to define the norm. The discrete series representations obviously extend to bounded holomorphic representations of the semigroup $G_{\mathbb{C}}^{<}$ by contraction operators, while the principal series representations cannot extend because when $|a| < 1$ the element $w \mapsto aw$ (here $w = (z - \mathrm{i})/(z + \mathrm{i})$ is the coordinate in the unit-disc model $|w| < 1$ of $\mathbb{U}$) of the semigroup $G_{\mathbb{C}}^{<}$ would be represented by an operator whose eigenvalues are a^n for *all* $n \in \mathbb{Z}$. But let us notice that, though the discrete series representations are unitary on the boundary group $G = \mathrm{PSL}_2(\mathbb{R})$,

[10] We shall ignore the "supplementary" series, which is of measure zero in the space of representations.

the degenerate elements of $G_{\mathbb{C}}^{<}$, which collapse $\mathbb{U}$ to a point $p \in \mathbb{U}$, are represented by bounded operators of rank 1. So these unitary representations of $\mathrm{PSL}_2(\mathbb{R})$ do not extend unitarily to the whole Shilov boundary: the degenerate elements correspond to rank 1 operators $\xi \mapsto \langle \zeta, \xi \rangle \eta$, where η and ζ are "non-normalizable elements" of the Hilbert space — i.e. they belong to an appropriate completion of it.

The group G is a subgroup of the group $\mathrm{Diff}^+(S^1)$ of orientation-preserving diffeomorphisms of the circle. This infinite-dimensional Lie group does not possess a complexification, though 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 quite independently ([N],[S1]) by the two authors, and also by Yu. Neretin, in the 1980s that there is an infinite-dimensional complex semigroup $\mathcal{A}$ which has exactly the same relation to $\mathrm{Diff}^+(S^1)$ as $G_{\mathbb{C}}^{<}$ has to $G = \mathrm{PSL}_2(\mathbb{R})$. 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 the complex plane. The annuli form a complex semigroup when concatenated as cobordisms, and the lowest-weight or "positive-energy" representations of $\mathrm{Diff}^+(S^1)$ which arise in 2-dimensional conformal field theory are precisely those which are boundary values of holomorphic representations of the semigroup $\mathcal{A}$ by trace-class operators.

The discussion of $\mathrm{PSL}_2(\mathbb{R})$ generalizes to the symplectic group $G = \mathrm{Sp}(V) \cong \mathrm{Sp}_{2n}(\mathbb{R})$ 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[11] Lagrangian subspaces of the complexification $V_{\mathbb{C}}$. The group G lies on the boundary of a semigroup $G_{\mathbb{C}}^{<}$ which is the Siegel domain $\mathbb{U}(\tilde{V} \oplus V)$, where $\tilde{V}$ 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 $\mathrm{PSL}_2(\mathbb{C})$, there are degenerate elements which map $\mathbb{U}(V)$ non-injectively into itself. The complex

[11] A complex Lagrangian subspace W is *positive* if $\mathrm{i}\omega(w, \bar{w}) > 0$ for every non-zero $w \in W$, where ω is the symplectic form of $\hat{V}_{\mathbb{C}}$. If we choose a decomposition $V = Q \oplus Q^*$ of V as a sum of real Lagrangian subspaces then positive Lagrangian subspaces W of $V_{\mathbb{C}}$ can be identified with complex-valued symmetric forms on Q whose imaginary part is positive-definite.

semigroup $G_{\mathbb{C}}^{<}$ has been carefully studied by Roger Howe [H], who called it the *oscillator semigroup*.

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

The group $G = \mathrm{Sp}(V)$ has a discrete series of unitary representations generalizing those of $\mathrm{PSL}_2(\mathbb{R})$. The most important is the *metaplectic representation* — actually a representation of a double covering $\tilde{G}$ of $\mathrm{Sp}(V)$ — 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 $(\bigwedge^n(W))^{\otimes(1/2)}$ for each point W of the domain $\mathbb{U}(V)$ — the square-root of the natural hermitian holomorphic line bundle $\{\bigwedge^n(W)\}$ on $\mathbb{U}(V)$ is canonical up to multiplication by ± 1, and is holomorphically embedded in $\mathcal{H}_V$. It is acted on by $\tilde{G}$ rather than G.

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

$$\mathcal{H}_{\tilde{V}\oplus V} \;\cong\; \mathcal{H}_V^* \otimes \mathcal{H}_V \;\cong\; \mathrm{End}_{HS}(\mathcal{H}_V),$$

where End_{HS} denotes the Hilbert-Schmidt endomorphisms[12].

When $n = 1$ the group $\mathrm{Sp}(V)$ is $\mathrm{SL}_2(\mathbb{R})$, a double covering of the group $\mathrm{PSL}_2(\mathbb{R})$ of Möbius transformations we considered before. To relate the cases of $\mathrm{PSL}_2(\mathbb{R})$ and $\mathrm{Sp}(V)$, recall that $\mathrm{PSL}_2(\mathbb{C})$ is an open subspace of the complex projective space $\mathbb{P}^3_{\mathbb{C}}$ formed from the vector space of 2×2 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, i.e. by the matrices of rank 1. The double covering group $\mathrm{SL}_2(\mathbb{C})$ sits inside the Grassmannian of complex Lagrangian subspaces of $\mathbb{C}^4$, which is a quadric 3-fold $Q^3_{\mathbb{C}}$ in $\mathbb{P}^4_{\mathbb{C}}$: it is a non-singular

[12] A more careful discussion shows that $G_{\mathbb{C}}^{<}$ is represented by operators of trace class.

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-matrices, branched along the quadric $Q^2_{\mathbb{C}}$ of rank 1 matrices. The contractible semigroup $\mathrm{SL}_2^<(\mathbb{C})$ 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 $\mathrm{PSL}_2^<(\mathbb{C})$.

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 $\mathcal{C}_d^{\mathrm{Lor}}$ lies on the boundary of $\mathcal{C}_d^{\mathbb{C}}$, and the role of the open dense subgroup of the Shilov boundary is played by the subcategory of *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 $\mathcal{C}_d^{\mathrm{Lor}}$ may be smaller than the optimal category we could put on the boundary of $\mathcal{C}_d^{\mathbb{C}}$. 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 $\mathcal{C}_d^{\mathbb{C}}$ let us first consider how to extend the assignment $\Sigma \mapsto E_\Sigma$ to a Lorentzian germ $\Sigma \subset U$ with Σ co-oriented in U. We can identify U with $\Sigma \times (-\varepsilon, \varepsilon)$ by exponentiating the geodesic curves emanating perpendicularly from Σ. The metric then takes the form $h_t - \mathrm{d}t^2$, where $t \mapsto h_t$ is a smooth map from $(-\varepsilon, \varepsilon)$ to the manifold of Riemannian metrics on Σ. If the germ is time-symmetric then we can define E_Σ by replacing the Lorentzian metric by the 'Wick rotated' Riemannian metric $h_{\mathrm{i}t} + \mathrm{d}t^2$, which makes sense because if $h_t = h_{-t}$ then h_t is a function of t^2, so that $h_{\mathrm{i}t}$ is defined and real. But 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 below.

It is less easy to assign an operator $Z_M : E_{\Sigma_0} \to E_{\Sigma_1}$ to a Lorentzian cobordism $M : \Sigma_0 \rightsquigarrow \Sigma_1$. Even if M is a cylinder topologically, it can be complicated in its "causal" structure. Consider, for example, a 2-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 which 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. [S2] 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 2-dimensional conformal field theory comes with a projective unitary representation of the group $\mathrm{Diff}^+(S^1) \times \mathrm{Diff}^+(S^1)$.

But 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, though, unlike a black hole, the Lorentzian metric here has no singularity.[13]

In works on general relativity a Lorentzian cobordism $M : \Sigma_0 \rightsquigarrow \Sigma_1$ between Riemannian manifolds is called *globally hyperbolic*[14] if every maximally-

[13]In this example, the "blocked" foliation is conformally the same as the "degenerate annulus" obtained by collapsing the closed light-line to a point, i.e. a pair of discs with their centre-points identified. This is usually regarded as an "annulus of infinite length", and it acts on an irreducible positive-energy representation of $\mathrm{Diff}^+(S^1)$ by a projection operator of rank one, just as a degenerate complex Möbius transformation acts by a rank 1 operator in a discrete-series representation of $\mathrm{PSL}_2(\mathbb{R})$.

[14]Of course we are only considering *compact* time-slices, which is not the usual focus in relativity theory.

extended time-like geodesic in M travels from Σ_0 to Σ_1. Such an M must be diffeomorphic to $\Sigma_0 \times \mathrm{i}[0,1]$, and in our compact situation we can take the definition of global hyperbolicity to be the existence of a smooth time-function $\tau : M \to \mathrm{i}[0,1]$ which is a fibration with Riemannian fibres. It is only for globally hyperbolic manifolds that, for example, the Cauchy problem for the wave-equation on M is soluble.

The globally hyperbolic cobordisms form an open subcategory $\mathcal{C}_d^{\mathrm{gh}}$ of $\mathcal{C}_d^{\mathrm{Lor}}$ which should play the role of the real Lie group to which the holomorphic contraction representations of Section 4 can be extended (though the result (5.2) we prove below is unfortunately weaker).

If M is a globally hyperbolic cobordism equipped with a time-function, we define a diffeomorphism $M \to \Sigma_0 \times \mathrm{i}[0,1]$ by following the orthogonal trajectories to the time-slices. In this description the metric takes the form $h_\tau + c^2 \mathrm{d}\tau^2$ for some function $c : \Sigma_0 \times [0,1] \to \mathbb{R}$. A small deformation δc 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 δc, but, more importantly, on the choice of the time-function, which should be irrelevant to the operator U_M. Happily, there is a better point of view.

The passage from $\mathcal{C}_d^{\mathbb{C}}$ to $\mathbb{C}_d^{\mathrm{Lor}}$ is already interesting when $d = 1$, i.e. 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.

The main point is to understand why a holomorphic representation of the category $\mathcal{C}_1^{\mathbb{C}}$ is just a 1-parameter contraction semigroup, where the parameter runs through the open half-plane $\mathbb{C}_+ = \{z \in \mathbb{C} : \mathrm{Re}(z) > 0\}$. Recall that we began this work with the relation of such semigroups to 1-parameter unitary groups.

Any complex metric on the interval $I = [0,1]$ can be pulled back from the holomorphic quadratic differential $\mathrm{d}z^2$ on $\mathbb{C}$ by means of a smooth embedding $f : I \to \mathbb{C}$ such that $f(0) = 0$ and $\mathrm{Re}\ f'(t) > 0$ for all $t \in I$. In fact the space $\mathrm{Emb}(I;\mathbb{C})$ of such embeddings is isomorphic to $\mathrm{Met}_{\mathbb{C}}(I)$ *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 $\mathcal{C}_1^{\mathbb{C}}$, 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 $f(1) \in \mathbb{C}_+$, where $\mathbb{C}_+$ denotes the open right half-plane of $\mathbb{C}$. We see that as follows. First, Z_f does not change if f is replaced by $\tilde{f} = f \circ \phi$ where ϕ 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 $\mathrm{Emb}(I;\mathbb{C})$ whose tangent vector at each point is the action of an element of the Lie algebra $\mathrm{Vect}(\mathring{I})$ of compactly supported vector fields on the interior of I, and hence — 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 $\mathrm{Vect}_{\mathbb{C}}(\mathring{I})$. But if $f, \tilde{f} \in \mathrm{Emb}(I;\mathbb{C})$ 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 $\mathrm{Vect}_{\mathbb{C}}(\mathring{I})$.

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, the vector space $\check{E}_P$ is a pre-Hilbert space because the object P of $\mathcal{C}_1^{\mathrm{gh}}$ is time-symmetric, and the unitarity condition tells us that $u(\bar{z})$ is the hermitian transpose of $u(z)$.

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

$$\begin{aligned} u(-\mathrm{i}T)u(\mathrm{i}T)\xi &= u(-\mathrm{i}T)u(\varepsilon + \mathrm{i}T)\eta \\ &= u(-\mathrm{i}T)u(\varepsilon/2)u(\varepsilon/2 + \mathrm{i}T)\eta \\ &= u(\varepsilon/2 - \mathrm{i}T)u(\varepsilon/2 + \mathrm{i}T)\eta \\ &= u(\varepsilon)\eta \;=\; \xi. \end{aligned}$$

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 $M_{\mathbb{C}}$, and M has an allowable complex metric induced from a holomorphic symmetric form g on the tangent bundle $TM_{\mathbb{C}}$, then the linear map Z_M does not change when M is moved around smoothly inside*

$M_{\mathbb{C}}$ (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 $M_{\mathbb{C}}$ is given by a complex vector field on M, while Z_M depends holomorphically on M and, being invariant under the action of $\mathrm{Diff}(M \text{ rel } \partial M)$, 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 $M_{\mathbb{C}}$, and the only natural way to ensure this is to pass from the smooth Lorentzian category $\mathcal{C}_d^{\mathrm{Lor}}$ to the corresponding *real-analytic* cobordism category $\mathcal{C}_d^{\mathrm{Lor},\omega}$, where both the manifolds and their metrics are assumed real-analytic. Inside this category there is the subcategory $\mathcal{C}_d^{\mathrm{gh},\omega}$ of globally hyperbolic cobordisms: we shall also assume that the time-function $\tau : M \to \mathrm{i}[0,1]$ is real-analytic, though 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. But one could also think that 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 $\mathcal{C}_d^{\mathbb{C}}$ induces a functor from $\mathcal{C}_d^{\mathrm{gh},\omega}$ 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 Σ is a time-symmetric $(d-1)$-manifold germ then there is a Hilbert space E_Σ^{Hilb} with

$$\check{E}_\Sigma \quad \subset \quad E_\Sigma^{Hilb} \quad \subset \hat{E}_\Sigma,$$

and a real-analytic globally hyperbolic cobordism $\Sigma_0 \rightsquigarrow \Sigma_1$ between time-symmetric hypersurfaces induces a unitary isomorphism $E_{\Sigma_0}^{Hilb} \to E_{\Sigma_1}^{Hilb}$ which also maps $\check{E}_{\Sigma_0}$ to $\check{E}_{\Sigma_1}$ and $\hat{E}_{\Sigma_0}$ to $\hat{E}_{\Sigma_1}$.

Proof of 5.2 Given a real-analytic globally hyperbolic cobordism $M : \Sigma_0 \rightsquigarrow \Sigma_1$ 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 $M_{\mathbb{C}}$ of M to which both t and g can be extended holomorphically, and we can assume that $\tau = \mathrm{i}t : M_{\mathbb{C}} \to U \subset \mathbb{C}$ is a holomorphic fibre bundle over a neighbourhood U of the interval $\mathrm{i}[0,1]$. Furthermore, the isomorphism $\Sigma_0 \times [0,1] \to M$ extends to a holomorphic trivialization of the bundle $M_{\mathbb{C}} \to U$. For any smooth curve $f : [0,1] \to U$ such that $f(0) = 0$ and $\mathrm{Re}\ f'(s) > 0$ for $s \in [0,1]$ this gives us a totally real submanifold M_f of $M_{\mathbb{C}}$ sitting over the curve. We can use the morphism associated to the cobordism M_f in exactly the way we used Z_f in discussing the 1-dimensional case, to obtain a unitary operator Z_M associated to 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 $\Sigma_0 \times [0,1] \to M_{\mathbb{C}}$ by a real-analytic diffeomorphism of $\Sigma_0 \times [0,1]$.

Remark 5.3 We can apply the principle 5.1 to understand better how a theory defined on $\mathcal{C}_d^{\mathbb{C}}$ assigns a vector space E_Σ 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 $\mathrm{Re}\ f'(t) > 0$ for all t. In particular, for $\lambda \in \mathbb{C}_+$ we have the radial paths $f_\lambda \in \mathcal{J}$ for which $f_\lambda(t) = \lambda t$. But recall that $\hat{E}_f$ is the inverse-limit of a sequence of spaces associated to 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 $\{t_k\}$, i.e. $t_k > t'_k > t_{k+1}$. We can choose a path $f \in \mathcal{J}$ which lies in the sector bounded by the rays f_λ and $f_{\lambda'}$ and coincides with them alternately in the neighbourhoods of the points λt_k and $\lambda' t'_k$. This f gives us a family of cobordisms from the germ at $\lambda' t'_k$ to the germ at λt_k, and from the germ at λt_{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 $\hat{E}_{\bar{f}}$ is the complex-

conjugate space to $\hat{E}_f$, so that reversing the complex time-direction conjugates the identification of $\hat{E}_\Sigma$ with the Euclidean choice $\hat{E}_{f_1}$. 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 $\check{E}_\Sigma$.

Field operators

Finally, we come to the Wick rotation of field operators, though our account will be sketchy. The first step is to understand how the vector space $\mathcal{O}_x$ 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 $\mathcal{O}_x$ to a germ at x of a complex metric on a manifold containing x: it is the fibre of a bundle on the space $\mathrm{Met}_{\mathbb{C}}(\hat{x})$ of such germs. If we embed a Lorentzian M in a complexification $M_{\mathbb{C}}$ there will be a holomorphic exponential map from a neighbourhood of 0 in the complexified tangent space $T_x^{\mathbb{C}} = T_x M \otimes \mathbb{C}$ to $M_{\mathbb{C}}$. Inside $T_x^{\mathbb{C}}$ we can consider the d-dimensional real vector subspaces V on which the metric induced from the complex bilinear form of $T_x^{\mathbb{C}}$ is allowable. We saw in (2.6) that these V form a contractible open subset $\mathcal{U}$ of the real Grassmannian $\mathrm{Gr}_d(T_x^{\mathbb{C}})$. Exponentiating V will give us a germ of a d-manifold with a complex metric, and hence a map $\mathcal{U} \to \mathrm{Met}_{\mathbb{C}}(\hat{x})$. 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 (5.3), we see to be trivial. Identifying its fibres gives us our definition of $\mathcal{O}_x$ for Lorentzian M.

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

$$\mathcal{O}_{x_1} \otimes \ldots \otimes \mathcal{O}_{x_k} \quad \to \quad \mathrm{Hom}(\check{E}_\Sigma; \hat{E}_\Sigma)$$

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. For the spaces $\check{E}_{\Sigma_t} \subset E^{Hilb}_{\Sigma_t} \subset E_{\Sigma_t}$ for all times $t_0 \leq t \leq t_1$ can be identified with those at time t_0 by the unitary propagation $U_{t,t'}$ from time t to a later time t' to get a single rigged Hilbert space $\check{E} \subset E^{Hilb} \subset E$, and we can define an unbounded operator

$$\tilde{\psi} \quad = \quad Z^{-1}_{t_0,t} \circ \psi \circ Z_{t_0,t} : \check{E} \to \hat{E}$$

for any $\psi \in \mathcal{O}_x$ 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 ψ, ψ' 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 Σ_t which contains them both, and we see that the composed operator, in either order, is $Z^{-1}_{t_0,t} \circ (\psi \otimes \psi') \circ Z_{t_0,t}$.

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 $\mathbb{M}$ of dimension d. There, the vacuum expectation values, initially defined as distributions or other generalized functions on k-fold products $\mathbb{M} \times \ldots \times \mathbb{M}$, are regarded as boundary values of holomorphic functions defined in an open domain in the complexified space $\mathbb{M}_{\mathbb{C}} \times \ldots \times \mathbb{M}_{\mathbb{C}}$. The Wightman axioms imply that the vacuum expectation values are the boundary values of holomorphic functions defined in the domain $\mathcal{U}_k$ known as the 'permuted extended tube', whose definition was given in Section 2.

If $k > 2$, however, $\mathcal{U}_k$ cannot be the largest domain where the expectation values are holomorphic, for it is known not to be holomorphically convex. It is an old problem to describe the maximal possible domain, or even the holomorphically convex hull of $\mathcal{U}_k$.

The ideas of this paper suggest that a candidate for the maximal domain is the simply-connected covering of the open subset $\mathcal{U}^*_k$ of all k-tuples $\mathbf{x} = \{x_1, \ldots, x_k\}$ of distinct points in $\mathbb{M}_{\mathbb{C}}$ which lie on a totally real submanifold M (of dimension d) of $\mathbb{M}_{\mathbb{C}}$ for which the induced complex metric belongs to $\mathrm{Met}_{\mathbb{C}}(M)$ and is constant outside of a compact region. By the results of Section 2 we know that any such submanifold M must project diffeomorphically to $\mathbb{E}$, the standard Euclidean subspace of $\mathbb{M}_{\mathbb{C}}$ obtained by rotating the time-axis of $\mathbb{M}$ by $\sqrt{-1}$.

To motivate the conjecture 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 a d-manifold M with an allowable complex metric which outside a compact subset is isomorphic to the part of an allowable totally real affine d-plane in $\mathbb{M}_{\mathbb{C}}$ cut off between two parallel $(d-1)$-planes.

A field theory 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 of the pair $(M, \mathbf{x})$. In particular we can restrict to the space $\mathcal{F}$ of those $(M, \mathbf{x})$ for which, as in the conjecture, M is embedded in $\mathbb{M}_{\mathbb{C}}$ and is flat in the neighbourhood of each point of $\mathbf{x}$. Then $\mathcal{F}$ is an infinite-dimensional bundle over the open subset $\mathcal{U}_k^*$ of $(\mathbb{M}_{\mathbb{C}})^k$, and the complex Poincaré group acts on it by bundle maps. The expectation values will be a holomorphic map

$$\mathcal{E} : \mathcal{F} \ \rightarrow \ \mathrm{Hom}(\mathcal{O}^{\otimes k}; \mathbb{C}),$$

where $\mathcal{O}$ is the space of observables at a point of $\mathbb{M}$.

Our much-used Principle 5.1 tells us that the value of the function $\mathcal{E}$ does not change if, while holding the marked points $\mathbf{x}$ fixed in $\mathbb{M}_{\mathbb{C}}$, we move M smoothly in the allowable class. So in fact we have a holomorphic function on a *covering space* of the open domain of possible marked points $\mathbf{x}$, where the sheets of the covering over a set of points $\mathbf{x}$ correspond to the isotopy classes of allowable manifolds M containing $\mathbf{x}$, i.e. to the connected components of the fibre $\mathcal{F}_{\mathbf{x}}$ of $\mathcal{F} \rightarrow \mathcal{U}_k^*$.

The conjecture is certainly correct in the simple case $k = 2$, for then $\mathcal{U}_2^*$ is just the same as $\mathcal{U}_2$: any two points (x_1, x_2) for which $||x_1 - x_2||^2$ is not real and negative lie, by the results of Section 2, on an allowable real affine linear subspace of $\mathbb{M}_{\mathbb{C}} \times \mathbb{M}_{\mathbb{C}}$.

References

[BF] Brunetti, R., and K. Fredenhagen, *Quantum field theory on curved backgrounds.* Lecture Notes in Physics **786**, 129 – 155, Springer 2009.

[C] Costello, Kevin, *Renormalization and effective field theory.* Math. Surveys and Monographs **170**, Amer.Math. Soc. 2011.

[H] Howe, Roger, *The Oscillator Semigroup.* Proc. Symp. Pure Math. **48**, 1196 – 1200, Amer. Math. Soc. 1988.

[Ka] Kazhdan, David, *Introduction to QFT.* In *Quantum fields and strings: a course for mathematicians.* Vol 1 (Princeton 1996/1997) 377–418. Amer. Math. Soc. 1999.

[Ke] Kelnhofer, Gerald, *Functional integration and gauge ambiguities in generalized abelian gauge theories.* J. Geom. Physics **59** (2009) 1017–1035. (arXiv:0711.4085 [hep-th])

[N] Neretin, Yu. A., *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).

[PS] Pressley, A., and G. Segal, *Loop Groups.* Oxford U.P. 1986.

[Se1] Segal, Graeme, *The definition of conformal field theory.* In: *Differential geometrical methods in theoretical physics. (Como 1987)* NATO Adv. Sci. Inst. Ser C, Math Phys. Sci. **250**, 165 – 171, Kluwer 1988.

[Se2] Segal, Graeme, *The definition of conformal field theory.* In: *Topology, Geometry, and Conformal Field Theory*, ed. U. Tillmann, London Math. Soc. Lecture Notes **308** (2004), 421 – 577.

[Sz] Szabo, R., *Quantization of higher abelian gauge theory in generalized differential cohomology.* In: *Proc. 7th Internat. Conf. on Math. Methods in Physics* (ICMP2012) (arXiv:1209.2530 [hep-th]).

[SW] Streater, R. F., and A. S. Wightman, *PCT, Spin and Statistics, and all that.* Princeton U.P. 2000. (1st edn Benjamin 1964)