The quantum theory of free automorphic fields



1980 J. Phys. A: Math. Gen. 13 2179

(http://iopscience.iop.org/0305-4470/13/6/039)

J. Phys. A: Math. Gen. **13** (1980) 2179–2203. Printed in Great Britain

# The quantum theory of free automorphic fields

Richard Banach†

Theoretical Physics Department, Schuster Laboratory, University of Manchester, Manchester, England

Received 21 August 1979, in final form 28 November 1979

**Abstract.** Heuristic spectral theory is developed for a symmetric operator on the universal covering space of a multiply connected static spacetime and is used to construct the quantum field theory of a multiplet of scalar fields in the customary sum-over-modes fashion. The non-local symmetries necessary to the theory are explicitly constructed, as are the projections on the field operators. The non-existence of a standard charge conjugation for certain types of representation is noted. Gauge transformations are used to give a simple and complete classification of automorphic field theories. The relationship between the unprojected and projected field algebras is clarified, and the implications for Fock space (vacuum degeneracy, etc) are discussed—earlier work being criticised. The analogy to black hole physics is pointed out, and the possible role of the Reeh–Schlieder theorems is speculated upon.

## 1. Introduction

It is well known that non-trivial topology in a spacetime produces interesting effects in quantised fields defined on it. Casimir stresses are one familiar example (DeWitt *et al* 1978, Dowker and Banach 1978, Banach and Dowker 1979b). The occurrence of topologically inequivalent field configurations—twisted fields (Isham 1978, Avis and Isham 1978) and automorphic fields (Banach and Dowker 1979a)—is another. More recently, interacting fields have produced novel and unexpected features such as spontaneous mass generation (Ford 1979a) and non-causal photon propagation in QED (Ford 1979b).

This paper is devoted to the quantum theory of free automorphic fields. Now the *sine qua non* of automorphic field theory is the fact that one studies fields on a non-simply connected spacetime. The non-simple connectivity of the space associates to it a simply connected universal covering space, and automorphic field theory aims to understand field theory on the former in terms of field theory on the latter. The fact that the two spaces are locally indistinguishable (from a purely geometrical point of view) means that one can envisage an observer, confined to a small region of the space, assuming that the physical space actually *is* the universal covering space and attempting to understand his physics in these terms. This lends a purely physical motive to the study.

We particularise to the following: (*a*) Spacetime topology is $T \otimes M$. $T$ is time, $M$ is multiply connected, $T \otimes \tilde{M}$ is the universal cover, where $\tilde{M}$ is the universal cover of $M$.

† Research supported by an SRC Research Studentship.

  

$\pi:\tilde{M}\to M$ is the covering projection, and we have that $\Gamma := \pi_1(M)$ is discrete and acts effectively and transitively on $\pi^{-1}(x)$ for $x\in M$. (*b*) We take the metric $ds^2 = dt^2 - d\sigma^2$ to be compatible with the topology in the obvious way; in particular we assume that $\Gamma$ induces isometries. For simplicity we take $d\sigma^2$ to be time-independent. (*c*) We study a multiplet of complex scalar fields possessing a $U(n)$ internal symmetry restricted to rigid gauge actions for simplicity.

The geometrical framework outlined above encompasses all the situations studied in practice so far. Most of the work has centred on the two-point function $\langle 0|\phi(x)\phi^{\dagger}(x')|0\rangle$ (time-ordered where appropriate) and on the physical quantities derivable from it (e.g. $\langle T_{\mu\nu}\rangle$). Calculations proceed either by going straight to mode sum expansions of the quantities of interest or by use of the method of images (effectively an automorphic projection) on the two-point function. The two methods must agree as has been shown (Banach and Dowker 1979a), but the image sum method in particular, using as it does the full field theory of the covering space to write down the covering space two-point function, obscures the field theoretic analogue of image summation. Thus we would like to pull back the topological issues in the theory into the field theory itself, rather than relegate them to the status of mathematical afterthoughts to be applied to the two-point functions—which are classical singular functions and thus easily dealt with. The fact that one *can* relegate the topological issues in this fashion owes its existence to the linearity of the differential equations that the two-point functions satisfy and the uniqueness theorems for them, and not to any deep understanding of the effects of topology on quantum field theory. The further examination of the 'pulling back' is therefore a pertinent matter.

In § 2 we develop the spectral theory of a symmetric operator acting on functions on $\tilde{M}$ in a more complete form than was given previously (Banach and Dowker 1979a). Not all of the details given are necessary to the development of the field theory (some are demoted to the Appendix), but they do nevertheless answer some very obvious questions left open by the earlier work. In § 3 the results of § 2 are used to quantise the field on $\tilde{M}$. Non-local symmetries and gauge symmetries are constructed in the operator algebra and used to project the quantised field. The Wightman functions are considered, and charge conjugation is also briefly discussed. The results of § 3 are used to give a simple classification scheme for automorphic field theories in § 4. In contrast to local gauge theories, each sector in the theory defines an acceptable field theory without the need to go to the dual space of the classifying set.

In § 5 the results of § 3 are used to discuss the operator algebra as a whole and the corresponding Fock space. We see that, in fact, the terminology of 'projecting' the field algebra is misleading and that a subalgebra and partial states language is more appropriate. This of course is analogous to the late-times situation in gravitational collapse and other physical situations where a horizon is present. One can only measure aspects of the state depending on the relevant subalgebra, and the precise amount of freedom in the rest of the state depends on the physical situation in hand. In automorphic field theory we see that the amount of freedom is maximal.

A certain amount of speculation on the role of the Reeh–Schlieder theorems is included.

*Notational convention*: In the following, it will frequently be necessary to employ subscripts, both to indicate a function name and to label its components. Where this happens, the subscripts are separated by a comma: thus $f_{i,\mu}$ is the $\mu$th component of $f_i$. Commas therefore have no connection with derivatives in this paper.

## 2. Spectral theory

The object of interest to us is the Hilbert space $L^2_n(\tilde{M})$ of (square integrable) complex $n$-tuple-valued functions on $\tilde{M}$ with the usual inner product

$$(f, g) = \int_{\tilde{M}} f^{\dagger}_{\mu} g_{\mu} \, \mathrm{d}\mu, \tag{1}$$

where $\mathrm{d}\mu$ is the $\Gamma$-invariant volume element of the metric $\mathrm{d}\sigma^2$ on $\tilde{M}$.

An automorphic function on $\tilde{M}$ is one satisfying

$$f_{\mu}(\gamma x) = a_{\mu\nu}(\gamma) f_{\nu}(x), \qquad \forall x, \tag{2}$$

where $a(\Gamma)$ is an $n$-dimensional representation of $\Gamma$ which we take to be unitary in this paper. For an arbitrary $f \in L^2_n(\tilde{M})$ we set $\overleftarrow{f}^{a}_{\mu}(x)$ to be given by (for $|\Gamma| < \infty$)

$$\overleftarrow{f}^{a}_{\mu}(x) = \frac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} a_{\mu\nu}(\gamma^{-1}) f_{\nu}(\gamma x), \tag{3}$$

which satisfies (2). Equation (3) defines a linear subspace of $L^2_n(\tilde{M})$, $L^2_n(a)$ in which all the functions satisfy $f = \overleftarrow{f}^{a}$. Thus $L^2_n(\tilde{M}) = L^2_n(a) \oplus L^2_n(a)^{\perp}$, and the purpose of this section is to clarify how a symmetric linear operator in $L^2_n(\tilde{M})$ respects this split.

Let $K$ be a linear operator (bounded for simplicity), and $P$ the projection operator $f \to \overleftarrow{f}^{a}$. Then we have the following textbook result (proof omitted):

*Lemma 1.* $KP = PK \Leftrightarrow K(L^2_n(a)) \subset L^2_n(a); K(L^2_n(a)^{\perp}) = 0.$

If $K$ is a Hilbert–Schmidt operator with kernel $K_{\mu\nu}(x, x')$, we have

$$KP = PK \Leftrightarrow \frac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} a(\gamma^{-1}) K(\gamma x, x') = \frac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} K(x, \gamma x') a(\gamma). \tag{4}$$

The right-hand side of (4) can be written $\overleftarrow{K} = \overrightarrow{K}$ in the notation of Banach and Dowker (1979a).

A sufficient set of conditions for the satisfaction of (4) (and one that can usually be satisfied in practice) is

$$a(\gamma) K(x, x') = K(x, x') a(\gamma), \qquad \forall \gamma \in \Gamma, \tag{5}$$

$$K(\gamma x, x') = K(x, \gamma^{-1} x'). \tag{6}$$

Now suppose that $K$ is symmetric—in the Hilbert–Schmidt case this means

$$K_{\mu\nu}(x, x') = K^{*}_{\nu\mu}(x', x), \tag{7}$$

and suppose it has a relatively simple spectral theory with a purely discrete real spectrum and eigenfunction expansion

$$K_{\mu\nu}(x, x') = \sum_{\lambda_i} \lambda_i f_{i,\mu}(x) f^{*}_{i,\nu}(x'), \tag{8}$$

and each eigenvalue has finite degeneracy. Then lemma 1 says (and an explicit proof in Banach and Dowker (1979a) confirms) that (in an appropriate basis) each $f_i$ lies entirely either in $L^2_n(a)$ or in $L^2_n(a)^{\perp}$ if and only if (4) is true.

Since the behaviour of $L_n^2(a)$ under the action of $\Gamma$ is well understood—it just consists of functions satisfying (2)—one naturally asks that the behaviour of $L_n^2(a)^\perp$ is under $\Gamma$. It must clearly contain any functions automorphic by representations $b(\Gamma)$ inequivalent to $a(\Gamma)$, since

$$\overleftarrow{f}^{a}_{\mu}(x)=\frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}a_{\mu\nu}(\gamma^{-1})f_\nu(\gamma x)=\frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}a_{\mu\nu}(\gamma^{-1})b_{\nu\lambda}(\gamma)f_\lambda(x)=0 \tag{9}$$

if $a(\Gamma)$ and $b(\Gamma)$ are inequivalent, by the orthogonality relations. On the other hand suppose that $b(\Gamma)=Ra(\Gamma)R^{-1}$, then, if $f$ is automorphic by $b(\Gamma)$, we have

$$\overleftarrow{f}^{a}=\frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}a_{\mu\nu}(\gamma^{-1})f_\nu(\gamma x)=\frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}a_{\mu\nu}(\gamma^{-1})R_{\nu\lambda}a_{\lambda\rho}(\gamma)R^{-1}_{\rho\sigma}f_\sigma(x)=(\operatorname{Tr}R)R^{-1}f, \tag{10}$$

which can also be zero if $\operatorname{Tr}R=0$. Lemma 1 then tells us that, for any $f_i$ in the expansion (8) which is automorphic by some $b(\Gamma)=Ra(\Gamma)R^{-1}$, we must have $\operatorname{Tr}R=0$. There remains the possibility that there are $f_i$ in (8) not automorphic by any representation of $\Gamma$ whatsoever. These we can call anarchic modes.

In general then, the structure of $L_n^2(a)^\perp$ is quite complicated. However, in the case when $K$ is one-dimensional, i.e.

$$K_{\mu\nu}(x,x')=k(x,x')\otimes\mathbb{1}^{(n)}_{\mu\nu}\equiv K^{(n)}, \tag{11}$$

where $\mathbb{1}^{(n)}_{\mu\nu}$ is the $(n\times n)$ unit matrix and $k(x,x')$ satisfies (6), we can give a complete decomposition of the eigenfunction expansion into automorphic and non-automorphic modes.

Firstly, we know that (6) applies by assumption, and (5) is obviously true for any $a(\Gamma)$; hence lemma 1 applies.

Now $K^{(1)}$ has eigenfunctions which are an orthonormal basis in $L_1^2(\tilde{M})$; call them $e_i(x)$ and suppose they are all automorphic by some (necessarily one-dimensional) representation of $\Gamma$, i.e.

$$e_i(\gamma x)=b_i(\gamma)e_i(x). \tag{12}$$

In addition, suppose $f$ is a function in $L_n^2(\tilde{M})$ automorphic by some irreducible $a(\Gamma)$. Then its components each lie in $L_1^2(\tilde{M})$ and so can be expanded in terms of the $e_i$; thus

$$f_\mu=\beta_\mu{}^{i}e_i. \tag{13}$$

Expressing the automorphic character of $f$ in two ways, we obtain

$$f_\mu(\gamma x)=a_{\mu\nu}(\gamma)\beta_\nu{}^{i}e_i(x)=\beta_\mu{}^{i}b_i(\gamma)e_i(x). \tag{14}$$

Thus $\beta$ intertwines $a(\Gamma)$ and the one-dimensional $b_i$ and so is necessarily zero. What we have shown is a special case of the following (the proof of which is essentially identical to the above):

*Lemma 2.* A function automorphic by an irreducible $a(\Gamma)\subset U(n)$ cannot have any proper subset of $m$ of its $n$ components expanded in terms of functions $\{e_i\}$ automorphic by $\{b_i(\Gamma)\subset U(m)\}$. Thus the only candidates for expanding components of higher-dimensional automorphic functions in terms of lower-dimensional modes are the lower-dimensional anarchic modes.

We now show how this takes place.

Let $f_i$ be any mode of $K^{(1)}$ whatsoever. Then the standard theory of symmetry coupled with (6), which says that $K^{(1)}$ commutes with the action of $\Gamma$, implies that the eigenspace of $f_i$ carries at least one irreducible representation of $\Gamma$. Restricting attention to the one relevant to $f_i$, we have a representation of $\Gamma$ in the unitary operators in the eigenspace: $\Gamma \to \hat{\Gamma}$, where $\hat{\gamma}f(x) = f(\gamma^{-1}x)$. We can write this action as a matrix representation of $\Gamma$ in the eigenspace

$$\hat{\gamma}f_i = u_{ji}(\gamma)f_j. \tag{15}$$

Consider now the alternative action of $\Gamma$ in the eigenspace, namely $\Gamma \to \hat{\hat{\Gamma}}$, where $\hat{\hat{\gamma}}f = f(\gamma x)$. This is an antirepresentation of $\Gamma$ in the unitary operators in the eigenspace which we can antirepresent in terms of unitary matrices

$$\hat{\hat{\gamma}}f_i = w_{ij}(\gamma)f_j. \tag{16}$$

The composition of these two antirepresentations means that $w(\Gamma)$ is a representation of $\Gamma$, and in fact it is obvious that the representations $u(\Gamma)$ and $w(\Gamma)$ are contragradient to each other.

So (16) says that

$$f_i(\gamma x) = w_{ij}(\gamma)f_j(x), \tag{17}$$

which is just (2) except that $i$ and $j$ are not internal indices. If we now turn to $K^{(n)}$ (where $n = \dim w(\Gamma)$), then the function $\Phi$, where $\Phi_{,j} = f_j$, is an eigenfunction, $j$ is now an internal index, and $\Phi$ is automorphic by $w(\Gamma)$.

This construction canonically associates a representation of $\Gamma$ to each mode of $K^{(1)}$, and automorphic functions automorphic by a given representation (up to equivalence) can have components only within the subspace of $L_1^2(\tilde{M})$ spanned by functions to which this is the associated representation. We have shown:

*Lemma 3.* For each eigenfunction $f$ of $K^{(1)}$, we can find an $n$, an $n$-dimensional irreducible representation of $\Gamma$, $a(\Gamma)$, and an eigenfunction $\Phi$ of $K^{(n)}$, such that $f$ is the first component of $\Phi$ and $\Phi$ is automorphic by $a(\Gamma)$.

One can repeat the above construction with anarchic modes of $K^{(m)}$, the result being an $mn$-dimensional representation of $\Gamma$, since the $i$ index above will be tensored with the already existing internal index. We see that the representation is in fact $w(\Gamma)\otimes \mathbb{1}^{(m)}$, and so is simply reduced to $m$ copies of $w(\Gamma)$, each component of the original $f$ generating its own automorphic piece. The uniqueness of the association of $f$ with the first component of $\Phi$ in lemma 3 will thus be lost. Nevertheless, we see that the anarchic designation of an eigenfunction is a purely dimension-dependent thing. If we embed the anarchic eigenfunction in $L_n^2(\tilde{M})$ for sufficiently high $n$, we can build it up into an automorphic mode.

We can confirm that the above mechanisms indeed give us the 'right' automorphic mode as follows. Let $\{f_{i,\mu}\}$ be a set of modes transforming under $\hat{\hat{\Gamma}}$ by some irreducible representation $w(\Gamma)$ as in (17). We project out the automorphic part of these modes for some irreducible $a(\Gamma)$; thus† (we assume enough zero-valued components have been

† Compare the treatment here with the treatment of the analogous quantity in Banach and Dowker (1979a), where the contragradient representation $u(\Gamma)$ is used.

added to $\{f_{i,\mu}\}$ for this to make sense)

$$\overset{\leftarrow a}{f_{i,\mu}} = \frac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} a_{\mu\nu}(\gamma^{-1}) f_{i,\nu}(\gamma x) = \frac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} a_{\mu\nu}(\gamma^{-1}) w_{ij}(\gamma) f_{j,\nu}(x). \tag{18}$$

Now unless $a(\Gamma) \equiv w(\Gamma)$, the coefficient of $f_{j,\nu}$ is zero. This shows the uniqueness of the automorphic representation associated to $f_{i,\mu}$. If $a \equiv w$, the orthogonality relations say that

$$\overset{\leftarrow a}{f_{i,\mu}} = (1/n) f_{\mu,i} \qquad (n = \dim a = \dim w). \tag{19}$$

Thus the set $\{f_i\}$ has been reflected along its diagonal, clearly making it automorphic. If we do the same thing again with the roles of the two indices reversed, i.e. we find that part of $\overset{\leftarrow a}{f_{i,\mu}}$ which transforms under $\hat{\hat{\Gamma}}$ as $w(\Gamma)$, we find $n^{-2} f_{i,\mu}$. Thus we get back to our starting point, but having lost out on the norm of $\{f_i\}$ by $n^2$. This illustrates that the two projections involved are not compatible in the sense that a set $\{f_i\}$ cannot simultaneously display correctly automorphic behaviour and the right $\hat{\hat{\Gamma}}$ transformations. Thus there must be a change of basis involved in the change between the two, and the final task of this section will be to exhibit this change of basis.

Suppose then we are given an $n$-fold degenerate eigenspace of $K^{(1)}$ spanned by $\{f_i\}$ which behave under $\hat{\hat{\Gamma}}$ according to some irreducible $a(\Gamma)$ as in (17). The above results tell us that we must go to $K^{(n)}$ to find automorphic behaviour. The basis $\{f_i\}$ induces an $n^2$-fold degenerate basis $\{f_{i\alpha,\mu}\}$ of the corresponding eigenspace of $K^{(n)}$, where we set

$$f_{i\alpha,\mu} = \delta_{\alpha\mu} f_i. \tag{20}$$

We next split up the basis into $n$ subsets of $n$, $\Sigma_i$, where

$$\Sigma_i = \{f_{j\{j+i(\mathrm{mod}\ n)\},\mu}\}. \tag{21}$$

Thus, for example, $\Sigma_0$ is the 'diagonal' subset $\{f_{ii,\mu}\}$, $\Sigma_1$ the 'subdiagonal' (mod $n$), etc. The modes in $\Sigma_i$ are obtainable from those in $\Sigma_0$ by applying the gauge transformation $S_i$, where

$$S_{i,\mu\nu} = \delta_{\mu\{\nu+i(\mathrm{mod}\ n)\}}. \tag{22}$$

Therefore the structure of $\Sigma_0$ implies that of $\Sigma_i$, and we restrict our attention to $\Sigma_0$.

Clearly the function $g$, where

$$g_{,\mu} = \frac{1}{\sqrt{n}} \sum_i f_{ii,\mu} \tag{23}$$

is in the span of $\Sigma_0$, $\langle\Sigma_0\rangle$, is of norm 1 and is automorphic by $a(\Gamma)$, since $g_{,\mu} = f_\mu$. Its orthogonal complement in $\langle\Sigma_0\rangle$ consists of functions of the form $h$, where

$$h = \sum_i \alpha_i f_{ii} \tag{24}$$

and

$$\sum_i \alpha_i = 0. \tag{25}$$

The $a(\Gamma)$-automorphic part of $h$ vanishes, since

$$\begin{aligned}\overleftarrow{h}^{a}_{\mu}(x) &= \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma} a_{\mu\nu}(\gamma^{-1})\sum_i \alpha_i f_{ii,\nu}(\gamma x)\\ &= \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma} a_{\mu\nu}(\gamma^{-1})\sum_{ij} \alpha_i a_{ij}(\gamma) S_{\{\nu-j(\text{mod } n)\},\nu\lambda} f_{jj,\lambda}(x)\\ &= \frac{1}{n}\delta_{\mu j}\delta_{i\nu}\sum_{ij}\alpha_i S_{\{\nu-j(\text{mod } n)\},\nu\lambda} f_{jj,\lambda}(x)\\ &= \frac{1}{n}\left[\sum_i \alpha_i S_{\{i-\mu(\text{mod } n)\},i\lambda}\right] f_{jj,\lambda}(x)\\ &= 0, \end{aligned} \tag{26}$$

since the expression in square brackets clearly vanishes. If we apply $S_k (k \neq 0)$ to either $g$ or $h$ and then project, we find by similar calculations that

$$\overleftarrow{S_k g}^{a} = \overleftarrow{S_k h}^{a} = 0. \tag{27}$$

Thus, since $g$ and all possible $h$'s span $\langle\Sigma_0\rangle$, we find that for $k \neq 0$

$$\overleftarrow{\langle\Sigma_k\rangle}^{a} = 0. \tag{28}$$

Therefore, of the $n^2$ independent modes in the eigenspace, only one survives. Automorphic projection turns out to be a singularly wasteful process.

What of the modes that are annihilated? They are not automorphic by $a(\Gamma)$, but we know that the representation associated with them is $a(\Gamma)$. What happens is a case of the remarks following lemma 3. The canonical construction puts the modes into an $(n^2)^2$-dimensional space, and a reduction is needed to regain pieces automorphic by $a(\Gamma)$. In this process the original modes are scrambled up and all come out looking like $g$ for an appropriate reducing operator.

This suggests that in some sense all the modes in the eigenspace are gauge equivalent to $g$. We now show that we can find a basis for the eigenspace for which this is actually the case.

*Lemma 4.* There exist $n$ diagonal unitary matrices $T_i$ in $U(n)$ (one of which, $T_0$, may be taken to be the identity), orthogonal in the trace inner product: $(U_1, U_2) = \text{Tr } U_1^\dagger U_2$.

This is proved in the Appendix, since the methods are rather inappropriate for the present discussion and would detract from our main goal.

Apply the $T_i$ to $g$. Since they are diagonal, they give linear combinations of the $f_{ii,\mu}$ and so map $g$ into $\langle\Sigma_0\rangle$. Orthogonality in the trace norm becomes orthogonality in $L^2_n(\tilde{M})$ for the $T_i g$, and so the $T_i g$ span $\langle\Sigma_0\rangle$. Now apply $S_k$ to the $T_i g$. It is trivial to observe that the $S_k T_i g$ span $\langle\Sigma_k\rangle$ in precisely the same way that the $T_i g$ span $\langle\Sigma_0\rangle$, and this gives us the basis we have been looking for.

Thus the $n^2$ eigenfunctions $R_{ij} g$, where

$$R_{ij} = S_i \circ T_j, \tag{29}$$

are linear combinations of the $\{f_{i\alpha,\mu}(x)\}$ and constitute a basis. Since they are all gauge relatives of $g$, they are all automorphic, $R_{ij}g$ being automorphic by $R_{ij}a(\Gamma)R^\dagger_{ij}$. It is trivial to confirm that, for $(i, j) \neq (0, 0)$, $\mathrm{Tr}\, R_{ij} = 0$, as must be true by the remarks following (28) and the remarks following (10).

We now have enough information at our disposal to be able to give a complete decomposition of the eigenfunction expansion (8) for any $K^{(n)}$ and any unitary representation $a(\Gamma)$.

Gauge symmetry will be seen to be a symmetry of the theory we want to construct in the next section, so we may, without loss of generality, take $a(\Gamma)$ to be in fully reduced form $a(\Gamma) = \oplus_i a_i(\Gamma)$, and we can then deal with each of the irreducible summands $a_i(\Gamma)$ separately. Let $a_1(\Gamma)$ be $m_1$-dimensional and occupy the upper left $m_1 \times m_1$ part of $a(\Gamma)$. Then it will only affect the first $m_1$ components of any $f \in L^2_n(\tilde{M})$, and so we can ignore the others.

We pass among the eigenspaces of $K^{(n)}$ and temporarily disregard any which do not transform under $\hat{\Gamma}$ as $a_1(\Gamma)$. An eigenspace belonging to $a_1(\Gamma)$ will contain $nm_1$ modes of the form $f_{i\alpha,\mu}$ (by analogy with (20)), and we fix on the $m_1^2$ of them for which the index $\alpha \leq m_1$, temporarily disregarding the others. The $m_1^2$ we subject to the change of basis constructed above; we extract and keep the one automorphic mode and discard the $m_1^2 - 1$ others.

We now return to the temporarily disregarded modes and look for eigenspaces behaving like $a_2(\Gamma)$, repeating the procedure above for the $f_{i\alpha}$ having $m_1 < \alpha \leq m_2 + m_1$. Continuing in this manner, we eventually exhaust $a(\Gamma)$, and any modes remaining which have neither been kept nor positively discarded are now discarded. They will consist of (i) anarchic modes, (ii) modes automorphic by none of the $a_i(\Gamma)$ present in $a(\Gamma)$, and (iii) modes whose canonically associated representation is one of the $a_i(\Gamma)$, but whose non-zero components do not match correctly with the position of $a_i(\Gamma)$ in $a(\Gamma)$.

Thus the kept modes span $L^2_n(a)$ and the discarded modes span $L^2_n(a)^\perp$, achieving the desired decomposition.

## 3. Field quantisation

Consider the following operator on smooth complex functions on $\tilde{M}$:

$$L(x, x') = (-\nabla_i \nabla^i + m^2)\delta(x, x'). \tag{30}$$

It is symmetric and elliptic in the inner product given by

$$(f, g) = \int_{\tilde{M}} f^* g \, \mathrm{d}\mu, \tag{31}$$

and so, particularly when $\tilde{M}$ is compact, it is reasonable to assume that $L$ has a discrete spectrum with each eigenspace finite-dimensional:

$$LY_i = \lambda_i Y_i, \qquad \lambda_i > 0, \tag{32}$$

$$(Y_i, Y_j) = \delta_{ij}. \tag{33}$$

Of course, $L$ is not Hilbert–Schmidt†, although the notation in (30) is suggestive, and we will continue to treat it as if it were, since the only significant property of Hilbert–Schmidt operators of real use to us is the discrete nature and finite-dimensionality of the eigenspaces.

† Its inverse usually will be.

Extending to $T\otimes\tilde{M}$, we see that the $Q_{\omega i}$, where

$$Q_{\omega i}=[\exp(\mathrm{i}\omega t)/(2\pi)^{1/2}]Y_i, \tag{34}$$

are orthonormal in the obvious extension of (31) to $T\otimes\tilde{M}$,

$$(f, g)=\int \mathrm{d}t\int \mathrm{d}\mu\, f^*g, \tag{35}$$

$$(Q_{\omega i}, Q_{\sigma j})=\delta_{ij}\delta(\omega, \sigma), \tag{36}$$

and thus form a 'complete set'.

Considering now the wave operator

$$K(x, x')=(\partial_t^2+L)\delta(x, x'), \tag{37}$$

we see that the $Q_{\omega i}$ are in fact eigenfunctions, and those for which

$$\omega^2=\lambda_i \tag{38}$$

are solutions to the wave equation. Setting

$$f_i=[\exp(-\mathrm{i}\omega_{\lambda_i}t)/(2\omega_{\lambda_i})^{1/2}]Y_i=(2\pi/2\omega_{\lambda_i})^{1/2}Q_{-\omega_{\lambda_i}i}, \tag{39}$$

where

$$\omega_{\lambda_i}=+\sqrt{\lambda_i}, \tag{40}$$

we see that

$$\langle f_i, f_j\rangle=\delta_{ij}, \qquad \langle f_i, f_j^*\rangle=0, \qquad \langle f_i^*, f_j^*\rangle=-\delta_{ij}, \tag{41}$$

where

$$\langle g, h\rangle=\mathrm{i}\int_{\tilde{M}} \mathrm{d}\mu[g^*(\partial_t h)-(\partial_t g^*)h] \tag{42}$$

is the usual spacelike-hypersurface-independent inner product for solutions of the wave equation particularised to the $t=$ constant surfaces for simplicity. The completeness of the $\{Q_{\omega i}\}$ implies that the $\{f_i, f_j^*\}$ are complete in the space of solutions to

$$K\phi=0 \tag{43}$$

with any solution satisfying

$$\phi(x)=\langle \mathrm{i}G(x, x'), \phi(x')\rangle, \tag{44}$$

where

$$\mathrm{i}G(x, x')=\sum_i f_i^*(x)f_i(x')-f_i(x)f_i^*(x'). \tag{45}$$

In general we are interested in $n$-tuple-valued fields, to which end we define $f_{i\alpha}$ by

$$f_{i\alpha,\mu}=\delta_{\alpha\mu}f_i \tag{46}$$

as before. Extending (42) in the obvious way we find

$$\langle f_{i\alpha}, f_{j\beta}\rangle=\delta_{ij}\delta_{\alpha\beta}, \qquad \langle f_{i\alpha}, f_{j\beta}^*\rangle=0, \qquad \langle f_{i\alpha}^*, f_{j\beta}^*\rangle=-\delta_{ij}\delta_{\alpha\beta}. \tag{47}$$

$K$ given by (37) obviously extends to $K^{(n)}$, and solutions to $K^{(n)}\phi=0$ have the property

$$\phi(x)=\langle \mathrm{i}G^{(n)}(x, x'), \phi(x')\rangle, \tag{48}$$

with

$$\mathrm{i}G^{(n)}(x, x') = \sum_{i,\alpha} f^*_{i\alpha}(x) f_{i\alpha}(x') - f_{i\alpha}(x) f^*_{i\alpha}(x'). \tag{49}$$

Having completed these preliminaries, we are in a position to quantise the fields. Following the usual procedure, we set

$$\phi(x) = \sum_{i,\alpha} \boldsymbol{a}_{i\alpha} f_{i\alpha} + \boldsymbol{b}^\dagger_{i\alpha} f^*_{i\alpha}, \tag{50}$$

$$\phi^\dagger(x) = \sum_{i,\alpha} \boldsymbol{b}_{i\alpha} f_{i\alpha} + \boldsymbol{a}^\dagger_{i\alpha} f^*_{i\alpha} \tag{51}$$

and

$$[\boldsymbol{a}_{i\alpha}, \boldsymbol{a}^\dagger_{j\beta}] = \delta_{ij}\delta_{\alpha\beta} = [\boldsymbol{b}_{i\alpha}, \boldsymbol{b}^\dagger_{j\beta}], \tag{52}$$

$$0 = [\boldsymbol{a}_{i\alpha}, \boldsymbol{a}_{j\beta}] = [\boldsymbol{b}_{i\alpha}, \boldsymbol{b}_{j\beta}] = [\boldsymbol{a}_{i\alpha}, \boldsymbol{b}_{j\beta}] = [\boldsymbol{a}_{i\alpha}, \boldsymbol{b}^\dagger_{j\beta}] \ldots (+\textsc{hc}), \tag{53}$$

whence we have the covariant commutation relations

$$[\phi(x), \phi^\dagger(x')] = [\phi^\dagger(x), \phi(x')] = -\mathrm{i}G^{(n)}(x, x'), \tag{54}$$

$$[\phi(x), \phi(x')] = [\phi^\dagger(x), \phi^\dagger(x')] = 0 \tag{55}$$

which are equivalent (at this non-rigorous level) to the equal-time relations

$$[\partial_t\phi(x), \phi^\dagger(x')]|_{t=t'} = [\partial_t\phi^\dagger(x), \phi(x')]|_{t=t'} = -\mathrm{i}\partial_t G^{(n)}(x, x')|_{t=t'} = -\mathrm{i}\delta^{(n)}_{\tilde{M}}(x, x'), \tag{56}$$

where $\delta^{(n)}_{\tilde{M}}(x, x')$ is the delta function on $\tilde{M}$.

Our main goal in this section is to implement the projection (3) on the quantised fields (50), (51) with the algebra generated by the $\{\boldsymbol{a}^{(\dagger)}_{i\alpha}, \boldsymbol{b}^{(\dagger)}_{j\beta}\}$ which we call $\mathscr{R}$. First we need some preliminaries.

Let $G$ be a compact matrix group. Familiar representation theory tells us that we may regard $G$ as a subgroup of a unitary group without loss of generality. Now it is well known (according to Markus 1973) that the exponential map is onto for $GL(n, \mathbb{C})$, or, to put it another way, every invertible matrix has a (complex) logarithm. If $A \in GL(n, \mathbb{C})$ happens to be unitary, this logarithm must necessarily satisfy

$$[\log A]^\dagger = -\log A. \tag{57}$$

The point of the above is the following. $G$ may be (and usually is) a Lie group with its own Lie algebra and associated exponential map, but the exponential map is, in general, not onto (e.g. $O(n)$), and so arbitrary $g \in G$ may not have logarithms within the Lie algebra. By embedding in $U(n)$ we can, however, find a logarithm for any $g \in G$ in the bigger Lie algebra of $U(n)$. Since quantum theory (in any concrete representation) always takes place within a complex Hilbert space, this embedding is always available to us in practice. In the following, '$\log A$' is always to be understood in this generalised sense.

Now let $g_{ij} \in G$ and let $\{\boldsymbol{a}^{(\dagger)}_i\}$ be a set of creation/annihilation operators satisfying the usual CCR. Then the above discussion allows us to formally write

$$g_{ij}\boldsymbol{a}_j = U(g)\boldsymbol{a}_i U^\dagger(g), \tag{58}$$

where

$$U(g) = \exp(-\boldsymbol{a}^\dagger_j [\log g]_{jk} \boldsymbol{a}_k). \tag{59}$$

By virtue of (57), $U(g)$ is a unitary operator in the field algebra and provides a unitary implementation of the automorphism $\boldsymbol{a}_i \to g_{ij}\boldsymbol{a}_j$ of the algebra.

We have two classes of such automorphisms in mind for automorphic field theory: the isometries $\phi(x) \to \phi(\gamma x)$ and the gauge transformations $\phi(x) \to g\phi(x)$. For the isometries, we assumed that $L$ given by (30) is such that we can decompose its action into that on finite-dimensional eigenspaces. This split is clearly preserved when we pass to the solution space of $K^{(n)}$. Thus the solution space consists of finite-dimensional subspaces $\{\langle\omega_i\rangle\}$ corresponding to specific frequencies. Pick one, $\langle\omega_i\rangle$ say. It is spanned by certain of the $f_{i\alpha}$, upon which we fix our attention. Because $\Gamma$ induces isometries, $L$ formally satisfies (6) and therefore so does $K^{(n)}$, and therefore $\langle\omega_i\rangle$ carries at least one irreducible representation of $\Gamma$ according to the action $\hat{\hat{\Gamma}}$; thus

$$f_{i\alpha}(\gamma x) = w_{ij}^{\langle\omega_i\rangle}(\gamma) f_{j\alpha}(x), \tag{60}$$

where $w_{ij}^{\langle\omega_i\rangle}(\Gamma)$ is a unitary representation of $\Gamma$ as in (17). Looking at the expression for $\phi(x)$ in (50) we see that (60) is equivalent to

$$\boldsymbol{a}_{k\alpha} \to \boldsymbol{a}_{j\alpha} w_{jk}^{\langle\omega_i\rangle}(\gamma) = w_{kl}^{\langle\omega_i\rangle\mathrm{T}} \boldsymbol{a}_{l\alpha}. \tag{61}$$

Thus

$$T^{\langle\omega_i\rangle\alpha}(\gamma) = \exp(-\boldsymbol{a}_{i\alpha}^{\dagger}[\log w^{\langle\omega_i\rangle\mathrm{T}}]_{ij}\boldsymbol{a}_{j\alpha})$$

implements $\hat{\hat{\Gamma}}$ in $\langle\omega_i\rangle$ and is clearly an antirepresentation of $\Gamma$. Taking adjoints, we find for $\langle-\omega_i\rangle = \langle\{f_{i\alpha}^*\}\rangle$ that (61) corresponds to $\boldsymbol{a}_{k\alpha}^{\dagger} \to w_{kl}^{\langle\omega_i\rangle *}\boldsymbol{a}_{l\alpha}^{\dagger} \equiv w_{kl}^{\langle-\omega_i\rangle\mathrm{T}}\boldsymbol{a}_{l\alpha}^{\dagger}$ and a corresponding definition of $T^{\langle-\omega_i\rangle\alpha}(\gamma)$ which must coincide with that of $T^{\langle\omega_i\rangle\alpha}(\gamma)$ by taking the adjoint of (58). The $\{\boldsymbol{b}_{i\alpha}\}$ are treated analogously.

Taking products over the $\langle\pm\omega_i\rangle\alpha$ we can implement the isometries thus:

$$T(\gamma) = \prod_{\{\pm\omega_i,\alpha\}} T^{\langle\omega_i\rangle\alpha}(\gamma), \tag{62}$$

$$T(\gamma)\phi(x)T^{\dagger}(\gamma) = \phi(\gamma x), \qquad T(\gamma)\phi^{\dagger}(x)T^{\dagger}(\gamma) = \phi^{\dagger}(\gamma x). \tag{63}$$

For the gauge transformations, we proceed in an analogous manner. $K^{(n)}$ is a diagonal operator and therefore commutes with gauge transformations $g \in U(n)$, i.e.

$$K^{(n)}(g\phi) = g(K^{(n)}\phi). \tag{64}$$

Therefore, again, each eigenspace $\langle\omega_i\rangle$ carries the (same) antirepresentation of $U(n)$,

$$(gf)_{i\alpha}(x) = g_{\alpha\beta} f_{i\beta}(x), \tag{65}$$

which implies

$$\boldsymbol{a}_{i\gamma} \to \boldsymbol{a}_{i\beta} g_{\beta\gamma} = g_{\gamma\beta}^{\mathrm{T}} \boldsymbol{a}_{i\beta} \tag{66}$$

implemented by

$$A^{i(\boldsymbol{a})}(g) = \exp(-\boldsymbol{a}_{i\alpha}^{\dagger}[\log g^{\mathrm{T}}]_{\alpha\beta}\boldsymbol{a}_{i\beta}), \tag{67}$$

which is a representation of $U(n)$. The $\{\boldsymbol{b}_{i\alpha}\}$ must be treated slightly differently. Taking adjoints of (65)–(67) would lead to the $\boldsymbol{b}_{i\alpha}^{\dagger} f_{i\alpha}^*$ part of $\phi$ transforming as $g^*$ rather than $g$, so we set

$$(gf^*)_{i\alpha}(x) = g_{\alpha\beta} f_{i\beta}^*, \tag{68}$$

giving

$$\boldsymbol{b}_{i\gamma}^{\dagger} \to \boldsymbol{b}_{i\beta}^{\dagger} g_{\beta\gamma} = g_{\gamma\beta}^{\mathrm{T}} \boldsymbol{b}_{\beta}^{\dagger} \tag{69}$$

and

$$A^{i(b)}(g) = \exp(-\boldsymbol{b}^{\dagger}_{i\alpha}[\log g^{\mathrm{T}}]_{\alpha\beta}b_{i\beta}). \tag{70}$$

Setting

$$A(g) = \prod_{\{i,(\boldsymbol{a}/\boldsymbol{b})\}} A^{i(\boldsymbol{a}/\boldsymbol{b})}(g), \tag{71}$$

we find

$$A(g)\phi(x)A^{\dagger}(g) = g\phi(x), \tag{72}$$

and taking adjoints of (66) and (69) we also find

$$A(g)\phi^{\dagger}(x)A^{\dagger}(g) = \phi^{\dagger}(x)g^{\dagger} = g^{*}\phi^{\dagger}(x). \tag{73}$$

We note that, since they act on different indices, $T(\Gamma)$ and $A(U(n))$ clearly commute.

The projections $\phi \to \overset{\leftarrow a}{\phi}$ are now easy to implement. $a(\Gamma)$ is a subgroup of $U(n)$, hence implementable by (73). Setting

$$A(\gamma) = A(a(\gamma)), \qquad U(\gamma) = A(\gamma)T(\gamma^{-1}), \tag{74}$$

we find

$$\overset{\leftarrow a}{\phi}(x) = \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma} a(\gamma)\phi(\gamma^{-1}x) = \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma} U(\gamma)\phi(x)U^{\dagger}(\gamma). \tag{75}$$

$U(\Gamma)$ is clearly a representation of $\Gamma$ in the Fock space $\mathscr{F}(\mathscr{H})$, where $\mathscr{H}$ is (the Hilbert completion of†) the space of positive-energy (positive Klein–Gordon norm) solutions of the wave equation $K^{(n)}\phi = 0$.

Up to now we have done everything using the basis $\{f^{(*)}_{i\alpha}\}$, which was obviously the easiest to work with thus far, but clearly it is not the best for dealing with automorphic fields. Accordingly we note that, since $L$ satisfies (6), its extension $K^{(n)}$ satisfies (6) and (5), and so we can apply the spectral theory of § 2 to the solution space of $K^{(n)}$. Thus we can write

$$\phi(x) = \sum_i \boldsymbol{p}_i g_i + \boldsymbol{q}^{\dagger}_i h^{*}_i, \tag{76}$$

$$\phi^{\dagger}(x) = \sum_i \boldsymbol{q}_i h_i + \boldsymbol{p}^{\dagger}_i g^{*}_i. \tag{77}$$

The $\{g_i\}$ are supposed to be related to the $\{f_{i\alpha}\}$ by the machinations of § 2. For the $\{h_i\}$ we have three cases to consider.

I. $a(\Gamma)$ is real. This means that if $g$ is automorphic by $a(\Gamma)$, so is $g^*$; and if $g$ is not, then neither is $g^*$. Thus we may set $h_i = g_i$ in (76), (77). Charge conjugation is the interchange $\boldsymbol{p}^{(\dagger)}_i \leftrightarrow \boldsymbol{q}^{(\dagger)}_i$ (implementable by analogy with previous results), and since $g$ and $g^*$ live or die together in the projection, charge conjugation survives in its standard form for the projected fields.

II. $a(\Gamma)$ is equivalent to $a^*(\Gamma)$ but not reducible to real form. Thus $a^*(\Gamma) = R^{\dagger}a(\Gamma)R$. In this case, if $g$ is automorphic by $a(\Gamma)$, then so is $Rg^*$—not $g^*$. Thus we set $h_i = Rg_i$ in (76), (77). Charge conjugation comes out modified for the projected fields, since it is now clearly mixed up with a gauge transformation which cannot be eliminated. It is still the interchange $\boldsymbol{p}^{(\dagger)}_i \leftrightarrow \boldsymbol{q}^{(\dagger)}_i$, but the wavefunctions of conjugate particles are not the same.

† We resolutely avoid all discussion of domains of definition, details of convergence and similar technicalities.

III. $a(\Gamma)$ is not equivalent to $a^*(\Gamma)$. Then if $g$ is automorphic by $a(\Gamma)$, neither $g^*$ nor any gauge transform of it is, and so it will disappear in automorphic projection; so we may as well set $h_i = g_i$ for automorphic $g_i$. If $g$ is not automorphic, then $g^*$ (or some gauge relative of it) *may* be, and so we set $h_i^* = (R)g_i^*$ in this case. To preserve the $g/h$ symmetry we can transform the original $g_i$ to $(h_i^*)^*$ if we wish, giving $h_i = g_i$ for all modes. Charge conjugation as such disappears for case III under projection, since no $\boldsymbol{p}_i^{(\dagger)}$ or $\boldsymbol{q}_i^{(\dagger)}$ has a conjugate operator that survives—regarding $\phi$ as a complex field in this case is perhaps misleading, and it is best to rewrite $\phi$ as a purely real field multiplet in terms of its real and imaginary parts, automorphic behaviour being given by

$$\begin{bmatrix} \mathrm{Re}\ \overset{\leftarrow a}{\phi} \\ \mathrm{Im}\ \overset{\leftarrow a}{\phi} \end{bmatrix}(\gamma x) = \begin{bmatrix} Re\, a(\gamma), & -\mathrm{Im}\ a(\gamma) \\ \mathrm{Im}\ a(\gamma), & \mathrm{Re}\ a(\gamma) \end{bmatrix} \begin{bmatrix} \mathrm{Re}\ \overset{\leftarrow a}{\phi} \\ \mathrm{Im}\ \overset{\leftarrow a}{\phi} \end{bmatrix}(x). \tag{78}$$

In this *real* representation, the field is self-conjugate (one could adopt a similar attitude in case II of course), and we see that in this case the connection between the interpretations of $\phi$ and $\overset{\leftarrow a}{\phi}$, straightforward in case I, somewhat more tenuous in case II, disappears almost completely.

Automorphic modes, it must be added, do not exist in isolation, but are part of a set of $k^2$ modes as in § 2. Therefore, where a gauge transformation has to be performed in cases II and III above, it is to be understood that all $k^2$ modes are subjected to the same transformation to ensure their linear independence is preserved.

With these thoughts in mind, we see that we can find a basis in which the fields can be written as in (76), (77), with each mode having yes/no behaviour under automorphic projection, $\phi \to \overset{\leftarrow a}{\phi}$, $\phi^\dagger \to \overset{\leftarrow a^*}{\phi^\dagger}$. Since the change of basis involved does not mix positive- and negative-energy solutions, it is a trivial Bogoliubov transformation and so preserves the commutation relations and definition of the vacuum $|0\rangle$,

$$[\boldsymbol{p}_i, \boldsymbol{p}_j^\dagger] = [\boldsymbol{q}_i, \boldsymbol{q}_j^\dagger] = \delta_{ij}, \qquad [\boldsymbol{p}_i, \boldsymbol{p}_j] = [\boldsymbol{q}_i, \boldsymbol{q}_j] = \ldots 0, \tag{79}$$

$$\boldsymbol{a}_{i\alpha}|0\rangle = \boldsymbol{b}_{i\alpha}|0\rangle = 0 \Leftrightarrow \boldsymbol{p}_i|0\rangle = \boldsymbol{q}_i|0\rangle = 0, \tag{80}$$

so that the two bases are entirely equivalent.

Let us consider the effect of gauge transformations a little further. Normally one uses Noether's theorem to derive a conserved current and integrates the time component to obtain a conserved charge which, when normal ordered and exponential, generates the gauge transformation. All of these operations can be written in terms of the field variables themselves and so are local. To say the same thing again another way, the charges are all actually constructed from the Lie algebra of the gauge group and so the whole of the one-parameter family of transformations they generate lies in the gauge group. By contrast, the discreteness of $\Gamma$ means that any representation $a(\Gamma)$ need not be exponentiable within the Lie algebra. Thus, although we can implement it as done above, the one-parameter families generated by the $\{\log(a(\Gamma))\}$ need not all lie within the gauge group and may only intersect it in a discrete number of points. This fact expresses a possible essential non-locality in the situation we consider. A more extreme example of the same thing is provided by the isometries $x \to \gamma x$, for which there is no possibility of deriving 'charges' unless there are Killing vector fields whose integral curves connect $x$ and $\gamma x$, in which case the 'charges' are essentially the momentum operators of the field $\phi$ in the direction of the Killing vectors.

Now consider the effect of a gauge transformation $g$ implemented by $A(g)$ on an automorphic field $\overleftarrow{\phi}^{a}$. Normally we have

$$
\begin{aligned}
\overleftarrow{\phi}(\gamma x) &= T(\gamma)\overleftarrow{\phi}T^{\dagger}(\gamma) = \frac{1}{|\Gamma|}\sum_{\gamma'\in\Gamma} T(\gamma)U(\gamma')\phi U^{\dagger}(\gamma')T^{\dagger}(\gamma) \\
&= A(\gamma)\left\{\frac{1}{|\Gamma|}\sum_{\gamma'\in\Gamma} U(\gamma^{-1}\gamma')\phi U^{\dagger}(\gamma^{-1}\gamma')\right\}A^{\dagger}(\gamma) \\
&= a(\gamma)\overleftarrow{\phi}(x)
\end{aligned}
\tag{81}
$$

as required. Acting now by $A(g)$,

$$
\begin{aligned}
&A(g)\overleftarrow{\phi}(\gamma x)a^{+}(g) \\
&\qquad = A(g)T(\gamma)\overleftarrow{\phi}T^{\dagger}(\gamma)A^{\dagger}(g) = A(g)A(\gamma)\overleftarrow{\phi}A^{\dagger}(\gamma)A(g) \\
&\qquad = A(ga(\gamma)g^{-1})A(g)\overleftarrow{\phi}A^{\dagger}(g)A^{\dagger}(ga(\gamma)g^{-1}) = [ga(\gamma)g^{-1}][g\overleftarrow{\phi}],
\end{aligned}
\tag{82}
$$

we find that $A(g)$ implements the inner automorphism of the representation as well as implementing the gauge transformation, as we would expect. If $g$ is generated by a *bona fide* charge in the Lie algebra, we can say that the charge generates the unitary equivalence between the gauge-related automorphic fields.

We close this section with some remarks about two-point functions of the projected fields: firstly, the commutator, or projected analogue of (54). We can write (49) as

$$
\mathrm{i}G^{(n)}(x, x') = \sum_i g_i^*(x)g_i(x') - h_i(x)h_i^*(x'),
\tag{83}
$$

which leads to

$$
[\overleftarrow{\phi}^{a}(x), \overleftarrow{\phi}^{\dagger\,a*}(x')] = [\overleftarrow{\phi}^{\dagger\,a*}(x), \overleftarrow{\phi}^{a}(x')] = -\mathrm{i}\,\overleftrightarrow{G^{(n)*}(x, x')}^{a},
\tag{84}
$$

where

$$
\begin{aligned}
\mathrm{i}\,\overleftrightarrow{G^{(n)*}(x, x')}^{a} &= \frac{\mathrm{i}}{|\Gamma|^2}\sum_{\gamma,\gamma'\in\Gamma} a(\gamma^{-1})G^{(n)*}(\gamma x, \gamma' x')a(\gamma') \\
&= \sum_i \overleftarrow{g_i}^{a}(x)\,\overleftarrow{g_i^*}^{a*}(x') - \overleftarrow{h_i^*}^{a}(x)\,\overleftarrow{h_i}^{a*}(x') \\
&= \sum_{ij} \delta_{i,a(\Gamma)}g_i(x)g_i^*(x') - \delta_{j,a(\Gamma)}h_j^*(x)h_j(x')
\end{aligned}
\tag{85}
$$

and

$$
\delta_{i,a(\Gamma)} = \begin{cases} 1 & \text{if } g_i(h_j^*) \text{ is automorphic by } a(\Gamma) \\ 0 & \text{otherwise.} \end{cases}
\tag{86}
$$

The equal-times version of this is clearly

$$
[\partial_t\overleftarrow{\phi}^{a}(x), \overleftarrow{\phi}^{\dagger\,a*}(x')]|_{t=t'} = [\partial_t\overleftarrow{\phi}^{\dagger\,a*}(x), \overleftarrow{\phi}^{a}(x')]|_{t=t'} = -\mathrm{i}\overleftrightarrow{\delta}^{(n)*\,a}_{\tilde{M}}(x, x').
\tag{87}
$$

The Wightman functions are just as easily dealt with. For example,

$$
\langle 0|\phi(x)\phi^{\dagger}(x')|0\rangle = \sum_i g_i(x)g_i^*(x') = W_2(x, x').
\tag{88}
$$

Hence

$$\langle 0|\overleftarrow{\phi}^{a}(x)\overleftarrow{\phi}^{\dagger a^*}(x')|0\rangle$$

$$=\frac{1}{|\Gamma|^2}\sum_{\gamma,\gamma'\in\Gamma}\langle 0|U(\gamma)\phi(x)U^\dagger(\gamma)U(\gamma')\phi^\dagger(x')U(\gamma')|0\rangle$$

(using the invariance of the vacuum)

$$=\frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}\langle 0|\phi(x)U(\gamma)\phi^\dagger(x')|0\rangle \tag{89}$$

$$=\sum_i \overleftarrow{g}_i^{a}(x)\overleftarrow{g}_i^{*a^*}(x')=\sum_i \delta_{i,a(\Gamma)}g_i(x)g_i^*(x')=\overleftrightarrow{W}_2^{a}(x,x'). \tag{90}$$

One can obviously do the same thing for higher-order functions, time-ordered products, and so on.

## 4. Classification

The results of the previous section allow us to classify quantised automorphic fields in a simple manner. So let us consider field theories on a multiply connected $T\otimes M$ having an internal symmetry group $G$ which we take to be compact in order to avoid infinite-mass multiplets—as is customary. Clearly, upon pulling back to $T\otimes\tilde{M}$ we see that every representation $a(\Gamma)$ gives an automorphic field theory, so that the various different possibilities are exhausted by the elements of $\mathrm{Hom}(\pi_1(M), G)$. Furthermore, the compactness of $G$ means that an analysis similar to that given in § 3 can be carried out for the implementation of the gauge symmetries in $G$, and so $a(\Gamma)$ and $ga(\Gamma)g^{-1}$ determine the same theory. Thus we form the equivalence classes under inner automorphism in $\mathrm{Hom}(\pi_1(M), G)$, and the set of equivalence classes, written $\mathrm{Hom}_G(\pi_1(M), G)$, classifies the automorphic field theories on $\tilde{M}$ associated with the particular covering map $\pi:\tilde{M}\to M$ and particular gauge group $G$.

Computing $\mathrm{Hom}_G(\pi_1(M), G)$ is not too difficult in principle. We assume that $\pi_1(M)\equiv\Gamma$ has been given in generator–relation format with generators $\gamma_1\ldots\gamma_n$ and relations $r_1(\{\gamma_i\})\ldots r_k(\{\gamma_i\})$. Any map $a:\{\gamma_i\}_{i=1\ldots n}\to G$ extends by definition to the free group on the $\{\gamma_i\}$ by

$$a(\gamma_{i_1}^{n_1}\gamma_{i_2}^{n_2}\ldots\gamma_{i_l}^{n_l})=[a(\gamma_{i_1})]^{n_1}[a(\gamma_{i_2})]^{n_2}\ldots[a(\gamma_{i_l})]^{n_l}, \tag{91}$$

and gives a representation in $G$ of the free group, which then factors down into a representation of $\Gamma$ if and only if

$$a(r_j(\{\gamma_i\}))=r_j(\{a(\gamma_i)\})=e. \tag{92}$$

So $\mathrm{Hom}(\pi_1(M), G)$ is given by those functions $a:\{\gamma_i\}\to G$ which satisfy (92)†. Knowledge of the structure of $G$ then enables us to determine the action of any inner

† Banach and Dowker (1979b) gives somewhat more detail on this point and computes some specific examples.

automorphism by $g \in G$ on the $a(\gamma_i)$; $a(\gamma_i) \to ga(\gamma_i)g^{-1}$, and thus to find which elements of $\mathrm{Hom}(\pi_1(M), G)$ are equivalent—hence determining $\mathrm{Hom}_G(\pi_1(M), G)$.

In practice, we use coordinates for $G$ which will normally be a Lie group, and so if $\Gamma$ has $n$ generators and $G$ is $m$-dimensional, the representations of the free group on $\{\gamma_i\}_{i=1\ldots n}$ will be in 1–1 correspondence with an open set in $\mathbb{R}^{nm}$. The $k$ relations provide a number of equations of constraint via (92) (naïve dimension-counting does not usually work), and this determines a subspace $H$ of $\mathbb{R}^{nm}$ which is 1–1 correspondence with $\mathrm{Hom}(\pi_1(M), G)$. Finally, the action of inner automorphisms $a \to gag^{-1}$ yields an equivalence relation on $H$, and the set of equivalence classes $H_G$ is in 1–1 correspondence with $\mathrm{Hom}_G(\pi_1(M), G)$. There may even be a subspace $\check{H}$ of $H$ which intersects each element of $H_G$ precisely once, in which case its points label the elements of $H_G$ and so coordinatise $\mathrm{Hom}_G(\pi_1(M), G)$. Of course, judicious choice of coordinates for $G$ greatly eases these calculations and can provide transparent expressions for the subspace $\check{H}$ (if it exists). The coordinates in $\check{H}$ of a particular theory can be said to express the 'topological quantum numbers' of the theory.

How does all this compare with what is done when one admits local gauge symmetries into the theory? Local gauge symmetries are of two types, large and small. Small gauge transformations are those which are homeomorphic to the identity transformation, others being large. Two field configurations are equivalent if they differ by a small transformation and the equivalence classes are the elements of the first cohomology set of $M$ with coefficients in $G_\partial$, $H^1(M, G_\partial)$, where $G_\partial$ is the sheaf of germs of differentiable $G$-valued functions on $M$ (see Hirzebruch (1966) for details, especially p 41). Further lengthy analysis is required to re-express this classification in terms of more familiar quantities such as the standard cohomology groups of $M$ (see Avis and Isham (1978) for relevant details). Each equivalence class (or element of $H^1(M, G_\partial)$ is said to define a sector of the theory. Now restricting the theory to one sector breaks local gauge symmetry, because large transformations are excluded. The usual solution to this is to form linear combinations of the sectors which give states in the theory invariant (up to a phase factor) under large and small transformations. Now the large transformations permute the sectors, and this leads to a group structure on equivalence classes of large transformations. The linear combinations required then turn out to be elements of the dual group of this group. See Jackiw (1977) and Avis and Isham (1979) for specific examples.

We can now contrast this with the rigid gauge framework of this paper. The most salient feature is that rigid gauge transformations do not split into large and small types. Even discrete transformations are exponentiable from the identity in the sense discussed in § 3. Thus the whole business of forming linear combinations of sectors is avoided. In automorphic rigid gauge theories each sector defines a complete theory, where by a sector we mean a member of $\mathrm{Hom}_G(\pi_1(M), G)$.

## 5. The field algebra

In this section we consider the so-called projection operator $\phi \to \overset{\leftarrow a}{\phi}$ further—in particular its extension to $\mathcal{R}$ and consequences and applications of this. Now $\mathcal{R}$ is generated by the $\{p_i^{(\dagger)}, q_i^{(\dagger)}\}$, and the map $\overset{\leftarrow a}{}$ is defined for these by

$$\overset{\leftarrow a}{p_i} = \delta_{i,a(\Gamma)} p_i, \text{ etc} \tag{93}$$

according to the various cases considered in § 3. We extend this to the whole of $\mathcal{R}$ in the obvious way:

$$\overleftarrow{(\boldsymbol{p}_1^{(\dagger)}\ldots\boldsymbol{p}_n^{(\dagger)})}^{a}=\overleftarrow{(\boldsymbol{p}_1^{(\dagger)})}^{a}\ldots\overleftarrow{(\boldsymbol{p}_n^{(\dagger)})}^{a}. \tag{94}$$

Suppose for some $\boldsymbol{p}_i$ that $\overleftarrow{\boldsymbol{p}_i}^{a}=0$. Then we have that

$$0=\overleftarrow{\boldsymbol{p}_i}^{a}\overleftarrow{\boldsymbol{p}_i^{\dagger}}^{a}=\overleftarrow{\boldsymbol{p}_i^{\dagger}}^{a}\overleftarrow{\boldsymbol{p}_i}^{a}=\overleftarrow{[\boldsymbol{p}_i,\boldsymbol{p}^{\dagger}]}^{a}=\overleftarrow{\mathbb{I}}^{a}. \tag{95}$$

Thus, for any $A\in\mathcal{R}$,

$$\overleftarrow{A}^{a}=\overleftarrow{\mathbb{I}A}^{a}=\overleftarrow{\mathbb{I}}^{a}\overleftarrow{A}^{a}=0, \tag{96}$$

and the only projection on $\mathcal{R}$ that $\overleftarrow{\ }^{a}$ extends to is the zero map. This is just a reflection of the fact that $\mathcal{R}$ is simple, since the kernel of $\overleftarrow{\ }^{a}$ must be a two-sided ideal, and it is well known that $\mathcal{R}$ does not possess any of these† (other than $\mathcal{R}$ and 0).

Suppose indeed that $\overleftarrow{\ }^{a}$ *was* a non-trivial projection. A state of the field is (if nothing else) a positive linear functional on $\mathcal{R}$. If $\rho$ is a state, we can write the expectation of $A$ as $(\rho, A)$ and that of $\overleftarrow{A}^{a}$ as $(\rho,\overleftarrow{A}^{a})$. Defining then $\overleftarrow{\rho}^{a}$ as $\rho\circ\overleftarrow{\ }^{a}$, we find

$$(\overleftarrow{\rho}^{a}, A)=(\rho,\overleftarrow{A}^{a}), \tag{97}$$

and we have pulled back the projection to the space of states. The discussion of the previous paragraph thus says that there is *no* pullback to the states.

However, $\overleftarrow{\overleftarrow{\phi}^{a}}^{a}=\overleftarrow{\phi}^{a}$, which is obvious from (75); thus we *do* have a projection on the linear subspace of $\mathcal{R}$ spanned by *single* creation/annihilation operators $\mathcal{S}$,

$$A\in\mathcal{S}\Rightarrow A=\Sigma\,\alpha_i\boldsymbol{p}_i+\beta_i\boldsymbol{p}_i^{\dagger}+\gamma_i\boldsymbol{q}_i+\delta_i\boldsymbol{q}_i^{\dagger}\qquad(\alpha_i\ldots\delta_i\in\mathbb{C}). \tag{98}$$

(Actually, we *could* extend $\overleftarrow{\ }^{a}$ to $\mathcal{R}$ provided we apply (94) only to normal ordered products (or some other suitable convention), the result for non-normal products being given by the CCR; however, this extension would simply define a linear map on $\mathcal{R}$ which does not preserve the multiplicative structure (e.g. if $\overleftarrow{\boldsymbol{p}_i}^{a}=0$, $\overleftarrow{\boldsymbol{p}_i\boldsymbol{p}_i^{\dagger}}^{a}=\mathbb{I}$) and so would be neither natural nor relevant to the physics at hand.)

The non-zero part of $\overleftarrow{\mathcal{S}}^{a}$ (where $\overleftarrow{\mathcal{S}}^{a}$ has its obvious meaning) generates a subalgebra of $\mathcal{R}$, $\mathcal{R}_a$, the automorphic field algebra. As a general rule both $\mathcal{S}$ and $\overleftarrow{\mathcal{S}}^{a}$ are infinite-dimensional with a countable basis, and so there will be a map $V:\mathcal{S}\to\overleftarrow{\mathcal{S}}^{a}$ defined by

$$V\boldsymbol{p}_i=\overleftarrow{\boldsymbol{p}_j}^{a}, \tag{99}$$

where $\boldsymbol{p}_i$ is any $\boldsymbol{p}_i$ in $\mathcal{S}$ and $\overleftarrow{\boldsymbol{p}_j}^{a}$ is some annihilation operator which survives automorphic projection. The infinite dimensionality of both $\mathcal{S}$ and $\overleftarrow{\mathcal{S}}^{a}$ ensures the 1–1 nature of $V$ is

† For suppose $A$ is in the ideal and $A$ contains $\boldsymbol{p}_i$ (say) as a factor in one of its terms. Then $[A,\boldsymbol{p}_i^{\dagger}]$ is a 'smaller' element of $\mathcal{R}$. We can continue in this manner until we get a multiple of the identity as in (95); hence (96) follows. All of this can be regarded as a crude bastardisation of the sophisticated results of Borchers (1967) (for Minkowski space) and Slawny (1972).

possible. The CCR and the obvious rule

$$V(\boldsymbol{p}_1 \dots \boldsymbol{p}_l) = (V\boldsymbol{p}_1) \dots (V\boldsymbol{p}_l), \text{ etc} \tag{100}$$

define an extension of $V$ to an isomorphism between $\mathscr{R}$ and $\mathscr{R}_a$. Note that $V$ is not an automorphism of $\mathscr{R}$, since it is not onto. This isomorphism $V$ is actually implementable by a pseudo-unitary operator $\mathscr{V}: \mathscr{F}(\mathscr{H}) \to \mathscr{F}(\overleftarrow{\mathscr{H}}^{a})$ (1–1, but again not onto) by analogy with § 3,

$$V\boldsymbol{p}_i^{(+)} = \mathscr{V}\boldsymbol{p}_i^{(+)}\mathscr{V}^{+}, \tag{101}$$

$$\mathscr{V}^{+}\mathscr{V} = \mathbb{1}_{\mathscr{F}(\mathscr{H})}, \qquad \mathscr{V}\mathscr{V}^{+} = \mathbb{1}_{\mathscr{F}(\overleftarrow{\mathscr{H}}^{a})}, \tag{102}$$

and what we have shown is essentially a case of the well-known result that the Fock representation of the CCR over a separable Hilbert space (of infinite dimension) is unique up to unitary equivalence (e.g. Berezin 1966, theorem 1.4).

Now our spectral theory says that for the one-particle states $\mathscr{H}$ we have

$$\mathscr{H} = \overleftarrow{\mathscr{H}}^{a} \oplus \overleftarrow{\mathscr{H}}^{a\perp} \tag{103}$$

reflected at the one-particle operator level by

$$\mathscr{S} = \overleftarrow{\mathscr{S}}^{a} \oplus \overleftarrow{\mathscr{S}}^{a\perp}. \tag{104}$$

Thus we have for Fock space $\mathscr{F}(\mathscr{H}) = \mathbb{C} \oplus \mathscr{H} \oplus (\mathscr{H} \otimes \mathscr{H})_s \oplus (\mathscr{H} \otimes \mathscr{H} \otimes \mathscr{H})_s \oplus \dots$

$$\mathscr{F}(\mathscr{H}) = \mathscr{F}(\overleftarrow{\mathscr{H}}^{a} \oplus \overleftarrow{\mathscr{H}}^{a\perp}) \cong \mathscr{F}(\overleftarrow{\mathscr{H}}^{a}) \otimes \mathscr{F}(\overleftarrow{\mathscr{H}}^{a\perp}), \tag{105}$$

where the isomorphism (105) is given by

$$[(k+l)!/k!\,l!]^{1/2}|\Phi_k \otimes_s \Psi_l\rangle \leftrightarrow |\Phi_k\rangle \otimes |\Psi_l\rangle, \tag{106}$$

where $\otimes_s$ is the symmetrised tensor product, $|\Phi_k\rangle$ is a $k$-particle state in $\mathscr{F}(\overleftarrow{\mathscr{H}}^{a})$, and $|\Psi_l\rangle$ is an $l$-particle state in $\mathscr{F}(\overleftarrow{\mathscr{H}}^{a\perp})$. Thus we can write for some $|\Psi\rangle \in \mathscr{F}(\mathscr{H})$

$$|\Psi\rangle = \sum_{k,l} \psi_{kl} |\theta_k\rangle \otimes |\chi_l\rangle, \tag{107}$$

where the $|\theta_k\rangle$ are a basis for $\mathscr{F}(\overleftarrow{\mathscr{H}}^{a})$ and the $|\chi_l\rangle$ a basis for $\mathscr{F}(\overleftarrow{\mathscr{H}}^{a\perp})$. Now automorphic observables lie in $\mathscr{R}_a$; thus, if $A \in \mathscr{R}_a$, then

$$\begin{aligned}\langle\psi|A|\psi\rangle &= \sum_{k,k',l,l'} \psi^*_{k'l'}\psi_{kl}\langle\chi_{l'}|\chi_l\rangle\langle\theta_{k'}|A|\theta_k\rangle \\ &= \mathrm{Tr}(|\theta_k\rangle\rho_{kk'}\langle\theta_{k'}|, A),\end{aligned} \tag{108}$$

where

$$\rho_{kk'} = \sum_l \psi_{kl}\psi^*_{k'l}, \tag{109}$$

and so we see that the $\mathscr{F}(\overleftarrow{\mathscr{H}}^{a\perp})$ part of the state only enters into the expectation value through the expression $\rho_{kk'}$, and consequently any state $|\psi'\rangle$ yielding the same $\rho_{kk'}$ is indistinguishable from $|\psi\rangle$ in the automorphic field algebra. Thus the states on $\mathscr{R}$ divide into equivalence classes under the equivalence relation of having the same $\rho_{kk'}$. In the language of local $C^*$-algebras, the states on $\mathscr{R}_a$ are partial states. From the form of

(109) it is clear that $|\psi\rangle$ and $|\psi'\rangle$ are equivalent if and only if they differ by a unitary transformation of $\mathscr{F}(\overleftarrow{\mathscr{H}}^{a\perp})$, i.e.

$$|\psi\rangle \sim |\psi'\rangle \Leftrightarrow |\psi\rangle = \mathscr{W}_{\perp}|\psi'\rangle, \tag{110}$$

where

$$\mathscr{W}_{\perp}|\theta_k\rangle = |\theta_k\rangle, \qquad \mathscr{W}_{\perp} \text{ unitary in } \mathscr{F}(\overleftarrow{\mathscr{H}}^{a\perp}). \tag{111}$$

In particular, the vaccum of $\mathscr{R}_a$, $|0\rangle_a$, is

$$|0\rangle_a = \{|0_a\rangle \otimes |\chi\rangle|;\ |0_a\rangle, \text{ the vacuum of } \mathscr{F}(\overleftarrow{\mathscr{H}}^{a}),\ |\chi\rangle \text{ any normalised state in } \mathscr{F}(\overleftarrow{\mathscr{H}}^{a\perp})\}, \tag{112}$$

which in particular contains the vacuum of $\mathscr{F}(\mathscr{H})$, i.e.

$$|0\rangle \in |0\rangle_a; \tag{113}$$

thus we can regard the two vacuua as expressible by the same element $|0\rangle$.

We would now like to re-examine some simple models prevalent in the literature, in the light of the general comments made above.

In the Appendix to his paper, Fulling (1973) considers two periodic boxes of lengths $L$ and $2L$ and the two field theories on the resulting topologies of $T \otimes S^1$. Clearly the larger box is a covering space for the smaller and so the formalism of this paper applies. He then takes a fixed function $g$ with compact $S^1_{(L)}$ support ($\sigma$ say) and considers the smeared fields $\phi(g)$ in the two $T \otimes S^1$ universes. Now $\phi(g)$ has Cauchy data restricted to $\sigma$ in each case, and so the evolution of $\phi$ in the Cauchy development $\mathscr{D}$ of $\sigma$ is the same in both cases; hence (argues Fulling) local quantities in $\mathscr{D}$ ought to depend only on $\mathscr{D}$ itself and not on the embedding space. Working out $\langle 0|\phi(g)\phi(g)|0\rangle_L$ and $\langle 0|\phi(g)\phi(g)|0\rangle_{2L}$ and finding them different, Fulling concludes that the vacuua $|0\rangle_L$ and $|0\rangle_{2L}$ are different (since the field operators are, according to him, the same).

The conclusions of this section force us to disagree with Fulling's reasoning. To start with we have seen that we can take $|0\rangle_L$ and $|0\rangle_{2L}$ to be represented by the same state $|0\rangle_{2L}$ ($|0\rangle_L$ is an equivalence class containing $|0\rangle_{2L}$); therefore the differing values of $\langle 0|\phi(g)\phi(g)|0\rangle$ must mean that the operators $\phi(g)$ are different in the two cases. How can this be? We note that Fulling's argument that the evolution in $\mathscr{D}$ is independent of anything outside works fine for a *classical* field, but the $\phi(g)$ in question is a *quantum* operator. As such it lives in $\mathscr{R}$ (actually $\mathscr{S}$) and can be regarded as having no direct relationship with $\mathscr{D}$ at all. To be more explicit, we can write a quantised real field smeared with $f$ as

$$\begin{aligned}\phi(g) &= \sum_i \boldsymbol{a}_i((f_i, g)) + \boldsymbol{a}_i^{\dagger}((f_i^*, g)) \\ &= \sum_i \alpha_i \boldsymbol{a}_i + \beta_i \boldsymbol{a}_i^{\dagger},\end{aligned} \tag{114}$$

where the $\{f_i, f_i^*, g\}$ belong to some inner product space (in practice a function space and hence the connection with $\mathscr{D}$) and $((\ ,\ ))$ is a bilinear form on the space (the specific details are unimportant). All of the details of the specific form and domain of $g$ and the nature of the $f_i^{(*)}$ become coded into the numbers $\{\alpha_i, \beta_i\}$. As far as the field algebra is concerned, two $\phi(g)$ operators are only the same if the sets of numbers $\{\alpha_i, \beta_i\}$ are the same. The two $\phi(g)$ of Fulling are not the same, since $\phi_L(g)$ contains only those $\{\alpha_i, \beta_i\}$

of $\phi_{2L}(g)$ arising from modes in the expansion of $\phi$ which obey

$$f = \overset{\leftarrow a}{f} = \tfrac{1}{2}(f + f_{\text{diam}}), \tag{115}$$

where $f_{\text{diam}}(x) = f(x_{\text{diam}})$ ($x_{\text{diam}}$ being the diametrically opposite point to $x$ in $S^2_{(2L)}$), since we can regard $\phi_L$† as being an automorphic field on $S^1_{(2L)}$ under the trivial representation of the projection $S^1_{(2L)} \to S^1_{(L)}$ which identifies diametrically opposite points. It is clear that, when embedded in the $S^1_{(2L)}$ field algebra, $\phi_L(g)$ is in fact $\phi_{2L}(\overset{\leftarrow a^*}{g})$, and, since the $g$ considered by Fulling does not satisfy $g = \overset{\leftarrow a^*}{g}$ (in $S^1_{(2L)}$), the two operators $\phi_L(g) = \phi_{2L}(\overset{\leftarrow a^*}{g})$ and $\phi_{2L}(g)$ are obviously different; $\phi_L(g)$ describing a situation which in fact involves a disconnected Cauchy development $\Gamma\mathscr{D} = \bigcup_{\gamma\in\Gamma} \gamma\mathscr{D}$ when embedded in $T \otimes S^1_{(2L)}$. Thus the locality arguments used by Fulling lose their force when these non-local effects are considered.

To recapitulate, we saw above that $\mathscr{R}$ and $\mathscr{R}_a$ are isomorphic, and it is clear that the isomorphism preserves the values of the $\{\alpha_i, \beta_i\}$. Thus it makes sense to talk about a single abstract Fock representation of the field algebra. However, the way the isomorphism works means that field operators corresponding to locally defined quantities do not map to the same sets of $\{\alpha_i, \beta_i\}$—depending on whether the local phenomenon is viewed in the covering space or the factor space—hence do not map to the same abstract operator. This is a consequence of the non-local nature of the embedding of the process on the factor space in the covering space; a copy of the phenomenon in question has to be glued to each pre-image in the covering space—this embedding *must* in turn be considered in order to be able to identify two separate processes in two separate spaces as locally identical.

There is another way of looking at the situation just discussed, namely in the light of the Reeh–Schlieder theorems, one of which says, loosely speaking, that the subalgebra of smeared field operators $\mathscr{R}(\bigcirc)$ for which the smearing functions lie within a fixed bounded open set $\bigcirc$ of Minkowski space, acting on the vacuum, generate a dense set in Fock space. Generalising crudely, as we did above, to the case of arbitrary background geometry, we see that the subalgebra belonging to an open set should also send the vacuum to a dense set in $\mathscr{F}(\mathscr{H})$. Now $\mathscr{R}_a$ does *not* send the vacuum to a dense set in $\mathscr{F}(\mathscr{H})$, and so neither does $\mathscr{R}(\bigcirc) \cap \mathscr{R}_{ra}$. We see that few localised operators are automorphic. Bearing in mind that $\overset{\leftarrow a}{\phi}(g) = \phi(\overset{\leftarrow a^*}{g})$, which presumably is in $\mathscr{R}(\Gamma\bigcirc)$ if the support of $g$ is in $\bigcirc$, this should not surprise us. However, the $\mathscr{R}(\bigcirc) \cap \mathscr{R}_a$ argument works even if $\bigcirc = \Gamma\bigcirc$, and so we see that even $\mathscr{R}(\Gamma\bigcirc)$ contains a vast quantity of non-automorphic operators. This is only to be expected considering the large number of non-automorphic test functions having supports in $\Gamma\bigcirc$.

The other well-known Reeh–Schlieder theorem says that adjoining the vacuum projection to $\mathscr{R}(\bigcirc)$ makes it irreducible. Again an intersection with $\mathscr{R}_a$ causes a breakdown of this property.

These comments show that it is incorrect to identify uncritically the algebras $\mathscr{R}(\bigcirc)$ and $\mathscr{R}_a(\bigcirc) = \mathscr{R}(\bigcirc) \cap \mathscr{R}_a$.

Going a little further, the cyclicity of $|0\rangle$ for $\mathscr{R}(\bigcirc)$ means that $\mathscr{R}(\bigcirc)|0\rangle$ has a 'tail' in $\mathscr{F}(\mathscr{H})$. However, the overlap of this tail with states $\mathscr{R}(\bigcirc')|0\rangle$, with $\bigcirc'$ distant from $\bigcirc$, tends to fade smoothly. Taking now $\mathscr{R}_a(\bigcirc)$ in place of $\mathscr{R}(\bigcirc)$, we see that the overlap

† Minor technicalities involving the normalisation of the $f_i$ when actually embedding have been ignored here. They do not affect the conclusions.

must be periodically strengthened as $\mathrm{O}'$ passes copies of $\mathrm{O}$ in $\Gamma\mathrm{O}$. This kind of difference in the behaviour of the tails is the sort of phenomenon pointed to by Haag and Kastler (1964) in discussing unitarily inequivalent representations of the CCR, and hence, despite the fact that $\mathscr{R}$ and $\mathscr{R}_a$ are abstractly isomorphic, the isomorphism even being implementable by the pseudo-unitary $\mathscr{V}$, we find them to be unitarily inequivalent; essentially because there is no unitary $\mathscr{W}$ such that

$$\overleftarrow{\phi}^a = \mathscr{W}\phi\mathscr{W}^\dagger, \tag{116}$$

as can be seen from (75).

The non-fading of the tail of $\mathscr{R}_a(\mathrm{O})|0\rangle$ leads us to suspect that, in the $|\Gamma| = \infty$ case, the automorphic operators exist as strong limits of localised operators and are not members of the algebra of quasi-local operators at all, by analogy with the case of Minkowski space where operators which are global, affecting near and far regions equally strongly, are characterised thus.

The arguments presented above, sketchy though they are, bring us into conflict with the recent paper of Kay (1978), who seeks to explain the Casimir effects arising in automorphic situations as follows: (1) A small neighbourhood of the factor space looks like a small neighbourhood of the covering space so we identify the local algebras in the two cases. (2) Assume (using local quasi-equivalence ideas) that the covering space representation of the local algebra is sufficient to calculate everything, and thus that the factor space energy momentum ${}_a\langle 0|\overleftarrow{T}^a_{\mu\nu}|0\rangle_a$ may be obtained as the expectation value of the covering space energy momentum operator in some density matrix state on the covering space algebra, i.e. that

$${}_a\langle 0|\overleftarrow{T}^a_{\mu\nu}|0\rangle_a = \mathrm{Tr}(\overleftarrow{\rho}^a, T_{\mu\nu}). \tag{117}$$

(3) Concurrently with (2) assume (from the known inequality of $\langle 0|T_{\mu\nu}|0\rangle$ and ${}_a\langle 0|\overleftarrow{T}^a_{\mu\nu}|0\rangle_a$) that the factor space ground state $\overleftarrow{\rho}^a$ is not the same as the covering space ground state $|0\rangle\langle 0|$.

Now (1) is incorrect as we have seen, since a local operation on the factor space pulls back to a non-local operation on the covering space in general. (3) is incorrect as we have seen explicitly, since $|0\rangle \in |0\rangle_a$, and (2) is nothing more than an attempt to pull back the topology from the operator algebra to the space of states, which we have also shown to be impossible. Indeed, we see that, since $T_{\mu\nu}$ is a bilinear product of field operators, ${}_a\langle 0|\overleftarrow{T}^a_{\mu\nu}|0\rangle_a$ can eventually be expressed in terms of $\overleftrightarrow{W}^a_2$, and since no amount of manipulation will remove the $U(\gamma)$ factors from between the $\phi(x)$ and $\phi^\dagger(x')$ in (89), we cannot express $\overleftrightarrow{W}^a_2$ in terms of $W_2$. Were we able to do so, (116) would be true.

We will finish this section with some comments on perhaps more familiar matters, namely the accelerated observer in Minkowski space (AOMS), the eternal black hole (EBH) and the collapsing body (CB), with particular regard to the Reeh–Schlieder theorems. We will be very brief on notation and definition, referring to the article by Isham (1977) for an overview and extensive references.

A cursory glance at the Penrose diagrams for the AOMS and EBH shows a striking similarity. Both diagrams split into L(eft), R(ight), F(uture) and P(ast) regions, and a timelike observer in EBH or uniformly accelerated observer in AOMS are confined to, say, the R regions only. In both cases one sets up field decompositions in these regions (the Boulware and Rindler decompositions) which are incomplete in the sense that the field modes are not specified in the L regions. One then specifies a further set of modes in the L regions in the most convenient manner. Therefore the total field can be written

$$\phi = \phi_{\mathrm{L}} + \phi_{\mathrm{R}}, \tag{118}$$

and we see that the field (and hence Fock space) splits into two parts in precisely the same way as for automorphic field theory (103), (104), (105). As far as observables constructed from $\phi_{\mathrm{R}}$ go, the $\mathscr{F}(\mathscr{H}_{\mathrm{L}})$ part of states enters only through the density matrix obtained by tracing out $\mathscr{F}(\mathscr{H}_{\mathrm{L}})$ as in (108), (109). Hence we note that, as far as observers in R are concerned, the Boulware/Rindler vacua are degenerate. It is clear that this degeneracy is a consequence of the idealised nature and global properties of the models. It is not true in the CB case, since in the distant past there is no horizon, hence no field split.

Now in the case of AOMS and EBH one can define field modes by a different, global prescription yielding the Minkowski and Hartle–Hawking vacua respectively; it turns out that the Minkowski vacuum is full of thermally distributed Rindler quanta, and the Hartle–Hawking vacuum of Boulware quanta. This and other analogies lead us to identify the mathematical structure of these models very strongly.

Taking Minkowski space, we know that both Reeh–Schlieder theorems are true. They are true in particular for the region R, i.e. $\mathscr{R}(\mathrm{R})$ has $|0\rangle_{\mathrm{Mink}}$ as cyclic vector and $\mathscr{R}(\mathrm{R}) \cup (|0\rangle\langle 0|)_{\mathrm{Mink}}$ is irreducible. Going now to the EBH, we are tempted to likewise conjecture that $\mathscr{R}(\mathrm{R})$ has $|0\rangle_{\mathrm{H\text{-}H}}$ as cyclic vector and that $\mathscr{R}(\mathrm{R}) \cup (|0\rangle\langle 0|)_{\mathrm{H\text{-}H}}$ is irreducible. We are supported in this to some extent by the CB, since at early times and large radial distances the metric is hardly different from Minkowski space, and so, if we suppose the Reeh–Schlieder theorems true in that regime, the smooth evolution of the CB is unlikely to have as catastrophic effect as to invalidate them later on.

Two more points are worth making. Firstly, it would seem to matter which vacuum projections we add to $\mathscr{R}(\mathrm{R})$; in Minkowski space, the $(|0\rangle\langle 0|)_{\mathrm{Rin}}$ projection clearly commutes with $\boldsymbol{a}_{\mathrm{L}}^{\dagger}\boldsymbol{a}_{\mathrm{L}}$, where $\boldsymbol{a}_{\mathrm{L}}$ is any L mode, and so $\mathscr{R}(\mathrm{R}) \cup (|0\rangle\langle 0|)_{\mathrm{Rin}}$ is apparently not irreducible. Secondly, it is known that in Minkowski space $\mathscr{R}(\mathcal{O})$, for which there is some open $\mathcal{O}'$ totally spacelike to $\mathcal{O}$, contains no creation or annihilation operators. This is presumably due to the fact that creation and annihilation operators are formally obtainable by smearing the field with functions which are not good test functions. Now the subalgebra of Rindler R operators is constructed formally from creation and annihilation operators and so is assumed to contain them. Using the Bogoliubov transformation to express these in terms of Minkowski operators, we find that $\mathscr{R}(\mathcal{O})$ apparently contains Minkowski creation and annihilation operators. This must be untrue, and so the identification of $\mathscr{R}(\mathcal{O})$ with the formal algebra of Rindler R operators is cast into grave doubt. This problem evidently reflects back on the first point, weakening the status of our $\boldsymbol{a}_{\mathrm{L}}^{\dagger}\boldsymbol{a}_{\mathrm{L}}$ counter-example.

The above remarks have obvious repercussions for the EBH and CB, not to mention other situations. Whatever the ultimate truth or otherwise of the conjectures made here (and in the rest of this section for that matter), it seems fairly clear that the Reeh–Schlieder problem in a general spacetime holds the key to a much deeper understanding of field theory in a general background. For automorphic field theory, it can give crucial information on locality/non-locality questions, and in an event horizon context it can illuminate the precise role of the horizon operator algebra, the normal assumption being that large red shifts make its effects negligible.

## 6. Concluding remarks

The major points to have emerged from the preceding five sections are the following. (1) Lemma 1 always enables us to make a split in the spectral resolution of an operator into automorphic and non-automorphic parts. In the case of complex fields, this split can be further decomposed to completely clarify the structure of the resolution. (2) The spectral decomposition in turn permits a fairly simple-minded construction of automorphic fields in terms of the automorphic modes of the field. (3) Non-locality plays a significant role in the relationship of the automorphic and covering field algebras; a subalgebra and partial states formalism seems to be the best heuristic vehicle for the theory, and the whole area would gain immensely from a thorough rigorous investigation. (4) The Reeh–Schlieder problem would amply repay vigorous study in a general background, throwing much light on not only automorphic field theory, but on many other situations of great interest.

## Appendix

In this Appendix we show the existence of $n$ diagonal, $n$-dimensional unitary matrices, orthogonal in the trace inner product. We take one of them to be the identity.

The fact that they are unitary and diagonal means that each non-zero entry must be of absolute value one; and we can also take out an overall phase factor in each to make the (1, 1) element equal to one.

Mapping each diagonal unitary matrix to its sequence of diagonal elements, we see that we are looking for $n-1$ sequences of the form $(1, e^{i\lambda^i_1} \dots e^{i\lambda^i_{n-1}})$ satisfying

$$1+e^{i\lambda^i_1}+\dots+e^{i\lambda_{n-1}}=0, \tag{A1}$$

$$1+e^{i(\lambda^i_1-\lambda^j_1)}+\dots+e^{i(\lambda_{n-1}-\lambda_{n-1})}=0. \tag{A2}$$

The condition (A1) is the orthogonality to $(1, 1 \dots 1)$, while (A2) represents the mutual orthogonality of the $n-1$ additional sequences.

Clearly the $\lambda^i_j$ constitute $(n-1)^2$ independent real parameters. Counting constraints, we see that (A1) represents $2\times(n-1)$ real constraints, and (A2) represents $2\times\frac{1}{2}(n-1)(n-2)$ real constraints— $n(n-1)$ constraints in total, so that we have a highly overconstrained system.

If we picture the sequences required as sequences of unit vectors in the complex plane, we see that (A1) is the condition that the vectors fit together to form a closed $n$-gon. Thus, despite the fact that for any particular sequence (A1) represents but two constraints, we see that we already encounter subtle domain problems for the $\lambda_i$ to ensure that the endpoint of the $(n-2)$nd vector in the chain is within a distance 2 of the starting point of the chain, so that the last two vectors are actually able to close the gap.

In view of all of these difficulties, then, it is particularly gratifying to be able to present a special solution to the problem for any $n$.

For the moment, *let n be prime*, and consider the finite field $\mathbb{Z}_n$ of arithmetic modulo $n$; in particular its multiplication table. Thus let $n_{ij}=i\times j \pmod n$. For $i$ fixed and $\neq 0 \pmod n$, the numbers $n_{ij}$, $j=0\dots(n-1)$, exhaust $\mathbb{Z}_n$ in a 1–1 manner, a consequence of the primeness of $n$. Furthermore, $n_{i_1 j}-n_{i_2 j}=n_{\{i_1-i_2 \pmod n\}j}$, and so for $i_1\neq i_2$ we have that the $n_{i_1 j}-n_{i_2 j}$ $(j=0\dots(n-1))$ also exhaust $\mathbb{Z}_n$. Now let $e^{i\omega}$ be an $n$th root of unity. We know (even for non-prime $n$) that

$$\sum_{k\in\mathbb{Z}_n} e^{ik\omega}=0, \tag{A3}$$

so, setting

$$\Omega_{ij}=e^{in_{ij}\omega}, \tag{A4}$$

we have that

$$\sum_{j\in\mathbb{Z}_n}\Omega_{ij}=\sum_{j\in\mathbb{Z}_n}e^{in_{ij}\omega}=0, \tag{A5}$$

$$\sum_{j\in\mathbb{Z}_n}\Omega_{i_1j}\Omega^*_{i_2j}=\sum_{j\in\mathbb{Z}_n}e^{i(n_{i_1j}-n_{i_2j})\omega}=0, \tag{A6}$$

which we recognise as (A1) and (A2) respectively. Thus for prime $n$ the $\Omega_{ij}$ solve the problem (clearly $\Omega_{0j}=1$ for any $j$, so this is just the identity sequence).

For non-prime $n$ we simply take the tensor product of the solutions for its prime factorisation. Thus let $n=m_1\ldots m_k$ with the $m_i$ all prime, and let

$$\Omega^n_{(i_1\ldots i_k)(j_1\ldots j_k)}=\Omega^{m_1}_{i_1j_1}\ldots\Omega^{m_k}_{i_kj_k}, \tag{A7}$$

where $\Omega^{m_i}_{i_ij_i}$ are the solutions for the prime factors. Then for (A1) we have

$$\sum_{j_1\in\mathbb{Z}_{m_1}}\cdots\sum_{j_k\in\mathbb{Z}_{m_k}}\Omega^n_{(i_1\ldots i_k)(j_1\ldots j_k)}=\sum_{j_1\in\mathbb{Z}_{m_1}}\Omega^{m_1}_{i_1j_1}\cdots\sum_{j_k\in\mathbb{Z}_{m_k}}\Omega^{m_k}_{i_kj_k}=0 \tag{A8}$$

and similarly for (A2). This completes our existence proof.

Finally, it is interesting to speculate on the uniqueness of our solution. For $n=2, 3$ the solution is obviously unique (no degrees of freedom in an equilateral triangle). For $n=4$ it is easy to see that the general solution of (A1) is a rhombus and so has one free angle in it. After that, a little trial and error shows that the general solution of (A1) and (A2) is (identity), $(1, 1, -1, -1)$, $(1, -1, \pm e^{i\alpha}, \mp e^{i\alpha})$†, and so we actually see that there is a one-parameter family of solutions despite the apparent overconstrainedness of the system. For $n=5$ and above we see that we encounter the domain problems mentioned above, and the general implications of (A1) and (A2) are not at all easy to fathom out. The nature of the general solution for any $n$ would be an interesting, if rather useless, piece of information.

## References


Avis S J and Isham C J 1978 *Cargese Lecture Notes*
—— 1979 *Nucl. Phys.* B **156** 441
Banach R and Dowker J S 1979a *J. Phys. A: Math. Gen.* **12** 2527–43
—— 1979b *J. Phys. A: Math. Gen.* **12** 2545–62
Berezin F A 1966 *The Method of Second Quantisation* (New York: Academic)
Borchers H J 1967 *Commun. Math. Phys.* **4** 315
DeWitt B S, Hart C F and Isham C J 1979 *Physica* **96**A 197
Dowker J S and Banach R 1978 *J. Phys. A: Math. Gen.* **11** 2255
Ford L H 1979a *King's College Preprint*
—— 1979b *King's College Preprint*
Fulling S A 1973 *Phys. Rev.* D **7** 2850
Haag R and Kastler D 1964 *J. Math. Phys.* **5** 848


† The $\Omega^4$ solution is $\alpha=0$.

Hirzebruch F 1966 *Topological Methods in Algebraic Geometry* (Berlin: Springer)
Isham C J 1977 *Ann. NY Acad. Sci.* **302** 114
—— 1978 *Proc. R. Soc.* A **362** 383
Jackiw R 1977 *Rev. Mod. Phys.* **49** 681
Kay B G 1978 *Phys. Rev.* D in the press
Markus L 1973 *Adv. Math.* **11** 351
Slawny J 1972 *Commun. Math. Phys.* **24** 151