Automorphic field theory-some mathematical issues

1979 J. Phys. A: Math. Gen. 12 2527

(http://iopscience.iop.org/0305-4470/12/12/031)

J. Phys. A: Math. Gen., Vol. 12, No. 12, 1979. Printed in Great Britain

Automorphic field theory—some mathematical issues

Richard Banach† and J S Dowker

Department of Theoretical Physics, Schuster Laboratory, Manchester University, Manchester M13 9PL, UK

Received 16 February 1979

Abstract. The general structure of field theories on multiply connected spaces is presented using the universal covering space concept. Restriction to rigid gauge theories allows the use of representation theory to answer several relevant mathematical questions.

1. Introduction

Recently there has been a marked increase in interest in field theories on space–times carrying non-trivial topology. At present, two approaches seem to have emerged. In the first, initiated by Isham (1978a, b), the attack is via the various cohomology groups of the space–time and extensive classification schemes have emerged. In the second, the concern of ourselves and of the Texas group (Dowker 1972, Dowker and Banach 1978, B S De Witt and C F Hart, private communication), the non-trivial fundamental group of the space–time is used to pull back the field theory onto the universal covering space–time manifold. The resulting theory closely resembles modern automorphic function theory (although the emphasis is obviously different) and for this reason we tend to use the adjective 'automorphic' to describe the special features that arise. This rather conflicts with Isham's use of terms like 'standard' and 'twisted' field (Isham 1978a, b) but we think that the covering space approach has a sufficiently different flavour to warrant the use of the mathematician's terminology.

This paper is chiefly concerned with mathematical questions. We extend our earlier work on scalar fields to include the methods and problems relevant to multicomponent and higher spin fields. Our results vary from very practical calculational devices to the more special and technical. We concentrate almost entirely on linear operators since many relevant problems reduce to them and because they afford the most scope for investigation.

In §§ 2 and 3 we set up the relevant formalism and consider some of the basic problems that arise. Sections 4 and 5 are concerned with the more specialised topics of self-adjoint operators and traces. Section 6 considers the case of a free field and the simplifications present there. We reserve applications of the results presented here to forthcoming papers.

† Research supported by an SRC Research Studentship.

0305-4470/79/122527+17$01.00

2. Automorphic fields

It is by now well known that to carry a reasonably unambiguous field theory, a space–time needs to possess a global time-like Killing field. Accordingly, we restrict our attention to space–times having the topology $T \otimes M$. We are particularly interested in fields which may not be single-valued on M and to this end the fundamental group of M, $\pi_1(M)$, is significant. The familiar fact that a function which is sufficiently regular in a region is necessarily single-valued there if $\pi_1(M) = \{e\}$ leads us to demand that $\pi_1(M) \neq \{e\}$ or M is multiply connected. We next introduce the universal covering space of M, $\tilde{M}$, and again standard theory says that $\pi_1(M)$ (which from now on we call Γ) acts as a discrete group on $\tilde{M}$. Modulo some ridiculous technicalities about boundaries of fundamental domains, and so on, we now identify M with a fundamental domain of $\Gamma/\tilde{M}$ and regard field theories on M as field theories on $\tilde{M}$ obeying certain conditions to be described shortly.

Our basic postulate is that since the fundamental domain of a space $\tilde{M}$ factored by a discrete group Γ is not unique, changing fundamental domains is a symmetry operation. Accordingly, the action over arbitrary space–time volumes $E \subset \tilde{M}$ is unaffected by the action of an element of Γ; thus

$$\int_E \mathscr{L}\, \mathrm{d}\mu = \int_{\gamma E} (\gamma \mathscr{L})(\gamma \mathrm{d}\mu). \tag{1}$$

Now $\mathrm{d}\mu$ is a positive definite measure on $\tilde{M}$ so $\gamma \mathrm{d}\mu$ and $\mathrm{d}\mu$ must differ at worst by a positive real multiplier:

$$(\gamma \mathrm{d}\mu)_x = f_\gamma(x)\, \mathrm{d}\mu, \qquad \forall \gamma \in \Gamma. \tag{2}$$

The group properties of Γ now imply

$$f_{\gamma_1\gamma_2}(x) = f_{\gamma_1}(x) f_{\gamma_2}(x). \tag{3}$$

Consideration of the cyclic subgroup generated by any $\gamma \in \Gamma$ now shows that either f is identically equal to one or that the relation between the measures grows without restriction (if the subgroup is of infinite order). We regard this latter possibility as physically unacceptable and conclude that the group action is measure preserving.

Hence, immediately,

$$\gamma \mathscr{L} = \mathscr{L}(\gamma x) = \mathscr{L}(x), \tag{4}$$

which is the basic symmetry re-expressed at the Lagrangian level.

We could just as well have started from equation (4) as from (1) but (1) gives measure invariance for free.

What happens next depends on what is in the Lagrangian. Note that the symmetry (4) is not something that is evident from the form of $\mathscr{L}$ either explicitly or implicitly as occurs with normal symmetries of the Lagrangian, but something that is impressed from outside. In other words it is an imposed symmetry. The reason for this is of course clear. $\mathscr{L}$ can just as well be a Lagrangian for $T \otimes \tilde{M}$ as for $T \otimes \Gamma/\tilde{M}$ (provided that it is a local Lagrangian), and as a Lagrangian for $T \otimes \tilde{M}$ it need not satisfy (4).

Suppose we now have a mass term like $m(\phi^\dagger \phi)$ in $\mathscr{L}$ where ϕ is (for the sake of definiteness) a complex vector-valued quantity. Symmetry (4) then says that

$$\phi_\alpha(\gamma x) = a_{\alpha\beta}(\gamma, x) \phi_\beta(x), \tag{5}$$

where $a(\gamma, x)$ is a member of the matrix group that preserves the quadratic form $\phi^\dagger\phi$. The group properties of Γ then give us

$$a_{\alpha\beta}(\gamma_1\gamma_2, x) = a_{\alpha\lambda}(\gamma_1, \gamma_2 x)a_{\lambda\beta}(\gamma_2, x). \tag{6}$$

(Note that this is the opposite convention to that of Dowker and Banach (1978), equation (3)) because we have written the action of Γ on the left rather than on the right. Assuming associativity then makes the action of Γ a representation rather than an anti-representation. This is more convenient in the following when we use representation theory. In one dimension it makes no difference since all the $a(\gamma)$ are simply numbers.) Fixing x, the $a(\gamma, \gamma' x)$ with the co-cycle rule (6) give us a representation of the action of Γ in the orbit of x under Γ into the group of isometries of the fibre of the ϕ field. Of course, ϕ may be coupled to things other than its own adjoint. However the fact that $\mathscr{L}$ is a scalar means that such an object will transform like $\phi^\dagger$ and so the above remarks apply just as well in this case.

We have said nothing of derivatives yet. Derivatives of ϕ will not transform in an invariant manner if we allow $a(\gamma, x)$ to depend non-trivially on x. In this case we are forced to consider only gauge theories and must introduce appropriate gauge-covariant connection fields. This procedure leads us directly to the type of theory discussed by Isham (1978a, b) and it is gratifying to see that our considerations lead to similar conclusions although the starting point is different. There is, however, another possibility. Among all the representations (6) there is a particular distinguished class, namely, those satisfying

$$a(\gamma, x) = a(\gamma), \qquad \forall x. \tag{7}$$

If we restrict ourselves to gauge transformations of this type we find two advantages. First, we do not need to alter the derivatives in $\mathscr{L}$ (provided they are already suitable geometric objects) and, second, the representation of Γ in (7) is a representation in the conventional sense and we can use standard Frobenius–Schur representation theory. The distinction between the two types of theory is similar to that made between gauge theories proper and the earlier rigid gauge theories. In fibre bundle language, a gauge theory defines an equivalence class of bundles, the equivalence relation being bundle diffeomorphism or 'equivalence' in the terminology of Steenrod (1951); while a rigid gauge theory, with a smaller symmetry group, defines a smaller equivalence class of bundles corresponding to Steenrod's 'strict equivalence'. For reasons of calculational expediency we stick to the latter type of theory for the remainder of this paper. Incidentally, it may be pointed out that for discrete gauge groups (specifically, for real scalar fields) the two types of theory coincide since in this case $a(\gamma, x)$ must be locally constant anyway.

There is one more matter which we have to deal with before we can say that our treatment is satisfactory for higher spin fields, and that is the transformation properties of the fields. A higher spin field is a field of n-tuples obeying the appropriate transformation law under change of orthonormal tetrad, so it is not so much a function of position as a function of tetrad field. Thus, given an n-tuple of complex fields (as we essentially had in the preceding discussion) which we want to regard as a higher spin field, we need to be able to display in some relative sense the frame that the value of the n-tuple refers to at each point. Topological objections notwithstanding, we can then extend this to all the other frames by the action of the transformation group of the field. So the problem is reduced to that of fixing the action of Γ on some given tetrad field. Let this field be $\{X_a\}$ (a is a bein index). Now the action of Γ is a set of diffeomorphisms of $\tilde{M}$

and so is a set of regular maps allowing extensions to the tangent bundle. The natural thing to do is to demand that this action takes the X_a into themselves. More explicitly, if F_x is the ring of germs of C^∞ functions in some open neighbourhood of x, we demand that

$$\gamma^* X_a = X_a, \qquad \forall a, \tag{8}$$

where $\gamma^* X(f) = X(\gamma_* f)$ and $\gamma_* f(x) = f(\gamma x), \forall f \in F_{\gamma x}$. Since the X_a are also orthonormal, this action guarantees that inner products are preserved so that Γ acts as an isometry. This is something which we might well want to demand at the outset, since the action of Γ is a symmetry operation, but we make the point that it is an additional assumption and does not automatically follow from the presence of Γ as a discrete group acting on $\tilde{M}$ as does the measure invariance.

Depending on circumstances, it may happen that we cannot find an $\{X_a\}$ such that (8) is satisfied. In this case we can either abandon all hope of building a viable theory or we can proceed in a more intuitive manner. An example of this latter process will arise in the following paper when we consider spin on a Klein bottle space–time.

3. General properties of automorphic fields and operators

Automorphic fields on $\tilde{M}$ are those that satisfy (5) with the $a(\gamma)$ constant. An equivalent characterisation arises by considering the projection operator $\phi \to \overset{\leftarrow a}{\phi}$, where

$$\overset{\leftarrow a}{\phi}_\alpha(x) = \frac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} a(\gamma^{-1})_{\alpha\beta} \phi_\beta(\gamma x). \tag{9}$$

We easily see that ϕ is automorphic if and only if $\phi = \overset{\leftarrow a}{\phi}$. Now consider the integral over $\tilde{M}$ of the inner product of an automorphic and a non-automorphic field,

$$\begin{aligned}\int_{\tilde{M}} f^\dagger(x) \overset{\leftarrow a}{\phi}(x)\, \mathrm{d}\mu_x &= \frac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} \int_{\tilde{M}} f^\dagger(\gamma x) \overset{\leftarrow a}{\phi}(\gamma x)\, \mathrm{d}\mu_x \\ &= \frac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} \int_{\tilde{M}} f^\dagger(\gamma x) a(\gamma) \overset{\leftarrow a}{\phi}(x)\, \mathrm{d}\mu_x \\ &= \int_{\tilde{M}} \overset{\rightarrow a}{f^\dagger}(x) \overset{\leftarrow a}{\phi}(x)\, \mathrm{d}\mu_x ,\end{aligned} \tag{10}$$

where

$$\overset{\rightarrow a}{f^\dagger_\alpha}(x) = \frac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} f^\dagger_\beta(\gamma x) a_{\beta\alpha}(\gamma) \tag{11}$$

is the adjoint projection operator. More generally, if there is a Green function sandwiched between the two fields, we find by exactly the same reasoning

$$\int_{\tilde{M}} \int_{\tilde{M}} \overset{\rightarrow a}{\phi_1}(x) K(x, y) \overset{\leftarrow a}{\phi_2}\, \mathrm{d}\mu_x\, \mathrm{d}\mu_y = \int_{\tilde{M}} \int_{\tilde{M}} \overset{\rightarrow a}{\phi_1}(x) \overset{\longleftrightarrow a}{K(x, y)} \overset{\leftarrow a}{\phi_2}(y)\, d\mu_x\, \mathrm{d}\mu_y, \tag{12}$$

where

$$\overset{\longleftrightarrow a}{K_{\alpha\beta}(x, y)} = \frac{1}{|\Gamma|^2} \sum_{\gamma, \gamma' \in \Gamma} a_{\alpha\mu}(\gamma^{-1}) K_{\mu\nu}(\gamma x, \gamma' y) a_{\nu\beta}(\gamma'). \tag{13}$$

In order that operation by Γ is still a symmetry, we must demand that translation by some γ and operation by K commute, which reveals that K must satisfy

$$K_{\alpha\beta}(\gamma x, y) = K_{\alpha\beta}(x, \gamma^{-1}y). \tag{14}$$

We can now see the difference between what was done in our earlier work and what appears here. Before, the $a(\gamma)$ were just numbers and commuted with K, and hence we had

$$\begin{aligned}
\overset{\leftrightarrow a}{K}_{\alpha\beta}(x, y) &= \frac{1}{|\Gamma|^2} \sum_{\gamma,\gamma'\in\Gamma} a_{\alpha\mu}(\gamma^{-1}) K_{\mu\nu}(\gamma x, \gamma' y) a_{\nu\beta}(\gamma') \\
&= \frac{1}{|\Gamma|^2} \sum_{\gamma,\gamma'\in\Gamma} K_{\alpha\mu}(x, \gamma^{-1}\gamma' y) a_{\mu\nu}(\gamma^{-1}) a_{\nu\beta}(\gamma') \\
&= \frac{1}{|\Gamma|^2} \sum_{\gamma,\gamma'\in\Gamma} K_{\alpha\mu}(x, \gamma^{-1}\gamma' y) a_{\mu\beta}(\gamma^{-1}\gamma') \\
&= \frac{1}{|\Gamma|} \sum_{\gamma\in\Gamma} K_{\alpha\mu}(x, \gamma y) a_{\mu\beta}(\gamma) \\
&= \overset{\rightarrow a}{K}_{\alpha\beta}(x, y),
\end{aligned} \tag{15}$$

or, quite analogously,

$$\overleftrightarrow{K_{\alpha\beta}(x, y)}^{a} = \frac{1}{|\Gamma|} \sum_{\gamma\in\Gamma} a_{\alpha\mu}(\gamma^{-1}) K_{\mu\beta}(\gamma x, y) = \overset{\leftarrow a}{K}_{\alpha\beta}(x, y). \tag{16}$$

Conversely, by reversing steps in the above, we see that a single projection is equivalent to a double projection if the $a(\gamma)$ commute with K, which is what we had in the scalar case (Dowker and Banach 1978 and other references mentioned there).

We can ask more generally under what circumstances just one projection suffices to make a linear operator K automorphic. The answer is readily obtained since we must then have

$$\overset{\leftarrow}{K} = \overset{\leftrightarrow}{K} \tag{17}$$

and

$$\vec{K} = \overset{\leftrightarrow}{K} \tag{18}$$

(we will drop the a-superscripts from now on), so that

$$\vec{K} = \overset{\leftrightarrow}{K} = \overset{\leftarrow}{K}. \tag{19}$$

On the other hand if we have

$$\vec{K} = \overset{\leftarrow}{K}, \tag{20}$$

then a projection gives

$$\vec{K} = \overset{\rightrightarrows}{K} = \overset{\rightleftarrows}{K} = \overset{\leftrightarrow}{K}, \tag{21}$$

so we arrive at (19) again and thus see that (20) represents a necessary and sufficient condition for one projection to be adequate.

As an application of the foregoing let us discuss the relationship between problem-solving on $\tilde{M}$ and on $\Gamma/\tilde{M}$. Suppose we are given the following operator equation on $\tilde{M}$ to describe some physical situation:

$$H - AH = \mathbb{1}. \tag{22}$$

Here A will be some given linear operator (since our treatment is only required to be invariant under rigid gauge transformations, A may quite well be a differential operator). What is the corresponding equation on $\Gamma/\tilde{M}$? There are two possibilities; namely (a) that we simply project (22) giving

$$\overleftrightarrow{H} - \overleftarrow{A}\overrightarrow{H} = \overset{\leftrightarrow}{\mathbb{1}}, \tag{23}$$

or (b) we demand that each constituent symbol in (22) be automorphic, yielding

$$\overleftrightarrow{H} - \overleftrightarrow{A}\overleftrightarrow{H} = \overset{\leftrightarrow}{\mathbb{1}}. \tag{24}$$

Now (24) is more natural in the sense that it is constructed from quantities which are well defined on $\Gamma/\tilde{M}$; however (23) is more convenient since it says that we should simply solve (22) (something which we may have done already on a previous occasion) and then project the solution. We can check that this is so since (24) is solved by

$$\overleftrightarrow{H} = (\overset{\leftrightarrow}{\mathbb{1}} - \overleftrightarrow{A})^{-1} \tag{25}$$

and (23) by

$$\overrightarrow{H} = \overrightarrow{(\mathbb{1} - A)}^{-1} \tag{26}$$

and

$$\overleftrightarrow{H} = \overleftrightarrow{(\mathbb{1} - A)}^{-1}, \tag{27}$$

as can be verified by writing out the appropriate Neumann series.

When are the two the same? Obviously when

$$\overleftrightarrow{A} = \overleftrightarrow{A}. \tag{28}$$

In general, comparison of the two methods of approach will give conditions like (17) and (18), so to cover all possibilities we can say that provided all our given quantities satisfy (20) then we can solve our equations by either method.

Suppose however that (20) is not satisfied. Can we then advance arguments in favour of either (23) or (24) other than the somewhat aesthetic ones presented above? Both after all yield correctly automorphic solutions. The answer is that we can, at least for theories possessing a Lagrangian. Thus let $\mathscr{L}$ be the Lagrangian for $\tilde{M}$ and S the action, regarded as functionals of arbitrary fields on $\tilde{M}$. Now extremise S with respect to automorphic variations of the fields giving

$$(\delta/\overleftarrow{\delta}\phi)S = 0, \tag{29}$$

where the operator $\delta/\overleftarrow{\delta}\phi$ is a partial derivative in function space and is a 'partial derivative in the direction of automorphic fields'. What we get thereby are classical solutions ϕ_c (not necessarily automorphic) which extremise S against automorphic variations $\phi_c \to \phi_c + \delta\phi$ where $\delta\phi = \overleftarrow{\delta\phi}$. Now project ϕ_c; $\phi_c \to \overleftarrow{\phi_c}$. In most cases the resulting $\overleftarrow{\phi_c}$ will no longer satisfy the Euler–Lagrange equations and so must be rejected. However some will and by the arguments at the beginning of this section will cause any operators occurring in $\mathscr{L}$ (e.g. $(\nabla_\mu \nabla^\mu + m^2)\delta(x, y)$) to become correctly automorphic and will hence lead to equations like (24) rather than (23) between them. The point of all this rather oblique discussion is to show that equations like (24) can be characterised by some criteria or other from among a wider set than themselves and for this reason are to be preferred to those like (23).

Actually there is a surprising parallel between what we have done here and the customary gauge-fixing procedure in non-Abelian gauge theories since, despite the fact that (24) is the correct equation, we may for reasons of expediency choose to work via (23) (in the non-Abelian case, expediency means avoiding the infinite volume of the gauge group). We would then be forced to compensate for the error in some way (e.g. by putting a term like $\overleftrightarrow{A}\overleftrightarrow{H} - \overleftarrow{A}\overrightarrow{H}$ into (23)). The non-Abelian analogue of this would be the addition of Fadeev–Popov ghosts to the Lagrangian to compensate for gauge-fixing terms added 'by hand'.

As a final remark on this matter we extend the notation introduced in (29). We let $\overleftarrow{\delta}$ in a functional derivative mean 'automorphic partial functional derivative' as above, and $\overleftarrow{\phi}$ mean that only automorphic functions are to be differentiated; thus $(\delta/\overleftarrow{\delta}\phi)$ is as before, $(\delta/\delta\overleftarrow{\phi})$ would mean arbitrary variations of automorphic functions and $(\delta/\overleftarrow{\delta}\overleftarrow{\phi})$ would correspond to the intrinsic functional derivative on $\Gamma/\tilde{M}$.

We now turn our attention to a different question. Given a linear operator K as before, is there some way of splitting K up into K_1 and K_2 so that

$$K = K_1 + K_2, \tag{30}$$

$$\overleftrightarrow{K} = \overleftrightarrow{K}_1, \tag{31}$$

$$\overleftrightarrow{K}_2 = 0 \tag{32}$$

and, if possible, so that K_1 has a relatively simple form? Such a process obviously holds a lot of practical advantages as a short cut in calculations. The affirmative answer is given by the following manipulation (we have dropped the indices on K and a),

$$\begin{aligned}\overleftrightarrow{K(x, y)} &= \frac{1}{|\Gamma|^2} \sum_{\gamma, \gamma' \in \Gamma} a(\gamma^{-1}) K(\gamma x, \gamma' y) a(\gamma') \\ &= \frac{1}{|\Gamma|^2} \sum_{\gamma, \gamma' \in \Gamma} a(\gamma^{-1}) K(x, \gamma^{-1}\gamma' y) a(\gamma') \\ &= \frac{1}{|\Gamma|^2} \sum_{\gamma, \gamma'' \in \Gamma} a(\gamma^{-1}) K(x, \gamma'' y) a(\gamma) a(\gamma'') \\ &= (\overrightarrow{K_C})(x, y),\end{aligned} \tag{33}$$

where

$$K_C(x, y) = \frac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} a(\gamma^{-1}) K(x, y) a(\gamma). \tag{34}$$

Note that K_C satisfies (14) and so is an operator of the type considered. Furthermore K_C commutes with all the $a(\gamma)$ and automatically satisfies (20) so that we can simplify (31) to a condition like (19), a fact expressed by (33).

Since K_C commutes with the $a(\gamma)$ we can use Schur's lemma to find its form. Thus let R be a matrix that reduces the representation $a(\Gamma)$ to its irreducible constituents

$$a(\gamma)=R^{-1}[u_1(\gamma)\oplus u_2(\gamma)\oplus\ldots\oplus u_m(\gamma)]R, \qquad \forall\gamma\in\Gamma, \tag{35}$$

$$\dim u_j(\Gamma)=n_j, \tag{36}$$

$$\sum_j n_j = n = \dim a(\Gamma), \tag{37}$$

and furthermore assume that any of the u_j that are equivalent are actually the same (as is always possible).

Let

$$K_R = RKR^{-1} \tag{38}$$

and

$$K_{CR}=RK_CR^{-1}; \tag{39}$$

then (34) becomes

$$K_{CR}=\frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}\Big\{\bigoplus_j u_j(\gamma^{-1})\Big\}K_R\Big\{\bigoplus_k u_k(\gamma)\Big\}. \tag{40}$$

If now K_{CRij} and K_{Rij} are the sub-blocks in the ith row of blocks and the jth column of blocks (dimensions being given by (36)) we get

$$K_{CRij}=\frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}u_i(\gamma^{-1})K_{Rij}u_j(\gamma). \tag{41}$$

Thus the K_{CRij} intertwine the ith and the jth representations, and by Schur's lemma and standard arguments

$$K_{CRij}=(n_in_j)^{-1/2}(\operatorname{Tr}K_{Rij})\mathbb{1}_{ij}\delta_{i\sim j}, \tag{42}$$

where

$$\delta_{i\sim j}=\begin{cases}1 & \text{if } u_i\equiv u_j\\ 0 & \text{otherwise}\end{cases} \tag{43}$$

and $\mathbb{1}_{ij}$ is the unit matrix where appropriate. Hence

$$K_C(x,y)=R^{-1}\{(n_in_j)^{-1/2}(\operatorname{Tr}K_{Rij}(x,y))\mathbb{1}_{ij}\delta_{i\sim j}\}R. \tag{44}$$

If ϕ is a multiplet of real fields (invariant under the real orthogonal group) then the derivation above breaks down if the K_{Rij} fail to have at least one real eigenvalue since Schur's lemma is no longer true in this case. This is not something we will be faced with in this paper.

Up to now we have tacitly assumed that $|\Gamma|<\infty$, in which case all of our above deductions are unquestionably well defined. However if $|\Gamma|$ is infinite almost everything becomes meaningless as it stands. There are two ways to try to remedy this. The first is the most obvious, namely, to fix upon a procedure for evaluating 'partial sums' and take limits. On a technical level, this amounts to introducing additional topological structure on Γ, namely, a well-ordered (by inclusion) sub-topology of the co-finite topology of Γ. Thus we have a sequence of closed sets

$$\{\emptyset \subset \Gamma_1 \subset \Gamma_2 \subset \Gamma_3 \subset \ldots \subset \Gamma\}, \qquad |\Gamma_n| < \infty, \tag{45}$$

and can define a projection as the continuous limit (at the empty set) of the partial projections on Γ_n thus:

$$\frac{1}{|\Gamma|} \sum_{\gamma\in\Gamma} a(\gamma^{-1})\phi(\gamma x) \stackrel{\text{def}}{=} \lim_{\Gamma_n\to\Gamma} \frac{1}{|\Gamma_n|} \sum_{\gamma\in\Gamma_n} a(\gamma^{-1})\phi(\gamma x). \tag{46}$$

Since the series will not in general be absolutely convergent, the measure on Γ defined by the above procedure will not satisfy the usual group-theoretically desirable criteria such as translation invariance and so is of little practical value.

An alternative procedure is as follows. By the arguments at the beginning of the section, integrals involving automorphic fields and non-automorphic fields can be replaced by integrals involving automorphic fields alone. Doing this backwards we see that if all operators are automorphic, the fields are also effectively automorphic. Thus we reduce the problem to finding a good definition for (13).

Note now that one of the $|\Gamma|^{-1}$ factors in (13) belongs more properly with the integral over $\tilde{M}$ if we want to regard $\vec{K}$ as an operator on M, since

$$\int_{\tilde{M}} K(x, y)\tilde{\phi}(y)\, \mathrm{d}\mu_y = \int_M \left[\sum_{\gamma\in\Gamma} K(x, \gamma y)a(\gamma)\right]\tilde{\phi}(y)\, \mathrm{d}\mu_y. \tag{47}$$

For this to make sense we must have some asymptotic condition on $K(x, y)$ like

$$K(x, \gamma y) \sim \mathrm{O}[(|\Gamma_n| - |\Gamma_{n-1}|)^{-2}], \qquad \gamma \in \Gamma_n. \tag{48}$$

Since the operator in (47) is O(1) everywhere we still have to be careful in making the other projection; thus we have to define

$$K_M(x, y) = \frac{1}{|\Gamma|} \sum_{\gamma,\gamma'\in\Gamma} a(\gamma^{-1})K(\gamma x, \gamma' y)a(\gamma'). \tag{49}$$

We can write this as

$$K_M(x, y) = \sum_{\gamma\in\Gamma} K_C(x, \gamma y)a(\gamma) \tag{50}$$

and take K_C to be given by (44) *by definition*. Since no explicit factors of $|\Gamma|$ remain we adopt this as a satisfactory definition of K_M.

Despite the difficulties above, we continue to use $\overleftrightarrow{K}$ since the number of $|\Gamma|$ factors and summations over Γ are equal and remain so throughout all manipulations, making things a bit easier to follow.

We next establish the behaviour of our automorphic objects under translation by $\gamma \in \Gamma$ which is supposed to be a symmetry. We easily see that if we set

$$\phi_\gamma(x) = \phi(\gamma x) = a(\gamma)\phi(x), \tag{51}$$

then

$$\phi_\gamma(\gamma' x) = a(\gamma)\phi(\gamma' x) = a(\gamma)a(\gamma')a(\gamma^{-1})\phi(\gamma x) = a(\gamma\gamma'\gamma^{-1})\phi_\gamma(x), \tag{52}$$

so ϕ_γ is automorphic by an inner automorphism of $a(\Gamma)$. Thus the operation of translation by Γ permutes the various automorphic fields amongst each other.

Our final topic for this section will be the divergence theorem. We assume that $a(\Gamma)$ commutes with derivatives. Then the quantity F_μ,

$$F_\mu = \psi^\dagger \nabla_\mu \phi, \tag{53}$$

is a Γ-invariant 1-form and so may be regarded as a 1-form on M provided ψ and ϕ are automorphic. If we now assume that M is orientable and compact and ∂M is empty then the surface integral of F_μ vanishes since there is no surface. On the other hand when we lift the 1-form back to a fundamental domain of $\Gamma/\tilde{M}$ we *do* find a surface, which is the difference between the closure and interior of $\Gamma/\tilde{M}$. It will definitely be non-empty since $\partial(\Gamma/\tilde{M}) = \varnothing$ requires that $\Gamma/\tilde{M}$ be both open and closed and so is a component of $\tilde{M}$ and so is equal to $\tilde{M}$ ($\tilde{M}$ simply connected) contradicting the non-triviality of Γ. The precise nature of $\partial(\Gamma/\tilde{M})$ is in general the realm of the ridiculous technicalities we undertook to avoid in § 2 but the 'physical reasonableness' criterion prevents any awkward problems in practice. Thus

$$\int_{\partial(\Gamma/\tilde{M})} F_\mu \, \mathrm{d}\sigma^\mu = 0. \tag{54}$$

This usually arises because $\partial(\Gamma/\tilde{M})$ can be split into two sets, say A and B, which are disjoint and which are carried into each other by the action of Γ,

$$\gamma_x x \in B, \qquad \forall x \in A \text{ for some } \gamma_x \in \Gamma \tag{55}$$

(γ_x can vary with x in a reasonable manner), and $\mathrm{d}\sigma^\mu$ satisfies

$$(\gamma_x)_* \, \mathrm{d}\sigma^\mu = -\mathrm{d}\sigma^\mu. \tag{56}$$

The invariance of F_μ then gives (54). The obvious example of this process is any torus.

Having established (54) we apply the divergence theorem and obtain

$$\int_{\Gamma/\tilde{M}} \nabla^\mu F_\mu \, \mathrm{d}\mu_x = \int_{\Gamma/\tilde{M}} (\nabla^\mu \vec{\psi}^\dagger)(\nabla_\mu \overleftarrow{\phi}) \, \mathrm{d}\mu_x + \int_{\Gamma/\tilde{M}} \vec{\psi}^\dagger (\nabla^\mu \nabla_\mu) \overleftarrow{\phi} \, \mathrm{d}\mu_x = 0. \tag{57}$$

Obvious generalisations of this result enable us to recast the problem of finding automorphic solutions to equations like

$$-\nabla^\mu [p(x) \nabla_\mu \phi] + q(x)\phi = \lambda \phi \tag{58}$$

into a Sturm–Liouville-type problem with automorphic boundary conditions on $\Gamma/\tilde{M}$. This technique has already been used by Avis and Isham (1978) in their discussion of two-phase ϕ^4 theory on $T \otimes S^1$.

If we want to relax the orientability stipulation then the above no longer holds. Although M is not orientable $\tilde{M}$ is (the proof of this statement is a lengthy business) and so $\Gamma/\tilde{M}$ inherits an orientation from $\tilde{M}$. However the induced surface form on $\partial(\Gamma/\tilde{M})$ will no longer satisfy (56) and so the surface integral will not vanish. Viewed as a phenomenon on M we find 'surface integrals without surfaces' (cf Wheeler's charge without charge!). On M this is explained by the fact that the volume form can no longer be smooth everywhere.

4. The spectral decomposition of self-adjoint automorphic operators

For definiteness we now fix $a(\Gamma)$ to be a subgroup of a unitary group. Then, given K, it is natural to define the adjoint of K by

$$(K^\dagger)_{\alpha\beta}(x, y) = K^*_{\beta\alpha}(y, x) \qquad (\text{* is conjugation}). \tag{59}$$

A simple manipulation then shows that

$$\overrightarrow{(K^\dagger)} = (\overleftarrow{K})^\dagger \qquad \text{and} \qquad \overleftarrow{(K^\dagger)} = (\overrightarrow{K})^\dagger. \tag{60}$$

Thus if K is self-adjoint,

$$K^\dagger = K, \tag{61}$$

we see that

$$\overleftrightarrow{K} = \overleftrightarrow{(K^\dagger)} = \overleftarrow{[\overrightarrow{(K^\dagger)}]} = \overleftarrow{[(\overleftarrow{K})^\dagger]} = \overrightarrow{[(\overleftarrow{K})]}^\dagger = (\overleftrightarrow{K})^\dagger \tag{62}$$

and $\overleftrightarrow{K}$ is too (we could have verified this from (13) directly but (62) is quite amusing).

Also if K is self-adjoint then we can restate (20) in a different way, namely

$$\overrightarrow{K} = \overleftarrow{K} \Leftrightarrow \overrightarrow{K} = (\overrightarrow{K})^\dagger \qquad \text{or} \qquad \overleftarrow{K} = (\overleftarrow{K})^\dagger, \tag{63}$$

which follows from various rearrangements of the equality

$$\overleftarrow{K} = \overleftarrow{(K^\dagger)} = (\overrightarrow{K})^\dagger = \overrightarrow{K} \qquad \text{or} \qquad \overleftarrow{K} = (\overleftarrow{K})^\dagger = \overrightarrow{(K^\dagger)} = \overrightarrow{K}. \tag{64}$$

We now move on to the main result of this section which describes how eigenfunction expansions of linear operators K behave under the projection operators.

Spectral projection theorem. Let K be a self-adjoint operator satisfying (14); then the two following conditions are equivalent.

(1) $\overrightarrow{K} = \overleftarrow{K}$.

(2) K has an eigenfunction expansion $K = \Sigma_i \lambda_i \psi_i(x)\psi_i^*(y)$ such that each ψ_i satisfies

$$\overleftarrow{\psi_i} = \psi_i \qquad \text{or} \qquad 0. \tag{65}$$

Proof. (1)$\Rightarrow$(2). By ordinary spectral theory, K will have an eigenfunction expansion

$$K_{\alpha\beta}(x, y) = \sum_i \lambda_i \psi_{i,\alpha}(x)\psi^*_{i,\beta}(y). \tag{66}$$

Hence

$$\overleftarrow{K}_{\alpha\beta}(x, y) = \sum_i \lambda_i \overleftarrow{\psi}_{i,\alpha}(x)\psi^*_{i,\beta}(y) \tag{67}$$

and

$$\overrightarrow{K}_{\alpha\beta}(x, y) = \sum_i \lambda_i \psi_{i,\alpha}(x)(\overleftarrow{\psi})^*_{i,\beta}(y). \tag{68}$$

We assume now that our space of functions (fields) is normed in the usual manner,

$$\|\phi\| = \left[\int_{\tilde{M}} \phi^{\dagger} \phi \, \mathrm{d}\mu \right]^{1/2}, \tag{69}$$

in which case liberal use of the Cauchy and Minkowsky inequalities applied to (9) yields

$$\|\overleftarrow{\phi}\| \leqslant \|\phi\|, \tag{70}$$

so we can expand the $\overleftarrow{\psi}_i$ in terms of the ψ_j. Thus

$$\overleftarrow{\psi}_{i,\alpha}(x) = \sum_j P_{ij} \psi_{j,\alpha}(x), \tag{71}$$

and since the ψ_i are orthonormal, we have

$$\sup_i \left[\sum_j |P_{ij}|^2 \right] < \infty. \tag{72}$$

We can find P_{ij} by noting that (formally at least) the δ function

$$\delta_{\alpha\beta}(x, y) = \delta_{\alpha\beta} \delta(x, y) = \sum_i \psi_{i,\alpha}(x) \psi^*_{i,\beta}(y) \tag{73}$$

satisfies (14). Hence

$$P_{ij} = \frac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} \int_{\tilde{M}} \psi^*_{j,\alpha}(y) a_{\alpha\beta}(\gamma^{-1}) \psi_{i,\beta}(\gamma y) \, \mathrm{d}\mu_y. \tag{74}$$

We easily see that

$$P_{ij} = P^*_{ji}, \tag{75}$$

whence

$$\sup_j \left[\sum_i |P_{ij}|^2 \right] < \infty. \tag{76}$$

Introducing Λ_{ij} by

$$\Lambda_{ij} = \delta_{ij} \lambda_j \qquad \text{(no sum)}, \tag{77}$$

we can rewrite (67) and (68) as

$$\overleftarrow{K}_{\alpha\beta}(x, y) = \sum_{i,j,k} \psi_{i,\alpha}(x) P^*_{ij} \Lambda_{jk} \psi^*_{k,\beta}(y) \tag{78}$$

and

$$\overrightarrow{K}_{\alpha\beta}(x, y) = \sum_{i,j,k} \psi_{i,\alpha}(x) \Lambda_{ij} P^*_{jk} \psi^*_{k,\beta}(y), \tag{79}$$

so equality and independence of the ψ_i give

$$(P^* \Lambda)_{ik} = (\Lambda P^*)_{ik} \tag{80}$$

or

$$P^*_{ik} \lambda_k = \lambda_i P^*_{ik} \qquad \text{(no sums)}. \tag{81}$$

So P_{ik} is zero unless $\lambda_i = \lambda_k$ and the projected eigenfunctions remain within their own eigenspace.

Now P is a projection, so

$$P^2 = P, \tag{82}$$

and this is well defined by (72) and (76). So we can diagonalise P by a unitary transformation whence its eigenvalues are 0 or 1. We can use the same unitary transformation to display the basis of eigenfunctions satisfying (65).

(2)$\Rightarrow$(1). If we have a basis satisfying (65) then, whichever projection we impose on K, each dyadic in (66) either disappears or remains unaffected; hence the two projections yield the same result. This completes the proof.

It will not have escaped the reader's attention that the above 'proof' is in reality mathematically worthless. To obtain something valid we would have to go into the details of what sort of spectrum K has, the precise nature of the function space (if $|\Gamma| = \infty$ we get particularly acute troubles since an automorphic function can no longer be square integrable) and we have to straighten out the rather dubious use of the δ function. This paper is no place for such things. Let us content ourselves by saying that at least in the case that $\tilde{M}$ is compact, Γ finite and K the inverse of some elliptic operator, there are no serious difficulties and we may take the result as true.

Before moving on we can make some more comments about P_{ij}. Since the action of Γ and that of K commute, we see that all translates of eigenfunctions are likewise eigenfunctions. Define $\hat{\gamma}$ by

$$(\hat{\gamma}\phi)(x) = \phi(\gamma^{-1}x); \tag{83}$$

then the map $\gamma \to \hat{\gamma}$ is a representation of Γ in the unitary operators on the ϕ function space. Restricting this action to a finite dimensional eigenspace of K we find a matrix representation of Γ, i.e.

$$\hat{\gamma}\psi_{i,\alpha}(x) = \psi_{i,\alpha}(\gamma^{-1}x) = u_{ji}(\gamma)\psi_{j,\alpha}(x). \tag{84}$$

Thus

$$\begin{aligned} P_{ij} &= \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma}\int_{\tilde{M}} \psi^*_{j,\mu}(y)a_{\mu\nu}(\gamma)u_{ki}(\gamma)\psi_{k,\nu}(y)\,\mathrm{d}\mu_y \\ &= \int_{\tilde{M}} \psi^*_{j,\mu}(y)T_{i\mu,k\nu}\psi_{k,\nu}(y)\,\mathrm{d}\mu_y, \end{aligned} \tag{85}$$

where

$$T_{i\mu,k\nu} = \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma} a_{\mu\nu}(\gamma)u_{ki}(\gamma). \tag{86}$$

We cannot say much about $T_{i\mu,k\nu}$ in general but we can give a condition on whether it is zero or not. T is a set of inner products in the group ring between the trivial representation and the set of coefficient functions of the tensor product of the u and a representations. This will be identically zero if and only if the trivial representation is not contained in $a\otimes u$. Taking traces we find

$$T \equiv 0 \Leftrightarrow \langle \mathrm{tr}(a\otimes u), 1\rangle = 0 \Leftrightarrow \langle (\mathrm{tr}\ a)(\mathrm{tr}\ u), 1\rangle = 0. \tag{87}$$

The quantity $(\mathrm{tr}\ a)$ $(\mathrm{tr}\ u)$ can itself be regarded as an inner product in character space between the character of a and the character of the conjugate of u (or *vice versa*) and so an equivalent criterion for the vanishing of T is that the a and u^* representations have no common irreducible components.

A useful by-product of the spectral projection theorem is the fact that it allows an abstract treatment of the relationship between the image and mode sum methods of calculating the automorphic analogue of a given self-adjoint linear operator. The image expression corresponds to the projected operator while the mode sum is its eigenfunction expansion. If (20) is satisfied there is no problem. The theorem asserts that (in an appropriate basis) we can simply select the correctly automorphic modes with their corresponding eigenvalues and the resulting expression is equal to the image sum. If (20) is not satisfied, then (as usual) the situation is not so simple. The difficulty arises because the mode sum is no longer a well-defined concept. We cannot find a basis in which (65) is true and so we cannot in general answer the question 'Is a particular ψ_i automorphic or not?' Although the projected $\overleftarrow{\psi}_i$ will in general be non-zero, it will not be of norm 1 and so cannot qualify as an eigenfunction. In this sense we see that the image sum method is the more fundamental of the two.

If we persist in wanting to write a mode sum expression for $\overleftrightarrow{K}$ we can always do so by considering K_C (which will also be self-adjoint if K is). Its modes will not in general correspond to those of K but it does satisfy (20). The point is that there are many K which give the same $\overleftrightarrow{K}$ (all differing by operators in the kernel of the double projection), some of which satisfy (20) and others which do not. K_C is just a particularly convenient one.

We finally point out that if $a(\Gamma)$ is essentially a one-dimensional representation (i.e. $a(\gamma)_{\alpha\beta} \equiv a(\gamma) \otimes \mathbb{1}_{\alpha\beta}$) then we always have (20) and so can use the eigenfunctions of K directly. This illustrates the special position held by these representations in both this and other problems.

5. Traces and Selberg's trace formula

Various physical quantities can be expressed as traces of appropriate linear operators. We have in mind particularly the Zeta functions introduced recently for the evaluation of the effective Lagrangian and action (Dowker and Critchley 1976) and their fully integrated counterparts.

Our deductions will be aimed at evaluating the trace of a general linear operator $\overleftrightarrow{K}$ and will feature the Selberg trace formula (Selberg 1956) (still the clearest exposition),† an object of almost irresistible attraction to mathematicians these days. First of all we assume $\overleftrightarrow{K}$ has a trace (this is a non-trivial problem in general). Then

$$
\begin{aligned}
\operatorname{tr} \overleftrightarrow{K} &= \int_{\tilde{M}} \operatorname{tr} \overleftrightarrow{K}(x, x)\, \mathrm{d}\mu_x \\
&= \int_M \operatorname{tr} K_M(x, x)\, \mathrm{d}\mu_x \\
&= \int_M \operatorname{tr}\Big(\sum_{\gamma\in\Gamma} a(\gamma) K_C(x, \gamma x)\Big)\, \mathrm{d}\mu_x \\
&= \int_M \operatorname{tr}\Big(\sum_{\gamma\in\Gamma} [u_1(\gamma) \oplus u_2(\gamma) \oplus \ldots \oplus u_m(\gamma)]\,[(n_i n_j)^{-1/2} (\operatorname{tr} K_{Rij}(x, \gamma x))\mathbb{1}_{ij}\delta_{i\sim j}]\Big)\, \mathrm{d}\mu_x \\
&= \sum_i \sum_{\gamma\in\Gamma} \operatorname{tr}(u_i(\gamma)) \int_M \operatorname{tr} K_{Rii}(x, \gamma x)\, \mathrm{d}\mu_x.
\end{aligned} \tag{88}
$$

† See also Hurt (1976).

We are now in a position to apply the crux of Selberg's argument. Choose some fixed conjugacy class in Γ, $\{\gamma\}$ say, γ being a fixed member of it. Then the $\mathrm{tr}(u_i(\gamma'))$ are constant over $\{\gamma\}$ and are all equal to $\mathrm{tr}(u_i(\gamma))$. Now (writing f for K_{Rii})

$$\sum_{\gamma'\in\{\gamma\}} \int_M f(x, \gamma' x)\, \mathrm{d}\mu_x = \frac{1}{|\Gamma_\gamma|} \sum_{\gamma_1\in\Gamma} \int_M f(x, \gamma_1\gamma\gamma_1^{-1}x)\, \mathrm{d}\mu_x \tag{89}$$

because all elements of $\{\gamma\}$ are of the form $\gamma_1\gamma\gamma_1^{-1}$ ($\gamma_1 \in \Gamma$) and two different γ_1 yield the same member of $\{\gamma\}$ if and only if they differ by an element of Γ_γ (the centraliser of γ i.e. the subgroup of Γ that commutes with γ). Hence the sum over all $\gamma_1 \in \Gamma$ will give an answer $|\Gamma_\gamma|$ times too big. Also

$$\begin{aligned}\frac{1}{|\Gamma_\gamma|} \sum_{\gamma_1\in\Gamma} \int_M f(x, \gamma_1\gamma\gamma_1^{-1}x)\, \mathrm{d}\mu_x &= \frac{1}{|\Gamma_\gamma|} \sum_{\gamma_1\in\Gamma} \int_M f(\gamma_1^{-1}x, \gamma\gamma_1^{-1}x)\, \mathrm{d}\mu_x \\ &= \frac{1}{|\Gamma_\gamma|} \sum_{\gamma_1\in\Gamma} \int_{\gamma_1^{-1}M} f(x, \gamma x)\, \mathrm{d}\mu_x \\ &= \frac{1}{|\Gamma_\gamma|} \int_{\tilde{M}} f(x, \gamma x)\, \mathrm{d}\mu_x. \end{aligned}\tag{90}$$

Summing over conjugacy classes in Γ, (88) yields

$$\mathrm{tr}\, \overleftrightarrow{K} = \sum_i \sum_{\{\gamma\}} (\mathrm{tr}\, u_i(\gamma)) \frac{1}{|\Gamma_\gamma|} \int_{\tilde{M}} \mathrm{tr}\, K_{Rii}(x, \gamma x)\, \mathrm{d}\mu_x. \tag{91}$$

Note that if $|\{\gamma\}|$ is the number of elements in the $\{\gamma\}$ conjugacy class then

$$|\Gamma| = |\{\gamma\}|\, |\Gamma_\gamma| \tag{92}$$

so that we do not actually need to know Γ_γ to use (91). Knowledge of the conjugacy classes is sufficient.

Equation (91) is essentially Selberg's trace formula adapted to the case where (20) is not satisfied (Selberg's original formula was only for the case where K is a scalar i.e. $K_{\alpha\beta}(x, y) \equiv k(x, y) \otimes \mathbb{I}_{\alpha\beta}$). Selberg actually went further than (91) (which is problematic if $|\Gamma_\gamma|$ is infinite): he showed that

$$\frac{1}{|\Gamma_\gamma|} \int_{\tilde{M}} f(x, \gamma x)\, \mathrm{d}\mu_x = \mu_{G_\gamma}(G_\gamma/\Gamma_\gamma) \int_{\tilde{M}} f(x, \gamma x) p_\gamma(x)\, \mathrm{d}\mu_x, \tag{93}$$

where G is the full group of isometries of $\tilde{M}$ (Γ being a subgroup), G_γ is the centraliser of γ in G, Γ_γ is the centraliser of γ in Γ, μ_{G_γ} is the right Haar measure on G_γ and $p_\gamma(x)$ satisfies

$$\int_{n\in G_\gamma} p_\gamma(nx)\, \mathrm{d}\mu_{G_\gamma} = 1, \qquad \forall x \in \tilde{M}. \tag{94}$$

On a formal level, (93) is not difficult to show (we do not do it here but refer the reader to Selberg's paper for details). Selberg's achievement was to show that the effect of substituting (93) in (91) is to produce a formula valid at all times without any restriction other than the existence of the trace and finiteness of the volume of M.

Our main concern is to present (91) as a labour saving device. Knowledge of all the necessary bits and pieces required means fewer integrals to evaluate. There are however two further points worth making. Firstly, if K is self-adjoint then we know that the trace is simply the sum of the eigenvalues and the same will be true of $\overleftrightarrow{K}$. In the general case one sum is not a sub-sum of the other. This fact and the fact that we have to unscramble $a(\Gamma)$ into its irreducible components is reflected by the implicit presence of R in (91), a complication absent from Selberg's original treatment. Secondly we notice that (91) enables us to display a nice functorial property of the trace function on automorphic linear operators. Since by definition the trace is a sum of terms like $\int f(x, \gamma x)\, \mathrm{d}\mu$ it clearly factors through the linear functionals on the group ring. What (91) says is that this factorisation can be further factored through the linear functionals on the character functions. Since the character functions generate the centre of the group ring this says that the trace factors through the functionals on the centre of the group ring. In words of one syllable we are saying that the trace is a function of conjugacy class rather than group.

6. Zeta functions for free fields

In this section we to some extent pour cold water on the technicalities of the foregoing when we examine the specific case of a free field. In this case the Lagrangian is a bilinear form in the field,

$$\mathscr{L} = \phi^{\dagger}(x)K(x, y)\phi(y), \tag{95}$$

where K is an appropriate linear operator. We now apply prescription (4) and obtain

$$\begin{aligned} \gamma\mathscr{L} &= \phi^{\dagger}(\gamma x)K(\gamma x, \gamma y)\phi(\gamma y) \\ &= \phi^{\dagger}(x)a^{-1}(\gamma)K(\gamma x, \gamma y)a(\gamma)\phi(y) \\ &= \mathscr{L} = \phi^{\dagger}(x)K(x, y)\phi(y), \end{aligned} \tag{96}$$

so that

$$K(x, y) = a^{-1}(\gamma)K(\gamma x, \gamma y)a(\gamma), \tag{97}$$

but (14) says that

$$K(x, y) = K(\gamma x, \gamma y), \tag{98}$$

so that K actually commutes with $a(\Gamma)$ since we have (97) for all γ. Thus, provided we stick to quantities manufactured from K, we will always be in the happy situation of not having to make more than one projection. Moreover since we now have (20) satisfied we can leave the projection to the end of any operator calculation and can use the spectral projection theorem directly.

If K is a Laplacian, we can construct its Zeta function and since we have (20) we can project the Zeta function in the way indicated in Dowker and Banach (1978 § 2). The conclusion expressed there — that the group action filters out the automorphic eigenfunctions in the eigenfunction expansion of the Zeta function — is true by the spectral projection theorem, but we no longer have a convenient formula for the degeneracy, essentially because the form of (86) only gives information about irreducible components readily.

Having obtained the Zeta function, the effective one-loop action $W^{(1)}$ can be given by the standard formula (Dowker and Banach 1978, Dowker and Critchley 1976)

$$W^{(1)} = -\lim_{s\to 1} \frac{\mathrm{i}}{2}\left(\frac{\zeta_M(0)}{(s-1)} + \zeta'_M(0)\right), \tag{99}$$

where $\zeta_M(x', x''; s)$ is the sth power of the inverse of the $K(x, y)$ appearing in (95) projected according to rule (50) and $\zeta_M(s)$ is the integral over M of the traced coincidence limit of $\zeta_M(x', x''; s)$. We can also use the Zeta function for evaluating the vacuum averaged stress–energy $\langle T_{\mu\nu}\rangle$,

$$\langle T_{\mu\nu}\rangle = \lim_{\substack{x'\to x\\ x''\to x\\ s\to 1}} T_{\mu\nu}(x', x'')\zeta_M(x', x''; s), \tag{100}$$

where $T_{\mu\nu}$ is the operator relevant to the type of field considered (given by e.g. De Witt 1975), but it is often easier to regularise this quantity by other procedures such as dropping divergent terms in an image summation.

7. Concluding remarks

In the previous sections we have examined the properties of an arbitrary linear operator under projection. Although in most physical situations that we are likely to examine K actually commutes with $a(\Gamma)$, as remarked in the last section, we thought it worthwhile to look at the non-commuting case for its mathematical interest. In so doing, we found that it is in fact condition (20) rather than commutativity which affords the crucial simplifications in the theory. Condition (20) is strictly weaker than commutativity as a simple example can show. (Consider K a 2×2 function matrix and $a(\Gamma)$ the group $\{\pm\mathrm{diag}\,(1, 1),\ \pm\mathrm{diag}\,(1, -1)\}$; then (20) gives certain conditions on the off-diagonal elements of K but does *not* say that they are zero.)

Apart from the obvious desirability of putting all of this to some practical use which, as promised in the Introduction, forthcoming papers will do, it is also of considerable interest to determine whether the group theoretical methods of the present paper can be extended to the more general type of representation (6) and hence to see whether we can 'derive' local gauge theories by this route. This is a problem we will also consider in the future.

References

Avis S J and Isham C J 1978 *Imperial College Preprint* ICTP/77-78/9
De Witt B S 1975 *Phys. Rep.* **19** 295
Dowker J S 1972 *J. Phys. A: Gen. Phys.* **5** 936
Dowker J S and Banach R 1978 *J. Phys. A: Math. Gen.* **11** 2255
Dowker J S and Critchley R 1976 *Phys. Rev.* D **13** 3224
Hurt N E 1976 in *Lecture Notes in Physics* vol 50, ed. A Janner, T Jansen and M Boon (Berlin: Springer)
Isham C J 1978a *Proc. R. Soc.* A **362** 383
—— 1978b *Imperial College Preprint* ICTP/77-78/13
Selberg A 1956 *J. Ind. Math. Soc.* **20** 47
Steenrod N 1951 *The Topology of Fibre Bundles* (Princeton, NJ: UP)