Another discussion of the axial vector anomaly and the index theorem

1978 J. Phys. A: Math. Gen. 11 347

(http://iopscience.iop.org/0305-4470/11/2/012)

J. Phys. A: Math. Gen., Vol. 11, No. 2, 1978. Printed in Great Britain. © 1978

Another discussion of the axial vector anomaly and the index theorem†

J S Dowker

Department of Theoretical Physics, The University, Manchester M13 9PL, UK

Received 7 October 1977

Abstract. A derivation is presented of the axial anomaly in a background Riemannian manifold using zeta-function regularisation. This leads directly to the relation with the index theorem. A spin-1 index theorem is derived, giving the Hirzebruch signature theorem and the expression for the Euler number in terms of the B_4 coefficients. Boundary effects are briefly mentioned. Some extensions are suggested.

1. Introduction

The importance of anomalous Ward identities in particle physics is widely appreciated. On the one hand, they have been used to place restrictions on the possible unified weak–electromagnetic field theories and on the other, to discuss the decay of pseudoscalar mesons. More recently ('t Hooft 1976a,b) their relevance to vacuum tunnelling has been explored.

The anomalous divergence of the axial vector current in a background gravitational field was first discussed by Kimura (1969). His preferred result was

$$\nabla_\mu j_5^\mu = -2mj_5 - \frac{(-g)^{-1/2}}{384\pi^2}\epsilon^{\mu\nu\rho\sigma}R_{\alpha\beta\mu\nu}R^{\alpha\beta}_{\ \ ..\rho\sigma}. \tag{1}$$

My conventions are the $(+---)$ signature and the Dirac equation

$$(\mathrm{i}\gamma^\mu\nabla_\mu - m)\psi = 0, \qquad \nabla_\mu = \partial_\mu + \Gamma_\mu.$$

Then

$$j_5^\mu \equiv \mathrm{i}\bar{\psi}\gamma^\mu\gamma_5\psi, \qquad j_5 \equiv \bar{\psi}\gamma_5\psi$$

with

$$\gamma_5 = \frac{1}{4!}(-g)^{-1/2}\epsilon^{\mu\nu\rho\sigma}\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma, \qquad (\epsilon^{0123}=1).$$

Kimura obtained (1) both by perturbation theory, *à la* Adler (1969), and by a proper-time method following Schwinger (1951) and De Witt (1965). The former approach has been used by Delbourgo and Salam (1972) and Eguchi and Freund (1976) in more recent investigations.

† An extended and amended version of a talk given at the second Gregynog Workshop on Quantisation in General Relativity, 31st August–3rd September 1977.

In the present paper I wish to re-derive (1), in the Euclidean region, using the zeta-function regularisation method developed by Dowker and Critchley (1976, 1977b) and by Hawking (1977a). This has various advantages, the major ones being elegance and, at least for external fields, generality. It also indicates most clearly the connection with the index theorem and I shall try to explain what this means, or says, in a physicist's language rather than a mathematician's. In § 3 I present a simple-minded, numerological proof of the Hirzebruch signature theorem and relate it to the spin-1 index theorem. Section 4 briefly looks at boundary effects and in the final section I suggest some possible extensions.

2. Euclidean axial anomaly

From now on all quantities are defined in a *negative*-definite, Euclidean signatured $(----)$ space with indices labelled from 1 to 4. For convenience I give some pertinent definitions and relations

$$\begin{aligned} &\gamma^4 = \mathrm{i}\gamma^0, \qquad x^4 = \mathrm{i}x^0, \qquad \gamma_5^{\mathrm{E}} = \mathrm{i}\gamma_5 \\ &\gamma^{(\mu}\gamma^{\nu)} = -\delta^{\mu\nu} \qquad (\mu, \nu = 1, 2, 3, 4) \\ &j_5^{\mu} \equiv \bar\psi\gamma^{\mu}\gamma_5^{\mathrm{E}}\psi, \quad \bar\psi \equiv \psi^{+} \qquad j_5 \equiv \mathrm{i}\bar\psi\gamma_5^{\mathrm{E}}\psi. \end{aligned} \tag{2}$$

Dirac's equation is formally unchanged,

$$(\mathrm{i}\gamma^{\mu}\nabla_{\mu} - m)\psi = 0$$

and ψ is a four-component representation of the covering group of SO(4), Sp(4), isomorphic to SU(2)⊗SU(2). Specifically, in terms of two-component spinors, $\psi = \binom{\phi_+}{\phi_-}$, where $\phi_+(\phi_-)$ belongs to the left (right) SU(2) group and I sometimes write this

$$\psi = (\tfrac{1}{2}, 0) \oplus (0, \tfrac{1}{2})$$

with the specific representations,

$$\gamma^{\mu} = \mathrm{i}\begin{pmatrix} 0 & \bar\sigma^{\mu} \\ -\sigma^{\mu} & 0 \end{pmatrix}, \qquad \gamma_5^{\mathrm{E}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \bar\sigma^{(\mu}\sigma^{\nu)} = \sigma^{(\mu}\bar\sigma^{\nu)} = -\delta^{\mu\nu}.$$

In flat space $\sigma^{\mu} = (\mathrm{i}1, \sigma^i)$, $\bar\sigma^{\mu} = (\mathrm{i}1, -\sigma^i)$ where the σ^i are the standard Pauli matrices. In curved space the corresponding quantities can be defined via the *vierbein* method, if desired.

Under O(4) *vierbein* rotations, roughly speaking (see Dowker and Dowker 1966a), ϕ_+ and ϕ_- transform as

$$\phi(k, l) \to \exp(\mathrm{i}\epsilon_{\mu\nu}J^{\mu\nu}(k, l))\phi(k, l), \qquad (J^{\mu\nu})^{\dagger} = J^{\mu\nu}, \tag{3}$$

where

$$\phi_+ = \phi(\tfrac{1}{2}, 0), \qquad \phi_- = \phi(0, \tfrac{1}{2})$$

with algebraically,

$$\begin{aligned} &J^{ij}(j, 0) = J^{ij}(0, j) = \epsilon^{ijk}J_k \\ &J^{i4}(j, 0) = -J^{i4}(0, j) = J_i, \end{aligned}$$

J_i being the spin-j angular momentum matrices ($=\frac{1}{2}\sigma^i$ for $j=\frac{1}{2}$). My notation here is simply that (k, l) refers to the spin-k representation of the left SU(2) group times the spin-l representation of the right one.

The field transformations determine the form of the spinor connection Γ_μ. I shall not need its explicit form and only record here the generalised Ricci identity, which defines the spinor curvature,

$$\nabla_{[\mu}\nabla_{\nu]}\phi(k, l)=\tfrac{1}{4}\mathrm{i}R_{\mu\nu\alpha\beta}J^{\alpha\beta}(k, l)\phi(k, l)$$

where $\phi(k, l)$ transforms as $\phi(k, 0)\otimes\phi(0, l)$.

Consider now the axial vector current j_5^μ and, following Schwinger (1951), concentrate on its vacuum ('in–out') average $\langle j_5^\mu\rangle$. Then find that this is given as the coincidence limit,

$$\langle j_5^\mu\rangle=-\mathrm{i}\lim_{x'\to x}\operatorname{tr}(\gamma^\mu(x)\gamma_5^{\mathrm{E}}(x)S_{\mathrm{E}}(x, x')) \tag{4}$$

where S_{E} is the Euclidean spinor Green function satisfying

$$\mathrm{i}\nabla\!\!\!\!/\, S_{\mathrm{E}}=1\delta_{\mathrm{E}}, \qquad -\mathrm{i}S_{\mathrm{E}}\overleftarrow{\nabla\!\!\!\!/}'=1\delta_{\mathrm{E}}$$

with $\nabla\!\!\!\!/=\gamma^\mu(x)\nabla_\mu(x)$ and $\nabla\!\!\!\!/'=\gamma^\mu(x')\nabla_\mu(x')$. δ_{E} is the Euclidean δ-function related to the Minkowski one, δ, by $\delta=\mathrm{i}\delta_{\mathrm{E}}$ (e.g. Schwinger 1970, p 146). I have set the mass equal to zero.

In equation (4), and in the following manipulations, I have not been, nor shall I be, too careful to symmetrise the various expressions in x and x'. The reason is that, if S_{E} is imagined to be replaced by a regularised quantity, as it will be, the coincidence limit is unambiguous. Furthermore, no parallel propagators have been introduced into the definition of the point split j_5^μ. In fact, the correct way of doing this, if there is one, has never been clear to me. In Schwinger's (1951) original calculation a gauge-covariance preserving factor is effectively incorporated into j_5^μ by ensuring that any derivatives of the ψ and $\bar{\psi}$ are of the correct covariant form (Schwinger 1951, equation (5.16)). In the present case this seems to be superfluous since the covariant derivatives, ∇_μ, are always in evidence.

It might be worthwhile to point out, parenthetically, that the parallel propagator, $I(x, x')$, hidden in Schwinger's calculation, must be multivalued because

$$\partial_\mu I=\mathrm{i}eA_\mu I.$$

In other discussions of j_5^μ and j^μ it is not always clear whether $I(x, x')$ is multivalued or not. If not, then its gradient has an extra term involving an integral of the field strength along the fixed path used to define it. This amounts to a different definition of the current (cf Kimura 1969).

Returning to our problem, define G_{E} by

$$S_{\mathrm{E}}=-\mathrm{i}\nabla\!\!\!\!/\, G_{\mathrm{E}}=\mathrm{i}G_{\mathrm{E}}\overleftarrow{\nabla\!\!\!\!/}'$$

so that

$$\nabla\!\!\!\!/^2 G_{\mathrm{E}}=G_{\mathrm{E}}\overleftarrow{\nabla\!\!\!\!/}'^2=1\delta_{\mathrm{E}} \tag{5}$$

and

$$\langle j_5^\mu\rangle=-\lim_{x'\to x}\operatorname{tr}(\gamma^\mu\gamma_5^{\mathrm{E}}\nabla\!\!\!\!/\, G_{\mathrm{E}}(x, x')).$$

Then

$$\nabla_\mu \langle j_5^\mu \rangle = 2 \lim_{x' \to x} \mathrm{tr}(\gamma_5^{\mathrm{E}} \nabla\!\!\!\!/^2 G_{\mathrm{E}}(x, x')) \tag{6}$$

where the factor of 2 comes from differentiating with respect to x and x', and I have set x' equal to x wherever necessary, bearing in mind that G_{E} is really regularised.

It is now necessary to remark that if the equation

$$\nabla\!\!\!\!/^2 \psi_n = \lambda_n \psi_n$$

has any zero eigenvalues, G_{E} cannot be constructed and it is necessary to project out the corresponding zero eigenfunctions. Define $\bar{G}_{\mathrm{E}}$ by

$$\nabla\!\!\!\!/^2 \bar{G}_{\mathrm{E}} = 1\delta_{\mathrm{E}} - P$$

where P is the projection operator onto the null space of $\nabla\!\!\!\!/^2$. Then, instead of (6) there is

$$\nabla_\mu \langle j_5^\mu \rangle = 2 \lim_{x' \to x} \mathrm{tr}(\gamma_5^{\mathrm{E}} \nabla\!\!\!\!/^2 \bar{G}_{\mathrm{E}}(x, x')).$$

In terms of the two-component representation this equation reads

$$\nabla_\mu \langle j_5^\mu \rangle = 2 \lim_{x' \to x} \mathrm{tr}(\Box_+ \bar{G}_+(x, x') - \Box_- \bar{G}_-(x, x')) \tag{7}$$

where the 'E' has finally been dropped. The '+' and '−' refer to the upper, $(\frac{1}{2}, 0)$, and lower, $(0, \frac{1}{2})$, components of ψ, respectively, and the trace is a two-dimensional one. Thus

$$\nabla\!\!\!\!/^2 = \begin{pmatrix} \Box_+ & 0 \\ 0 & \Box_- \end{pmatrix} \quad \text{and} \quad \bar{G}_{\mathrm{E}} = \begin{pmatrix} \bar{G}_+ & 0 \\ 0 & \bar{G}_- \end{pmatrix}$$

and

$$\Box_i \bar{G}_i = 1 - P_i \qquad (i = + \text{ or } -) \tag{8}$$

$(\Box_+ = \bar{\sigma}^\mu \sigma^\nu \nabla_\mu \nabla_\nu, \Box_- = \sigma^\mu \bar{\sigma}^\nu \nabla_\mu \nabla_\nu)$.

The implied regularisation is now made explicit by replacing the $\bar{G}$ Green functions by their matrix powers, $\bar{G}^s$, as described in Dowker and Critchley (1976) where other, relevant references are given. The coincidence limit will be finite if $s > 2$ and the desired, physical quantity is obtained by continuing s into the complex plane and then down to $s = 1$. This continuation replaces the powers $\bar{G}^s$ by the zeta-function $\zeta(s)$. For simplicity I have not introduced the scaling parameter.

To implement this procedure I write firstly

$$\nabla_\mu \langle j_5^\mu \rangle = 2 \lim_{\substack{x' \to x \\ s \to 1}} \mathrm{tr}(\Box_+ \bar{G}_+^s(x, x') - \Box_- \bar{G}_-^s(x, x')). \tag{9}$$

If (8) is formally multiplied on the right by $\bar{G}_i^{s-1}$ there results

$$\Box_i \bar{G}_i^s = \bar{G}_i^{s-1} \qquad (s > 2)$$

if one uses the fact that

$$P_i \bar{G}_i = 0.$$

Then it is seen that the right-hand side of (9) consists of the difference of $\bar{G}_+^{s-1}$ and $\bar{G}_-^{s-1}$ so that the continuation to $s=1$ produces

$$\nabla_\mu \langle j_5^\mu \rangle = 2 \operatorname{tr} \operatorname{diag}_x(\zeta_+(0) - \zeta_-(0))$$

or, in another notation,

$$\nabla_\mu \langle j_5^\mu \rangle = 2\Big(\sum_a (x, a|\zeta_+(0)|x, a) - \sum_b (x, b|\zeta_-(0)|x, b)\Big). \tag{10}$$

Here, a and b are two-spinor indices and the operator zeta-functions $\zeta_i(s)$ are given in terms of the eigenvalue problems

$$\Box_i|n, i) = \lambda_n^i|n, i)$$

by

$$\zeta_i(s) = \sum_{\substack{n \\ (\lambda_n^i \neq 0)}} \frac{|n, i)(n, i|}{(\lambda_n^i)^s}.$$

The standard theory of zeta-functions (e.g. Minakshisundaram and Pleijel 1949, theorem on pp 252, 253, Seeley 1967, Atiyah *et al* 1973) gives the particular value

$$(x, a|\zeta_i(0)|x, b) = [E_4^i(x)]_a{}^b - (x, a|P_i|x, b), \tag{11}$$

in Gilkey's notation (Gilkey 1975a). The trace on the spinor (= 'vector, or spin, bundle') indices turns the $E_4(x)$ into $B_4(x)$ and it is simply a question of evaluating this quantity. One can use Gilkey's general expressions, but for spin-$\frac{1}{2}$ De Witt's are sufficient (De Witt 1965).

The results are, by now, fairly well known, but I write them out again for completeness. The general form of $B_4(x)$ $(=(16\pi^2)^{-1} \operatorname{tr} a_2)$ is

$$16\pi^2 B_4(x) = \alpha C_{\mu\nu\alpha\beta} C^{\mu\nu\alpha\beta} + \beta(R_{\alpha\beta}R^{\alpha\beta} - \tfrac{1}{3}R^2) + \gamma\Box R + \delta R^2$$
$$+ \epsilon (g)^{-1/2} R^{\alpha\beta}{}_{\rho\sigma} R_{\alpha\beta\mu\nu} \epsilon^{\rho\sigma\mu\nu}, \tag{12}$$

with the specific constants $(\alpha, \beta, \gamma, \delta, \epsilon)$ for $B_4^\pm$ equal to $(-\frac{7}{720}, -\frac{11}{360}, \frac{1}{60}, 0, \mp\frac{1}{96})$. Thus (10) reads

$$\nabla_\mu \langle j_5^\mu \rangle = -\frac{(g)^{-1/2}}{384\pi^2} R^{\alpha\beta}{}_{\rho\sigma} R_{\alpha\beta\mu\nu} \epsilon^{\rho\sigma\mu\nu} - 2\sum_1^{n_+} \mathring{\phi}_+^\dagger \mathring{\phi}_+ + 2\sum_1^{n_-} \mathring{\phi}_-^\dagger \mathring{\phi}_- \tag{13}$$

where the $\mathring{\phi}_\pm$ are normalised zero eigenvalue modes, with degeneracies $n_\pm$.

Equation (13) is a local statement. An integration over the manifold M, which is supposed to be closed (i.e. compact without boundary), yields the specific *index theorem* for the Dirac operator,

$$n_+ - n_- = -\frac{1}{768\pi^2} \int_M R^{\alpha\beta}{}_{\rho\sigma} R_{\alpha\beta\mu\nu} \epsilon^{\rho\sigma\mu\nu} \, d^4x = -\frac{1}{24} p, \tag{14}$$

p being the Pontrjagin number of M. This result can be derived directly of course (Atiyah and Singer 1968, Seeley 1967, p 292, Atiyah *et al* 1973).

I shall give a general definition of the index in the next section and so simply point out here that the index is defined as the left-hand side of (14), i.e. the difference of the numbers of positive and negative helicity, massless Dirac fields.

A simple mathematical consequence of (14) is that the Pontrjagin number of a compact manifold that admits a spin structure must be a multiple of 24 since the left-hand side is an integer. Examples are, apparently, hard to find, but there exists at least one, the famous K3 surface.

Physical consequences are not so easy to see. Unless we wish to remain in Euclidean signatured space, and perhaps argue along thermal lines (e.g. Charap and Duff 1977, unpublished report, Harrington and Shepard 1977), it is necessary to return to the $(+---)$ signature. Now equation (1) was derived in this signature. If it is integrated over all space–time and if equation (14) is used *ad hoc* for the right-hand side there results (cf 't Hooft 1976a, b)

$$\Delta Q^5 = n_+ - n_-$$

where ΔQ^5 is the change in axial charge. If this argument is valid it shows that the divergence $\nabla_\mu j_5^\mu$ in the Euclidean region should not be compared with the divergence in the $(+---)$ region. Rather, it is the zero eigenfunction projection operators in (13) that take over this role.

Equation (13) has also been derived by Nielsen *et al* (1977), using a point-splitting method. This would resolve a problem raised by Kimura (1969) who could not obtain agreement with the numerical factor of $\frac{1}{384}$ with the point-splitting technique.

The relation between the axial anomaly and the index theorem has also been considered, in the context of gauge theories, by Jackiw and Rebbi (1977). They take the massless limit of a massive theory and use Pauli–Villars regularisation as advocated by Hagen (1969) and used by Delbourgo and Salam (1972). It is perhaps worthwhile pointing out again that this regularisation method has the inestimable advantage of being applicable at the Lagrangian level.

3. Spin-1 index theorem

At the risk of repeating standard material, I should like to give a description of the index theorem. The simplest one that I could find is the following. With the necessary explanations and examples, it is not the shortest form, but I hope it will be useful.

Suppose that $\mathscr{D}$ is an elliptic operator from a bundle E to a bundle F. This means that, acting on a field with indices (= a section of E), $\mathscr{D}$ produces another field with indices (= a section of F). For example, $\mathscr{D}$ could be $\sigma^\mu \nabla_\mu$ of § 2. Then E would be the collection of $(\frac{1}{2}, 0)$ spinor fields, $\phi_+(x)$, and F that of $(0, \frac{1}{2})$ spinors, $\phi_-(x)$.

Assume further that there are Hermitian inner products in the fibres of E and F. In our example, these would be just the products $\phi_+^\dagger(x) \cdot \phi_+(x)$ and $\phi_-^\dagger(x) \cdot \phi_-(x)$ at each point x of the underlying manifold. (Each such point picks out a particular fibre from the bundle, i.e. picks out the set of values of the collection of fields at that point.) Then for two fields ('sections') there is defined a functional, or global, Hermitian inner product which is, for example, for the two spinor fields $\phi_1(x)$ and $\phi_2(x)$, both of either + or − type,

$$(\phi_1|\phi_2) = \sum_a \int_M (\phi_1|x, a)(x, a|\phi_2) g^{1/2}\, dx = \int_M \phi_1^\dagger(x) \cdot \phi_2(x) g^{1/2}\, dx.$$

With respect to this inner product $\mathscr{D}$ has a formal adjoint $\mathscr{D}^*$ in the usual way. $\mathscr{D}^*$ takes F into E. For our spin-$\frac{1}{2}$ example, $\mathscr{D}^* = \bar{\sigma}^\mu \nabla_\mu$.

It is now possible to define the 'squared' operators or Laplacians,

$$\Box_E = \mathscr{D}^* \mathscr{D}, \qquad \Box_F = \mathscr{D}\mathscr{D}^*,$$

which take E into E and F into F respectively. In the Dirac case $\Box_E = \Box_+$ and $\Box_F = \Box_-$ (see equation (8)).

Both $\mathscr{D}^*\mathscr{D}$ and $\mathscr{D}\mathscr{D}^*$ are *self-adjoint, elliptic* and *non-negative.*

Consider now the null spaces ('kernels') of $\mathscr{D}$ and $\mathscr{D}^*$, i.e. the spaces of those fields satisfying

$$\mathscr{D}\phi_E = 0, \qquad \mathscr{D}^*\phi_F = 0.$$

Such fields are called 'harmonic'. According to a general theorem of elliptic operators these linear vector spaces are finite dimensional and the index of $\mathscr{D}$ can be defined to be the difference in these dimensions. This is written as

$$\text{index}\,(\mathscr{D}) = \dim\ker\mathscr{D} - \dim\ker\mathscr{D}^*.$$

The operators $\mathscr{D}^*\mathscr{D}$ and $\mathscr{D}\mathscr{D}^*$ have discrete spectra with finite multiplicities (degeneracy) and their *non-zero* eigenvalues coincide, including multiplicities, as is easily shown. Therefore, we have the useful, and general, results

$$\text{index}\,(\mathscr{D}) = \operatorname{Tr}\exp(-\mathscr{D}^*\mathscr{D}t) - \operatorname{Tr}\exp(-\mathscr{D}\mathscr{D}^*t), \qquad \text{all } t, \tag{15}$$

and

$$\text{index}\,(\mathscr{D}) = \operatorname{Tr}(1+\mathscr{D}^*\mathscr{D})^s - \operatorname{Tr}(1+\mathscr{D}\mathscr{D}^*)^s, \qquad \text{all } s, \tag{16}$$

or we can make the seemingly vacuous statement that the traced zeta-functions for $\mathscr{D}\mathscr{D}^*$ and $\mathscr{D}^*\mathscr{D}$, after removing the zero eigenvalues, are identical:

$$\operatorname{Tr}\underset{\sim}{\zeta}_E(s) = \operatorname{Tr}\underset{\sim}{\zeta}_F(s). \tag{17}$$

The trace operation in (15), (16) and (17) stands for a trace over any convenient complete set of 'states'. Thus, for an operator $\underset{\sim}{A}$,

$$\operatorname{Tr}\underset{\sim}{A} = \sum_a \int_M (x, a|\underset{\sim}{A}|x, a) g^{1/2}\,\mathrm{d}x$$

or, equivalently,

$$\operatorname{Tr}\underset{\sim}{A} = \sum_n (n|\underset{\sim}{A}|n).$$

The formal relation between the operator (i.e. untraced) zeta-functions is

$$\underset{\sim}{\zeta}_F(s) = \mathscr{D}\underset{\sim}{\zeta}_E(s+1)\mathscr{D}^*,$$

which generalises (17).

So far, the underlying manifold of the x, M, has been tacitly assumed compact, with or without boundary. If M has a boundary it is necessary to say something about boundary conditions. This is deferred until the following section. If M is closed, $\partial M = \varnothing$, we do not have to worry about boundary conditions and can immediately use the asymptotic expansion (e.g. Atiyah *et al* 1973),

$$\operatorname{Tr}\exp(-\Box_i t) = t^{-d/2} \sum_{n=0,1,\dots} B^i_{2n} t^n, \qquad t\downarrow 0,\ (i = E, F), \tag{18}$$

where d is the dimension of M and the B_{2n} are the integrated $B_{2n}(x)$,

$$B_{2n} = \int_M B_{2n}(x) g^{1/2}\,\mathrm{d}x$$

in order to obtain the index as the constant term in (15)

$$\text{index}\,(\mathscr{D}) = B_d^E - B_d^F. \tag{19}$$

The same result follows by setting s equal to zero in (17) in accordance with equation (11), for $d=4$.

The quantities $B_{2n}(x)$ are constructed out of the objects occurring in the operator $\mathscr{D}$, or $\Box$. If $\Box$ is a purely geometric operator, such as a covariant Laplacian on a Riemannian space, then the $B_{2n}(x)$ are local combinations of quantities constructed from the metric, such as the curvature. This is the only case I am interested in and for spin-$\frac{1}{2}$ in four dimensions equation (19) is just the previous result (14).

Equation (19) is one statement of the index theorem. It can be shown, and this is important (e.g. Atiyah and Singer 1968), that the index of an operator is *unchanged* under continuous deformations of that operator. In our case this would amount to a continuous change in the metric and then the index theorem says that the particular geometric quantity found on the right-hand side must also be unchanged if the metric is altered. In other words, it must be a topological invariant. Equation (14) of course agrees with this. One can argue on dimensional and scaling grounds that the topological invariant has to be either the Euler or the Pontrjagin number and the latter is selected for parity reasons.

After this standard discussion of the elementary index theorem I wish to apply it to the four-dimensional massless spin-1 case (the Euclidean photon). Instead of using the notation and terminology of p-forms I prefer, at least here, to use the description in terms of spinors and tensors more familiar to physicists. I shall begin by writing Maxwell's equations in a form similar to the neutrino equations, thus,

$$\alpha^\mu \nabla_\mu \Phi = 0 \qquad \bar{\alpha}^\mu \nabla_\mu \Upsilon = -\Phi. \tag{20}$$

The first equation corresponds to a positive-helicity equation.

The history of this form is given in earlier works (Dowker and Dowker 1966b, Dowker 1967a). It dates back to Rumer (1930).

Φ is a spinor transforming as the reduced representation $(j,0)\oplus(j-1,0)$ in general, here as $(1,0)\oplus(0,0)$, i.e. as a self-dual Maxwell field, $(1,0)$, plus a scalar, $(0,0)$. I write

$$\Phi = \begin{pmatrix} \phi \\ \phi_0 \end{pmatrix}, \qquad \phi \sim (1,0),\ \phi_0 \sim (0,0).$$

Υ is a potential field transforming as $(j-\frac{1}{2},\frac{1}{2})$ in general, here as $(\frac{1}{2},\frac{1}{2})$, i.e. as a four-vector. The α^μ are $4j\times 4j$ matrices, in general, here 4×4 and $\alpha^\mu\nabla_\mu$ acting on $(1,0)\oplus(0,0)$ produces $(\frac{1}{2},\frac{1}{2})$. The precise form and properties of the α^μ and $\bar{\alpha}^\mu$ are given in the cited references, but are not needed here. Suffice it to say that they possess all the algebraic properties, except completeness, that the σ^μ and $\bar{\sigma}^\mu$ have. Also, $\alpha^i\nabla_i$ corresponds to the operator

$$\begin{pmatrix} \text{curl} & \text{grad} \\ \text{div} & 0 \end{pmatrix}$$

(Atiyah *et al* 1975a).

As well as the pair of equations (20) there will be a conjugate set involving $(0,1)\oplus(0,0)$ and $(\frac{1}{2},\frac{1}{2})$ fields, and the index theorem is to be applied to these two sets in turn.

Choose $\mathscr{D}=\alpha^{\mu}\nabla_{\mu}$ and $\mathscr{D}^{*}=\tilde{\alpha}^{\mu}\nabla_{\mu}$. It is found (Dowker and Dowker 1966b) that the equation

$$\mathscr{D}^{*}\mathscr{D}\Phi=0$$

decouples into one for ϕ only,

$$(\nabla^{\mu}\nabla_{\mu}-\tfrac{1}{4}R_{\mu\nu\rho\sigma}J^{\mu\nu}(1,0)J^{\rho\sigma}(1,0))\phi=0, \tag{21}$$

and one for ϕ_0 only, the minimally coupled scalar equation,

$$\nabla^{\mu}\nabla_{\mu}\phi_0=0. \tag{22}$$

The index theorem says, with obvious notation,

$$n[(1,0)\oplus(0,0)]-n(\tfrac{1}{2},\tfrac{1}{2})=B_4[(1,0)\oplus(0,0)]-B_4(\tfrac{1}{2},\tfrac{1}{2}). \tag{23}$$

The decoupling (21), (22) means that

$$B_4[(1,0)\oplus(0,0)]=B_4(1,0)+B_4(0,0)$$

and if we set $\phi_0=0$ (the only solutions of (22) are constants) the index theorem becomes

$$n(1,0)-n(\tfrac{1}{2},\tfrac{1}{2})=B_4(1,0)+B_4(0,0)-B_4(\tfrac{1}{2},\tfrac{1}{2}). \tag{24}$$

$n(1,0)$ is the number of regular self-dual harmonic two-forms ($(1,0)$ is a self-dual two-form, $\underset{\sim}{H}-\mathrm{i}\underset{\sim}{E}$ in Minkowski space) and $n(\frac{1}{2},\frac{1}{2})$ is the number of regular harmonic vector fields in the Lorentz gauge, $\nabla_{\mu}A^{\mu}=0$.

The conjugate set of Maxwell equations produce the conjugate index theorem,

$$n(\tfrac{1}{2},\tfrac{1}{2})-n(0,1)=B_4(\tfrac{1}{2},\tfrac{1}{2})-B_4(0,1)-B_4(0,0), \tag{25}$$

where $n(0,1)$ is the number of regular anti-self-dual harmonic two-forms.

The B_4 coefficients are all given by integrals of expression (12). The particular values for the constants (α, β, γ, δ, ϵ) can be obtained in principle by substituting the relevant spinor curvatures and endomorphisms 'E' into Gilkey's formulae (Gilkey 1975a) and doing the traces over products of angular momentum matrices. (For $(1,0)$ the endomorphism, E, is the second term in brackets in (21)). Actually, for the simple cases needed here one can avoid angular momentum theory. Some useful values are given in table 1. Our curvature conventions are those of Schouten (1954) (which are those of Gilkey 1975a and De Witt 1965). To obtain the Misner–Thorne–Wheeler convention reverse the sign of R.

If (24) and (25) are added there results, using the tabulated values,

$$n(1,0)-n(0,1)=\tfrac{1}{3}p \tag{26}$$

which is a statement of the Hirzebruch signature theorem (e.g. Hirzebruch 1966, Atiyah and Singer 1963, 1968, Atiyah *et al* 1973), derived here in a lowbrow way.

On subtraction (24) and (25) yield

$$n(1,0)+n(0,1)-2n(\tfrac{1}{2},\tfrac{1}{2})=B_4(1,0)+B_4(0,1)+2B_4(0,0)-2B_4(\tfrac{1}{2},\tfrac{1}{2}) \tag{27}$$

and we would expect that the right-hand side would be a topological invariant. This is indeed the case. By a fundamental identity (e.g. Gilkey 1975b) it is the Euler characteristic, χ, of M, as can be checked numerically from the values in table 1.

Note that the Euler characteristic appears as the difference of two index theorems. This is in agreement with the general treatment given by Atiyah, Patodi and Singer

Table 1. Values of the coefficients ($\alpha, \beta, \gamma, \delta, \epsilon$) in equation (12) for $16\pi^2 B_4(k, l)$. Also included are the corresponding numbers for the Euler and Pontrjagin classes ($\times 16\pi^2$).

(k, l)	α	β	γ	δ	ϵ
(0, 0) Minimal	$\frac{1}{180}$	$\frac{1}{180}$	$-\frac{1}{30}$	$\frac{1}{72}$	0
(0, 0) Conformal	$\frac{1}{180}$	$\frac{1}{180}$	$-\frac{1}{180}$	0	0
$(\frac{1}{2}, \frac{1}{2})$	$-\frac{11}{180}$	$\frac{16}{45}$	$\frac{1}{30}$	$\frac{1}{36}$	0
$(\frac{1}{2}, 0)$ / $(0, \frac{1}{2})$	$-\frac{7}{720}$	$-\frac{11}{360}$	$\frac{1}{60}$	0	$-\frac{1}{96}$ / $+\frac{1}{96}$
$(1, 0)$ / $(0, 1)$	$\frac{11}{60}$	$-\frac{3}{20}$	$\frac{1}{15}$	$\frac{1}{72}$	$+\frac{1}{12}$ / $-\frac{1}{12}$
χ Euler class	$\frac{1}{2}$	-1	0	0	0
p Pontrjagin class	0	0	0	0	$\frac{1}{2}$

(APS) (1975a, pp 66–7). I feel that the present discussion might assist those unfamiliar with the techniques used by APS.

For consistency (26) says that the Pontrjagin number, p, of a closed manifold must be three times an integer. The only condition on the Euler number is that it is odd or even according as p is odd or even.

For further interesting comments on the zero eigenvalue modes and their significance, see the work of Gibbons ('Functional integrals in curved spacetime', Munich, 1977).

In the purely gravitational case it is not possible to derive 'vanishing theorems' as easily as for external gauge fields. That is, to prove that only one of n_+, n_- or of $n(1, 0)$, $n(0, 1)$ is non-zero (Jackiw and Rebbi 1977). Consider equation (21) for the $(1, 0)$ spinor (self-dual two-form) and choose a closed Ricci-flat manifold, $R_{\mu\nu} = 0$. Because $J^{\mu\nu}(1, 0)$ is self-dual, $R_{\mu\nu\alpha\beta}$ ($= C_{\mu\nu\alpha\beta}$, the Weyl tensor) can be replaced by its traceless and completely self-dual part, $W_{\mu\nu\alpha\beta}$, say. Further, assume that this is zero, i.e. that the space is 'half-flat'. Then

$$\nabla^\mu \nabla_\mu \phi = 0$$

and so, in the usual manner,

$$\int_M \phi^\dagger \nabla^\mu \nabla_\mu \phi = -\int_M (\nabla_\mu \phi)^\dagger (\nabla^\mu \phi) = 0$$

which implies that ϕ has vanishing covariant derivative, $\nabla_\mu \phi = 0$.

Now construct the commutator $\nabla_{[\mu} \nabla_{\nu]} \phi$ and use the generalised Ricci identity to give

$$C_{\mu\nu\alpha\beta} J^{\alpha\beta}(1, 0)\phi = 0.$$

Because of the properties of the Weyl tensor this quantity is also self-dual on the $\mu\nu$ pair of indices and so one finds

$$W_{\mu\nu\alpha\beta}J^{\alpha\beta}\phi = 0$$

which is automatically satisfied. One cannot conclude that $n(1, 0)$ is zero.

In the gauge field case, however, the transformation properties of ψ do not select the (anti) self-dual part of the gauge field, $\underset{\sim}{F}_{\mu\nu}$, in the commutator of covariant derivatives, as they did above, and the corresponding condition makes the *whole* of $\underset{\sim}{F}_{\mu\nu}$ vanish, which is not interesting, or makes ψ vanish, which means that the related degeneracy is zero. (Note that it is not enough to say that *obviously* $\nabla^\mu\nabla_\mu$ is a positive-definite operator and hence has no normalisable zero eigenfunction.)

In a combined situation, for example, a charged Dirac particle in a half-flat manifold, one would expect to obtain vanishing theorems. The explicit calculations of Pope bear this out (see § 4).

The Ricci-flat condition is, in fact, too strong. All that is required is that the scalar curvature, R, should be zero.

If R is not zero, but still with the space being half-flat, the equation for ϕ is

$$\nabla^\mu\nabla_\mu\phi - \tfrac{1}{3}R\phi = 0$$

and the analysis proceeds along the lines indicated by Lichnerowicz (1963) with identical conclusions. These are left for the reader to draw.

Similar considerations apply to the vector potential, Y. Thus, for a Ricci-flat space, whether half-flat or not, it is easy to show that $n(\frac{1}{2}, \frac{1}{2})$ is zero, i.e. that there are no regular harmonic vectors (or 'one-forms'). Another way of saying this is that if M is Ricci-flat, but not flat, the first Betti number is zero. This is sometimes called Myers' theorem (see the work of Gibbons, already cited). When combined with the index theorems, (24) and (25), it gives

$$n(1,0) = \tfrac{1}{2}(\chi + \tfrac{1}{3}p), \qquad n(0,1) = \tfrac{1}{2}(\chi - \tfrac{1}{3}p)$$

in Ricci-flat spaces.

If, in addition, the space is half-flat, $\chi = \pm\frac{1}{2}p$, with the sign depending on whether the curvature is self-dual or anti-self-dual, and so

$$n(1,0) = \tfrac{5}{12}p \text{ or } -\tfrac{1}{12}p, \qquad n(0,1) = \tfrac{1}{12}p \text{ or } -\tfrac{5}{12}p$$

for Ricci-flat, half-flat spaces. Thus p must be a multiple of 12. If M is a spin manifold this is guaranteed since it then has to be a multiple of 24.

4. Boundary effects

If M is not closed, $\partial M \neq \varnothing$, a number of problems arise and I wish to make a few general comments on these.

Firstly, if it is wished to integrate equation (10) over the manifold, including the boundary, it is not correct to use expression (11) everywhere since this is true only if x is an interior point, i.e. $x \notin \partial M$. However, a formal integration of (10) yields

$$\int_M \nabla_\mu \langle j_5^\mu \rangle = 2(\zeta_+(0) - \zeta_-(0)), \tag{28}$$

where $\zeta_i(s)$ is the traced and integrated zeta-function,

$$\zeta_i(s)=\sum_{\substack{n\\(\lambda_n^i\neq 0)}}(\lambda_n^i)^{-s}\qquad (i=+,-).$$

Since the non-zero eigenvalues of $\mathscr{D}^*\mathscr{D}$ and $\mathscr{D}\mathscr{D}^*$ coincide, the right-hand side of (28) vanishes and for consistency it would seem that the boundary conditions would have to be such as to make the left-hand side also zero, even though $\partial M\neq\varnothing$.

Apparently, it is not possible to impose local boundary conditions, such as Dirichlet or Neumann, and obtain a well posed elliptic problem. Instead APS (1975a, b, 1976) introduce a non-local condition which is essentially a restriction on the spectral decomposition of the fields on the boundary in terms of the eigenfunctions of the 'boundary part' of $\mathscr{D}$. Thus, for ϕ_+, negative eigenvalues only are allowed, which implies, roughly speaking, that ϕ_+ vanishes at the infinite limit of the cylinder that can be attached to ∂M in order to extend the normal coordinate at the boundary, u, to the *outward* range 0 to $-\infty$. The spectral restriction means that ϕ_+ contains only terms like $\exp(-\lambda u)$ with $\lambda<0$.

For such boundary conditions $\zeta_i(0)$ is given, in four dimensions, by

$$\zeta_i(0)=B_4^i+C_4^i-n_i-S_i$$

where B_4 has been given before, C_4 is an integral over the boundary of local quantities involving the second fundamental form of ∂M (e.g. Greiner 1971, Gilkey 1975b) and S is an additional term, a spectral invariant associated with ∂M.

In the general case the equation

$$\zeta_E(0)=\zeta_F(0)$$

becomes

$$B_4^E+C_4^E-B_4^F-C_4^F-\tfrac{1}{2}(h+\eta(0))=n_E-n_F=\text{index}\,(\mathscr{D}),\tag{29}$$

where h is the degeneracy of the zero eigenvalue of the boundary part of $\mathscr{D}$ and $\eta(s)$ is the APS spectral invariant defined by

$$\eta(s)=\sum_{\substack{n\\(\nu_n\neq 0)}}\epsilon(\nu_n)|\nu_n|^{-s},\qquad \epsilon(x)=\begin{cases}1, & x>0\\ -1, & x<0\end{cases}$$

in terms of the eigenvalues, ν_n, of the boundary part of $\mathscr{D}$.

Applied to the Dirac equation, charged and uncharged, equation (29) has been investigated and checked explicitly in a number of spaces by C N Pope (Cambridge, unpublished) following work by Hawking (1977b) on the Schwarzschild and Taub–NUT spaces who pointed out that there should be a residual boundary effect in these non-compact spaces to make up the defect between the left- and right-hand sides of (14) when applied to these spaces.

A related calculation has been performed by Römer and Schroer (1977). However, for them the boundary terms in (29) arise from the left-hand side of (28) while the $B_4^++n_--B_4^--n_+$ contribution is just the integrated right-hand side of (10) using (11) everywhere.

5. Discussion and extensions

I now make some disconnected remarks on the preceding theory.

(a) Instead of constructing j_5^μ as a bilinear expression in the spinor fields and obtaining $\langle j_5^\mu \rangle$ and $\nabla_\mu \langle j_5^\mu \rangle$ therefrom, it is possible to use a variational approach. One could define an effective action $W[A_\mu]$ as a functional of a background axial vector field A_μ. Then $\langle j_5^\mu \rangle$ would be defined as

$$\langle j_5^\mu \rangle = \frac{\delta W[A]}{\delta A_\mu}\bigg|_{A=0}$$

and the divergence will be

$$\nabla_\mu \langle j_5^\mu \rangle = \frac{\delta W[\partial \Lambda]}{\delta \Lambda}\bigg|_{\Lambda=0}.$$

This approach has certain attractions, but I will not elaborate on them here.

(b) It is possible to derive an index theorem for arbitrary spin, j. Equations (20) remain valid for any spin. Not only this, but there is a series of such pairs of equations (Dowker 1967b) and so there will be a series of index theorems, on the right-hand sides of which one would expect to find combinations of p and χ. It is not likely that these relations would place any restriction on p stronger than that implied by the spin-$\frac{1}{2}$ case, equation (14), since the coefficients increase with increasing spin.

(c) The APS invariant $\eta(0)$ has been evaluated for the spin-$\frac{1}{2}$ case in various spaces. There is no difficulty in performing the corresponding calculation for the spin-1 theory, especially in view of proposition 4.20 of APS (1975a).

(d) It is straightforward to extend the results to the case when the fields ψ and ϕ belong to the carrier space of some internal gauge group (cf Jackiw and Rebbi 1977, Nielsen and Schroer 1977).

(e) In d dimensions the index is given by equation (19) and, in general, the anomalies will be determined by the B_d coefficients. *Explicit* forms for the coefficient B_6 are available (Gilkey 1975a), so that the present analysis can be repeated in six dimensions.

(f) Since the axial anomaly, (13), is the local form of the spin-$\frac{1}{2}$ index theorem, (14), it is reasonable to suppose that the spin-1 index theorem, (24) and (25), would also have local forms. Presumably one would need two axial vector currents, possibly conserved classically; one composed from Φ and Υ and the other from the (independent) conjugate fields. In the Minkowski region this is not possible if the currents are bilinear in the fields and their Hermitian conjugates, although if one uses the fields and their *transposes* it is easy to construct currents from $\Psi = \binom{\Phi}{\Upsilon}$ and the new 'γ-matrices', β^μ,

$$\beta^\mu = i\begin{pmatrix} 0 & \tilde{\alpha}^\mu \\ -\alpha^\mu & 0 \end{pmatrix}, \qquad \beta^5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

However, these currents are not conserved classically.

The brief note by Nielsen *et al* (1977) contains a description of two currents—the 'Euler' and 'Hirzebruch' currents—which appear to be classically conserved four-vectors and which, when integrated over M, produce equations (27) and (26) respectively.

It is not yet clear to me why the algebra of the present paper does not also yield these currents naturally.

References

Adler S L 1969 *Phys. Rev.* **177** 2426
Atiyah M F, Bott R and Patodi V K 1973 *Invent. Math.* **19** 279
Atiyah M F, Patodi V K and Singer I M 1975a *Math. Proc. Camb. Phil. Soc.* **77** 43
—— 1975b *Math. Proc. Camb. Phil. Soc.* **78** 405
—— 1976 *Math. Proc. Camb. Phil. Soc.* **79** 71
Atiyah M F and Singer I M 1963 *Bull. Am. Math. Soc.* **69** 422
—— 1968 *Ann. Math., NY* **87** 546
Delbourgo R and Salam A 1972 *Phys. Lett.* **40B** 381
De Witt B S 1965 *Dynamical Theory of Groups and Fields* (New York: Gordon and Breach)
Dowker J S 1967a *Proc. R. Soc.* A **297** 351
—— 1967b *Suppl. Nuovo Cimento*, Ser. I **5** 734
Dowker J S and Critchley R 1976 *Phys. Rev.* D **13** 3224
—— 1977a *Phys. Rev.* D **15** 1484
—— 1977b *Phys. Rev.* D in the press
Dowker J S and Dowker Y P 1966a *Proc. Phys. Soc.* **87** 65
—— 1966b *Proc. R. Soc.* A **294** 175
Eguchi T and Freund P G O 1976 *Phys. Rev. Lett.* **37** 1251
Gilkey P B 1975a *J. Diff. Geom.* **10** 601
—— 1975b *Adv. Math.* **15** 334
Greiner P 1971 *Archs Ration. Mech. Analysis* **41** 163
Hagen C R 1969 *Phys. Rev.* **188** 2423
Harrington B J and Shepard H F 1977 *Nucl. Phys.* B **124** 409
Hawking S W 1977a *Commun. Math. Phys.* **55** 133
—— 1977b *Phys. Lett.* **60A** 81
Hirzebruch F 1966 *Topological Methods in Algebraic Geometry* (Berlin: Springer)
't Hooft G 1976a *Phys. Rev. Lett.* **37** 8
—— 1976b *Phys. Rev.* D **14** 3422
Jackiw R and Rebbi C 1977 *Phys. Rev.* D **16** 1052
Kimura T 1969 *Prog. Theor. Phys.* **42** 1191
Lichnerowicz A 1963 *C.R. Acad. Sci., Paris* A–B **257** 7
Minakshisundaram S and Pleijel A 1949 *Can. J. Math.* **1** 320
Nielsen N K, Römer H and Schroer B 1977 *CERN Report* TH2352
Nielsen N K and Schroer B 1977 *CERN Report* TH2317
Römer H and Schroer B 1977 *CERN Report* TH2357
Rumer G 1930 *Z. Phys.* **65** 244
Schouten J A 1954 *Ricci-Calculus* (Berlin: Springer)
Schwinger J 1951 *Phys. Rev.* **82** 664
—— 1970 *Particles, Sources and Fields* (New York: Addison-Wesley)
Seeley R T 1967 *Proc. Symp. on Pure Mathematics* vol. 10 (Providence, RI: American Mathematical Society) p 288