Commun. math. Phys. 55, 133—148 (1977)

Communications in Mathematical Physics

Zeta Function Regularization of Path Integrals in Curved Spacetime

S. W. Hawking

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 9EW, England

Abstract. This paper describes a technique for regularizing quadratic path integrals on a curved background spacetime. One forms a generalized zeta function from the eigenvalues of the differential operator that appears in the action integral. The zeta function is a meromorphic function and its gradient at the origin is defined to be the determinant of the operator. This technique agrees with dimensional regularization where one generalises to n dimensions by adding extra flat dimensions. The generalized zeta function can be expressed as a Mellin transform of the kernel of the heat equation which describes diffusion over the four dimensional spacetime manifold in a fith dimension of parameter time. Using the asymptotic expansion for the heat kernel, one can deduce the behaviour of the path integral under scale transformations of the background metric. This suggests that there may be a natural cut off in the integral over all black hole background metrics. By functionally differentiating the path integral one obtains an energy momentum tensor which is finite even on the horizon of a black hole. This energy momentum tensor has an anomalous trace.

1. Introduction

The purpose of this paper is to describe a technique for obtaining finite values to path integrals for fields (including the gravitational field) on a curved spacetime background or, equivalently, for evaluating the determinants of differential operators such as the four-dimensional Laplacian or D'Alembertian. One forms a gemeralised zeta function from the eigenvalues λ_n of the operator

$$\zeta(s) = \sum_n \lambda_n^{-s} . \tag{1.1}$$

In four dimensions this converges for $\mathrm{Re}(s) > 2$ and can be analytically extended to a meromorphic function with poles only at $s=2$ and $s=1$. It is regular at $s=0$. The derivative at $s=0$ is formally equal to $-\sum_n \log \lambda_n$. Thus one can define the determinant of the operator to be $\exp(-d\zeta/ds)|_{s=0}$.

In situations in which one knows the eigenvalues explicitly one can calculate the zeta function directly. This will be done in Section 3, for the examples of thermal radiation or the Casimir effect in flat spacetime. In more complicated situations one can use the fact that the zeta function is related by an inverse Mellin transform to the trace of the kernel of the heat equation, the equation that describes the diffusion of heat (or ink) over the four dimensional spacetime manifold in a fifth dimension of parameter time t. Asymptotic expansions for the heat kernel in terms of invariants of the metric have been given by a number of authors [1–4].

In the language of perturbation theory the determinant of an operator is expressed as a single closed loop graph. The most commonly used method for obtaining a finite value for such a graph in flat spacetime is dimensional regularization in which one evaluates the graph in n spacetime dimensions, treats n as a complex variable and subtracts out the pole that occurs when n tends to four. However it is not clear how one should apply this procedure to closed loops in a curved spacetime. For instance, if one was dealing with the four sphere, the Euclidean version of de Sitter space, it would be natural to generalize that S^4 to S^n [5, 6]. On the other hand if one was dealing with the Schwarzschild solution, which has topology $R^2 \times S^2$, one might generalize to $R^2 \times S^{n-2}$. Alternatively one might add on extra dimensions to the R^2. These additional dimensions might be either flat or curved. The value that one would obtain for a closed loop graph, would be different in these different extensions to n dimensions so that dimensional regularization is ambiguous in curved spacetime. In fact it will be shown in Section 5 that the answer given by the zeta function technique agrees up to a multiple of the undetermined renormalization parameter with that given by dimensional regularization where the generalization to n dimensions is given by adding on extra flat dimensions.

The zeta function technique can be applied to calculate the partition functions for thermal gravitons and matter quanta on black hole and de Sitter backgrounds. It gives finite values for these despite the infinite blueshift of the local temperature on the event horizons. Using the asymptotic expansion for the heat kernel, one can relate the behaviour of the partition function under changes of scale of the background spacetime to an integral of a quadratic expression in the curvature tensor. In the case of de Sitter space this completely determines the partition function up to a multiple of the renormalization parameter while in the Schwarzschild solution it determines the partition function up to a function of r_0/M where r_0 is the radius of the box containing a black hole of mass M in equilibrium with thermal radiation. The scaling behaviour of the partition function suggests that there may be a natural cut off at small masses when one integrates over all masses of the black hole background.

By functional differentiating the partition function with respect to the background metric one obtains the energy momentum tensor of the thermal radiation. This can be expressed in terms of derivatives of the heat kernel and is finite even on the event horizon of a black hole background. The trace of the energy momentum is related to the behaviour of the partition function under scale transformations. It is given by a quadratic expression in the curvature and is non zero even for conformally invariant fields [7–12].

The effect of the higher order terms in the path integrals is discussed in Section 9. They are shewn to make an insignificant contribution to the partition function for thermal radiation in a black hole background that is significantly bigger than the Planck mass. Generalised zeta functions have also been used by Dowker and Critchley [11] to regularize one-loop graphs. Their approach is rather different from that which will be given here.

2. Path Integrals

In the Feynmann sum over histories approach to quantum theory one considers expressions of the form

$$Z=\int d[g]d[\phi]\exp\{iI[g,\phi]\}\,, \tag{2.1}$$

where $d[g]$ is a measure on the space of metrics g, $d[\phi]$ is a measure on the space of matter fields ϕ and $I[g,\phi]$ is the action. The integral is taken over all fields g and ϕ that satisfy certain boundary or periodicity conditions. A situation which is of particular interest is that in which the fields are periodic in imaginary time on some boundary at large distance with period β [13]. In this case Z is the partition function for a canonical ensemble at the temperature $T=\dfrac{1}{\beta}$.

The dominant contribution to the path integral (2.1) will come from fields that are near background fields g_0 and ϕ_0 which satisfy the boundary or periodicity conditions and which extremise the action i.e. they satisfy the classical field equations. One can expand the action in a Taylor series about the background fields:

$$I[g,\phi]=I[g_0,\phi_0]+I_2[\tilde{g}]+I_2[\tilde{\phi}]+\text{higher order terms}, \tag{2.2}$$

where

$$g=g_0+\tilde{g}\,,\qquad \phi=\phi_0+\tilde{\phi}$$

and $I_2[\tilde{g}]$ and $I_2[\tilde{\phi}]$ are quadratic in the fluctuations $\tilde{g}$ and $\tilde{\phi}$. Substituting (2.2) into (2.1) and neglecting the higher order terms one has

$$\begin{aligned}\log Z=iI[g_0,\phi_0]&+\log\int d[\tilde{g}]\exp iI_2[\tilde{g}]\\&+\log\int d[\tilde{\phi}]\exp iI_2[\tilde{\phi}]\,.\end{aligned} \tag{2.3}$$

The background metric g_0 will depend on the situation under consideration but in general it will not be a real Lorentz metric. For example in de Sitter space one complexifies the spacetime and goes to a section (the Euclidean section) on which the metric is the real positive definite metric on a four sphere. Because the imaginary time coordinate is periodic on this four sphere, Z will be the partition function for a canonical ensemble. The action $I[g_0,\phi_0]$ of the background de Sitter metric gives the contribution of the background metric to the partition function while the second and third terms in Equation (2.3) give the contributions of thermal gravitons and matter quanta respectively on this background. In the case of the canonical ensemble for a spherical box with perfectly reflecting walls the background metric can either be that of a Euclidean space or it can be that of a section (the Euclidean section) of the complexified Schwarzschild solution on which the metric is real positive definite. Again the action of the background metric gives the contribution of the background metric to the partition function. This corresponds to an entropy equal to one quarter of the area of the event horizon in units in which $G=c=\hbar=k=1$. The second and third terms in Equation (2.3) give the contributions of thermal gravitons and matter quanta on a Schwarzschild background. In the case of the grand canonical ensemble for a box with temperature $T=\beta^{-1}$ and angular velocity Ω one considers fields which, on the walls of the box, have the same value at the point (t, r, θ, ϕ) and at the point $(t+i\beta, r, \theta, \phi+i\beta\Omega)$. This boundary cannot be filled in with any real metric but it can be filled in with a complex flat metric or with a complex section (the quasi Euclidean section [13]) of the Kerr solution. In both cases the metric is strongly elliptic (I am grateful to Dr. Y. Manor for this point) [14] if the rotational velocity of the boundary is less than that of light. A metric g is said to be strongly elliptic if there is a function f such that $\mathrm{Re}(fg)$ is positive definite. It seems necessary to use such strongly elliptic background metrics to make the path integrals well defined. One could take this to be one of the basic postulates of quantum gravity.

The quadratic term $I_2[\phi]$ will have the form

$$I_2[\tilde{\phi}] = -\int \tfrac{1}{2}\tilde{\phi} A \tilde{\phi}(-g_0)^{1/2} d^4x\,, \tag{2.4}$$

where A is a second order differential operator constructed out of the background fields g_0, ϕ_0. (In the case of the fermion fields the operator A is first order. For simplicity I shall deal only with boson fields but the results can easily be extended to fermions.) The quadratic term $I_2[\tilde{g}]$ in the metric fluctuations can be expressed similarly. Here however, the second order differential operator is degenerate i.e. it does not have an inverse. This is because of the gauge freedom to make coordinate transformations. One deals with this by taking the path integral only over metrics that satisfy some gauge condition which picks out one metric from each equivalence class under coordinate transformations. The Jacobian from the space of all metrics to the space of those satisfying the gauge condition can be regarded in perturbation theory as introducing fictitious particles known as Feynmann-de Witt [15, 16] or Fadeev-Popov ghosts [17]. The path integral over the gravitational fluctuations will be treated in another paper by methods similar to those used here for matter fields without gauge degrees of freedom.

In the case when the background metric g_0 is Euclidean i.e. real and positive definite the operator A in the quadratic term $I_2[\phi]$ will be real, elliptic and self-adjoint. This means that it will have a complete spectrum of eigenvectors ϕ_n with real eigenvalues λ_n:

$$A\phi_n = \lambda_n \phi_n\,. \tag{2.5}$$

The eigenvectors can be normalized so that

$$\int \phi_n \phi_m (g_0)^{1/2} d^4x = \delta_{nm}\,. \tag{2.6}$$

Note that the volume element which appears in the (2.6) is $(g_0)^{1/2}$ because g_0 is positive definite. On the other hand the volume element that appears in the action I is $(-g)^{1/2} = -i(g)^{1/2}$ where the minus sign corresponds to a choice of the direction of Wick rotation of the time axis into the complex plane.

If the background metric g_0 is not Euclidean, the operator A will not be self-adjoint. However I shall assume that the eigen functions ϕ_n are still complete. If this is so, one can express the fluctuation $\tilde{\phi}$ in terms of the eigen functions.

$$\tilde{\phi} = \sum a_n \phi_n \,. \tag{2.7}$$

The measure $d[\phi]$ on the space of all fields $\tilde{\phi}$ can then be expressed in terms of the coefficients a_n:

$$\mathrm{d}[\phi] = \prod_n \mu da_n \,, \tag{2.8}$$

where μ is some normalization constant with dimensions of mass or inverse length. From (2.5)–(2.8) it follows that

$$\begin{aligned} Z[\tilde{\phi}] &= \int d[\phi] \exp iI_2[\tilde{\phi}] \\ &= \prod_n \int \tfrac{1}{2}\mu da_n \exp(-\lambda_n a_n^2) \\ &= \prod_n \tfrac{1}{2}\mu\pi^{1/2}\lambda_n^{-1/2} \\ &= (\det(4\mu^{-2}\pi^{-1}A))^{-1/2} \,. \end{aligned} \tag{2.9}$$

3. The Zeta Function

The determinant of the operator A clearly diverges because the eigenvalues λ_n increase without bound. One therefore has to adopt some regularization procedure. The technique that will be used in this paper will be called the zeta function method. One forms a generalized zeta function from the eigenvalues of the operator A:

$$\zeta(s) = \sum_n \lambda_n^{-s} \,. \tag{3.1}$$

In four dimensions this will converge for $\mathrm{Re}(s) > 2$. It can be analytically extended to a merophorphic function of s with poles only at $s=2$ and $s=1$ [18]. In particular it is regular at $s=0$. The gradient of zeta at $s=0$ is formally equal to $-\sum_n \log\lambda_n$. One can therefore *define* $\det A$ to be $\exp(-d\zeta/ds|_{s=0})$ [19]. Thus the partition function

$$\log Z[\tilde{\phi}] = \tfrac{1}{2}\zeta'(0) + \tfrac{1}{2}\log(\tfrac{1}{4}\pi\mu^2)\zeta(0) \,. \tag{3.2}$$

In situations in which the eigenvalues are known, the zeta function can be computed explicitly. To illustrate the method, I shall treat the case of a zero rest mass scalar field ϕ contained in a box of volume V in flat spacetime at the temperature $T = \beta^{-1}$. The partition function will be defined by a path integral over all fields ϕ on the Euclidean space obtained by putting $\tau = it$ which are zero on the walls of the box and which are periodic in τ with period β. The operator A in the action is the negative of the four dimensional Laplacian on the Euclidean space. If the dimensions of the box are large compared to the characteristic wavelength β, one can approximate the spatial dependence of the eigenfunctions by plane waves with periodic boundary conditions. The eigenvalues are then

$$\lambda_n = (2\pi\beta^{-1}n)^2 + k^2 \tag{3.3}$$

and the density of eigenvalues in the continuum limit is

$$\frac{2V}{(2\pi)^3}\int d^3k \tag{3.4}$$

when $n>0$ and half that when $n=0$. The zeta function is therefore

$$\zeta(s) = \frac{4\pi V}{(2\pi)^3}\left\{\int dk k^{2-2s} + 2\sum_{n=1}^{\infty}\int dk k^2(4\pi^2\beta^{-2}n^2 + k^2)^{-s}\right\}. \tag{3.5}$$

The second term can be integrated by parts to give

$$-\frac{8\pi V}{(2\pi)^3}\sum_{n=1}^{\infty}\int dk(4\pi^2\beta^{-2}n^2 + k^2)^{-s+1}(2-2s)^{-1}. \tag{3.6}$$

Put $k = 2\pi n\beta^{-1}\sinh y$. This gives

$$\begin{aligned} &-\frac{8\pi V}{(2\pi)^3}\sum_{n=1}^{\infty}\int dy(2\pi\beta^{-1}n)^{-2s+3}(2-2s)^{-1}(\cosh y)^{-2s+3} \\ &= -\frac{8\pi V}{(2\pi)^3}(2\pi\beta^{-1})^{3-2s}\times\zeta_R(2s-3) \\ &\cdot(2-2s)^{-1}\times\frac{1}{2}\frac{\Gamma(1/2)\Gamma(s-3/2)}{\Gamma(s-1)}, \end{aligned} \tag{3.7}$$

where ζ_R is the usual Riemann zeta function $\sum_n n^{-s}$. The first term in (3.5) seems to diverge at $k=0$ when s is large and positive. This infra red divergence can be removed if one assumes that the box containing the radiation is large but finite. In this case the k integration has a lower cut off at some small value ε. If s is large, the k integration then gives a term proportional to ε^{3-2s}. When analytically continued to $s=0$, this can be neglected in the limit $\varepsilon \to 0$, corresponding to a large box.

The gamma function $\Gamma(s-1)$ has a pole at $s=0$ with residue -1. Thus the generalised zeta function is zero at $s=0$ and

$$\begin{aligned} \zeta'(0) &= 2\pi V\beta^{-3}\zeta_R(-3)\Gamma(1/2)\Gamma(-3/2) \\ &= \frac{\pi^2}{45}VT^3 \end{aligned} \tag{3.8}$$

thus the partition function for scalar thermal radiation at temperature T in a box of volume V is given by

$$\log Z = \frac{\pi^2 V T^3}{90}. \tag{3.9}$$

Note that because $\zeta(0)=0$, the partition function does not depend on the undetermined normalization parameter μ. However, this will not in general be the case in a curved space background.

From the partition function one can calculate the energy, entropy and pressure of the radiation.

$$E=-\frac{d}{d\beta}\log Z=\frac{\pi^2}{30}VT^4\,, \tag{3.10}$$

$$S=\beta E+\log Z=\frac{2\pi^2}{45}VT^3\,, \tag{3.11}$$

$$P=\beta^{-1}\frac{d}{dV}\log Z=\frac{\pi^2}{90}T^4\,. \tag{3.12}$$

One can calculate the partition functions for other fields in flat space in a similar manner. For a charged scalar field there are twice the number of eigenfunctions so that $\log Z$ is twice the value given by Equation (3.9). In the case of the electromagnetic field the operator A in the action integral is degenerate because of the freedom to make electromagnetic gauge transformations. One therefore has, as in the gravitational case, to take the path integral only over fields which satisfy some gauge condition and to take into account the Jacobian from the space of all fields satisfying the gauge condition. When this is done one again obtains a value $\log Z$ which is twice that of Equation (3.9). This corresponds to the fact that the electromagnetic field has two polarization states.

One can also use the zeta function technique to calculate the Casimir effect between two parallel reflecting planes. In this case instead of summing over all field configurations which are periodic in imaginary time, one sums over fields which are zero on the plates. Defining Z to be the path integral over all such fields over an interval of imaginary time τ one has

$$\log Z=\frac{\pi^2 A\tau b^{-3}}{720}\,, \tag{3.13}$$

where b is the separation and A the area of the plates. Thus the force between the plates is

$$F=\tau^{-1}\frac{d}{db}\log Z=-\frac{\pi^2 A b^{-4}}{240}\,. \tag{3.14}$$

4. The Heat Equation

In situations in which one does not know the eigenvalues of the operator A, one can obtain some information about the generalized zeta function by studying the heat equation.

$$\frac{d}{dt}F(x,y,t)+AF(x,y,t)=0 \tag{4.1}$$

here x and y represent points in the four dimensional spacetime manifold, t is a fifth dimension of parameter time and the operator A is taken to act on the first argument of F. With the initial conditions

$$F(x,y,0)=\delta(x,y) \tag{4.2}$$

the heat kernel F represents the diffusion over the spacetime manifold in parameter time t of a unit quantity of heat (or ink) placed at the point y at $t=0$. One can express F in terms of the eigenvalues and eigenfunctions of A:

$$F(x,y,t)=\sum_n \exp(-\lambda_n t)\,\phi_n(x)\phi_n(y)\,. \tag{4.3}$$

In the case of a field ϕ with tensor or spinor indices, the eigenfunctions will carry a set of indices at the point x and a set at the point y. If one puts $x=y$, contracts over the indices at x and y and integrates over all the manifold one obtains

$$Y(t)\equiv \int \mathrm{Tr}F(x,x,t)(g_0)^{1/2}d^4x=\sum_n \exp(-\lambda_n t)\,. \tag{4.4}$$

The generalized zeta function is related to $Y(t)$ by a Mellin transform:

$$\zeta(s)=\sum_n{}' \lambda_n^{-s}=\frac{1}{\Gamma(s)}\int_0^\infty t^{s-1}Y(t)dt\,. \tag{4.5}$$

A number of authors e.g. [1–4] have obtained asymptotic expansions for F and Y valid as $t\to 0^+$. In the case that the operator A is a second order Laplacian type operator on a four dimensional compact manifold.

$$Y(t)\sim\sum_n B_n t^{n-2}\,, \tag{4.6}$$

where the coefficients B_n are integrals over the manifold of scalar polynomials in the metric, the curvature tensor and its covariant derivatives, which are of order $2n$ in the derivatives of the metric

i.e. $$B_n=\int b_n(g_0)^{1/2}d^4x\,. \tag{4.7}$$

DeWitt [1,2] has calculated the b_n for the operator $-\Box+\xi R$ acting on scalars,

$$\begin{aligned} b_0&=(4\pi)^{-2}\\ b_1&=(4\pi)^{-2}(\tfrac{1}{6}-\xi)R\\ b_2&=(2880\pi^2)^{-1}\\ &\quad\cdot[R^{abcd}R_{abcd}-R^{ab}R_{ab}+30(1-6\xi)^2R^2+(6-30\xi)\Box R]\,. \end{aligned} \tag{4.8}$$

Note that b_1 is zero when $\xi=\frac{1}{6}$ which corresponds to a conformally invariant scalar field.

In the case of a non-compact spacetime manifold one has to impose boundary conditions on the heat equation and on the eigenfunctions of the operator A. This can be done by adding a boundary to the manifold and requiring the field or its normal derivative to be zero on the boundary. An example is the case of a black hole metric such as the Euclidean section of the Schwarzschild solution in which one adds a boundary at some radius $r=r_0$. This boundary represents the walls of a perfectly reflecting box enclosing the black hole. For a manifold with boundary the asymptotic expansion for Y takes the form [20].

$$Y(t)=\sum_n (B_n+C_n)t^{n-2}\,, \tag{4.9}$$

where, as before, B_n has the form (4.7) and

$$C_n=\int c_n(h)^{1/2}d^3x\,,$$

where c_n is a scalar polynomial in the metric, the normal to the boundary and the curvature and their covariant derivatives of order $2n-1$ in the derivatives of the metric and h is the induced metric on the boundary. The first coefficient c_0 is zero because their is no polynomial of order -1. McKean and Singer [3] showed that $c_1=\frac{-1}{48\pi^2}K$ when $\xi=0$ where K is the trace of the second fundamental form of the boundary. In the case of a Schwarzschild black hole in a spherical box of radius r_0, c_2 must be zero in the limit of large r_0 because all polynomials of degree 3 in the derivatives of the metric go down faster than r_0^{-2}.

In a compact manifold with or without boundary with a strongly elliptical metric g_0 the eigenvalues of a Laplacian type operator A will be discrete. If there are any zero eigenvalues they have to be omitted from the definition of the generalized zeta function and dealt with separately. This can be done by defining a new operator $\tilde{A}=A-P$ where P denotes projection on the zero eigenfunctions. Zero eigenvalues have important physical effects such as the anomaly in the axial vector current conservation [21, 22]. Let $\varepsilon>0$ be the lowest eigenvalue of $\tilde{A}$ (from now on I shall simply use A and assume that any zero eigenfunctions have been projected out). Then

$$\zeta(s)=\frac{1}{\Gamma(s)}\left[\int_0^1 t^{s-1}Y(t)dt+\int_1^\infty t^{s-1}Y(t)dt\right]. \tag{4.10}$$

As $t\to\infty$, $Y\to e^{-\varepsilon t}$. Thus the second integral in Equation (4.10) converges for all s. In the first integral one can use the asymptotic expression (4.9). This gives

$$\sum_n \frac{B_n+C_n}{n+s-2}\,. \tag{4.11}$$

Thus ζ has a pole at $s=2$ with residue B_0 and a pole at $s=1$ with residue B_1+C_1. There would be a pole at $s=0$ but it is cancelled out by the pole in $\Gamma(s)$. Thus $\zeta(0)=B_2+C_2$. Similarly the values of ζ at negative integer values of s are given by (4.11) and (4.10).

5. Other Methods of Regularization

A commonly used method to evaluate the determinant of the operator A is to start with the integrated heat kernel

$$Y(t)=\sum_n \exp(-\lambda_n t)\,. \tag{5.1}$$

Multiply by $\exp(-m^2 t)$ and integrate from $t=0$ to $t=\infty$

$$\int_0^\infty Y(t)\exp(-m^2 t)dt = \sum_n{}' (\lambda_n + m^2)^{-1} \tag{5.2}$$

then integrate over m^2 from $m^2=0$ to $m^2=\infty$ and interchange the orders of integration to obtain

$$\int_0^\infty t^{-1} Y(t)dt = \left[\sum_n{}' \log(\lambda_n + m^2)\right]_0^\infty . \tag{5.3}$$

One then throws away the value of the righthand side of (5.3) at the upper limit and claims that

$$\begin{aligned} \log \det A &= \sum \log \lambda_n \\ &= -\int_0^\infty t^{-1} Y(t)dt . \end{aligned} \tag{5.4}$$

This is obviously a very dubious procedure. One can obtain the same result from the zeta function method in the following way. One has

$$\begin{aligned} \log \det A &= -\zeta'(0) \\ &= -\frac{d}{ds}\left[\frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} Y(t)dt\right] . \end{aligned} \tag{5.5}$$

Near $s=0$

$$\frac{1}{\Gamma(s)} = s + \gamma s^2 + O(s^3) , \tag{5.6}$$

where γ is Euler's constant.

Thus

$$\begin{aligned} \log \det A = -\lim_{s\to 0} \Bigg[&(1+2\gamma s) \int_0^\infty t^{s-1} Y(t)dt \\ &+ (s+\gamma s^2) \int_0^\infty t^{s-1} \log t\, Y(t)dt \Bigg] . \end{aligned} \tag{5.7}$$

If one ignores the fact that the two integrals in Equation (5.7) diverged when $s=0$, one would obtain Equation (5.4). Using the asymptotic expansion for Y, one sees that the integral in Equation (5.4) has a t^{-2}, t^{-1}, and a $\log t$ divergence at the lower limit with coefficients $\frac{1}{2}B_0$, B_1, and B_2 respectively. The first of these is often subtraced out by adding an infinite cosmological constant to the action while the second is cancelled by adding an infinite multiple of the scalar curvature which is interpreted as a renormalization of the gravitational constant. The logarithmic term requires an infinite counter term of a new type which is quadratic in the curvature.

To obtain a finite answer from Equation (5.4) dimensional regularization is often used. One generalizes the heat equation from $4+1$ dimensions to $2\omega+1$ dimensions and then subtracts out the pole that occurs in (5.4) at $2\omega=4$. As mentioned in the introduction, this is ambiguous because there are many ways that one could generalize a curved spacetime to 2ω dimensions. The simplest generalization would be to take the product of the four dimensional spacetime manifold with $2\omega-4$ flat dimensions. In this case the integrated heat kernel Y would be multiplied by $(4\pi t)^{2-\omega}$. Then (5.4) would become

$$\log \det A = -\int_0^\infty t^{1-\omega}(4\pi)^{2-\omega}Y(t)dt\,. \tag{5.8}$$

This has a pole at $2\omega=4$ with residue $\zeta(0)$ and finite part $-\zeta'(0)+(2\gamma+\log 4\pi)\times\zeta(0)$. Thus, the value of the $\log Z$ derived by the dimensional regularization using flat dimensions agrees with the value obtained by the zeta function method up to a multiple of $\zeta(0)$ which can be absorbed in the normalization constant. However, if one extended to $2\omega+1$ dimensions in some more general way than merely adding flat dimensions, the integrated heat kernel would have the form

$$Y(t_0\omega)=\sum_n B_n(\omega)t^{n-\omega}\,, \tag{5.9}$$

where the coefficients $B_n(\omega)$ depend on the dimensions 2ω. The finite part at $\omega=2$ would then acquire an extra term $B_2'(2)$. This could not be absorbed in the normalization constant μ. One therefore sees that the zeta function method has the conceptual advantages that it avoids the dubious procedures used to obtain Equation (5.4), it does not require the subtraction of any pole term or the addition of infinite counter terms, and it is unambiguous unlike dimensional regularization which depends on how one generalizes to 2ω dimensions.

6. Scaling

In this Section I shall consider the behaviour of the partition function Z under a constant scale transformation of the metric

$$\tilde{g}_{ab}=kg_{ab}\,. \tag{6.1}$$

If A is a Laplacian type operator for a zero rest mass field, the eigenvalues transform as

$$\lambda_n=k^{-1}\lambda_n\,. \tag{6.2}$$

Thus the new generalized zeta function is

$$\tilde{\zeta}(s)=k^s\zeta(s) \tag{6.3}$$

and

$$\log \det \tilde{A}=\log \det A-\log k\zeta(0)\,. \tag{6.4}$$

Thus

$$\begin{aligned}\log \tilde{Z}=&\log Z+\tfrac{1}{2}\log k\zeta(0)\\&+(\log\tilde{\mu}-\log\mu)\zeta(0)\,.\end{aligned} \tag{6.5}$$

If one assumed that the normalization constant μ remained unchanged under a scale transformation, the last term would vanish. This assumption is equivalent to assuming that the measure in the path integral over all configurations of the field ϕ is defined not on a scalar field but on a scalar density of weight $\frac{1}{2}$. This is because the eigenfunctions of the operator A would have to transform according to

$$\tilde{\phi}_n = k^{-1}\phi_n \tag{6.6}$$

in order to maintain the normalization condition (2.6). The coefficients a_n of the expansion of a given scalar field ϕ would therefore transform according to

$$\tilde{a}_n = k a_n \tag{6.7}$$

and the normalization constant μ would transform according to

$$\tilde{\mu} = k^{-1}\mu \tag{6.8}$$

if the measure is defined on the scalar field itself, i.e. if

$$d[\phi] = \Pi_x d\phi(x)\,. \tag{6.9}$$

However if the measure is defined on densities of weight $\frac{1}{2}$, i.e.

$$d[\phi] = \Pi_x (g(x))^{1/4} d\phi(x) \tag{6.10}$$

then the normalization parameter μ is unchanged.

The weight of the measure can be deduced from considerations of unitarity. In the case of a scalar field one can use the manifestly unitary formalism of summing over all particle paths. This gives the conformally invariant scalar wave equation if the fields are taken to be densities of weight $\frac{1}{2}$ [23]. By contrast, the "minimally coupled" wave equation $\Box\phi = 0$ will be obtained if the weight is 1. In the case of a gravitational field itself one can use the unitary Hamiltonian formalism. From this Fadeev and Popov [17] deduce that the measure is defined on densities of weight $\frac{1}{2}$ and is scale invariant. Similar procedures could be used to find the weight of the measure for other fields. One would expect it to be $\frac{1}{2}$ for massless fields.

These scaling arguments give one certain amounts of information about the partition function. In DeSitter space they determine it up to the arbitrariness of the normalization parameter μ because DeSitter space is completely determined by the scale. Thus

$$\log Z = B_2 \log r/r_0\,, \tag{6.11}$$

where r is the radius of the space and r_0 is related to μ. In the case of a Schwarzschild black hole of mass M in a large spherical box of radius r_0,

$$\log Z = B_2 \log M/M_0 + f(r_0 M^{-1})\,, \tag{6.12}$$

where again M_0 is related to μ. If the radius of the box is large compared to M, one would expect that the partition function should approach that for thermal radiation at temperature $T = (8\pi M)^{-1}$ in flat space. Thus one would expect

$$f = \frac{r_0^3}{34560 M^3} + O\left(\frac{r_0^2}{M^2}\right). \tag{6.13}$$

It should be possible to verify this and to calculate the lower order terms by developing suitable approximations to the eigenvalues of the radial equation in the Schwarzschild solution. In particular f and $\log Z$ will be finite. This contrasts with the result that one would obtain if one naively assumed that the thermal radiation could be described as a fluid with a density of $\log Z$ equal to $\pi^2/_{90}\bar{T}^3$ where $\bar{T} = T(1-2Mr^{-1})^{-1/2}$ is the local temperature. Near the horizon $\bar{T}$ would get very large because of a blueshift effect and so $\log Z$ would diverge.

For a conformally invariant scalar field $B_2 = -\frac{2}{45}$ for DeSitter space and $\frac{1}{45}$ for the Schwarzschild solution. The fact that B_2 is positive in the latter case may provide a natural cut off in the path integral when one integrates over background metrics will all masses M. If the measure on the space of gravitational fields is scale invariant then the action of the background fields will give an integral of the form

$$\int_0^\infty \exp(-4\pi M^2)M^{-1}dM\,. \tag{6.14}$$

This converges nicely at large M but has a logarithmic divergence at $M=0$. However if one includes a contribution of the thermal radiation the integral is modified to

$$\int_0^\infty \exp(-4\pi M^2)M^{-1+B_2}dM\,. \tag{6.15}$$

This converges if B_2 is positive. Such a cut off can however be regarded as suggestive only because it ignores the contributions of high order terms which will be important near $M=0$. One might hope that these terms might in turn be represented by further black hole background metrics.

7. Energy-Momentum Tensor

By functionally differentiating the partition function one obtains the energy momentum tensor of the thermal radiation

$$T_{ab} = -2(g_0)^{-1/2}\frac{\delta \log Z}{\delta g_0^{ab}}\,. \tag{7.1}$$

The energy momentum tensor will be finite even on the event horizon of a black hole background metric despite the fact that the blueshifted temperature $\bar{T}$ diverges there. This shows that the energy momentum tensor cannot be that of a perfect fluid with pressure equal to one third the energy density.

One can express the energy momentum tensor in terms of derivatives of the heat kernel F:

$$\delta \log Z = \tfrac{1}{2}\delta\zeta'(0) - \mu^{-1}\delta\mu\zeta(0) - \tfrac{1}{2}\log(\tfrac{1}{4}\pi\mu^2)\delta\zeta(0)\,. \tag{7.2}$$

The second term on the right of (7.2) will vanish if one assumes that μ does not change under variations of the metric. This will be the case if the measure is defined on densities of weight $\frac{1}{2}$. The third term can be expressed as the variation of an integral quadratic in the curvature tensor and can be evaluated directly. To calculate the first term one writes

$$\zeta'(0)=\frac{d}{ds}\left[\frac{1}{\Gamma(s)}\int\int_0^\infty t^{s-1}F(x,x,t)\,(g_0)^{1/2}d^4x\,dt\right]\Bigg|_{s=0}. \tag{7.3}$$

Therefore

$$\delta\zeta'(0)=\frac{d}{ds}\left[\frac{1}{\Gamma(s)}\int\int_0^\infty t^{s-1}\delta[F(x,x,t)\,(g_0)^{1/2}d^4x\,dt\right]\Bigg|_{s=0}. \tag{7.4}$$

To calculate δF one uses the varied heat equation

$$\left(A+\frac{\partial}{\partial t}\right)\delta F(x,y,t)+\delta AF(x,y,t)=0 \tag{7.5}$$

with $\delta[(g_0(y))^{1/2}F(x,y,0)]=0$. The solution is

$$\delta[(g_0(y)^{1/2}F(x,y,t)]=-\int\int_0^t F(x,z,t-t')\delta AF(z,y,t')g_0(y)g_0(z)^{1/2}d^4z\,dt'. \tag{7.6}$$

Therefore

$$\delta\int F(x,x,t)\,(g_0)^{1/2}d^4x=-t\int\delta AF(z,z,t)\,(g_0)^{1/2}d^4z\,. \tag{7.7}$$

Where the operator δA acts on the first argument of F.

The operator δA involves δg^{ab} and its covariant derivatives in the background metric. Integrating by parts, one obtains an expression for T^{ab} in terms of F and its covariant derivatives. For a conformally invariant scalar field.

$$\begin{aligned}T_{ab}=\frac{d}{ds}\Bigg[\frac{1}{\Gamma(s)}\int_0^\infty t^s\big(&\tfrac{2}{3}{}_aF_b-\tfrac{1}{6}g_{ab\,c}F^c-\tfrac{1}{3}F_{ab}\\&+\tfrac{1}{3}g_{ab}F^c_c+\tfrac{1}{6}R_{ab}F-\tfrac{1}{12}g_{ab}RF\big)dt\Bigg]\\&-\log(\tfrac{1}{4}\pi\mu^2)\frac{\delta B_2}{\delta g_{ab}}(g_0)^{-1/2}\,.\end{aligned} \tag{7.8}$$

Where indices placed before or after F indicates differentiation with respect to the first or second arguments respectively and the two arguments are taken at the point x at which the energy momentum tensor is to be evaluated. In an empty spacetime the quantity B_2 is the integral of a pure divergence so B_2 vanishes.

8. The Trace Anomaly

Naively one would expect T^a_a, the trace of the energy momentum tensor, would be zero for a zero rest mass field. However this is not the case as can be seen either directly from (7.8) or by the following simple argument. Consider a scale transformation in which the metric is multiplied by a factor $k=1+\varepsilon$. Then $\delta g_{ab}=\varepsilon g_{ab}$ and

$$\begin{aligned}\int T^a_a(g_0)^{1/2}d^4x &= 2\frac{d\log Z}{dk}\\ &= B_2(1+\tfrac{1}{2}\mu^{-1}d\mu/dk)\\ &= B_2\end{aligned} \tag{8.1}$$

if the measure is defined on densities of weight $\frac{1}{2}$. Thus for the case of a conformally invariant scalar field

$$T^a_a=\frac{1}{2880\pi^2}\left[R_{abcd}R^{abcd}-R_{ab}R^{ab}+\Box R\right]. \tag{8.2}$$

The trace anomalies for other zero rest mass fields can be calculated in a similar manner.

These results for the trace anomaly agree with those of a number of other authors [7–12]. However, they disagree with some calculations by the point separation method [24] which do not obtain any anomaly. The trace anomaly for DeSitter completely determines the energy momentum because it must be a multiple of the metric by the symmetry. In a two dimensional black hole in a box the trace anomaly also determines the energy momentum tensor and in the four dimensional case it determines it up to one function of position [25].

9. Higher Order Terms

The path integral over the terms in the action which are quadratic in the fluctuations about the background fields are usually represented in perturbation theory by a single closed loop without any vertices. Functionally differentiating with respect to the background metric to obtain the energy momentum tensor corresponds to introducing a vertex coupling the field to the gravitational field. If one then feeds this energy momentum tensor as a perturbation back into the Einstein equations for the background field, the change in the $\log Z$ would be described by a diagram containing two closed loops each with a gravitational vertex and with the two vertices joined by a gravitational propagator. Under a scale transformation in which the metric was multiplied by a constant factor k, such a diagram would be multiplied by k^{-2}. Another diagram which would have the same scaling behaviour could be obtained by functionally differentiating $\log Z$ with respect to the background metric at two different points and then connecting these points by a gravitational propagator. In fact all the higher order terms have scaling behaviour k^{-n} where $n\geqq 2$. Thus one would expect to make a negligible contribution to the partition function for black holes of significantly more than the Planck mass. The higher order terms will however be important near the Planck mass and will cause the scaling argument in Section 6 to break down. One might nevertheless hope that just as a black hole background metric corresponds to an infinite sequence of higher order terms in a perturbation expansion around flat space, so the higher order terms in expansion about a black hole background might in turn be represented by more black holes.

Acknowledgement. I am grateful for discussions with a number of colleagues including G. W. Gibbons, A. S. Lapedes, Y. Manor, R. Penrose, M. J. Perry, and I. M. Singer.

References

1. DeWitt, B.S.: Dynamical theory of groups and fields in relativity, groups and topology (eds. C. M. and B. S. DeWitt). New York: Gordon and Breach 1964
2. DeWitt, B.S.: Phys. Rep. **19** C, 295 (1975)
3. McKean, H.P., Singer, J.M.: J. Diff. Geo. **5**, 233—249 (1971)
4. Gilkey, P.B.: The index theorem and the heat equation. Boston: Publish or Perish 1974
5. Candelas, P., Raine, D.J.: Phys. Rev. D **12**, 965—974 (1975)
6. Drummond, I.T.: Nucl. Phys. **94** B, 115—144 (1975)
7. Capper, D., Duff, M.: Nuovo Cimento **23** A, 173 (1974)
8. Duff, M., Deser, S., Isham, C.J.: Nucl. Phys. **111** B, 45 (1976)
9. Brown, L.S.: Stress tensor trace anomaly in a gravitational metric: scalar field. University of Washington, Preprint (1976)
10. Brown, L.S., Cassidy, J.P.: Stress tensor trace anomaly in a gravitational metric: General theory, Maxwell field. University of Washington, Preprint (1976)
11. Dowker, J.S., Critchley, R.: Phys. Rev. D **13**, 3224 (1976)
12. Dowker, J.S., Critchley, R.: The stress tensor conformal anomaly for scalar and spinor fields. University of Manchester, Preprint (1976)
13. Gibbons, G.W., Hawking, S.W.: Action integrals and partition functions in quantum gravity. Phys. Rev. D (to be published)
14. Manor, Y.: Complex Riemannian sections. University of Cambridge, Preprint (1977)
15. Feynman, R.P.: Magic without magic, (eds. J. A. Wheeler and J. Klaunder). San Francisco: W. H. Freeman 1972
16. DeWitt, B.S.: Phys. Rev. **162**, 1195—1239 (1967)
17. Fadeev, L.D., Popov, V.N.: Usp. Fiz. Nauk **111**, 427—450 (1973) [English translation in Sov. Phys. Usp. **16**, 777—788 (1974)]
18. Seeley, R.T.: Amer. Math. Soc. Proc. Symp. Pure Math. **10**, 288—307 (1967)
19. Ray, D.B., Singer, I.M.: Advances in Math. **7**, 145—210 (1971)
20. Gilkey, P.B.: Advanc. Math. **15**, 334—360 (1975)
21. 't Hooft, G.: Phys. Rev. Letters **37**, 8—11 (1976)
22. 't Hooft, G.: Computation of the quantum effects due to a four dimensional pseudoparticle. Harvard University, Preprint
23. Hartle, J.B., Hawking, S.W.: Phys. Rev. D **13**, 2188—2203 (1976)
24. Adler, S., Lieverman, J., Ng, N.J.: Regularization of the stress-energy tensor for vector and scalar particles. Propagating in a general background metric. IAS Preprint (1976)
25. Fulling, S.A., Christensen, S.: Trace anomalies and the Hawking effect. Kings College London, Preprint (1976)

Communicated by R. Geroch

Received February 10, 1977