On the Euclidean Approach to Quantum Field Theory in Curved Spacetime⋆

Robert M. Wald⋆⋆

Enrico Fermi Institute, University of Chicago, Chicago, Illinois 60637, USA

Abstract. Quantum field theory in curved spacetime is examined from the Euclidean approach, where one seeks to define the theory for metrics of positive (rather than Lorentzian) signature. Methods of functional analysis are used to give a proof of the heat kernel expansion for the Laplacian, which extends the well known result for compact manifolds to all non-compact manifolds for which the Laplacian and its powers are essentially self-adjoint on the initial domain of smooth functions of compact support. Using this result, precise prescriptions of the zeta-function, dimensional, and point-splitting type are given for renormalizing the action of a Klein-Gordon scalar field. These prescriptions are shown to be equivalent up to local curvature terms. It is also shown that for static spacetimes, the Euclidean prescription for defining the Feynman propagator agrees with the definition of Feyman propagator obtained by working directly on the spacetime.

I. Introduction

In quantum field theory, it is common to derive in a formal way expressions for quantities of physical interest. Taken literally, these expressions frequently are either meaningless or infinite. Thus, a major task in quantum field theory is to give a well defined, finite meaning to these formal expressions.

The theory of free quantum fields in curved spacetime is sufficiently simple that much of it – for example, the derivation of the S-matrix – can be mathematically formulated in a perfectly well defined manner (see e.g. [1]). However, even in this simple case there are two major problems that confront one: (1) the definition of "in" and "out" particle states (or, equivalently, the definition of the Feynman propagator) in general circumstances – in particular, for spacetimes with initial and/or final singularities; (2) the definition of finite, mathematically well defined expressions for physical quantities whose formal expressions are nonlinear in the

---

quantum field. A prime example of such a quantity is the stress-energy tensor of the quantum field.

One approach toward dealing with these issues is to work directly in the (Lorentz signature) spacetime under consideration. This approach has the advantage of yielding a direct, physical interpretation of the results obtained. However, with regard to problem (1) above, although some proposals have been given (see, e.g. [2, 3]), no natural generally applicable prescription for the Feynman propagator has been proven to exist. With regard to problem (2), Fulling et al. [4] have recently shown that if the two-point function $\langle \phi(x)\phi(x') + \phi(x')\phi(x)\rangle$ (where ϕ denotes the quantum field) is initially a distribution of the Hadamard form, then it is of the Hadamard form throughout the spacetime. Hence, in this case, the point-splitting prescription for renormalizing the stress-energy tensor is a mathematically well defined algorithm. That this prescription yields physically correct results (up to local curvature terms) can be seen from the fact that it satisfies conditions 1–4 of [5, 6] and that these conditions uniquely determine a prescription for $T_{\mu\nu}$ up to local curvature terms. Thus, provided the definition of states is such that the two-point function is of the Hadamard form, the problem of renormalizing the stress-energy tensor has been solved (up to local curvature terms) in this approach.

An alternative approach – which will be the subject of this paper – is to attempt to define all quantities on a Riemannian manifold (i.e., a positive definite metric). This has considerable mathematical advantages since elliptic operators are much easier to deal with than hyperbolic operators. However, the physical interpretation is much less direct since our universe is a Lorentzian rather than Riemannian spacetime. The point of view usually taken is that the Lorentz spacetime results are to be obtained by analytic continuation of the Riemannian results. This viewpoint has obvious difficulties, since even if one restricts oneself to analytic spacetimes, many Lorentz spacetimes of interest may not be sections of 4-dimensional complex manifolds which also have Riemannian sections (and even if they do, the Riemannian section might not be unique). Thus, the general issue of extracting physics from the Euclidean approach is a difficult one for which much further investigation is needed. Nevertheless, use of this approach in special cases where the analytic continuation can be done – most notably the Schwarzschild spacetime [7] – has elegantly reproduced results obtained with much greater labor by the direct spacetime approach. Furthermore, although the theory of quantum fields in curved spacetime is sufficiently simple that most results can be obtained by the direct spacetime approach, this is not true of quantum gravity itself, where most progress has been made via the Euclidean approach [8].

In this paper, we will focus on the formulation of quantum field theory in curved spacetime in the Euclidean approach, specifically on the two problems mentioned above. In the Euclidean approach, the definition of the "Feynman propagator" (i.e., the Green's function) is much more straightforward than in the Lorentz case. As discussed in Sect. II below, in many cases the "wave" operator for the field will uniquely determine a Green's function and even in those cases where a unique Green's function is not determined, there still exists a natural choice for it. Thus, the main problem on which we will concentrate is the renormalization of quantities of physical interest; specifically we will consider the effective

Lagrangian. (The stress-energy tensor can be obtained from the effective Lagrangian by functional differentiation with respect to the metric.) Our procedure for doing this is as follows: We express the effective Lagrangian as an integral involving the "heat operator" of the "wave" operator of the field. We show in Sect. III that if the "wave" operator and its powers are essentially self-adjoint on smooth functions of compact support – as is always the case for complete Riemannian manifolds – then the heat operator can be expressed as an integral kernel with singularity structure of the standard DeWitt-Schwinger form. (This is well known for the case of compact manifolds [9, 10].) Using this result, mathematically well defined prescriptions for the renormalized effective action can be given and equivalence of them (up to local curvature terms) will be shown in Sect. IV.

Finally, in Sect. V we shall examine static spacetimes. In this case, a "Euclidean section" exists, so the Euclidean approach can be used to define the Feynman propagator on the Lorentzian spacetime. On the other hand, since one has a natural notion of "positive frequency", the Feynman propagator could also be defined directly on the spacetime. We will show that these two definitions agree.

Throughout the paper we will consider only the Klein-Gordon scalar field, whose "wave" operator is

$$A = -\nabla_a \nabla^a + m^2$$

i.e., A is the Laplacian of the manifold plus a constant. We do not assume $m \neq 0$ (although if $m = 0$ and the manifold has finite volume we must project out the constant functions to make A positive definite). However, the results of this paper should be applicable to any other field whose Euclideanized "wave" operator is self-adjoint, elliptic, and positive definite.

II. Self-Adjointness, the Feynman Propagator, and Some Preliminaries

Let M be a manifold with Riemannian metric g_{ab} defined on it. We can view the operator

$$A = -\nabla_a \nabla^a + m^2$$

as an operator on the Hilbert space, $L^2(M)$, of square integrable functions on M. Initially, we may define A in the obvious way on the dense domain $C_0^\infty(M)$ of smooth (C^∞) functions of compact support. Defined in this way, A will be symmetric but not self-adjoint[1]. Our first task is to extend the domain of definition of A to make it self-adjoint.

Since A is unbounded, by the Hellinger-Toeplitz theorem [11, 12] no self-adjoint extension can act on all vectors in $L^2(M)$. The choice of functions that A acts on, i.e., the choice of self-adjoint extension of A, is intimately related to the choice of boundary conditions on the Green's function associated with A. Namely, once we have extended A to a self-adjoint operator (and assuming that the

---

1   The adjoint, $B^\dagger$, of an operator $B: H \to H$ is defined as follows: If $(y, Bx) = (z, x)$ for all $x \in \text{dom}(B)$, one says $y \in \text{dom}(B^\dagger)$ and defines $B^\dagger y = z$. B is *symmetric* if $\text{dom}(B^\dagger) \supset \text{dom}(B)$ and the restriction of $B^\dagger$ to $\text{dom}(B)$ equals B. B is *self-adjoint* if it is symmetric and $\text{dom}(B^\dagger) = \text{dom}(B)$

extended operator is positive), we may use the spectral theorem to define the "Green's operator" A^{-1} as a densely defined self-adjoint operator. [It will follow from later results (see Sect. IV) that if A and its powers are essentially self-adjoint (defined below) then in most cases Green's operator A^{-1} can be realized as a kernel $\Delta_F(x, x')$ with singularities only at $x = x'$.] Since the range of A^{-1} is, of course, precisely the domain of A, the specification of what functions A acts on is equivalent to the specification of what functions can result from action by the Green's function. Thus, the problem of defining the "Feynman propagator" (i.e., the Green's function A^{-1}) is equivalent to the problem of defining a self-adjoint extension of A.

Since A is positive on its initial domain C_0^∞, there always exist self-adjoint extensions of A – indeed, positive self-adjoint extensions of A always exist [13]. Thus, the only obstacle to defining the Feynman propagator is that more than one self-adjoint extension may exist. However, in one very important case – namely, when A is essentially self-adjoint on C_0^∞ – no ambiguity can arise. The notion of essential self-adjointness is defined as follows: The graph of A is defined as the subset of $L^2(M) \times L^2(M)$ consisting of the pairs $\langle x, Ax \rangle$ for all $x \in \mathrm{dom}(A)$. If we take the closure of this subset, it will – since A is symmetric – define the graph of an operator, denoted $\bar{A}$ (called the *closure* of A). $\bar{A}$ is clearly an extension of A. If $\bar{A}$ is self-adjoint, then A is said to be essentially self-adjoint and $\bar{A}$ is the unique self-adjoint extension of A. Note that the domain of $\bar{A}$ is the Cauchy completion of the domain of A in the norm $\|\chi\|^2 + \|A\chi\|^2$.

Thus, if A is essentially self-adjoint, there is a unique natural solution to the problem of defining the Feynman propagator. When is A essentially self-adjoint? Gaffney [15] has shown the essential self-adjointness of A for all manifolds with "negligible boundary". This includes all complete Riemannian manifolds as well as some incomplete ones (such as Euclidean space with a point removed). On the other hand, essential self-adjointness does not hold manifolds with boundary, nor would it be expected to hold for most manifolds with singularities.

If A is not essentially self-adjoint, then a choice of self-adjoint extensions must be made to define the "Feynman propagator". (In more intuitive terms, further boundary conditions must be imposed.) One condition to impose on the self-adjoint extension is that it preserve the positivity of A, but this condition alone need not uniquely determine a self-adjoint extension. It appears that the most natural choice of self-adjoint extension is the Friedrich's extension [13]. This extension preserves the positivity of A and is the unique extension with domain contained within the Cauchy completion of the original domain of A under the norm $\|\chi\|^2 + (\chi, A\chi)$. Since there are at present no strong grounds for arguing which definition of the "Feynman propagator" is physically correct, it cannot be claimed that this solves the Euclidean version of the problem of defining the Feynman propagator on an arbitrary curved spacetime (possibly with singularities), but it is interesting that a natural, mathematically well defined prescription always exists.

For the remainder of this paper we will focus on the second problem mentioned in the introduction, namely, the definition of finite, mathematically well defined expressions for quantities of physical interest, in particular, for the effective

action. For this purpose, we will restrict ourselves to the case where A is essentially self-adjoint on C_0^∞ and will further assume that powers A^k of A are also essentially self-adjoint. By a theorem of Cordes [16], this holds for all complete Riemannian manifolds. It would be interesting to see to what extent the results of this paper would also hold without these assumptions. Essential self-adjointness of A is certainly not necessary for the validity of the heat kernel expansion (see Sect. III) as it holds for compact manifolds with boundary [9] where A is not essentially self-adjoint. Indeed, it is possible that (choosing, say the Friedrich's extension of A) the heat kernel expansion is valid in all cases. The essential self-adjointness assumption will be used only in the proof of the following lemma. *Since only the self-adjoint extension $\bar{A}$ of A will be considered in the remainder of the paper, from this point on the bar will be dropped and A will now denote the self-adjoint extension.*

Lemma. *Let A and its powers be essentially self-adjoint on C_0^∞. Suppose $\psi \in L^2(M)$ is in the domain of A^k for all positive integers k. Then $\psi \in C^\infty$ and for every compact subset $\mathscr{C}$ of M and for k sufficiently large (specifically $4k > n = \dim(M)$) there is a constant K (independent of ψ) such that for all $x \in \mathscr{C}$ $|\psi(x)| \leq K(\|\psi\| + \|A^k \psi\|)$. Similar bounds also hold for the derivatives of ψ.*

Proof. We prove first that if $\psi \in \mathrm{dom}(A)$ and $f \in C_0^\infty$ then $f\psi \in \mathrm{dom}(A)$. Using the essential self-adjointness of A, we know that ψ is in the Cauchy completion of C_0^∞ under the norm $\|\chi\|^2 + \|A\chi\|^2$. Hence, there is a sequence $\{g_n\}$ of C_0^∞ functions which is Cauchy in this norm and converges to ψ in the L^2-norm. Consider the sequence $\{fg_n\}$. We have

$$\|fg_n - fg_m\|^2 \leq \max|f|^2 \|g_n - g_m\|^2. \tag{2.2}$$

Using the explicit form of A, Eq. (2.1), we also have,

$$A(fg_n) = f(Ag_n) + (\nabla_a \nabla^a f)g_n + 2\nabla_a f \nabla^a g_n. \tag{2.3}$$

Note that

$$
\begin{aligned}
\|\nabla_a f \nabla^a g_n\|^2 &= \int |\nabla_a f \nabla^a g_n|^2 \\
&\leq \max|\nabla_a f|^2 \int \nabla_a g_n^* \nabla^a g_n \\
&= \max|\nabla_a f|^2 (g_n, (A - m^2)g_n) \\
&\leq \max|\nabla_a f|^2 \|g_n\| \|Ag_n\|.
\end{aligned}
\tag{2.4}
$$

Consequently, we have

$$
\begin{aligned}
\tfrac{1}{3}\|A(fg_n - fg_m)\|^2 &\leq \max|f|^2 \|A(g_n - g_m)\|^2 \\
&\quad + \max|\nabla_a \nabla^a f|^2 \|g_n - g_m\|^2 \\
&\quad + 2\max|\nabla_a f|^2 \|g_n - g_m\| \|A(g_n - g_m)\|.
\end{aligned}
\tag{2.5}
$$

From Eqs. (2.2) and (2.5) it follows that the sequence $\{fg_n\}$ is Cauchy in the norm $\|\chi\|^2 + \|A\chi\|^2$ and consequently its L^2-limit, namely, $f\psi$, is in the domain of A.

By similar brute force arguments, one can show more generally that if $\psi \in \mathrm{dom}(A^k)$, then $f\psi \in \mathrm{dom}(A^k)$.

If we choose the support of f to be contained within a coordinate patch, we may view $f\psi$ as a function on $\mathbb{R}^n$. Since for functions with support in a fixed compact set the norm $\|\chi\|^2 + \|A^k\chi\|^2$ is equivalent to the $2k$-Sobolev norm on $\mathbb{R}^n$ (see [14]), it follows that $f\psi$ is in every Sobolev space. Thus, $f\psi$, and hence ψ itself, must be C^∞. The bounds on $\psi(x)$ and its derivatives follow in a straightforward way from the fact that the $2k^{\text{th}}$-Sobolev norm of $f\psi$ is bounded by $\|\psi\| + \|A^k\psi\|$ together with the standard Sobolev bounds [14] on $\mathbb{R}^n$.

The main utility of this lemma for our purposes is as follows. Using the spectral theorem [11, 12], we can define the "heat operator" $\exp(-\tau A)$ associated with the operator A by

$$\exp(-\tau A) = \int \exp(-\tau\lambda) dE_\lambda, \tag{2.6}$$

where E_λ is the spectral family of A. Since A is positive, $\exp(-\tau A)$ is a bounded operator and, indeed, for all positive k, $A^k \exp(-\tau A)$ is also bounded. Consequently, for every $\psi \in L^2(M)$, $\exp(-\tau A)\psi$ is in the domain of A^k for all k. Therefore, by the above lemma for any $\psi \in L^2(M)$, the function

$$h(\tau, x) = \exp(-\tau A)\psi \tag{2.7}$$

is C^∞ in x for each $\tau > 0$. Furthermore, by the type of argument used in the proof of Stone's theorem [11] it follows that, viewed as a τ-dependent vector in $L^2(M)$, h (as well as powers of A applied to h) is differentiable in τ with derivative

$$\frac{dh}{d\tau} = -Ah = -Ae^{-\tau A}\psi. \tag{2.8}$$

The bound given in the above lemma shows that this Hilbert space differentiability implies differentiability in the ordinary function sense i.e., for fixed x, $h(\tau, x)$ is differentiable in τ. More generally, we find that $h(\tau, x)$ is C^∞ in τ, and its τ-derivatives, all are of the form $A^j e^{-\tau A}\psi$ and therefore are all in the domain of A^k for all k. Consequently, they are C^∞ functions of x and thus $h(\tau, x)$ is jointly C^∞ for $\tau > 0$. If ψ itself is in the domain of A^k for every k (e.g., if $\psi \in C_0^\infty$), this result can be extended to include $\tau = 0$. Thus, noting that $\|e^{-\tau A}\psi\| \le \|\psi\|$ we have shown the following result which will be used in Sect. III.

Proposition. *If A and its powers are essentially self-adjoint on C_0^∞ and $\psi \in C_0^\infty$, then, for $\tau \ge 0$, $h(\tau, x) = e^{-\tau A}\psi$ is C^∞ in τ and x and for any compact set $\mathscr{C}$ and $k > n/4$ there exists a constant K' such that for all $x \in \mathscr{C}$ and all $\tau \ge 0$, $|h(\tau, x)| \le K'(\|\psi\| + \|A^k\psi\|)$.*

The reason for introducing the heat operator $e^{-\tau A}$ is that quantities of physical interest can be expressed in terms of it. For example, using the functional calculus of self-adjoint operators (see [11, 12]), the "Feynman propagator" A^{-1}

can be expressed as,

$$A^{-1} = \int\limits_0^\infty e^{-\tau A} d\tau . \tag{2.9}$$

[Equation (2.9) – as well as the other similar equations below – is to be understood in the following sense: For each $\psi \in \mathrm{dom}(A^{-1})$, the integral of the vector $e^{-\tau A}\psi$ converges strongly to $A^{-1}\psi$.] In Sect. III we shall show that $e^{-\tau A}$ can be expressed as a smooth integral kernel for $\tau > 0$ with a calculable asymptotic behavior as $\tau \to 0$. Thus, using this result and barring the possibility of "infra-red" divergences discussed in Sect. IV, one can show that A^{-1} can be expressed as a kernel $\Delta_F(x, x')$ which is smooth except for $x = x'$ and the divergence behavior as $x \to x'$ can be calculated.

Our main focus of attention for the remainder of the paper will be on the effective action, S, of the quantum field. In the path integral approach, the effective action is formally given by an integral over field configurations

$$e^{-S} = \int [d\phi] e^{-1/2(\phi, A\phi)}. \tag{2.10}$$

If we pretend that the space of field configurations is finite dimensional (so that ϕ is a vector in a finite dimensional vector space and A is a self-adjoint linear map on that space), the integral would simply be a Gaussian, and ignoring overall numerical factors (which affect S only by an additive constant), we would have

$$e^{-S} = (\det A)^{-1/2} = e^{-1/2 \operatorname{tr} \ln A}. \tag{2.11}$$

If, at this point, we return to infinite dimensions but now attempt to view $\ln A$ as an integral kernel $L(x, x')$, we have,

$$S = \tfrac{1}{2} \int L(x, x) \sqrt{g} \, d^4 x . \tag{2.12}$$

Thus, the effective Lagrangian is

$$\mathscr{L}(x) = \tfrac{1}{2} L(x, x) . \tag{2.13}$$

Quantities of physical interest can be obtained from $\mathscr{L}$. In particular, the stress-energy tensor is given by,

$$\frac{\langle 0_{\mathrm{out}} | T_{\mu\nu} | 0_{\mathrm{in}} \rangle}{\langle 0_{\mathrm{out}} | 0_{\mathrm{in}} \rangle} = \frac{2}{\sqrt{g}} \frac{\delta(\mathscr{L} \sqrt{g})}{\delta g^{\mu\nu}} . \tag{2.14}$$

The derivation above is, of course, entirely formal. As we shall see below $\ln A$ can indeed be viewed as an integral kernel for $x \neq x'$ but, like A^{-1}, is singular for $x = x'$. Thus, Eq. (2.13) is meaningless and our aim is to give it a finite, well-defined meaning. We will do so by expressing it in terms of the heat operator. From the functional calculus [11, 12] together with the corresponding equation for positive real numbers, we have

$$\ln A = \lim_{\varepsilon \to 0} \left\{ - \int\limits_\varepsilon^\infty e^{-\tau A} \frac{d\tau}{\tau} + (\gamma - \ln \varepsilon) I \right\}, \tag{2.15}$$

where I denotes the identity operator and γ is Euler's constant,

$$\gamma = \int_0^\infty e^{-\lambda} \ln \lambda \, d\lambda . \tag{2.16}$$

Thus, by determining the structure of the heat kernel we can develop prescriptions to give Eq. (2.13) well defined meaning.

III. The Heat Kernel

In this section we will show that the heat operator $e^{-\tau A}$ can be expressed as a smooth integral kernel and derive an asymptotic expansion for it near $\tau = 0$. These results are well known for the case of a compact manifold; only the method of proof and the extension to non-compact cases where A and its powers are essentially self-adjoint are in any way new.

We begin by explicitly constructing what mathematicians would call a parametrix and what physicists would call a WKB approximation to $e^{-\tau A}$. Following the mathematicians Minakshisundaram and Pleijel [9], Patodi [10], and others and the physicists Schwinger [17], DeWitt [18], and others, we define for $\tau > 0$,

$$F(\tau, x, x') = \chi(x, x') e^{-m^2\tau} \frac{e^{-\sigma/4\tau}}{(4\pi\tau)^{n/2}} \sum_{j=0}^N A_j(x, x') \tau^j . \tag{3.1}$$

Here σ is the square of the geodesic distance between x and x', $\chi(x, x')$ is a smooth function which, for each fixed x' is identically 1 in a neighborhood of x' but vanishes outside a normal neighborhood of x', and the $A_j(x, x')$ are determined by recursion relations obtained by formally substituting F (with χ set equal to 1) into the heat equation. The presence of χ is necessary to keep the expression for F (which involves σ) well defined and smooth when more than one geodesic connects x and x'. Note also that the series is truncated at a finite integer N. The infinite series probably does not converge in general.

The function F has the following properties: It is smooth in x, x', and τ. If we view F as an operator on $L^2(M)$ via

$$(F\psi)(x) = \int F(\tau, x, x') \psi(x') dV' \tag{3.2}$$

then F approaches the identity operator as $\tau \to 0$. Finally, F nearly satisfies the heat equation in the sense that if we define $p(\tau, x, x')$ by

$$p(\tau, x, x') = \left(\frac{\partial}{\partial \tau} + A \right) F(\tau, x, x') \tag{3.3}$$

(where A acts on the x-variable) then p is C^∞ in x, x', τ, and for fixed x', is of compact support in x and, most importantly, for fixed x and x', $p(\tau, x, x')$ vanishes as $\tau \to 0$ at least as rapidly as $\tau^{N-n/2-2}$,

$$|p(\tau, x, x')| \leq C\tau^{N-n/2-2} . \tag{3.4}$$

The L^2-norm of p (viewed as a function of x with x' fixed) as well as powers of A applied to p are also bounded in τ in a similar manner. The full details of the

construction of F and the demonstration of the above statement can easily be extracted from the above-cited references and will not be reproduced here. Note that the construction of F does not depend on the nature of the manifold and would also work for a metric of non-positive signature.

It is sometimes assumed (either implicitly or explicitly) by physicists that F (with $\chi = 1$ and $N = \infty$) *is* the heat operator $e^{-\tau A}$. However, except in trivial cases (such as flat space) this cannot be true. In the first place, without the presence of χ and a finite N, the expression for F is ill defined. More fundamentally, the construction of F is entirely local while the heat operator $e^{-\tau A}$ is highly non-local, i.e., a bump of curvature in one place will affect the heat kernel everywhere. Nevertheless, we shall now show that the difference between $e^{-\tau A}$ and F is a smooth function which vanishes rapidly as $\tau \to 0$. Thus, $e^{-\tau A}$ is given by a smooth integral kernel and F provides an asymptotic expansion for $e^{-\tau A}$ about $\tau = 0$. The idea of the proof is to analyze the correction term Q (defined below) which needs to be subtracted from F to make it satisfy the heat equation. Then we use the heat equation evolution[2] to show that $F - Q$ must in fact be the heat operator $e^{-\tau A}$.

To begin, we define $q(\tau, \tau', x, x')$ for $\tau \geq \tau' \geq 0$ by

$$q(\tau, \tau', x, x') = e^{-(\tau - \tau')A} p(\tau', x, x'), \tag{3.5}$$

where here and in the following the operator acts on the x-variable. We claim that q is C^∞ in all its variables. Namely, for fixed τ' and x', smoothness of q in τ and x follows directly from the proposition of Sect. II. Smoothness of q in the remaining variables τ', x' follows from arguments similar to those used in the proof of the proposition, using also the smoothness of p in these variables. Furthermore, the proposition of Sect. II implies that for each τ' and x' and for x in a fixed compact set $\mathscr{C}$, we have

$$|q(\tau, \tau', x, x')| \leq K_1(\tau', x')(\| p(\tau', \cdot, x')\| + \| A^k p(\tau', \cdot, x')\|). \tag{3.6}$$

Using the bounds on p for small τ, we find that given any integer $M > 0$ we can always choose the series truncation integer N sufficiently large that,

$$|q(\tau, \tau', x, x')| \leq K_2(\tau', x') \tau'^M. \tag{3.7}$$

Letting $K = \max K_2(\tau', x')$ over all $x' \in \mathscr{C}$ and $\tau' \in [0, \tau]$ we find that for all $x, x' \in \mathscr{C}$,

$$|q(\tau, \tau', x, x')| \leq K \tau^M. \tag{3.8}$$

Next, we define the "correction term" Q by

$$Q(\tau, x, x') = \int_0^\tau q(\tau, \tau', x, x') d\tau'. \tag{3.9}$$

Thus Q is smooth and vanishes as $\tau \to 0$; in particular, for $x, x' \in \mathscr{C}$,

$$|Q(\tau, x, x')| \leq K \tau^{M+1}. \tag{3.10}$$

---

2   The idea of using the heat equation evolution was first suggested to me by M. Sweeny

Furthermore Q satisfies

$$\frac{\partial Q}{\partial \tau} = q(\tau, \tau, x, x') + \int_0^\tau \frac{\partial q}{\partial \tau} \, d\tau' \tag{3.11}$$

$$= p(\tau, x, x') - AQ. \tag{3.12}$$

Consequently, using the definition of p, Eq. (3.3), we find that

$$\frac{\partial}{\partial \tau}(F - Q) = -AF + p - p + AQ = -A(F - Q) \tag{3.13}$$

that is,

$$H(\tau, x, x') \equiv F(\tau, x, x') - Q(\tau, x, x') \tag{3.14}$$

satisfies the heat equation, and thus is an excellent candidate to be the heat operator $e^{-\tau A}$.

To prove that H is indeed the heat operator, we define the operator $\mathcal{O}(\tau)$ by

$$\mathcal{O}(\tau) = e^{-\tau A} - H(\tau), \tag{3.15}$$

where H is now interpreted as an integral kernel operator as in Eq. (3.2). Our aim is to prove that $\mathcal{O}(\tau) = 0$. We shall do so by showing that sufficiently many of its matrix elements vanish. Letting $\psi \in C_0^\infty$ (to avoid any potential domain problems or convergence problems with the integrals), we have

$$\frac{d}{d\tau}(\mathcal{O}(\tau)\psi) = -Ae^{-\tau A}\psi - \frac{\partial H}{\partial \tau}\psi$$

$$= -A\mathcal{O}(\tau)\psi. \tag{3.16}$$

Furthermore, since $e^{-\tau A}$ and F both approach the identity operators as $\tau \to 0$, whereas using Eq. (3.5) and the bounds on p, it is not difficult to show that $\|Q\psi\|$ goes to zero, we have $\|\mathcal{O}(\tau)\psi\| \to 0$ as $\tau \to 0$. Thus, if we fix $\tau_0 > 0$, let $\psi' \in L^2(M)$, and define

$$b(\tau) = (e^{-(\tau_0 - \tau)A}\psi', \mathcal{O}(\tau)\psi) \tag{3.17}$$

we have $b(0) = 0$. Furthermore,

$$\frac{db}{d\tau} = (Ae^{-(\tau_0 - \tau)A}\psi', \mathcal{O}(\tau)\psi) + (e^{-(\tau_0 - \tau)A}\psi', -A\mathcal{O}(\tau)\psi)$$

$$= 0, \tag{3.18}$$

Thus, we have

$$b(\tau) = 0 \tag{3.19}$$

for all $\tau < \tau_0$. Since this is true for all $\psi \in C_0^\infty$, all $\psi' \in L^2(M)$, and all τ_0, it follows that $\mathcal{O}(\tau) = 0$, i.e.,

$$e^{-\tau A} = H(\tau, x, x'). \tag{3.20}$$

Thus, recalling the definition of H and the fact that Q rapidly goes to zero as $\tau \to 0$, we have the following theorem:

Theorem. *If A and its powers are essentially self-adjoint on C_0^∞, then for $\tau > 0$ the heat operator $e^{-\tau A}$ is given by a smooth integral kernel $H(\tau, x, x')$ defined by Eq. (3.14). Furthermore, $F(\tau, x, x')$ defined by Eq. (3.1) yields an asymptotic expansion of the heat kernel about $\tau = 0$.*

(Please see "Note Added in Proof" at end of article.)

IV. Renormalization of the Effective Action

As already mentioned in Sect. II, A^{-1} (as well as all powers A^s of A) and $\ln A$ can be expressed in terms of the heat operator by integrals of the form

$$Z(s) = \int_0^\infty e^{-\tau A} \tau^{s-1} d\tau. \tag{4.1}$$

Specifically, for $s \neq 0$, we have

$$A^s = Z(-s)/\Gamma(-s) \tag{4.2}$$

whereas $Z(0)$ is related to $\ln A$ by Eq. (2.15). Since for $\tau > 0$, $e^{-\tau A}$ is given by a smooth integral kernel $H(\tau, x, x')$, the operator $Z(s)$ can fail to be given by a smooth integral kernel $Z(s, x, x')$ only on account of divergences in the integral at the upper and lower ends. We will refer to divergences at the upper end $(\tau \to \infty)$ as "infra-red divergences", while divergences at the lower end $(\tau \to 0)$ will be called "ultra-violet". We first examine the infra-red divergences.

There are two important cases where infra-red divergences cannot occur in $Z(s, x, x')$ for any value of s: (a) when $m \neq 0$ or (b) when the manifold M is compact. The reason is that in both these cases the spectrum of A will have a positive lower bound. In case (a), $\lambda_0 = m^2$, is such a bound. In case (b), since $e^{-\tau A}$ is given by an integral kernel and M is compact, it follows that $e^{-\tau A}$ is a compact operator [12] and since it is self-adjoint by the Hilbert-Schmidt theorem [12] it has a discrete spectrum (with zero the only possible accumulation point). Thus the spectrum of A itself is discrete. Since zero is not an eigenvalue of A – if $m = 0$ then as remarked at the end of Sect. I, we project the constants out of $L^2(M)$ – A has a minimum eigenvalue $\lambda_0 > 0$ which yields a positive bound for A. Consequently, in both cases (a) and (b) we have,

$$\|e^{-\tau A}\| \leq e^{-\tau \lambda_0}. \tag{4.3}$$

Using the type of estimates derived in Sect. II, we can exponentially bound the function $H(\tau, x, x')$ in τ and thus assure convergence of the upper limit $(\tau \to \infty)$ of the integral, Eq. (4.1), for any value of s.

If $m = 0$ and the manifold is non-compact, we can still bound $e^{-\tau A}$ by a constant and thereby bound $H(\tau, x, x')$ by a constant as $\tau \to \infty$. This ensures that no infra-red divergences can occur if $s < 0$. If $s = 0$ (the case of interest for the effective Lagrangian, $\ln A$) it seems highly unlikely that infra-red divergences ever occur, as only a very modest fall-off in τ of $H(\tau, x, x')$ is required for convergence of this integral. However, infra-red divergences for the Green's function or "Feynman propagator" $A^{-1} = Z(1)$ can occur. The heat kernel in n-dimensional Euclidean space in the massless case is simply

$$H_E(\tau, x, x') = \frac{e^{-\sigma/4\tau}}{(4\pi\tau)^{n/2}}. \tag{4.4}$$

Consequently, if $n=1$ or 2, the integral for $Z(1, x, x')$ will fail to converge. This type of infra-red divergence behavior for a massless quantum field in two dimensions has been noted previously [19]. However, by the same reasoning, infra-red divergence will occur for the Green's function in 4-dimensions if the Riemannian manifold M is a product $M_3 \times \mathbb{R}$ or $M_2 \times \mathbb{R}^2$ (with a flat metric on $\mathbb{R}^2$) with M_3 and M_2 compact (unless all functions which are constant on M_3 and M_2 are projected out of $L^2(M)$). For larger values of s, infra-red divergences in $Z(s)$ would, of course, become a much more common phenomenon in the massless non-compact case.

We now turn to the ultra-violet divergences. The asymptotic expansion $F(\tau, x, x')$, Eq. (3.1), for the heat kernel gives us all the information we need to calculate these divergences. First note that for $x \neq x'$, F vanishes as $\tau \to 0$ faster than any power of τ. Thus, the integral [Eq. (4.1)] is well behaved as $\tau \to 0$ for $x \neq x'$. Consequently, *except for the possibility of the infra-red divergences discussed above,* $Z(s)$ *(and thus in particular A^{-1} and $\ln A$) is given by an integral kernel $Z(s, x, x')$ which can be singular only in the coincidence limit $x = x'$.* Furthermore, the formula for F shows that if $s > n/2$ the integral is well behaved even at $x = x'$, so in this case (barring infra-red divergences) $Z(s, x, x')$ is continuous everywhere.

For $s \leqq n/2$, $Z(s, x, x')$ is singular when $x = x'$, but the divergence behavior as $x \to x'$ can easily be calculated from F. This has been done for the "Feynman propagator" $A^{-1} = \Delta_F(x, x')$ by DeWitt [18] and others although an asymptotic expansion in powers of $1/m^2$ (which, for the purpose of studying the singularities structure is entirely unnecessary) is also made there. It is found that $\Delta_F(x, x')$ has singularity structure of the Hadamard form [4] as $x \to x'$. Since, in spacetime, the Feynman propagator is equal to the anti commutator of the field operator at space-like separations, this suggests – but does not prove – that the anticommutator will be a Hadamard distribution in spacetime if definitions are made by analytic continuation from a Euclidean section. The importance of this remark is discussed in [4] and at the end of Sect. V below.

We now consider the problem of renormalizing the effective Lagrangian, Eq. (2.13), i.e., giving finite, well defined meaning to the values of $\ln A$ viewed as an integral kernel $L(x, x')$ in the coincidence limit $x = x'$. Our first step is to throw away the δ-function (i.e., the multiple of the identity operator) term in Eq. (2.15) thereby writing,

$$L(x, x') = - Z(0, x, x').$$
(4.5)

In particular, we have formally,

$$\mathscr{L}(x) = \tfrac{1}{2} L(x, x) = - \tfrac{1}{2} \int_0^\infty H(\tau, x, x) \frac{d\tau}{\tau}.$$
(4.6)

As remarked above, it is extremely unlikely that infra-red divergences can ever occur for L (even for non-compact manifolds with $m = 0$) so we will focus only on the ultra-violet divergences. From the formula for F, it can be seen that (for $n = 4$) only the first three terms ($j = 0, 1, 2$) of the series contribute divergences to $L(x, x)$. The important point to notice is that, in this well defined sense, the divergences in $\mathscr{L}(x)$ are proportional to $A_0(x, x)$, $A_1(x, x)$, and $A_2(x, x)$. Since the recursion

relations for the $A_j(x, x')$ determine these quantities in a completely local manner, the divergences in the effective action are thus proportional to local curvature terms. Indeed, $A_0(x, x) = 1$, and on dimensional grounds it is easy to see that $A_1(x, x)$ is proportional to the scalar curvature R, whereas $A_2(x, x)$ is quadratic in the curvature. The prescriptions we shall consider for making $\mathscr{L}(x)$ finite all involve the subtraction of terms proportional to A_0, A_1, and A_2 in the heat kernel. The subtraction of terms proportional to A_0 and A_1 can be interpreted, respectively, as a renormalization of the cosmological constant and the gravitational constant, as these terms are of the same form as appear in the action of the gravitational field. The A_2 term, however, does not appear in the Einstein gravitational action.

Perhaps the simplest prescription for defining $\mathscr{L}(x)$ is to subtract from $H(\tau, x, x)$ the term

$$e^{-\tau m^2}(4\pi\tau)^{-2}[A_0(x, x) + A_1(x, x)\tau + A_2(x, x)\tau^2]$$

thereby removing the divergence of the integral, Eq. (4.6), as $\tau \to 0$. The factor of $e^{-\tau m^2}$ (which is not essential for cancelling the ultraviolet divergences) ensures that the entire heat kernel will be subtracted in flat spacetime and also improves the infra-red convergence. However, if $m = 0$, the subtraction of the term $A_2(x, x)$ will result in an infra-red divergence in the integral for $\tau \to \infty$. To avoid this difficulty, we introduce a cutoff in the last term at $\tau = \mu$ by multiplying $A_2(x, x)$ by $\theta_\mu(\tau)$ where $\theta_\mu(\tau) = 1$ for $\tau \leq \mu$ and $\theta_\mu(\tau) = 0$ for $\tau > \mu$. We cannot simply set μ equal to a definite numerical value (e.g., $\mu = 1$) because we want $\mathscr{L}$ to change only by an overall scale factor when we change the metric by a scale factor $g_{ab} \to \Lambda^2 g_{ab}$ (with Λ constant) so that physical predictions will remain unchanged. This requires that μ must transform as $\mu \to \Lambda^2 \mu$, i.e., μ has dimensions of (length)^2. If $m \neq 0$ we get no infra-red divergence even if we do not introduce the cutoff μ. However, this would still be unacceptable because the limit as $m \to 0$ of the massive theory would be singular rather than continuously approaching the $m = 0$ theory. Thus, we define a prescription for $\mathscr{L}(x)$, which we will call *heat kernel renormalization* as follows: We define

$$K_\mu(\tau, x) = H(\tau, x, x) - e^{-\tau m^2}(4\pi\tau)^{-2}\{A_0(x, x)$$

$$+ A_1(x, x)\tau + \theta_\mu(\tau)A_2(x, x)\tau^2\} \qquad (4.7)$$

and set

$$\mathscr{L}(x) = -\tfrac{1}{2}\int_0^\infty K_\mu(\tau, x)\frac{d\tau}{\tau}. \qquad (4.8)$$

This yields a mathematically well defined formula for $\mathscr{L}(x)$.

A change in the value of μ merely changes $\mathscr{L}(x)$ by adding a multiple of $A_2(x, x)$. However, an important contribution is made by μ when we vary $\mathscr{L}$ with respect to $g^{\mu\nu}$ to obtain the stress energy tensor. Assuming that μ obeys the above scaling law when the metric is scaled but remains unchanged when the metric is

infinitesimally varied by a tensor whose integrated trace is zero, we have $\dfrac{\delta\mu}{\delta g^{\mu\nu}}$ $=\mu g_{\mu\nu}$. Consequently, an "anomalous term" proportional to $g_{\mu\nu}A_2(x,x)$ will appear in the expression for $T_{\mu\nu}$ due to the variation of μ. This is identical to the "correction term" added in to a point-splitting renormalization prescription for $T_{\mu\nu}$ which was required in order to make $T_{\mu\nu}$ conserved [6]. The trace of this term yields the well known trace anomaly.

The above heat kernel renormalization prescription is rather ad hoc. We can reformulate it in a possibly more appealing manner as follows: As mentioned above, except possibly for infra-red divergences, $Z(s,x,x')$ is well defined for all x,x' (in particular for $x=x'$) if $s>2$. Indeed, allowing s to be complex, $Z(s,x,x)$ is an analytic function of s for $\operatorname{Re} s>2$. We can extend the definition of $Z(s,x,x)$ to $\operatorname{Re} s\leqq2$ by analytic continuation. We consider first the case where no infra-red divergences occur for any value of s (i.e., $m\neq0$ or M compact). Since $\tau^2 H(\tau,x,x)$ is smooth in τ, for $\operatorname{Re} s>2$ we may integrate by parts to obtain,

$$Z(s,x,x)=\int\limits_0^\infty (\tau^2 H(\tau,x,x))\tau^{s-3}\,d\tau$$

$$=-\frac{1}{s(s-1)(s-2)}\int\limits_0^\infty (\tau^2 H)'''\tau^s d\tau, \tag{4.9}$$

where the primes denote derivatives with respect to τ. The right-hand side of Eq. (4.9) defines an analytic function of s for $\operatorname{Re} s>-1$ with simple poles at $s=0,1,2$. We may define $-2\mathscr{L}(x)$ by subtracting the pole at $s=0$ from this function to define an analytic function $\mathscr{L}(s,x)$ and then setting $s=0$. Noting that

$$\lim_{s\to0}\left(\frac{\tau^s}{s}-\frac{1}{s}\right)=\ln\tau \tag{4.10}$$

we obtain

$$\mathscr{L}(s=0,x)=-\tfrac{1}{2}\int\limits_0^\infty (\tau^2 H)'''\ln\tau\,d\tau+\tfrac{3}{4}\alpha_2(x), \tag{4.11}$$

where $\alpha_2(x)=(\tau^2 H)''|_{\tau=0}$. We may make Eq. (4.11) more recognizable by breaking up the integral into an integral from 0 to μ plus an integral from μ to ∞ and integrating each integral by parts, choosing initial conditions in each case on integrals of derivatives of $\tau^2 H$ so as to get vanishing contribution from boundary terms, respectively, at 0 and ∞. The final result is,

$$\mathscr{L}(0,x)=\int\limits_0^\infty \left\{H-\frac{\alpha_0}{\tau^2}-\frac{\alpha_1}{\tau}-\tfrac{1}{2}\theta_\mu(\tau)\alpha_2\right\}\frac{d\tau}{\tau}+\tfrac{1}{2}\alpha_2\ln\mu, \tag{4.12}$$

where $\alpha_0=\tau^2 H|_{\tau=0}$, $\alpha_1=(\tau^2 H)'|_{\tau=0}$. The right-hand side of Eq. (4.12) is μ-independent. As noted above, this is unsatisfactory since one wants $\mathscr{L}(x)$ scale homogeneously when the metric is scaled. To accomodate this requirement, we modify the prescription by dropping the term $\tfrac{1}{2}\alpha_2\ln\mu$. With this term dropped, it is easily seen that the analytic continuation prescription for defining $-2\mathscr{L}(x)$, Eq.

(4.12) is identical to the heat kernel renormalization prescription when $m=0$ and differs only in a trivial manner – namely, by finite multiples of A_0, A_1, and A_2 – when $m \neq 0$. (Thus when $m \neq 0$ a correction must be made to the analytic continuation prescription in order to get vanishing effective Lagrangian in flat spacetime.)

If infra-red divergences do occur for large enough s, $Z(s, x, x)$ may not be defined for any value of s. However, we may break-up the integral for Z into an integral from 0 to μ plus an integral from μ to ∞. The former integral is analytic for $\mathrm{Re}\,s > 2$, while, as discussed above, the latter integral is almost certainly always finite for $s = 0$. Thus, we may leave the second integral unchanged and analytically continue the first integral as above. The final result is again Eq. (4.12). Thus, infra-red divergence for $s > 0$ pose no difficulties for this prescription.

The above analytic continuation prescription for defining $\mathscr{L}(x)$ can be interpreted as dimensional renormalization [20] as follows. If the self-adjoint operators B_1 and B_2 commute, then $e^{-\tau(B_1 + B_2)} = e^{-\tau B_1} e^{-\tau B_2}$. Consequently, if the Riemannian manifold (M, g) is a product of the manifolds (M_1, g_1) and (M_2, g_2), then the heat kernel for the Laplacian on (M, g) will be the product of the heat kernels on (M_1, g_1) and (M_2, g_2). Since the heat kernel for the Laplacian in k-dimensional flat space in the coincidence limit $x = x'$ is simply

$$H_E(\tau, x, x) = (4\pi\tau)^{-k/2} \tag{4.13}$$

we see that – except for the extra k-dependence of the factor 4π (which leads to entirely trivial consequences for the prescription) – adding on extra flat dimensions to the original manifold is equivalent to changing the value of s in the integral for $Z(s)$, Eq. (4.1). Thus we may view the procedure of considering values of s for which the integral for Z is well defined followed by analytic continuation to $s = 0$ as adding on extra flat dimensions followed by analytic continuation to $n = 4$. Unfortunately for this interpretation the number of flat dimensions that must be "added on" to make the integral well defined is negative ($k < -2$). It should be emphasized that this "dimensional renormalization" prescription is equivalent to the above analytic continuation prescription only when done precisely in the manner described here. If the dimension is varied in a different manner (e.g., non-flat dimensions are "added on") or if other quantities (such as the form of the equation) are allowed to vary with dimension, different results may be obtained.

A slight variant of dimensional regularization is the following zeta-function procedure [21]: Instead of considering $Z(s, x, x)$ we consider the ζ-kernel defined by

$$\zeta(s, x) = \frac{1}{\Gamma(s)} Z(s, x, x). \tag{4.14}$$

In the same manner as for Z, it follows that except for infra-red divergences if M is non-compact and $m = 0$, $\zeta(s, x)$ is analytic in s for $\mathrm{Re}\,s > 2$ and can be defined for all s by analytic continuation. Whereas $Z(s)$ has poles at all integer values of $s \leq 2$, $\zeta(s)$ will have poles only at $s = 1, 2$ since the poles in $\Gamma(s)$ cancel those of $Z(s)$ at the

other values of s. Using Eq. (4.9) for $Z(s)$, we calculate the derivative of $\zeta(s, x)$ at $s=0$ to be,

$$-\zeta'(s=0, x) \leqq -\tfrac{1}{2} \int_0^\infty (\tau^2 H)''' \ln \tau \, d\tau$$
$$+ \tfrac{3}{4}\alpha_2 - \tfrac{1}{2}\gamma\alpha_2 , \tag{4.15}$$

where γ is Euler's constant, Eq. (2.16). Comparing this with Eq. (4.11), we see that if we define the zeta-function renormalization [21] prescription for $2\mathscr{L}(x)$ to be $-\zeta'(0, x)$ plus $\tfrac{1}{2}\alpha_2 \ln \mu$ (to get the proper scaling behavior of $\mathscr{L}$), this prescription differs from the analytic continuation prescription only by the entirely trivial term $\tfrac{1}{2}\gamma\alpha_2$.

The intuitive motivation for the zeta-function prescription is as follows: If M is compact, we may define the ζ-function by

$$\zeta(s) = \int_M \zeta(s, x). \tag{4.16}$$

For $\operatorname{Re} s > 2$ we have,

$$\zeta(s) = \operatorname{tr} A^{-s} = \sum_j \lambda_j^{-s} \tag{4.17}$$

where λ_j are the eigenvalues of A. For $\operatorname{Re} s \leqq 2$ the sum does not in general converge but formally we have

$$\zeta'(0) = -\sum_j \ln \lambda_j = -\ln \det A . \tag{4.18}$$

Thus, formally $-\tfrac{1}{2}\zeta'(0)$ is the effective action, Eq. (2.11). It is gratifying therefore that the zeta-function renormalization yields essentially the same prescription for $\mathscr{L}(x)$ as dimensional renormalization and heat kernel renormalization.

The heat kernel $H(\tau, x, x')$ becomes singular when both $\tau=0$ and $x=x'$. In the prescriptions given above, we set $x=x'$ and modified the τ behavior of H to yield finite results for $\mathscr{L}(x)$. An alternative approach is initially to keep $x \neq x'$. The τ-integral is then finite but divergences occur when the coincidence limit $x \to x'$ is taken. However, the "divergent part" of this expression can be explicitly calculated from the asymptotic expansion of H near $\tau=0$ and we may obtain a prescription for $\mathscr{L}(x)$ by subtraction of this divergent part from $L(x, x')$ before taking the coincidence limit $x \to x'$. This defines the point-splitting renormalization prescription [23], which may be stated more explicitly in this case as follows.

For $x \neq x'$, we have,

$$L(x, x') = -\int_0^\infty H(\tau, x, x') \frac{d\tau}{\tau} . \tag{4.19}$$

By inspection, the divergences in $L(x, x')$ when $x \to x'$ can be attributed to the first three terms in the asymptotic expansion F of H. Using the definite integral [22]

$$\int_0^\infty x^{\nu-1} e^{-\beta/x - \gamma x} dx = 2\left(\frac{\beta}{\gamma}\right)^{\nu/2} K_\nu(2\sqrt{\beta\gamma}), \tag{4.20}$$

where K_ν denotes the modified Hankel function [22], we find that the first three terms of F contribute to $L(x, x')$ the terms,

$$- \frac{1}{(4\pi)^2} \left\{ A_0(x, x') \frac{8m^2}{\sigma} K_{-2}(\sqrt{m^2\sigma}) + A_1(x, x') 4 \left(\frac{m^2}{\sigma} \right)^{1/2} K_{-1}(\sqrt{m^2\sigma}) \right.$$

$$\left. + A_2(x, x') \int_0^\mu e^{-m^2\tau} e^{-\sigma/4\tau} \frac{d\tau}{\tau} \right\}, \qquad (4.21)$$

where we have cut off the last integral at $\tau = \mu$ to prevent infra-red divergences when $m \to 0$. The point-splitting renormalization prescription for defining $\mathscr{L}(x)$ consists of subtracting the above expression from $L(x, x')$ and then taking the coincidence limit $x \to x'$. Defined in this manner it is clear, that point-splitting is identical to heat kernel renormalization, which, as shown above, differs from dimensional renormalization and zeta-function renormalization by at most finite multiples of the local curvature terms A_0, A_1, and A_2.

As discussed in Sect. II above, the stress-energy tensor is obtained by functional differentiation of $\mathscr{L}(x)$ with respect to $g^{\mu\nu}$. Most of the point-splitting prescriptions have been given directly for renormalization of the stress-energy tensor $T_{\mu\nu}$ rather than for the action [23, 6]. A formal calculation indicates that the above point-splitting prescription for $\mathscr{L}(x)$ followed by functional differentiation with respect to $g^{\mu\nu}$ is essentially equivalent to the direct prescriptions for $T_{\mu\nu}$. [Note that the functional derivative of the subtraction term for $L(x, x')$ yields both the subtraction terms needed to make $T_{\mu\nu}$ finite plus the "correction term" mentioned above arising from the functional derivative of μ and required for a conserved stress-tensor.] Thus, the point-splitting prescription for $\mathscr{L}(x)$ given above should be equivalent (up to finite multiples of functional derivatives of $A_0(x, x)$, $A_1(x, x)$, and $A_2(x, x)$ with respect to $g^{\mu\nu}$) to the direct stress-energy point-splitting prescription. However, a proof of this statement would require a careful treatment of functional derivatives of operators, which we shall not attempt to do here.

Up to this point, we have focused on the question of whether the various prescriptions for defining $\mathscr{L}(x)$ are mathematically well defined, not on the question of whether they are physically correct. This latter question is difficult to analyze, since the prescriptions here are defined for Riemannian manifolds while the physics takes place on Lorentzian spacetimes. However, the point-splitting prescription for $T_{\mu\nu}$ is also mathematically well defined on a wide class of Lorentzian spacetimes [4]. There it yields a renormalized $T_{\mu\nu}$ which is consistent with causality, conservation, and other requirements [5, 6]. Since these requirements uniquely determine $T_{\mu\nu}$ up to local curvature terms [5, 6], we have strong grounds for believing that the point-splitting prescription on spacetimes is physically correct (up to the local curvature terms). If the field theory on spacetime is defined by analytic continuation from a Euclidean section, the point-splitting prescription for $T_{\mu\nu}$ defined directly on the spacetime should be equivalent to analytic continuation of the point-splitting prescription for $T_{\mu\nu}$ defined on the Euclidean section, which should be equivalent to functional differentiation of the

point-splitting prescription $\mathscr{L}(x)$, which, in turn, is equivalent (up to local curvature terms) to heat kernel, dimensional, and zeta-function renormalization. Thus, we have strong grounds for believing that all the prescriptions considered above are physically correct (up to addition of local curvature terms). In particular, this argument strongly suggests that if heat kernel, dimensional, or zeta-function renormalization is performed on a Euclidean section and continued back to spacetime, it will satisfy the causality axiom of [5]. This is rather remarkable in view of the fact that the notion of causality cannot be meaningfully expressed in the Euclidean section.

V. Static Spacetimes

In the previous sections we have discussed the definition of the "Feynman propagator" and the renormalization of the effective action in a "Euclidean section" of curved spacetime, i.e., on a Riemannian manifold (M, g_{ab}). We have avoided the question of how one finds this Euclidean section and how one continues the results back to the Lorentz spacetime. As remarked in Sect. I, these are very difficult issues. However, in the case of a globally static spacetime, there is a natural prescription for obtaining a Euclidean section. If we label the hypersurfaces orthogonal to the timelike Killing field by the Killing parameter t, then the "Wick rotation" $t \rightarrow -it$ yields a Riemannian metric. Thus, in the static case we may attempt to define the Feynman propagator by Wick rotating $t \rightarrow -it$, obtaining A^{-1} as in Sect. II, and then Wick rotating back $t \rightarrow it$. On the other hand, in static spacetimes one has a natural motion of "positive frequency" and can define the Feynman propagator directly on the spacetime. It is of interest, therefore, to see if these two definitions agree. In this section we shall show that they do. We begin by deriving an expression for the Feynman propagator obtained directly in the Lorentz spacetime from the natural notion of positive frequency.

The wave operator in a static curved spacetime can be written as,

$$-V_a V^a + m^2 = V^{-2} \frac{\partial^2}{\partial t^2} + B, \tag{5.1}$$

where $V^2 = -g_{00} > 0$ is the squared norm of the timelike Killing field and B is a positive elliptic operator which is independent of t and thus may be viewed as acting on $L^2(M_3)$ where M_3 denotes the 3-manifold orthogonal to the Killing field. We define the operator C by

$$C = V^2 B. \tag{5.2}$$

If we let $\tilde{L}^2(M_3)$ be the Hilbert space of square-integrable functions on M_3 with the volume element $\sqrt{g_3}/V$ (where $\sqrt{g_3}$ denotes the natural volume element obtained from the induced metric on M_3), then $C : \tilde{L}^2(M_3) \rightarrow \tilde{L}^2(M_3)$ is a positive symmetric operator on $C_0^\infty(M_3)$, and hence can be extended [13] to a positive self-adjoint operator, which we shall also denote by C.

We define the one-particle Hilbert space $\mathscr{H}$ of positive frequency solutions to the wave equation as follows: Let $f \in \operatorname{dom} C^{1/2} \subset \tilde{L}^2(M_3)$. Then

$$F(t, x) = e^{-iC^{1/2}t} f \tag{5.3}$$

will be a solution of the wave equation, which we shall call "positive frequency". [Note, incidentally, that if C and its powers are essentially self-adjoint on $C_0^\infty(M_3)$ and $f \in C_0^\infty$, then the arguments of Sect. II show that $F(t, x)$ is C^∞ in t and x]. The Klein-Gordon inner product of F and $G = e^{-iC^{1/2}t} g$ with $g \in \operatorname{dom} C^{1/2}$ is

$$
\begin{aligned}
(F, G)_{KG} &= i \int_{M_3} (F^* \nabla_a G - G \nabla_a F^*) n^a \sqrt{g} \, d^3x \\
&= i \int_{M_3} \left(F^* \frac{\partial G}{\partial t} - G \frac{\partial F^*}{\partial t} \right) \frac{\sqrt{g}}{V} \, d^3x \\
&= \int_{M_3} (f^* C^{1/2} g + (C^{1/2} g) f^*) \frac{\sqrt{g}}{V} \, d^3x \\
&= 2(f, C^{1/2} g)_{\tilde{L}^2(M_3)} .
\end{aligned}
\tag{5.4}
$$

Thus, the Klein-Gordon norm (F, F) is positive definite. We define the Hilbert space $\mathscr{H}$ of all positive frequency solutions to be the Cauchy completion of this space in the Klein-Gordon norm. The Hilbert space of all states of the quantum field is taken to be the Fock space constructed from $\mathscr{H}$.

The quantum field operator ϕ may be defined by

$$\phi(t, x) = \sum_i [F_i(t, x) a_i + F_i^*(t, x) a_i^\dagger], \tag{5.5}$$

where $\{F_i\}$ is an orthonormal basis of positive frequency solutions, a_i and $a_i^\dagger$ denote the annihilation and creation operators associated with F_i, and both sides of the equation are to be interpreted as distributions. In the following we will ignore technical issues relating to the convergence of the sums. [In the unlikely event that any convergence problems arise in the derivation, we could simply take Eq. (5.8) below to be the definition of the Feynman propagator.] In terms of the field operator ϕ, the Feynman propagator is,

$$
\begin{aligned}
\Delta_F(t, x; t', x') &= \langle 0 | T(\phi(t', x') \phi(t, x)) | 0 \rangle \\
&= \begin{cases}
\displaystyle\sum_i F_i(t', x') F_i^*(t, x) & t' \geq t \\
\displaystyle\sum_i F_i^*(t', x') F_i(t, x) & t' \leq t .
\end{cases}
\end{aligned}
\tag{5.6}
$$

For each value of t, we may view each test function $\psi(t, x)$ on which Δ_F acts as a vector $\psi_t \in \tilde{L}^2(M_3)$. We may also take the orthonormal basis $\{F_i\}$ to be of the form $F_i = e^{-iC^{1/2}t} f_i$ with $f_i \in \operatorname{dom} C^{1/4}$ [so that $\{\sqrt{2} C^{1/4} f_i\}$ is an orthonormal basis of

$\tilde{L}^2(M_3)$ according to Eq. (5.4)]. Assuming that $V^2\psi_t \in \text{dom}\, C^{-1/2}$, we find

$$\Delta_F(t', x'; \psi) = \sum_i{}' F_i(t', x') \int_{t \leqq t'} F_i^*(t, x)\psi(t, x)V\sqrt{g_3}\,d^3x\,dt$$

$$+ t \geqq t' \text{ term}$$

$$= \sum_i e^{-iC^{1/2}t'}f_i(x') \int_{-\infty}^{t'} dt(e^{-iC^{1/2}t}f_i, V^2\psi_t)_{\tilde{L}^2(M_3)}$$

$$+ t \geqq t' \text{ term}$$

$$= \tfrac{1}{2}C^{-1/4}e^{-iC^{1/2}t'} \int_{-\infty}^{t'} dt\left\{\sum_i [\sqrt{2}C^{1/4}f_i] \cdot (\sqrt{2}C^{1/4}f_i, e^{iC^{1/2}t}C^{-1/4}V^2\psi_t)\right\}$$

$$+ t \geqq t' \text{ term}$$

$$= \tfrac{1}{2} \int_{-\infty}^{t'} dt\, e^{-iC^{1/2}(t'-t)}C^{-1/2}(V^2\psi_t)$$

$$+ t \geqq t' \text{ term}. \tag{5.7}$$

Thus, if for each t and t' we view Δ_F as a map from $\tilde{L}^2(M_3)$ into $\tilde{L}^2(M_3)$, we obtain

$$\Delta_F = \begin{cases} \tfrac{1}{2}e^{-i(t'-t)C^{1/2}}C^{-1/2}V^2 & t' \geqq t \\ \tfrac{1}{2}e^{+i(t'-t)C^{1/2}}C^{-1/2}V^2 & t' \leqq t, \end{cases} \tag{5.8}$$

where here V^2 denotes the operation of multiplication by the function V^2. Our aim now is to compare this expression with the Euclidean expression for the Feynman propagator.

Making the "Wick rotation" $t \to -it$, we convert the wave equation into the Laplace operator

$$A = -V^{-2}\frac{\partial^2}{\partial t^2} + B. \tag{5.9}$$

If we define the operator $\tilde{A}$ by

$$\tilde{A} = V^2 A = -\frac{\partial^2}{\partial t^2} + C \tag{5.10}$$

then the Feynman propagator A^{-1} is given by

$$A^{-1} = \tilde{A}^{-1}V^2. \tag{5.11}$$

Since the operator A is self-adjoint with the natural volume element $V\sqrt{g_3}$ on the 4-dimensional manifold M (where $\sqrt{g_3}$ is the natural volume element on M_3), $\tilde{A}$ will be self-adjoint with the volume element $\sqrt{g_3}/V$ on M.

We can express $\tilde{A}^{-1}$ in terms of its heat kernel by,

$$\tilde{A}^{-1} = \int_0^\infty e^{-\tau\tilde{A}}\,d\tau. \tag{5.12}$$

The reason for introducing $\tilde{A}$ is that it is the sum of the two commuting self-adjoint operators $-\partial^2/\partial t^2$ and C. Hence, its heat kernel is the product of the heat kernels of the individual operators. But the heat kernel of $-\partial^2/\partial t^2$ is well known to be

$$e^{-\tau(-\partial^2/\partial t^2)} = \frac{e^{-(t-t')^2/4\tau}}{(4\pi\tau)^{1/2}}.\tag{5.13}$$

Thus, we have,

$$\tilde{A}^{-1} = \frac{1}{\sqrt{4\pi}} \int_0^\infty e^{-(t-t')^2/4\tau} e^{-\tau C} \tau^{-1/2}\,d\tau.\tag{5.14}$$

Using the fact that for α and β positive real numbers,

$$\int_0^\infty e^{-\alpha/4\tau} e^{-\beta\tau} \tau^{-1/2}\,d\tau = \sqrt{\frac{\pi}{\beta}}\, e^{-\sqrt{\alpha\beta}}\tag{5.15}$$

(see § 3.472 of [22]) we obtain,

$$A^{-1} = \tilde{A}^{-1} V^2 = \begin{cases} \frac{1}{2} e^{-(t'-t)C^{1/2}} C^{-1/2} V^2 & t' \geqq t \\ \frac{1}{2} e^{+(t'-t)C^{1/2}} C^{-1/2} V^2 & t' \leqq t. \end{cases}\tag{5.16}$$

Hence, if we Wick rotate Eq. (5.16) back to the spacetime by setting $t \to it$, we obtain precisely, Δ_F, Eq. (5.8). Thus, in the static case, the Euclidean procedure for defining the Feynman propagator agrees with the direct spacetime construction of it.

It is worth remarking that, as already noted in Sect. IV above, except for the possibility of infrared divergences when $m = 0$, the Euclidean Feynman propagator always is given by a kernel with singularity structure of the Hadamard form as $x \to x'$. This strongly suggests that in static spacetimes (at least when $m \neq 0$) the two-point function $\langle \phi(x)\phi(x') + \phi(x')\phi(x) \rangle$ – which is equal to the Feynman propagator when x and x' are spacelike separated – will be a Hadamard distribution. If this holds, then the result of Fulling et al. [4] proves that in any initially static spacetime, the two point function will be a Hadamard distribution everywhere, and thus the point-splitting prescription for renormalizing $T_{\mu\nu}$ (defined directly on the spacetime) will be well defined. Furthermore, the proof of [1] could be directly taken over to prove existence of the S-matrix in a spacetime which is initially and finally static and has compact spacial sections (i.e., a closed universe).

References

1. Wald, R.M.: Ann. Phys. **118**, 490–510 (1979)
2. Candelas, P., Raine, D.J.: Phys. Rev. D **15**, 1494–1500 (1977)
3. Rumpf, H.: Phys. Lett. **61** B, 272–274 (1976)
4. Fulling, S.A., Sweeny, M., Wald, R.M.: Commun. Math. Phys. **63**, 257–264 (1978)
5. Wald, R.M.: Commun. Math. Phys. **54**, 1–19 (1977)
6. Wald, R.M.: Phys. Rev. D **17**, 1477–1484 (1978)
7. Hartle, J.B., Hawking, S.W.: Phys. Rev. D **13**, 2188–2203 (1976)

8. Hawking, S.W.: In: General relativity: an Einstein centennary survey (eds. S.W. Hawking, W. Israel). Cambridge, UK: Cambridge University Press 1979
9. Minakshisundaram, S., Pleijel, A.: Canad. J. Math. **1**, 242–256 (1949)
10. Patodi, V.K.: J. Diff. Geo. **5**, 233–249 (1971)
11. Riesz, F., Sz.-Nagy, B.: Functional analysis. New York: Ungar 1955
12. Reed, M., Simon, B.: Functional analysis. New York: Academic Press 1972
13. Reed, M., Simon, B.: Fourier analysis. New York: Academic Press 1975
14. Gilkey, P.: The index theorem and the heat equation. Boston: Publish or Perish 1974
15. Gaffney, M.P.: Proc. Nat. Acad. Sci. **37**, 48–50 (1951); Ann. Math. **60**, 140–145 (1954)
16. Cordes, H.O.: Math. Ann. **195**, 257–272 (1972)
17. Schwinger, J.: Phys. Rev. **82**, 664–679 (1951)
18. DeWitt, B.: Phys. Rep. **19**C, 295–357 (1975)
19. Wightman, A.S.: In: High energy electromagnetic interactions and field theory (ed. M. Levy). New York: Gordon and Breach 1967
20. Brown, L.S.: Phys. Rev. D **15**, 1469–1483 (1977)
21. Hawking, S.W.: Commun. Math. Phys. **55**, 133–148 (1977)
22. Gradshteyn, I.S., Ryzhik, I.W.: Table of integrals series and products. New York: Academic Press 1965
23. Christensen, S.M.: Phys. Rev. D **14**, 2490–2501 (1976); **17**, 946–963 (1978)

Note Added in Proof

The essential self-adjointness restriction on A and its powers can be removed and the theorem can be generalized to prove the validity of the local heat kernel expansion of the operator $e^{-\tau A_E}$ where A_E is any self-adjoint extension of A which is bounded from below. To do so, we construct F as before, or better yet (to avoid the technical inconvenience of working with a finite N) pick a small neighborhood $\mathcal{N}$ of x', construct a compact manifold (for which the above result already applies) with a neighborhood isometric to $\mathcal{N}$, and let $\tilde{F}$ denote the heat kernel of this manifold multiplied by a smooth cutoff function $\chi(x, x')$ which is 1 near x' but vanishes outside $\mathcal{N}$. The failure, $\tilde{p}(\tau, x, x')$, of $\tilde{F}$ to satisfy the heat equation will be a smooth function which vanishes as $\tau \to 0$ faster than any power of τ and thus can be extended to a C^∞ function for $\tau \leq 0$ by setting $\tilde{p} = 0$ there. If we again define the correction term $\tilde{Q}$ by $\tilde{Q} = \int_0^\tau e^{-(\tau - \tau)A_E}\tilde{p}d\tau'$ for $\tau \geq 0$ and now set $\tilde{Q} = 0$ for $\tau < 0$, then we can no longer use our estimates to prove directly that $\tilde{Q}$ is smooth, but viewing $\tilde{Q}$ as an L^2-vector for each τ and x', it will satisfy $d\tilde{Q}/d\tau + A_E\tilde{Q} = \tilde{p}$, where the τ-derivative is understood in the (strong) Hilbert space sense. Consequently, if we view $\tilde{Q}$ as a distribution acting on test functions of (τ, x) in $\mathbb{R} \times \mathcal{N}$, we see that $\tilde{Q}$ is a weak solution of the heat equation with smooth source $\tilde{p}$. Since the heat equation is hypoelliptic (see L. Hormander, Commun. Pure Appl. Math. **11**, 197–218 (1958)) this implies that, in fact, $\tilde{Q}$ is given by a C^∞ function. Since $\tilde{Q} = 0$ for $\tau < 0$, this function vanishes faster than any power of τ as $\tau \to 0$ from above. Since the argument that $e^{-\tau A_E} = \tilde{F} - \tilde{Q}$ goes through as before and the above theorem holds for $\tilde{F}$, this establishes the result for $e^{-\tau A_E}$. (The bounded from below property of A_E was used here only to ensure that the heat operator is bounded, thereby avoiding possible domain problems.) I wish to thank M. Atiyal for helpful discussions on this argument.