Massive quantum field theory in two-dimensional Robertson-Walker space-time

T. S. Bunch
Department of Physics, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201

S. M. Christensen
Lyman Laboratory of Physics and Center for Astrophysics, Harvard University, Cambridge, Massachusetts 02138

S. A. Fulling
Department of Mathematics, Texas A&M University, College Station, Texas 77843
(Received 4 May 1978)

The stress tensor of a massive scalar field, as an integral over normal modes (which are not mere plane waves), is regularized by covariant point separation. When the expectation value in a Parker-Fulling adiabatic vacuum state is expanded in the limit of small curvature-to-mass ratios, the series coincides in each order with the Schwinger-DeWitt-Christensen proper-time expansion. The renormalization ansatz suggested by these expansions (which applies to arbitrary curvature-to-mass ratios and arbitrary quantum state) can be implemented at the integrand level for practical computations. The renormalized tensor (1) passes in the massless limit, for appropriate choice of state, to the known vacuum stress of a massless field, (2) agrees with the explicit results of Bernard and Duncan for a special model, (3) has a nonzero vacuum expectation value in the two-dimensional "Milne universe" (flat space in hyperbolic coordinates). Following Wald, we prove that the renormalized tensor is conserved and point out that there is no arbitrariness in the renormalization procedure. The general approach of this paper is applicable to four-dimensional models.

I. INTRODUCTION

Let $\phi(x) \equiv \phi(t, y)$ be a Hermitian scalar field satisfying canonical commutation relations and the wave equation

$$\Box\phi \equiv g^{\mu\nu}\phi_{;\mu\nu} = -m^2\phi \qquad (1.1)$$

in a two-dimensional space-time, and possessing the classical stress tensor

$$T_{\mu\nu} = \phi_{;\mu}\phi_{;\nu} - \tfrac{1}{2}\phi^{;\rho}\phi_{;\rho}g_{\mu\nu} + \tfrac{1}{2}m^2\phi^2 g_{\mu\nu} \,. \qquad (1.2)$$

We consider the Robertson-Walker metrics

$$ds^2 \equiv g_{\mu\nu}dx^\mu dx^\nu = C(\eta)(d\eta^2 - dy^2)\,, \qquad (1.3a)$$

where $-\infty < y < \infty$. Equation (1.3a) is related to the more familiar form

$$ds^2 = dt^2 - a(t)^2 dy^2 \qquad (1.3b)$$

by

$$\eta = \int^t a(t)^{-1}dt, \quad C(\eta) = a(t)^2 \,.$$

Our aim is to construct, as explicitly as possible, the physical (renormalized) stress-energy-momentum tensor operator, $T_{\mu\nu}(x)$, for this quantum field theory. The stress tensor must have finite matrix elements with respect to a suitable class of physical (nonpathological) states, and therefore cannot literally have the form (1.2). Nevertheless, the operator should in some way be deducible from that expression. It seems generally agreed that these goals can be achieved in two steps. The first is *regularization*: The expression (1.2) is modified ("cut off"), in a generally covariant manner, so that it is finite and unambiguous even when the ϕ's are the quantum operators. It is essential to the success of the program that the regularized object can be divided into two parts. The first part (a) has a finite, unambiguous limit as the cutoff is removed. The second part (b) is a c-number (state-independent quantity) built entirely from a small, sharply defined class of local, covariant functionals of the external gravitational field. The second step is *renormalization*: The divergent (more precisely, discontinuous) term (b), which depends on the regularization method, must be discarded to obtain the physically meaningful stress tensor. This step leaves undetermined the part of the tensor proportional to functionals of the types which appear in (b) because the division into (a) and (b) is not unique. However, only finitely many terms of these types are physically acceptable, and their arbitrary coefficients can be identified as renormalizations of coupling constants in the equation of motion of the gravitational field, to which $T_{\mu\nu}$ will ultimately be coupled in some semiclassical or fully quantum theory.

The regularized stress tensor will be finite and admit the decomposition (a) + (b) only if one is working within the class of physically acceptable quantum states. In practice it is sometimes necessary to tackle the problem of determining these states simultaneously with the calculation of the stress tensor.

Various methods of regularization and renormalization (some of which deal with the Lagrangian instead of $T_{\mu\nu}$) differ in calculational simplicity, in the apparent naturalness of the renormalization ansatz, and in their ability to produce explicit, concrete answers for wide classes of space-time geometries. Regularization by separation of the space-time points in the operator products in Eq. (1.2) has the third of these desirable properties as its strong point. For massless fields in two-dimensional space-time a complete solution of the problem was obtained, starting from the expansion of the field operator in normal modes.[1-3] The regularization was carried out by the same methods for a conformally invariant massless scalar field in a spatially Euclidean four-dimensional Robertson-Walker universe.[4] Explicit calculations in these models were facilitated by "conformal triviality": The classical normal-mode solutions are plane waves in a suitable coordinate system. On the other hand, the divergent and direction-dependent terms which must be removed in the renormalization of $T_{\mu\nu}$ were determined for a scalar field in an arbitrary four-dimensional geometry from the proper-time expansion of the two-point function.[5] At this stage the research reported here was begun in order to demonstrate that regularization by point separation can be used to calculate the renormalized $T_{\mu\nu}$ in theories which are not conformally trivial. We obtain a formula for the renormalized $\langle T_{\mu\nu}\rangle$ in any state of the theory (1.1)–(1.3) as a convergent integral involving the mode functions, suitable for numerical computation or for further analytical work based on suitable approximations to the mode functions. From this work it is clear how, in principle at least, such calculations can be done for any model in which the field equation can be solved by separation of variables.

Because the time dependence of the modes in this system is not of the plane-wave form, the problem of choosing physically acceptable states arises at the outset. Our (fulfilled) expectation was that "good" states would have the ultraviolet behavior prescribed in Ref. 6. Since the mass provides a built-in infrared cutoff, it is possible to impose the adiabatic condition of Ref. 6 on *all* modes, obtaining a narrower class of "strictly adiabatic vacuum states." For such a state we verify that $\langle T_{\mu\nu}\rangle$ coincides with the two-dimensional version of the proper-time expansion of Ref. 5 up through the first order (in ratios of geometrical quantities to powers of the mass) in which divergences and discontinuities do not occur; it is clear that the quantities coincide to all finite orders. In particular, the strictly adiabatic vacuum stress is a purely local geometrical quantity—a polynomial in covariant derivatives of the curvature tensor—to any finite order in the curvature. For more general states, built from mode functions satisfying the adiabatic condition only in the ultraviolet limit, the divergent and discontinuously direction-dependent terms in $\langle T_{\mu\nu}\rangle$ will still be those given by the proper-time analysis, and the proper renormalization is clear.

It is possible to pass to the massless limit in the expectation values for suitably defined states (*not* in the expansion for a strictly adiabatic vacuum, which refers to the opposite limit). The theory goes over continuously to the well-established massless one (Refs. 1, 2, 7, 8, and 15). In fact, extrapolating to the four-dimensional situation, we believe (see remarks in Refs. 4, 9, 10, 11, and 12) that the clarity of the renormalization process for a massive field in a strictly adiabatic vacuum state casts considerable light on the renormalization of a massless field, where the separation of the leading divergences from the physical remainder seems somewhat ambiguous at first glance. In particular, the proper-time analysis correctly gives the divergences and discontinuities in $\langle T_{\mu\nu}\rangle$ and similar quantities even when the mass is zero. Together with the profound work of Wald,[13-16] these considerations point to a prescription for constructing the physical $T_{\mu\nu}$ which, in our opinion, suffers from no arbitrariness.

The plan of the paper is as follows. In Sec. II and Appendix A the proper-time analysis of Ref. 5 is redone for a two-dimensional space-time. Section III discusses the construction of the space of "physical" state vectors, following Ref. 6. The principal calculation, summarized from Ref. 17, is presented in Sec. IV and Appendices B and C. This is the determination of the regularized stress tensor's expectation value in a state which is "vacuum" in the sense of Sec. III. The final, renormalized stress tensor is attained in Sec. V. Section VI deals with the massless field in the context of the present work. The relation between massive and massless theories is explored further in Sec. VII, where the renormalized $\langle T_{\mu\nu}\rangle$ is examined in two special cases, flat space with a "vacuum" defined by naive quantization relative to a curvilinear coordinate system, and the model recently studied by Bernard and Duncan,[18] whose results are found to be in agreement with ours. In Sec. VIII we examine the renormalization ansatz critically. The analysis of Wald[14] is adapted to prove that our stress tensor is conserved and causal, and we point out that various formulations (those of Refs. 9, 11, 14, and the present paper) of the renormalization procedure are equivalent. In that section we also summarize the assumptions which go into our renormalization procedure and discuss their implications for the problem of uniqueness of the result.

Our notational conventions are those of Ref. 2, except that

(1) D means C'/C rather than $\frac{1}{2}C'/C$ (the prime indicates differentiation with respect to η),

(2) in Sec. II and Appendix A, for compatibility with Ref. 5 and the other literature on proper-time expansions, we use the metric with $g_{00}<0$. Equations which must be interpreted in this convention[19] are marked with as asterisk. In the rest of the paper the convention $g_{00}>0$ is followed. In both cases the sign of the two-dimensional curvature scalar is such that

$$R=D'/C \tag{1.4}$$

in our Robertson-Walker metrics and R is positive in the de Sitter and reduced Schwarzschild[1,34] models. A key to previous publications on stress tensors in two dimensions may be helpful:

1. $g_{00}>0$, $R>0$: Refs. 2, 7, 11, 17, 20,
2. $g_{00}<0$, $R>0$: Refs. 9, 19,
3. $g_{00}>0$, $R<0$: Refs. 1, 3, 8,
4. $g_{00}<0$, $R<0$: Ref. 18.

In conventions 1 and 4 the trace anomaly of a massless scalar field is

$$T^{\alpha}_{\alpha}=-(24\pi)^{-1}R \tag{1.5}$$

[clerical errors in Eq. (4.5) of Ref. 20 and Eq. (16) of Ref. 7 notwithstanding], while in conventions 2 and 3 this sign is positive.

II. PROPER-TIME EXPANSIONS

The central object in our calculations is the symmetrized two-point function

$$\langle\phi^2\rangle(x,x')=\tfrac{1}{2}\langle\phi(x)\phi(x')+\phi(x')\phi(x)\rangle\,, \tag{2.1}$$

where the ϕ's are quantized field operators and the expectation value is with respect to a "physically acceptable" quantum state. The meaning of "physically acceptable" will be clarified later; intuitively, the point is that the physical situation considered must be one in which the density of matter is finite. Then $\langle\phi^2\rangle(x,x')$ and the "separated" stress tensor formed from it [Eq. (2.5)] will be well defined when the geodesic separation of the points is not null. Furthermore, any terms which diverge as the points approach each other, or which depend in that limit on the direction of separation, will be c numbers depending only on the local geometry. In fact, such terms in $2\langle\phi^2\rangle$ will be identical to the terms of the same type in

$$G^{(1)}(x,x')=\frac{\langle\text{out, vac}|[\phi(x)\phi(x')+\phi(x')\phi(x)]|\text{in, vac}\rangle}{\langle\text{out, vac}|\text{in, vac}\rangle}\,, \tag{2.2}$$

where $|\text{in, vac}\rangle$ and $|\text{out, vac}\rangle$ are the initial and final vacuum states. If necessary, the geometry of the model considered can be modified in the distant past and future to make those states well defined, without changing the local terms in question. (See Ref. 21. An alternative characterization of $G^{(1)}$, $|\text{in, vac}\rangle$, etc. has been developed by Rumpf.[22-25]) $G^{(1)}$ in turn is a linear combination of the retarded, advanced, and causal (Feynman) Green's functions.

When $m^2>0$, an asymptotic series for $G^{(1)}$ can be found by methods due to Schwinger, Hadamard, Synge, and DeWitt. The four-dimensional case is treated in Ref. 5. The two-dimensional analog of that work is relatively simple (see Appendix A) and results in

$$\begin{aligned}\tfrac{1}{2}G^{(1)}(x,x')=(2\pi)^{-1}\{&-L[1+\tfrac{1}{2}m^2\sigma+O(\sigma^2)]+[\tfrac{1}{12}m^{-2}R+\tfrac{1}{120}m^{-4}(R^2+2\Box R)+O(m^{-6})]-\tfrac{1}{24}m^{-2}R_{;\alpha}\sigma^{\alpha}+O(m^{-4}\sigma^{1/2})\\&+\sigma[\tfrac{1}{2}m^2-\tfrac{1}{24}R+\tfrac{1}{240}m^{-2}(R^2-\Box R)]+\tfrac{1}{80}m^{-2}R_{;\alpha\beta}\sigma^{\alpha}\sigma^{\beta}+O(m^{-4}\sigma)+O(\sigma^{3/2})\}\,.\end{aligned} \tag{2.3*}$$

Here σ is half the square of the geodesic distance from x to x', and

$$L=\gamma+\tfrac{1}{2}\ln|\tfrac{1}{2}m^2\sigma| \tag{2.4}$$

(γ = Euler's constant). Also, $-\sigma^{\alpha}\equiv-g^{\alpha\rho}\sigma_{;\rho}$ is tangent to the geodesic from x to x' and has magnitude equal to the geodesic distance; thus $\sigma_{\alpha}\sigma^{\alpha}=2\sigma$, and σ^{α} is of the order $O(\sigma^{1/2})$. The methods of Ref. 5 also produce from Eq. (1.2) a two-dimensional point-separated stress tensor,

$$T^{(1)}_{\mu\nu}(x,x') = (2\pi)^{-1}\Big\{(\sigma_\beta\sigma^\beta)^{-1}\Big(g_{\mu\nu} - 2\,\frac{\sigma_\mu\sigma_\nu}{\sigma_\alpha\sigma^\alpha}\Big) + \tfrac{1}{2}m^2 L g_{\mu\nu} + O(\sigma L) + \Big[-\tfrac{1}{4}m^2\Big(g_{\mu\nu} - 2\,\frac{\sigma_\mu\sigma_\nu}{\sigma_\alpha\sigma^\alpha}\Big) - \tfrac{1}{12}R\,\frac{\sigma_\mu\sigma_\nu}{\sigma_\alpha\sigma^\alpha}\Big] + \tfrac{1}{60}m^{-2}(R_{;\mu\nu} - \Box R g_{\mu\nu} - \tfrac{1}{4}R^2 g_{\mu\nu}) + O(m^{-4}) + O(\sigma^{1/2})\Big\}. \qquad (2.5*)$$

For its definition see Eqs. (2.8)–(2.10) and Refs. 2, 4, and 5. The superscript "(1)" is a reminder that this tensor is formed from $\frac{1}{2}G^{(1)}$, not from $\langle\phi^2\rangle$, and thus is a normalized matrix element between in and out vacuum states, not an expectation value in a single state.

We note some properties of the expansions (2.3) and (2.5). First, for dimensional reasons the factors in every term satisfy

$$d_m - d_\epsilon + d_\partial = \begin{cases} 0 \text{ in } G^{(1)} \\ 2 \text{ in } T^{(1)}_{\mu\nu} \end{cases} \qquad (2.6)$$

(in two dimensions), where d_m is the order of the term in m, d_ϵ is the order in $\sigma^{1/2}$, and d_∂ is the number of differentiations acting on the metric tensor [e.g., $d_\partial(R) = 2$, $d_\partial(\Box R) = d_\partial(R^2) = 4$]. Thus the remainder terms of higher order in m^{-1} or σ are also of high order in the curvature.

Second, these remainder terms are smooth functions of x and x' (even in the coincidence limit). The same statement is true of the explicitly exhibited coefficients of negative powers of m, and of the coefficient of the logarithm. The logarithmic coefficient itself does not contain negative powers of m. It follows that the terms can be divided (not quite uniquely) into two classes: those which are well behaved in the limit $m \to 0$ and those which are well behaved in the limit of null separation ($\sigma \to 0$), which includes the coincidence limit ($x' \to x$). This requires breaking up the logarithmic factor as

$$L = L_\epsilon + L_m\,,$$
$$L_\epsilon = \gamma + \tfrac{1}{2}\ln|\tfrac{1}{2}\mu^2\sigma|\,, \qquad (2.7)$$
$$L_m = \tfrac{1}{2}\ln(m^2\mu^{-2})\,,$$

where μ is an arbitrary number with dimensions of mass. All the terms which are badly behaved in the coincidence limit have been displayed explicitly in Eqs. (2.3) and (2.5). [We regard a term in $\langle T^{(1)}_{\mu\nu}\rangle$ as badly behaved if it is not continuous in the coincidence limit; this encompasses both divergent terms, such as $(\sigma_\beta\sigma^\beta)^{-1}g_{\mu\nu}$, and terms which are direction dependent and do not vanish in the limit, such as $R\sigma_\mu\sigma_\nu(\sigma_\alpha\sigma^\alpha)^{-1}$. We regard a term in $G^{(1)}$ as badly behaved if it yields a badly behaved term in $\langle T^{(1)}_{\mu\nu}\rangle$ when the latter is formed from $G^{(1)}$ by covariant differentiation. (The relation between the two expansions is explained in Ref. 5 and in Appendix D of Ref. 4.) Thus $\sigma\ln|\frac{1}{2}m^2\sigma|$ is badly behaved in the context of $G^{(1)}$ because, although it vanishes as $\sigma \to 0$, it does not have continuous second derivatives.]

Finally, the expansion could (if the metric is sufficiently smooth) in principle be carried out to an arbitrary finite order in m^{-2} or σ, and the coefficients in each order would always be local geometrical objects formed polynomially from the curvature scalar and its covariant derivatives. The terms of order m^{-2} in $T^{(1)}_{\mu\nu}$ and those of orders m^{-4} and $m^{-2}\sigma$ in $G^{(1)}$ are typical of the higher-order terms; we are carrying them along in the two-dimensional discussion to demonstrate their completely harmless nature. [The term of order $m^{-2}\sigma^0$ in $G^{(1)}$ *does* enter into the renormalization problem; it appears in the wave equation (1.1) and stress tensor (1.2) multiplied by m^2. The reader may find it instructive to verify, using Eq. (A.5d), that the expression (2.3) satisfies the wave equation in x up to the indicated orders in m and σ.]

For comparison with the massless case later, it is important to note that the part of Eq. (2.5) which arises from the first two terms in Eq. (1.2) (which together are traceless) is

$$\tilde{T}^{(1)}_{\mu\nu}(x,x') \equiv \tfrac{1}{2}(\overline{G}^{(1)}_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}\overline{G}^{(1)\rho}_{\rho}) = (2\pi)^{-1}\Big[[(\sigma_\beta\sigma^\beta)^{-1} - \tfrac{1}{4}m^2 + \tfrac{1}{24}R]\Big(g_{\mu\nu} - 2\,\frac{\sigma_\mu\sigma_\nu}{\sigma_\alpha\sigma^\alpha}\Big) + \tfrac{1}{60}m^{-2}(R_{;\mu\nu} - \tfrac{1}{2}\Box R g_{\mu\nu}) + O(m^{-4}) + O(\sigma^{1/2})\Big]\,. \qquad (2.8*)$$

[As in Ref. 5, the bar indicates that a tensor index representing covariant differentiation at x' has been transported back to x:

$$\overline{G}^{(1)}_{\mu\nu} \equiv g_\nu^{\nu'} G^{(1)}_{;\mu\nu'}\,, \qquad (2.9)$$

where $g_\nu^{\nu'}$ is the bivector of parallel transport.

Strictly speaking, the derivatives should be symmetrized with respect to μ and ν, but this is not necessary in dimension 2 unless one wants to examine terms in $\tilde{T}_{\mu\nu}$ which vanish in the coincidence limit.] The other contribution to the stress tensor is, from Eq. (2.3) in the coincidence limit,

$$-\tfrac{1}{4}m^2 g_{\mu\nu}G^{(1)} = (2\pi)^{-1}g_{\mu\nu}[\tfrac{1}{2}m^2 L - \tfrac{1}{24}R - \tfrac{1}{240}m^{-2}(R^2 + 2\Box R) + O(m^{-4}) + O(\sigma^{1/2})] \,. \tag{2.10*}$$

The sum of expressions (2.8) and (2.10) is (2.5).

III. PHYSICAL STATES

The Schwinger-DeWitt series does not provide a complete answer to the problem of calculating $\langle\phi^2\rangle(x,x')$ [or the corresponding regularized stress tensor, $\langle T_{\mu\nu}\rangle(x,x')$] for two reasons. First to obtain $\langle\phi^2\rangle$ from $G^{(1)}$ or $\langle T_{\mu\nu}\rangle$ from $T^{(1)}_{\mu\nu}$ requires a knowledge of how the quantum state under investigation is related to $|\text{in, vac}\rangle$ and $|\text{out, vac}\rangle$. For example, even if the state in the expectation value is $|\text{in, vac}\rangle$ itself, to eliminate $|\text{out, vac}\rangle$ one needs the coefficients of the Bogolubov transformation which relates the in and out states. [See Eqs. (175) and (176) and preceding text in Ref. 21, and also Ref. 18.] The same remark applies to approaches based on effective actions.

Second, the series expansion techniques tell us nothing about the *remainder* in the expanded quantity after the series has been calculated to a given order, except that it is a smooth function in the coincidence limit and that it falls off rapidly (i.e., to the appropriate order) as a functional of dimensionless ratios of geometrical quantities to the mass of the field. The remainder depends on global boundary conditions and, in general, is not a local functional of the curvature tensor. The local terms in the series for $\langle T_{\mu\nu}\rangle$ cannot describe the "real" matter present at a point, which is a datum independent of the local geometry. (Recall that we are studying the dynamics of a field against a given background metric; gravitational field equations are not yet relevant.) Particles (or energy, etc.) can be produced in the past and then propagate to the space-time point concerned. Hence, even when definite boundary conditions are imposed to fix the state of the matter field, the stress tensor at x depends on the geometry in other regions of space-time. Thus one of the most interesting parts of the quantity being investigated is being lost in the local expansion process.

Consequently, it is desirable to calculate $\langle\phi^2\rangle$ or $\langle T_{\mu\nu}\rangle$ directly, with respect to a particular quantum state. This cannot be done in full generality, but it is feasible for particular classes of space-time metrics, especially those where the classical field equation can be solved by separation of variables. These "mode sum" calculations generate, at least in the intermediate stages, complicated expressions which are not manifestly covariant in structure. It is therefore not easy to see which of the divergences inevitably encountered are due to an unphysical choice of state (containing, intuitively speaking, an infinite density of matter) and which constitute a ubiquitous "vacuum" effect which can and must be removed by some kind of renormalization. Furthermore, even if the state is known to by physically acceptable, it is difficult to isolate the divergent terms covariantly from the finite, physical remainder. The knowledge of the generally covariant leading terms in $G^{(1)}$ and $T^{(1)}_{\mu\nu}$ is very helpful in bringing order into this chaos. Thus the proper-time and mode-sum approaches are complementary.

In the case of Robertson-Walker and similar metrics, a solution to the problem of choosing a physically realizable state was proposed in Ref. 6. This approach is based on a study of the behavior of the solutions of the field equation in the limit of high frequency, which is where the divergences a appear.

For the two-dimensional Robertson-Walker metric (1.3) the general operator-valued solution of Eq. (1.1) is

$$\phi(\eta, y) = (2\pi)^{-1/2}\int_{-\infty}^{\infty} dk[a_k\psi_k(\eta)e^{iky} + a_k^\dagger\psi_k(\eta)^* e^{-iky}] \,, \tag{3.1}$$

where

$$\frac{d^2\psi_k}{d\eta^2} + \omega_k{}^2\psi_k = 0, \quad \omega_k{}^2 = k^2 + m^2 C \,, \tag{3.2}$$

and ψ_k is normalized to yield the creation-annihilation commutation relations, $[a_k, a_l^\dagger] = \delta(k-l)$, etc. For each complete set of such ψ_k's one has a "vacuum" state defined by $a_k|0\rangle = 0$. If C is independent of the time coordinate η, then the usual theory of a free field is obtained by taking

$$\psi_k(\eta) = (2\omega_k)^{-1/2}e^{-i\omega_k\eta} \,;$$

but when C is not constant, there is no unique, natural choice of ψ_k. One might try to choose creation and annihilation operators at each fixed time η_0 so as to diagonalize the instantaneous Hamiltonian:

$$H(\eta_0) = \int dk\, \omega_k(\eta_0)a_k(\eta_0)^\dagger a_k(\eta_0) + \text{infinite } c \text{ number} \,,$$

and then define finite observables at time η_0 by normal-ordering the formal expression with respect to these operators. However, the ansatz leads to unphysical results.[26-29] The trouble is that the Bogolubov transformations relating the operators at different times,

$$a_k(\eta_1)=\alpha_k a_k(\eta_0)+\beta_k a_{-k}(\eta_0)^\dagger ,$$

contain pair-creation amplitudes β_k which fall off only as some low inverse power of k (more precisely, of the ratios of k to derivatives of C) as $|k|\to\infty$. In the generic four-dimensional case one has $\beta_k\sim k^{-1}$, resulting in infinite densities of particles and of energy created from vacuum. An exception is the case where all derivatives of C vanish at η_0 and again at η_1; then β_k falls off faster than any power of k^{-1}, and the particle interpretation of the theory seems as sound and straightforward as in flat space. [Here we assume that $C(\eta)$ is differentiable to all orders in the interval between η_0 and η_1.] In particular, this is true of the Bogolubov transformation relating the in and out particle structures for a space-time asymptotically static in past and future. One can say that the "real" particle creation is of *transcendental order* in k, or in the global curvature (more precisely, in the intermediate time dependence of C).

The aim of Ref. 6 was to characterize "vacuum" physical conditions at each time in such a way that (1) the definition applies to an *arbitrary* Robertson-Walker geometry, whether or not asymptotically static, and (2) when there *are* initial and final vacuum states, they are "vacuous" in the ultraviolet modes, by the new definition, at *all* times, including intermediate times when the universe is expanding or contracting. The result is a notion of vacuum which is only approximate (being imposed to some finite order in the ratios of k to derivatives of C), but is independent of time, since any particle creation which occurs is not of finite order in k. The key observation is that when C had a nonzero derivative of any order, the natural frequency of oscillation of the solution ψ_k of Eq. (3.2) is not ω_k, but something shifted slightly away from ω_k. An analogy may make this clearer. Consider the damped oscillator equation with constant coefficients,

$$\frac{d^2\psi}{dt^2}+2\gamma\frac{d\psi}{dt}+\omega^2\psi=0 .$$

If $\gamma=0$, the solution are linear combinations of $e^{\pm i\omega t}$. When $0<\gamma<\omega$, the basis solutions are

$$e^{-\gamma t}e^{\pm i(\omega^2-\gamma^2)^{1/2}t} .$$

In addition to the obvious damping, γ has caused a shift in frequency away from the "bare" value ω. A naive attempt to calculate the solution as a power series in γ results in $e^{\pm i\omega t}$ times a power series in t; this result is very nonuniform in t and is a classic example of how not to do perturbation theory.[30] The situation is similar when $\gamma=0$ but ω is a function of time, as in Eq. (3.2). There is a basis of solutions of the form

$$\psi_k^\pm(\eta)=(2W_k)^{-1/2}\exp\left[\mp\int^\eta W_k(\eta')d\eta'\right] +\text{remainder} , \qquad (3.3)$$

where the effective frequency $W_k(\eta)$ is chosen to have an asymptotic series of the schematic form

$$W_k(\eta)=\omega_k(\eta)[1+\delta_2(\eta)\omega_k^{-2}+\delta_4(\eta)\omega_k^{-4}+\cdots] . \qquad (3.4)$$

The $\delta_{2n}(\eta)$ are functions of $C(\eta)$ and its derivatives (at the point η only) and of k and m. The remainder in Eq. (3.3) vanishes to a high order in k^{-2} when W_k satisfies Eq. (3.4) to a corresponding order.

Now let us use ψ_k^+ for the ψ_k in the field expansion (3.1) for all large k. [More precisely, let ψ_k be an exact solution of Eq. (3.2) with initial values $\psi_k(\eta_0)$ and $\psi_k'(\eta_0)$ (at some arbitrary η_0) equal to the respective values from the expression in Eq. (3.3) with the upper sign.] The corresponding $|0\rangle$ represents a state of the universe which (at *all* times) is empty of matter in the ultraviolet modes to a given order in k: an *adiabatic vacuum.*

The old strategy of Hamiltonian diagonalization corresponds to expanding out the factor

$$e^{i\omega\eta}\exp\left(-i\int_{\eta_0}^{\eta}W\,d\eta'\right)$$

in ψ_k^+ as a power series in k^{-1} (obtaining a result highly nonuniform in η) and using the initial data of the lowest-order term in that expression to define a solution which appears to have "positive frequency" at η_0. The example of the damped harmonic oscillator helps show why that procedure for defining a vacuum is physically wrong.

As adiabatic vacuum is not unique, since Eq. (3.4) does not uniquely define W_k. However, the construction yields a *class* of physically acceptable vacuum states and corresponding particle notions, which are related among themselves by Bogolubov transformations with $\beta_k\sim k^{-2N+1}$ for some N. The class can be made narrower by carrying the adiabatic expansion (3.4) to higher order, thereby increasing N. Whenever they exist, $|\text{in, vac}\rangle$ and $|\text{out, vac}\rangle$ are among the adiabatic vacuum states. The main significance of the construction is that it determines the space of "physically acceptable" states, a dense domain of vectors in a Hilbert space, with respect to which the expectation values of observables, such as $T_{\mu\nu}(x)$, are finite, or should become finite after renormalization. (When $m=0$, so that ω_k does not have a positive lower bound, one must also restrict the choice of ω_k as $k\to 0$ to avoid infrared divergences.[31]) These states are obtained from any adiabatic vacuum of sufficiently high order by acting on it with combinations of creation-operator monomials

$$\prod_{j=1}^{n} a_{k_1}^\dagger \cdots a_{k_j}^\dagger$$

with coefficients that fall off sufficiently rapidly in k_j and n.

IV. CALCULATION OF THE REGULARIZED STRESS TENSOR IN A STRICTLY ADIABATIC VACUUM STATE

An adiabatic vacuum is a natural choice of state for which to investigate the point-separated quantities $\langle\phi^2\rangle(x,x')$ and $\langle T_{\mu\nu}\rangle(x,x')$. For details of the study reported here, see Ref. 17.

We define the point separation (and choose the metric sign convention) as in Refs. 1, 2, and 4; the relation between this formalism and that of Ref. 5 is discussed thoroughly in Ref. 4. In summary, $\langle T_{\mu\nu}\rangle(x,x')$ is regarded as the expectation value of $T_{\mu\nu}(x_0;\epsilon,t^\rho)$, a function of the point x_0 midway between x and x' and of the tangent vector to the geodesic through the three points. The distance from x_0 to x or x' is ϵ, and t^ρ is normalized ($t_\alpha t^\alpha\equiv\Sigma=\pm1$). Thus we have (within either sign convention)

$$\sigma^\rho=-2\epsilon t^\rho,\quad \sigma=2\Sigma\epsilon^2\,, \tag{4.1}$$

and t^ρ is now thought of as rooted at x_0, not x. Furthermore, the coefficients in the geodesic power series for $\langle\phi^2\rangle$ and $\langle T_{\mu\nu}\rangle$ are to be expanded about x_0, rather than x as in the Hadamard-DeWitt approach. In two dimensions this symmetrization does not affect any terms which survive when $\epsilon\to 0$, but it eliminates the terms of order $\sigma^{1/2}$. The (ϵ, t^ρ) notation has the advantage of making the behavior of a term in the coincidence limit obvious at a glance, while the σ notation has technical advantages in covariant calculations.

The Robertson-Walker calculation is most easily done in the null coordinate system ($v=\eta+y$, $u=\eta-y$), where the components of the stress tensor (1.2) are

$$\begin{aligned} &T_{uu}=(\partial_u\phi)^2,\quad T_{vv}=(\partial_v\phi)^2\,,\\ &T_{uv}=T_{vu}=\tfrac{1}{4}m^2C\phi^2\,. \end{aligned} \tag{4.2}$$

Note that

$$\tilde T_{uu}=T_{uu},\quad \tilde T_{vv}=T_{vv}\,,\quad \tilde T_{uv}=0\,. \tag{4.3}$$

Because of the u-v symmetry, the only essentially independent objects to be calculated are the point-separated expectation values of ϕ^2 and $(\partial_u\phi)^2$.

The field operator ϕ has the expansion (3.1), where the ψ_k are solutions of Eq. (3.2). The adiabatic approximation (3.3) is good when the coefficient function $\omega_k{}^2$ in Eq. (3.2) is *large* or *slowly varying*, that is, when either k^2 or m^2 is large compared to the derivatives of C. The discussion is simplified by introducing a single formal adiabatic parameter, T: One studies the asymptotics of

$$d^2\psi/d\eta^2+T^2\omega^2(\eta)\psi=0$$

in the limit $T\to\infty$, but sets $T=1$ in the results.

We determine the square of a suitable $W_k(\eta)$ to order T^{-4} by the Chakraborty method[32] in Appendix B. This truncated $W_k(\eta)^2$ is positive for all sufficiently large T, and the exact solutions of Eq. (3.2) have the form

$$\psi_k(\eta)=(2W_k)^{-1/2}\Big[\alpha_k(\eta)\exp\Big(-i\int^\eta W_k d\eta'\Big) +\beta_k(\eta)\exp\Big(i\int^\eta W_k d\eta'\Big)\Big] \tag{4.4}$$

with $\alpha_k(\eta)$ and $\beta_k(\eta)$ constant to order T^{-4}. The state annihilated by the a_k in Eq. (3.1) is, by definition, a fourth-order adiabatic vacuum if the mode functions have $\alpha_k\simeq 1$ and $\beta_k\simeq 0$ as $|k|\to\infty$. The mode functions for small k (indeed, for any finite k) are arbitrary; only the behavior of the whole family $\{\psi_k\}$ as $|k|\to\infty$ is prescribed.

In the present calculation we can exploit the dual meaning of the adiabatic limit to specify ψ_k further. If $C(\eta)$ varies sufficiently slowly on the time scale determined by the mass of the field, then $W_k(\eta)^2$ will be strictly positive (and finite) for *all* k, even $k=0$. Then it is meaningful to require Eq. (4.4) with $\alpha_k\simeq 1$ and $\beta_k\simeq 0$ to hold for all k, with $\alpha_k(\eta)$ and $\beta_k(\eta)$ becoming identically equal to 1 and 0 in the limit of large m or very slowly varying C. We call such a state a fourth-order *strictly adiabatic vacuum*. For example, one could determine the state completely by setting $\alpha_k(\eta_0)=1$ and $\beta_k(\eta_0)=0$ at some initial time η_0; other strictly adiabatic vacuum states include those corresponding to different choices of η_0 or to choices of W_k differing from the fourth-order Chakraborty form (B4) by terms of higher order.

We shall find that the strictly adiabatic vacuum states are those whose expectation values are directly given by the proper-time expansions of Sec. II. However, it is of the utmost importance to understand that the introduction of a strictly adiabatic vacuum is not essential to our program of covariant isolation and removal of the ultraviolet divergences of the stress tensor. The validity of the adiabatic characterization of the physically realizable states is by no means limited to situations of large mass or of small curvature (slow time variation of the metric). Since the divergences arise entirely from the large-k end of the integration over normal modes, in order to isolate those divergences from the contribution of "real matter" it is necessary only to impose the adiabatic vacuum condition for all k greater than some K, which can be chosen so that the dimensionless ratios of derivatives of C to powers of ω_k are small in the vicinity of any given nonsingular point of space-time. After renormalization the stress tensor will be finite in all adiabatic vacuum states and in all of the states associated with them which represent finite densities of the number and energy of real particles.

In massless theories which are not conformally trivial, such as a minimally coupled massless scalar field in a four-dimensional Robertson-Walker universe or a conformally coupled massless field in an anisotropic universe,[28] there is no strictly adiabatic vacuum because $W_k(\eta)$ becomes infinite, zero, or negative at sufficiently small $|k|$. In other words, the adiabatic expansion of the mode functions in the limit of slow variation of $C(\eta)$ is not uniform in k and hence does not lead to an adiabatic expansion of the integrated stress tensor. This is reflected in the inapplicability in the massless case (also, of course, the case of small positive mass) of the high-order terms of the proper-time series (Sec. II), which involve inverse powers of the mass.[5] Nevertheless, the divergences can be removed by comparison of the proper-time series with the contribution of the ultraviolet modes alone, and it is for this reason that subtraction of the leading terms of the adiabatic (or proper-time) expansion is the correct way to renormalize, even for zero mass or for strongly curved space-time (see Refs. 4, 9, 10, 11, 12, 13, 14, 16, and 37). It is always permissible (though perhaps technically messy) to treat high and low modes on separate footings; in the general massless case this is mandatory if the adiabatic approximation is to be used. (On the other hand, a few models can be solved in closed form.[12]

Let us return to the two-dimensional massive Robertson-Walker problem and consider expectation values with respect to $|0\rangle$, a strictly adiabatic vacuum (of at least fourth order, one order higher than needed to isolate all divergences). The two-point function

$$\langle 0|\phi^2(x_0;\epsilon,t^\rho)|0\rangle = \mathrm{Re}\langle 0|\phi(x)\phi(x')|0\rangle$$

can be calculated from Eqs. (3.1) and (4.4) [in effect, (3.3)] and the properties of annihilation and creation operators:

$$\langle\phi^2\rangle = (4\pi)^{-1}\,\mathrm{Re}\int_{-\infty}^{\infty} dk[W_k(\eta)W_k(\eta')]^{-1/2}\exp\left(i\int^{\eta} W_k d\bar\eta\right)\exp\left(-i\int^{\eta'} W_k d\bar\eta'\right)e^{ik(y'-y)} + \text{remainder}. \tag{4.5}$$

The remainder contains all effects of the deviation of α_k from 1 and β_k from 0; by construction of the strictly adiabatic vacuum, the integrand of this term is of adiabatic order higher than T^{-4}, and hence falls off fast enough in k to make the integral converge even when $\epsilon = 0$. Since the adiabatic expansion is uniform, the integral is also of order higher than T^{-4}. This term is a nonvanishing physical contribution to the expectation value of ϕ^2 (and hence $T_{\mu\nu}$). However, it is not relevant to the isolation of the terms which are discontinuous in the coincidence limit ($\epsilon \to 0$), so we may neglect it in the following discussion.

Even with the points separated, Eq. (4.5) is only a conditionally convergent, oscillatory integral. However, $\langle\phi(x)\phi(x')\rangle$ has a rigorous meaning as a distribution, and in Appendix C we show that in regions of timelike or spacelike separation of the points this distribution coincides with the function we obtain in this section by formal manipulations of the integral.

We seek an expansion of $\langle\phi^2\rangle$ in ϵ, the separation of x and x' from x_0 along the geodesic generated by t^ρ. The coordinates of x and x' can be expanded as in Ref. 2; see our Eqs. (B8) and (B9). One then finds

$$\int^{\eta} W_k d\bar\eta - \int^{\eta'} W_k d\bar\eta' = \int_{\eta'}^{\eta} W_k d\bar\eta$$
$$= 2\epsilon\eta_1 W_k + \tfrac{1}{3}\epsilon^3(\eta_3 W_k + 3\eta_1\eta_2 W_k' + \eta_1^3 W_k'') + O(\epsilon^5) \tag{4.6}$$

and a similar expression for $k(y'-y)$. [On the right-hand side of Eq. (4.6) W_k and its derivatives are evaluated at η_0.] We also have

$$[W_k(\eta)W_k(\eta')]^{-1/2}$$
$$= W_k^{-1} - \tfrac{1}{2}\epsilon^2 W_k^{-2}\{\eta_1^2[W_k'' - (W_k')^2] + \eta_2 W_k'\} + O(\epsilon^4). \tag{4.7}$$

The expansions of W_k and W_k^{-1} are Eqs. (B4) and (B6). One sees that derivatives (with respect to η_0) of W_k are of order k^{-1}. It follows that the terms beyond the first in Eq. (4.7) and in the power-series expansion (in ϵ) of

$$\exp\left[i\int_{\eta'}^{\eta} W_k d\bar{\eta} + ik(y' - y)\right]$$

yield convergent integrals of positive order in ϵ. Since these terms vanish in the coincidence limit, they may be ignored. Thus the interesting part of Eq. (4.5) becomes

$$\langle\phi^2\rangle = (4\pi)^{-1}\,\mathrm{Re}\int_{-\infty}^{\infty} dk\, W_k^{-1}\exp[2i\epsilon(\eta_1 W_k - y_1 k)] + \cdots . \tag{4.8}$$

When Eqs. (B4) and (B6) are put into Eq. (4.8), one finds that

$$\langle\phi^2\rangle = (4\pi)^{-1}\,\mathrm{Re}\int_{-\infty}^{\infty} dk\,\omega_k^{-1}\exp[2i\epsilon(\eta_1\omega_k - y_1 k)] + \text{continuous terms.} \tag{4.9}$$

This is, all the integrals that arise except the one from the leading term define finite, continuous functions of ϵ. Consequently, one can take the limit $\epsilon \to 0$ before performing the integrations, so that the exponential factors no longer appear, and the continuous terms in Eq. (4.9) become

$$-(4\pi)^{-1}\int_{-\infty}^{\infty}[A^2\omega_k^{-5} + B^2\omega_k^{-7} + K^4\omega_k^{-7} + (L^4 - A^4)\omega_k^{-9} + (M^4 - 2A^2B^2)\omega_k^{-11} + (N^4 - B^4)\omega_k^{-13}]dk + O(T^{-6}). \tag{4.10}$$

These integrals can be evaluated, since they are all of the form

$$\int_{-\infty}^{\infty}(k^2 + m^2C)^{-n-(1/2)}\,dk .$$

The results are included in Eq. (4.14).

Let I_1 stand for the remaining integral in Eq. (4.9), which contains all the divergences. Suppose first that t^ρ is spacelike, so that $y_1 > \eta_1$ and the normalization condition implies $C(\eta_1^2 - y_1^2) = -1$. Make the following change of integration variable:

$$\begin{aligned} z &= \eta_1\omega_k - y_1 k, \quad k = -Cy_1 z + C\eta_1(z^2 + m^2)^{1/2}, \\ (z^2+m^2)^{1/2}\frac{dk}{dz} &= C\eta_1 z - Cy_1(z^2+m^2)^{1/2} \\ &= -\omega_k . \end{aligned} \tag{4.11}$$

Then (taking ϵ, y_1, and η_1 all positive without loss of generality) one obtains

$$\begin{aligned} I_1 &= (2\pi)^{-1}\int_0^{\infty}(z^2+m^2)^{-1/2}\cos 2\epsilon z\, dz \\ &= -\tfrac{1}{4}\mathrm{Re}Y_0(2i\epsilon m) = (2\pi)^{-1}K_0(2\epsilon m), \end{aligned} \tag{4.12a}$$

according to Ref. 33, Eqs. (3.387.7) and (8.550.1) or Eq. (3.771.1) (see Appendix C for further discussion).

Now let t^ρ be timelike, so that $\eta_1 > y_1$ and $C(\eta_1^2 - y_1^2) = 1$. Setting $z = \eta_1\omega_k - y_1 k$ as before gives $k = Cy_1 z \pm C\eta_1(z^2 - m^2)^{1/2}$. It is no longer possible to choose the square-root sign consistently over the whole range of integration. The interval must be split up at $z = m$ (that is, $k = Cmy_1$):

$$I_1 = (4\pi)^{-1}\,\mathrm{Re}\left\{\int_{-\infty}^{Cy_1 m}\omega_k^{-1}\exp[2i\epsilon(\eta_1\omega_k - y_1 k)]dk + \int_{Cy_1 m}^{\infty}\omega_k^{-1}\exp[2i\epsilon(\eta_1\omega_k - y_1 k)]dk\right\}.$$

In the first integral set $k = Cy_1 z - C\eta_1(z^2 - m^2)^{1/2}$, and in the second set $k = Cy_1 z + C\eta_1(z^2 - m^2)^{1/2}$. One obtains two copies of the same z integral:

$$\begin{aligned} I_1 &= 2(4\pi)^{-1}\int_m^{\infty}(z^2 - m^2)^{-1/2}\cos 2\epsilon z\, dz \\ &= (2\pi)^{-1}\mathrm{Re}K_0(2i\epsilon m) = -\tfrac{1}{4}Y_0(2\epsilon m) \end{aligned} \tag{4.12b}$$

according to Ref. 33, Eq. (3.387.6) or Eq. (3.771.9).

In both the timelike and the spacelike case we find from the series representations of the Bessel functions that

$$I_1 = -(2\pi)^{-1}(\ln|m\epsilon| + \gamma) \tag{4.13}$$

plus terms which vanish as $\epsilon \to 0$. Adding Eqs. (4.10) and (4.13), we have

$$\begin{aligned} \langle 0|\phi^2(x_0;\epsilon,t^\rho)|0\rangle = &-(2\pi)^{-1}L + (24\pi m^2C)^{-1}D' \\ &+ (120\pi m^4C^2)^{-1}[-D^{(3)} + 2D''D + \tfrac{3}{2}(D')^2 - D'D^2] \\ &+ O(T^{-6}) + O(\epsilon^2\ln|\epsilon|), \end{aligned} \tag{4.14}$$

where $L = \ln|m\epsilon| + \gamma$ is defined by Eq. (2.4). This can be put into geometrical form through Eqs. (1.4), (B11), and (B13):

$$\langle 0|\phi^2(x_0;\epsilon,t^\rho)|0\rangle=(2\pi)^{-1}[-L+\tfrac{1}{12}m^{-2}R+\tfrac{1}{120}m^{-4}(R^2-2\Box R)]+O(m^{-6})+O(\epsilon^2\ln|\epsilon|)\,. \tag{4.15}$$

This result is in complete agreement with the proper-time series (2.3) when the differences in conventions are taken into account. [Note that the term $-\frac{1}{24}m^{-2}R_{;\alpha}\sigma^\alpha=\frac{1}{12}m^{-2}R_{;\alpha}t^\alpha\epsilon$ in Eq. (2.3) merely compensates for the fact that there $\frac{1}{12}m^{-2}R$ is evaluated at the end point x instead of the midpoint x_0.]

The calculation of $\langle T_{uu}\rangle=\langle(\partial_u\phi)^2\rangle$ [with points separated as in Eq. (2.4) of Ref. 2] is similar. The matrix element is expressed in terms of the derivatives of the mode function (3.3), and hence

$$\langle T_{uu}\rangle=U_\epsilon U_{-\epsilon}(16\pi)^{-1}\operatorname{Re}\int_{-\infty}^{\infty}dk\,W_k^{-1}[(W_k+k)^2+\tfrac{1}{4}(W_k'/W_k)^2]\exp[iW_k(2\epsilon\eta_1+\tfrac{1}{3}\epsilon^3\eta_3)-ik(2\epsilon y_1+\tfrac{1}{3}\epsilon^3y_3)]+\cdots\,, \tag{4.16}$$

where $U_{\pm\epsilon}$ are the parallel-transport scale factors [Ref. 2, Eq. (2.20a)]. The dots in Eq. (4.16) represent (1) the "remainder" terms, which have smooth coincidence limits and are of high adiabatic order, and (2) terms which are being dropped because they vanish in the coincidence limit. As regards terms of the second type, note that, as before, it is permissible to replace $[W_k(\eta)W_k(\eta')]^{-1/2}$ by $W_k(\eta_0)^{-1}$, but that this time an extra term proportional to ϵ^3 must be retained from Eq. (4.6), because it modifies the leading quadratic divergence [see Eq. (4.19)] so as to produce a finite, nonzero contribution to $\langle T_{uu}\rangle$.

When the adiabatic expansion of W_k is put into Eq. (4.16), one finds the quadratic divergence to be concentrated in

$$I_2=U_\epsilon U_{-\epsilon}(16\pi)^{-1}\operatorname{Re}\int_{-\infty}^{\infty}dk\,\omega_k^{-1}(\omega_k+k)^2\exp[i(\omega_k\delta-k\delta')]\,,$$
$$\delta\equiv2\epsilon\eta_1+\tfrac{1}{3}\epsilon^3\eta_3,\quad \delta'\equiv2\epsilon y_1+\tfrac{1}{3}\epsilon^3y_3\,. \tag{4.17}$$

As before, we let the exponent $z=\omega_k\delta-k\delta'$ be the new variable of integration, treat spacelike and timelike separations separately, and obtain oscillatory integrals which can be identified with Bessel functions, the spacelike and timelike results being related by analytic continuation from real to imaginary argument:

$$I_2=\tfrac{1}{16}U_\epsilon U_{-\epsilon}Cm^2(\delta+\delta')(\delta-\delta')^{-1}\operatorname{Re}Y_2([Cm^2(\delta^2-\delta'^2)]^{1/2}) \tag{4.18}$$

regardless of the sign of $\delta^2-\delta'^2$. (The spacelike case is discussed rigorously in Appendix C.) As $\epsilon\to0$ it suffices to keep the first two terms in the series for Y_2:

$$\begin{aligned}I_2&=-(16\pi)^{-1}U_\epsilon U_{-\epsilon}(\delta-\delta')^{-2}[4+Cm^2(\delta^2-\delta'^2)+\cdots]\\&=(2\pi)^{-1}\{-[8\epsilon^2(t^u)^2]^{-1}+\tfrac{1}{96}(2D'-D^2)+(\tfrac{1}{48}D'-\tfrac{1}{8}Cm^2)t^v/t^u+\cdots\}\,.\end{aligned} \tag{4.19}$$

In the last step Eqs. (B9) and (B10) have been used.

To I_2 must be added the remaining, continuous terms of low adiabatic order in the expansion of Eq. (4.16). These are integrals of the same type as in Eq. (4.10). Adding them to Eq. (4.19), we obtain as the leading terms of $\langle 0|T_{uu}(x_0;\epsilon,t^\rho)|0\rangle$ a quadratic divergence of adiabatic order T^0,

$$-[16\pi\epsilon^2(t^u)^2]^{-1}\,, \tag{4.20a}$$

a finite term of order T^{-2},

$$(16\pi)^{-1}[-\tfrac{1}{12}D'+(\tfrac{1}{6}D'-Cm^2)t^v/t^u]\,, \tag{4.20b}$$

and a finite term of order T^{-4},

$$(480\pi Cm^2)^{-1}[D^{(3)}-3D''D-(D')^2+2D'D^2]=(120\pi m^2)^{-1}R_{;uu}\,. \tag{4.20c}$$

By comparing Eqs. (4.20) with Eqs. (B11)–(B14) one can obtain a geometrical expression for $\langle 0|\tilde T_{\mu\nu}(x_0;\epsilon,t^\rho)|0\rangle$ [defined above Eq. (2.8)]. In the first two adiabatic orders all terms must be proportional to either $g_{\mu\nu}$ or $t_\mu t_\nu$, the only available tensors. The $g_{\mu\nu}$ term is left arbitrary by Eqs. (4.20), since $g_{uu}=0$, but it is determined by the condition $\tilde T_{\nu\nu}=0$: Only the traceless combination $g_{\mu\nu}-2t_\mu t_\nu/\Sigma$ can appear. Equation (2.8) is thus reproduced. Adding $\frac{1}{2}m^2\langle\phi^2\rangle g_{\mu\nu}$ with $\langle\phi^2\rangle$ given in Eq. (4.15), we obtain Eq. (2.5) as the expansion of $\langle 0|T_{\mu\nu}(x_0;\epsilon,t^\rho)|0\rangle$: We have the quadratic divergence

$$(2\pi)^{-1}(4\epsilon^2\Sigma)^{-1}(g_{\mu\nu}-2t_\mu t_\nu/\Sigma)\,, \tag{4.21a}$$

the logarithmic term

$$-(2\pi)^{-1}\tfrac{1}{2}m^2 g_{\mu\nu}L\,, \tag{4.21b}$$

and the first three adiabatic orders of the finite term

$$(2\pi)^{-1}\{\tfrac{1}{4}m^2(g_{\mu\nu}-2t_\mu t_\nu/\Sigma)+\tfrac{1}{12}Rt_\mu t_\nu/\Sigma+\tfrac{1}{60}m^{-2}(R_{;\mu\nu}-\Box R g_{\mu\nu}+\tfrac{1}{4}R^2 g_{\mu\nu})\}\,. \tag{4.21c}$$

Here we have reproduced all the discontinuous terms in the proper-time series for the stress tensor, and also the first of the higher-order continuous terms proportional to negative powers of the mass. It is clear that agreement to arbitrarily high adiabatic order could be attained if one started with a W_k correctly specified to the corresponding order—i.e., with a "better vacuum." This result has a certain polemical significance. First, it confirms that the Schwinger-DeWitt integral (A2) is relevant to the expectation values of genuine field operators in definite quantum states. (This was questioned in, for instance, Refs. 34 and 35.) Also, it demonstrates the necessity of the *high-order* adiabatic definition of "vacuum" at a time when the universe is expanding.[6] If, in the definition of creation and annihilation operators, W_k had been replaced by ω_k (the lowest-order WKB approximation, corresponding to diagonalization of the Hamiltonian), then the covariant finite terms of orders m^0 and m^{-2} in $\langle T_{\mu\nu}\rangle$ would not have been reproduced; they would have been submerged in nongeometrical contributions ("remainder terms") of the same order, representing excess mass present in the high-frequency modes of the field. The long-term significance of the work, however, is that it points the way to practical calculation of the actual value of the renormalized $\langle T_{\mu\nu}\rangle$ for a given state (which may or may not be an adiabatic vacuum). How this is to be done is described at the end of Sec. V.

V. RENORMALIZATION

In this section we determine (and discuss the properties of) the expectation values of the "renormalized" one-point operators, $\langle 0|\phi(x)^2|0\rangle$ and $\langle 0|T_{\mu\nu}(x)|0\rangle$, starting from the short-distance expansions, determined above, of the corresponding two-point quantities. The ansatz employed is that developed in Refs. 1, 2, and especially 4; its justification is discussed in Sec. VIII.

The "adiabatic order" d_∂ was introduced in connection with Eq. (2.6). Observe that all terms with $d_\partial > 2$ in the expansion (2.5) or (4.21) of $\langle 0|T_{\mu\nu}(x;\epsilon,t^\rho)|0\rangle$, and also the remainder of the series, have smooth limits as $\epsilon\to 0$. The term involving $Rt_\mu t_\nu/\Sigma$, although finite, is not continuous at $\epsilon=0$ because the limit depends on the direction of approach. In $\langle 0|\phi^2(x;\epsilon,t^\rho)|0\rangle$ [Eq. (2.3) or (4.15)] the only discontinuous term is $-(2\pi)^{-1}L$, the term with $d_\partial=0$.

We take the renormalized quantities to consist of the terms with unambiguous limits, plus arbitrary linear combinations of the local geometrical c numbers of the orders d_∂ in which discontinuities appear. In the case of the stress tensor, these ambiguous local terms are restricted by the requirement of conservation ($T_\mu{}^\nu{}_{;\nu}=0$). Thus we obtain in the two-dimensional massive scalar theory

$$\langle 0|\phi(x)^2|0\rangle=(2\pi)^{-1}[\alpha+\tfrac{1}{12}m^{-2}R+\tfrac{1}{120}m^{-4}(R^2-2\Box R)] + \text{remainder }[O(m^{-6})]\,, \tag{5.1}$$

$$\langle 0|T_{\mu\nu}(x)|0\rangle=(2\pi)^{-1}[\tfrac{1}{2}\beta m^2 g_{\mu\nu} + \tfrac{1}{60}m^{-2}(R_{,\mu\nu}-\Box R g_{\mu\nu}+\tfrac{1}{4}R^2 g_{\mu\nu})] + \text{remainder }[O(m^{-4})]\,, \tag{5.2}$$

where α and β are arbitrary dimensionless numbers. [From now on we shall be dealing only with quantities for which $d_\epsilon=0$ in Eq. (2.6); therefore we can use the order in m as a handy way of indicating the adiabatic order of a term.] We obtain renormalized *operators* $\phi^2(x)$ and $T_{\mu\nu}(x)$ by adding their normal-ordered parts to the c-number parts (5.1) and (5.2).

The effect of the arbitrary constants is to renormalize coupling constants in all equations of motion involving ϕ^2 and $T_{\mu\nu}$. In particular, such coupling constants must be assumed to exist. For example, suppose we enlarged the theory to include another scalar field ψ satisfying on the formal level the equation

$$\Box\psi=\xi R\psi+c\phi^2\psi\,.$$

The presence of an arbitrary constant α in ϕ^2 would require us to modify the postulated theory to allow ψ to have a mass, which could be determined only by experiment:

$$\Box\psi=\mu^2\psi+\xi R\psi+c\phi^2\psi\,.$$

[In fact, to get a definite value for μ^2 requires *both* an experiment and a definite choice of α, since the $c\phi^2\psi$ term must be taken into account in interpreting the experimental data. In the present case it would be natural to choose $\alpha = 0$, making $\langle\phi^2\rangle$ vanish to lowest order in derivatives of $g_{\mu\nu}$; but in some situations there is no distinguished choice of the constants (the "renormalization-group ambiguity").] On the other hand, the $(24\pi)^{-1}m^{-2}R$ in $\langle\phi^2\rangle$ could be taken to renormalize ξ, but since that term is finite and unambiguous, there is really no need to introduce a ξR coupling into a theory which does not have one already. Therefore, we take the $(24\pi)^{-1}m^{-2}R$ to be definite and physically meaningful, since it could potentially appear in a context where there is no coupling constant to absorb it, like

$$\Box\chi = \mu^2\chi + c\phi^2\chi .$$

The same holds for local terms of all higher orders in $\langle\phi^2\rangle$ or $\langle T_{\mu\nu}\rangle$.

Next we must discuss the conservation and trace properties of $\langle 0|T_{\mu\nu}(x)|0\rangle$. Since covariant differentiation increases the d_∂ of a term by precisely 1, $\langle T_{\mu\nu}\rangle$ can be conserved only if its series is conserved order by order in ∂ (equivalently, in m). This is what forbids a term proportional to $Rg_{\mu\nu}$ in Eq. (5.2). It is easy to check that the object $R_{;\mu\nu} - \Box R g_{\mu\nu} + \frac{1}{4}R^2 g_{\mu\nu}$ is conserved, and it would be extremely surprising if the same were not true of the higher-order terms and of the remainder at any stage. In Sec. VIII we prove this conservation law by Wald's method.[14]

At the classical level the trace of the stress tensor (1.2) is $m^2\phi^2$. The trace of the tensor (5.2) is

$$\langle 0 | T^\alpha_\alpha(x)|0\rangle = (2\pi)^{-1}[\beta m^2 + \tfrac{1}{80}m^{-2}(-\Box R + \tfrac{1}{2}R^2)] + O(m^{-4}) , \tag{5.3}$$

whereas

$$m^2\langle 0|\phi(x)^2|0\rangle = (2\pi)^{-1}[\alpha m^2 + \tfrac{1}{12}R + \tfrac{1}{120}m^{-2}(R^2 - 2\Box R)] + O(m^{-4}) , \tag{5.4}$$

so that

$$\langle 0|T^\alpha_\alpha(x)|0\rangle = m^2\langle 0|\phi(x)^2|0\rangle + (2\pi)^{-1}(\beta-\alpha)m^2 - (24\pi)^{-1}R + O(m^{-4}) . \tag{5.5}$$

Again, the analysis of Sec. VIII shows that the $O(m^{-4})$ term in Eq. (5.5) is exactly zero. For simplicity we assume henceforth that $\beta = \alpha$, so that [cf. Ref. 36]

$$\langle 0|T^\alpha_\alpha(x)|0\rangle = m^2\langle 0|\phi(x)^2|0\rangle - (24\pi)^{-1}R . \tag{5.6}$$

The origin of the anomaly term, $-(24\pi)^{-1}R$, is clear: The term of order $d_\partial = 2$ in $T_{\mu\nu}$ was destroyed in the renormalization, while that in ϕ^2 was retained.

More profound, however, is the observation that this anomaly is *inevitable* as long as (1) $\langle\phi^2\rangle$ is given by Eq. (5.1), (2) $\langle T_{\mu\nu}\rangle$ is conserved, (3) the series for $\langle T_{\mu\nu}\rangle$ is purely local. As remarked above, conservation of $\langle T_{\mu\nu}\rangle$ implies that the terms in the series are conserved individually. But in dimension 2 there is *no* conserved local tensor with $d_\partial = 2$, since $R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}$ vanishes. Hence $\langle T^\alpha_\alpha\rangle$ cannot contain any terms with $d_\partial = 2$, and the term of that order in $m^2\langle\phi^2\rangle$ [see Eq. (5.4)] must be canceled by an anomaly in the trace relation.

Equation (5.6) is actually an operator identity for the renormalized observables (i.e., the "$\langle 0|\cdots|0\rangle$" is gratuitous). This is so because the difference between two properly normalized matrix elements of an operator (such as $\langle 0|\phi^2|0\rangle - \frac{1}{2}G^{(1)}$) is well defined in the coincidence limit; the divergent and direction-dependent terms are always c numbers. Therefore, such a quantity can be calculated from its formal classical expression, which is identically zero in the case of the operator $T^\alpha_\alpha - m^2\phi^2$.

The anomaly can be removed (in the massive case only) by redefining ϕ^2 to delete the c number term $(24\pi)^{-1}m^{-2}R$. This seems unjustified unless, as remarked previously, one is prepared to insert curvature-scalar coupling terms into all interactions involving ϕ^2.

The goal of the entire foregoing discussion, of course, is the actual calculation of the expectation values of the stress tensor in particular states. The calculation in Sec. IV has not brought us to that goal; the replacement of ψ_k by its adiabatic approximation (3.3), just like the proper-time expansion process discussed at the beginning of Sec. III, amounts to neglecting the nonlocal remainder terms of "transcendental order," which are of great physical interest. Nevertheless, a local expansion based on normal modes brings one a step closer to an explicit answer than an expansion based on the proper-time representation (A2). By inspecting the details of the calculation, one can determine precisely which terms of the integrand of the "mode sum" give rise to the leading terms of the covariant expansion (2.5) or (4.21), the terms which are *discarded* in the renormalization process yielding Eq. (5.2). These terms can then be subtracted from the integrand before integration, and one thus obtains a finite integral (with $\epsilon = 0$) to evaluate for the physical expectation value of $\phi(x)^2$ or $T_{\mu\nu}(x)$. Evaluating the integral may not be trivial—in particular, the adiabatic approximation must *not* be used for all modes, lest the nonlocal contribution be lost again. However, the calculation *is* a routine problem of classical applied mathematics; the questions of physical principle have been settled. In this sense we may regard the problem of a massive scalar field in a two-dimensional Robertson-Walker universe as *solved*.

To obtain Eq. (5.2) (with $\beta=0$) one must subtract from the regularized stress tensor all the terms which appear explicitly in Eqs. (4.21) except the final, continuous, $O(m^{-2})$ term of Eq. (4.21c), and then set $\epsilon=0$ in the remainder. Looking back through the details of the calculation of Sec. IV, one can see that these subtraction terms are *precisely* the contributions (except those that vanish as $\epsilon\to 0$) of the terms of order T^0 and T^{-2} in the adiabatic expansion of the integrands of the mode sums [in $\langle T_{\mu\nu}\rangle$, for example, the integral I_1 and the first two terms of Eq. (4.10)]. Therefore, if we set $\epsilon=0$ in the integrand at the outset and subtract these two leading adiabatic orders before integrating, then the resulting convergent integral must be equal to the renormalized stress tensor [the last term of Eq. (4.21c) plus the remainder of order m^{-4}].

The terms to be subtracted can be written as

$$(16\pi)^{-1}[\omega_k^{-1}m^2C+\tfrac{1}{8}\omega_k^{-5}m^4C''C-\tfrac{5}{32}\omega_k^{-7}m^6(C')^2C] \tag{5.7a}$$

for $\langle T_{uv}\rangle$ and

$$(16\pi)^{-1}\{(\omega_k\pm k)^2\omega_k^{-1}-\tfrac{1}{8}\omega_k^{-3}m^2C''+\tfrac{1}{32}\omega_k^{-5}[4m^2k^2C''+7m^4(C')^2]-\tfrac{5}{32}\omega_k^{-7}m^4k^2(C')^2\} \tag{5.7b}$$

for $\langle T_{uu}\rangle$ (upper sign) and $\langle T_{vv}\rangle$ (lower sign). These expressions are to be subtracted from the integrands of the formal expressions for the components of the stress tensor in terms of mode functions, obtained by substituting the field expansion (3.1) into Eqs. (4.2). The expressions are given for integrals over the interval $-\infty<k<\infty$. Since the mode functions are even in k, the integrations would in practice be reduced to $0<k<\infty$; in this case Eqs. (5.7) should be multiplied by 2, and the cross term $(\pm 2k)$ in $(\omega_k\pm k)^2\omega_k^{-1}$ omitted.

From a computational standpoint, this prescription is the principal result of the present work. It reduces the evaluation of the physical expectation value of $T_{\mu\nu}(x)$ with respect to a given quantum state to a feasible integration, provided that the mode functions $\psi_k(\eta)$ can be obtained. The latter, of course, depend on the particular metric function $C(\eta)$ considered, and in general one must resort to a variety of analytic or numerical approximations to obtain representations of them suitable for use in the various regions of the k integration.

In Refs. 17 and 37 it was shown that the subtraction (5.7) is also valid for renormalizing the $T_{\mu\nu}$ of a massless scalar field, with the understanding that the fictitious mass m is to be taken to 0 in the end. In that limit the terms in Eqs. (5.7) of adiabatic order T^0 precisely cancel the divergent unrenormalized massless stress tensor, and the terms of order T^{-2} form m-independent integrals which reproduce the anomalous physical massless stress tensor (6.3).

VI. THE MASSLESS FIELD

We briefly review Refs. 1 and 2. In any two-dimensional space-time there are many "conformal" coordinate systems, in which the metric takes the form $C(t,x)(dt^2-dx^2)$. In such coordinates the normal-mode solutions of the massless field equation $\Box\phi=0$ are plane waves, as in flat space, and have an obvious classification into positive- and negative-frequency functions. Thus for each conformal coordinate system there is a natural vacuum state, $|\text{vac}\rangle$. For the Robertson-Walker metric in the form (1.3a), we have $|\text{in},\text{vac}\rangle=|\text{out},\text{vac}\rangle=|\text{vac}\rangle$, and the field operator is given by Eq. (3.1) with $\psi_k=e^{-i|k|\eta}$.

The representation of $\frac{1}{2}G^{(1)}=\langle\text{vac}|\phi^2|\text{vac}\rangle(x,x')$ as an integral over k diverges at the infrared end. In fact, the two-point function of a massless scalar field in two dimensions does not exist as a distribution unless one modifies the mathematical formalism of quantum field theory considerably.[38,39] However, there is no obstacle to computing

$$\langle\text{vac}|T_{\mu\nu}|\text{vac}\rangle(x,x')=T^{(1)}_{\mu\nu}(x,x')=\tilde{T}^{(1)}_{\mu\nu}(x,x')$$

directly. One obtains

$$\langle\text{vac}|T_{\mu\nu}(x;\epsilon,t^\rho)|\text{vac}\rangle=(2\pi)^{-1}[(4\Sigma\epsilon^2)^{-1}(g_{\mu\nu}-2t_\mu t_\nu/\Sigma)-\tfrac{1}{24}R(g_{\mu\nu}-2t_\mu t_\nu/\Sigma)]+\theta_{\mu\nu}+O(\epsilon^2)\,, \tag{6.1}$$

where the tensor $\theta_{\mu\nu}$ is expressed in terms of second derivatives of the metric at x in the conformal coordinate system defining $|\text{vac}\rangle$, but is not determined by the local geometry at x. In the Robertson-Walker case it is given by

$$\theta_{\eta\eta}=\theta_{yy}=(48\pi)^{-1}(D'-\tfrac{1}{2}D^2),\quad \theta_{\eta y}=0\,. \tag{6.2}$$

Renormalization is more subtle here than in the massive case because $d_\mathfrak{d}(\theta_{\mu\nu})=2$. That is, the nonlocal, state-dependent part of the stress tensor, which must be kept, is of the same order as the ambiguous local terms. Nevertheless, the nonlocal part itself has the smooth coincidence limit $\theta_{\mu\nu}$ to which one must add arbitrary amounts of the local objects of the two lowest orders. The conservation law for the complete tensor, including $\theta_{\mu\nu}$, uniquely determines the local term of order $d_\mathfrak{d}=2$, and one finally obtains

$$\langle\text{vac}|T_{\mu\nu}(x)|\text{vac}\rangle=(4\pi)^{-1}Bg_{\mu\nu}+\theta_{\mu\nu}-(48\pi)^{-1}Rg_{\mu\nu}\,. \tag{6.3}$$

B is an arbitrary quantity with dimension $[\text{mass}]^2$; presumably it can be set equal to zero in this scale-invariant model, even without the assumption of a "cosmological constant" for it to renormalize. The uniqueness and physical correctness of Eq. (6.3) are now well established.[7,8,9,15]

The discontinuous low-order geometrical terms which are subtracted from Eq. (6.1) to get Eq. (6.3) are identical to the massless limit of the terms subtracted from Eqs. (4.21) to yield Eq. (5.2). [In four dimensions the latter terms will have a finite massless limit only if the logarithmic part of $\langle T_{\mu\nu}\rangle(x,x')$ is decomposed by means of Eq. (2.7) and only the L_ϵ part is subtracted. The L_m terms remaining in the renormalized stress tensor are just absorbed in the renormalization of the arbitrary coupling constants.] In this sense the renormalization subtraction is the same for both cases (massive and massless), and it can be described as subtraction of the leading terms (i.e., the first two adiabatic orders in dimension 2, the first three orders in dimension 4) of the Schwinger-DeWitt proper-time series (2.5). We emphasize that this statement of the prescription is not an *ad hoc* assumption, but a demonstrated consequence of the fundamental principles of the method: the isolation of the divergences into geometrical terms of low adiabatic order, the removal of those terms (more precisely, their replacement by renormalization terms with arbitrary coefficients), and the requirement of conservation.

The stress tensor (6.3) is conserved, by construction. If $B=0$, its trace is $-(24\pi)^{-1}R$ [Eq. (1.5)]. Like Eq. (5.6), this equation holds on the operator level as well. Formally it is the massless limit of Eq. (5.6). However, the significance of this anomalous trace relation in the case of zero (or very small) mass is quite different from that in the case of very large mass. As $m\to\infty$, as we have seen, there is a natural definition of vacuum state, and the expectation values of $T_{\mu\nu}$ and T^α_α in that state vanish in the limit (except for the arbitrary term proportional to the metric tensor, which we can ignore). From a calculational point of view, this vanishing occurs because the extra term $\frac{1}{2}m^2\langle 0|\phi^2|0\rangle g_{\mu\nu}$ in the strictly adiabatic vacuum expectation value of $T_{\mu\nu}$ contains a contribution of order m^0 [cf. Eq. (5.4)], which cancels the anomalous trace of $\langle\tilde{T}_{\mu\nu}\rangle$. The trace anomaly in the massless case *requires* the *existence* of the nonlocal term $\theta_{\mu\nu}$, since there is no conserved, purely geometrical tensor with the trace (1.5). In contrast, we have seen that the massive anomaly [i.e., validity of Eq. (5.6)] is *required by* the absence of nonlocal terms. In a sense, the massless trace anomaly arises because of the absence of the term $\frac{1}{2}m^2\phi^2g_{\mu\nu}$ from the formula for the massless stress tensor; this point of view was emphasized in Ref. 9 and is also strongly suggested by a Pauli-Villars approach to the renormalization (Vilenkin, private communication and Ref. 44).

It would be wrong, however, to conclude that there is any physical discontinuity in the theory at $m=0$. Although the adiabatic definition of vacuum and the associated asymptotic expansion of the integrand of $\langle T_{\mu\nu}\rangle$ are always applicable to the high-frequency modes (where they suffice to determine the physically realizable states and to put the ultraviolet divergences into a form amenable to renormalization), the definition of a *strictly* adiabatic vacuum and the asymptotic expansions such as Eq. (5.2) are applicable only when the derivatives of C are small compared with m (i.e., the limit of large m for a fixed geometry, or the limit of a nearly static metric for a fixed m). To take the limit of small m with the geometry fixed in Eq. (5.2) yields no information about the expectation value of the stress tensor in any particular state. (Consequently, the inverse powers of m are no cause for alarm.) When the time variation of the metric is rapid compared with m^{-1}, the low-frequency modes generally must be given special treatment, both in the definition of the state and in the actual calculation of their contribution to $\langle T_{\mu\nu}\rangle$. For such modes the particle concept is inappropriate, and hence there is no way to define vacuum for them, even approximately. In some situations, however, it is possible to distinguish a state (e.g., an $|\text{in, vac}\rangle$) which smoothly reduces to $|\text{vac}\rangle$ in the limit $m\to 0$. The expectation value of $T_{\mu\nu}$ in such a state passes smoothly to $\langle\text{vac}|T_{\mu\nu}|\text{vac}\rangle$ in that limit, as the examples in Sec. VII show.

VII. EXAMPLES

The foregoing remarks should become clearer after consideration of the following special cases, in which explicit expressions for the renormalized stress tensor have been obtained.

In this section we write "$|\text{in}\rangle$" for "$|\text{in, vac}\rangle$" and "$|\text{out}\rangle$" for "$|\text{out, vac}\rangle$".

A. The two-dimensional "Milne universe"

If in Eqs. (1.3a) and (1.3b) we have

$$C(\eta)=e^{2\eta}, \quad t=e^{\eta}, \tag{7.1}$$

then the manifold is a region of two-dimensional flat space-time, as shown by the change of variables

$$\begin{aligned} &\xi^0=e^{\eta}\cosh y, \quad \xi^1=e^{\eta}\sinh y, \\ &ds^2=(d\xi^0)^2-(d\xi^1)^2. \end{aligned} \tag{7.2}$$

In Ref. 2 it was observed that the renormalized massless stress tensor (6.3) becomes in this model

$$\langle \text{vac}|T_{tt}|\text{vac}\rangle=-(24\pi t^2)^{-1}, \text{ etc.} \tag{7.3}$$

That this is not zero is at first sight surprising, since one would expect the vacuum expectation value to vanish in Minkowski space-time. However, as pointed out in Ref. 2, the vacuum state with respect to which Eq. (7.3) is calculated is not the same as the conventional Minkowski-space vacuum. It is the massless limit of a vacuum for massive particles which is different from the usual one. This will now be illustrated by considering the massive field explicitly.

In this model the time-dependent part of the wave equation, Eq. (3.2) is equivalent to

$$t^2\frac{d^2\psi_k}{dt^2}+t\frac{d\psi_k}{dt}+(k^2+m^2t^2)\psi_k=0, \tag{7.4}$$

whose solutions are Bessel or Hankel functions. It is well known[39-41] that the solutions

$$\psi_k^{\text{out}}\equiv\tfrac{1}{2}\sqrt{\pi}e^{\pi k/2}H_{ik}^{(2)}(mt) \tag{7.5}$$

have positive frequency with respect to *Minkowski* time, ξ^0; if they are used in the field expansion (3.1), the corresponding vacuum state is the usual one for a free field in flat space. On the other hand, for small mt the solutions

$$\psi_k^{\text{in}}\equiv[\tfrac{1}{2}\pi/\sinh(\pi|k|)]^{1/2}J_{-i|k|}(mt) \tag{7.6}$$

display positive-frequency behavior with respect to the time coordinate η, and so the vacuum state based on these solutions will correspond, in the limit $m\to 0$, to the state appearing in Eq. (6.3). The mode functions (7.5) and (7.6) have been given the canonical normalization ($\psi_k\partial_\eta\psi_k^*-\psi_k^*\partial_\eta\psi_k=i$). The coefficients of the Bogolubov transformation

$$\psi_k^{\text{in}}=\alpha_k\psi_k^{\text{out}}+\beta_k\psi_k^{\text{out}*} \tag{7.7}$$

relating them are

$$\alpha_k=\left(\frac{e^{\pi k}}{2\sinh\pi k}\right)^{1/2}, \quad \beta_k=\left(\frac{e^{-\pi k}}{2\sinh\pi k}\right)^{1/2}. \tag{7.8}$$

[The "in-out" notation reflects the fact that ψ_k^{in} has positive frequency at early times (near the coordinate singularity) and ψ_k^{out} has positive frequency at late times, even though the metric is not asymptotically static. An alternative interpretation is expressed in Eqs. (7.14).]

We know that $\langle\text{out}|T_{\mu\nu}|\text{out}\rangle=0$ in Minkowski space after renormalization. Within any regularization scheme, with a parameter ϵ, the unrenormalized $\langle\text{out}|T_{\mu\nu}(\epsilon)|\text{out}\rangle$ must be merely the divergent (as $\epsilon\to 0$) term which must be subtracted in renormalizing any expectation value of $T_{\mu\nu}$ (see Sec. V). This is easily verified for point separation.[17] Consequently, to find the renormalized $\langle\text{in}|T_{\mu\nu}|\text{in}\rangle$ it is not necessary to separate points in the unrenormalized expression; we need only to subtract, mode by mode, the unrenormalized $\langle\text{out}|T_{\mu\nu}|\text{out}\rangle$ (equivalent to normal ordering with respect to Minkowski annihilation and creation operators). It makes no difference whether the integrands are regularized, since the integral of their difference is a finite, continuous function of ϵ as $\epsilon\to 0$. From Eq. (7.7) one finds[17]

$$\begin{aligned}\langle\text{in}|\phi^2|\text{in}\rangle&=\langle\text{in}|\phi^2|\text{in}\rangle_{\text{unren}}-\langle\text{out}|\phi^2|\text{out}\rangle_{\text{unren}}\\ &=2\pi^{-1}\int_0^\infty[\text{Re}[\alpha_k^*\beta_k^*\psi_k^{\text{in}}(mt)^2]-|\beta_k|^2|\psi_k^{\text{in}}|^2]dk\\ &=\tfrac{1}{2}\int_0^\infty(\sinh\pi k)^{-2}\{\text{Re}[J_{ik}(mt)^2]-e^{-\pi k}|J_{ik}(mt)|^2\}dk,\end{aligned} \tag{7.9}$$

$$\langle\text{in}|(\partial_t\phi)^2|\text{in}\rangle=\tfrac{1}{2}\int_0^\infty(\sinh\pi k)^{-2}\left(\text{Re}\left\{\left[\frac{d}{dt}J_{ik}(mt)\right]^2\right\}-e^{-\pi k}\left|\frac{d}{dt}J_{ik}(mt)\right|^2\right)dk, \tag{7.10a}$$

$$\langle\text{in}|(\partial_y\phi)^2|\text{in}\rangle=\tfrac{1}{2}\int_0^\infty(\sinh\pi k)^{-2}\{\text{Re}[J_{ik}(mt)^2]-e^{-\pi k}|J_{ik}(mt)|^2\}k^2\,dk, \tag{7.10b}$$

from which $\langle\text{in}|T_{\mu\nu}|\text{in}\rangle$ may be formed according to Eq. (1.2).

The limit of $\langle\text{in}|T_{tt}|\text{in}\rangle$ as $t\to 0$ or $m\to 0$ may be obtained from the power series of the Bessel function

$$\langle\text{in}|T_{tt}|\text{in}\rangle = -\frac{1}{\pi t^2}\int_0^\infty \frac{k\,dk}{e^{2\pi k}-1} + O(m^2)$$
$$= -\frac{1}{24\pi t^2} + O(m^2) \qquad (7.11)$$

in precise agreement with Eq. (7.3) for $m=0$. The Planckian integral in Eq. (7.11) may be written in the form

$$-\frac{1}{\pi}\int_0^\infty \frac{q\,dq}{e^{2\pi q t}-1}, \qquad (7.12)$$

representing the negative of the energy density of radiation at temperature $(2\pi t)^{-1}$, and also as

$$-(\pi t^2)^{-1}\int_0^\infty k|\beta_k|^2 dk, \qquad (7.13)$$

suggesting that the Minkowski vacuum consist of a density of Milne particles of energy $q=k/t$.

Since the critical parameter in this problem is mt, for fixed t the state $|\text{in}\rangle$ corresponds to the massless limit and the state $|\text{out}\rangle$ corresponds to the opposite, adiabatic limit: In the notation of the previous sections, we make the identifications

$$|\text{in}\rangle \equiv |\text{vac}\rangle, \quad |\text{out}\rangle \equiv |0\rangle. \qquad (7.14)$$

B. The model of Bernard and Duncan

An even clearer demonstration of the relation of the $\langle T_{\mu\nu}\rangle$ for a finite mass to the massless and adiabatic limits is achieved by examining the work of Bernard and Duncan.[18] They treat the sigmoid cosmological model

$$C(\eta) = 1 + \tfrac{1}{2}b(1+\tanh\rho\eta) \qquad (7.15)$$

(b and ρ constants), especially in the initial region of slow expansion, $e^{\rho\eta}\ll 1$. The adiabatic parameter ρ may be identified with the reciprocal of the T in Sec. IV; in an expansion in positive powers of ρ, our d_ϑ will be a term's order in ρ. As $\eta\to\pm\infty$ the metric is essentially static, so $|\text{in}\rangle$ and $|\text{out}\rangle$ states are defined. When η is finite but $\eta\ll 0$, they find [see their Eqs. (4.4a), (4.11), (4.9c), (4.25b), and (4.23)]

$$\langle\text{in}|T^{\mu\nu}|\text{in}\rangle = -\eta_1^\mu\eta_1^\nu(4\pi)^{-1}be^{2\rho\eta}\left[m^2-\tfrac{2}{3}\rho^2-\frac{1}{2}\left(\frac{m^4}{\rho(m^2+\rho^2)^{1/2}}\right)\ln\left(\frac{(m^2+\rho^2)^{1/2}+\rho}{(m^2+\rho^2)^{1/2}-\rho}\right)\right] \qquad (7.16)$$

to first order in $be^{2\rho\eta}$ (but to *all* orders d_ϑ in derivatives of the metric). Here $\eta_{\mu\nu}$ is the underlying Minkowski metric of Eq. (1.3a). (Thus to lowest order in b the vacuum stress is purely "pressure," as predicted in Ref. 42.)

The regularization-renormalization procedure which led to Eq. (7.16) was of the Pauli-Villars variety, but there is no reason to doubt that point separation would yield the same answer. The effect of a Pauli-Villars calculation is that all terms in the stress tensor of *non-negative order in the mass* are canceled, or at least replaced by renormalization terms with arbitrary coefficients. These are precisely the discontinuous terms which are removed by our procedure. Related observations have been made by Vilenkin.[43]

In the adiabatic limit, $m\gg\rho$, one finds from Eq. (7.16)

$$\langle\text{in}|T^{\mu\nu}|\text{in}\rangle = \eta_1^\mu\eta_1^\nu(4\pi)^{-1}b\rho^2e^{2\rho\eta}$$
$$\times\left(\tfrac{8}{15}\rho^2/m^2-\tfrac{67}{70}\rho^4/m^4+\cdots\right) \qquad (7.17)$$

and hence

$$\langle\text{in}|T_\alpha^\alpha|\text{in}\rangle = -2(15\pi)^{-1}b\rho^2e^{2\rho\eta}\rho^2/m^2+\cdots. \qquad (7.18)$$

From Eqs. (B11)–(B13) one can show that Eqs. (7.17) and (7.18) are precisely the lowest-order (in R or b) terms of the strictly adiabatic vacuum stress and its trace, Eqs. (5.2) and (5.3) (with $\beta=0$). This is a very welcome demonstration of consistency between Ref. 18 and the present work.

On the other hand, in the limit $m\ll\rho$, Eq. (7.16) becomes (with neglect of terms of order $m^4\ln m$ and higher)

$$\langle\text{in}|T^{\mu\nu}|\text{in}\rangle = \eta_1^\mu\eta_1^\nu(6\pi)^{-1}b\rho^2e^{2\rho\eta}$$
$$= (24\pi)^{-1}R\eta_1^\mu\eta_1^\nu. \qquad (7.19)$$

This is in complete agreement (to first order in b) with the results of the general massless theory, Eqs. (6.2), (6.3), (1.4) and (1.5). It follows from remarks in Secs. V and VI that the nonvanishing trace proportional to R is associated with a *nonlocal* term of order ρ^2 in $\langle T_{\mu\nu}\rangle$, which, of course, is present for small positive m as well as for $m=0$. We would say that the expansion of the universe has created enough "matter" in the low-k modes to build up this nongeometrical contribution to the stress tensor. (This is a strange, "virtual" type of matter, which goes away again when the expansion stops. Recall that the "real" particle creation vanishes entirely when $m=0$.) Thus the $\langle T_{\mu\nu}\rangle$ of the massive theory goes smoothly over to that of the massless theory as $m\to 0$, provided that one looks at the right states.

Since the massive trace relation (5.6) is supposed to hold for all states, the second-order (in ρ) nonlocal term in $\langle T_{\mu\nu}\rangle$ must be accompanied by a new second-order term in $\langle\phi^2\rangle$ to preserve the anomaly [and, coincidentally, to remove the divergence in $\langle\phi^2\rangle$, Eq. (5.1), as $m\to 0$]. From the adiabatic point of view this also is a contribution of "matter," even though it is proportional to R. For any $m\neq 0$ we think of Eq. (5.1) as the basic vacuum c number and the rest of $\langle\phi^2\rangle$ as peculiar to the state considered. It is not possible to check Eq. (5.6) in the context of Ref. 18, since a renormalized ϕ^2 is not present there. We presume that a systematic Pauli-Villars renormalization of ϕ^2 would yield a ϕ^2 operator equivalent to ours and consistent with Eq. (5.6) for all states.

Finally, Ref. 18 [Eq. (4.3), (4.17), (4.15c), (4.25a), and (4.24)] provides the expression

$$T^{(1)\mu\nu}-\langle \mathrm{in}|T^{\mu\nu}|\mathrm{in}\rangle=\eta_1^\mu\eta_1^\nu(16\rho)^{-1}bm^4 \times\int_{-\infty}^{\infty}e^{2i(k^2+m^2)^{1/2}\eta}(k^2+m^2)^{-1}\{\sinh[\pi(k^2+m^2)^{1/2}\rho^{-1}]\}^{-1}dk\,, \tag{7.20}$$

which is an indication of the amount of particle creation in the model. It is of order m^2 for $m\ll\rho$ and of order $e^{-\pi m/\rho}$ for $m\gg\rho$. Thus there is no creation of massless particles, and the particle creation shuts off exponentially in the adiabatic limit, as expected from the discussion in Sec. III.

VIII. CONSERVATION AND UNIQUENESS OF THE RENORMALIZED STRESS TENSOR

Although the procedures leading to the renormalized vacuum expectation values (5.2) and (6.3) are very "natural" (at least in the eyes of their developers), questions have been raised concerning their uniqueness, both internally (i.e., whether a modified formulation of point separation might yield a different answer) and externally (i.e., whether completely different regularization methods, such as dimensional regularization, must necessarily give equivalent results). Also, in the case of a massive field we have not yet shown (except to lowest adiabatic order) that the renormalized stress tensor satisfies the conservation law, $T^{\nu}_{\mu,\nu}=0$. Fortunately, recent work of Wald[13-15] establishes conservation (also causality with respect to variations in the metric) for stress tensors obtained by a point-separation procedure equivalent to ours, shows that the trace anomaly follows inevitably from the requirement of conservation and the restriction of the renormalization subtractions in that procedure to polynomial functionals of the local geometry, and comes very close to settling the uniqueness question as well.

The analysis in Ref. 14 (which extends and corrects Refs. 35 and 13; see also Ref. 44) is presented for a conformally coupled massless scalar field in dimension 4, but it is actually quite general. Here we take advantage of the simplicity of two dimensions to present a very explicit version of the argument for a massive two-dimensional scalar field with $\xi=0$. The notational framework is that of Sec. II and Appendix A.

A smooth function of two nearby space-time points possesses a covariant power-series expansion about the midpoint, x_0, of the geodesic joining them:

$$F(x,x')=p_0+p_{1\bar\alpha}\sigma^{\bar\alpha}+\tfrac{1}{2}p_{2\bar\alpha\bar\beta}\sigma^{\bar\alpha}\sigma^{\bar\beta}+\tfrac{1}{8}p_{3\bar\alpha\bar\beta\bar\gamma}\sigma^{\bar\alpha}\sigma^{\bar\beta}\sigma^{\bar\gamma}+\cdots\,. \tag{8.1}$$

Here the bars distinguish midpoint tensor indices from end-point ones. We may require the p_n's to be symmetric in all their Greek indices. They are functions of x_0. Note that $F(x,x')=F(x',x)$ if and only if $p_n=0$ for all odd n. As in Ref. 4, Appendix D, the midpoint expansion can be converted to an end-point expansion by expanding p_n about x according to

$$p_0(x_0) = p_0 - \tfrac{1}{2}p_{0;\alpha}\sigma^\alpha + \tfrac{1}{8}p_{0;\alpha\beta}\sigma^\alpha\sigma^\beta - \tfrac{1}{48}p_{0;\alpha\beta\gamma}\sigma^\alpha\sigma^\beta\sigma^\gamma + \cdots , \quad (8.2*)$$

etc., the coefficients now being functions of x. (Recall that σ^α is twice as long as the distance ϵ from x to x_0 and points in the opposite direction.) Thus we obtain

$$F = p_0 + (p_{1\alpha} - \tfrac{1}{2}p_{0;\alpha})\sigma^\alpha + \tfrac{1}{2}(p_{2\alpha\beta} - p_{1\alpha;\beta} + \tfrac{1}{4}p_{0;\alpha\beta})\sigma^\alpha\sigma^\beta + \tfrac{1}{6}(p_{3\alpha\beta\gamma} - \tfrac{3}{2}p_{2\alpha\beta;\gamma} + \tfrac{3}{4}p_{1\alpha;\beta\gamma} - \tfrac{1}{8}p_{0;\alpha\beta\gamma})\sigma^\alpha\sigma^\beta\sigma^\gamma + O(\sigma^2). \quad (8.3*)$$

Let us require that F satisfies the field equation $\Box F = m^2 F$ (*) to all orders in σ ($\Box$ acting on the variable x). Substituting the expression (8.3) into the equation, we get a sequence of recursion relations beginning with

$$p_{2\mu}{}^{\mu} = m^2 p_0 - \tfrac{1}{4}\Box p_0 - p_{1\mu}{}^{;\mu} , \quad (8.4a*)$$

$$p_{3\alpha\mu}{}^{\mu} = \tfrac{1}{8}Rp_{0;\alpha} + m^2 p_{1\alpha} + \tfrac{5}{24}Rp_{1\alpha} - \tfrac{1}{4}\Box(p_{1\alpha}) - p_{2\alpha\mu}{}^{;\mu} . \quad (8.4b*)$$

These determine the traces of the successive coefficients, but p_0, p_1, and the traceless parts of p_2 and p_3 can be chosen freely. [In this calculation one uses the two-dimensional derivative-commutation rule

$$V_{\alpha;\nu\mu} = V_{\alpha;\mu\nu} + \tfrac{1}{2}R(V_\nu g_{\alpha\mu} - V_\mu g_{\alpha\nu}) \quad (8.5a*)$$

for a vector V_α, which leads to

$$S_{;\alpha\nu}{}^{\alpha} = S_{;\nu\alpha}{}^{\alpha} = (\Box S)_{;\nu} + \tfrac{1}{2}RS_{;\nu} \quad (8.5b*)$$

for a scalar S. Also, we used the first recursion relation to simplify the second.]

If we take $p_0 = 1$ and set p_1 and the traceless parts of p_2 and p_3 equal to 0, our solution becomes

$$F_{\text{min}}(x, x') = 1 + \tfrac{1}{2}m^2\sigma + O(\sigma^2) . \quad (8.6*)$$

Let us call this the *minimal* solution with $p_0 = 1$. We note that the series multiplying the logarithm in the proper-time expansion of $\tfrac{1}{2}G^{(1)}$, Eq. (2.3), is $-(2\pi)^{-1}F_{\text{min}}$. Thus $\tfrac{1}{2}G^{(1)}$ is of the form

$$H = -(2\pi)^{-1}LF_{\text{min}} + F , \quad (8.7)$$

where F is another series of the form (8.3). By a calculation like the previous one, we find that such an H is a solution of $\Box H = m^2 H$ if the coefficients of F satisfy

$$p_{2\mu}{}^{\mu} = \text{Eq. (8.4a)} + (2\pi)^{-1}(m^2 - \tfrac{1}{6}R) , \quad (8.8a*)$$

$$p_{3\alpha\mu}{}^{\mu} = \text{Eq. (8.4b)} - (48\pi)^{-1}R_{;\alpha} . \quad (8.8b*)$$

For $H = \tfrac{1}{2}G^{(1)}$ as given by Eq. (2.3), we have [cf. Eq. (4.15)]

$$p_0 = (24\pi)^{-1}[m^{-2}R + \tfrac{1}{10}m^{-4}(R^2 + 2\Box R) + O(m^{-6})] , \quad (8.9a*)$$

$$p_{2\alpha\beta} = (4\pi)^{-1}\{m^2 g_{\alpha\beta} - \tfrac{1}{12}Rg_{\alpha\beta} + \tfrac{1}{120}m^{-2}[(R^2 - \Box R)g_{\alpha\beta} + R_{;\alpha\beta}] + O(m^{-4})\} , \quad (8.9b*)$$

$$p_{1\alpha} = p_{3\alpha\beta\gamma} = 0 . \quad (8.10)$$

The odd coefficients vanish because $G^{(1)}$ is by definition symmetric in x and x'; the terms of fractional order in σ in Eq. (2.3) arise entirely from the asymmetric choice of expansion point. These equations also apply to $H = \langle\phi^2\rangle(x, x')$, since it is the same as $\tfrac{1}{2}G^{(1)}$ to finite order.

Let us also consider the *minimal* solution with such a logarithmic singularity, H_{min}, which is defined by requiring

$$p_0^{\text{min}} = p_{1\alpha}^{\text{min}} = 0 \quad (8.11)$$

and $p_{2\alpha\beta}^{\text{min}}$ to be a pure trace, and then deriving from Eqs. (8.8) that

$$p_{2\alpha\beta}^{\text{min}} = (4\pi)^{-1}(m^2 g_{\alpha\beta} - \tfrac{1}{6}Rg_{\alpha\beta}) , \quad (8.12*)$$

$$p_{3\alpha\mu}^{\text{min}\,\mu} = (48\pi)^{-1}R_{;\alpha} . \quad (8.13*)$$

Since $p_{3\alpha\beta\gamma}^{\text{min}} \neq 0$, $H_{\text{min}}(x, x')$ is not symmetric, and hence not a solution of $\Box' H = m^2 H$, since that equation would, by the analogous argument, force $p_{3\alpha\mu}{}^{\mu}$ to have the opposite sign.

In the terminology of Ref. 35, H_{min} is the *local-part* of the two-point function $\langle\phi^2\rangle(x, x')$, and

$$F_{\text{ren}} \equiv \langle\phi^2\rangle - H_{\text{min}} \quad (8.14)$$

is the *boundary-condition-dependent part*. The latter is a smooth (singularity-free) solution of the field equation in x, but not a solution in x'.

From any smooth function $F(x, x')$ one can form a "stress tensor"

$$T^F_{\mu\nu} = \lim_{x'\to x} \tfrac{1}{2}(F_{;\mu\nu'} + F_{;\nu\mu'} - F_{;\rho}{}^{\rho'} g_{\mu\nu} - m^2 F g_{\mu\nu})$$
$$\equiv \tilde{T}^F_{\mu\nu} - \tfrac{1}{2} m^2 F g_{\mu\nu} . \qquad (8.15^*)$$

(This defintion can be made covariant by inserting—at least implicitly—matrices of parallel transport.[5] Such details do not affect the limit when it is nonsingular.) From Eq. (8.15) one finds

$$T^F_{\mu\nu}{}^{;\mu} = \lim_{x'\to x} \tfrac{1}{2}\{[(\Box - m^2)F]_{;\nu'} + [(\Box' - m^2)F]_{;\nu}\} . \qquad (8.16^*)$$

Thus the covariant divergence of $T^F_{\mu\nu}$ vanishes if F is a solution of the field equation in both variables. The prototype situation is where $F(x,x') = \phi(x)\phi(x')$ for a classical field ϕ, and $T^F_{\mu\nu}$ is the corresponding classical stress tensor.

If F is given by Eq. (8.3), one finds directly from that series that

$$T^F_{\mu\nu} = \tfrac{1}{4} p_{0;\mu\nu} - \tfrac{1}{8} \Box p_0 g_{\mu\nu} - p_{2\mu\nu}$$
$$+ \tfrac{1}{2} p_{2\rho}{}^{\rho} g_{\mu\nu} - \tfrac{1}{2} m^2 p_0 g_{\mu\nu} . \qquad (8.17^*)$$

It follows that

$$T^F_{\rho}{}^{\rho} = -m^2 p_0 = -m^2 \lim_{x'\to x} F \qquad (8.18^*)$$

(the classical, nonanomalous trace identity), and that

$$T^F_{\mu\nu}{}^{;\mu} = \tfrac{1}{8}(\Box p_0)_{;\nu} + \tfrac{1}{8} R p_{0;\nu}$$
$$- p_{2\nu\mu}{}^{;\mu} + \tfrac{1}{2} p_{2\rho}{}^{\rho}{}_{;\nu} - \tfrac{1}{2} m^2 p_{0;\nu} . \qquad (8.19^*)$$

Ironically, Eq. (8.19) depends only on the *even* coefficients. However, if we now use (for the first time in the present context) the equation $\Box F = m^2 F$ in the form (8.4), we get

$$T^F_{\mu\nu}{}^{;\mu} = -m^2 p_{1\nu} - \tfrac{5}{24} R p_{1\nu} + \tfrac{1}{4}\Box(p_{1\nu}) - \tfrac{1}{2} p_{1\mu}{}^{;\mu}{}_{\nu} + p_{3\nu\mu}{}^{\mu} , \qquad (8.20a^*)$$

$$T^F_{\mu\nu}{}^{;\mu} = p_{3\nu\mu}{}^{\mu} \text{ if } p_{1\alpha} = 0 . \qquad (8.20b^*)$$

Equations (8.20) relate the divergence of $T^F_{\mu\nu}$ to the asymmetry in F.

Thus for $F = F_{\text{ren}}$ one has [see Eqs. (8.10) and (8.13)]

$$T^F_{\mu\nu}{}^{;\mu} = -(48\pi)^{-1} R_{;\nu} . \qquad (8.21^*)$$

It is easy to check that no nonsingular modification of the (polynomial, geometrical) coefficients of $H_{\min}$ (e.g., taking $p_{1\alpha} \neq 0$) would yield a conserved $T^F_{\mu\nu}$. However,

$$T^{(\text{ren})}_{\mu\nu} \equiv T^F_{\mu\nu} + (48\pi)^{-1} R g_{\mu\nu} \qquad (8.22^*)$$

is a conserved tensor; it has the expected anomalous trace

$$T^{(\text{ren})\mu}_{\mu} = -m^2 \lim_{x'\to x} F_{\text{ren}} + (24\pi)^{-1} R . \qquad (8.23^*)$$

The causality of $T^{(\text{ren})}_{\mu\nu}$ (in the sense of Ref. 13) is obvious from the construction. This completes the proof that $T^{(\text{ren})}_{\mu\nu}$ has the properties one would expect of the renormalized $\langle 0|T_{\mu\nu}|0\rangle$.

In Ref. 14 it is proposed that the vacuum expectation value of the physical, renormalized stress tensor be obtained from a renormalized two-point function F_{ren} according to (four-dimensional analogs of) Eqs. (8.14) and (8.22), with the possible addition of conserved, local, geometrical terms such as $\beta g_{\mu\nu}$. The last-minute addition of the "anomaly term" $(48\pi)^{-1} R g_{\mu\nu}$ is slightly *ad hoc*, but this approach has the tremendous advantage of providing a proof of the conservation (and trace anomaly) of the entire renormalized stress tensor, rather than just an order-by-order verification for its adiabatic series. In the present work, on the other hand, we have obtained a point-separated stress tensor from the unrenormalized two-point function and *then* subtracted off the singular terms, along with all other local terms of the same adiabatic orders. (This qualitative description of the renormalization process is accurate regardless of whether end-point or midpoint expansions are used—see Ref. 4, Sec. VII.) Let us establish the equivalence of the two prescriptions, that is, that

$$T^{(\text{ren})}_{\mu\nu}(x) = \langle 0|T_{\mu\nu}(x)|0\rangle \qquad (8.24)$$

up to possible local terms of adiabatic orders 0 and 2.

Since the formation of a stress tensor from a two-point function (singular or not) is a linear operation,

$$T^F_{\mu\nu} \equiv T_{\mu\nu}[\langle\phi^2\rangle - H_{\min}] = T_{\mu\nu}[\langle\phi^2\rangle] - T_{\mu\nu}[H_{\min}] , \qquad (8.25)$$

it suffices to show that the terms which are subtracted from $\langle 0|T_{\mu\nu}(x;\epsilon,t^\rho)|0\rangle$ $(\equiv T_{\mu\nu}[\langle\phi^2\rangle(x,x')])$ to produce $\langle 0|T_{\mu\nu}(x)|0\rangle$ (see Sec. V) are equal, modulo terms which vanish in the coincidence limit, to the point-separated stress tensor formed from $H_{\min}$, minus the last term in Eq. (8.22). Now $T_{\mu\nu}[\langle\phi^2\rangle]$ and $T_{\mu\nu}[H_{\min}]$ can be calculated from the series for $\langle\phi^2\rangle(x,x')$ and $H_{\min}$ by covariant differentiation and use of Eqs. (A5d) and (A5e), along the lines of Ref. 4, Appendix D. (In fact, this process provides the *definition* of $T_{\mu\nu}[H_{\min}]$.) The logarithmic terms of the two series are identical, and their contributions to the respective stress tensors all either vanish in the coincidence limit, or have adiabatic order 0 and hence are completely removed in the renormalization subtraction, as we desired to prove. So we can concentrate on the nonsingular terms of the series of order σ^{+1} or lower, which are given by Eqs. (8.9)–(8.12) with Eq. (8.3). Such a term makes two kinds of contributions to the stress tensor: the $\tilde{T}$ type, in which the original term has suffered two covariant differentiations, and the $m^2\phi^2$ type, in which it appears undifferentiated.

Consider first the $\tilde{T}$ terms of $T_{\mu\nu}[H_{\min}]$. Those which survive the coincidence limit come entirely from the term $\frac{1}{2}p_{2\alpha\beta}^{\min}\sigma^\alpha\sigma^\beta$, with, moreover, both derivatives acting on σ's. Their adiabatic orders d_∂ are 0 and 2 [see Eq. (8.12)], so that if identical terms appeared in $T_{\mu\nu}[\langle\phi^2\rangle]$ they *would* be subtracted during renormalization. But $\tilde{T}$ terms in $T_{\mu\nu}[\langle\phi^2\rangle]$ can arise [look at Eq. (8.3)] from the second derivative of the p_0, from the $-\frac{1}{2}p_{0;\alpha}\sigma^\alpha$ with one derivative acting on σ^α and the other on $p_{0;\alpha}$, or from the $\frac{1}{2}(p_{2\alpha\beta}+\frac{1}{4}p_{0;\alpha\beta})\sigma^\alpha\sigma^\beta$ with both derivatives acting on σ's. Of these, only the p_2 terms contain the adiabatic orders affected by renormalization. Comparing Eq. (8.9b) with Eq. (8.12), we see that, *a priori*, the relevant terms of $\tilde{T}_{\mu\nu}[\langle\phi^2\rangle]$ will differ from the smooth part of $\tilde{T}_{\mu\nu}[H_{\min}]$ by a term formed from $(48\pi)^{-1}Rg_{\mu\nu}$. However, since the latter is a pure trace, it will in fact be annihilated when the traceless combination $\tilde{T}_{\mu\nu}$ is formed. Therefore, the two prescriptions are still consistent, so far. The critical facts used in this part of the argument are that $p_{2\alpha\beta}$ is a pure trace (by construction of $H_{\min}$) and that $p_{0;\alpha\beta}$ contains no terms of adiabatic order 2, since the zeroth-order part of p_0 is covariantly constant. These features will persist in four dimensions, as long as $H_{\min}$ can be constructed as a Hadamard elementary solution (see Ref. 13) whose arbitrary first term (the quantity called w_0 in Refs. 35, 13, and 14) is a geometrical object.

Finally, we examine the contributions to the stress tensors from the term $-\frac{1}{2}m^2Fg_{\mu\nu}$ (F = smooth part of $\langle\phi^2\rangle$ or $H_{\min}$). Clearly, only p_0 can contribute in the coincidence limit. From $H_{\min}$ we get nothing, since $p_0^{\min}=0$. From $\langle\phi^2\rangle$, according to Eq. (8.9a), we have $-(48\pi)^{-1}Rg_{\mu\nu}$, and this is precisely the term which is subtracted by hand in Eq. (8.22). This completes the surprisingly cumbersome bookkeeping exercise needed to verify explicitly that the terms subtracted in our renormalization are the same as those subtracted in Wald's. (A simpler but less constructive argument is to note that only terms with $d_\partial=0$ or 2 are involved in either subtraction, and that the terms of those orders in the two answers are manifestly identical.) The close relation of the m^2F term to the trace anomaly was noted in Ref. 9, Sec. IV, where the derivation of Eq. (5.2) and its relevance to the massless case were first qualitatively described. (See also Ref. 44.)

A third renormalization prescription was advocated in Refs. 11 and 12. This, in its two-dimensional version, is to subtract from $\langle\phi^2\rangle(x,x')$ all the terms in its series of adiabatic orders 0 and 2 (which removes all divergences) and then to form the stress tensor by Eq. (8.15); one keeps track of all terms in the answer which have $d_\partial=4$ and originated from the subtracted function, and in the end one discards these terms—i.e., nullifies that part of the subtraction. This is easily shown to be equivalent to Wald's method. Comparing Eqs. (8.9)–(8.12), we see that the function subtracted from $\langle\phi^2\rangle$ by Bunch and Davies differs from $H_{\min}$ by an expression of form (8.3) with

$$p_0=(24\pi)^{-1}m^{-2}R,\quad p_{1\alpha}=0,\quad p_{2\alpha\beta}=+(48\pi)^{-1}Rg_{\alpha\beta}, \tag{8.26*}$$

plus terms of irrelevantly high order in σ. One contribution of this object to the stress tensor, the $-\frac{1}{2}m^2p_0g_{\mu\nu}$ term, provides the anomaly term in Eq. (8.22). The only other nonvanishing contribution (from $p_{0;\mu\nu}$) has $d_\partial=4$ and hence is not to be subtracted after all, according to the rules of Ref. 11. (The contribution of $p_{2\alpha\beta}$ to $\tilde{T}$ vanishes because, again, $p_{2\alpha\beta}$ is a pure trace.)

The "rules of the game" for calculating stress tensors by covariant separation of points are unambiguous. Indeed, we have just seen that various orderings of the algebraic operations involved are equivalent, provided that each method is patched up in the end, if necessary, to satisfy the conservation law. Some other possible sources of ambiguity were disposed of in Ref. 4. (The question of a possible arbitrariness in the *procedure* should be distinguished from the arbitrariness in the *result*, associated with renormalization constants. In two dimensions the latter reduces to a "cosmological" term in the stress tensor and a constant term in the renormalized ϕ^2; these can be fixed by requiring the vacuum expectation values to vanish in Minkowski space—i.e., setting α, β, and B in Secs. V and VI equal to zero, as we have more or less tacitly done at various points in our discussion.) The only assumptions which have gone into the determination of this unique renormalization prescription are as follows: (1) The renormalized stress tensor must be conserved. (2) The terms subtracted from the point-separated tensor must be local polynomial c-number functionals of the metric and curvature tensors, of adiabatic orders no higher than those in which the divergences and discontinuities of the point-separated tensor reside. The second of these is perhaps suspect, since it refers to a particular method of regularization. Wald's uniqueness and existence theorems show that any other proposed renormalized stress-tensor operator satisfying mild physical criteria (conservation, causality, and agreement with the formal expression when it makes sense) must agree with the result of the point-separation method modulo c-number terms which depend on the geometry only at the space-time point concerned. (The existence theorem is needed to guarantee that point separation itself yields a physically acceptable answer. It has been established for a large class of space-time in Ref. 16.) Unfortunately, this leaves open the question of whether different regularization schemes, each equipped with a "minimal" or "most natural" renormalization ansatz, might yield stress tensors which differ by conserved, local polynomial, c-number tensors of higher adiabatic order (containing terms such as $R^3 g_{\mu\nu}$ or $RR_{;\mu\nu}$), or even by local but nonpolynomial functionals of the geometry.[45] All we can say now is that one would not expect that to happen, since all plausible methods of calculating $\langle T_{\mu\nu}\rangle$ in a given state seem to be based on the same symmetric two-point function, or on a Green's function, effective Lagrangian, or similar object directly related to it. To our knowledge the only non-naive calculation which conflicts with the results of our approach is that of Brown and Dutton,[46] which yields no tract anomaly. Their method, applied to our two-dimensional metric (1.3), yields a $\langle 0|T_{\mu\nu}|0\rangle$ which is nonvanishing and *nonlocal* in the limit $m\to\infty$, in contrast to Eq. (5.2). [It is, in fact, the negative of our massless tensor (6.3).] That result appears to us to violate general covariance, since it is nonzero for the state built on the mode functions (7.5) for the Milne universe.

ACKNOWLEDGMENTS

Research performed primarily at King's College London with the support of the Science Research Council (U.K.) through a Studentship (T.S.B.) and Grant No. B/RG/68807 (S.M.C. and S.A.F.), and reported in part in a thesis submitted by T. S. B. to the University of London for the degree of Doctor of Philosophy.

Since our departures from King's College, this research has been supported by the National Science Foundation under Grants Nos. PHY 77-07111 (University of Wisconsin—Milwaukee), MPS 74-16311 (University of Utah), PHY 14852 (Harvard University), and PHY 77-01432 (Texas A & M University), and by the College of Science of Texas A & M University.

T. S. B. thanks his thesis supervisor, P. C. W. Davies, for encouragement and assistance. Appendix C and other improvements in the manuscript grew out of discussions with participants at the Workshop on Quantum Mechanics and Gravity at the University of British Columbia, 1977. We thank R. M. Wald for communicating his work before publication and for many helpful discussions, and M. R. Brown for discussing Refs. 45 and 46 with S. A. F. Finally, we thank the Center for Relativity Theory of the University of Texas at Austin for hospitality extended to S. M. C. and S. A. F. while this work was being completed.

APPENDIX A: DERIVATION OF THE TWO-DIMENSIONAL PROPER-TIME SERIES

We briefly indicate the origin of expressions (2.3) and (2.5). Only a few changes are needed in the four-dimensional derivation in Ref. 5.

$G^{(1)}(x,x')$ [Eq. (2.2)] is equal, except for a remainder term of transcendental order, to twice the imaginary part of Feynman's time-ordered Green's function,

$$G(x,x')\equiv\frac{i\langle\text{out, vac}|\mathcal{T}\{\phi(x),\phi(x')\}|\text{in, vac}\rangle}{\langle\text{out, vac}|\text{in, vac}\rangle}\,. \tag{A1}$$

The latter, according to a well-known formal argument, has the form

$$-\int_0^\infty ds(4\pi i s)^{-n/2}\Delta^{1/2}(x,x')\times e^{-i(m^2 s-\sigma/2s)}\Omega(x,x',s) \tag{A2*}$$

with

$$i\frac{\partial\Omega}{\partial s}+\frac{i}{s}\Omega^{;\mu}\sigma_{;\mu}=-\Delta^{-1/2}\Box(\Delta^{1/2}\Omega)+\xi R\Omega,$$

$$\Omega(x,x',0)=1\,. \tag{A3*}$$

Here n is the dimension of space-time, $\Delta=g^{-1/2}(x)\times g^{-1/2}(x')\det(-\sigma_{;\mu\nu'})$, and ξ is the conformal coupling constant in the field equation $(-\Box+m^2+\xi R)\phi=0(*)$, taken to be 0 elsewhere in this paper.

All terms in G and its second derivatives which are discontinuous as $x'\to x$ come from the first few terms in an asymptotic expansion based on the series

$$\Omega(x,x',s)=\sum_{k=0}^\infty a_k(x,x')(is)^k\,. \tag{A4}$$

The power series for $\Delta^{1/2}$, a_k, etc. given in Ref. 5 are valid for all n, if ξ is treated as an independent variable. When $n=2$ the most important ones simplify to

$$\Delta^{1/2}=1+\tfrac{1}{12}R\sigma-\tfrac{1}{24}R_{;\alpha}\sigma^\alpha\sigma+\cdots\,, \tag{A5a*}$$

$$a_1=(\tfrac{1}{6}-\xi)R-\tfrac{1}{2}(\tfrac{1}{6}-\xi)R_{;\alpha}\sigma^\alpha+\tfrac{1}{360}(R^2+3\Box R)\sigma+(\tfrac{1}{40}-\tfrac{1}{6}\xi)R_{;\alpha\beta}\sigma^\alpha\sigma^\beta+\cdots\,, \tag{A5b*}$$

$$a_2=\tfrac{1}{2}[\tfrac{1}{180}+(\tfrac{1}{6}-\xi)^2]R^2+\tfrac{1}{6}(\tfrac{1}{5}-\xi)\Box R+\cdots=\tfrac{1}{60}(R^2+2\Box R)+O(\xi)+\cdots\,, \tag{A5c*}$$

$$\sigma_{;\mu\nu}=g_{\mu\nu}-\tfrac{1}{6}R(2\sigma g_{\mu\nu}-\sigma_\mu\sigma_\nu)+\tfrac{1}{24}R_{;\alpha}\sigma^\alpha(2\sigma g_{\mu\nu}-\sigma_\mu\sigma_\nu)+\cdots\,, \tag{A5d*}$$

$$\sigma_{;\mu\nu'} = -g_{\mu\nu} - \tfrac{1}{12}R(2\sigma g_{\mu\nu} - \sigma_\mu\sigma_\nu) + \tfrac{1}{24}R_{;\alpha}\sigma^\alpha(2\sigma g_{\mu\nu} - \sigma_\mu\sigma_\nu) + \cdots . \qquad \text{(A5e*)}$$

Substitution of Eq. (A4) into Eq. (A2) with $n = 2$ leads to

$$G(x, x') = \tfrac{1}{4}\Delta^{1/2}\sum_{k=0}^{\infty} a_k \left(-\frac{\partial}{\partial m^2}\right)^k \times H_0^{(2)}[(-2m^2\sigma)^{1/2}] . \qquad \text{(A6)}$$

Inserting the power series of the Hankel function and taking the imaginary part of G, one obtains

$$\pi G^{(1)}(x, x') = \Delta^{1/2}\{-L[1 + \tfrac{1}{2}m^2\sigma - \tfrac{1}{2}a_1\sigma + O(\sigma^2)] + \tfrac{1}{2}[m^{-2}a_1 + m^{-4}a_2 + O(m^{-6})] + \tfrac{1}{4}\sigma[2m^2 - a_1 - m^{-2}a_2 + O(m^{-4})] + O(\sigma^2)\} , \qquad \text{(A.7*)}$$

where $L = \gamma + \frac{1}{2}\ln|\frac{1}{2}m^2\sigma|$ and the other quantities are given in Eqs. (A5). This yields Eq. (2.3) when $\xi = 0$.

Differentiating $G^{(1)}$ and using Eq. (A5e), one forms $\bar{T}_{\mu\nu}^{(1)}(x, x')$ and hence $T_{\mu\nu}^{(1)}(x, x')$ as given in Sec. II.

APPENDIX B: FORMULAS USED IN THE CALCULATION OF SEC. IV

A function W_k which makes the expression (3.3) satisfy Eq. (3.2) to a certain order in an adiabatic parameter may be found by substituting a formal power series for W_k (or W_k^2 or $W_k^{1/2}$) as in Eq. (3.4) and solving for the coefficients recursively. Alternately, Chakraborty's iterative procedure[32,6] yields

$$W_k^2 = \omega_k^2(1 + \epsilon_2)(1 + \epsilon_4) \cdots , \qquad \text{(B1)}$$

where

$$\epsilon_2 = -\omega_k^{-3/2}\partial_\eta(\omega_k^{-1}\partial_\eta\omega_k^{1/2}) , \qquad \text{(B2)}$$

$$\epsilon_4 = -(1 + \epsilon_2)^{-3/4}\omega_k^{-1}\partial_\eta[\omega_k^{-1}(1 + \epsilon_2)^{-1/2}\,\partial_\eta(1 + \epsilon_2)^{1/4}] .$$

In our case ($\omega_k^2 = k^2 + m^2C$) we obtain

$$\epsilon_2 = -\tfrac{1}{4}m^2\omega_k^{-4}C'' + \tfrac{5}{16}m^4\omega_k^{-6}(C')^2 , . \qquad \text{(B3a)}$$

and the part of ϵ_4 which is of order T^{-4} is

$$\epsilon_{4(4)} = \tfrac{1}{16}\{m^2\omega_k^{-6}C^{(4)} - m^4\omega_k^{-8}[7C'C^{(3)} + \tfrac{9}{2}(C'')^2] + 27m^6\omega_k^{-10}(C')^2C'' - \tfrac{135}{8}m^8\omega_k^{-12}(C')^4\}. \qquad \text{(B3b)}$$

To adiabatic order T^{-4}, therefore, we have

$$W_k = \omega_k + A^2\omega_k^{-3} + B^2\omega_k^{-5} + K^4\omega_k^{-5} + L^4\omega_k^{-7} + M^4\omega_k^{-9} + N^4\omega_k^{-11} , \qquad \text{(B4)}$$

$$A^2 \equiv -\tfrac{1}{8}m^2C'', \quad B^2 \equiv \tfrac{5}{32}m^4(C')^2, \quad K^4 \equiv \tfrac{1}{32}m^2C^{(4)},$$
$$L^4 \equiv -\tfrac{1}{128}m^4[28C'C^{(3)} + 19(C'')^2],$$
$$M^4 \equiv \tfrac{221}{256}m^6(C')^2C'' ,$$
$$N^4 \equiv -\tfrac{1105}{2048}m^8(C')^4 . \qquad \text{(B5)}$$

The exponents 2 and 4 on these coefficients indicate their adiabatic order and do not imply positivity. We also need the reciprocal of the series

$$W_k^{-1} = \omega_k^{-1} - A^2\omega_k^{-5} - B^2\omega_k^{-7} - K^4\omega_k^{-7} + (A^4 - L^4)\omega_k^{-9} + (2A^2B^2 - M^4)\omega_k^{-11} + (B^4 - N^4)\omega_k^{-13} + O(T^{-6}) . \qquad \text{(B6)}$$

From this point on it is convenient to use the logarithmic derivative of C as the basic quantity:

$$C'/C \equiv D, \quad C''/C = D' + D^2, \quad C^{(3)}/C = D'' + 3D'D + D^3 ,$$
$$C^{(4)}/C = D^{(3)} + 4D''D + 3(D')^2 + 6D'D^2 + D^4 . \qquad \text{(B7)}$$

The coordinates of the displaced points are expanded in power series as in Sec. 2(c) of Ref. 2:

$$\eta \equiv \eta(\epsilon) = \eta_0 + \epsilon\eta_1 + \tfrac{1}{2}\epsilon^2\eta_2 + \tfrac{1}{6}\epsilon^3\eta_3 + \cdots \qquad \text{(B8)}$$

and similarly for $\eta' \equiv \eta(-\epsilon)$ and for y and y'. Thus $\eta_1 \equiv t^\eta = \frac{1}{2}(t^v + t^u)$ and $y_1 \equiv t^y = \frac{1}{2}(t^v - t^u)$. The higher coefficients are given by [cf. Ref. 2, Eq. (2.21)]

$$\eta_2 = -\tfrac{1}{4}D[(t^v)^2 + (t^u)^2] ,$$
$$y_2 = -\tfrac{1}{4}D[(t^v)^2 - (t^u)^2] ,$$
$$\eta_3 = \tfrac{1}{8}\{(2D^2 - D')[(t^v)^3 + (t^u)^3] - D'[(t^v)^2t^u + (t^u)^2t^v]\} ,$$
$$y_3 = \tfrac{1}{8}\{(2D^2 - D')[(t^v)^3 - (t^u)^3] - D'[(t^v)^2t^u - (t^u)^2t^v]\} . \qquad \text{(B9)}$$

Similarly [see Ref. 2, Sec. 2(c)] one obtains

$$U_\epsilon U_{-\epsilon} = 1 + \tfrac{1}{4}\epsilon^2[(D^2 - D')(t^u)^2 - D't^ut^v] + O(\epsilon^4) . \qquad \text{(B10)}$$

Finally, from the curvature scalar (1.4) we have

$$R^2 = C^{-2}(D')^2 , \qquad \text{(B11)}$$

$$R_{,uu} = R_{,vv} = (4C)^{-1}[D^{(3)} - 3D''D - (D')^2 + 2D'D] , \qquad \text{(B12)}$$

$$R_{,uv} = R_{,vu} = (4C)^{-1}[D^{(3)} - 2D''D - (D')^2 + D'D] ,$$
$$\Box R = C^{-2}[D^{(3)} - 2D''D - (D')^2 + D'D]. \qquad \text{(B13)}$$

We also use

$$t^v = (Ct^u)^{-1}\Sigma, \quad (t^u)^{-2} = 4t_u^2/\Sigma^2,$$
$$g_{uv} = \tfrac{1}{2}C = 2t_ut_v/\Sigma . \qquad \text{(B14)}$$

APPENDIX C: POINT SEPARATION AS DISTRIBUTION THEORY

In Sec. IV the oscillatory integrals I_1 and I_2 were evaluated in terms of Bessel functions. This required either analytically continuing formulas [Ref. 33, Eqs. (3.387.6, 7)] from real to imaginary exponentials, or applying related formulas [Ref. 33, Eqs. (3.771.2, 9)] at values of the Bessel-function index for which they are not obviously valid. The real point which this uncertainty reflects is that the integrals, not being absolutely convergent, are slightly ambiguous in the first place. To guarantee that the meaning we are giving to them is the covariant, physically correct one, it is necessary to return to the very definition of the point-separated expectation values $\langle\phi^2\rangle(x,x')$ and $\langle T_{\mu\nu}\rangle(x,x')$ and examine why they are expected to be finite functions of x and x'.

A priori, one expects a quantum field to be an operator-valued distribution over test functions defined on space-time (or on space at a fixed time, depending on formulation). Consequently, $\langle\phi^2\rangle$ and $\langle T_{\mu\nu}\rangle$ should be distributions in two variables. [It seems unnecessary to review the well-known details of Fock quantization in order to establish that these expectations are fulfilled in the rigorous field theory based on the formal expansion (3.1). The special problems of the *massless* two-dimensional scalar field do not concern us here.] Point separation "works" whenever these two-point distributions actually turn out to be *functions* for x and x' with timelike or spacelike separation (cf. Refs. 13 and 14). We believe that this is always true when the expectation value is with respect to a "physically acceptable" state. In the present model it is easy to see from the naive calculations that $\langle\phi^2\rangle(x,x')$ and $\langle T_{\mu\nu}\rangle(x,x')$ make sense as distributions in one coordinate (such as ϵ) of (x,x') space; the other three coordinates may be given definite values with impunity. To make reliably the final step to a function of four coordinates, however, one must take explicit account of the distributional definition of the object being studied. We show how this can be done for $\langle T_{uu}\rangle(x,x')$ in the spacelike region.

Lemma 1.

$$G_\mu(a)\equiv\cos(\pi\mu)\pi^{-1/2}\Gamma(\mu+\tfrac{1}{2})2^\mu|a|^{-\mu}K_\mu(|a|) \qquad (C1)$$

satisfies, at $a\neq 0$,

$$-d^2G_\mu/da^2+G_\mu=G_{\mu+1}\,. \qquad (C2)$$

Proof. Direct computation, using the differential equation and recursion relation satisfied by modified Bessel functions [Ref. 33, Eqs. (8.494.1) and and (8.486.12)].

Lemma 2. In the distribution sense,

$$\int_0^\infty (x^2+1)^{\nu-1/2}\cos ax\,dx=G_\nu(a) \text{ if } a\neq 0\,. \qquad (C3)$$

That is, the integral defines a distribution on test functions (in C_0^∞, say) whose support does not include $a=0$, and this distribution is the same as the one defined by the *function* $G_\nu(a)$ on such test functions:

$$\int_0^\infty (x^2+1)^{\nu-1/2}\left\{\int_{-\infty}^\infty g(a)\cos ax\,da\right\}dx = \int_{-\infty}^\infty g(a)G_\nu(a)\,da \qquad (C4)$$

for every allowed test function, $g(a)$.

Proof. Integration by parts converts the left side of Eq. (C4) to

$$\int_0^\infty (x^2+1)^{\nu-1/2}\left\{(-1)^n(x^2+1)^{-n}\int_{-\infty}^\infty\left[\left(\frac{d^2}{da^2}-1\right)^n g(a)\right]\cos ax\,da\right\}dx$$

for any positive integer n. If n is sufficiently large, the new integrand falls off rapidly in all directions in the (x,a) plane and so the order of integrations can be interchanged:

$$(-1)^n\int_{-\infty}^\infty\left[\left(\frac{d^2}{da^2}-1\right)^n g(a)\right]\left[\int_0^\infty (x^2+1)^{\nu-n-1/2}\cos ax\,dx\right]da.$$

Moreover, the inner integral now converges in the ordinary sense and is equal (as a function) to $G_{\nu-n}(a)$ [Ref. 33, Eq. (3.771.2)]. Integrating by parts again yields

$$(-1)^n\int_{-\infty}^\infty g(a)\left(\frac{d^2}{da^2}-1\right)^n G_{\nu-n}(a)\,da\,, \qquad (C5)$$

which equals the right side of Eq. (C4) by virtue of Eq. (C2).

The role of the requirement that $g(a)$ and its derivatives vanish at $a=0$ is to make the integral (C5) meaningful despite the singularity of the Bessel function $G_\mu(a)$ at $a=0$ when μ is not negative. With extra work one could determine how the distribution $\int (x^2+1)^{\nu-1/2}\cos ax\,dx$ differs from the principal-value integral over $G_\nu(a)$ for test functions whose support does not include 0 [e.g., there might be additional terms involving $\delta(a)$ and its derivatives], but this information is not necessary for our purpose.

Lemma 3. The integral

$$\mathrm{Re}\int_{-\infty}^{\infty} \omega_k^{-1}(\omega_k+k)^2 e^{i(\omega_k\delta-k\delta')}dk \qquad \text{(C6)}$$

defines a distribution in the variables δ and δ', which, acting on test functions with support in the "spacelike" region $\delta'>|\delta|$, coincides with the *function* defined by the integral

$$2m^4C^2\frac{\delta'+\delta}{\delta'-\delta}\int_0^\infty \frac{2z^2+1}{(z^2+1)^{1/2}}\cos\gamma z\,dz, \qquad \text{(C7)}$$

$$\gamma^2 \equiv m^2C(\delta'^2-\delta^2). \qquad \text{(C8)}$$

Proof. Equation (C6) is converted to Eq. (C7) by the formal change of variable $z=\gamma^{-1}(\omega\delta-k\delta')$. According to lemma 2, the integral in Eq. (C7) equals as a distribution $2G_1(\gamma)-G_0(\gamma)$, which is $-\frac{1}{3}\gamma^2G_2(\gamma)$ by a Bessel recursion relation [Ref. 33, Eq. (8.486.10)]. This result is equivalent to Eq. (4.18), spacelike case.

However, it is not immediately obvious that such a change of variable is valid. To make the argument rigorous we must first write down a representation of the action of the distribution (C6) on a test function in terms of convergent integrals, in analogy to the proof of lemma 2. Then we make the change of variable in this well-defined integral and hope to arrive at an expression equivalent to Eq. (C7) smeared with a test function of γ. (This will give a distribution in one of the δ's with the other fixed, and *a fortiori* a distribution in both.)

The neatest way to do this is to arrange things so that the "well-defined integral" turns out to have the structure of Eq. (C5) with $a=\gamma$, so that Eq. (C7) is obtained immediately by integration by parts. In other words, we want to make the change of variable in an integral whose integrand differs from that of Eq. (C6) [hence (C7)] only by a convergence factor, $(z^2+1)^{-n}$. This will be accomplished if we treat Eq. (C6) along the lines of the proof of lemma 2, but use in the role of d/da a differential operator in δ and δ' which is equivalent to $\partial/\partial\gamma$ with $\zeta\equiv(\delta'+\delta)/(\delta'-\delta)$ held fixed. Calculating the Jacobian matrix $\partial(\delta,\delta')/\partial(\gamma,\zeta)$, we find that

$$\frac{\partial}{\partial\gamma}=\frac{\delta}{\gamma}\frac{\partial}{\partial\delta}+\frac{\delta'}{\gamma}\frac{\partial}{\partial\delta'}, \qquad \text{(C9)}$$

and from this verify that $\partial^2/\partial\gamma^2-1$ acts on $\exp(i\omega_k\delta-ik\delta')$ to produce a factor $-(z^2+1)$, as expected and required.

It is clear that the change of variables could be conducted similarly in the timelike region $(\delta>|\delta'|)$. This completes the justification of Eq. (4.18), since a distribution in δ and δ' obviously induces one in ϵ. Equations (4.12) can be treated in the same way.

[1]P. C. W. Davies, S. A. Fulling, and W. G. Unruh, Phys. Rev. D 13, 2720 (1976).

[2]P. C. W. Davies and S. A. Fulling, Proc. R. Soc. London A354, 59 (1977).

[3]P. C. W. Davies and W. G. Unruh, Proc. R. Soc. London A356, 259 (1977).

[4]P. C. W. Davies, S. A. Fulling, S. M. Christensen, and T. S. Bunch, Ann. Phys. (N.Y.) 109, 108 (1977).

[5]S. M. Christensen, Phys. Rev. D 14, 2490 (1976). See also S. M. Christensen, Ph.D. dissertation, University of Texas at Austin, 1975 (unpublished).

[6]L. Parker and S. A. Fulling, Phys. Rev. D 9, 341 (1974). See also L. Parker, Ph.D. thesis, Harvard University, 1966 (unpublished).

[7]P. C. W. Davies, Proc. R. Soc. London A354, 529 (1977).

[8]L. S. Brown and J. P. Cassidy, Phys. Rev. D 16, 1712 (1977).

[9]S. M. Christensen and S. A. Fulling, Phys. Rev. D 15, 2088 (1977).

[10]S. M. Christensen, Phys. Rev. D 17, 946 (1978).

[11]T. S. Bunch and P. C. W. Davies, Proc. R. Soc. London A360, 117 (1978).

[12]T. S. Bunch and P. C. W. Davies, J. Phys. A 11, 1315 (1978).

[13]R. M. Wald, Commun. Math. Phys. 54, 1 (1977).

[14]R. M. Wald, Phys. Rev. D 17, 1477 (1978).

[15]R. M. Wald, Ann. Phys. (N.Y.) 110, 472 (1978).

[16]S. A. Fulling, M. Sweeny, and R. M. Wald, Commun. Math. Phys. (to be published).

[17]T. S. Bunch, Ph.D. thesis, University of London (King's College), 1977 (unpublished).

[18]C. Bernard and A. Duncan, Ann. Phys. (N.Y.) 107, 201 (1977).

[19]C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973), inside front cover.

[20]S. A. Fulling, J. Phys. A 10, 917 (1977).

[21]B. S. DeWitt, Phys. Rep. 19C, 295 (1975).

[22]H. Rumpf, Phys. Lett. 61B, 272 (1976).

[23]H. Rumpf, Nuovo Cimento 35B, 321 (1976).

[24]H. Rumpf, dissertation, University of Vienna, 1977 (unpublished).

[25]H. Rumpf and H. K. Urbantke (unpublished).

[26]L. Parker, Phys. Rev. 183, 1057 (1969).

[27]A. A. Grib and S. G. Mamayev, Yad. Fiz. 10, 1276 (1969) [Sov. J. Nucl. Phys. 10, 722 (1969)].

[28]Ya. B. Zel'dovich and A. A. Starobinsky, Zh. Eksp. Teor. Fiz. 61, 2161 (1971) [Sov. Phys.—JETP 34, 1159 (1971)].

[29]M. Castagnino, A. Verbeure, and R. A. Weder, Nuovo Cimento 26B, 396 (1975).

[30]J. D. Cole, *Perturbation Methods in Applied Mathematics* (Blaisdell, Waltham, Mass., 1968), pp. 4–7, 79 and 80.
[31]L. H. Ford and L. Parker, Phys. Rev. D 16, 245 (1977).
[32]B. Chakraborty, J. Math. Phys. 14, 188 (1973).
[33]I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic, New York, 1965).
[34]W. G. Unruh, Phys. Rev. D 14, 870 (1976).
[35]S. L. Adler, J. Lieberman, and Y. J. Ng, Ann. Phys. (N.Y.) 106, 279 (1977).
[36]L. S. Brown, Phys. Rev. D 15, 1469 (1977).
[37]T. S. Bunch, J. Phys. A 11, 603 (1978).
[38]A. S. Wightman, in *Cargèse Lectures in Theoretical Physics: High Energy Electromagnetic Interactions and Field Theory*, edited by M. Lévy (Gordon and Breach, New York, 1967), pp. 204–207.
[39]A. di Sessa, J. Math. Phys. 15, 1892 (1974).
[40]C. M. Sommerfield, Ann. Phys. (N.Y.) 84, 285 (1974).
[41]D. G. Boulware, Phys. Rev. D 11, 1404 (1975).
[42]Ya. B. Zel'dovich and L. P. Pitaevsky, Commun. Math. Phys. 23, 185 (1971).
[43]A. Vilenkin (unpublished).
[44]S. L. Adler and J. Lieberman (unpublished).
[45]M. R. Brown (unpublished).
[46]M. R. Brown and C. R. Dutton, preceding paper, Phys. Rev. D 18, 4422 (1978).