Unitarity of interacting fields in curved spacetime

John L. Friedman,* Nicolas J. Papastamatiou,† and Jonathan Z. Simon‡
Department of Physics, University of Wisconsin, Milwaukee, Wisconsin 53201
(Received 29 June 1992)

On globally hyperbolic spacetimes, each foliation by spacelike hypersurfaces corresponds to a Hamiltonian description of field theory, and unitarity follows formally from the Hermiticity of the Hamiltonian. For a renormalizable theory, unitarity at each order in perturbation theory follows from the corresponding Hermiticity of each term in the time-ordered product of interaction Hamiltonians. For more general spacetimes, one can still use the path integral to obtain a generalized Lehmann-Symanzik-Zimmermann reduction formula for S-matrix elements and the corresponding perturbative expansion. Unitarity imposes an infinite set of identities on the scattering amplitudes, which are the generalizations of the flat-spacetime Cutkosky rules. We find these explicitly to $O(\lambda^3)$ in a $\lambda\varphi^4$ theory, and show how to find the relations to any order. For globally hyperbolic spacetimes the unitarity identities are satisfied [at least to $O(\lambda^3)$] because of a single property of the configuration-space propagator that reflects the causal structure of the spacetime.

PACS number(s): 03.70.+k, 04.60.+n

I. INTRODUCTION

There is now a substantial literature on the behavior of quantum fields in curved spacetime [1]. Because of both conceptual and calculational difficulties, most investigations have been restricted to the behavior of free fields. There has been more limited progress in the study of interacting quantum fields in curved spacetime, focused on the renormalizability of theories [2,3]. Although unitarity of interacting quantum fields in curved spacetime has been assumed, the scattering identities which follow from unitarity (analogues of the Cutkosky rules) have apparently not yet been explicitly obtained for general curved backgrounds.

A treatment of the unitarity relations that govern interacting fields in curved spacetime differs in two key ways from the standard discussion of flat-space unitarity. First, in flat spacetime the number of S-matrix elements contributing to each unitarity relation is substantially restricted by conservation of energy and momentum. In a generic curved spacetime no analogous restrictions persist. Second, flat-space quantum field theory is ordinarily treated in a momentum-space context, and the unitarity relations are obtained from analyticity properties of the momentum-space propagator. In a curved spacetime one has no natural way to define a global momentum space, and the unitarity relations are obtained in a configuration-space context. In some ways this turns out to be an advantage: the role that causal structure plays in enforcing unitarity is clarified by the configuration-space form of the propagator.

The plan of the paper is as follows. In Sec. II we review the status of unitarity for free field theories in curved spacetime, emphasizing the relation between free-field unitarity and the preservation of the inner product on the one-particle Hilbert space (for scalar fields this is the Klein-Gordon inner product). We present the argument in a way that does not rely on global hyperbolicity of the spacetime. As we discuss in the following paper, free fields are unitary on a class of spacetimes with closed timelike curves, primarily because the inner product is conserved on these spacetimes as well [4].

Next, in Sec. III, we consider unitarity for interacting field theories on globally hyperbolic spacetimes. When the corresponding free-field theory is unitary on a spacetime, perturbative unitarity of the interacting theory follows from the existence at each order of a self-adjoint Hamiltonian. For a perturbatively defined theory shown to be renormalizable to a particular order, we show that the theory is unitary to that order.

Even for spacetimes which are not globally hyperbolic, one can define a perturbative expansion for S-matrix elements. In Sec. IV we obtain a generalized Lehmann-Symanzik-Zimmermann (LSZ) reduction formula [5] for general spacetimes by means of the path integral (following a suggestion of J. B. Hartle) and show how Feynman rules can be defined. As in Sec. II, this derivation does not rely on global hyperbolicity of the spacetime. Unitarity of the S matrix entails a set of identities which must hold order by order in perturbation theory. In flat spacetime, these identities are the configuration-space counterparts of the momentum-space Cutkosky rules [6]. We present a diagrammatic method for deriving the unitarity identities to any order in perturbation theory and show that they can be expressed as identities involving only the Feynman propagator. Thus, for any given spacetime, one can address the issue of unitarity by examining properties of the Feynman propagator. We explicitly obtain these identities for a $\lambda\varphi^4$ theory to order λ^3, and also show that they are satisfied for all propagators of the form

$$i\Delta_F(x,y)=\theta(x^0-y^0)D(x,y)+\theta(y^0-x^0)\overline{D(x,y)}\ ,$$

*Electronic address: friedman@thales.phys. uwm.edu
†Electronic address: njp@csd4.csd.uwm.edu
‡Electronic address: jsimon@csd4.csd.uwm.edu

where $D(x,y)$ is the two-point Wightman function. We use lower-case Latin letters as spacetime indices and a signature of $(-+++)$.

II. UNITARITY FOR FREE QUANTUM FIELDS IN CURVED SPACETIME

A. The field representation

The presentation of free-field unitarity given below is written in terms of a path-integral formalism intended to make sense regardless of whether the spacetime can be foliated by spacelike hypersurfaces. In globally hyperbolic spacetimes, free-field unitarity is well understood [7], and readers interested only in the results for interacting fields in a globally hyperbolic setting can skip to the next section. The formalism introduced here will be used in the companion paper on the loss of unitarity for interacting fields on spacetimes with closed timelike curves [4].

Let M, g_{ab} be an asymptotically flat spacetime that is static in the past and future. (This could be relaxed to stationary as $t \to \pm\infty$.) Instead of introducing asymptotic scattering states (limits as $t \to \pm\infty$), we choose particular spacelike hypersurfaces $\Sigma_{\rm in}$ in the past static region and $\Sigma_{\rm out}$ in the future static region, and define in states and out states on these two hypersurfaces.

A static metric with timelike Killing vector t^a can be written in the form

$$g_{ab} = -e^{+2\nu}\nabla_a t \nabla_b t + h_{ab} , \tag{1}$$

where $e^{-\nu} = (-\nabla_a t \nabla^a t)^{1/2}$ is the lapse function on a t=const hypersurface $\Sigma, t^a = g^{ab}\nabla_b t$, and h_{ab} is the three-metric on Σ. (That is, the components of h_{ab} orthogonal to Σ vanish, and the part of h_{ab} tangent to Σ is its three-metric.) The Klein-Gordon operator

$$K = -\nabla_a \nabla^a + m^2 + \xi R , \tag{2}$$

with mass m and curvature coupling ξ, has the form

$$K = e^{-2\nu}(\pounds_t^2 + \Omega^2) , \tag{3}$$

with Ω the operator satisfying $\Omega f = i\pounds_t f$, or $\Omega f = \omega f$ for a solution f with positive frequency ω. If we denote by D_a the covariant derivative operator with respect to the three-metric h_{ab} on Σ, then

$$\Omega = (-e^{-\nu}D^a e^{\nu} D_a + m^2 + \xi R)^{1/2} . \tag{4}$$

We will also need the Klein-Gordon inner product on the space of complex solutions to $Kf = 0$:[1]

$$\langle f | g \rangle = \frac{1}{i}\int_\Sigma d\Sigma_a (\bar{f} \overleftrightarrow{\nabla}^a g) . \tag{5}$$

[1]A positive definite inner product is given by

$$\langle f | g \rangle = \frac{1}{i}\int_\Sigma d\Sigma_a (\overline{f_+} \overleftrightarrow{\nabla}^a g_+ - \overline{f_-}\overleftrightarrow{\nabla}^a g_-) ,$$

where f_+ and f_- are the positive- and negative-frequency parts of f. This would eliminate the negative sign in Eq. (7) but would be inappropriate for the discussion in Sec. II B below.

A path integral over configuration space describes the evolution of states in the field representation, in which state vectors on a hypersurface Σ are functions $\Psi(\varphi)$ of field configurations $\varphi(x)$ on Σ. (Detailed descriptions of the field representation in flat spacetime are given by Glimm and Jaffe and by Symanzik [8]. The field representation in curved spacetime has been used by Freese *et al.* and Floreannini *et al.* [9]) The vacuum state corresponding to t^a is

$$|0\rangle = \Psi_0(\varphi) = N \exp\left[-\int_\Sigma d\Sigma \tfrac{1}{2}\varphi\Omega\varphi\right] , \tag{6}$$

where N is a normalization constant. The Schrödinger field operator $\Phi_S(x)$ is multiplication by $\varphi(x)$, and the corresponding momentum operator is $\pi_S(x) = -i\delta/\delta\varphi(x)$.

To relate the usual Fock space states to functions $\Psi(\varphi)$, we need to find, in terms of $\Phi_S(x)$ and $\pi_S(x)$, the creation operator $a^\dagger(f)$, corresponding to a positive-frequency solution f. In the Fock representation one ordinarily uses Heisenberg operators $\Phi_H(x)$ and $\pi_H(x)$, where $\pi_H = h^{1/2}e^{-\nu}\pounds_t\Phi_H$ is the canonically conjugate momentum and h is the determinant of the three-metric. For the Heisenberg operators,

$$a^\dagger(f) = \langle -\bar{f} | \Phi_H \rangle . \tag{7}$$

By choosing $\Phi_H(x)$ and $\pi_H(x)$ to coincide on Σ with $\Phi_S(x)$ and $\pi_S(x)$, we can rewrite $a^\dagger(f)$ in terms of the Schrödinger field operators,

$$a^\dagger(f) = i\int_\Sigma d\Sigma_a f \overleftrightarrow{\nabla}^a \Phi_H = \int_\Sigma d^3x (h^{1/2}e^{-\nu}\Omega f \Phi_S - i f \pi_S) , \tag{8}$$

or in the field representation,

$$a^\dagger(f)\Psi(\varphi) = \int_\Sigma d^3x [h^{1/2}e^{-\nu}\Omega f \varphi(x) - f\delta/\delta\varphi(x)]\Psi(\varphi) . \tag{9}$$

The usual Fock space is thereby isomorphic to the Fock space of the field representation. The n-particle state of the usual Fock space,

$$a_i^\dagger \cdots a_j^\dagger |0\rangle , \tag{10}$$

is the symmetrized tensor product of n one-particle states,

$$f_{(i} \otimes \cdots \otimes f_{j)} , \tag{11}$$

and the corresponding state of the field representation is

$$a^\dagger(f_i) \cdots a^\dagger(f_j)\Psi_0(\varphi) . \tag{12}$$

In terms of creation and annihilation operators, the operator algebra takes the form

$$\begin{aligned} [a(f), a^\dagger(g)] &= \langle f | g \rangle , \\ [a(f), a(g)] &= 0 , \\ [a^\dagger(f), a^\dagger(g)] &= 0 . \end{aligned} \tag{13}$$

Let $\mathcal{F}^{\text{in}}$ and $\mathcal{F}^{\text{out}}$ be the Fock spaces associated with Σ_{in} and Σ_{out}, respectively. Because the final static geometry is not isometric to the initial geometry, the Fock spaces $\mathcal{F}^{\text{in}}$ and $\mathcal{F}^{\text{out}}$ are not naturally isomorphic, and the Schrödinger field operators are distinct. There is no natural way to associate a field operator $\Phi(x_1)$ at a point x_1 of Σ_{in} with a field operator $\Phi(x_2)$ at a point x_2 of Σ_{out}.

The S-matrix $\mathcal{S}$ describing the evolution from $\mathcal{F}^{\text{in}}$ to $\mathcal{F}^{\text{out}}$ can be formally expressed as a path integral:

$$ {}_{\text{out}}\langle\beta|\mathcal{S}|\alpha\rangle_{\text{in}}=\int \mathcal{D}\varphi\,\overline{\Psi}_\beta(\varphi|_{\Sigma_{\text{out}}})e^{iS[\varphi]}\Psi_\alpha(\varphi|_{\Sigma_{\text{in}}})\ , \tag{14}$$

where $S[\varphi]$ is the action. For the free-field theory $S[\varphi]=S_0[\varphi]$, where

$$S_0[\varphi]=-\tfrac{1}{2}\int dz\,\varphi K\varphi\ . \tag{15}$$

As usual, when K appears in a path integral, it is understood as $K+i\epsilon$, in the limit $\epsilon\to 0+$. A Heisenberg field operator $\Phi_H(x)$ satisfying the Klein-Gordon equation, agreeing with the Schrödinger field operator Φ on Σ_{in}, and acting on $\mathcal{F}^{\text{in}}$, can be defined by writing

$$ {}_{\text{out}}\langle\beta|\mathcal{S}\Phi_H|\alpha\rangle_{\text{in}}=\int \mathcal{D}\varphi\,\overline{\Psi}_\beta(\varphi|_{\Sigma_{\text{out}}})\varphi(x)e^{iS}\Psi_\alpha(\varphi|_{\Sigma_{\text{in}}})\ . \tag{16}$$

When one can foliate the spacetime with spacelike hypersurfaces Σ_t (with $\Sigma_{\text{in}}=\Sigma_{t_1}$, $\Sigma_{\text{out}}=\Sigma_{t_2}$), the Heisenberg field operator $\Phi_H(x), x\in\Sigma_t$ can be related to the Schrödinger field operator $\Phi_S(x)$ in the usual manner:

$$\Phi_H(x)=\mathcal{U}_{tt_1}^{-1}\Phi_S(x)\mathcal{U}_{tt_1}\ . \tag{17}$$

Here $\mathcal{U}_{tt_1}$ is the time evolution operator mapping a state Ψ_α on Σ_{in} to a state $\mathcal{U}_{tt_1}\Psi_\alpha$ on Σ_t,

$$\mathcal{U}_{tt_1}\Psi_\alpha=\int_{(t,t_1]}\mathcal{D}\varphi\, e^{iS(t,t_1)}\Psi_\alpha(\varphi|_{\Sigma_{\text{in}}})\ . \tag{18}$$

Then from Eq. (16) one obtains Eq. (17) as

$$\begin{aligned}\langle\beta|\mathcal{S}\Phi_H(x)|\alpha\rangle&=\int \mathcal{D}\varphi\overline{\Psi}_\beta(\varphi|_{\Sigma_{\text{out}}})\varphi(x)e^{iS}\Psi_\alpha(\varphi|_{\Sigma_{\text{in}}})\\&=\int_{(t_2,t]}\mathcal{D}\varphi\,\overline{\Psi}_\beta(\varphi|_{\Sigma_{\text{out}}})e^{iS(t_2,t)}\varphi(x)\int_{(t,t_1]}\mathcal{D}\varphi\, e^{iS(t,t_1)}\Psi_\alpha(\varphi|_{\Sigma_{\text{in}}})\\&=\langle\beta|\mathcal{U}_{t_2t}\Phi_S(x)\mathcal{U}_{tt_1}|\alpha\rangle\\&=\langle\beta|\mathcal{S}\mathcal{U}_{tt_1}^{-1}\Phi_S(x)\mathcal{U}_{tt_1}|\alpha\rangle\ .\end{aligned} \tag{19}$$

We have used in the last equality the relation $\mathcal{S}=\mathcal{U}_{t_2t_1}$.

In a region where there are closed timelike curves, no Schrödinger field operator exists, but $\Phi_H(x)$ is defined by the path integral of Eq. (16), or, equivalently, by the value of Φ and π for $x\in\Sigma_{\text{in}}$ together with the Klein-Gordon equation

$$K\Phi_H=0\ . \tag{20}$$

Equation (20) is formally implied by Eq. (16):

$$\begin{aligned}K\int\mathcal{D}\varphi\,\overline{\Psi}_\beta e^{iS}\varphi(x)\Psi_\alpha&=\int\mathcal{D}\varphi\,\overline{\Psi}_\beta K\varphi(x)\exp(-i\tfrac{1}{2}\int dz\,\varphi K\varphi)\Psi_\alpha\\&=i\int\mathcal{D}\varphi\,\overline{\Psi}_\beta\frac{\delta}{\delta\varphi(x)}\exp(-i\tfrac{1}{2}\int dz\,\varphi K\varphi)\Psi_\alpha\\&=i\int\mathcal{D}\varphi\frac{\delta}{\delta\varphi(x)}[\overline{\Psi}_\beta e^{iS}\Psi_\alpha]\\&=0\ .\end{aligned} \tag{21}$$

B. Unitarity

For free fields, when the S matrix exists, unitarity follows from the fact that the classical time evolution preserves the Klein-Gordon inner product,

$$\langle Uf|Ug\rangle_{\text{out}}=\langle f|g\rangle_{\text{in}}\ , \tag{22}$$

where f and g are solutions to the Klein-Gordon equation in the past, Uf and Ug the corresponding solutions evolved in the future,

$$\langle f|g\rangle_{\text{in}}:=-i\int_{\Sigma_{\text{in}}}\overline{f}\overleftrightarrow{\nabla}^a g\,dS_a \tag{23}$$

and

$$\langle Uf|Ug\rangle_{\text{out}}:=-i\int_{\Sigma_{\text{out}}}\overline{Uf}\overleftrightarrow{\nabla}^a Ug\,dS_a\ . \tag{24}$$

The S matrix $\mathcal{S}$ is defined up to an overall constant by the requirement that it map Φ and π on the initial hypersurface Σ_{in} to Φ and π on the final hypersurface Σ_{out}. The statement that Φ_H on Σ_{out} is evolved in time from Φ_H on Σ_{in} by the Klein-Gordon equation can be expressed in the manner $\langle f|\Phi_H\rangle=\langle Uf|\Phi_H\rangle$ or

$$\mathcal{S}\Phi_S(f)\mathcal{S}^{-1}=\Phi_S(Uf) \tag{25}$$

for any complex solution f in the past.

Define a^{out} by

$$\mathcal{S} a^{\text{in}}(f)\mathcal{S}^{-1}=a^{\text{out}}(Uf) \ . \tag{26}$$

In other words, a^{out} is an annihilation operator that kills the image $\mathcal{S}|0\rangle$ of the in vacuum. If we define $a^{\dagger\,\text{out}}$ by

$$\Phi_S(x)=a^{\text{out}}(x)+a^{\dagger\,\text{out}}(x) \ , \tag{27}$$

Hermiticity of Φ implies $a^{\dagger\,\text{out}}=(a^{\text{out}})^{\dagger}$. Then preservation of the Klein-Gordon inner product implies that the commutation relations of a and $a^{\dagger}$ are preserved by $\mathcal{S}$.

$$[a^{\text{out}}(Uf),a^{\dagger\,\text{out}}(Ug)]=\mathcal{S}[a^{\text{in}}(f),a^{\text{in}\,\dagger}(g)]\mathcal{S}^{-1}=\mathcal{S}\langle f|g\rangle\mathcal{S}^{-1}=\langle f|g\rangle=\langle Uf|Ug\rangle \ . \tag{28}$$

Finally, if we fix the overall normalization of $\mathcal{S}$ by

$${}_{\text{in}}\langle 0|\mathcal{S}^{\dagger}\mathcal{S}|0\rangle_{\text{in}}=1 \tag{29}$$

(see caveats below), then

$$\mathcal{S}^{\dagger}\mathcal{S}=1 \ . \tag{30}$$

To obtain this last equation, write an n-particle state in the form

$$|i\cdots j\rangle := a_i^{\text{in}\,\dagger}\cdots a_j^{\text{in}\,\dagger}|0\rangle_{\text{in}} \ . \tag{31}$$

We have

$$\mathcal{S}|i\cdots j\rangle=\mathcal{S}a_i^{\text{in}\,\dagger}\mathcal{S}^{-1}\cdots\mathcal{S}a_j^{\text{in}\,\dagger}\mathcal{S}^{-1}\mathcal{S}|0\rangle_{\text{in}}=a_i^{\dagger\,\text{out}}\cdots a_j^{\dagger\,\text{out}}\mathcal{S}|0\rangle_{\text{in}} \ . \tag{32}$$

Then the inner product ${}_{\text{in}}\langle k\cdots l|i\cdots j\rangle_{\text{in}}$ of two in states is preserved by $\mathcal{S}^{\dagger}\mathcal{S}$:

$$\begin{aligned}
{}_{\text{in}}\langle k\cdots l|\mathcal{S}^{\dagger}\mathcal{S}|i\cdots j\rangle_{\text{in}} &= {}_{\text{in}}\langle 0|\mathcal{S}^{\dagger}a_k^{\text{out}}\cdots a_l^{\text{out}}a_i^{\dagger\,\text{out}}\cdots a_j^{\dagger\,\text{out}}\mathcal{S}|0\rangle_{\text{in}} \\
&= {}_{\text{in}}\langle 0|0\rangle_{\text{in}}\sum_{\sigma}[\alpha_{\sigma(k)}^{\text{out}},a_i^{\dagger\,\text{out}}]\cdots[a_{\sigma(l)}^{\text{out}},a_j^{\dagger\,\text{out}}] \\
&= {}_{\text{in}}\langle 0|0\rangle_{\text{in}}\sum_{\sigma}[a_{\sigma(k)}^{\text{in}},a_i^{\text{in}\,\dagger}]\cdots[a_{\sigma(l)}^{\text{in}},a_j^{\text{in}\,\dagger}] \\
&= {}_{\text{in}}\langle k\cdots l|i\cdots j\rangle_{\text{in}} \ ,
\end{aligned} \tag{33}$$

where Eqs. (28) and (29) were used to obtain the next-to-last equality and σ is a permutation of the indices $i\cdots j$.

Two caveats: First, if the image of the in vacuum does not have finite norm in $\mathcal{F}^{\text{out}}$, Eq. (29) will fail to define an S matrix (see, e.g., DeWitt or Wald [7]). Second, unitarity requires not simply that the image of the in vacuum be finite, but that its norm be preserved. One ordinarily obtains the S matrix only up to an overall normalization from the Bogoliubov coefficients that describe the evolution of solutions to the Klein-Gordon equation; unitarity is then used to fix the normalization. Requiring that the S matrix be unitary up to an over all normalization is sufficient to guarantee that the algebra of observables generated by $\mathcal{S}\Phi\mathcal{S}^{-1}$ and $\mathcal{S}\pi\mathcal{S}^{-1}$ be unitarily equivalent to that generated by Φ and π. In principle, however, by solving the functional Schrödinger equation, one could decide whether the norm of the vacuum was preserved.

Note that the discussion in this section does not rely on specific properties of the scalar field, but, rather, on the existence of an inner product. The discussion can be straightforwardly generalized to two-component spinors, Dirac spinors, and electromagnetic fields (antisymmetric tensors), with corresponding inner products

$$\langle\mu|\nu\rangle=\int_{\Sigma}d\Sigma_{A\dot{B}}\bar{\mu}^{\dot{B}}\nu^{A}=\int_{\Sigma}d\Sigma_a\sigma^{a}{}_{A\dot{B}}\bar{\mu}^{\dot{B}}\nu^{A} \ ,$$

$$\langle\chi|\psi\rangle=\int_{\Sigma}d\Sigma_a\bar{\chi}\gamma^{a}\psi \ ,$$

and

$$\langle F_2|F_1\rangle=-i\int_{\Sigma}d\Sigma_a(\bar{A}_{2b}F_1^{ab}-\bar{F}_2^{ab}A_{1b}) \ .$$

In the last relation, A_a is any asymptotically regular vector field such that $F_{ab}=\nabla_a A_b-\nabla_b A_a$, and the product is gauge invariant.

III. UNITARITY FOR INTERACTING FIELDS IN CURVED SPACETIME

A formal proof of unitarity for interacting fields in globally hyperbolic spacetimes can be given by a straightforward extension of the corresponding proof for flat spacetime. As in the case of free fields we require that the spacetime be static in future and past regions containing the spacelike hypersurfaces Σ_{in} and Σ_{out}, respectively. Let t be a universal time, i.e., a scalar for which the t =

const surfaces foliate the spacetime, and for which t increases to the future. In each static region with timelike Killing vector t^a, successive hypersurfaces are chosen to be Lie derived by t^a, i.e., $t^a \nabla_a t = 1$.

Consider an action of the form

$$S[\varphi]=S_0[\varphi]-\int dz\, V(\varphi(z)) , \tag{34}$$

where S_0 is the free action of Eq. (15). The defining equation for the elements of the S matrix, Eq. (14), can be written in the form

$$_{\rm out}\langle\beta|\mathcal{S}|\alpha\rangle_{\rm in}=\int \mathcal{D}\varphi\, \overline{\Psi}_\beta \left[\sum_{n=0}^{\infty}\frac{(-i)^n}{n!}\left[\int dz\, V(\varphi(z))\right]^n\right] e^{iS_0}\Psi_\alpha . \tag{35}$$

For spacetimes of the type described above, we have [cf. Eq. (16)]

$$\int \mathcal{D}\varphi\, \overline{\Psi}_\beta \varphi(x_1)\cdots\varphi(x_n)e^{iS_0}\Psi_\alpha = {}_{\rm out}\langle\beta|\mathcal{S}^{(0)}T\Phi(x_1)\cdots\Phi(x_n)|\alpha\rangle_{\rm in} , \tag{36}$$

where $\mathcal{S}^{(0)}$ is the free-field S matrix and the evolution is governed by the free action S_0 (interaction picture). Then one can formally describe the S matrix as

$$\mathcal{S}=T\exp\left[-i\int dz\, V(\varphi)\right] = T\exp\left[-i\int dt\, H_I(t)\right] , \tag{37}$$

where

$$H_I(t)=\int_{\Sigma_t} d\Sigma\, e^{\nu}V(\varphi) , \tag{38}$$

with $e^{\nu}=(-\nabla_a t\nabla^a t)^{-1/2}$ the lapse function.

For Hermitian $V(\varphi)$, one can use Eq. (37) to establish the unitarity of $\mathcal{S}$ to each order of perturbation theory, provided that the theory is renormalizable. It is widely believed that theories which are renormalizable in flat space are renormalizable in curved spacetimes as well, and renormalizability has been shown for spacetimes that can be analytically continued to spaces with Riemannian metrics [3]. If a given theory is renormalizable, one can systematically construct a finite theory by first regularizing and then introducing a set of counterterms for which matrix elements remain finite when the regularization is removed. Let μ denote the regulator in some particular scheme for which the regularization is removed when $\mu\to 0$. We assume that the following hold for a renormalizable theory.

(1) There is a family $H(\mu)$ of regularized Hamiltonians which are Hermitian on the free-field Fock space for $\mu\neq 0$.

(2) There is a finite set of regularized counterterms $Z_j H_j(\mu)$, each of which is Hermitian on the free-field Fock space for $\mu\neq 0$.

(3) At each order m of perturbation theory, one can choose renormalization constants $Z(\mu,m)$ so that the corresponding renormalized Hamiltonian $H_r^{(m)} = H(\mu)+Z_j H_j$ has, as a weak limit,

$$\lim_{\mu\to 0} H_r(\mu) , \tag{39}$$

an Hermitian operator on the free-field Fock space. After renormalization, H_I appearing in Eq. (37) is replaced by $H_{I,r}=H-H_{0,r}$, where $H_{0,r}$ is the free Hamiltonian expressed in terms of renormalized quantities. Similarly, the mth-order approximation to the S matrix,

$$\mathcal{S}^{(m)}=I+i\int dt\, H_{I,r}^{(m)}(t)+\cdots+\frac{i^m}{m!}T\int dt_1\cdots dt_m H_{I,r}^{(m)}(t_1)\cdots H_{I,r}^{(m)}(t_m) , \tag{40}$$

has a weak limit obeying the relations

$$\left[\lim_{\mu\to 0}\mathcal{S}^{(m)}\right]^{\dagger}=\lim_{\mu\to 0}[\mathcal{S}^{(m)\dagger}] , \quad \left[\lim_{\mu\to 0}\mathcal{S}^{(m)}\right]^{-1}=\lim_{\mu\to 0}[\mathcal{S}^{(m)-1}] . \tag{41}$$

Then for $\mu\neq 0$, the Hermiticity of $H_r(\mu)$ implies that $\mathcal{S}^{(m)}(\mu)$ is unitary. Finally, the existence of a well-defined limit, as the regularization parameter is allowed to approach zero, expressed by conditions (41) implies that $\mathcal{S}^{(m)}$ is unitary.

Because this proof of perturbative unitarity requires a globally defined Hamiltonian, it holds only in spacetimes that allow a 3+1 decomposition. On the other hand, Eq. (35) holds independently of this restriction. As we will show in the next section, it can be used to derive a diagrammatic expansion of the S matrix which allows a direct verification of the unitarity conditions to any order in perturbation theory.

IV. UNITARITY AND FEYNMAN DIAGRAMMATIC EXPANSION

A. Background

For a unitary S matrix, $\mathcal{S}^{\dagger}\mathcal{S}=1$. When the matrix elements of $\mathcal{S}^{\dagger}\mathcal{S}$ are expanded in powers of the interaction coupling constant, the consequence is an infinite number of unitary scattering identities. If the background spacetime is Minkowski and Feynman diagrams are evaluated in momentum space, these identities are known as Cutkosky rules [6]. The specification of the flat-spacetime Cutkosky rules is straightforward because of the conservation of energy and momentum. For an interaction with strength λ,

$$\mathcal{S}=\mathcal{S}^{(0)}+\lambda\mathcal{S}^{(1)}+\lambda^2\mathcal{S}^{(2)}+\cdots \tag{42}$$

we have

$$\langle\alpha'|\alpha\rangle=\langle\alpha'|(\mathcal{S}^{(0)}+\lambda\mathcal{S}^{(1)}+\lambda^2\mathcal{S}^{(2)}+\cdots)^\dagger(\mathcal{S}^{(0)}+\lambda\mathcal{S}^{(1)}+\lambda^2\mathcal{S}^{(2)}+\cdots)|\alpha\rangle \tag{43}$$

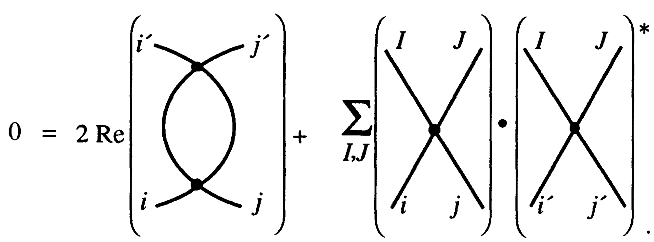
and so, order by order,

$$\begin{aligned}
&\lambda^0{:}\langle\alpha'|\alpha\rangle=\langle\alpha'|\mathcal{S}^{(0)\dagger}\mathcal{S}^{(0)}|\alpha\rangle\ ,\\
&\lambda^1{:}0=\langle\alpha'|\mathcal{S}^{(0)\dagger}\mathcal{S}^{(1)}|\alpha\rangle+\langle\alpha'|\mathcal{S}^{(1)\dagger}\mathcal{S}^{(0)}|\alpha\rangle\ ,\\
&\lambda^2{:}0=\langle\alpha'|\mathcal{S}^{(0)\dagger}\mathcal{S}^{(2)}|\alpha\rangle+\langle\alpha'|\mathcal{S}^{(1)\dagger}\mathcal{S}^{(1)}|\alpha\rangle+\langle\alpha'|\mathcal{S}^{(2)\dagger}\mathcal{S}^{(0)}|\alpha\rangle,\ \text{etc}\ .
\end{aligned} \tag{44}$$

One example of a Cutkosky rule comes from the matrix element $\langle i',j'|\mathcal{S}^\dagger\mathcal{S}|i,j\rangle$, the overlap of $\mathcal{S}^\dagger\mathcal{S}$ between two-particle states. For the $\lambda\varphi^4$ theory the first nontrivial equation is second order in λ, and we have

$$\begin{aligned}
0&=\lambda^2[\langle i',j'|\mathcal{S}^{(0)\dagger}\mathcal{S}^{(2)}|i,j\rangle+\langle i',j'|\mathcal{S}^{(1)\dagger}\mathcal{S}^{(1)}|i,j\rangle+\langle i',j'|\mathcal{S}^{(2)\dagger}\mathcal{S}^{(0)}|i,j\rangle]\\
&=\lambda^2[2\,\mathrm{Re}\langle i',j'|\mathcal{S}^{(2)}|i,j\rangle+\sum_{I\le J}\langle I,J|\mathcal{S}^{(1)}|i',j'\rangle^*\langle I,J|\mathcal{S}^{(1)}|i,j\rangle]\ ,
\end{aligned} \tag{45}$$

where we have used energy conservation, the scalar nature of the field, and the fact that the free S-matrix $\mathcal{S}^{(0)}$ is trivial in flat spacetime. The pair of indices (I,J) label a complete, orthonormal set of two-particle states, arising from the sum over the complete, orthonormal set of all n-particle states inserted between $\mathcal{S}^{(1)\dagger}$ and $\mathcal{S}^{(1)}$. In terms of Feynman diagrams,[2] this equation takes the form



(46)

The second term of Eq. (46) can be viewed as having cut the first term into two pieces and sewn them back together by summing over physical states.

In curved spacetime, the generalization of this infinite tower of unitarity identities must be made somewhat carefully. Energy and momentum conservation, which limits the number of contributing Feynman diagrams in Eq. (46), is no longer available. The triviality of the free field S-matrix $\mathcal{S}^{(0)}$ is also lost, since even free fields undergo particle creation in general curved spacetimes. We cannot (in general) transform the propagator and Green's functions to momentum space and must be content with configuration-space path integrals.

The loss of conservation of energy and momentum means that more diagrams contribute nontrivially, but this does not make the generalized unitarity identities significantly more difficult; we simply must include all possible terms in deriving them. The nontriviality of the free-field S-matrix $\mathcal{S}^{(0)}$ can also be overcome. In flat spacetime, the middle term of the first line of Eq. (45) is simplified by the insertion of a complete set of states between $\mathcal{S}^\dagger$ and $\mathcal{S}$. In general curved spacetimes, all terms can similarly be simplified by inserting a complete set of states. For a unitary S matrix,

$${}_{\rm in}\langle\alpha'|\alpha\rangle_{\rm in}=\sum_A {}_{\rm in}\langle\alpha'|\mathcal{S}^\dagger|A\rangle_{\rm out}\,{}_{\rm out}\langle A|\mathcal{S}|\alpha\rangle_{\rm in}\ , \tag{47}$$

and, in particular,

[2] As described later, our Feynman diagrams represent the S matrix, not the T matrix, where $\mathcal{S}=I+i\mathcal{T}\delta(\Delta p^\mu)$. With this convention, Eq. (46) and related Feynman diagrammatic equations involve the real, rather than the imaginary, part of the amplitude. The dot product is defined in Sec. IV D.

$$\lambda^0:{}_{\rm in}\langle \alpha'|\alpha\rangle_{\rm in}=\sum_A {}_{\rm in}\langle \alpha'|\mathcal{S}^{(0)\dagger}|A\rangle_{\rm out\,out}\langle A|\mathcal{S}^{(0)}|\alpha\rangle_{\rm in}\ ,$$

$$\lambda^1:0=\sum_A [{}_{\rm in}\langle \alpha'|\mathcal{S}^{(1)\dagger}|A\rangle_{\rm out\,out}\langle A|\mathcal{S}^{(0)}|\alpha\rangle_{\rm in}+{}_{\rm in}\langle \alpha'|\mathcal{S}^{(0)\dagger}|A\rangle_{\rm out\,out}\langle A|\mathcal{S}^{(1)}|\alpha\rangle_{\rm in}]\ ,$$

$$\lambda^2:0=\sum_A [{}_{\rm in}\langle \alpha'|\mathcal{S}^{(2)\dagger}|A\rangle_{\rm out\,out}\langle A|\mathcal{S}^{(0)}|\alpha\rangle_{\rm in}+{}_{\rm in}\langle \alpha'|\mathcal{S}^{(1)\dagger}|A\rangle_{\rm out\,out}\langle A|\mathcal{S}^{(1)}|\alpha\rangle_{\rm in} + {}_{\rm in}\langle \alpha'|\mathcal{S}^{(0)\dagger}|A\rangle_{\rm out\,out}\langle A|\mathcal{S}^{(2)}|\alpha\rangle_{\rm in}],\ \eta\ . \qquad (48)$$

Equation (47) has a simple interpretation. The overlap between an initial state $|\alpha\rangle$ with another initial state $|\alpha'\rangle$ is given by the joint amplitude to go from the state $|\alpha\rangle$ to some final basis state $\langle A|$ via the scattering process $\mathcal{S}$, and then backward in time (via the same scattering process) to the state $|\alpha'\rangle$, summed over all possible intermediate orthonormal states. If we view $\mathcal{S}|\alpha\rangle$ as the evolution in time of state $|\alpha\rangle$, as in the Schrödinger picture, then the function of unitarity is to enforce the normalization of states over time. Of course the particular basis we choose when inserting the set of states $\{|A\rangle\}$ does not affect the results, so long as the basis is complete, but, as we will see, there is a class of bases that considerably simplify the formalism.

B. LSZ reduction

To compute transition amplitudes from past states to future states we use a generalized Lehmann-Symanzik-Zimmermann (LSZ) reduction [5]. This generalization relies only on the Feynman path integral and does not require the spacetime to be foliated with spacelike hypersurfaces.

We present here a variant of the standard path-integral reduction of S-matrix elements to products of Feynman propagators. If one defines the free Feynman propagator by a path integral in the manner

$$i\Delta_F(x,y) := \int \mathcal{D}\varphi\, \overline{\Psi_0^{\rm out}}\varphi(x)\varphi(y)\exp\left[-i\tfrac{1}{2}\int dz\, \varphi K\varphi\right]\Psi_0^{\rm in}\ , \qquad (49)$$

then the steps in the reduction make no mention of the causal structure of the spacetime. If one can foliate the spacetime, Eq. (49) is equivalent to a vacuum expectation value of time-ordered Heisenberg field operators. There are, however, spacetimes with closed timelike curves in which the Cauchy problem is well defined [10]. The free-field Feynman propagator is also well defined, and the reduction given below can be regarded as a path-integral derivation of the usual Feynman rules for the scattering of interacting fields on this more general class of spacetimes.

Let

$$S=\int_{\Sigma_{\rm in}}^{\Sigma_{\rm out}} dz\,\mathcal{L}\ , \qquad (50)$$

where

$$\mathcal{L}=-\tfrac{1}{2}\varphi K\varphi - V(\varphi) \qquad (51)$$

and denote by S_0 the free-field action. We write the general matrix element ${}_{\rm out}\langle k\cdots|\mathcal{S}|i\cdots\rangle_{\rm in}$ in the form

$${}_{\rm out}\langle k\cdots|\mathcal{S}|i\cdots\rangle_{\rm in}=\int\mathcal{D}\varphi\,\overline{a_k^{\dagger\,\rm out}\cdots\Psi_0^{\rm out}}e^{iS}a_i^{\rm in\,\dagger}\cdots\Psi_0^{\rm in}$$
$$=\int_{\Sigma_{\rm out}} d\Sigma_a \overline{f_k^{\rm out}}(x)\overleftrightarrow{\nabla}^a\cdots\int_{\Sigma_{\rm in}} d\Sigma_b f_i^{\rm in}(y)\overleftrightarrow{\nabla}^b\cdots\int\mathcal{D}\varphi\,\overline{\Psi_0^{\rm out}}\varphi(x)\cdots\varphi(y)\cdots e^{iS}\Psi_0^{\rm in}\ . \qquad (52)$$

This is the generalization of the usual LSZ reduction formula.

To replace the product of field operators by products of propagators, we introduce, as usual, a generating functional

$$W(J)=\int\mathcal{D}\varphi\,\overline{\Psi_0^{\rm out}}\exp\left[i\int dx[-\tfrac{1}{2}\varphi K\varphi - V(\varphi)+J\varphi]\right]\Psi_0^{\rm in}\ , \qquad (53)$$

for which

$$\int\mathcal{D}\varphi\,\overline{\Psi_0^{\rm out}}\varphi(x)\cdots\varphi(y)e^{iS}\Psi_0^{\rm in}=\frac{1}{i}\frac{\delta}{\delta J(x)}\cdots\frac{1}{i}\frac{\delta}{\delta J(y)}W(J)\bigg|_{J=0}\ . \qquad (54)$$

Then Eq. (53) can be rewritten as

$$W(J)=\exp\left[-i\int dx\ V\left[\frac{1}{i}\frac{\delta}{\delta J}\right]\right]\int\mathcal{D}\varphi\,\overline{\Psi_0^{\rm out}}\exp\left[i\int dx[-\tfrac{1}{2}\varphi K\varphi+J\varphi]\right]\Psi_0^{\rm in}$$
$$=N\exp\left[-i\int dx\ V\left[\frac{1}{i}\frac{\delta}{\delta J}\right]\right]\exp\left[-\frac{i}{2}\int dx\,dy\,J(x)\Delta_F(x,y)J(y)\right]\ , \qquad (55)$$

after performing the Gaussian integration. Finally, up to an overall normalization,

$$
{}_{\rm out}\langle k \cdots |\mathcal{S}| i \cdots \rangle_{\rm in} = \int_{\Sigma_{\rm out}} d\Sigma_a \overline{f_k^{\rm out}}(x) \overleftrightarrow{\nabla}^a \cdots \int_{\Sigma_{\rm in}} d\Sigma_b f_i^{\rm in}(y) \overleftrightarrow{\nabla}^b \cdots
\times \frac{1}{i}\frac{\delta}{\delta J(x)} \cdots \frac{1}{i}\frac{\delta}{\delta J(y)} \exp\left[-i \int dx\, V\left[\frac{1}{i}\frac{\delta}{\delta J}\right]\right] \exp\left[-\frac{i}{2}\int dx\, dy\, J(x)\Delta_F(x,y)J(y)\right]\Bigg|_{J=0} . \tag{56}
$$

This is the desired expression for $\mathcal{S}$ in terms of Δ_F. From it we will obtain the usual Feynman rules in configuration space.

The description of scattering for interacting fields is greatly simplified by a choice of basis of the final Fock space $\mathcal{F}^{\rm out}$ for which the free-field scattering matrix is trivial. Our spacetime has static epochs in the past and future, and in the past we can choose a basis $\{|A\rangle_{\rm in}\}$ for $\mathcal{F}^{\rm in}$ associated with the timelike Killing vector in the past. That is, each past basis vector $|A\rangle_{\rm in}$ is an n-particle state of the form given in Eqs. (10) and (11), the symmetrized tensor product of n solutions $F_I, \ldots, F_J$, to the free Klein-Gordon equation, where each F_I has a positive frequency with respect to the past Killing vector, and $\langle I|J\rangle = \delta_{IJ}$.

The future basis vectors, however, will *not* be n-particle states with respect to the future Killing vector. Instead, by choosing the basis $\{|A\rangle_{\rm out}\}$, where $|A\rangle_{\rm out} = \mathcal{S}^{(0)}|A\rangle_{\rm in}$, we have ${}_{\rm out}\langle A'|\mathcal{S}^{(0)}|A\rangle_{\rm in} = \delta_{AA'}$ for unitary $\mathcal{S}^{(0)}$. Note that each basis vector of $\mathcal{F}^{\rm out}$ has the form $|A\rangle_{\rm out} = |I \cdots J\rangle_{\rm out} = a_I^{\dagger\,\rm out} \cdots a_J^{\dagger\,\rm out}|0\rangle_{\rm out}$, with $a_I^{\dagger\,\rm out} = \mathcal{S}^{(0)} a_I^{\dagger\,\rm in} \mathcal{S}^{(0)-1}$ as in Sec. II. In particular, we will choose the future "vacuum" state in Eq. (49) such that $|0\rangle_{\rm out} = \mathcal{S}^{(0)}|0\rangle_{\rm in}$.

We shall refer to a state $|i_1 \cdots i_n\rangle_{\rm out}$ as an n-particle state, because this terminology corresponds to the Feynman diagrams of the interaction picture: diagrams with n future external lines can then be called amplitudes for n-particle out states. A clean break is thereby made between particle creation due to the curved background and particle creation due to interactions. All remaining calculations, Feynman rules, and Feynman diagrams will use this expedient basis and terminology. To recapitulate, a future "n-particle state" has n particles with respect to the vacuum $\mathcal{S}^{(0)}|0\rangle_{\rm in}$, not to the natural vacuum of the future Killing vector.

Our notation uses lower-case Latin letters from the middle of the alphabet to indicate one-particle states created by arbitrary (not necessarily positive frequency) solutions to the Klein-Gordon equation [e.g., $|i\rangle = a_i^\dagger|0\rangle$, $a_i = a(f_i)$]. Upper-case Latin letters from the middle of the alphabet correspond to one-particle *basis* states [e.g., $|I\rangle = a_I^\dagger|0\rangle$, $a_I = a(F_I)$, $\langle I|J\rangle = \delta_{IJ}$]. Lower-case letters from the beginning of the Greek alphabet correspond to multiparticle states composed of arbitrary one-particle states (e.g., $|\alpha\rangle = a_i^\dagger \cdots a_j^\dagger|0\rangle$). Upper-case letters from the beginning of the alphabet correspond to multiparticle *basis* states (e.g., $|A\rangle = a_I^\dagger \cdots a_J^\dagger|0\rangle$).

C. Feynman diagrams

For $V(\varphi) = (\lambda/4!)\varphi^4$, the configuration-space Feynman rules are as follows.

(i) Draw all topologically distinct diagrams with $l+n$ external lines, labeled by

l in states $f_1, \ldots, f_l$,

n out states $f_{l+1}, \ldots, f_{l+n}$,

m internal vertices $y_1, \ldots, y_m$.

The contribution of a diagram to the S-matrix element $\langle f_{l+1} \cdots f_{l+n}|\mathcal{S}|f_1 \cdots f_1\rangle$ is computed as follows.

(ii) To each line from y_i to y_j assign a factor $i\Delta_F(y_i, y_j)$:

$$y_i \bullet\!\!-\!\!-\!\!\bullet\; y_j \quad (\text{internal}) = i\Delta_F(y_i, y_j) . \tag{57}$$

(iii) To each external line f_j from y_i assign a factor $f_j(y_i)$ for an in state and $\overline{f_j}(y_i)$ for an out state:

$$f_j \bullet\!\!-\!\!-\!\!\bullet\; y_i \quad (\text{initial external}) = f_j(y_i) ,$$
$$y_i \bullet\!\!-\!\!-\!\!\bullet\; f_j \quad (\text{final external}) = \overline{f}_j(y_i) . \tag{58}$$

(iv) To each vertex y_i assign $-i\lambda \int dy_i$:

$$y_i \times = -i\lambda \int dy_i . \tag{59}$$

(v) Assign a symmetry factor, the reciprocal of the number of permutations of internal lines that leave the diagram unchanged for fixed vertices (just as for flat spacetime).

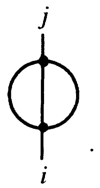
(vi) To each line unattached to a vertex, from an in state f_i to an out state f_j, assign the overlap value $\langle j|i\rangle = -i\int_\Sigma d\Sigma_a \overline{f_j}(x)\overleftrightarrow{\nabla}^a f_i(x)$:

$$f_i \bullet\!\!-\!\!-\!\!\bullet\; f_j \quad (\text{initial} \rightarrow \text{final}) = \langle j|i\rangle . \tag{60}$$

For example, the second order connected one-particle to one-particle transition amplitude for the $\lambda\varphi^4$ theory

$$\langle j|\lambda^2\mathcal{S}^{(2)}|i\rangle = \frac{-i\lambda^2}{6}\int dx\, dy\, \overline{f_j}(y)[\Delta_F(x,y)]^3 f_i(x) \tag{61}$$

is represented by



(62)

The usual convention is to use Feynman diagrams to represent the T matrix, not the S matrix, where $\mathcal{S}=I+i\mathcal{T}\delta(\Delta p^{\mu})$. We cannot follow this convention, since energy-momentum is not conserved, and we shall use Feynman diagrams to represent the elements of the S matrix (not the T matrix). Some equations will therefore disagree with the T-matrix convention by factors of i.

D. Identities governing unitary scattering

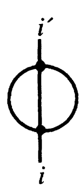
The generalization of the infinite tower of unitarity identities to curved spacetime can be neatly described by the Feynman diagrams just outlined. The process of diagrammatically constructing the scattering unitarity identities in configuration space is quite straightforward. Several steps are similar to those used in constructing the Cutkosky rules in flat spacetime, in the momentum-space representation. We will obtain the identities corresponding to Eqs. (48) at order λ, λ^2, and λ^3. First, however, as a representative example, we begin with a detailed discussion of a particular diagram corresponding to one-particle–one-particle scattering at order λ^2.

The one-particle–one-particle scattering identity is

$$0=\sum_{A}\lambda^2[\langle i'|\mathcal{S}^{(2)\dagger}|A\rangle\langle A|\mathcal{S}^{(0)}|i\rangle$$
$$+\langle i'|\mathcal{S}^{(1)\dagger}|A\rangle\langle A|\mathcal{S}^{(1)}|i\rangle$$
$$+\langle i'|\mathcal{S}^{(0)\dagger}|A\rangle\langle A|\mathcal{S}^{(2)}|i\rangle]\ . \tag{63}$$

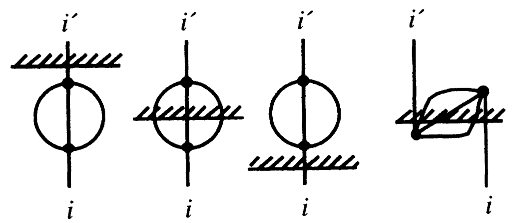
Very few of the infinite number of intermediate basis states $|A\rangle$ contribute nontrivially at this order, in fact only the one-, three-, and five-particle states.

The first step in the diagrammatic approach is to construct a (fictitious) connected, configuration-space Feynman diagram that scatters the state $|\alpha\rangle$ (e.g., $|i\rangle$) into the other state $|\alpha'\rangle$ (e.g., $|i'\rangle$). Using the example of the second order, connected, one-particle–one-particle scattering process, the relevant diagram is



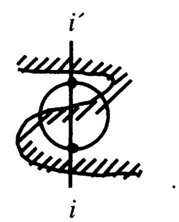
(64)

If there is more than one diagram contributing to the scattering at a given order (as in scattering processes described below), then the procedure must be applied to all diagrams. Draw a continuous curve that cuts the diagram into two parts, one part containing the initial state $|\alpha\rangle$ and the other containing the initial state $|\alpha'\rangle$, that does not pass through any interaction points. The cut will correspond to the final surface in the future. Any nontrivial scattering process from the initial state $|\alpha\rangle$ to the final surface (and from the initial state $|\alpha'\rangle$ to the final surface) will contribute. In our example there are four topologically distinct ways of making this cut that will contribute:



(65)

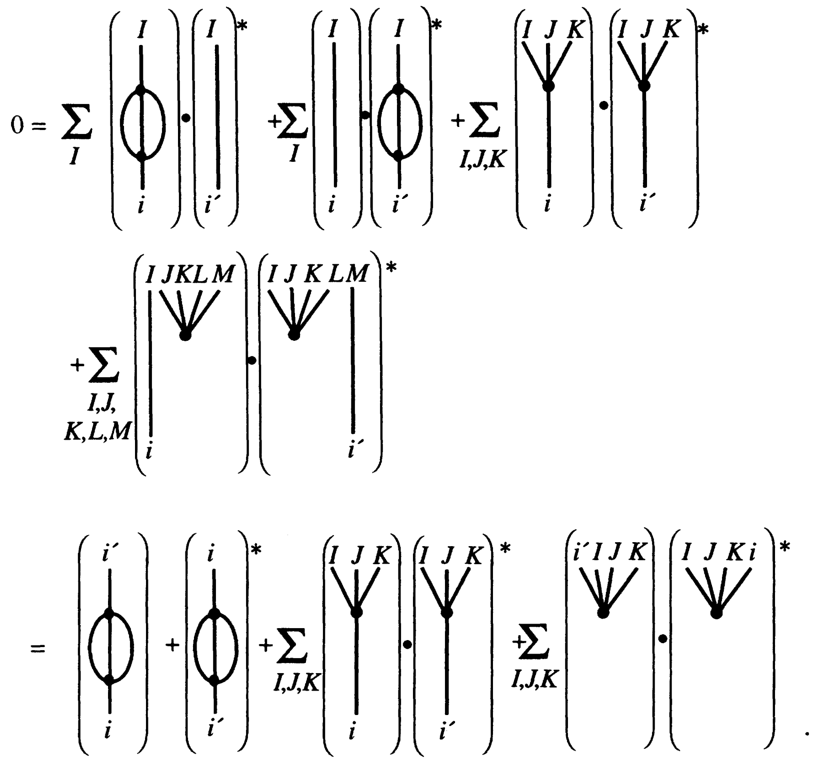
The last individual diagram of (65) can also be represented by



(66)

The final step is to break up each of the cut diagrams into two factors. The first factor is the Feynman diagram made of the bottom part of the cut diagram with the initial state $|\alpha\rangle$ scattering to an intermediate state of n particles on the final (cutting) surface. The second factor is the remaining top part of the cut diagram, with the initial state $|\alpha'\rangle$ scattering to the same intermediate state on the final surface, turned upside down and complex conjugated. All n-particle states in the final surface are summed over (with a factor of $1/n!$ to avoid overcounting identical states), and the two diagrams are multiplied together with a combinatoric factor p. The combinatoric factor is necessary when either of the diagrams gives multiple contributions, and it is calculated as follows: for a precut diagram with a symmetry factor of q [from Feynman rule (v)] and for cut diagrams with symmetry factors of r and s, respectively, then $p=rs/q$. Diagram multiplication can be combined with multiplication by the combinatoric factor p into a single operation, denoted by a dot product.

When this cutting process is applied to all diagrams and all terms are summed over, the resulting expression must vanish by Eq. (48). The four diagrams of (65) lead immediately to the second-order unitary scattering identity:



(67)

or, more explicitly,

$$0=\frac{\lambda^2}{6}\int dx\,dy\,f_i(x)\overline{f_{i'}}(y)\{-[i\Delta_F(x,y)]^3-[\overline{i\Delta_F(x,y)}]^3+D(x,y)^3+\overline{D(x,y)^3}\}\ , \tag{68}$$

where the two-point Wightman function is given by

$$D(x,y)=\sum_I F_I(x)\overline{F_I}(y)\ . \tag{69}$$

Note that, in general, terms contribute to Eqs. (63) and (67) that would not appear in the flat-spacetime case. This includes the amplitude for one particle to decay into three particles of the same species, and for four particles to appear out of the vacuum. In curved spacetime we lose both energy-momentum conservation and stability of the vacuum, allowing such amplitudes to contribute nontrivially. In this relation, as in the following unitarity relations, the effect of changing the diagram by switching any external legs from past to future is simply to change the corresponding f's to $\bar{f}$'s (in fact, the f's need not be positive frequency at any rate).

The smeared unitarity identity (68) will hold for arbitrary particle states if the pointwise unitarity relation

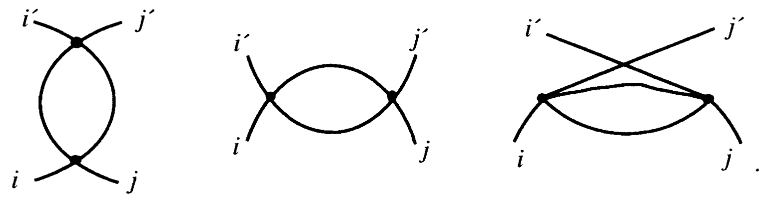
$$0=-[i\Delta_F(x,y)]^3-[\overline{i\Delta_F(x,y)}]^3+D(x,y)^3+\overline{D(x,y)}^3 \tag{70}$$

also holds. We believe that any unitary, renormalizable theory which satisfies the smeared unitarity relation (68) also satisfies the pointwise unitarity identity (70) [though Eq. (68) certainly does not imply Eq. (70)]. The pointwise relation follows trivially if $i\Delta_F(x,y)=\theta(x^0-y^0)D(x,y)+\theta(y^0-x^0)\overline{D}(x,y)$, but for spacetimes which cannot be foliated with spacelike hypersurfaces, Eqs. (68) and (70) are nontrivial restrictions on the theory.

The identities following from the second order, connected, two-particle–two-particle scattering process are derived analogously. The two-particle–two-particle scattering identity is

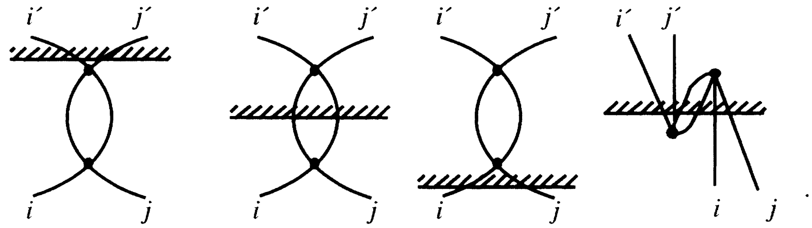
$$0=\sum_A \lambda^2[\langle i',j'|\mathcal{S}^{(2)\dagger}|A\rangle\langle A|\mathcal{S}^{(0)}|i,j\rangle+\langle i',j'|\mathcal{S}^{(1)\dagger}|A\rangle\langle A|\mathcal{S}^{(1)}|i,j\rangle+\langle i',j'|\mathcal{S}^{(0)\dagger}|A\rangle\langle A|\mathcal{S}^{(2)}|i,j\rangle]\ . \tag{71}$$

Only the two-, four-, and six-particle states contribute to the intermediate states $|A\rangle$ at this order. There are three (fictitious) connected, precut diagrams contributing:



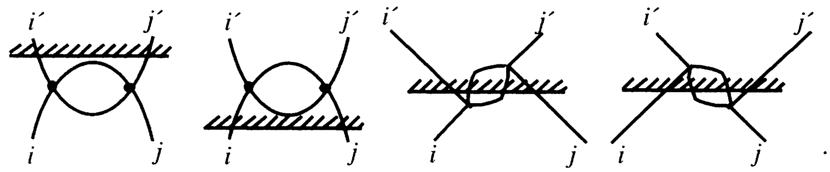
(72)

The four contributing cuts from the first individual diagram of (72) are very similar to the four contributing cuts of (65):



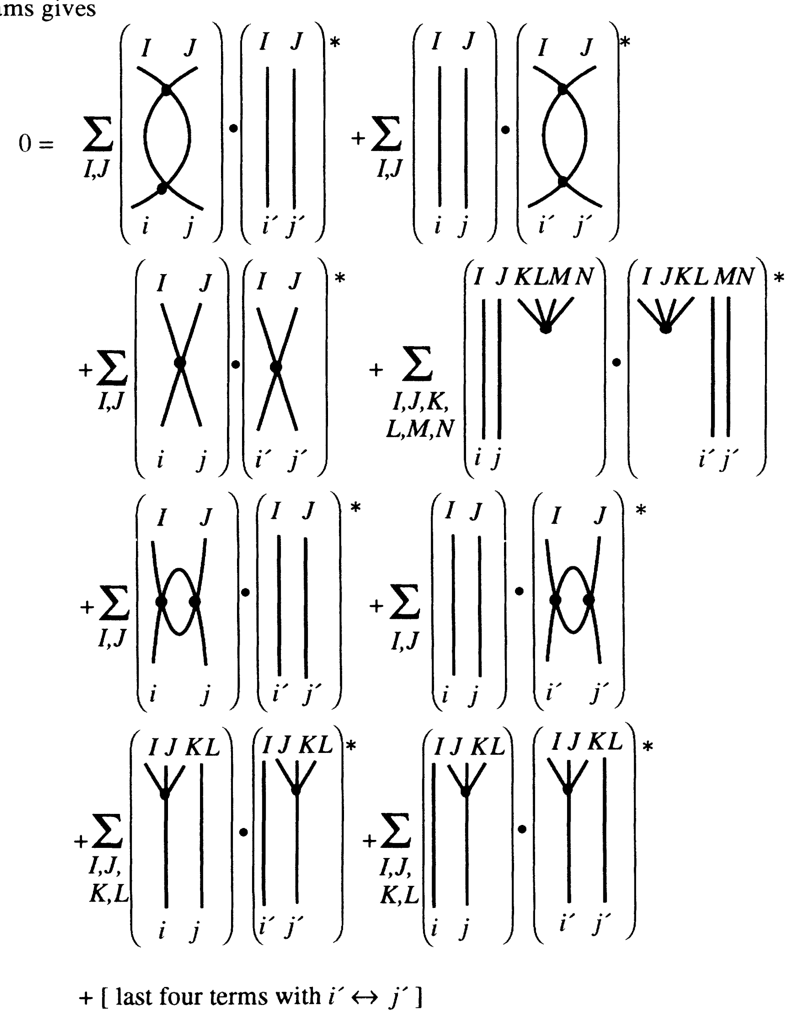
(73)

The four contributing cuts from the second individual diagram of (72) are given by



(74)

The four contributing cuts from the third individual diagram of (72) can be obtained by exchanging i' and j' in the diagram of (74). When summed over the intermediate states allowed on the final surfaces, the sum of the resulting 12 diagrams gives



+ [last four terms with $i' \leftrightarrow j'$] (75)

or, the smeared unitarity identity

$$0=\frac{\lambda^2}{2}\int dx\,dy\{-[i\Delta_F(x,y)]^2-[\overline{i\Delta_F(x,y)}]^2+[D(x,y)]^2+[\overline{D(x,y)}]^2\}$$
$$\times\{f_i(x)f_j(x)\overline{f_{i'}}(y)\overline{f_{j'}}(y)+f_i(x)f_j(y)\overline{f_{i'}}(x)\overline{f_{j'}}(y)+f_i(x)f_j(y)\overline{f_{i'}}(y)\overline{f_{j'}}(x)\}\ . \tag{76}$$

This will hold for arbitrary particle states if the pointwise relation

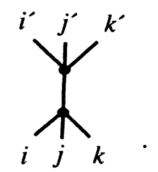
$$0=-[i\Delta_F(x,y)]^2-[\overline{i\Delta_F(x,y)}]^2+[D(x,y)]^2+[\overline{D(x,y)}]^2 \tag{77}$$

is satisfied.

The three-particle–three-particle scattering identity is

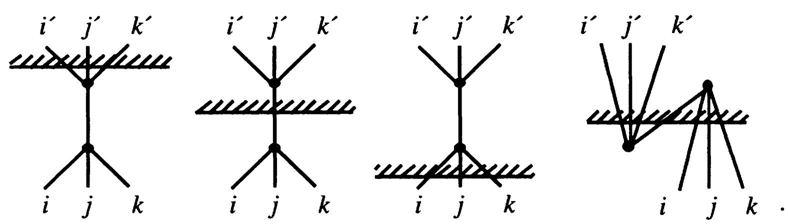
$$0=\sum_A \lambda^2[\langle i',j',k'|\mathcal{S}^{(2)\dagger}|A\rangle\langle A|\mathcal{S}^{(0)}|i,j,k\rangle+\langle i',j',k'|\mathcal{S}^{(1)\dagger}|A\rangle\langle A|\mathcal{S}^{(1)}|i,j,k\rangle+\langle i',j',k'|\mathcal{S}^{(0)\dagger}|A\rangle\langle A|\mathcal{S}^{(2)}|i,j,k\rangle]\ .$$

Only the one-, three-, five-, and seven-particle states contribute to the intermediate states $|A\rangle$ at second order. There are several contributing diagrams, of which only one will be cut apart here:



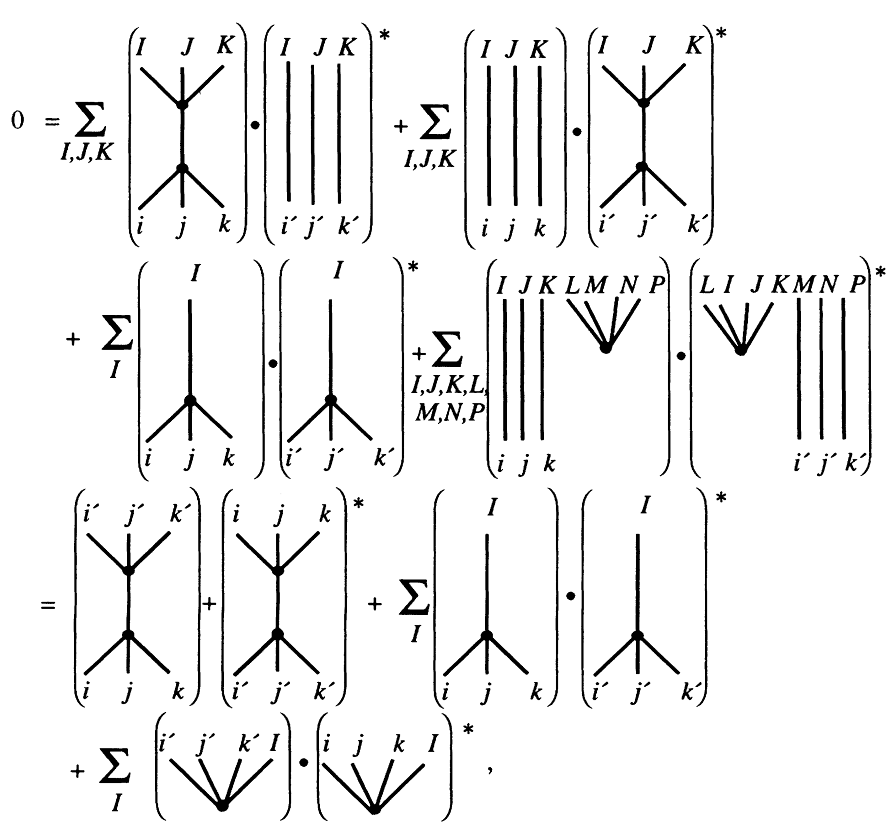
(79)

The four topologically distinct ways of making this cut that contribute are



(80)

This leads to the vanishing expression



(81)

and to an analogue of pointwise unitarity relation (70) and (77):

$$0=-i\Delta_F(x,y)-\overline{i\Delta_F(x,y)}+D(x,y)+\overline{D(x,y)}\ . \tag{82}$$

If $i\Delta_F(x,y)=\theta(x^0-y^0)D(x,y)+\theta(y^0-x^0)\overline{D(x,y)}$, then Eq. (82) follows trivially. Because of its linearity, Eq. (82) is true under even weaker conditions. If $i\Delta_F(x,y)=\theta(x^0-y^0)D(x,y)+\theta(y^0-x^0)\overline{D(x,y)}$ on the initial hypersurface, then because both $\mathrm{Re}[i\Delta_F(x,y)]$ and $D(x,y)$ each satisfy the Klein-Gordon equation throughout all spacetime, Eq. (82) also holds for the entire spacetime, even if the spacetime does not possess a foliation of spacelike hypersurfaces.

A theory with higher-order interactions (in lower dimensions) would be unitary if it satisfied analogues of the pointwise relations (82), (77), and (70) with higher powers of n:

$$0=-[i\Delta_F(x,y)]^n-[\overline{i\Delta_F(x,y)}]^n+[D(x,y)]^n+\overline{D(x,y)}]^n\ . \tag{83}$$

The connected $O(\lambda)$ diagrams have four external lines (four-point vertex), or two (tadpole). For four external lines, unitarity is trivial: for a diagram with two legs in the past and two in the future, it has the form

$$\begin{aligned}0&=\langle i',j'|\mathcal{S}^{(1)}|i,j\rangle+\langle i',j'|\mathcal{S}^{(1)\dagger}|i,j\rangle\\&=i\lambda\int dz\,[-\overline{f_{i'}}(z)\overline{f_{j'}}(z)f_i(z)f_j(z)+\overline{f_{i'}}(z)\overline{f_{j'}}(z)f_i(z)f_j(z)]\ .\end{aligned} \tag{84}$$

The unitarity of the tadpole diagram depends on the reality of the regularized coincidence limit of the two-point function $\Delta_F(x,x)=D(x,x)+\overline{D}(x,x)$:

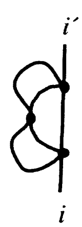
$$0=\langle i'|\mathcal{S}^{(1)}|i\rangle+\langle i'|\mathcal{S}^{(1)\dagger}|i\rangle=-i\lambda\int dz\,\overline{f_{i'}}(z)f_i(z)[\Delta_F(z,z)-\overline{\Delta_F}(z,z)]\ . \tag{85}$$

In the presence of closed timelike curves, the coincidence limit of the propagator will in general be complex, because it includes contributions from multiple loops around closed timelike (or null) curves, and the phases of these contributions depend on the global structure of the spacetime, not just on the short-distance behavior [4,12].

Higher-order unitarity relations are not of the form of Eq. (83). For instance, the pointwise unitarity relation from the third order, connected, one-particle–one-particle scattering diagram is

$$\begin{aligned}0=&[i\Delta_F(x,y)]^p[i\Delta_F(y,z)]^q[i\Delta_F(x,z)]^r-[\overline{i\Delta_F(x,y)}]^p[\overline{i\Delta_F(y,z)}]^q[\overline{i\Delta_F(x,z)}]^r\\&-[i\Delta_F(x,y)]^p[D(y,z)]^q[D(x,z)]^r+[D(x,y)]^p[\overline{i\Delta_F(y,z)}]^q[D(x,z)]^r\\&-[\overline{D(x,y)}]^p[i\Delta_F(y,z)]^q[\overline{D(x,z)}]^r+[\overline{i\Delta_F(x,y)}]^p[\overline{D(y,z)}]^q[\overline{D(x,z)}]^r\\&-[D(x,y)]^p[\overline{D(y,z)}]^q[i\Delta_F(x,z)]^r+[\overline{D(x,y)}]^p[D(y,z)]^q[\overline{i\Delta_F(x,z)}]^r\ ,\end{aligned} \tag{86}$$

for $(p,q,r)=(2,2,1)$. It cannot be reexpressed by equations of the form of Eq. (83). This relation is obtained by cutting the diagram



(87)

in its eight possible ways (one above all three vertices, three with one vertex above the cut, three with two vertices above the cut, and one below all three vertices).

All remaining third order, connected, scattering diagrams, when cut, give the same pointwise relation as Eq. (86), but with various non-negative integer values for p, q, and r. All these relations are trivially satisfied when the propagator is of the form

$$i\Delta_F(x,y)=\theta(x^0-y^0)D(x,y)+\theta(y^0-x^0)\overline{D(x,y)}\ .$$

The failure of any of the smeared identities constitutes a disproof of unitarity for the theory. Note that all the identities are consequences of unitarity of the interacting field, but are themselves statements about the free-field propagator. A free field could violate any of these identities and would still be unitary, but any attempt to make the field interact with itself would destroy the unitarity. This appears to be the case for the $\lambda\varphi^4$ theory in the presence of closed timelike curves [4].

In quantum electrodynamics in curved spacetime [11], the unitarity relations again rely on a set of identities analogous to those relating the Feynman propagators to two-point Wightman functions in Eqs. (83) and (86). In a covariant gauge, however, relations between the propagator and the Wightman functions are not quite sufficient to ensure unitarity. One must also verify that only physical degrees of freedom contribute to the sum over intermediate states appearing in the unitarity identities.

V. SUMMARY

Free-field theories enjoy unitarity in a large class of background spacetimes, including some spacetimes with closed timelike curves. Our demonstration of unitarity for interacting fields, however, required global hyperbolicity. Unitarity also requires renormalizability of the particular quantum theory, which is difficult to prove for general curved backgrounds, but it is believed that this difficulty is technical, not fundamental.

In curved spacetime, an interacting quantum field theory, defined by its perturbative expansion in powers of an interaction coupling constant, possesses an LSZ reduction similar to the case of flat spacetime. By using the Feynman path-integral formulation, this generalization of LSZ reduction does not require the spacetime to be globally foliated with spacelike hypersurfaces. As in flat spacetime, Feynman diagrams simplify scattering calculations and also allow a diagrammatic representation of the interaction process. Unlike flat spacetime, we must, in general, restrict the calculations and diagrams to configuration space, since a momentum space representation is not available.

A unitary, perturbative, interacting field theory satisfies an infinite sequence of identities obtained by expanding matrix elements of the identity $\mathcal{S}^\dagger\mathcal{S}=1$ in powers of the coupling constant. The momentum-space representation versions of these identities in flat spacetime are known as Cutkosky rules. These identities have a straightforward interpretation: any field state that evolves into the future, if then evolved backward to the past, and summed over all future states, should result in the initial field state. Using Feynman diagrams one can represent the process pictorially as a sewing together of cut versions of the same diagram, as in flat spacetime. This pictorial nature of the Feynman diagrams makes these rules particularly straightforward to calculate.

ACKNOWLEDGMENTS

We would like to thank Jim Hartle, Kip Thorne, John Preskill, Rafael Sorkin, Leonard Parker, and Bruce Allen for helpful conversations. This work was supported by NSF Grant No. PHY-91-05935.

[1] See, e.g., N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, Cambridge, England, 1982); S. A. Fulling, *Aspects of Quantum Field Theory in Curved Space-Time* (Cambridge University Press, Cambridge, England, 1989), and references in both.

[2] S. L. Adler, Phys. Rev. D **6**, 334 (1972); D. Z. Freedman and E. H. Weinberg, Ann. Phys. (N.Y.) **87**, 354 (1974); I. T. Drummond, Nucl. Phys. **B94**, 115 (1975); I. T. Drummond and G. M. Shore, Phys. Rev. D **19**, 1134 (1979); B. S. Kay, *ibid.* **20**, 3052 (1979); N. D. Birrell and L. H. Ford, Ann. Phys. (N.Y.) **122**, 1 (1979); Phys. Rev. D **22**, 330 (1980); N. C. Birrell and J. G. Taylor, J. Math. Phys. **21**, 1740 (1980); I. Ichinose, Phys. Rev. D **25**, 365 (1982). For a good review article, see L. H. Ford, in *Quantum Theory of Gravity: Essays In Honor of The 60th Birthday of Bryce S. DeWitt*, edited by S. M. Christensen (Hilger, Bristol, 1984).

[3] T. S. Bunch, P. Panangaden, and L. Parker, J. Phys. A **13**, 901 (1980); T. S. Bunch and P. Panangaden, *ibid.* **13**, 919 (1980); T. S. Bunch, Ann. Phys. (N.Y.) **131**, 118 (1981); D. J. Toms, Phys. Rev. D **26**, 2713 (1982); T. S. Bunch and L. Parker, *ibid.* **20**, 2499 (1979).

[4] J. Friedman, N. Papastamatiou, and J. Simon, following paper, Phys. Rev. D **46**, 4456 (1992).

[5] H. Lehmann, K. Symanzik, and W. Zimmermann, Nuovo Cimento **1**, 205 (1955).

[6] R. E. Cutkosky, J. Math. Phys. **1**, 429 (1960).

[7] B. S. DeWitt, Phys. Rep. **19**, 295 (1975); R. M. Wald, Ann. Phys. (N.Y.) **118**, 490 (1979).

[8] J. Glimm and A. Jaffe, in *Mathematics of Contemporary Physics,* edited by R. F. Streater (Academic, London, 1972); K. Symanzik, Nucl. Phys. **B190**, 1 (1981).

[9] K. Freese, C. Hill, and M. Mueller, Nucl. Phys. **B255**, 693 (1985); R. Floreanini, C. T. Hill, and R. Jackiw, Ann. Phys. (N.Y.) **175**, 345 (1987).

[10] J. L. Friedman and M. S. Morris, Phys. Rev. Lett. **66**, 401 (1990); J. L. Friedman, M. S. Morris, I. D. Novikov, F. Echeverria, G. Klinkhammer, K. S. Thorne, and U. Yurtsever, Phys. Rev. D **42**, 1915 (1990).

[11] I. L. Buchbinder, E. S. Fradkin, and D. M. Gitman, Fortschr. Phys. **29**, 187 (1981).

[12] D. Boulware, preceding paper, Phys. Rev. D **46**, 4421 (1992).