

Construction of states for quantum fields in nonstatic space-times

David Deutsch

Department of Astrophysics, South Parks Road, Oxford, OX1 3RQ United Kingdom

Amir-Homayoon Najmi

Department of Physics, University of Utah, Salt Lake City, Utah 84102

(Received 16 August 1982)

We set out the problem of constructing states for quantum field theories in nonstationary space-time backgrounds. We stress that this problem is a fundamental one, not to be confused with the inessential question "what is a particle." We present a formalism, related to that of Ashtekhar and Magnon, in which the problem can be attacked more easily than at present. By way of illustration we express the ansatz of "energy minimization" in our formalism and thereby find its general solution for noninteracting scalar fields.

I. INTRODUCTION

The process of "quantization" of a classical field theory yielding quantum dynamical equations and an algebra of quantum operators is not sufficient to determine a quantum field theory. Even if the problems of factor ordering and renormalization were solved, it would still be necessary to find a general procedure for constructing a space \mathcal{F} of physical states.

Such a procedure is essential for the extraction of physical information from the formalism of the quantum field theory. For example, one may follow Schwinger¹ and relate all physical quantities to S -matrix elements

$$S(t_2, k_2; t_1, k_1) = \langle \hat{O}(t_2), k_2 | \hat{O}(t_1), k_1 \rangle . \quad (1.1)$$

Here $\hat{O}(t)$ is a "fundamental observable". That is,

(i) $\hat{O}(t)$ is constructed solely from field operators in an infinitesimal neighborhood of the instant t , and from locally determined c numbers,

(ii) $\hat{O}(t_1)$ and $\hat{O}(t_2)$ at different instants are constructed *in an analogous way*, and

(iii) the spectrum of $\hat{O}(t)$ is preserved by the dynamical evolution. $|\hat{O}(t), k\rangle$ is the k th eigenstate of $\hat{O}(t)$. We shall pay no attention to the question whether k ranges over a discrete or a continuous set; thus we shall use the symbol \sum_k to denote any appropriate summations and/or integrations and δ_{kk_1} to denote the corresponding δ function.

In asymptotically static space-times (see below), $\hat{O}(t)$ may be taken to commute with a particle-number operator $\hat{N}(t)$. This gives the states $|\hat{O}(t), k\rangle$ a simple interpretation as many-particle

states. This has sometimes led to an unfortunate identification of the problem of constructing states with the problem of defining what particles are. The latter is a subsidiary problem of nomenclature and interpretation which will arise in quite a different form in nonstatic backgrounds since, as we shall see, the correct $\hat{O}(t)$ will then not commute with $\hat{N}(t)$ and the latter may not exist.

One reason why S matrices (1.1) with a temporally local (i) form, rather than the more general

$$\langle \hat{A}[\hat{\Phi}], k_2 | \hat{B}[\hat{\Phi}], k_1 \rangle , \quad (1.2)$$

are customarily used to make contact between theory and experiment may be that most practical measurement apparatuses select states $|\hat{A}[\hat{\Phi}], k\rangle$, where the function \hat{A} has effective support over only a small (space)-time domain. A large body of experience and understanding has been built up for the interpretation of (1.1), and hardly any for (1.2). Nevertheless, we may eventually be forced² to abandon the requirement (i) and embrace (1.2) with all its interpretational difficulties.

The physical content of requirement (ii), that a fundamental observable be constructed *in an analogous way* at each instant, is that it should correspond to measurements made with the same apparatus at different times. In this paper, we shall follow the usual practice of assuming that an interpretation with this property exists whenever (iii) is satisfied, so we shall ignore (ii) henceforward. We note in passing that the physical intuitions motivating (i)–(iii) may not be appropriate to a nonstatic expanding universe. For example, the spectrum of the (smeared) field operators would not be constant if the volume of the universe changed.

For noninteracting fields in Minkowski space-time, the fundamental observables $\hat{O}(t)$ may be chosen to be the particle-number operator $\hat{N}(t)$ and a set of operators commuting with it. [The construction of $\hat{N}(t)$ is discussed in Sec. II.] In this case, $\hat{N}(t)$ is of course constant and the conditions (i)–(iii) are trivially satisfied. If interactions and space-time curvature of a reasonably regular nature are introduced, *but confined to a compact space-time region*, then the same construction applied in either static region yields a particle-number operator $\hat{N}(t)$ which is no longer constant but, as Wald³ has shown, still satisfies (i)–(iii). Wald's result holds in any space-time which is static in the remote⁴ past and future and where the interactions and nonstaticity have effectively⁴ compact support. Thus particle production and related phenomena in such space-times are in principle completely understood: The S matrix $S(+\infty, k_1; -\infty, k_0)$ is the amplitude that a system initially in a state with past-attributes k_0 evolves into one with future-attributes k_1 . Our inability to extract from a quantum field theory physical information *inside* the compact interaction region causes no practical problem at the present state of technology, though it should be borne in mind that in principle the amplitudes $S(+\infty, k_1; -\infty, k_0)$ do not express the entire physical content of a quantum theory.

When dealing with general cosmological space-time backgrounds (say, for the sake of argument, recollapsing models with no static region and no event horizon), it is not yet known how to construct the space of physical states. This problem is separate from that of the ambiguities⁵ arising when horizons are present. No fundamental observable with the properties (i)–(iii) is known. In particular, no generalization of the static particle-number operator is known to satisfy (i)–(iii) in nonstatic space-times. Since (iii) implies that corresponding elements of the state spaces constructed at different times are related by a unitary transformation, i.e.,

$$\begin{aligned} \text{(i)} \quad & \sum_{k_0} S(t_1, k_1; t_0, k_0) S^*(t_1, k_2; t_0, k_0) = \delta_{k_1 k_2} \quad , \\ \text{(ii)} \quad & \sum_{k_1} S^*(t_1, k_1; t_0, k_0) S(t_1, k_1; t_0, k_2) = \delta_{k_0 k_2} \quad , \end{aligned} \quad (1.3)$$

the failure of a particular construction to satisfy (i)–(iii) is sometimes referred to as nonexistence of the S matrix or a lack of unitary equivalence between state spaces at different times. We shall call it the “Fulling pathology,” since Fulling² was the first to give a detailed account of the difficulties generally encountered in trying to make the existence of a particle-number operator at each instant consistent with (i)–(iii).

Several related formalisms have been used by various authors to express procedures for constructing states for noninteracting fields. Three of these are directed towards specifying a complex structure⁶ on the space V of real c -number solutions of the dynamical equations. This is used to distinguish between “positive-frequency” and “negative-frequency” complex solutions which in turn (at least via canonical quantization), determine a (Fock) space of states. These three formalisms are discussed in Sec. II and their equivalence demonstrated in Sec. III. They are as follows. (1) The specification of a set of physical positive- and negative-frequency modes in terms of some reference set. Such expansions are called Bogoliubov expansions. They are useful only when the space-time contains a region, such as a static region, where a “preferred” complex structure exists. (2) Ashtekar's J , a linear operator on V such that $\frac{1}{2}(1 \mp iJ)$ are projection operators for positive- and negative-frequency functions. This is cumbersome in practical calculations since J must satisfy the constraint $J^2 = -1$. (3) Boundary conditions imposed on some spacelike hypersurface \mathcal{S}_t . This is the method which we are elaborating in this paper. Generalizing the method of Brown and Dutton,⁷ we specify the linear operator \vec{M} on a suitable function space on \mathcal{S}_t for which positive-frequency functions Φ_k^+ satisfy

$$[n_\mu \nabla^\mu \Phi_k^+(x) - \vec{M} \Phi_k^+(x)]|_{x \in \mathcal{S}_t} = 0 \quad , \quad (1.4)$$

where n_μ is the unit future-pointing vector field normal to \mathcal{S}_t .

It is appropriate to mention two approaches to the construction of states which do not necessarily lead to Fock spaces. One is to specify sufficient boundary or other conditions to fix the Feynman Green's function⁸

$$G^F(x, x') = i \frac{\langle \hat{O}(t_2), \text{vac} | T \hat{\Phi}(x) \hat{\Phi}(x') | \hat{O}(t_1), \text{vac} \rangle}{\langle \hat{O}(t_2), \text{vac} | \hat{O}(t_1), \text{vac} \rangle} \quad , \quad (1.5)$$

where T is the time-ordering symbol and the points x and x' are at the instants t and t' , respectively.

The other is to construct the direct sum of Fock spaces, each based on a different complex structure.⁹ Under such constructions no particle-number operator exists although physically measurable quantities all have finite matrix elements.

In this paper we shall restrict ourselves to single Fock spaces, although our \vec{M} formalism can also be used to express these wider constructions.

If \mathcal{F} is to be a Fock space the problem of constructing states reduces to finding a general criterion for distinguishing between positive and negative fre-

quency. Many such criteria have been proposed, most¹⁰ in the hope (to date, vain) that they might satisfy conditions such as (i)–(iii). By “general” one means geometrical and applicable in arbitrary space-time backgrounds.

Parker’s suggestion² of an “adiabatic definition of states,” a temporally nonlocal construction, can unfortunately be applied only in spatially homogeneous cosmologies. However, the physical motivation of Parker’s construction does give simple physical interpretations to his states, so if he is right in advocating the abandonment of requirement (i) and the interpretation of the theory via amplitudes such as (1.2) rather than (1.1) it is necessary to seek a generalization of his adiabatic construction to general space-times. We do not attempt that in this paper, though the formalism we set up in Sec. II provides tools with which one could attempt it.

The procedure to which we shall be devoting most of our attention (though more as a testing ground for our formalism and general approach than because we especially advocate it) is energy minimization. The idea here is that one considers the renormalized instantaneous expectation value of the Hamiltonian operator (or some other total-energy operator) in the vacuum state of the Fock space generated by *each* possible construction of states. The choice which minimizes this value is then supposed to be the physical state space. The difficulty here is that all known versions of this method in general generate a different space for each instant and thus suffer from the Fulling pathology.

Ashtekar’s criterion, a “correspondence principle” requirement that the energy of a one-particle state should be equal to its classical energy, yields the same complex structure as energy minimization in the case of minimal coupling (see the next section).

II. FIELD QUANTIZATION AND COMPLEX STRUCTURES

For simplicity we shall consider only the real (Hermitian) free scalar field $\hat{\Phi}(x)$ propagating in a curved globally hyperbolic space-time \mathcal{M} . Thus the problems of alternative quantizations in space-times with horizons, though interesting and important in themselves, will not concern us. The stationarity of the action functional

$$\hat{S}[\hat{\Phi}] = -\frac{1}{2} \int_{\mathcal{M}} [\hat{\Phi}_{;\mu} \hat{\Phi}^{;\mu} + (m^2 + \xi R) \hat{\Phi}^2] g^{1/2} d^4x \quad (2.1)$$

with respect to infinitesimal c -number changes $\delta\phi(x)$ in the form of $\hat{\Phi}(x)$ implies the dynamical equation

$$\hat{\Phi}_{;\mu}^{;\mu} - (m^2 + \xi R) \hat{\Phi} = 0, \quad (2.2)$$

where m is the mass of a field quantum and ξ is the conformal coupling constant, $\xi = \frac{1}{6}$ for the new improved scalar field, which transforms as a density under conformal transformations and $\xi = 0$ for the minimally coupled scalar field.

Let \mathcal{M} be foliated by a family \mathcal{S}_t of spacelike hypersurfaces with normals

$$n^\mu(x) = n^\mu(t, \underline{x}) = ||\nabla t||^{-1} \nabla t, \quad (2.3)$$

where t is a parameter labeling the hypersurfaces and the \underline{x} label points in \mathcal{S}_t .

The canonical commutation relations for $\hat{\Phi}(x)$ may be expressed as follows:

$$\begin{aligned} \text{(i)} \quad & [\hat{\Phi}(x'), n \cdot \nabla \hat{\Phi}(x)] |_{x, x' \in \mathcal{S}_t} = i \delta(\underline{x}, \underline{x}') \hat{1}, \\ \text{(ii)} \quad & [\hat{\Phi}(x'), \hat{\Phi}(x)] |_{x, x' \in \mathcal{S}_t} = \hat{0}. \end{aligned} \quad (2.4)$$

We now solve (2.2) subject to (2.4). To this end consider the auxiliary c -number field equation

$$\phi_{;\mu}^{;\mu} - (m^2 + \xi R) \phi = 0. \quad (2.5)$$

Because of the Cauchy completeness of \mathcal{M} every solution $\phi(x)$ of (2.5) may be identified with its Cauchy data on some \mathcal{S}_t ,

$$\phi(x) \leftrightarrow \{ \phi(t, \underline{x}), n \cdot \nabla \phi(t, \underline{x}) \}. \quad (2.6)$$

Let $U_k(\underline{x})$ be any complete and orthonormal basis for real functions on \mathcal{S}_t , in the sense that

$$\begin{aligned} \text{(i)} \quad & \sum_k U_k(\underline{x}) U_k(\underline{x}') = \delta(\underline{x}, \underline{x}'), \\ \text{(ii)} \quad & \int_{\mathcal{S}_t} U_{k_1}(\underline{x}) U_{k_2}(\underline{x}) d\underline{x} = \delta_{k_1 k_2}, \end{aligned} \quad (2.7)$$

where $d\underline{x}$ is the covariant volume element on \mathcal{S}_t .

If M_{kk_1} is any complex matrix whose anti-Hermitian part has a negative spectrum, then the functions

$$\begin{aligned} u_k(x) \leftrightarrow & \sum_{k_2} [i(M - M^\dagger)]^{-1/2}_{k_2 k} \\ & \times \left\{ U_{k_2}(\underline{x}), \sum_{k_1} M_{k_1 k_2} U_{k_1}(\underline{x}) \right\} \end{aligned} \quad (2.8)$$

form a set of solutions of (2.5), complete and orthonormal according to

$$\begin{aligned} i \sum_k & [u_k^*(x') n \cdot \nabla u_k(x) \\ & - u_k(x') n \cdot \nabla u_k^*(x)] |_{x, x' \in \mathcal{S}_t} = \delta(\underline{x}, \underline{x}') \end{aligned} \quad (2.9)$$

and

$$\begin{aligned} \text{(i)} \quad & i \int_{\mathcal{S}_t} u_{k_1}^* n \cdot \nabla u_{k_2} d\underline{x} = \delta_{k_1 k_2}, \\ \text{(ii)} \quad & i \int_{\mathcal{S}_t} u_{k_1} n \cdot \nabla u_{k_2} d\underline{x} = 0. \end{aligned} \quad (2.10)$$

Alternative sets of solutions $\{u_k(x)\}$ and $\{v_k(x)\}$ constructed via different M_{kk_1} 's may always be expressed in terms of each other by a Bogoliubov transformation

$$v_{k_1}(x) = \sum_k [\alpha_{kk_1} u_k(x) + \beta_{kk_1} u_k^*(x)] . \quad (2.11)$$

If these also satisfy (2.9) and (2.10), then

$$(i) \sum_k (\alpha_{kk_1}^* \alpha_{kk_2} - \beta_{kk_1}^* \beta_{kk_2}) = \delta_{k_1 k_2} , \quad (2.12)$$

$$(ii) \sum_k (\alpha_{kk_1} \beta_{kk_2} - \beta_{kk_1} \alpha_{kk_2}) = 0 .$$

The quantity M_{kk_1} determines a so-called complex structure on the space V of real solutions of (2.5) (some restrictions on M_{kk_1} will be discussed in the following section). It is to be thought of as a geometrical object on \mathcal{S}_t . It is convenient to consider it both in the k representation, as M_{kk_1} , and in the x representation, as \vec{M} , a linear integral operator on a suitable function space on \mathcal{S}_t .

Since $\hat{\Phi}(x)$ is Hermitian, the general solution of (2.2) is

$$\hat{\Phi}(x) = \sum_k [\hat{a}_k u_k(x) + \hat{a}_k^* u_k^*(x)] , \quad (2.13)$$

where the operators \hat{a}_k are chosen so as to satisfy the commutation relations (2.4), which amount to

$$(i) [\hat{a}_{k_1}, \hat{a}_{k_2}^*] = \delta_{k_1 k_2} \hat{1}$$

and (2.14)

$$(ii) [\hat{a}_{k_1}, \hat{a}_{k_2}] = [\hat{a}_{k_1}^*, \hat{a}_{k_2}^*] = \hat{0} .$$

Everything we have said so far is independent of the properties of the physical states, the construction of which has been reduced to the choice of a representation for (2.14), to which we now turn. As is well known,¹¹ there exist uncountably many such representations not related by unitary transformations. From an algebraic point of view it might seem natural to require that the construction of states be invariant under changes in the choice of basis functions $u_k(x)$ or of \vec{M} , since these were introduced apparently only as technical aids in the solution of the dynamical equations. Although such constructions are possible,^{9,12} they have not been widely explored and their interpretation (nonparticle states) is problematic. In this paper we are concerned almost entirely with the conventional Fock-space or many-particle-states construction. To this end, consider a single choice \vec{M} of complex structure, and let \hat{a}_k be the operator coefficients in the expansion (2.12) of $\hat{\Phi}(x)$ in terms of the corresponding solutions of (2.5). Then it is easily verified that

the formal operator

$$\hat{N}_{\vec{M}} = \sum_k \hat{a}_k^*(\vec{M}) \hat{a}_k(\vec{M}) \quad (2.15)$$

has in virtue of (2.14) a non-negative-integer spectrum. The requirement that $\hat{N}_{\vec{M}}$ exist is sufficient to determine uniquely a space of many-particle states on which $\hat{N}_{\vec{M}}$ is the particle-number operator and $\hat{a}_k^*(\vec{M})$ and $\hat{a}_k(\vec{M})$ are, respectively, creation and annihilation operators for \vec{M} -particles of type k . We call this the \vec{M} -Fock space.

The Bogoliubov relations (2.11) between sets of solutions corresponding to different complex structures imply the following relation between the corresponding creation and annihilation operators:

$$(i) \hat{a}_k(\vec{M}) = \sum_{k_1} [\alpha_{kk_1} \hat{a}_{k_1}(\vec{N}) + \beta_{kk_1}^* \hat{a}_{k_1}^*(\vec{N})] , \quad (2.16)$$

$$(ii) \hat{a}_k(\vec{N}) = \sum_{k_1} [\alpha_{k_1 k}^* \hat{a}_{k_1}(\vec{M}) - \beta_{k_1 k}^* \hat{a}_{k_1}^*(\vec{M})] .$$

Thus the condition that $\hat{N}_{\vec{M}}$ and $\hat{N}_{\vec{N}}$ exist simultaneously is

$$\langle \vec{M}, \text{vac} | \hat{N}_{\vec{N}} | \vec{M}, \text{vac} \rangle = \sum_{k, k_1} \beta_{kk_1}^* \beta_{kk_1} < \infty . \quad (2.17)$$

In this case the \vec{N} - and \vec{M} -Fock spaces are unitarily equivalent; that is, they are the same space (*qua* Hilbert space) with different nomenclature for the states.

As we shall see, \vec{M} is a most convenient characterization of the complex structure. In the following section we demonstrate the equivalence of our \vec{M} formalism with the other formalisms.

III. FORMALISMS FOR EXPRESSING COMPLEX STRUCTURES

In this section, we elaborate on the connection between the three characterizations of the complex structure, namely our \vec{M} , Ashtekar and Magnon's J , and Bogoliubov transformations. We first obtain the operator J corresponding to a given \vec{M} .

Let $\Phi(x)$ be a real solution of (2.5) with Cauchy data $\{\Phi(\underline{x}), \Pi(\underline{x})\}$ on a chosen hypersurface \mathcal{S}_t , where $\underline{x} \in \mathcal{S}_t$. Then the complex structure J may be thought of as a real linear operator on the space of all real pairs of functions $\{f(\underline{x}), g(\underline{x})\}$:

$$J\{f(\underline{x}), g(\underline{x})\} = \{\vec{A}f(\underline{x}) + \vec{B}g(\underline{x}) , \vec{C}f(\underline{x}) + \vec{D}g(\underline{x})\} . \quad (3.1)$$

The positive- and negative-frequency parts of $\Phi(x)$ are defined to be

$$\Phi^\pm(x) = \frac{1}{2}(1 \mp iJ)\Phi(x) = P^\pm \Phi(x) , \quad (3.2)$$

where P^\pm are the positive- and negative-frequency projection operators. In order that P^\pm be projection operators, J^2 must equal -1 , or equivalently,

$$\begin{aligned} \text{(i)} \quad & \vec{A}^2 + \vec{B}\vec{C} = -\vec{I} \ , \\ \text{(ii)} \quad & \vec{D}^2 + \vec{C}\vec{B} = -\vec{I} \ , \\ \text{(iii)} \quad & \vec{A}\vec{B} + \vec{B}\vec{D} = \vec{0} \ , \\ \text{(iv)} \quad & \vec{C}\vec{A} + \vec{D}\vec{C} = \vec{0} \ . \end{aligned} \quad (3.3)$$

Φ^+ has Cauchy data (on \mathcal{S}_t)

$$\begin{aligned} & \left\{ \frac{1}{2}(1-i\vec{A})\Phi(\underline{x}) - \frac{1}{2}i\vec{B}\Pi(\underline{x}) \ , \right. \\ & \left. \frac{1}{2}(1-i\vec{D})\Pi(\underline{x}) - \frac{1}{2}i\vec{C}\Phi(\underline{x}) \right\} \ . \end{aligned} \quad (3.4)$$

In terms of the \vec{M} formalism,

$$n \cdot \nabla \Phi^+(x) \big|_{x \in \mathcal{S}_t} = \vec{M} \Phi^+(x) \big|_{x \in \mathcal{S}_t} \ . \quad (3.5)$$

Therefore,

$$\begin{aligned} (1-i\vec{D})\Pi(\underline{x}) - i\vec{C}\Phi(\underline{x}) \\ = \vec{M}[(1-i\vec{A})\Phi(\underline{x}) - i\vec{B}\Pi(\underline{x})] \ . \end{aligned} \quad (3.6)$$

Since $\Phi(x)$ and $\Pi(x)$ are arbitrary real functions,

$$\begin{aligned} \text{(i)} \quad & -i\vec{C} = \vec{M}(1-i\vec{A}) \ , \\ \text{(ii)} \quad & -i\vec{M}\vec{B} = 1-i\vec{D} \ , \end{aligned} \quad (3.7)$$

which can be solved for \vec{M} to give

$$\vec{M} = -i\vec{C}(1-i\vec{A})^{-1} \ . \quad (3.8)$$

Thus we have obtained the \vec{M} corresponding to a given J . More convenient forms can be obtained from (3.14) below.

Substituting the above in (3.7 ii) and equating real and imaginary parts leads to

$$\begin{aligned} \text{(i)} \quad & 1 = -\vec{C}(1+\vec{A}^2)^{-1}\vec{B} \ , \\ \text{(ii)} \quad & \vec{D} = \vec{C}\vec{A}(1+\vec{A}^2)^{-1}\vec{B} \ . \end{aligned} \quad (3.9)$$

We may now solve (3.3) for \vec{B} , \vec{C} , and \vec{D} :

$$\begin{aligned} \text{(i)} \quad & \vec{B} = -(1+\vec{A}^2)\vec{C}^{-1} = i(1+i\vec{A})\vec{M}^{-1} \ , \\ \text{(ii)} \quad & \vec{C} = i\vec{M}(1-i\vec{A}) \ , \\ \text{(iii)} \quad & \vec{D} = -\vec{M}\vec{A}\vec{M}^{-1} \ . \end{aligned} \quad (3.10)$$

Now, writing $\vec{M} = \vec{R} + i\vec{I}$, where

$$\begin{aligned} \text{(i)} \quad & \vec{R} = \frac{1}{2}(\vec{M} + \vec{M}^*) \ , \\ \text{(ii)} \quad & \vec{I} = -\frac{1}{2}i(\vec{M} - \vec{M}^*) \ , \end{aligned} \quad (3.11)$$

and demanding that J should map real functions into real functions, we obtain

$$\vec{A} = -\vec{I}^{-1}\vec{R} \ . \quad (3.12)$$

In view of the identity

$$\vec{M}^{-1} \equiv -i(1+i\vec{I}^{-1}\vec{R})(1+\vec{I}^{-1}\vec{R}\vec{I}^{-1}\vec{R})^{-1}\vec{I}^{-1} \ , \quad (3.13)$$

(3.10) and (3.12) are equivalent to

$$\begin{aligned} \text{(i)} \quad & \vec{A} = -\vec{I}^{-1}\vec{R} \ , \\ \text{(ii)} \quad & \vec{B} = \vec{I}^{-1} \ , \\ \text{(iii)} \quad & \vec{C} = -(\vec{I} + \vec{R}\vec{I}^{-1}\vec{R}) \ , \\ \text{(iv)} \quad & \vec{D} = \vec{R}\vec{I}^{-1} \ . \end{aligned} \quad (3.14)$$

Thus J has been expressed in terms of \vec{M} .

There are two further conditions on J which follow from the requirement that $\Phi^+(x)$ be a one-particle wave function with positive classical energy, which we shall now translate into constraints on \vec{M} . They are

$$\int_{\mathcal{S}_t} \Phi_1 \vec{\nabla}_\mu \Phi_2 d\Sigma^\mu = \int_{\mathcal{S}_t} \Phi_2 \vec{\nabla}_\mu \Phi_1 d\Sigma^\mu \quad (3.15)$$

and

$$\int_{\mathcal{S}_t} \Phi \vec{\nabla}_\mu J \Phi d\Sigma^\mu > 0 \text{ (for all } \Phi \neq 0) \ . \quad (3.16)$$

In terms of our real basis functions $U_k(\underline{x})$, these conditions are

$$\begin{aligned} \text{(i)} \quad & \int_{\mathcal{S}_t} U_{k_1}(\underline{x}) \vec{I}^{-1} \vec{R} U_{k_2}(\underline{x}) d\underline{x} \\ & = \int_{\mathcal{S}_t} U_{k_2}(\underline{x}) \vec{R} \vec{I}^{-1} U_{k_1}(\underline{x}) d\underline{x} \ , \\ \text{(ii)} \quad & \int_{\mathcal{S}_t} U_{k_1}(\underline{x}) (\vec{I} + \vec{R} \vec{I}^{-1} \vec{R}) U_{k_2}(\underline{x}) d\underline{x} \\ & = \int_{\mathcal{S}_t} U_{k_2}(\underline{x}) (\vec{I} + \vec{R} \vec{I}^{-1} \vec{R}) U_{k_1}(\underline{x}) d\underline{x} \ , \\ \text{(iii)} \quad & \int_{\mathcal{S}_t} U_{k_1}(\underline{x}) \vec{I}^{-1} U_{k_2}(\underline{x}) d\underline{x} \\ & = \int_{\mathcal{S}_t} U_{k_2}(\underline{x}) \vec{I}^{-1} U_{k_1}(\underline{x}) d\underline{x} \ , \end{aligned} \quad (3.17)$$

and

$$\begin{aligned} \text{(i)} \quad & \int_{\mathcal{S}_t} U_k(\underline{x}) \vec{I}^{-1} U_k(\underline{x}) < 0 \ , \\ \text{(ii)} \quad & \int_{\mathcal{S}_t} U_k(\underline{x}) (\vec{I} + \vec{R} \vec{I}^{-1} \vec{R}) U_k(\underline{x}) < 0 \ . \end{aligned} \quad (3.18)$$

These imply that \vec{R} and \vec{I} are symmetric and that \vec{I} is negative definite.

This completes our discussion of the connection between the J and \vec{M} characterizations of the complex structure.

In order to discuss the relations between Bogoliubov transformations and J , we require properly normalized positive-frequency functions; that is, given J , we wish to find the real solution $\Phi_k(x)$ whose positive- and negative-frequency parts, viz.

$$\Phi_k^\pm(x) = \frac{1}{2}(1 \mp iJ)\Phi_k(x) \quad (3.19)$$

satisfy the orthonormality relations (2.10). It is necessary and sufficient to have

$$(i) \int_{\mathcal{S}_t} \Phi_{k_1} \vec{\nabla}_\mu \Phi_{k_2} d\Sigma^\mu = 0$$

and

$$(3.20)$$

$$(ii) \int_{\mathcal{S}_t} \Phi_{k_1} \vec{\nabla}_\mu J \Phi_{k_2} d\Sigma^\mu = 2\delta_{k_1 k_2} .$$

We may identify $\Phi_k(x)$ with its Cauchy data on \mathcal{S}_t

$$\sum_{k_1} \{ \lambda_{k_1 k} U_{k_1}(\underline{x}), \mu_{k_1 k} U_{k_1}(\underline{x}) \} , \quad (3.21)$$

where $U_k(\underline{x})$ are the real basis functions on \mathcal{S}_t previously described.

Then the real coefficients λ_{kk_1} and μ_{kk_1} will depend on the real linear operator J . Equation (3.20) is equivalent to

$$(i) \sum_k (\lambda_{kk_1} \mu_{kk_2} - \mu_{kk_1} \lambda_{kk_2}) = 0 , \quad (3.22)$$

$$(ii) \sum_{k', k''} (\lambda_{k' k_1} C_{k' k''} \lambda_{k'' k_2} + \lambda_{k' k_1} D_{k' k''} \mu_{k'' k_2} - \mu_{k' k_1} A_{k' k''} \lambda_{k'' k_2} - \mu_{k' k_1} B_{k' k''} \mu_{k'' k_2}) = 2\delta_{k_1 k_2} .$$

Equation (3.22) admits many solutions for λ and μ . For example,

$$(i) \lambda_{kk_1} = 0, \mu_{kk_1} = 2(-B)_{kk_1}^{-1/2} ,$$

$$(ii) \lambda_{kk_1} = 2C_{kk_1}^{-1/2}, \mu_{kk_1} = 0 . \quad (3.23)$$

We now prove that all such solutions lead to identical definitions of positive frequency (connected to each other by trivial Bogoliubov transformations with $\beta_{kk_1} = 0$).

To this end let

$$(i) \Phi_{k_1}^1(x) \leftrightarrow \sum_k \{ \lambda_{kk_1}^1 U_k(\underline{x}), \mu_{kk_1}^1 U_k(\underline{x}) \} ,$$

$$(ii) \Phi_{k_1}^2(x) \leftrightarrow \sum_k \{ \lambda_{kk_1}^2 U_k(\underline{x}), \mu_{kk_1}^2 U_k(\underline{x}) \} , \quad (3.24)$$

where both pairs $\{\lambda^1, \mu^1\}$ and $\{\lambda^2, \mu^2\}$ are solutions to (3.22). We now show that if

$$(i) u_k^1(x) = \frac{1}{2}(1 - iJ)\Phi_k^1(x) ,$$

$$(ii) u_k^2(x) = \frac{1}{2}(1 - iJ)\Phi_k^2(x) \quad (3.25)$$

are the corresponding positive-frequency solutions, where

$$i \int u_k^{1*} \vec{\nabla}_\mu u_{k_1}^1 d\Sigma^\mu = i \int u_k^{2*} \vec{\nabla}_\mu u_{k_1}^2 d\Sigma^\mu = \delta_{kk_1} , \quad (3.26)$$

then

$$\beta_{kk_1} = -i \int u_k^1 \vec{\nabla}_\mu u_{k_1}^2 d\Sigma^\mu = 0 . \quad (3.27)$$

We have

$$i \int u_k^1 \vec{\nabla}_\mu u_{k_1}^2 d\Sigma^\mu = \frac{1}{4} i [\Omega(\Phi_k^1, \Phi_{k_1}^2) - \Omega(J\Phi_k^1, J\Phi_{k_1}^2) - i\Omega(\Phi_k^1, J\Phi_{k_1}^2) + i\Omega(\Phi_k^1, J\Phi_{k_1}^2)] , \quad (3.28)$$

where the ‘‘symplectic structure’’ is defined by

$$\Omega(\Psi^1, \Psi^2) = \int_{\mathcal{S}_t} \Psi^1 \vec{\nabla}_\mu \Psi^2 d\Sigma^\mu . \quad (3.29)$$

But in view of (3.3) and (3.15),

$$\Omega(J\Phi^2, J\Phi^2) = \Omega(\Phi^1, \Phi^2) . \quad (3.30)$$

Equation (3.27) follows.

We now determine the Bogoliubov coefficients relating mode functions corresponding to two distinct complex structures J^1 and J^2 .

The positive-frequency functions, i.e., u_k^1 and u_k^2 are constructed as in (3.19). Since both $\{u_k^1\}$ and $\{u_k^2\}$ are complete and orthonormal, the Bogoliubov coefficients α_{kk_1} and β_{kk_1} which connect $\{u_k^1\}$ and $\{u_k^2\}$ according to

$$u_k^2(x) = \sum_{k_1} [\alpha_{k_1 k} U_{k_1}^1(x)] \quad (3.31)$$

are given by

$$(i) \alpha_{kk_1} = i \int u_k^{1*} \vec{\nabla}_\mu u_{k_1}^2 d\Sigma^\mu ,$$

$$(ii) \beta_{kk_1} = -i \int u_k^1 \vec{\nabla}_\mu u_{k_1}^2 d\Sigma^\mu , \quad (3.32)$$

or, in terms of real solutions Φ_k^1 and Φ_k^2 ,

$$(i) \alpha_{kk_1} = \frac{1}{4} i [\Omega(\Phi_k^1, \Phi_{k_1}^2) + \Omega(J^1 \Phi_k^1, J^2 \Phi_{k_1}^2) - i\Omega(\Phi_k^1, J^2 \Phi_{k_1}^2) + i\Omega(J^1 \Phi_k^1, \Phi_{k_1}^2)] ,$$

$$(ii) \beta_{kk_1} = -\frac{1}{4} i [\Omega(\Phi_k^1, \Phi_{k_1}^2) - \Omega(J^1 \Phi_k^1, J^2 \Phi_{k_1}^2) - i\Omega(\Phi_k^1, J^2 \Phi_{k_1}^2) - i\Omega(J^1 \Phi_k^1, \Phi_{k_1}^2)] . \quad (3.33)$$

The inverse problem to the above is essentially that of determining J given a complete set $\{u_k(x)\}$ of positive-frequency solutions. Since an arbitrary real solution $\Psi(x)$ can be expressed in the form

$$\Psi(x) = \sum_k [A_k u_k(x) + A_k^* u_k^*(x)] , \quad (3.34)$$

$$A_k = -i \int_{\mathcal{S}_t} \Psi \vec{\nabla}_\mu u_k^* d\Sigma^\mu , \quad (3.35)$$

where

the effect of J on such a function is easily calculated:

$$J\Psi = \sum_k (iA_k u_k - iA_k^* u_k^*) = \sum_k \left[\left[\int_{\mathcal{S}_t} \Psi(y) \vec{\nabla}_y^\mu u_k^*(y) d\Sigma_\mu(y) \right] u_k(x) + \left[\int_{\mathcal{S}_t} \Psi(y) \vec{\nabla}_y^\mu u_k(y) d\Sigma_\mu(y) \right] u_k^*(x) \right] . \quad (3.36)$$

IV. GENERAL SOLUTION OF THE ANSATZ OF ENERGY MINIMIZATION

A historically important approach to the problem of constructing fundamental observables was ‘‘Hamiltonian diagonalization.’’ The most physical way of stating this criterion is that the space of physical states be a Fock space \mathcal{F}_{\min} whose vacuum state has the lowest possible total energy. This approach now appears to be defunct.² All known versions of it suffer from the Fulling pathology in general space-times. However, it is a useful testing ground for our \vec{M} formalism, as we shall now demonstrate by calculating the general solution of the energy-minimization ansatz—indeed of a wide class of related criteria of which the historical one is a special case.

The renormalized stress-energy tensor $\hat{T}_{\text{ren}}^{\mu\nu}$ for the scalar field $\hat{\Phi}(x)$ of (2.1) is formally given by

$$\begin{aligned} \hat{T}_{\text{ren}}^{\mu\nu}(x) = & 2g^{-1/2} \frac{\delta}{\delta g_{\mu\nu}} \hat{S}[\hat{\Phi}] - T_\infty^{\mu\nu}(x) \hat{1} = g^{\mu\nu} \left[(2\xi - \frac{1}{2}) \hat{\Phi}_{;\sigma} \hat{\Phi}^{;\sigma} + 2\xi \hat{\Phi} \hat{\Phi}_{;\sigma}{}^{;\sigma} - \frac{1}{2} (m^2 - \xi R) \hat{\Phi}^2 \right] \\ & + (1 - 2\xi) \hat{\Phi}^{;\mu} \hat{\Phi}^{;\nu} - 2\xi \hat{\Phi} \hat{\Phi}^{;\mu\nu} - \xi R^{\mu\nu} \hat{\Phi}^2 - T_\infty^{\mu\nu}(x) \hat{1} . \end{aligned} \quad (4.1)$$

The counterterm $T_\infty^{\mu\nu}(x)$ is a formally infinite c number. The total energy which we shall wish to minimize is the spatial integral of the expectation value of an energy density constructed from (4.1),

$$E_{\text{ren}} = \int_{\mathcal{S}_t} \langle \vec{M}, \text{vac} | \hat{T}_{\text{ren}}^{\mu\nu}(\underline{x}) | \vec{M}, \text{vac} \rangle l_\mu d\Sigma_\nu , \quad (4.2)$$

where $d\Sigma_\nu = n_\nu d\underline{x}$ is the (future-pointing) volume element on \mathcal{S}_t . If \mathcal{S}_t is not compact, (4.2) is to be understood as the result of some formal procedure to factor out the infrared divergence in E_{ren} due to an infinite volume of \mathcal{S}_t . $l^\mu(\underline{x})$ is some timelike vector field. Different choices of l^μ correspond to different definitions of total energy. Unconventional choices of l^μ await exploration in the literature. The conventional ones are (i) $l^\mu = n^\mu$, which makes $T_{\mu\nu} l^\mu n^\nu$ the local energy density for observers at rest in \mathcal{S}_t ; (ii) l^μ is some timelike Killing vector field, which makes E_{ren} a conserved total energy; and (iii) l^μ is normal to the hypersurfaces of homogeneity (if such are present). To obtain the minimum E_{ren} we set

$$\frac{\delta}{\delta \vec{M}} E_{\text{ren}} = 0 . \quad (4.3)$$

Because the expectation value of $T_\infty^{\mu\nu} \hat{1}$ is formally independent of \vec{M} , E_{ren} will attain its *bona fide* minimum when the divergent quantity

$$E = \int_{\mathcal{S}_t} \langle \vec{M}, \text{vac} | \hat{T}^{\mu\nu} | \vec{M}, \text{vac} \rangle l_\mu d\Sigma_\nu \quad (4.4)$$

attains a formal stationary value with respect to variations

$$\vec{M} \rightarrow \vec{M} + \delta \vec{M} . \quad (4.5)$$

When $\xi = 0$ the energy E is formally a sum of squares, and may be considered non-negative definite. It follows that a formal minimum with respect to variations (4.5) will exist at some value of \vec{M} . When $\xi \neq 0$ (and, in particular, for the conformally invariant case $\xi = \frac{1}{6}$) this need no longer be the case. As we shall see, in a general background geometry, and for a general hypersurface \mathcal{S}_t , the ansatz of energy minimization possesses no solution for $\xi \neq 0$ scalar fields. We now obtain the solution for the class of cases where it exists.

We have

$$\begin{aligned} \langle \vec{M}, \text{vac} | \hat{T}^{\mu\nu}(x) | \vec{M}, \text{vac} \rangle = & \lim_{x' \rightarrow x} \left[(2\xi - \frac{1}{2}) g^{\mu\nu} (g^{\lambda\sigma'} \nabla_\lambda \nabla_{\sigma'} + m^2 + \xi R) + (1 - 2\xi) g^{\nu\sigma'} \nabla^\mu \nabla_{\sigma'} \right. \\ & \left. - \xi (\nabla^\mu \nabla^\nu + g^{\mu\sigma'} g^{\nu\delta'} \nabla_{\sigma'} \nabla_{\delta'}) + \frac{1}{2} \xi (R_\sigma^\mu g^{\sigma\nu} + g^{\mu\sigma'} R_{\sigma'}^\nu) \right] \sum_k u_k(x) u_k^*(x') . \end{aligned} \quad (4.6)$$

Now we use the field equation in the form

$$(n \cdot \nabla)^2 u_k(x) = u_{k|\alpha} + n^\mu n^\nu{}_{;\mu} u_{k;\nu} - \chi n \cdot \nabla u_k(x) - (m^2 + \xi R) u_k(x) , \quad (4.7)$$

where the vertical bar denotes the covariant derivative with respect to the intrinsic metric

$$h_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu \quad (4.8)$$

of the hypersurface \mathcal{S}_t and

$$\chi_{\alpha\beta} = h_{\mu\alpha} h_{\nu\beta} n^{\mu;\nu} \quad (4.9)$$

is the second fundamental form on \mathcal{S}_t ($\chi = \chi_\alpha^\alpha$), to eliminate all second-order time derivatives.

Next we set

$$\vec{M} = \vec{R} + i \vec{I} \quad (4.10)$$

with \vec{R} and \vec{I} Hermitian [in view of (2.10) we may in fact take them both real and symmetric] and write

$$E = -\frac{1}{4} \text{Tr}(l_\mu n^\mu) \{ [D^\alpha D_\alpha - m^2 - \xi R + 2\xi(l_\mu n^\mu)^{-1}(l_\mu n^\mu)_{|\alpha}{}^\alpha - \xi(l_\mu n^\mu)^{-1}(2l_\mu n^\mu n^\nu n^\beta{}_{;\nu} + 2l_\mu n^{\beta;\mu} + l_\mu \chi^{\mu\beta})_{|\beta} + 2\xi(l_\mu n^\mu)^{-1} l_\mu n_\nu R^{\mu\nu}] \vec{I}^{-1} - 2\xi \chi \vec{I}^{-1} \vec{R} - 2\xi \chi \vec{R} \vec{I}^{-1} - \vec{R} \vec{I}^{-1} \vec{R} - \vec{I} \} , \quad (4.11)$$

where Tr denotes a trace over space coordinates

$$\text{Tr}f(\underline{x}, \underline{x}') = \int_{\mathcal{S}_t} f(\underline{x}, \underline{x}') n_\nu d\Sigma^\nu . \quad (4.12)$$

Under a variation $\delta\vec{M} = \delta\vec{R}$,

$$\delta E = \frac{1}{4} \text{Tr}(l_\mu n^\mu) (2\xi \chi \vec{I}^{-1} \delta\vec{R} + 2\xi \chi \delta\vec{R} \vec{I}^{-1} + \delta\vec{R} \vec{I}^{-1} \vec{R} + \vec{R} \vec{I}^{-1} \delta\vec{R}) , \quad (4.13)$$

whence we infer that E is stationary when

$$\vec{R} = \vec{R}_{\min} = -2\xi \chi \vec{I} \quad (4.14)$$

and the stationary value of E is

$$E = \frac{1}{4} \text{Tr}[\frac{1}{2}(Q \vec{\Delta} + \vec{\Delta} Q) \vec{I}^{-1} + Q \vec{I}] , \quad (4.15)$$

where

$$Q = l_\mu n^\mu , \quad \vec{\Delta} = -D^\alpha D_\alpha + m^2 + \xi R - 2\xi Q^{-1} Q_{|\alpha}{}^\alpha + \xi Q^{-1} (2Q n^\nu n^\beta{}_{;\nu} + 2l^\nu n^\beta{}_{;\nu} + l^\nu \chi_\nu{}^\beta)_{|\beta} - 2\xi Q^{-1} l_\mu n_\nu R^{\mu\nu} - 4\xi^2 \chi^2 . \quad (4.16)$$

Under a further variation $\delta\vec{M} = i\delta\vec{I}$,

$$\delta E = -\frac{1}{4} \text{Tr}[\frac{1}{2} \vec{I}^{-1} (Q \vec{\Delta} + \vec{\Delta} Q) \vec{I}^{-1} \delta\vec{I} - Q \delta\vec{I}] . \quad (4.17)$$

Hence E is stationary when

$$\vec{I} = \vec{I}_{\min} = 2^{-1/2} (-Q)^{-1/2} [(-Q)^{-1/2} (-Q \vec{\Delta} - \vec{\Delta} Q) (-Q)^{-1/2}]^{-1/2} (-Q)^{-1/2} . \quad (4.18)$$

It is easily verified using the connecting relations (3.14) that in the case $l^\nu = n^\nu$ (and setting $\xi = 0$), the complex structure (3.1) coincides with that of Ashtekar and Magnon.⁶

The solution (4.13) exists whenever the square roots exist, which they will if and only if the operator $\vec{\Delta}$ has a positive spectrum. When $\xi = 0$ it always does, regardless of the geometry of the space-time, hypersurface \mathcal{S}_t , and the vector field l^μ . When $\xi \neq 0$ the positivity of $\vec{\Delta}$ depends on all these things.

In order to verify that the value (4.15) is indeed a minimum (and not a saddle point), we evaluate the second variations

$$\frac{\delta^2 E}{\delta R_{k_0 k_1} \delta I_{k_2 k_3}} \Big|_{\vec{R}_{\min}, \vec{I}_{\min}} = \frac{\delta^2 E}{\delta I_{k_0 k_1} \delta R_{k_2 k_3}} \Big|_{\vec{R}_{\min}, \vec{I}_{\min}} = 0, \quad (4.19)$$

$$\frac{\delta^2 E}{\delta R_{k_0 k_1} \delta R_{k_2 k_3}} \Big|_{\vec{R}_{\min}, \vec{I}_{\min}} = \frac{\delta^2 E}{\delta I_{k_0 k_1} \delta I_{k_2 k_3}} \Big|_{\vec{R}_{\min}, \vec{I}_{\min}} = \frac{1}{4} (Q_{k_0 k_3} \vec{I}_{\min k_1 k_2}^{-1} + Q_{k_0 k_3} \vec{I}_{\min k_1 k_2}^{-1})$$

at the turning point (4.14) and (4.18). Considered as a matrix with clumped indices $(k_0 k_1)$ and $(k_2 k_3)$, the second variations are positive definite whenever \vec{I}_{\min} exists and is negative definite. We conclude that the turning point is always a minimum.

In the case when \mathcal{M} is a Robertson-Walker space-time with metric

$$g_{\mu\nu} = a^2(t) \eta_{\mu\nu} \quad (4.20)$$

($\eta_{\mu\nu}$ is the metric of Minkowski space-time) with the \mathcal{S}_t chosen to be the hypersurfaces of homogeneity and $l^\nu = n^\nu$, we have

$$\vec{M}f(\underline{x}) = (2\pi)^{-3} \int \int d^3 \underline{k} d^3 \underline{y} \exp[i \underline{k} \cdot (\underline{x} - \underline{y})] \left[-6\xi \dot{a} / a^2 - \frac{i}{a} [k^2 + m^2 + 6\xi(1 - 6\xi)(a/a)^2]^{1/2} \right] f(\underline{y}). \quad (4.21)$$

This agrees with the results of Brown and Dutton and others^{7,1} for energy minimization performed directly. The resulting construction of states suffers from the Fulling pathology. It is conceivable that different choices of \mathcal{S}_t might remove the pathology, though this possibility must be viewed with little confidence since such choices would necessarily violate the symmetries present in the space-time.

ACKNOWLEDGMENTS

We are grateful to Professor A. Ashtekar for interesting conversations and to Dr. M. R. Brown for much helpful criticism during the course of this work.

- ¹J. Schwinger, in *Les Houches Lectures, 1955* (unpublished); *Phys. Rev.* **82**, 914 (1951); **91**, 713 (1953).
²S. A. Fulling, *Gen. Relativ. Gravit.* **10**, 807 (1979); L. Parker, *Phys. Rev.* **183**, 1057 (1969).
³R. M. Wald, *Ann. Phys. (N.Y.)* **118**, 490 (1979).
⁴B. S. Dewitt, in *Dynamical Theory of Groups and Fields* (Gordon and Breach, New York, 1965).
⁵S. A. Fulling, Ph.D thesis, Princeton University, 1972 (unpublished).
⁶A. Ashtekar and A. Magnon, *Proc. R. Soc. London* **A346**, 375 (1975).
⁷M. R. Brown and C. R. Dutton, *Phys. Rev. D* **20**, 3052 (1979).

- ⁸P. Candelas and D. J. Raine, *Phys. Rev. D* **15**, 1494 (1977).
⁹D. Deutsch, in *Quantum Gravity, A Second Oxford Symposium, 1981*, edited by C. J. Isham (Oxford University Press, London, 1982).
¹⁰A. A. Grib, S. G. Mamayev, and V. M. Mostepanenko, *Gen. Relativ. Gravit.* **7**, 535 (1976); D. J. Raine and C. P. Winlove, *Phys. Rev. D* **12**, 946 (1975).
¹¹R. Haag, *K. Dan. Vidensk. Selsk. Mat. Fys. Medd.* **29**, No. 12 (1955).
¹²M. R. Brown, A. C. Ottewill, and S. T. C. Siklos, *Phys. Rev. D* **26**, 1881 (1982).