Commun. math. Phys. 45, 9—34 (1975)

Communications in Mathematical Physics

On Particle Creation by Black Holes

Robert M. Wald*

Laboratory for Astrophysics and Space Research, University of Chicago, Chicago, Ill. 60637, USA

Abstract. Hawking's analysis of particle creation by black holes is extended by explicitly obtaining the expression for the quantum mechanical state vector Ψ which results from particle creation starting from the vacuum during gravitational collapse. (Hawking calculated only the expected number of particles in each mode for this state.) We first discuss the quantum field theory of a Hermitian scalar field in an external potential or in a curved but asymptotically flat spacetime with no horizon present. In agreement with previously known results, we find that we are led to a unique quantum scattering theory which is completely well behaved mathematically provided a certain condition is satisfied by the operators which describe the scattering of classical positive frequency solutions. In terms of these operators we derive the expression for the state vector describing particle creation from the vacuum, and we prove that S-matrix is unitary. Making the necessary modification for the case when a horizon is present, we apply this theory for a massless Hermitian scalar field to get the state vector describing the steady state emission at late times for particle creation during gravitational collapse to a Schwarzschild black hole. There is some ambiguity in the theory in this case arising from freedom involved in defining what one means by "positive frequency" at the future event horizon. However, it is proven that the expression for the density matrix formed from Ψ describing the emission of particles to infinity is independent of this choice, and thus unambiguous predictions for the results of all possible measurements at infinity are obtained. We find that the state vector describing particle creation from the vacuum decomposes into a simple product of state vectors for each individual mode. The density matrix describing emission of particles to infinity by this particle creation process is found to be identical to that of black body emission. Thus, black hole emission agrees in complete detail (i.e., not only in expected number of particles) with black body emission.

I. Introduction

In a recent paper, Hawking [1] analyzed the problem of particle creation caused by the gravitational collapse of a body to form a black hole. In this theory the

* Supported in part by the National Science Foundation under grant MPS 74-17456 to the University of Chicago.

gravitational field is treated classically, i.e. the spacetime geometry is taken to be that of a body undergoing complete gravitational collapse in general relativity. In this background geometry one quantizes the field under consideration (e.g. a scalar or Dirac field) in the same manner as in quantum field theory with an external potential present (see Ref. [2]). Thus, the analysis is semiclassical and is closely analogous to semiclassical radiation theory in ordinary quantum mechanics where one treats the electrons in an atom quantum mechanically but treats the perturbing electric field classically. The justification for making this semiclassical approximation here is that one believes that the quantum effects of the gravitational field are not likely to play an important role in the particle creation process. Of course, the question of whether or not this is indeed the case cannot be answered until one has a satisfactory quantum theory of the gravitational field. However, it seems very reasonable that the semiclassical analysis of particle creation by gravitational collapse should at the very least give a good indication of the type of effects which will occur in an exact quantum treatment.

Using this semiclassical treatment, Hawking calculated the expected number of particles $\langle N \rangle$ emitted in each mode at late times following the gravitational collapse of a body, assuming that at early times the quantized field was in the vacuum state. In the case of spherical gravitational collapse he obtained the truly remarkable result that there is a steady rate of emission of particles in each mode at late times, with $\langle N \rangle$ given by precisely the black body formula with temperature $kT = \hbar\kappa/2\pi$, where κ is the surface gravity of the black hole formed by the collapse.

However, there are many more properties of the final state of the quantized field than simply the expected number of emitted particles $\langle N \rangle$. The full information is contained in the state vector, or – if we are interested only in making measurements of particles which escape to infinity and not particles which go down the black hole – the density matrix formed from this state vector. The purpose of this paper is to obtain the complete description of the quantum mechanical particle creation effects by obtaining the explicit expression for this state vector.

The main result of the analysis is the following: *the density matrix for emission of particles to infinity at late times by spontaneous particle creation resulting from spherical gravitational collapse to a black hole is identical in all aspects to that of black body thermal emission at temperature* $kT = \hbar\kappa/2\pi$. The particles emitted to infinity at late times are completely uncorrelated with each other (although they are correlated with particles that enter the black hole at early times). The probability distribution for observing N particles in a given mode at infinity is identical to what one would obtain by starting with the black body Boltzmann distribution $\exp(-N\hbar\omega/kT)$ for emission of particles from the black hole and assuming that each particle has a probability of $|t|^2$ of reaching infinity, where t is the classical transmission amplitude (and thus $|t|^2$ is the absorption cross section of the black hole). Thus, if a black hole were placed in a thermal cavity at exactly the same temperature, there would be no way of determining where the black hole is by observing particle emission, since the black hole would emit exactly as much thermal radiation by spontaneous quantum particle creation as it would absorb by classical processes.

The state vector describing spontaneous particle creation from the vacuum is found to decompose into a simple product of state vectors for each individual mode. Each of these individual state vectors describes multiple pair creation in which one of the particles in each pair enters the black hole near the formation of the event horizon, while the other particle is emitted to infinity or gets scattered back into the black hole at late times. Thus, our expression for the state vector supports the intuitive picture of the particle creation process given by Hawking [1, 3].

In Section II we briefly review the quantum theory of free fields in order to define the basic quantities and introduce our notation. In Section III we discuss the quantum field theory of the Hermitian scalar field in an external potential or in curved, asymptotically flat spacetime with no horizon present. Most of the results of Section III are re-derivations of results which have previously appeared in the literature. In Section IV we discuss the modifications of the theory appropriate to the case where a horizon is present, and we apply the theory to particle creation resulting from spherical gravitational collapse. We obtain the explicit expression for the quantum mechanical state vector of the created particles and discuss its properties. The modifications to the emission which occur if the black hole is rotating and some further remaining issues are discussed in Section V.

II. Free Field Theory

In this section we shall briefly review the standard quantum theory of the free, real scalar field ϕ in order to establish the notation and framework of ideas for the following sections. The notation used here follows closely that used by Geroch [4].

One wants the Hilbert space of states in the quantum theory of the real scalar field to have particle interpretations. One would also like the quantized scalar field $\phi(x)$ to be a self-adjoint operator on this Hilbert space satisfying

$$(\Box + m^2)\phi(x) = 0 . \tag{2.1}$$

As is well known, however, the attempt to define ϕ for each point x in Minkowski space runs into serious mathematical difficulties. These difficulties are overcome by "smearing" ϕ with test functions f thus making ϕ an operator valued distribution $\phi(f)$. In place of Eq. (2.1) we require

$$\phi(g) = 0 \tag{2.2}$$

for all g of the form

$$g = (\Box + m^2) f \tag{2.3}$$

where f is a test function. [Equation (2.2) is just the integration by parts version of the smeared Eq. (2.1).] We now proceed to present the standard quantum theory of the real scalar field which, as is easily seen, satisfies the above minimal requirements.

The Hilbert space of one particles states $\mathscr{H}$ is taken to be $L^2(M_+)$ where M_+ is the positive mass shell [i.e. M_+ is the submanifold of (Fourier transformed) Minkowski space defined by $k^\mu k_\mu + m^2 = 0$ with k^μ future directed]. The Hilbert space of states is taken to be the symmetric Fock space $\mathscr{F}(\mathscr{H})$ defined by,

$$\mathscr{F}(\mathscr{H}) = \mathbb{C} \oplus \mathscr{H} \oplus (\mathscr{H} \otimes \mathscr{H})_s \oplus (\mathscr{H} \otimes \mathscr{H} \otimes \mathscr{H})_s \oplus \ldots$$

where the subscript s denotes the symmetric tensor product. Vectors in Fock space will be denoted by capital Greek letters. We shall use the following tensor-like notation: elements of the symmetrized tensor product of n copies of $\mathscr{H}$ will be denoted by lower case Greek letters with n latin upper indices. Elements of the dual (complex conjugate) Hilbert space $\bar{\mathscr{H}}$ and its tensor products will be denoted by barred Greek letters with the corresponding number of lower latin indices. ($\bar{\mathscr{H}}$ and $\mathscr{H}$ are, of course, naturally isomorphic but this isomorphism is antilinear. We shall explicitly deal with $\bar{\mathscr{H}}$ throughout this paper so that all the maps we consider will be linear.) Thus, for example, an element $\Phi \in \mathscr{F}(\mathscr{H})$ will be written

$$\Phi = (c, \xi^a, \xi^{ab}, \xi^{abc}, \ldots) \tag{2.4}$$

while an element of $\bar{\mathscr{H}}$ will be written $\bar{\sigma}_a$. A contraction of indices, e.g. $\xi^a \bar{\sigma}_a$, will denote the complex number obtained by applying $\bar{\sigma}_a$ to ξ^a. ($\xi^a \bar{\sigma}_a$ is, of course, the same as the scalar product of σ^a and ξ^a, where σ^a is the element fo $\mathscr{H}$ corresponding to the element $\bar{\sigma}_a$ of $\bar{\mathscr{H}}$ under the natural isomorphism.) When no confusion will arise, the upper or lower index will be omitted when writing elements of $\mathscr{H}$ or $\bar{\mathscr{H}}$ respectively, e.g. ξ denotes an element of $\mathscr{H}$, while $\bar{\tau}$ denotes an element of $\bar{\mathscr{H}}$.

For every element $\bar{\tau} \in \bar{\mathscr{H}}$ we define the annihilation oprator $a(\bar{\tau}): \mathscr{F}(\mathscr{H}) \to \mathscr{F}(\mathscr{H})$ as follows: for $\Phi \in \mathscr{F}(\mathscr{H})$ given by Eq. (2.4) we define,

$$a(\bar{\tau})\Phi = (\xi^a \bar{\tau}_a, \sqrt{2}\,\xi^{ab}\bar{\tau}_a, \sqrt{3}\,\xi^{abc}\bar{\tau}_a, \ldots)\,. \tag{2.5}$$

Similarly, for every $\sigma \in \mathscr{H}$ we define the creation operator $a^\dagger(\sigma): \mathscr{F}(\mathscr{H}) \to \mathscr{F}(\mathscr{H})$ by

$$a^\dagger(\sigma)\Phi = (0, c\sigma^a, \sqrt{2}\,\sigma^{(a}\xi^{b)}, \sqrt{3}\,\sigma^{(a}\xi^{bc)}, \ldots) \tag{2.6}$$

where the round brackets around the indices denotes the symmetrized tensor product. Then $a^\dagger(\sigma)$ is indeed the adjoint of $a(\bar{\sigma})$. Of course, $a(\bar{\tau})$ and $a^\dagger(\sigma)$ are unbounded operators, defined only on a dense domain.

Before proceeding further, it will be useful to establish some correspondences between solutions of the classical Klein-Gordon equation, states in $\mathscr{H}$, and test functions. Let F and G be two solutions of the classical Klein-Gordon equation, Eq. (2.1). The Klein-Gordon scalar product of F and G is defined by

$$(F, G)_{\mathrm{KG}} = i \int (\bar{F} \nabla_\mu G - G \nabla_\mu \bar{F}) d\Sigma^\mu \tag{2.7}$$

where Σ is an asymptotically flat spacelike hypersurface. The value of $(F, G)_{\mathrm{KG}}$ is independent of the choice of Σ by virtue of the Klein-Gordon equation. We note the following correspondences:

(I) Every positive frequency solution F of finite Klein-Gordon norm is associated in a one-to-one linear manner with an element σ_F of $\mathscr{H}$ via,

$$\hat{F}(k^\mu) = \sigma_F(k^\mu)\delta(k^\nu k_\nu + m^2) \tag{2.8}$$

where   denotes Fourier transform. Furthermore,

$$(F, G)_{\mathrm{KG}} = (\sigma_F, \sigma_G) \tag{2.9}$$

where the right-hand side of Eq. (2.9) denotes the Hilbert space scalar product.

(II) Similarly, every negative frequency solution $\bar{F}$ of finite Klein-Gordon norm is associated in a one-to-one linear manner with an element $\bar{\sigma}_F$ of $\bar{\mathscr{H}}$.

Furthermore,

$$(\bar{F}, \bar{G})_{KG} = -(\bar{\sigma}_F, \bar{\sigma}_G) = -(\sigma_G, \sigma_F) \tag{2.10}$$

where the middle term of Eq. (2.10) denotes the scalar product in $\bar{\mathscr{H}}$ and the right-hand side is the scalar product of the corresponding elements of $\mathscr{H}$.

(III) To every test function f we can linearly associate an element σ_f of $\mathscr{H}$ by Fourier transforming f and restricting $\hat{f}$ to the positive mass shell to get an element of $L^2(M_+)$.

We are now in a position to complete our discussion of the free field by defining the field operator $\phi(f)$ acting on $\mathscr{F}(\mathscr{H})$. For every test function f we define,

$$\phi(f) = a(\bar{\sigma}_f) + a^\dagger(\sigma_f) \tag{2.11}$$

where σ_f is the state associated with f by correspondence (III) above. Note that Eq. (2.2) is trivially satisfied because $\sigma_g = 0$ for any of the form $g = (\Box + m^2)f$ since $\hat{g}$ clearly vanishes on the mass shell. Finally, we observe that we can write Eq. (2.11) in a more familiar (though less elegant) form as follows: let $\{\sigma_i\}$ be an orthonormal basis of $\mathscr{H}$ and let $\{F_i\}$ be the corresponding positive frequency solutions of the Klein-Gordon equation. Write $a_i = a(\bar{\sigma}_i)$ and $a_i^\dagger = a^\dagger(\sigma_i)$. Then,

$$\phi = \sum_i (F_i a_i + \bar{F}_i a_i^\dagger) \tag{2.12}$$

where the meaning of Eq. (2.12) is to be understood as follows: for every test function f,

$$\phi(f) = a(\textstyle\sum_i \bar{\sigma}_i \int F_i f) + a^\dagger(\sum_i \sigma_i \int \bar{F}_i f) \tag{2.13}$$

where the integrals are taken over Minkowski space. It is not difficult to show that Eqs. (2.11) and (2.13) are identical.

III. Quantum Field Theory in an External Potential or Curved, Asymptotically Flat Spacetime (No Horizon)

In this section we shall consider the quantum field theory associated with the real scalar field ϕ satisfying,

$$(-\nabla_\mu \nabla^\mu + m^2 + V(x))\phi = 0 \tag{3.1}$$

where ∇_μ denotes the covariant derivative. In this theory both the spacetime curvature and external potential $V(x)$ are treated classically. For simplicity and definiteness we shall assume that both $V(x)$ and the spacetime curvature have compact support. We shall find that very minimal assumptions suffice to lead us to a unique quantum scattering theory. Furthermore, provided that a certain condition is satisfied, the theory is completely well behaved mathematically. Much of the material of this section is a re-working of results for the external potential case reported by Seiler [2]. (See also Fulling [5] and DeWitt [6].)

We proceed to construct this quantum field theory. By analogy with the free field case, we want a theory where ϕ is an operator valued distribution $\phi(f)$ acting on some Hilbert space of states, $\mathscr{F}$. However, we now want ϕ to satisfy Eq. (3.1) rather than Eq. (2.1) in the sense that now $\phi(g) = 0$ for all g of the form $g = (-\nabla_\mu \nabla^\mu +$

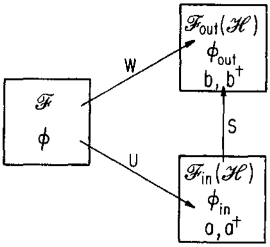

Fig. 1. The relationships between the Hilbert space of states $\mathscr{F}$ and the "in" and "out" Hilbert spaces $\mathscr{F}_{in}(\mathscr{H})$, $\mathscr{F}_{out}(\mathscr{H})$

$m^2+V)f$. Furthermore, in the distant past, prior to the interactions (i.e. more precisely, outside the future of the union of the supports of V and the curvature) we want the states of $\mathscr{F}$ to "look like" the states of the free field $\mathscr{F}(\mathscr{H})$ and we want ϕ to look like the free field operator. This is a minimal assumption if we are to have an interpretation of the theory in terms of scattering. We can state this assumption more precisely as follows: we require there to be an isomorphism (i.e. a unitary map) $U:\mathscr{F}\to\mathscr{F}_{in}(\mathscr{H})$ (where $\mathscr{F}_{in}(\mathscr{H})$ is a copy of the free field Hilbert space described in Section II above) such that for every test function f with support in the distant past, we have

$$U\phi(f)U^{-1}=\phi_{in}(f)=a(\bar{\sigma}_f)+a^{\dagger}(\sigma_f) \tag{3.2}$$

where ϕ_{in} is the free field operator on $\mathscr{F}_{in}(\mathscr{H})$ and we have denoted the annihilation and creation operators on $\mathscr{F}_{in}(\mathscr{H})$ by a and $a^{\dagger}$. Similarly, we require there to be another isomorphism W from $\mathscr{H}$ into another copy $\mathscr{F}_{out}(\mathscr{H})$ of the free field Hilbert space such that for any test function f with support in the distant future

$$W\phi(f)W^{-1}=\phi_{out}(f)=b(\bar{\sigma}_f)+b^{\dagger}(\sigma_f) \tag{3.3}$$

where we have denoted the annihilation and creation operators on $\mathscr{F}_{out}(\mathscr{H})$ by b and $b^{\dagger}$. The states of $\mathscr{F}_{in}(\mathscr{H})$ and $\mathscr{F}_{out}(\mathscr{H})$ are interpreted, respectively, as the incoming and outgoing particle states. The S-matrix, $S=WU^{-1}$, relates $\mathscr{F}_{in}(\mathscr{H})$ to $\mathscr{F}_{out}(\mathscr{H})$ and thus gives all the relevant information concerning scattering experiments. The relations between $\mathscr{F}$, $\mathscr{F}_{in}(\mathscr{H})$, and $\mathscr{F}_{out}(\mathscr{H})$ are summarized in Fig. 1.

We shall now show that the above, very minimal, assumptions already suffice to determine the S-matrix. The assumptions that ϕ satisfies Eq. (3.1) and agrees with ϕ_{in} in the past implies that ϕ takes the form,

$$U\phi U^{-1}=\textstyle\sum_i(G_i a_i+\bar{G}_i a_i^{\dagger}) \tag{3.4}$$

where the meaning of Eq. (3.4) is to be understood in the same manner as Eq. (2.12) above and where G_i is the solution of Eq. (3.1) which agrees in the past with the free field solution F_i appearing in Eq. (2.12). [The function G_i may be constructed by choosing a spacelike hypersurface which lies entirely outside the future of the support of the curvature and V and assigning the value and time derivative of F_i on that slice as initial data for a solution of Eq. (3.1).]

In a similar manner, we also must have,

$$W\phi W^{-1}=\textstyle\sum_j(H_j b_j+\bar{H}_j b_j^{\dagger}) \tag{3.5}$$

where H_i is the solution of Eq. (3.1) which agrees with F_i in the future. Combining Eqs. (3.4) and (3.5) we find,

$$S[\sum_i (G_i a_i + \bar{G}_i a_i^\dagger)]S^{-1} = \sum_j (H_j b_j + \bar{H}_j b_j^\dagger)\,. \tag{3.6}$$

Equation (3.6) implies,

$$S a_i S^{-1} = \sum_j \{(G_i, H_j)_{\mathrm{KG}} b_j + (G_i, \bar{H}_j)_{\mathrm{KG}} b_j^\dagger\} \tag{3.7}$$

where $(,)_{\mathrm{KG}}$ denotes the Klein-Gordon scalar product defined by Eq. (2.7). [The Klein-Gordon scalar product is easily seen to be independent of choice of spacelike hypersurface for two solutions of Eq. (3.1).]

We shall now define several operators on the Hilbert space $\mathscr{H}$ of single particle free field states which will enable us to re-write Eq. (3.7) in a much cleaner form [thus manifestly demonstrating that Eq. (3.7) is indeed both meaningful and basis independent]. Let F be a positive frequency solution of the free Klein-Gordon Eq. (2.1). Let H be the solution of Eq. (3.1) which agrees with F in the future. In the past, H will again agree with some solution of the free field Eq. (2.1). Decomposing this classical free field solution into its positive and negative frequency components, we find that we may uniquely write,

$$H = G' + \bar{G}'' \tag{3.8}$$

where G' and G'' are solutions of Eq. (3.1) which agree in the past with positive frequency free field solutions, denoted F' and F'', respectively. We define the operators $A: \mathscr{H} \to \mathscr{H}$ and $B: \mathscr{H} \to \bar{\mathscr{H}}$ by

$$A\sigma_F = \sigma_{F'} \tag{3.9}$$

$$B\sigma_F = \bar{\sigma}_{F''} \tag{3.10}$$

where we have made use of the correspondences between positive frequency classical free field solutions F and elements of $\mathscr{H}$ described in Section II above. In a similar manner, we define the operators $C: \mathscr{H} \to \mathscr{H}$ and $D: \mathscr{H} \to \bar{\mathscr{H}}$ by interchanging the roles of past and future in the definitions of A and B, respectively.

We shall assume that A, B, C, and D are everywhere defined bounded operators on $\mathscr{H}$, though we will not use this assumption in an essential way below (i.e. there does not appear to be any essential reason why the analysis could not be carried out even if these operators are only densely defined unbounded operators; the boundedness assumption is basically used only to ensure the existence of compositions of these operators). We now establish the properties of these operators which play a key role in the subsequent discussions.

In the notation of Eq. (3.8) we have,

$$(H_1, H_2)_{\mathrm{KG}} = (G_1' + \bar{G}_1'', G_2' + \bar{G}_2'')_{\mathrm{KG}}\,. \tag{3.11}$$

Using the independence of $(,)_{\mathrm{KG}}$ on the choice of slice and the fact that the scalar product of a positive frequency and a negative frequency free field solution vanishes, we obtain

$$(F_1, F_2)_{\mathrm{KG}} = (F_1', F_2')_{\mathrm{KG}} + (\bar{F}_1'', \bar{F}_2'')_{\mathrm{KG}}\,. \tag{3.12}$$

Thus, we obtain,

$$(\sigma_{F_1}, \sigma_{F_2}) = (A\sigma_{F_1}, A\sigma_{F_2}) - (B\sigma_{F_1}, B\sigma_{F_2}) \tag{3.13}$$

i.e. we have

$$A^{\dagger}A - B^{\dagger}B = I\,. \tag{3.14}$$

Similarly, the equation,

$$(\bar{H}_1, H_2)_{\mathrm{KG}} = (\bar{G}_1' + G_1'', G_2' + \bar{G}_2'') \tag{3.15}$$

implies,

$$B^{\dagger}\bar{A} = A^{\dagger}\bar{B} \tag{3.16}$$

where the operator $\bar{A}:\bar{\mathscr{H}} \to \bar{\mathscr{H}}$ is defined by $\bar{A}\bar{\tau} = (\overline{A\tau})$ for all $\bar{\tau} \in \bar{\mathscr{H}}$. In a similar manner, we obtain,

$$C^{\dagger}C - D^{\dagger}D = I \tag{3.17}$$

and,

$$D^{\dagger}\bar{C} = C^{\dagger}\bar{D}\,. \tag{3.18}$$

Finally, writing the decomposition analogous to Eq. (3.8) for a past positive frequency solution G,

$$G = H' + \bar{H}'' \tag{3.19}$$

we have,

$$(G, G' + \bar{G}'')_{\mathrm{KG}} = (H' + \bar{H}'', H)_{\mathrm{KG}} \tag{3.20}$$

which implies that for all states σ, τ,

$$(\sigma, A\tau) = (C\sigma, \tau) \tag{3.21}$$

i.e.

$$A = C^{\dagger}\,. \tag{3.22}$$

Note that Eqs. (3.14), (3.17), and (3.22) then imply that A^{-1} and C^{-1} exist as everywhere defined, bounded operators (see Ref. [7]). In a similar manner, we obtain,

$$B = -\bar{D}^{\dagger}\,. \tag{3.23}$$

Returning now to Eq. (3.7), it may be verified that Eq. (3.7) is just the basis expanded version of the following statement: for all states $\sigma \in \mathscr{H}$, we have,

$$Sa(\bar{\sigma})S^{-1} = b(\overline{C\sigma}) - b^{\dagger}(\overline{D\sigma})\,. \tag{3.24}$$

We are now in a position to solve for the image, $\Psi = S\Psi_0$, of the "in"-vacuum state $\Psi_0 \in \mathscr{F}_{\mathrm{in}}(\mathscr{H})$. Physically, Ψ gives us the complete information on particle creation from the vacuum. Setting $\tau = C\sigma$ and defining the operator $E:\bar{\mathscr{H}} \to \mathscr{H}$ by

$$E = \bar{D}\bar{C}^{-1} \tag{3.25}$$

we have for all $\tau \in \mathscr{H}$,

$$Sa(\overline{C^{-1}\tau})\Psi_0 = 0 = \{b(\bar{\tau}) - b^{\dagger}(E\bar{\tau})\}S\Psi_0\,. \tag{3.26}$$

Writing,

$$\Psi = (c, \eta^a, \eta^{ab}, \eta^{abc}, \eta^{abcd}, \ldots) \tag{3.27}$$

and explicitly writing out Eq. (3.26) component by component, we find for the first four terms,

$$\eta^a \bar{\tau}_a = 0\,, \tag{3.28a}$$

$$\sqrt{2}\eta^{ab}\bar{\tau}_a = c(E\bar{\tau})^b\,, \tag{3.28b}$$

$$\sqrt{3}\eta^{abc}\bar{\tau}_a = \sqrt{2}(E\bar{\tau})^{(b}\eta^{c)}\,, \tag{3.28c}$$

$$\sqrt{4}\eta^{abcd}\bar{\tau}_a = \sqrt{3}(E\bar{\tau})^{(b}\eta^{cd)}\,. \tag{3.28d}$$

But, the only way Eq. (3.28a) can be satisfied for all τ is to have $\eta^a = 0$. Equation (3.28c) then implies $\eta^{abc} = 0$. By induction, the amplitude for being in a state with an odd number of particles vanishes. In other words, particles are created in pairs. Equation (3.28b) states that E and η^{ab}, viewed as operators from $\bar{\mathscr{H}}$ into $\mathscr{H}$, must be proportional. However, this is possible only if the following conditions are satisfied: (1) since every two-particle state η^{ab} is symmetric, E must be a symmetric operator, $\bar{E}^\dagger = E$; (2) since, η^{ab} must have finite Hilbert space norm, we must have $\mathrm{tr}(E^\dagger E) < \infty$, i.e. E must be a Hilbert-Schmidt operator. If these conditions are not satisfied, there is no solution of Eq. (3.28b) (except for the trivial solution $\eta^{ab} = 0$ and $c = 0$ which will imply $\Psi = 0$; but this is unacceptable since S must be unitary). If the above conditions are satisfied, we may view E as an element of $(\mathscr{H} \otimes \mathscr{H})_s$. We shall denote this two-particle state associated with E as ε^{ab}. Equation (3.28b) yields,

$$\eta^{ab} = (c/\sqrt{2})\varepsilon^{ab}\,. \tag{3.29}$$

Equation (3.28d) then gives,

$$\eta^{abcd} = c((3 \cdot 1)/(4 \cdot 2))^{\frac{1}{2}} \varepsilon^{(ab}\varepsilon^{cd)} \tag{3.30}$$

and by induction, we obtain for the n-particle amplitude (n even),

$$\eta^{abcd\ldots yz} = c((2n)!^{\frac{1}{2}}/(2^n \cdot n!))\varepsilon^{(ab}\varepsilon^{cd}\ldots\varepsilon^{yz)}\,. \tag{3.31}$$

Assuming the above conditions on E are valid, one can show that the norm of Ψ is finite, so we may chose c to make $\|\Psi\| = 1$.

Thus, we have found the following: if the operator E – which is constructed entirely from the behavior of the classical solutions of Eq. (3.1) – does not satisfy either condition (1) or (2) above, no quantum field theory satisfying our minimal requirements exists. If E satisfies conditions (1) and (2), we have explicitly solved for the image $\Psi = S\Psi_0$ of the in-vacuum state. This solution is unique up to a phase factor. As will be seen below, the remainder of the S-matrix is also uniquely determined and a consistent theory exists.

Are conditions (1) and (2) satisfied? It is an immediate consequence of Eq. (3.18) and the definition of E that condition (1) (namely, $\bar{E}^\dagger = E$) must always be satisfied. Note that it is rather remarkable that this works out so well. If we had used a quantum field theory for spin zero with the "wrong" statistics – i.e. if we had used the free field theory of Section II but with the antisymmetric rather than the symmetric Fock space – the analysis could be carried through as before but when Eq. (3.28b) is reached we would have found that we needed E to be antisymmetric. Thus, there exists no reasonable quantum theory of spin zero particles with the "wrong" statistics in an external potential. This is independent of the much more subtle arguments which are ordinarily invoked to rule out theories with the "wrong" statistics.

Condition (2) has been proven to hold when one has only a potential (of compact support) by Seiler [2]. While we shall not investigate here whether condition (2) must also hold when spacetime curvature is present, we shall show below that condition (2) is equivalent to the condition of finiteness of the expectation value for the total number of particles created from the vacuum as computed by the algorithm used by Hawking.

Assuming now that condition (2) is satisfied, we demonstrate that the entire S-matrix is uniquely determined. Taking the adjoint of Eq. (3.24) we have,

$$Sa^{\dagger}(\sigma)S^{-1} = b^{\dagger}(C\sigma) - b(D\sigma)\,. \tag{3.32}$$

Applying both sides of this equation to the state $\Psi = S\Psi_0$, we obtain,

$$S(a^{\dagger}(\sigma)\Psi_0) = \{b^{\dagger}(C\sigma) - b(D\sigma)\}\Psi \tag{3.33}$$

and thus knowing Ψ, we find that the image under S of an arbitrary one-particle state σ is uniquely determined and explicitly given by the right-hand side of Eq. (3.33). By induction, the image of a simple n-particle state is also uniquely determined and since the various simple n-particle states are dense in $\mathscr{F}_{\text{in}}(\mathscr{H})$, S is uniquely determined.

Now that S has been defined, we can ask the crucial question necessary for the consistency of the theory: is S unitary? Since the scalar product of two states in $\mathscr{F}_{\text{in}}(\mathscr{H})$ having the form of a product of creation operators acting on Ψ_0 can be expressed in terms of the commutators of a and $a^{\dagger}$, it is easily seen that scalar products of these states – and, hence, scalar products of all states since these states are dense – will be preserved under S if and only if the commutators of the right-hand sides of Eqs. (3.24) and (3.32) are the same as those of the corresponding a and $a^{\dagger}$. But, we have

$$[a(\bar{\sigma}), a^{\dagger}(\varrho)] = (\sigma, \varrho)I \tag{3.34}$$

whereas,

$$\begin{aligned}
&[b(\overline{C\sigma}) - b^{\dagger}(\overline{D\sigma}), b^{\dagger}(C\varrho) - b(D\varrho)]\\
&= \{(C\sigma, C\varrho) - (\overline{D\varrho}, \overline{D\sigma})\}I\\
&= \{(C\sigma, C\varrho) - (D\sigma, D\varrho)\}I\\
&= (\sigma, \varrho)I
\end{aligned} \tag{3.35}$$

where we have used Eq. (3.17) in the last step. In a similar manner, using Eq. (3.18) we find that the commutators corresponding to $[a(\bar{\sigma}), a(\bar{\tau})]$ and $[a^{\dagger}(\sigma), a^{\dagger}(\tau)]$ both vanish, as they should. This, together with the fact that the range of S is dense in $\mathscr{F}_{\text{out}}(\mathscr{H})$, proves that S is indeed unitary. The unitarity of S completes the proof of the existence of a theory satisfying the requirements stated at the beginning of this section.

As a final remark, we note that the expected number of particles created from the vacuum (as well as everything else about this state) can be computed directly from our explicit solution for $\Psi = S\Psi_0$ given above. However, there is an alternative procedure for computing this quantity which is the one used by Hawking [1]. It follows from Eq. (3.24) and its adjoint Eq. (3.32) that the "time reversed" equation

$$S^{-1}b(\bar{\sigma})S = a(\overline{A\sigma}) - a^{\dagger}(\overline{B\sigma}) \tag{3.36}$$

must also hold. Applying both sides to the state Ψ_0 and taking norms, one obtains (using $S^{-1} = S^{\dagger}$),

$$(S\Psi_0, b^{\dagger}(\sigma)b(\bar{\sigma})S\Psi_0) = (\overline{B\sigma}, \overline{B\sigma}) = (\sigma, B^{\dagger}B\sigma)\,. \tag{3.37}$$

But the left-hand side of Eq. (3.37) is expectation value of the number of particles in the state σ which are created from the vacuum. This is the formula used by Hawking [1]. To get the expected total number of particles $\mathcal{N}$, we sum over an orthonormal basis and obtain,

$$\langle \mathcal{N} \rangle = \mathrm{tr}(B^{\dagger}B)\,. \tag{3.38}$$

But notice that $B = -\bar{D}^{\dagger}$ and C^{-1} is a bounded operator, so if $\mathrm{tr}(B^{\dagger}B) < \infty$ then $\mathrm{tr}(E^{\dagger}E) < \infty$, i.e. our condition (2) above is satisfied. Thus, if $\langle \mathcal{N} \rangle$ as computed by Eq. (3.38) is finite we are assured that our theory is free of all mathematical difficulties. Conversely, if C is bounded, the condition $\mathrm{tr}(E^{\dagger}E) < \infty$ implies that $\langle \mathcal{N} \rangle$ is finite.

In summary, we have found that the very minimal requirements stated at the beginning of this section for a quantum field theory in an external potential or in curved spacetime are sufficient to uniquely determine what the S-matrix must be. Furthermore, a consistent theory satisfying these requirements does indeed exist provided a certain condition [namely, $\mathrm{tr}(E^{\dagger}E) < \infty$] is satisfied by the operator E which is constructed from the behavior of the classical solutions. In the next section we shall apply this theory – making the necessary modifications to allow for the presence of a horizon – to the problem of particle creation occurring when a spherical body undergoes complete gravitational collapse and forms a black hole.

IV. Particle Creation by Gravitational Collapse

In this section we wish to extend the theory described in the previous section to the case treated by Hawking [1]. The spacetime we wish to consider is that of a body (e.g. a star) which undergoes complete gravitational collapse and forms a black hole. For simplicity, we shall assume that the collapse is spherically symmetric; in particular, this implies that the black hole formed by the collapse is a non-rotating Schwarzschild black hole. The modifications which occur when one drops the assumption of spherical symmetry are briefly discussed in Section V. In this curved spacetime geometry, we wish to consider the quantum field theory of a real scalar field ϕ. We shall consider only the massless case,

$$\nabla_{\mu}\nabla^{\mu}\phi = 0 \tag{4.1}$$

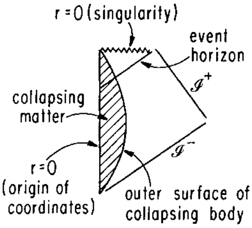

Fig. 2. A conformal diagram of a spacetime in which a body collapses and forms a black hole

since this will allow us to speak in precise terms of the asymptotic behavior of the field. [Solutions of the classical Eq. (4.1) in the spacetime we consider are presumably determined by their data at past null infinity $\mathscr{I}^-$ or by their data at future null infinity $\mathscr{I}^+$ and the future event horizon. This is not true of massive fields.] A spacetime diagram depicting gravitational collapse is given in Fig. 2.

As discussed in the previous section, the quantum theory should be such that the states of the quantum system are represented by vectors in a Hilbert space and the field ϕ is represented by an operator-valued distribution acting on this Hilbert space and satisfying Eq. (4.1). Furthermore, in the distant past the states of the system should asymptotically "look like" states in the free field Hilbert space $\mathscr{F}_{\text{in}}(\mathscr{H})$ and the field operator ϕ should approach the free field operator ϕ_{in}. As in the previous section, these assumptions lead to [see Eq. (3.4)],

$$U\phi U^{-1} = \sum_i (G_i a_i + \bar{G}_i a_i^\dagger) \tag{4.2}$$

where G_i is the solution of Eq. (4.1) with the same data at past null infinity as the free field solution F_i (where $a_i = a(\bar{\sigma}_{F_i})$).

However, in the asymptotic future the situation is different from that of the previous section. When no horizon is present all classical wave solutions of Eq. (4.1) propagate out to future null infinity and in the quantum theory it is natural to assume – as we did in Section III – that in the asymptotic future all states can be interpreted as free particles propagating out to infinity. However, when a horizon is present part of the classical waves can propagate into the black hole and never reach future null infinity. Hence, in this case it seems most natural to assume that states of the system have interpretations in terms of free particles at infinity and particles which have gone into the black hole. This suggests that the "out" Hilbert space in this case should be $\mathscr{F}_{\text{out}}(\mathscr{H} \oplus \mathscr{H}')$ where $\mathscr{H}$ is the usual single particle free field Hilbert space ("particles at infinity") and $\mathscr{H}'$ is the single particle Hilbert space of particles which have entered the black hole; i.e. the "out" Hilbert space should be the symmetric Fock space of the Hilbert space of all possible one-particle states.

Since vectors in $\mathscr{H}$ can be put into one-to-one correspondence with positive frequency data for classical solutions of Eq. (4.1) at future null infinity, it is natural to associate $\mathscr{H}'$ with positive frequency data for classical solutions of Eq. (4.1) at the horizon. The term "positive frequency" is well defined at future null infinity $\mathscr{I}^+$, since one has an asymptotic time translation parameter u defined there, with respect to which one can take Fourier transforms. [Note that time translations are well defined at $\mathscr{I}^+$ although there is considerable (supertranslation) ambiguity in defining asymptotic rotations and boosts.] In the case of a static, vacuum

Schwarzschild black hole (i.e. one not created by gravitational collapse) there is also a time translation parameter (i.e. Killing parameter) v running along the future horizon, thus enabling one to unambiguously define positive frequency on the horizon in that case. However, in the case of a black hole formed by gravitational collapse, the horizon is the static Schwarzschild horizon only outside of the collapsing matter in the case of exact spherical symmetry and it is only asymptotically stationary in a generic collapse. Thus, one does not have a time translation vector defined everywhere on the future horizon and this results in ambiguity in the definition of positive frequency. One can still define "positive frequency" as follows: choose a set of solutions K_i of Eq. (4.1) which vanish at $\mathscr{I}^+$, which are orthonormal (with positive norm) in the Klein-Gordon scalar product, and are such that the $\{K_i\}$ and their complex conjugates $\{\bar{K}_i\}$ span all solutions which vanish at $\mathscr{I}^+$. A solution of Eq. (4.1) will be called "positive frequency at the horizon" if it can be expressed as a sum of the $\{K_i\}$ (without using their complex conjugates). There is, of course, considerable ambiguity in the choice of the $\{K_i\}$. However, we shall show below that the predictions of the theory with regard to all measurements made at infinity will be independent of the choice of definition of positive frequency at the future event horizon.

If we accept the above arguments, it is natural to postulate that the field operator ϕ, when brought to the "out" Hilbert space by the isomorphism W, will take the form,

$$W\phi W^{-1} = \sum_i (H_i b_i + \bar{H}_i b_i^\dagger + K_i c_i + \bar{K}_i c_i^\dagger) \tag{4.3}$$

where H_i is the solution of Eq. (4.1) with the same data at $\mathscr{I}^+$ as F_i and vanishing data on the horizon, and where c_i and $c_i^\dagger$ are the annihilation and creation operators for the state in $\mathscr{H}'$ associated with the "positive frequency" solution K_i. It should be noted that the arguments leading to Eq. (4.3) are not as compelling as those leading to Eq. (4.2). Equations (4.2) and (4.3) are precisely Hawking's [1] Eqs. (2.3) and (2.4).

It is not clear what, if any, physical interpretation should be placed on the "horizon particle states". One might be tempted to say that if the quantum mechanical state is described by the vector σ' in $\mathscr{H}'$, then an "observer at the horizon" making the appropriate measurements would detect a single particle in the state σ'. However, there is no such thing as "an observer at the horizon" since the horizon is a null surface whereas all observers move on timelike worldlines. Furthermore, it is not at all clear what the "appropriate measurements" would be. In any case, one is primarily interested in the results of measurements made at infinity. To describe the results of all possible measurements at infinity, one constructs a density matrix in the following manner.

The "out" Hilbert space $\mathscr{F}_{\text{out}}(\mathscr{H} \oplus \mathscr{H}')$ is naturally isomorphic to the Hilbert space $\mathscr{F}(\mathscr{H}) \otimes \mathscr{F}(\mathscr{H}')$ as follows: let $\varrho^{ab\ldots z} \in (\bigotimes_n \mathscr{H})_s$ and let $\sigma'^{a'b'\ldots z'} \in (\bigotimes_m \mathscr{H}')_s$, i.e. $\varrho^{a\ldots z}$ is an n-particle "infinity state" and $\sigma'^{a'\ldots z'}$ is an m-particle "horizon state". The isomorphism associates each vector in $\mathscr{F}(\mathscr{H} \oplus \mathscr{H}')$ of the form

$$\varrho^{(a\ldots z}\sigma'^{a'\ldots z')} \tag{4.4}$$

with the vector

$$(n!m!/(n+m)!)^{\frac{1}{2}}\varrho^{a\ldots z} \otimes \sigma'^{a'\ldots z'} \tag{4.5}$$

in $\mathscr{F}(\mathscr{H})\otimes\mathscr{F}(\mathscr{H}')$. This map is easily seen to be norm preserving on these vectors, and since states of the form (4.4) are dense in $\mathscr{F}(\mathscr{H}\oplus\mathscr{H}')$ whereas states of the form (4.5) are dense in $\mathscr{F}(\mathscr{H})\otimes\mathscr{F}(\mathscr{H}')$, we have indeed defined an isomorphism. Let $\Phi\in\mathscr{F}(\mathscr{H}\oplus\mathscr{H}')$. To get the density matrix in $\mathscr{F}(\mathscr{H})$ associated with Φ we consider the state $\Phi\otimes\bar{\Phi}\in\mathscr{F}(\mathscr{H}\oplus\mathscr{H}')\otimes\overline{\mathscr{F}(\mathscr{H}\oplus\mathscr{H}')}$ and use the above isomorphism to view it as an element of $(\mathscr{F}(\mathscr{H})\otimes\overline{\mathscr{F}(\mathscr{H})})\otimes(\mathscr{F}(\mathscr{H}')\otimes\overline{\mathscr{F}(\mathscr{H}')})$. One then takes the trace of this element with respect to a basis of $\mathscr{F}(\mathscr{H}')$ to get a vector in $\mathscr{F}(\mathscr{H})\otimes\overline{\mathscr{F}(\mathscr{H})}$, which one views as an operator $\mathscr{D}:\mathscr{F}(\mathscr{H})\to\mathscr{F}(\mathscr{H})$, called the density matrix. If $\mathscr{A}:\mathscr{F}(\mathscr{H})\to\mathscr{F}(\mathscr{H})$ is any observable on $\mathscr{F}(\mathscr{H})$, its expectation value is given by,

$$\langle\mathscr{A}\rangle=\operatorname{tr}(\mathscr{A}\mathscr{D}) \tag{4.6}$$

so the density matrix $\mathscr{D}$ gives one the complete information concerning the results of measurements at infinity.

As noted above, there is ambiguity and arbitrariness in the choice of definition of positive frequency on the future event horizon. It would be very disturbing if any predictions of the theory with regard to measurements made at infinity were to depend upon this choice. However, the following argument establishes that the density matrix $\mathscr{D}$ for measurements at infinity is indeed independent of the definition of positive frequency on the horizon:

If we change the definition of positive frequency on the horizon, the annihilation and creation operators for the horizon states will undergo a transformation of precisely the same form as Eq. (3.24). This induces a linear transformation $\mathscr{S}$ on the "out" Hilbert space $\mathscr{F}(\mathscr{H})\otimes\mathscr{F}(\mathscr{H}')$. By the same arguments as given at the end of Section III, $\mathscr{S}$ is unitary. Furthermore, since only the annihilation and creation operators for $\mathscr{H}'$ were transformed, $\mathscr{S}$ will act only on $\mathscr{F}(\mathscr{H}')$, i.e. $\mathscr{S}$ takes a vector in $\mathscr{F}(\mathscr{H})\otimes\mathscr{F}(\mathscr{H}')$ of the form $\Gamma\otimes\Gamma'$ with $\Gamma\in\mathscr{F}(\mathscr{H})$, $\Gamma'\in\mathscr{F}(\mathscr{H}')$ into the vector $\Gamma\otimes\mathscr{S}\Gamma'$. But from these facts and from the definition given above of the density matrix $\mathscr{D}$ for $\mathscr{F}(\mathscr{H})$, it is clear that $\mathscr{D}$ remains unchanged when the "out" Hilbert space undergoes the transformation $\mathscr{S}$ resulting from a redefining of positive frequency at the horizon[1].

Thus, all predictions of the theory with regard to measurements made at infinity must be independent of the definition of positive frequency on the horizon. As a consequence, if we are concerned only with the results of measurements made at infinity, we can pick any convenient choice of $\{K_i\}$ for performing our calculations.

Our aim now is to calculate explicitly the state $\Psi\in\mathscr{F}_{\text{out}}(\mathscr{H}\oplus\mathscr{H}')$ which results when the complete gravitational collapse of a body occurs with no particles initially present (i.e. starting with the vacuum "in" state). The mathematical structure of the theory described above is the same as that of the previous section and we may directly take over the results of the analysis given there. Hence, we obtain,

$$\Psi=\Psi(\varepsilon^{ab})=c(1,0,2^{-\frac{1}{2}}\varepsilon^{ab},0,((3\cdot 1)/(4\cdot 2))^{\frac{1}{2}}\varepsilon^{(ab}\varepsilon^{cd)},0,\ldots) \tag{4.7}$$

[1] We should point out that for the strict validity of the above arguments the operators which describe the change in definition of positive frequency at the horizon must satisfy the analog of condition (2) of Section III (i.e. $\operatorname{tr}(E^{\dagger}E)<\infty$). If this condition is not satisfied, the argument given above is only a formal one.

where ε^{ab} is the 2-particle state associated with the operator $E=\overline{DC^{-1}}$, where the operators $A:(\mathscr{H}\oplus\mathscr{H}')\to\mathscr{H}$, $B:(\mathscr{H}\oplus\mathscr{H}')\to\overline{\mathscr{H}}$, $C:\mathscr{H}\to(\mathscr{H}\oplus\mathscr{H}')$, and $D:\mathscr{H}\to\overline{(\mathscr{H}\oplus\mathscr{H}')}$ are defined in the same manner as previously except that "in the past" now means "at $\mathscr{I}^-$" and "in the future" now means "at $\mathscr{I}^+$ and the future horizon". In order to explicitly determine ε^{ab} we must determine the action of these operators. Fortunately, the bulk of the analysis required for this purpose has already been carried out by Hawking [1].

Following Ref. [1], our first step is to introduce an orthonormal basis for $\mathscr{H}$ and an orthonormal basis for $\mathscr{H}'$ as follows: for each ω, l, m, let $P_{\omega lm}$ denote the solution generated by the data $\omega^{-\frac{1}{2}}\exp(i\omega u)Y_{lm}(\theta,\varphi)$ at future null infinity. Fix a real number E with $0<E\ll 1$ and define

$$P_{jnlm}=E^{-\frac{1}{2}}\int_{jE}^{(j+1)E}\exp(-2\pi in\omega/E)P_{\omega lm}d\omega\,. \tag{4.8}$$

Then the $\{P_{jnlm}\}$ with $j\geqq 0$ yield an orthonormal basis of $\mathscr{H}$ (i.e. they yield basis for all solutions generated from positive frequency data on $\mathscr{I}^+$ and they are orthonormal in the Klein-Gordon scalar product). The wave packets P_{jnlm} are made up of frequencies within E of $\omega=jE$. They are peaked around the retarded time $u=2\pi n/E$ and have a time spread $\sim 2\pi/E$. Hence, the following physical apparatus should give a good approximation to a measurement of the projection operator onto the state P_{jnlm}: a particle detector sensitive only to frequencies within E of $\omega=jE$ and angular dependence Y_{lm} which is turned on for a time interval $2\pi/E$ at time $u=2\pi n/E$. Thus, the use of the basis P_{jnlm} not only simplifies calculations but also allows one to get a direct physical interpretation of the state Ψ in terms of the outcomes of particle detection experiments. To conform to our previously established notation, we will use the symbol ${}_i\varrho^a$ to denote a typical element of this basis. Here the index i stands for $jnlm$ and the role of this index – which enumerates the members of the basis – should not be confused with that of the index a which tells us that ${}_i\varrho^a$ is an element of $\mathscr{H}$.

For the vacuum Schwarzschild solution one can construct a similar basis $\{Q_{jnlm}\}$ for the "horizon states" $\mathscr{H}'$ by precisely the same procedure starting from the solutions $Q_{\omega lm}$ generated by the data $\omega^{-\frac{1}{2}}\exp(i\omega v)Y_{lm}$ at the future horizon. In the case of a black hole formed by spherical gravitational collapse, we can also perform this construction after we define a time coordinate v on the horizon which agrees with the Killing parameter outside the collapsing matter. For large n (i.e. late times) the ambiguity in defining Q_{jnlm} resulting from the ambiguity in defining v will be negligible. We will use our freedom in choosing the elements of the positive frequency basis $\{K_i\}$ at the horizon so that the $\{Q_{jnlm}\}$ for large n form part of this basis. We will use the symbol ${}_i\sigma^a$ to denote such a basis element corresponding to Q_{jnlm}.

Next we construct new late times basis elements of $\mathscr{H}\oplus\mathscr{H}'$ as follows: consider the vacuum Schwarzschild solution and prescribe data for the solution X_{jnlm} at the *past* horizon in the same manner as we prescribed data for P_{jnlm} and Q_{jnlm} at $\mathscr{I}^+$ and the future horizon. Assuming that the transmission and reflection amplitudes $t=t_{lm}(\omega)$ and $r=r_{lm}(\omega)$ vary negligibly over the frequency interval E, we have

$$X_{jnlm}=tP_{jnlm}+rQ_{jnlm}\,. \tag{4.9}$$

Defining Y_{jnlm} in a similar manner at $\mathscr{I}^-$, we have

$$Y_{jnlm} = TQ_{jnlm} + RP_{jnlm}. \tag{4.10}$$

We define the new basis elements, ${}_i\lambda^a$ and ${}_i\gamma^a$ by,

$${}_i\lambda^a = t_i\,{}_i\varrho^a + r_i\,{}_i\sigma^a, \tag{4.11}$$

$${}_i\gamma^a = T_i\,{}_i\sigma^a + R_i\,{}_i\varrho^a. \tag{4.12}$$

Our aim now is to find the action of the operator DC^{-1} on these basis vectors. First, at late times (i.e. for large n), if we propagate the wave packet $(TQ_{jnlm} + RP_{jnlm})$ corresponding to the state ${}_i\gamma^a$ backward in time, it will be almost entirely scattered back to $\mathscr{I}^-$ by the static Schwarzschild geometry. Hence, it cannot pick up any negative frequency part and the resulting wave packet at $\mathscr{I}^-$ will be the purely positive frequency wave packet Y_{jnlm}. This implies,

$$DC^{-1}\,{}_i\gamma^a = 0. \tag{4.13}$$

On the other hand, at late times, the wave packet $(tP_{jnlm} + rQ_{jnlm})$ corresponding to ${}_i\lambda^a$ will be almost entirely scattered through the dynamically collapsing body and thence back to $\mathscr{I}^-$. The major effect which occurs is that the wave will suffer a very large blueshift upon entering the collapsing body (near the formation of the horizon). This blueshift will not be compensated by a correspondingly large redshift when the wave leaves the body since the body is in a less collapsed state at earlier times. Since the effective frequency of the wave is very high when it enters the collapsing body and propagates to $\mathscr{I}^-$, the geometrical optics approximation will be valid in this regime. Almost all of this wave packet will reach $\mathscr{I}^-$ just prior to the advanced time v_0 corresponding to the formation of the event horizon. In fact, in the geometrical optics approximation, it follows immediately from forming wave packets using Eq. (2.18) of Ref. [1], that the v dependence of the wave at $\mathscr{I}^-$ is given by

$$Z_{jnlm}(v) \sim \begin{cases} 0 & v > v_0 \\ \exp(-i\omega L/E)\sin(L/2)/L & v < v_0 \end{cases} \tag{4.14}$$

where $\omega = (j + \frac{1}{2})E$ is the effective frequency of the original wave packet at $\mathscr{I}^+$ and the future horizon, and

$$L = 2\pi n + (E/\kappa)\ln(v_0 - v) \tag{4.15}$$

where κ is the surface gravity of the black hole. (For a Schwarzschild black hole, $\kappa = 1/(4M)$ in geometrized units $G = c = 1$, where M is the mass of the black hole.)

For convenience, we shall now set $v_0 = 0$. As shown in Appendix A, the Fourier transform, $\hat{Z}_{jnlm}(\omega')$, of $Z_{jnlm}(v)$ satisfies the following relation for $\omega' > 0$ (assuming $E \ll 1$),

$$\hat{Z}_{jnlm}(-\omega') = -\exp(-\pi\omega/\kappa)\hat{Z}_{jnlm}(\omega'). \tag{4.16}$$

Consider now the "time reflected" wave packet $\tilde{Z}_{jnlm}$ at $\mathscr{I}^-$ given by

$$\tilde{Z}_{jnlm}(v) \sim \begin{cases} \exp(-i\omega\tilde{L}/E)\sin(\tilde{L}/2)/\tilde{L} & v > v_0 = 0 \\ 0 & v < v_0 = 0 \end{cases} \tag{4.17}$$

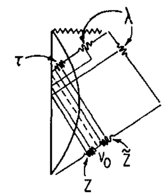

Fig. 3. The relationships between ${}_i\lambda^a$, ${}_i\tau^a$, Z_{jnlm}, and $\tilde{Z}_{jnlm}$. When the wave packet corresponding to the state ${}_i\lambda^a$ is propagated backward into the past, it gets scattered through the collapsing body and produces the data Z_{jnlm} at $\mathscr{I}^-$. The data $\tilde{Z}_{jnlm}$ at $\mathscr{I}^-$ is the time reflection of Z_{jnlm} about the advanced time v_0 corresponding to the formation of the event horizon. The wave packet defined by $\tilde{Z}_{jnlm}$ propagates into the black hole and defines the early time state ${}_i\tau^a$ as described in the text

where

$$\tilde{L} = 2\pi n + (E/\kappa)\ln v\,. \tag{4.18}$$

Since the $\{Z_{jnlm}\}$ are orthonormal in the Klein-Gordon scalar product, it is clear that the $\{\tilde{Z}_{jnlm}\}$ are also orthonormal, but with negative unit norm since the time reflection changes the sign of the Klein-Gordon scalar product. Furthermore, the scalar product of any $\tilde{Z}_{jnlm}$ with any Z_{jnlm} clearly vanishes. Also, the Fourier transform $\hat{\tilde{Z}}_{jnlm}(\omega')$ clearly satisfies,

$$\hat{\tilde{Z}}_{jnlm}(-\omega') = \hat{Z}_{jnlm}(\omega') \tag{4.19}$$

for all ω', since Eq. (4.19) holds for the time reflection of any function. Suppose now we propagate the wave packet $\tilde{Z}_{jnlm}$ into the future. The geometrical optics approximation will be valid as this wave packet propagates toward the collapsing body, since the effective frequency of $\tilde{Z}_{jnlm}$ is as high as Z_{jnlm}. The original wave packet Z_{jnlm} arrives at the center of the collapsing body just prior to the formation of the event horizon; it just barely escapes being captured by the newly formed black hole. However, the wave packet $\tilde{Z}_{jnlm}$ arrives just after the formation of the horizon and in the geometrical optics approximation it propagates entirely into the black hole. Let J_{jnlm} denote the data for this wave packet at the future event horizon. We shall use our freedom in defining positive frequency at the horizon to take the $\{\bar{J}_{jnlm}\}$ as part of our positive frequency basis $\{K_i\}$. We shall denote by ${}_i\tau^a$ the horizon state associated with the wave packet $\bar{J}_{jnlm}$. The relationships between ${}_i\lambda^a$, ${}_i\tau^a$, Z_{jnlm}, and $\tilde{Z}_{jnlm}$ are illustrated in Fig. 3.

It follows from Eqs. (4.16) and (4.19) that if we propagate the wave packet associated with the state $({}_i\lambda^a + \exp(-\pi\omega_i/\kappa){}_i\bar{\tau}_a)$ backward into the past, we obtain a purely positive frequency wave packet at $\mathscr{I}^-$. This implies,

$$DC^{-1}{}_i\lambda^a = \exp(-\pi\omega_i/\kappa){}_i\bar{\tau}_a\,. \tag{4.20}$$

Similarly, if we propagate the wave packet corresponding to $({}_i\bar{\lambda}_a + \exp(+\pi\omega_i/\kappa){}_i\tau^a)$ backward into the past we also get a purely positive frequency wave packet at $\mathscr{I}^-$. This implies

$$DC^{-1}(\exp(\pi\omega_i/\kappa){}_i\tau^a) = {}_i\bar{\lambda}_a\,. \tag{4.21}$$

From Eqs. (4.13), (4.20), and (4.21) it follows that

$$\varepsilon^{ab} = \textstyle\sum_i \exp(-\pi\omega_i/\kappa) 2\,{}_i\lambda^{(a}{}_i\tau^{b)} + \varepsilon_0^{ab} \tag{4.22}$$

where ε_0^{ab} is orthogonal to all the late time basis vectors $\{{}_i\lambda^a\}$ and $\{{}_i\gamma^a\}$ as well as the early time horizon states $\{{}_i\tau^a\}$. (Physically ε_0^{ab} gives the pair creation part of Ψ

which reaches infinity at early times while the summed term gives the final steady state emission.)

Equation (4.22), together with Eq. (4.7), gives the solution for the state vector Ψ which results from particle creation starting from the vacuum during gravitational collapse. The task that remains is to interpret our solution and derive its properties. However, we first should comment that, as noted in the previous section, $E = \overline{DC}^{-1}$ must be a Hilbert-Schmidt operator in order that ε^{ab} and $\Psi(\varepsilon^{ab})$ be normalized states and hence that the theory make rigorous mathematical sense. But it is clear from Eq. (4.22) that in this case ε^{ab} does not have finite norm since (recalling that the index i stands for $jnlm$) even for fixed j and m one has terms of finite norm (bounded away from zero) for infinitely many n and l. This infinity we get in the norm of ε^{ab} results from the steady rate of emission in all modes over all time. However, as we shall see very shortly below, if we restrict attention to measurements of a single or finite number of modes, then in a well defined sense the infinite norm factor (due to the infinity of other modes) factors out and we can obtain well defined predictions. Thus, the mathematical difficulties one encounters here are very minor compared with the difficulties one ordinarily encounters in the quantum theory of fields.

The fact which permits us to reduce the state vector $\Psi(\varepsilon^{ab})$ to a form where it can be easily interpreted is the following:

Lemma. *Let $\mathscr{H}_1$ and $\mathscr{H}_2$ be Hilbert spaces and let $\psi^{ab} \in (\mathscr{H}_1 \otimes \mathscr{H}_1)_s$, $\eta^{ab} \in (\mathscr{H}_2 \otimes \mathscr{H}_2)_s$. Consider the state $\Phi(\mu^{ab}) \in \mathscr{F}(\mathscr{H}_1 \oplus \mathscr{H}_2)$ defined by*

$$\Phi(\mu^{ab}) = (1, 0, 2^{-\frac{1}{2}}\mu^{ab}, 0, ((3 \cdot 1)/(4 \cdot 2))^{\frac{1}{2}}\mu^{(ab}\mu^{cd)}, 0, \ldots)$$

where $\mu^{ab} = \psi^{ab} + \eta^{ab}$. Then under the natural isomorphism (discussed above) between $\mathscr{F}(\mathscr{H}_1 \oplus \mathscr{H}_2)$ and $\mathscr{F}(\mathscr{H}_1) \otimes \mathscr{F}(\mathscr{H}_2)$ the state $\Phi(\mu^{ab})$ is mapped into the simple product state $\Phi_1(\psi^{ab}) \otimes \Phi_2(\eta^{ab})$ where

$$\Phi_1(\psi^{ab}) = (1, 0, 2^{-\frac{1}{2}}\psi^{ab}, 0, ((3 \cdot 1)/(4 \cdot 2))^{\frac{1}{2}}\psi^{(ab}\psi^{cd)}, 0, \ldots),$$
$$\Phi_2(\eta^{ab}) = (1, 0, 2^{-\frac{1}{2}}\eta^{ab}, 0, ((3 \cdot 1)/(4 \cdot 2))^{\frac{1}{2}}\eta^{(ab}\eta^{cd)}, 0, \ldots).$$

The proof of this lemma is straightforward, but since it plays a key role in our analysis we give an explicit proof in Appendix B.

We now apply this lemma to our state vector $\Psi(\varepsilon^{ab})$ setting $\mathscr{H}_1$ equal to the two-dimensional Hilbert space generated by the vectors ${}_i\lambda^a$ and ${}_i\tau^a$ and setting $\mathscr{H}_2 = (\mathscr{H}_1)^\perp$. We obtain,

$$\Psi(\varepsilon^{ab}) = \Psi_i(\exp(-\pi\omega_i/\kappa)2{}_i\lambda^{(a}{}_i\tau^{b)}) \otimes \Psi(\tilde{\varepsilon}^{ab}) \tag{4.23}$$

where

$$\begin{aligned}\Psi_i = (&1, 0, 2^{-\frac{1}{2}}\exp(-\pi\omega_i/\kappa)2{}_i\lambda^{(a}{}_i\tau^{b)}, 0,\\ &((3 \cdot 1)/(4 \cdot 2))^{\frac{1}{2}}\exp(-2\pi\omega_i/\kappa)4{}_i\lambda^{(a}{}_i\tau^{b}{}_i\lambda^{c}{}_i\tau^{d)}, \ldots)\end{aligned} \tag{4.24}$$

and where $\tilde{\varepsilon}^{ab}$ is defined by the same Eq. (4.22) as ε^{ab} except that the single term $\exp(-\pi\omega_i/\kappa)2{}_i\lambda^{(a}{}_i\tau^{b)}$ is omitted in the sum. Suppose now that we are interested only in measuring emission in the i-th mode. Since the state vector $\Psi(\varepsilon^{ab})$ is of the form of a simple product of a state vector Ψ_i for the i-th mode and a state vector for modes orthogonal to the i-th mode, the density matrix for emission in the i-th mode is the same as that of the pure state vector Ψ_i. In other words, *emission in the various modes is independent, i.e. there are no correlations between measurements of particles emitted in different modes. Each mode has its own state vector Ψ_i, defined by Eq.* (4.24).

Continuing the reduction process on $\Psi(\tilde{\varepsilon}^{ab})$, we may symbolically express $\Psi(\varepsilon^{ab})$ as,

$$\Psi(\varepsilon^{ab}) = (\bigotimes_i \Psi_i) \otimes (\bigotimes_k (\Psi_0)_k) \otimes \Psi(\varepsilon_0^{ab}) \tag{4.25}$$

where $(\Psi_0)_k$ is the vacuum state of the Fock space generated by the one-dimensional Hilbert space spanned by ${}_k\gamma^a$. Thus, the state vector describing particle creation during gravitational collapse decomposes into a product of a state vector $\Psi(\varepsilon_0^{ab})$ describing the early time emission multiplied by a product of state vectors describing emission in the various modes at late times. The natural physical interpretation of each state vector Ψ_i is that it describes multiple pair creation in which one of the particles (namely ${}_i\tau^a$) in each pair enters the black hole just after its formation while the other particle in the pair (namely, ${}_i\lambda^a = t_i\,{}_i\varrho^a + r_i\,{}_i\sigma^a$) reaches infinity (with amplitude t_i) or gets scattered back into the black hole (with amplitude r_i) at late times. Of course, it should be kept in mind that we do not claim any physical interpretation of "particles which go down the black hole" in terms of measurements which observers near the black hole can make. Indeed, the concept of "particles which go down the black hole" depends on the quite arbitrary choice of definition of positive frequency at the future event horizon.

As discussed above, our main interest is in describing the emission of particles that reach infinity. We have a physical interpretation of "particles at infinity" and we have already shown that the density matrix describing emission to infinity is independent of the choice of definition of positive frequency on the future horizon. To find the density matrix for particles in the i-th mode which reach infinity, we view $\Psi_i \otimes (\Psi_0)_i$ as an element of $\mathscr{F}(\mathscr{H}_i) \otimes \mathscr{F}(\mathscr{H}'_i)$ where $\mathscr{H}_i$ is the one-dimensional Hilbert space spanned by ${}_i\varrho^a$ and $\mathscr{H}'_i$ is the two-dimensional Hilbert space spanned by ${}_i\sigma^a$ and ${}_i\tau^a$. We then find the density matrix by "tracing out" the degrees of freedom corresponding to $\mathscr{H}'_i$ in the manner described above.

Let us first calculate the probability P_N for observing N particles at infinity for the simple case of a mode whose transmission amplitude t_i is unity, so that ${}_i\lambda^a = {}_i\varrho^a$. In this case, P_N is simply proportional to the squared norm of the vector appearing in the $2N$-particle entry in the expression for Ψ_i, Eq. (4.24), since exactly half of the particles are emitted to infinity in this case. We obtain,

$$\begin{aligned} P_N &\propto ((2N-1)(2N-3)\ldots 1)/((2N)(2N-2)\ldots 2) \\ &\cdot \exp(-N2\pi\omega/\kappa) 2^{2N} \| {}_i\varrho^{(a}\,{}_i\tau^b \ldots {}_i\varrho^y\,{}_i\tau^{z)} \|^2 \\ &= ((2N-1)(2N-3)\ldots 1)/((2N)(2N-2)\ldots 2) \\ &\quad \cdot \exp(-N2\pi\omega/\kappa) 2^{2N} (N!)(N!)/(2N)! \\ &= \exp(-N2\pi\omega/\kappa)\,. \end{aligned} \tag{4.26}$$

Thus, the probability of observing N particles at infinity in a mode with $t_i = 1$ is given by precisely the Boltzmann factor corresponding to the temperature T

given by

$$kT = \hbar\kappa/2\pi\,. \tag{4.27}$$

A further simple calculation reveals that the density matrix for emission to infinity in a mode with $t_i = 1$ is given by,

$$\mathscr{D}_i = \sum_N \exp(-N\hbar\omega/kT)\psi_N \otimes \bar{\psi}_N \tag{4.28}$$

where

$$\psi_N = \bigotimes_N {}_i\varrho^a\,. \tag{4.29}$$

Equation (4.28) *is precisely the density matrix for black body thermal radiation in the i-th mode.*

To find the density matrix (and probability distribution P_N) for emission to infinity in an arbitrary mode ($t_i \neq 1$), one must replace ${}_i\varrho^a$ in the above formula by ${}_i\lambda^a = t_i\,{}_i\varrho^a + r_i\,{}_i\sigma^a$ and "trace out" the degrees of freedom corresponding to ${}_i\sigma^a$. The result one obtains is precisely what one would get by starting with the previous probability distribution and density matrix for the case of unit transmission amplitude and assuming that each emitted particle has a probability of $|t_i|^2$ of reaching infinity. Since $|t_i|^2$ is the classical absorption cross section of the black hole for the given mode, this means that a black hole placed in a thermal cavity at temperature $kT = \hbar\kappa/2\pi$ would absorb precisely as much thermal black body radiation from the cavity as it would emit via quantum particle creation effects. Thus, a black hole placed in a thermal cavity at the same temperature would be in exact (though, perhaps, unstable) equilibrium; by measuring only particle emission, it would be impossible to tell where in the cavity the black hole is located. Thus, a black hole is a true black body emitter.

We should perhaps emphasize how remarkable it is that our calculation yields exactly the thermal black body density matrix. The quantum spontaneous particle creation process always produces particles in a pure state; it is only because some of these particles go down the black hole that one gets a mixed state for emission to infinity. Even so, it is quite remarkable that one gets a steady rate of uncorrelated emission at late times. Perhaps even more remarkable is the cancellation of the numerical factors in Eq. (4.26) to yield precisely the black body Boltzmann factor for the probability distribution P_N. (There are, of course, many other probability distributions whose first moment $\langle N \rangle$ agrees with black body emission.) It seems difficult to believe that the exact agreement of black hole emission with black body emission can be merely a chance coincidence; the possibility that there may be a deep reason why black holes are black body radiators appears worthy of future investigation.

V. Discussion

In this paper we have obtained the explicit solution for the quantum mechanical state vector which describes the massless scalar particle creation from the vacuum occurring when a spherical body undergoes complete gravitational collapse and forms a black hole. The results of all possible measurements on these particles at late times at infinity can be obtained from our solution. We found that the spontaneous quantum particle emission by a black hole agrees in complete detail with thermal emission by a black body at temperature $kT=\hbar\kappa/2\pi$. In this final section we shall discuss some remaining issues.

In the analysis of Section IV above, we considered only the case of spherical gravitational collapse to a Schwarzschild black hole. However, by the same arguments as given in Ref. [1], our results for the final steady state emission do not depend on the details of the collapse process but only on the final black hole produced by the collapse. Hence, the final steady state emission for any gravitational collapse which results in Schwarzschild black hole will be described by the state vector Ψ derived above. For the case when the collapse results in the formation of a Kerr black hole, the state vector is modified as follows: for non-superradiant modes $\omega > m\Omega_H$ (where ω is the frequency, m is the azimuthal quantum number and Ω_H is the angular velocity of the horizon) we may define our basis of $\mathscr{H} \oplus \mathscr{H}'$ as before. The arguments given in Ref. [1] show that the results of Section IV still hold, with the modification that the frequency ω_i of the i-th mode is replaced by $(\omega_i - m\Omega_H)$. Hence, for the non-superradiant modes all the discussion and formulae of Section IV apply, provided one makes the substitution $\omega \to (\omega - m\Omega_H)$ in the exponential factor occurring in these formulae. In the superradiant regime $0<\omega<m\Omega_H$, the solutions Q_{jnlm} have negative Klein-Gordon norm, so one must use the complex conjugates of these solutions rather than the solutions themselves to construct the late time positive frequency basis ${}_i\sigma^a$ for the horizon states. (This is equivalent to defining positive frequency at the horizon at late times via the Killing vector which points along the generators of the horizon rather than the Killing vector which is timelike at infinity.) Similarly, the wave packets J_{jnlm} now have positive Klein-Gordon norm, so one must use them (rather than their complex conjugates) to define the early time horizon states ${}_i\tau^a$. In place of Eq. (4.13) one obtains,

$$DC^{-1}(R_i\,{}_i\varrho^a) = T_i\,{}_i\bar{\sigma}_a\,. \tag{5.1}$$

In place of Eqs. (4.20) and (4.21), we now obtain,

$$DC^{-1}(t_i\,{}_i\varrho^a + \exp(\pi(m\Omega_H - \omega)/\kappa)_i\tau^a) = r_{ii}\bar{\sigma}_a\,, \tag{5.2}$$

$$DC^{-1}(\bar{r}_i\,{}_i\sigma^a) = \bar{t}_i\,{}_i\bar{\varrho}_a + \exp(-\pi(m\Omega_H - \omega)/\kappa)_i\bar{\tau}_a\,. \tag{5.3}$$

Note that in the superradiant regime we have $|R_i|^2 - |T_i|^2 = 1 = |r_i|^2 - |t_i|^2$. Equations (5.1), (5.2), and (5.3) yield,

$$\varepsilon^{ab}_{\text{superradiant}} = \sum_i (1/r_i) 2\,{}_i\sigma^{(a}[t_i\,{}_i\varrho + \exp(-\pi(m\Omega_H - \omega)/\kappa)_i\tau]^{b)} + \varepsilon_0^{ab} \tag{5.4}$$

where ε_0^{ab} is orthogonal to the late time superradiant modes as well as the early time horizon states $\{{}_i\tau^a\}$ associated with superradiant modes. One may analyze $\Psi(\varepsilon^{ab})$ by decomposing it into a product of individual state vectors for each mode as in Section IV. The particle emission to infinity is again uncorrelated but now it is not thermal. The expected emission to infinity in each mode is given by,

$$\langle N_i \rangle = |t_i|^2/(1 - \exp(-2\pi(m\Omega_H - \omega)/\kappa)) \tag{5.5}$$

in agreement with Hawking's [1] result. In the limit $\kappa \to 0$ corresponding to purely superradiant emission we have

$$\varepsilon^{ab}_{\text{superradiant}} \to \sum_i (t_i/r_i) 2\, {}_i\sigma^{(a}\, {}_i\varrho^{b)} + \varepsilon^{ab}_0 \,. \tag{5.6}$$

The natural physical interpretation of the state vector Ψ_i for the i-th mode in this case is that pairs of particles are being created at late times, with one of the particles going into the black hole and the other escaping to infinity. The probability of observing N particles in a superradiant mode at infinity in the limit $\kappa \to 0$ is given by

$$P_N \propto |t_i|^{2N}/|r_i|^{2N} \,. \tag{5.7}$$

We considered only the case of a massless scalar field in Section IV. However, Hawking [1] has argued that his results are also valid in the massive case, the only modification being that the frequency ω now includes a rest mass contribution. These arguments apply with equal validity to our analysis, so the expression for Ψ given in Section IV applies to the massive scalar case as well. One can also analyze the emission of charged scalar particles by a charged black hole after one obtains the analog of our solution for $S\Psi_0$, Eq. (3.31), for the case of a charged scalar field (see Ref. [2]). Presumably, a similar analysis could be carried out for the case of spin $\frac{1}{2}$ and spin 1 fields. However, for higher spin fields many difficulties arise when one attempts to construct a quantum field theory of these fields in an external potential or in curved spacetime (see Ref. [2] and the other articles in that book).

As described in the introduction, in the theory we have used here one presupposes a fixed spacetime geometry corresponding to a body undergoing gravitational collapse and calculates the particle emission in that spacetime. However, one certainly expects that the particle emission itself will affect the spacetime geometry. The calculation of the magnitude and nature of this back-reaction effect of particle creation on the spacetime metric is of considerable importance and interest. An understanding of this effect in the case of gravitational collapse would yield considerable insight into the similar effects which one expects to occur from particle creation in the early universe. Furthermore, it is important that the magnitude of the back-reaction effect be calculated in order to check the consistency of the theory we have used, for if particle creation causes large local changes in the metric we cannot expect our approach of calculating the particle creation in a fixed background geometry to be valid.

The simplest approach [1] to the back-reaction problem is to assume that the major back-reaction effect of particle creation is to cause the mass of the black hole to decrease at precisely the rate necessary to compensate for the expected energy flux at infinity of the created particles. Since the expected energy flux at infinity is proportional to (area) $xT^4 \sim M^2\kappa^4 \sim 1/M^2$ where M is the mass of the black hole, this is a runaway process, i.e. a mass decrease of the black hole causes an increased energy flux at infinity and hence an increased mass loss rate. Putting in the correct factors, one finds [1] that any black hole created in the early universe with mass $\lesssim 10^{15}$ grams would have completely evaporated by now, with perhaps 10^{30}–10^{35} ergs of its mass energy being emitted in a final flash of less than one second.

A more sophisticated approach to the back-reaction problem involves the calculation of the expectation value $\langle T_{\mu\nu}\rangle$ of the stress energy tensor of the created matter. One substitutes $\langle T_{\mu\nu}\rangle$ in Einstein's equations,

$$G_{\mu\nu}=8\pi(T_{\mu\nu}+\langle T_{\mu\nu}\rangle)$$

(where $T_{\mu\nu}$ denotes the stress energy tensor of the ordinary matter undergoing gravitational collapse) and seeks a self-consistent solution for $\langle T_{\mu\nu}\rangle$ and the spacetime metric, analogous to self-consistent approaches for finding the electron wave functions and the electromagnetic field of atoms with many electrons. Unfortunately, the formal expression for $\langle T_{\mu\nu}\rangle$ is infinite and must be regularized. Procedures for regularizing $\langle T_{\mu\nu}\rangle$ are discussed in Refs. [6] and [8]. Perhaps the approach and results of Section IV above will yield some insight into the regularization problem, since from the state vector Ψ one can unambiguously calculate finite values of $\langle T_{\mu\nu}\rangle$ at late times at infinity. Furthermore, one may be able to determine the energy density and energy flux at the future horizon from Ψ if one can obtain a valid physical interpretation of the horizon states.

One of the most intriguing issues concerning particle creation by black holes is the possibility of a deep connection between entropy and black hole surface area. In ordinary thermodynamics, the total entropy of the universe must increase with time. However, when black holes are present there are difficulties with this law because there is no way of determining the entropy of matter which has fallen into a black hole. One can define the total entropy of matter outside of black holes, but this quantity need not always increase, since it decreases everytime some matter falls into a black hole. On the other hand, in classical general relativity one has the law that the total surface area of all black holes must increase with time [9]. Indeed, the laws of classical black hole mechanics can be formulated in complete analogy with the laws of thermodynamics, with black hole area playing the role of entropy, and surface gravity κ playing the role of temperature [10]. However, on account of the quantum particle creation process, this classical law of black hole area increase will be violated according to the simple back-reaction estimate described above. But notice that black hole area will decrease only at the expense of particle creation and thus an increase in the entropy of matter outside of black holes. Furthermore, a decrease in the entropy of matter outside black holes due to matter falling into a black hole occurs at the expense of an increase in black hole area. This suggests a new law of physics (first proposed by Bekenstein [11] prior to the particle creation analysis): the total generalized entropy of the universe always increases with time, where the generalized entropy is the sum of the ordinary entropy of all matter outside of black holes plus an appropriate numerical factor times the surface area of all black holes. The fact proven in Section IV that quantum black hole emission agrees in complete detail with black body thermal emission greatly strenghtens the possibility that there may be deep significance in this generalized second law of thermodynamics.

Many fundamental questions remain for future investigations: is there indeed a deep connection between entropy and black hole area, or is the above analogy merely an accidental quirk? Is conservation of baryons and leptons violated in the collapse and particle creation process? It appears that it is since we can form a black hole purely out of baryons; but when this black hole evaporates by particle creation effects, it apparently produces equal numbers of baryons and anti-baryons. What happens when a black hole evaporates completely? Does it leave behind empty space or a singularity or perhaps even a more exotic object? It is possible that we may have to wait until we have a complete quantum theory of gravitation coupled to other fields before we can obtain satisfactory answers to these questions.

Acknowledgements. It is a pleasure to thank Robert Geroch for countless stimulating discussions throughout the course of this investigation. I also wish to thank my colleagues at the Seventh Texas Symposium for many valuable conversations which took place during that meeting. Special thanks are due William Unruh and an anonymous referee for pointing out an error in an earlier version of this manuscript.

Appendix A: Fourier Transform of $Z_{jnlm}(v)$

Consider the function [see Eq. (4.14)],

$$Z(v)=\begin{cases}0 & v>0\\ \exp(-i\omega L/E)\sin(L/2)/L & v<0\end{cases} \tag{A.1}$$

where

$$L=2\pi n+(E/\kappa)\ln(-v) \tag{A.2}$$

and $E \ll 1$. We show here that for $\omega'>0$ the Fourier transform, $\hat{Z}$, of Z satisfies

$$\hat{Z}(-\omega')=-\exp(-\pi\omega/\kappa)\hat{Z}(\omega') \tag{A.3}$$

Let $\omega'>0$. We have,

$$\begin{aligned}\hat{Z}(\omega')&=\int_{-\infty}^{\infty}\exp(-i\omega' v)Z(v)dv\\ &=\int_{-\infty}^{0}\exp(-i\omega' v-i\omega L/E)\sin(L/2)/L\,dv\,.\end{aligned} \tag{A.4}$$

Substitute $x=i\omega' v$. Then the integral is taken over the negative imaginary axis in the complex x-plane. If we define the logarithm function to be analytic in the first quadrant (i.e. if we choose the branch cut to lie outside the first quadrant), the integrand will be analytic in the fourth quadrant of the x-plane. We may close the contour in this quadrant to express $\hat{Z}(\omega')$ as a contour integral along the positive real axis,

$$\hat{Z}(\omega')=\int_0^{\infty}\exp(-x)\exp(-i\omega L_x/E)\sin(L_x/2)/(i\omega' L_x)dx \tag{A.5}$$

where

$$\begin{aligned}L_x&=2\pi n+(E/\kappa)\ln(ix/\omega')\\ &=2\pi n+(E/\kappa)(i\pi/2+\ln x-\ln\omega')\,.\end{aligned} \tag{A.6}$$

Thus,

$$\begin{aligned}\hat{Z}(\omega')&=(i\omega')^{-1}\exp(-i\omega 2\pi n/E+i(\omega/\kappa)\ln\omega')\\ &\quad\cdot\exp(\pi\omega/2\kappa)\int_0^{\infty}\exp(-x-i(\omega/\kappa)\ln x)\sin(L_x/2)/L_x dx\,.\end{aligned} \tag{A.7}$$

On the other hand,

$$\hat{Z}(-\omega')=\int_{-\infty}^{0}\exp(+i\omega'v-i\omega L/E)\sin(L/2)/L\,dv\,. \tag{A.8}$$

Substituting $y=-i\omega'v$ and proceeding in an exactly similar manner, we obtain,

$$\hat{Z}(-\omega')=-(i\omega')^{-1}\exp(-i\omega 2\pi n/E+i(\omega/\kappa)\ln\omega')\cdot\exp(-\pi\omega/2\kappa)\int_0^\infty\exp(-y-i(\omega/\kappa)\ln y)\sin(L_y/2)/L_y\,dy \tag{A.9}$$

where

$$L_y=2\pi n+E/\kappa(-i\pi/2+\ln y-\ln\omega')\,. \tag{A.10}$$

For $E\ll 1$ the difference in the integrals appearing in Eqs. (A.7) and (A.9) is negligible, and we obtain the desired result,

$$\hat{Z}(-\omega')=-\exp(-\pi\omega/\kappa)\hat{Z}(\omega')\,. \tag{A.11}$$

Appendix B: Decomposition of the State $\Psi(\varepsilon^{ab})$

We give here the proof of the lemma used in the analysis of Section IV.

Lemma. *Let $\mathscr{H}_1$ and $\mathscr{H}_2$ be Hilbert spaces and let $\psi^{ab}\in(\mathscr{H}_1\otimes\mathscr{H}_1)_s$, $\eta^{ab}\in(\mathscr{H}_2\otimes\mathscr{H}_2)_s$. Define $\Phi(\mu^{ab})\in\mathscr{F}(\mathscr{H}_1\oplus\mathscr{H}_2)$ by,*

$$\Phi(\mu^{ab})=(1,0,2^{-\frac{1}{2}}\mu^{ab},0,((3\cdot 1)/(4\cdot 2))^{\frac{1}{2}}\mu^{(ab}\mu^{cd)},0,\ldots)$$

where $\mu^{ab}=\psi^{ab}+\eta^{ab}$. Then the natural isomorphism between $\mathscr{F}(\mathscr{H}_1\oplus\mathscr{H}_2)$ and $\mathscr{F}(\mathscr{H}_1)\otimes\mathscr{F}(\mathscr{H}_2)$ takes the state $\Phi(\mu^{ab})$ into the simple product state $\Phi_1(\psi^{ab})\otimes\Phi_2(\eta^{ab})$, where

$$\Phi_1(\psi^{ab})=(1,0,2^{-\frac{1}{2}}\psi^{ab},0,((3\cdot 1)/(4\cdot 2))^{\frac{1}{2}}\psi^{(ab}\psi^{cd)},0,\ldots)\,,$$
$$\Phi_2(\eta^{ab})=(1,0,2^{-\frac{1}{2}}\eta^{ab},0,((3\cdot 1)/(4\cdot 2))^{\frac{1}{2}}\eta^{(ab}\eta^{cd)},0,\ldots)\,.$$

Proof. We can write $\Phi(\mu^{ab})$ as a sum of terms containing k factors of ψ^{ab} and l factors of η^{ab}, i.e. terms of the form

$$c_{kl}\underbrace{\psi^{(ab}\psi^{cd}\ldots\psi^{yz}}_{k\text{ factors}}\underbrace{\eta^{a'b'}\eta^{c'd'}\ldots\eta^{y'z')}}_{l\text{ factors}}\,. \tag{B.1}$$

The coefficient c_{kl} is simply the coefficient of the n-particle term in the expression for $\Phi(\mu^{ab})$ [where $n=2(k+l)$], times the appropriate binomial coefficient, i.e.,

$$c_{kl}=[([2(l+k)-1][2(l+k)-3]\ldots 1)/([2(l+k)][2(l+k)-2]\ldots 2)]^{\frac{1}{2}}\cdot(k+l)!/(k!l!)\,. \tag{B.2}$$

Under the natural isomorphism of $\mathscr{F}(\mathscr{H}_1\oplus\mathscr{H}_2)$ with $\mathscr{F}(\mathscr{H}_1)\otimes\mathscr{F}(\mathscr{H}_2)$ the vector (B.1) gets mapped into the vector

$$\tilde{c}_{kl}\psi^{(ab}\ldots\psi^{yz)}\otimes\eta^{(a'b'}\ldots\eta^{y'z')} \tag{B.3}$$

where [see Eqs. (4.4) and (4.5)],

$$\tilde{c}_{kl}=[(2l)!(2k)!/(2l+2k)!]^{\frac{1}{2}}c_{kl}=[((k-1)(k-3)\ldots 1)/(k(k-2)\ldots 2)]^{\frac{1}{2}}[((l-1)(l-3)\ldots 1)/(l(l-2)\ldots 2)]^{\frac{1}{2}}\,. \tag{B.4}$$

Thus, the state $\Phi(\mu^{ab})$ get mapped into the state

$$\Phi(\mu^{ab}) \to \sum_{k,l} \tilde{c}_{kl} \psi^{(ab} \dots \psi^{yz)} \otimes \eta^{(a'b'} \dots \eta^{y'z')}$$
$$= \Phi_1(\psi^{ab}) \otimes \Phi_2(\eta^{ab}) \tag{B.5}$$

where Φ_1 and Φ_2 are defined above.

References

1. Hawking, S.W.: Particle Creation by Black Holes. Commun. math. Phys. **43**, 199—220 (1975)
2. Seiler, R.: Remarks on the theory of particles with spin up to one in external fields. In: Troubles in the external field problem for invariant wave equations, ed. A. S. Wightman. New York: Gordon and Breach 1971
3. Hawking, S.W.: Black Holes are White Hot, Lecture given at the Seventh Texas Symposium in Relativistic Astrophysics (1974)
4. Geroch, R.: unpublished lecture notes
5. Fulling, S. A.: Scalar Quantum Field Theory in a Closed Universe of Constant Curvature, Ph. D. thesis, Princeton University (1972)
6. DeWitt, B.: Quantum Field Theory in Curved Spacetime, Phys. Reports (in press)
7. Riesz, F., Sz.-Nagy, B.: Functional Analysis. New York: Ungar Publishing Co. 1955
8. Parker, L., Fulling, S. A.: Phys. Rev. D **9**, 341 (1974)
9. Hawking, S.W.: Phys. Rev. Lett. **26**, 1344 (1971)
10. Bardeen, J. M., Carter, B., Hawking, S.W.: Commun. Math. Phys. **31**, 161 (1973)
11. Bekenstein, J. D.: Phys. Rev. D **9**, 3292 (1974)

Communicated by A. S. Wightman

Received March 4, 1975