PHYSICAL REVIEW D VOLUME 14, NUMBER 10 15 NOVEMBER 1976

Breakdown of predictability in gravitational collapse*

S. W. Hawking†

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, England and California Institute of Technology, Pasadena, California 91125

(Received 25 August 1975)

The principle of equivalence, which says that gravity couples to the energy-momentum tensor of matter, and the quantum-mechanical requirement that energy should be positive imply that gravity is always attractive. This leads to singularities in any reasonable theory of gravitation. A singularity is a place where the classical concepts of space and time break down as do all the known laws of physics because they are all formulated on a classical space-time background. In this paper it is claimed that this breakdown is not merely a result of our ignorance of the correct theory but that it represents a fundamental limitation to our ability to predict the future, a limitation that is analogous but additional to the limitation imposed by the normal quantum-mechanical uncertainty principle. The new limitation arises because general relativity allows the causal structure of space-time to be very different from that of Minkowski space. The interaction region can be bounded not only by an initial surface on which data are given and a final surface on which measurements are made but also a "hidden surface" about which the observer has only limited information such as the mass, angular momentum, and charge. Concerning this hidden surface one has a "principle of ignorance": The surface emits with equal probability all configurations of particles compatible with the observers limited knowledge. It is shown that the ignorance principle holds for the quantum-mechanical evaporation of black holes: The black hole creates particles in pairs, with one particle always falling into the hole and the other possibly escaping to infinity. Because part of the information about the state of the system is lost down the hole, the final situation is represented by a density matrix rather than a pure quantum state. This means there is no S matrix for the process of black-hole formation and evaporation. Instead one has to introduce a new operator, called the superscattering operator, which maps density matrices describing the initial situation to density matrices describing the final situation.

I. INTRODUCTION

Gravity is by far the weakest interaction known to physics: The ratio of the gravitational to electrical forces between two electrons is about one part in 10^{43}. In fact, gravity is so weak that it would not be observable at all were it not distinguished from all other interactions by having the property known as the principle of universality or equivalence: Gravity affects the trajectories of all freely moving particles in the same way. This has been verified experimentally to an accuracy of about 10^{-11} by Roll, Krotkov, and Dicke[1] and by Braginsky and Panov.[2] Mathematically, the principle of equivalence is expressed as saying that gravity couples to the energy-momentum tensor of matter. This result and the usual requirement from quantum theory that the local energy density should be positive imply that gravity is always attractive. The gravitational fields of all the particles in large concentrations of matter therefore add up and can dominate over all other forces. As predicted by general relativity and verified experimentally, the universality of gravity extends to light. A sufficiently high concentration of mass can therefore produce such a strong gravitational field that no light can escape. By the principle of special relativity, nothing else can escape either since nothing can travel faster than light. One thus has a situation in which a certain amount of matter is trapped in a region whose boundary shrinks to zero in a finite time. Something obviously goes badly wrong. In fact, as was shown in a series of papers by Penrose and this author,[3-6] a space-time singularity is inevitable in such circumstances provided that general relativity is correct and that the energy-momentum tensor of matter satisfies a certain positive-definite inequality.

Singularities are predicted to occur in two areas. The first is in the past at the beginning of the present expansion of the universe. This is thought to be the "big bang" and is generally regarded as the beginning of the universe. The second area in which singularities are predicted is the collapse of isolated regions of high-mass concentration such as burnt-out stars.

A singularity can be regarded as a place where there is a breakdown of the classical concept of space-time as a manifold with a pseudo-Reimannian metric. Because all known laws of physics are formulated on a classical space-time background, they will all break down at a singularity. This is a great crisis for physics because it means that one cannot predict the future: One does not know what will come out of a singularity.

Many physicists are very unwilling to believe that physics breaks down at singularities. The following attempts were therefore made in order to try to avoid this conclusion.

1. *General relativity does not predict singularities.* This was widely believed at one time (e.g., Lifshitz and Khalatnikov[8]). It was, however, abandoned after the singularity theorems mentioned above and it is now generally accepted that the classical theory of general relativity does indeed predict singularities (Lifshitz and Khalatnikov[9]).

2. *Modify general relativity.* In order to prevent singularities the modifications have to be such as to make gravity repulsive in some situations. The simplest viable modification is probably the Brans-Dicke theory,[10]. In this, however, gravity is always attractive so that the theory predicts singularities just as in general relativity.[7] The Einstein-Cartan theory[11] contains a spin-spin interaction which can be repulsive. This might prevent singularities in some cases but there are situations (such as a purely gravitational and electromagnetic fields) in which singularities will still occur. Most other modifications of general relativity appear either to be in conflict with observations or to have undesirable features like negative energy or fourth-order equations.

3. *The "cosmic censorship" hypothesis: Nature abhors a naked singularity.* In other words, if one starts out with an initially nonsingular asymptotically flat situation, any singularities which subsequently develop due to gravitational collapse will be hidden from the view of an observer at infinity by an event horizon. This hypothesis, though unproved, is probably true for the classical theory of general relativity with an appropriate definition of nontrivial singularities to rule out such cases as the world lines of pressure-free matter intersecting on caustics. If the cosmic censorship hypothesis held, one might argue that one could ignore the breakdown of physics at space-time singularities because this could never cause any detectable effect for observers careful enough not to fall into a black hole. This is a rather selfish attitude because it ignores the question of what happens to an observer who does fall through an event horizon. It also does not solve the problem of the big-bang singularity which definitely is naked. The final blow to this attempt to evade the issue of breakdown at singularities, however, has been the discovery by this author[12,13] that black holes create and emit particles at a steady rate with a thermal spectrum. Because this radiation carries away energy, the black holes must presumably lose mass and eventually disappear. If one tries to describe this process of black-hole evaporation by a classical space-time metric, there is inevitably a naked singularity when the black hole disappears. Even if the black hole does not evaporate completely one can regard the emitted particles as having come from the singularity inside the black hole and having tunnelled out through the event horizon on spacelike trajectories. Thus even an observer at infinity cannot avoid seeing what happens at a singularity.

4. *Quantize general relativity.* One would expect quantum gravitational effects to be important in the very strong fields near a singularity. A number of people have hoped, therefore, that these quantum effects might prevent the singularity from occurring or might smear it out in some way such as to maintain complete predictability within the limits set by the uncertainty principle. However, serious difficulties have arisen in trying to treat quantum gravity like quantum electrodynamics by using perturbation theory about some background metric (usually flat space). Usually in electrodynamics one makes a perturbation expansion in powers of the small parameter $e^2/\hbar c$, the charge squared. Because of the principle of equivalence, the quantity in general relativity that corresponds to charge in electrodynamics is the energy of a particle. The perturbation expansion is therefore really a series in powers of the various energies involved divided by the Planck mass $\hbar^{1/2}c^{1/2}G^{-1/2} \simeq 10^{-5}$ g.

This works well for low-energy tree-approximation diagrams but it breaks down for diagrams with closed loops where one has to integrate over all energies. At energies of the Planck mass, all diagrams become equally important and the series diverges. This is the basic reason why general relativity is not renormalizable.[14,15]

Each additional closed loop appears to involve a new infinite subtraction. There appears to be an infinite sequence of finite remainders or renormalization parameters which are not determined by the theory. One therefore cannot, as was hoped, construct an S matrix which would make definite predictions. The trouble with perturbation theory is that it uses the light cones of a fixed background space. It therefore cannot describe situations in which horizons or worm holes develop by vacuum fluctuations. This is not to say that one cannot quantize gravity, but that one needs a new approach.

One possible view of the failure of the above attempts to avoid the breakdown of predictability would be that we have not yet discovered the correct theory. The aim of this paper, however, is to show this cannot be the case if one accepts that quantum effects will cause a black hole to radiate. In this case there is a basic limitation on our ability to predict which is similar but additional to the usual quantum-mechanical uncertainty principle. This extra limitation arises because general relativity allows the causal structure of space-time to be very different from that of Minkowski space. For example, in the case of gravitational collapse which produces a black hole there is an event horizon which prevents observers at infinity from measuring the internal state of the black hole apart from its mass, angular momentum, and charge. This means that measurements at future infinity are insufficient to determine completely the state of the system at past infinity: One also needs data on the event horizon describing what fell into the black hole. One might think that one could have observers stationed just outside the event horizon who would signal to the observers at future infinity every time a particle fell into the black hole. However, this is not possible, just as one cannot have observers who will measure both the position and the velocity of a particle. To signal accurately the time at which a particle crossed the event horizon would require a photon of the same wavelength and therefore the same energy as that of the infalling particle. If this were done for every particle which underwent gravitational collapse to form the black hole, the total energy required to signal would be equal to that of the collapsing body and there would be no energy left over to form the black hole. It therefore follows that when a black hole forms, one cannot determine the results of measurements at past infinity from observations at future infinity. This might not seem so terrible because one is normally more concerned with prediction than postdiction. However, although in such a situation one could classically determine future infinity from knowledge of past infinity, one cannot do this if quantum effects are taken into account. For example, quantum mechanics allows particles to tunnel on spacelike or past-directed world lines. It is therefore possible for a particle to tunnel out of the black hole through the event horizon and escape to future infinity. One can interpret such a happening as being the spontaneous creation in the gravitational field of the black hole of a pair of particles, one with negative and one with positive energy with respect to infinity. The particle with negative energy would fall into the black hole where there are particle states with negative energy with respect to infinity. The particles with positive energy can escape to infinity where they constitute the recently predicted thermal emission from black holes. Because these particles come from the interior of the black hole about which an external observer has no knowledge, he cannot predict the amplitudes for them to be emitted but only the *probabilities* without the phases.

In Secs. III and IV of this paper it is shown that the quantum emission from a black hole is completely random and uncorrelated. Similar results have been found by Wald[16] and Parker.[17] The black hole emits with equal probability every configuration of particles compatible with conservation of energy, angular momentum, and charge (not every configuration escapes to infinity with equal probability because there is a potential barrier around the black hole which depends on the angular momentum of the particles and which may reflect some of the particles back into the black holes). This result can be regarded as a quantum version of the "no hair" theorems because it implies that an observer at infinity cannot predict the internal state of the black hole apart from its mass, angular momentum, and charge: If the black hole emitted some configuration of particles with greater probability than others, the observer would have some *a priori* information about the internal state. Of course, if the observer measures the wave functions of all the particles that are emitted in a particular case he can then *a posteriori* determine the internal state of the black hole but it will have disappeared by that time.

A gravitational collapse which produces an event horizon is an example of a situation in which the interaction region is bounded by an initial surface on which data are prescribed, a final surface on which measurements are made, and, in addition, a third "hidden" surface about which the observer can have only limited information such as the flux of energy, angular momentum, or charge. Such hidden surfaces can surround either singularities (as in the Schwarzschild solution) or "wormholes" leading to other space-time regions about which the observer has no knowledge (as in the Reissner-Nordström or other solutions). About this surface one has the *principle of ignorance*.

All data on a "hidden" surface compatible with the observer's limited information are equally probable.

So far the discussion has been in terms of quantized matter fields on a fixed classical background metric (the semiclassical approximation). However, one can extend the principle to treatments in which the gravitational field is also quantized by means of the Feynman sum over histories. In this one performs an integration (with an as yet undetermined measure) over all configuration of both matter and gravitational fields. The classical example of black-hole event horizons shows that in this integral one has to include metrics in which the interaction region (i.e., the region over which the action is evaluated) is bounded, not only by the initial and final surfaces, but by a hidden surface as well. Indeed, in any quantum gravitational situation there is the possibility of "virtual" black holes which arise from vacuum fluctuations and which appear out of nothing and then disappear again. One therefore has to include in the sum over histories metrics containing transient holes, leading either to singularities or to other space-time regions about which one has no knowledge. One therefore has to introduce a hidden surface around each of these holes and apply the principle of ignorance to say that all field configurations on these hidden surfaces are equally probable provided they are compatible with the conservation of mass, angular momentum, etc. which can be measured by surface integrals at a distance from the hole.

Let H_1 be the Hilbert space of all possible data on the initial surface, H_2 be the Hilbert space of all possible data on the hidden surface, and H_3 be the Hilbert space of all possible data on the final surface. The basic assumption of quantum theory is that there is some tensor S_{ABC} whose three indices refer to H_3, H_2, and H_1, respectively, such that if

$$\xi_C \in H_1\,, \quad \zeta_B \in H_2\,, \quad \chi_A \in H_3\,,$$

then

$$\sum \sum \sum S_{ABC}\chi_A \zeta_B \xi_C$$

is the amplitude to have the initial state ξ_C, the final state χ_A, and the state ζ_B on the hidden surface. Given only the initial state ξ_C one cannot determine the final state but only the element $\sum S_{ABC}\xi_C$ of the tensor product $H_2 \otimes H_3$. Because one is ignorant of the state on the hidden surface one cannot find the amplitude for measurements on the final surface to give the answer χ_A but one can calculate the probability for this outcome to be $\sum\sum\sum \rho_{CD}\bar{\chi}_C \chi_D$, where

$$\rho_{CD} = \sum \sum \sum \bar{S}_{CBE} S_{DBF} \bar{\xi}_E \xi_F$$

is the density matrix which completely describes observations made only on the future surface and not on the hidden surface. Note that one gets this density matrix from $\sum S_{ABC}\xi_C$ by summing with equal weight over all the unobserved states on the "hidden" surface.

One can see from the above that there will not be an S matrix or operator which maps initial states to final states, because the observed final situation is described, not by a pure quantum state, but by a density matrix. In fact, the initial situation in general will also be described not by a pure state but by a density matrix because of the hidden surface occurring at earlier times. Instead of an S matrix one will have a new operator called the superscattering operator $\$$, which maps density matrices describing the initial situation to density matrices describing the final situation. This operator can be regarded as a 4-index tensor $\$_{ABCD}$ where the first two indices operate on the final space $H_3 \otimes H_3$ and the last two indices operate on the space $H_1 \otimes H_1$. It is related to the 3-index tensor S_{ABC} by

$$\$_{ABCD} = \tfrac{1}{2} \sum (\bar{S}_{AEC} S_{BED} + \bar{S}_{BEC} S_{AED})\,.$$

The final density matrix ρ_{2AB} is given in terms of the initial density matrix ρ_{1CD} by

$$\rho_{2AB} = \sum \sum \$_{ABCD}\,\rho_{1CD}\,.$$

The superscattering operator is discussed further in Sec. V.

The fact that in gravitational interactions the final situation at infinity is described by a density matrix and not a pure state indicates that quantum gravity cannot, as was hoped, be renormalized to give a well-defined S matrix with only a finite number of undetermined parameters. It seems reasonable to conjecture that there is a close connection between the infinite sequence of renormalization constants that occur in perturbation theory and the loss of predictability which arises from hidden surfaces.

One can also appeal to the principle of ignorance to provide a possible explanation of the observations of the microwave background and of the abundances of helium and deuterium which indicate that the early universe was very nearly spatially homogeneous and isotropic and in thermal equilibrium. One could regard a surface very close to the initial big-bang singularity (say, at the Planck time 10^{-43} sec) as being a "hidden surface" in the sense that we have no *a priori* information about it. The initial surface would thus emit all configurations of particles with equal probability. To obtain a thermal distribution one would need to impose some constraint on the total energy of the configurations where the total energy is the rest-mass energy of the particles plus their kinetic energy of expansion minus their gravitational potential energy. Observationally this energy is very nearly, if not exactly, zero and this can be understood as a necessary condition for our existence: If the total energy were large and positive, the universe would expand too rapidly for galaxies to form, and if the total energy were large and negative, the universe would collapse before intelligent life had time to develop. We therefore do have some limited knowledge of the data on the initial surface from the fact of our own existence. If one assumes that the initial surface emitted with equal probability all configurations of particles with total energy (with some appropriate definition) nearly equal to zero, then an approximately thermal distribution is the most probable macrostate since it corresponds to the largest number of microstates. Any significant departure from homogeneity or isotropy could be regarded as the presence in some long-wavelength modes of a very large number of gravitons, a number greatly in excess of that for a thermal distribution and therefore highly improbable. It should be pointed out that this view of the generality of isotropic expansion is the opposite of that adopted by Collins and Hawking.[18] The difference arises from considering microscopic rather than macroscopic configurations.

One might also think to explain the observed net baryon number of the universe by saying that we, as observers, could result only from initial configurations that had a net baryon number. An alternative explanation might be that CP violations in the highly T-nonsymmetric early universe caused expanding configurations in which baryons predominated to have lower energies than similar expanding configurations in which antibaryons predominated. This would mean that for a given energy density there would be more configurations with a positive baryon number than with a negative baryon number, thus the expectation value of the baryon number would be positive. Alternatively, there might be a sort of spontaneous symmetry breaking which resulted in regions of pure baryons or pure antibaryons having lower energy densities than regions containing a mixture of baryons and antibaryons. In this case, as suggested by Omnès,[19] one would get a phase transition in which regions of pure baryons were separated from regions of pure antibaryons. Unlike the case considered by Omnès, there is no reason why the separation should not be over length scales larger than the particle horizon. Such a greater separation would overcome most of the difficulties of the Omnès model.

There is a close connection between the above proposed explanation for the isotropy of the universe and the suggestion by Zel'dovich[20] that it is caused by particle creation in anisotropic regions. In Zel'dovich's work, however, in order to define particle creation, one has to pretend that the universe was time-independent at early times (which is obviously not the case). The present approach avoids the difficulty of talking about early times; one merely has to count the configurations at some convenient late time.

The conclusion of this paper is that gravitation introduces a new level of uncertainty or randomness into physics over and above the uncertainty usually associated with quantum mechanics. Einstein was very unhappy about the unpredictability of quantum mechanics because he felt that "God does not play dice." However, the results given here indicate that "God not only plays dice, He sometimes throws the dice where they cannot be seen."

II. QUANTUM THEORY IN CURVED SPACE-TIME

In this section a brief outline is given of the formalism of quantum theory on a given space-time background which was used by Hawking[13] to derive the quantum-mechanical emission from black holes. This formalism will be used in Sec. III to show that the radiation which escapes to infinity is completely thermal and uncorrelated. In Sec. IV a specific choice of states for particles going into the black hole is used to calculate explicitly both the ingoing and the emitted particles. This shows that the particles are created in pairs with one member of the pair always falling into the hole and the other member either falling in or escaping to infinity. Section V contains a discussion of the superscattering operator $\mathcal{S}$ which maps density matrices describing the initial situation to density matrices describing the final situation.

For simplicity only a massless Hermitian scalar field ϕ and an uncharged nonrotating black hole will be considered. The extension to charged massive fields of higher spin and charged rotating black holes is straightforward along the lines indicated in Ref. 13. Throughout the paper units will be used in which $G=c=\hbar=k=1$.

Figure 1 is a diagram of the situation under consideration: A gravitational collapse creates a black hole which slowly evaporates and eventually disappears by the quantum-mechanical creation and emission of particles. Except in the final stages of the evaporation, when the black hole gets down to the Planck mass, the back reaction on the gravitational field is very small and it can be treated as an unquantized external field. The metric at late times can be approximated by a sequence of time-independent Schwarzschild solutions and the gravitational collapse can be taken to be spherically symmetric (it was shown in Ref. 13 that departures from spherical symmetry made no essential difference).

The scalar field operator ϕ satisfies wave equation

$$\Box\phi=0 \tag{2.1}$$

in this metric and the commutation relations

$$[\phi(x),\phi(y)]=iG(x,y), \tag{2.2}$$

where $G(x,y)$ is the half-retarded minus half-advanced Green's function. One can express the operator ϕ as

$$\phi=\sum(f_i a_i+\bar{f}_i a_i^\dagger), \tag{2.3}$$

where the $\{f_i\}$ are a complete orthonormal family

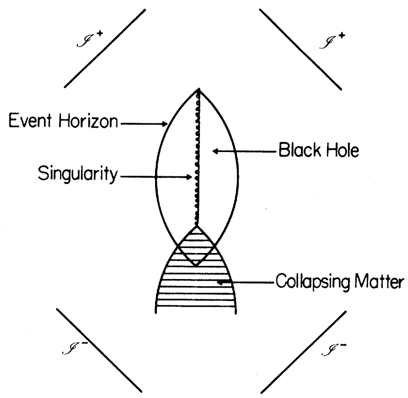

FIG. 1. A gravitational collapse produces a black hole which slowly evaporates by the emission of radiation to future null infinity $\mathscr{I}^+$. Because of the loss of energy, the black hole decreases in size and eventually disappears.

of complex-valued solutions of the wave equation $\Box f_i = 0$ which contain only positive frequencies at past null infinity $\mathscr{I}^-$. The operators a_i are position independent and obey the commutation relations

$$[a_i, a_j] = 0\,, \tag{2.4}$$

$$[a_i, a_j^\dagger] = \delta_{ij}\,. \tag{2.5}$$

The operators a_i and $a_i^\dagger$ are respectively the annihilation and creation operators for particles in the ith mode at past infinity. The initial vacuum state for scalar particles $|0_-\rangle$, i.e., the state which contains no scalar particles at past infinity, is defined by

$$a_i|0_-\rangle = 0 \quad \text{for all } i\,. \tag{2.6}$$

One can also express ϕ in the form

$$\phi = \sum_i (p_i b_i + \bar{p}_i b_i^\dagger + q_i c_i + \bar{q}_i c_i^\dagger)\,. \tag{2.7}$$

Here the $\{p_i\}$ are a complete orthonormal family of solutions of the wave equation which contain only positive frequencies at future null infinity $\mathscr{I}^+$ and which are purely outgoing, i.e., they have zero Cauchy data on the event horizon H. The $\{q_i\}$ are a complete orthonormal set of solutions of the wave equation which contain no outgoing component. The position-independent operators b_i and c_i obey the commutation relations

$$[b_i, b_j] = [c_i, c_j] = 0\,, \tag{2.8}$$

$$[b_i, c_j] = [b_i, c_j^\dagger] = 0\,, \tag{2.9}$$

$$[b_i, b_j^\dagger] = [c_i, c_j^\dagger] = \delta_{ij}\,. \tag{2.10}$$

The operators b_i and $b_i^\dagger$ are respectively the annihilation and creation operators for outgoing particles at future infinity. By analogy one could regard the operators c_i and $c_i^\dagger$ as the annihilation and creation operators for particles falling into the black hole. However, because one cannot uniquely define positive frequency for the $\{q_i\}$, the division into annihilation and creation parts is not unique and so one should not attach too much physical significance to this interpretation. The nonuniqueness of the $\{c_i\}$ and the $\{c_i^\dagger\}$ does not affect any observable at future infinity. In Sec. IV a particular choice of the $\{q_i\}$ will be made which will allow an explicit calculation of the particles going into the black hole. The final scalar-particle vacuum state $|0_+\rangle$, i.e., the state which contains no outgoing particles at future infinity or particles going into the black hole, is defined by

$$b_i|0_+\rangle = c_i|0_+\rangle = 0\,. \tag{2.11}$$

It can be represented as $|0_I\rangle|0_H\rangle$, where the b and c operators act on $|0_I\rangle$ and $|0_H\rangle$, respectively, which are the vacua for outgoing particles and for particles falling into the hole. $|0_I\rangle$ is uniquely defined by the positive-frequency condition on the $\{p_i\}$ but the ambiguity in the choice of the $\{q_i\}$ means that $|0_H\rangle$ is not unique.

Because massless fields are completely determined by their data on $\mathscr{I}^-$ one can express $\{p_i\}$ and $\{q_i\}_{\Phi}$ as linear combinations of the $\{f_i\}$ and $\{\bar{f}_i\}$:

$$p_i = \sum_j (\alpha_{ij} f_j + \beta_{ij}\bar{f}_j)\,, \tag{2.12}$$

$$q_i = \sum_j (\gamma_{ij} f_j + \eta_{ij}\bar{f}_j)\,. \tag{2.13}$$

These relations lead to corresponding relations between the operators:

$$b_i = \sum_j (\bar{\alpha}_{ij} a_j - \bar{\beta}_{ij} a_j^\dagger)\,, \tag{2.14}$$

$$c_i = \sum_j (\bar{\gamma}_{ij} a_j - \bar{\eta}_{ij} a_j^\dagger)\,. \tag{2.15}$$

In the situation under consideration the metric is spherically symmetric. This means the angular dependence of the $\{f_i\}$, $\{p_i\}$, and $\{q_i\}$ can be taken to be that of spherical harmonics Y_{lm}. The relations (2.12) and (2.13) will connect only solutions with the same values of l and $|m|$. (This is not true if the collapse is not exactly spherically symmetric but it was shown in Ref. 13 that this makes no essential difference.) For computational purposes it is convenient to use f and p solutions which have time dependence of the form $e^{i\omega' v}$ and $e^{i\omega u}$, respectively, where v and u are advanced and retarded times. The solutions will be denoted by

$\{f_{\omega'}\}$ and $\{p_\omega\}$ and will have continuum normalization. They can be superposed to form wave-packet solutions of finite normalization. The summations in Eqs. (2.3), (2.7), and (2.12) are replaced by integrations over frequency. The operators a_ω, b_ω, etc. obey similar commutation relations involving δ functions in the frequency.

The advantage of using Fourier components with respect to time is that one can calculate the coefficients $\alpha_{\omega\omega'}$ and $\beta_{\omega\omega'}$ in the approximation that the mass of the black hole is changing only slowly. One considers a solution p_ω propagating backwards in time from future infinity. A part $p_\omega^{(1)}$ is reflected by the static Schwarzschild metric and reaches past infinity with the same frequency. This gives a term $r_\omega\delta(\omega-\omega')$ in $\alpha_{\omega\omega'}$, where r_ω is the reflection coefficient of the Schwarzschild metric for the frequency ω and the given angular mode. More interesting is the behavior of the part $p_\omega^{(2)}$ which propagates through the collapsing body and out to past infinity with a very large blue-shift. This gives contributions to $\alpha_{\omega\omega'}$ and $\beta_{\omega\omega'}$ of the form

$$\alpha^{(2)}_{\omega\omega'} \simeq t_\omega(2\pi)^{-1}e^{i(\omega-\omega')v_0}\left(\frac{\omega'}{\omega}\right)^{1/2} \times\Gamma\left(1-\frac{i\omega}{\kappa}\right)(-i\omega')^{-1+i\omega/\kappa}, \tag{2.16}$$

$$\beta^{(2)}_{\omega\omega'} \simeq -i\alpha^{(2)}_{\omega(-\omega')}, \tag{2.17}$$

where $\kappa=(4M)^{-1}$ is the surface gravity of the black hole and where t_ω is the transmission coefficient for the given Schwarzschild metric, i.e.,

$$|t_\omega|^2=\Gamma_\omega$$

is the fraction of a wave with frequency ω and the given angular dependence which penetrates through the potential barrier into the hole,

$$|t_\omega|^2+|r_\omega|^2=1\,.$$

III. THE OUTGOING RADIATION

One assumes that there are no scalar particles present in the infinite past, i.e., the system is in the initial scalar-particle vacuum state $|0_-\rangle$. (It is not a complete vacuum because it contains the matter that will give rise to the black hole.) The state $|0_-\rangle$ will not coincide with the final scalar-particle vacuum state $|0_+\rangle$ because there is particle creation. One can express $|0_-\rangle$ as a linear combination of states with different numbers of particles going out to infinity and into the horizon:

$$|0_-\rangle=\sum\sum\lambda_{AB}|A_I\rangle|B_H\rangle, \tag{3.1}$$

where $|A_I\rangle$ is the outgoing state with n_{ja} particles in the jth outgoing mode and $|B_H\rangle$ is the horizon state with n_{kb} particles in the kth mode going into the hole. In other words,

$$|A_I\rangle=\prod_j(n_{ja}!)^{-1/2}(b_j^\dagger)^{n_{ja}}|0_I\rangle, \tag{3.2}$$

$$|B_H\rangle=\prod_k(n_{kb}!)^{-1/2}(c_k^\dagger)^{n_{kb}}|0_H\rangle. \tag{3.3}$$

An operator Q which corresponds to an observable at future infinity will be composed only of the $\{b_j\}$ and the $\{b_j^\dagger\}$ and will operate only on the vectors $|A_I\rangle$. Thus the expectation value of this operator will be

$$\langle 0_-|Q|0_-\rangle=\sum\sum\rho_{AC}Q_{CA} \tag{3.4}$$

where $Q_{CA}=\langle C_I|Q|A_I\rangle$ in the matrix element of the operator Q on the Hilbert space of outgoing states and $\rho_{AC}=\sum\lambda_{AB}\bar{\lambda}_{CB}$ is the density matrix which completely describes all observations which are made only at future infinity and do not measure what went into the hole. The components of ρ_{AC} can be completely determined from the expectation values of polynomials in the operators $\{b_j\}$ and $\{b_j^\dagger\}$. Thus the density matrix is independent of the ambiguity in the choice of the $\{q_j\}$ which describes particles going into the hole.

As an example of such a polynomial consider $b_j^\dagger b_j$, which is the number operator for the jth outgoing mode. Then

$$\begin{aligned}\langle n_j\rangle&=\sum n_{ja}\rho_{AA}\\&=\langle 0_-|b_j^\dagger b_j|0_-\rangle\\&=\sum_k|\beta_{jk}|^2.\end{aligned} \tag{3.5}$$

In order to calculate this last expression one expands the finite-normalization wave-packet mode p_j in terms of continuum-normalization modes p_ω,

$$p_j(u)=\int^\infty\tilde{p}_j(\omega)p_\omega(u)d\omega, \tag{3.6}$$

where

$$\int^\infty\bar{\tilde{p}}_j\tilde{p}_l d\omega=\delta_{jl}, \tag{3.7}$$

then

$$\langle n_j\rangle=\iiint\bar{\tilde{p}}_j(\omega_1)\tilde{p}_j(\omega_2)\,, \times\beta_{\omega_1\omega'}\bar{\beta}_{\omega_2\omega'}d\omega_1d\omega_2d\omega'\,.$$

If the wave packet is sharply peaked around frequency ω, one can use Eq. (2.14) to show that

$$\int \beta_{\omega_1\omega'}\bar{\beta}_{\omega_2\omega'}d\omega' = (2\pi)^{-2}|t_\omega|^2|\Gamma(1-i\omega\kappa^{-1})|^2 \times e^{i(\omega_1-\omega_2)v_0}\omega^{-1}e^{-\pi\omega\kappa^{-1}} \times \int_{-\infty}^{\infty} e^{iy\kappa^{-1}(\omega_1-\omega_2)}dy\,, \tag{3.8}$$

where $y=\ln(-\omega')$. The factor $e^{-\pi\omega\kappa^{-1}}$ arises from the analytic continuation of ω' to negative values in the expression (2.15) for $\beta_{\omega\omega'}$,

$$\text{Eq. (3.8)} = |t_\omega|^2(e^{2\pi\omega\kappa^{-1}}-1)^{-1}\delta(\omega_1-\omega_2)\,, \tag{3.9}$$

therefore

$$\langle n_j\rangle = |t_\omega|^2(e^{2\pi\omega\kappa^{-1}}-1)^{-1}\,. \tag{3.10}$$

This is precisely the expectation value for a body emitting thermal radiation with a temperature $T=\kappa/2\pi$. To show that the probabilities of emitting different numbers of particles in the jth mode and not just the average number are in agreement with thermal radiation, one can calculate the expectation values of n_j^2, n_j^3, and so on. For example,

$$\begin{aligned}\langle n_j^2\rangle &= \langle 0_-|b_j^\dagger b_j b_j^\dagger b_j|0_-\rangle \\ &= \langle n_j\rangle + \langle 0_-|(b_j^\dagger)^2(b_j)^2|0_-\rangle\,. \end{aligned}\tag{3.11}$$

One can evaluate the second term on the right-hand side of (3.11) using Eqs. (2.14) and (2.15) as above. The terms $\alpha^{(2)}_{\omega\omega'}$ give rise to expressions involving functions like $\delta(\omega_1+\omega_2)$ which do not contribute, since ω_1 and ω_2 are both positive. The terms in $\alpha^{(1)}_{\omega\omega'}$ give rise to expressions involving functions like $\int\tilde{p}_j^2(\omega)d\omega$ which vanish because for wave packets at late times the phase of $\tilde{p}_j(\omega)$ varies very rapidly with ω. Thus,

$$\langle n_j^2\rangle = \frac{x\Gamma[1+(2\Gamma-1)x]}{(1-x)^2}\,, \tag{3.12}$$

where $x=e^{-\omega T^{-1}}$ and $\Gamma=|t_\omega|^2$. Proceeding inductively one can calculate the higher moments $\langle n_j^3\rangle$, etc. These are all consistent with the probability distribution for n particles in the jth mode,

$$P(n_j) = \frac{(1-x)(x\Gamma)^n}{[1-(1-\Gamma)x]^{n+1}}\,. \tag{3.13}$$

This is just the combination of the thermal probability $(1-x)x^m$ to emit m particles in the given mode with the probability Γ that a given emitted particle will escape to infinity and not be reflected back into the hole by the potential barrier.

One can also investigate whether there is any correlation between the phases for emitting different numbers of particles in the same mode by examining the expectation values of operators like $b_j b_j$ which connect components of the density matrix with different numbers of particles in the jth mode. These expectation values are all zero. To see whether there are any correlations between different modes one can consider the expectation values of operators like $b_j^\dagger b_l$ which relate to other nondiagonal components of the density matrix. These are also all zero. Thus the density matrix is completely diagonal in a basis of states with definite particle numbers in modes which are sharply peaked in frequency. One can express the density matrix explicitly as

$$\rho_{AC} = \prod_j \delta_{n_{ja}n_{jc}}P(n_{ja})\,. \tag{3.14}$$

The density matrix (3.14) is exactly what one would expect for a body emitting thermal radiation.

As the black hole emits radiation its mass will go down and its temperature will go up. This variation will be slow except when the mass of the black hole has gone down to nearly the Planck mass. Thus to a good approximation the probability of n_j particles being emitted in the jth wave-packet mode will be given by Eq. (3.13) where the temperature corresponds to the mass of the black hole at the retarded time around which the jth mode is peaked. After the black hole has completely evaporated and disappeared, the only possible states $|A_I\rangle$ for the radiation at future infinity will be those for which the total energy of the particles is equal to the initial mass M_0 of the black hole. The probability of such a state occurring will be

$$\begin{aligned}P(A) &= \rho_{AA} \\ &= \prod_j P(n_{ja})\,. \end{aligned}\tag{3.15}$$

If Γ were 1 for all modes,

$$\ln[P(A)] = \sum_j \ln(1-x_{ja}) - \sum 8\pi n_{ja}\omega_j M_{ja}\,, \tag{3.16}$$

where M_{ja} is the mass to which the black hole has been reduced by the retarded time of the jth mode by emission of particles in configuration A. By conservation of energy $\sum n_{ja}\omega_j = M_0$ for all possible configurations A of the emitted particles. Because M_{ja} is only a slowly varying function of the mode number j, the last term in Eq. (3.16) will be nearly the same for all configurations A. Thus the black hole emits all configurations with equal probability. The probabilities of different configurations at future infinity are not equal because the Γ factors are different for different modes.

IV. THE INGOING PARTICLES

In this section a specific choice will be made of the ingoing solutions $\{q_i\}$ which will allow an explicit calculation to be made of the coefficients λ_{AB} so that the state of the system can be expressed in terms of particles falling into the black hole and particles escaping to infinity. The outgoing solutions $\{p_j\}$ are chosen to be purely positive frequency along the orbits of the approximate time-translation Killing vector K in the quasistationary region outside the black hole at late times. They therefore, correspond to particle modes that would be measured by an observer with a detector moving along a world-line at constant distance from the black hole. They do not correspond to what would be detected by nonstationary observers, in particular observers falling into the black hole, because they are not purely positive frequency along the world lines of such observers.

A stationary observer outside the black hole could regard a particle he detected in a mode $\{p_j\}$ as being one member of a pair of particles created by the gravitational yield of the collapse, the other member having negative energy and having fallen into the black hole. The horizon states $\{q_j\}$ will be chosen so that some of them describe those negative-energy particles which the stationary observer considers to exist inside the black hole. The remaining $\{q_j\}$ will describe those positive-energy particles which are reflected back by the potential barrier around the black hole and which fall through the event horizon. It should be emphasized that this choice of $\{q_j\}$ does not correspond to anything that an infalling observer would measure since they are not positive frequency along his world line. However, given the $\{p_j\}$, the choice of the $\{q_j\}$ that will be used is minimal in the sense that any other choice would describe the creation of extra pairs of particles, both of which fell into the black hole.

To calculate the coefficients α and β which relate the $\{p_i\}$ to the $\{f_i\}$ and $\{\bar{f}_i\}$ one decomposes the $\{p_i\}$ into Fourier components $\{p_\omega\}$ with time dependence of the form $e^{i\omega u}$, where $u=t-r-2M\ln(r-2M)$ is the retarded time coordinate in the Schwarzschild solution. Because u tends to $+\infty$ in the exterior region as one approaches the future horizon, the surfaces of constant phase of p_ω pile up just outside the future horizon (Fig. 2). In other words, p_ω is blue-shifted to a very high frequency near the future horizon. This means that it propagates by geometric optics back through the collapsing body and out to past null infinity $\mathscr{I}^-$ where it has time dependence of the form

$$e^{-i\omega\kappa^{-1}\ln(v_0-v)} \quad \text{for } v<v_0$$

and (4.1)

$$0 \quad \text{for } v>v_0,$$

where $v=t+r+2M\ln(r-2M)$ is the advanced time coordinate and v_0 is the last advanced time before which a null geodesic could leave $\mathscr{I}^-$, pass through the center of the collapsing object, and escape to $\mathscr{I}^+$. Similarly, to calculate the coefficients γ and η which express the $\{q_i\}$ in terms of the $\{f_i\}$ and the $\{\bar{f}_i\}$ one decomposes the $\{q_i\}$ into Fourier components $\{q_\omega\}$. In the quasistationary region the part $\{q_\omega^{(3)}\}$ that crosses the future horizon in the quasistationary region will have time dependence of the form $e^{i\omega v}$. The part $q_\omega^{(4)}$ which crosses the horizon just after its formation will have time dependence of the form $e^{-i\omega u}$ (the minus sign is because in the interior region the direction of increase of u is reversed). The surfaces of constant phase of $\{q_\omega^{(4)}\}$ pile up just inside the horizon (Fig. 2). One can therefore propagate them backwards also by geometric optics through the collapsing body and out to $\mathscr{I}^-$, where they will have time dependence of the form

$$e^{i\omega\kappa^{-1}\ln(v-v_0)} \quad \text{for } v>v_0$$

and (4.2)

$$0 \quad \text{for } v<v_0.$$

In order to calculate the coefficients α, β, γ, and η one can decompose (4.1) and (4.2) into positive- and negative-frequency components of the form $e^{i\omega' v}$ and $e^{-i\omega' v}$ in terms of the advanced time v at $\mathscr{I}^-$. However, one can obtain the same results

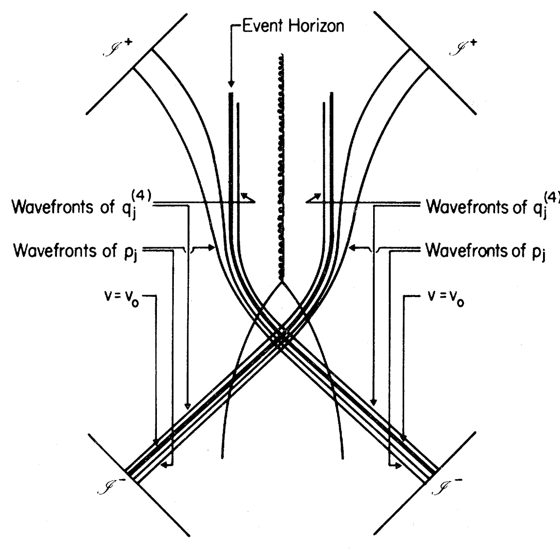

FIG. 2. The wave fronts or surfaces of constant phase of the solutions p_j pile up just outside the event horizon because of the large blue-shift. They propagate by geometric optics through the collapsing body and out to past null infinity $\mathscr{I}^-$ just before the advanced time $v=v_0$. Similarly the wave fronts of $q_j^{(4)}$ will pile up just inside the horizon and will propagate through the collapsing body out to $\mathscr{I}^-$ just after the advanced time $v=v_0$.

if one leaves out the collapsing body and analytically extends back to the past horizon the Schwarzschild solution that represents the quasistationary region. Instead of propagating p_ω and q_ω back through the collapsing body to $\mathscr{I}^-$ and analyzing them there into positive- and negative-frequency components with respect to the advanced time v, one propagates them back to the past horizon H^- and analyzes them into positive- and negative-frequency components with respect to an affine parameter U along the generators of H^-. (A similar construction has been used by Unruh.[21]) One can then discuss the creation of particles in terms of the Penrose diagram (Fig. 3) of the analytically extended Schwarzschild solution. The initial vacuum state $|0_-\rangle$ is now defined as the state which on $\mathscr{I}^-$ has no positive-frequency components with respect to the advanced time v and which on the past horizon H^- has no positive-frequency components with respect to affine parameter U. In other words, one can express the operator ϕ in the form

$$\phi = \int_0^\infty (a_\omega^{(1)} f_\omega^{(1)} + a_\omega^{(2)} f_\omega^{(2)} + \text{H.c.}) d\omega, \tag{4.3}$$

where $\{f_\omega^{(1)}\}$ are a family of solutions of the wave equation in the analytically extended Schwarzschild solution with continuum normalization which have zero Cauchy data on the past horizon and have time dependence of the form $e^{i\omega v}$ on $\mathscr{I}^-$, and $\{f_\omega^{(2)}\}$ are a family of solutions with continuum normalization which have zero Cauchy data on $\mathscr{I}^-$ and have time dependence of the form $e^{i\omega U}$ on the past horizon. The initial vacuum state is then defined by

$$a_\omega^{(1)}|0_-\rangle = a_\omega^{(2)}|0_-\rangle = 0. \tag{4.4}$$

This definition of the vacuum state is different from that used by Boulware[22] for the analytically extended Schwarzschild solution. The above definition, however, reproduces the results on particle creation by a black hole which was formed by a collapse.

The affine parameter U on the past horizon is related to the retarded time u by

$$u = -\kappa^{-1}\ln(-U), \tag{4.5}$$

where $-\infty < u < \infty$, $U<0$. One can analytically continue (4.5) past the logarithmetic singularity at $U=0$. In doing so, one picks up an imaginary part of $\pm\pi\kappa^{-1}$ depending on whether one passes above or below the singularity, respectively. Define the two analytic continuations u_+ and u_- by

$$\begin{aligned} &u_+ = u_- = -\kappa^{-1}\ln(-U) \quad \text{for } U<0, \\ &u_\pm = -\kappa^{-1}\ln U \pm i\pi\kappa^{-1} \quad \text{for } U>0. \end{aligned} \tag{4.6}$$

Because u_+ is holomorphic in the upper half U plane, the functions $e^{i\omega u_+}$ and $e^{-i\omega u_+}$ defined all the way up the past horizon from $U=-\infty$ to $U=+\infty$ both contain only positive frequencies with respect to U. This means that one can replace the family of solutions $\{f_\omega^{(2)}\}$, which have zero Cauchy data on $\mathscr{I}^-$ and only positive frequencies with respect to U on the past horizon, by two orthogonal families of solutions $\{f_\omega^{(3)}\}$ and $\{f_\omega^{(4)}\}$, with continuum normalization which have zero Cauchy data on $\mathscr{I}^-$, and which have time dependence on the past horizon of the form $e^{i\omega u_+}$ and $e^{-i\omega u_+}$, respectively. One can then express ϕ as

$$\phi = \int (a_\omega^{(1)} f_\omega^{(1)} + a_\omega^{(3)} f_\omega^{(3)} + a_\omega^{(4)} f_\omega^{(4)} + \text{H.c.}) d\omega. \tag{4.7}$$

Equation (4.4) then becomes

$$a_\omega^{(1)}|0_-\rangle = a_\omega^{(3)}|0_-\rangle = a_\omega^{(4)}|0_-\rangle = 0. \tag{4.8}$$

Equation (4.8) says that there are no scalar particles in the modes $\{f_\omega^{(3)}\}$ and $\{f_\omega^{(4)}\}$. However, these modes extend across both the interior and exterior regions of the analytically continued Schwarzschild solution. An observer at future null infinity $\mathscr{I}^+$ cannot measure these modes but only the part of them outside the future horizon. To correspond with what an observer sees, define a new basis consisting of three orthogonal families $\{w_\omega\}$, $\{y_\omega\}$, and $\{z_\omega\}$ of solutions with continuum normalization with the following properties:

$\{w_\omega\}$ have zero Cauchy data on $\mathscr{I}^-$ and on the past horizon for $U<0$. On the past horizon for $U>0$ they have time dependence of the form $e^{-i\omega u_+}$. (The minus sign is necessary in order for the $\{w_\omega\}$ to have positive Klein-Gordon norm and thus for the associated annihilation and creation operators to have the right commutation relations.)

$\{y_\omega\}$ have zero Cauchy data on $\mathscr{I}^-$ and the past horizon for $U>0$. On the past horizon for $U<0$ they have time dependence of the form $e^{i\omega u_+}$.

$\{z_\omega\}$ have zero Cauchy data on the past horizon and on $\mathscr{I}^-$ they have time dependence of the form $e^{i\omega v}$.

The modes $\{z_\omega\}$ represent particles which come in from $\mathscr{I}^-$ and pass through the future horizon with probability $|t_\omega|^2$ or are reflected back to $\mathscr{I}^+$ with probability $|r_\omega|^2$. The modes $\{y_\omega\}$ represent particles which, in the analytically extended Schwarzschild space, appear to come from the past horizon and which escape to $\mathscr{I}^+$ with probability $|t_\omega|^2$ or are reflected back to the future horizon with probability $|r_\omega|^2$. In the spacetime which includes the collapsing body, the outgoing and incoming solutions $\{p_\omega\}$ and $\{q_\omega\}$ in the quasistationary region outside the horizon correspond to linear combinations of the $\{y_\omega\}$ and the $\{z_\omega\}$:

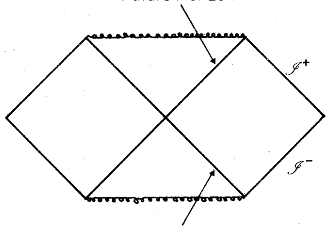

FIG. 3. The Penrose diagram of the r-t plane of the analytically extended Schwarzschild space. Null lines are at $\pm 45°$ and a conformal transformation has been made to bring infinity, represented by $\mathscr{I}^+$ and $\mathscr{I}^-$, to a finite distance. Each point in this diagram represents a sphere of area $4\pi r^2$.

$$p_\omega = t_\omega y_\omega + r_\omega z_\omega,$$
$$q_\omega = \bar{r}_\omega y_\omega - \bar{t}_\omega z_\omega. \tag{4.9}$$

The modes $\{w_\omega\}$ represent particles which, in the analytically extended Schwarzschild space, are always inside the future horizon and which do not enter the exterior region. In the real space-time with the collapsing body they correspond to particles which cross the event horizon just after its formation.

The modes $\{z_\omega\}$ have the same Cauchy data as the $\{f_\omega^{(1)}\}$, therefore they are the same everywhere, i.e.,

$$z_\omega = f_\omega^{(1)} \tag{4.10}$$

on the past horizon for $U<0$,

$$y_\omega = (1-x)^{1/2} f_\omega^{(3)}$$
$$= x^{-1/2}(1-x)^{1/2}\bar{f}_\omega^{(4)}, \tag{4.11}$$

where $x = e^{-2\pi\omega\kappa^{-1}}$. The factors $(1-x)^{1/2}$ and $x^{-1/2}(1-x)^{1/2}$ appear because of the normalization. On the past horizon for $U>0$

$$w_\omega = x^{-1/2}(1-x)^{1/2}\bar{f}_\omega^{(3)}$$
$$= (1-x)^{1/2} f_\omega^{(4)}. \tag{4.12}$$

This implies that $(1-x)^{-1/2}(y_\omega + x^{1/2}\bar{w}_\omega)$ has the same Cauchy data as $f_\omega^{(3)}$ and therefore is the same everywhere, i.e.,

$$\bar{f}_\omega^{(3)} = (1-x)^{-1/2}(y_\omega + x^{1/2}\bar{w}_\omega). \tag{4.13}$$

Similarly,

$$f_\omega^{(4)} = (1-x)^{-1/2}(w_\omega + x^{1/2}\bar{y}_\omega). \tag{4.14}$$

One can express the operator ϕ in terms of the basis $\{w_\omega, y_\omega, z_\omega\}$:

$$\phi = \int (g_\omega w_\omega + h_\omega y_\omega + j_\omega z_\omega + \text{H.c.})d\omega, \tag{4.15}$$

where the $\{g_\omega\}$ and the $\{g_\omega^\dagger\}$, etc., are the annihilation and creation operators for particles in the modes $\{w_\omega\}$, etc. Comparing (4.15) with (4.7) and using (4.13) and (4.14) one sees that

$$a_\omega^{(3)} = (1-x)^{-1/2}(h_\omega - x^{1/2}g_\omega^\dagger),$$
$$a_\omega^{(4)} = (1-x)^{-1/2}(g_\omega - x^{1/2}h_\omega^\dagger), \tag{4.16}$$
$$a_\omega^{(1)} = j_\omega.$$

One can superimpose the continuum-normalization solutions $\{f_\omega^{(1)}\}$, etc., $\{w_\omega\}$, etc. to form families of orthonormal wave-packet solutions $\{f_j^{(1)}\}$, $\{\bar{f}_j^{(3)}\}$, $\{f_j^{(4)}\}$, $\{w_j\}$, $\{y_j\}$, $\{z_j\}$. If the wave packets are sharply peaked around frequency ω, the corresponding operators $a_j^{(1)}$, etc., g_j, etc. will be related by Eq. (4.16), where the suffix ω is replaced by j and modes with the same suffix j are taken to be made up from continuum modes in the same way, i.e., they have the same Fourier transforms.

One can define a future vacuum state $|0_+\rangle$ by

$$g_j|0_+\rangle = h_j|0_+\rangle = j_j|0_+\rangle = 0. \tag{4.17}$$

One can then define states $|A;B;C\rangle$ which contain n_{1a} particles in the mode w_1, n_{2a} particles in the mode w_2, etc., n_{1b} particles in the mode y_1, etc., and n_{1c} particles in the mode z_1, etc. by

$$|A;B;C\rangle = \left[\prod (n_{ja}!)^{-1/2}(g_j^\dagger)^{n_{ja}}\right]$$
$$\times\left[\prod (n_{jb}!)^{-1/2}(h_j^\dagger)^{n_{jb}}\right]$$
$$\times\left[\prod (n_{jc}!)^{-1/2}(j_j^\dagger)^{n_{jc}}\right]|0_+\rangle. \tag{4.18}$$

The initial vacuum state $|0_-\rangle$ can be expressed as a linear combination of these states:

$$|0_-\rangle = \sum \mu(A;B;C)|A;B;C\rangle. \tag{4.19}$$

The coefficients $\mu(A;B;C)$ may be found by using Eqs. (4.8) and (4.16) which give

$$(g_k - x^{1/2}h_k^\dagger)|0_-\rangle = 0, \tag{4.20}$$

$$(h_k - x^{1/2}g_k^\dagger)|0_-\rangle = 0, \tag{4.21}$$

$$j_k|0_-\rangle = 0. \tag{4.22}$$

Equation (4.22) implies that the coefficients μ will be nonzero only for states with no particles in the $\{z_j\}$ modes, i.e., states for which $n_{jc}=0$ for all j. Equation (4.20) connects the coefficients μ for states with m particles in the w_k mode and s particles in the y_k mode, with the coefficients μ for states with $m-1$ particles in the w_k mode and $s-1$ particles in the y_k mode, i.e.,

$$(m_k)^{1/2}\mu(A[m_k];\ B[s_k];\ 0) - x^{1/2}(s_k)^{1/2}\mu(A[m-1_k];\ B[(s-1)_k];0)=0, \tag{4.23}$$

where $\mu(A[m_k];\ B[s_k];0)$ is the coefficient for the state $\{n_{1a}, n_{2a}, \ldots;\ n_{1b}, n_{2b}, \ldots;0\}$, where $n_{ka}=m$ and $n_{kb}=s$. By induction on (4.23) one sees that

$$\mu(A[m_k];\ B[s_k];0)=\delta_{ms}x^{m/2}\mu(A[0_k];\ B[0_k];0). \tag{4.24}$$

In other words, if one compares states with the same numbers of particles in all modes except the w_k mode and the y_k mode, the relative probabilities of having m and s particles, respectively, in these modes is zero unless $m=s$, in which case it is proportional to x^m. One can interpret this as saying that the particles are created in pairs in the corresponding w and y modes. The particle in the w mode enters the black hole shortly after its formation. The particle in the y mode is emitted from the black hole and will escape to infinity with probability $|t_\omega|^2$ or be reflected back into the black hole with probability $|r_\omega|^2$. The relative probabilities of different numbers of particles being emitted in the y modes correspond exactly to the probability distribution for thermal radiation.

By applying (4.24) to each value of k one obtains

$$\mu(A;B;0)=\exp\left(-\pi\kappa^{-1}\sum n_{ja}\omega_j\right)\mu(0;0;0) \tag{4.25}$$

if $\{n_{1a}, n_{2a}, \ldots\}=\{n_{1b}, n_{2b}, \ldots\}$, $\mu(A;B;0)=0$ otherwise. Strictly speaking, $\mu(0;0;0)$ is zero because in the approximation that has been used the back reaction of the created particles has been ignored and the space-time has been represented by a Schwarzschild solution of constant mass. This means that the black hole goes on emitting at a steady rate for an infinite time and therefore the probability of emitting any given finite number of particles is vanishingly small. However, if one considers the emission only over some finite period of time in which the mass of the black hole does not change significantly, Eq. (4.25) gives the correct relative probabilities of emitting different configurations of particles. Again one sees that the probabilities of emitting all configurations with some given energy are equal.

If one puts in the angular dependence Y_{lm} of the modes, one finds that because (4.13) and (4.14) connect w_ω and $\bar{y}_\omega$, they connect modes with the opposite angular momenta, (l, m) and $(l, -m)$. This means that the particles are created in pairs in the w and y modes with opposite angular momenta. Because the w modes have time dependence of the form $e^{-i\omega u}$ while the y modes have time dependence of the form $e^{i\omega u}$, there is also a sense in which they have opposite signs of energy: The y particles have positive energy and can escape to infinity while the w particles have negative energy and reduce the mass of the black hole.

The particle creation that is observed at infinity comes about because an observer at infinity divides the modes of the scalar field in a manner which is discontinuous at the event horizon and loses all information about modes inside the horizon. An observer who was falling into the black hole would not make such a discontinuous division. Instead, he would analyze the field into modes which were continuous against the event horizon. When propagated back to the past horizon, these modes would merely be blue-shifted by some constant factor and therefore would still be purely positive frequency with respect to the affine parameter U on the past horizon. Thus the observer falling into the black hole would not see any created particles.

V. THE SUPERSCATTERING OPERATOR $\$$

It was shown in Sec. III that observations at future infinity had to be described in terms of a density operator or matrix rather than a pure quantum state. The reason for this was that part of the information about the quantum state of the system was lost down the black hole. One might think that this information might reemerge during the final stages of the evaporation and disappearance of the black hole so that what one would be left with at future infinity would be a pure quantum state after all. However, this cannot be the case; there must be nonconservation of information in black-hole formation and evaporation just as there must be a nonconservation of baryon number. A large black hole formed by the collapse of a star consisting mainly of baryons will have a very low temperature. It will therefore emit most of its rest-mass energy in the form of particles of zero rest mass. By the time it becomes hot enough to emit baryons it will have lost all but a small fraction of its original mass and there will be insufficent energy avilable to emit the number of baryons that went into forming the black hole. Thus, if the black hole disappears completely, there will be nonconservation of baryon number. The situation with regard to information nonconservation is similar. The black hole is formed by the collapse of some well-ordered body with low entropy. During the quasistationary emission phase the black hole sends out random thermal radiation with a large amount of entropy. In order to end up in a pure quantum state the black hole would have to emit a similar amount of negative entropy or information in the final stages of the evaporation. However, information like baryon number requires energy and there is simply not enough energy available in the final stages of the evaporation. To carry the large amount of information needed would require the emission in the final stages of about the same number of particles as had already been emitted in the quasistationary phase.

Because one ends up with a density operator rather than pure quantum space, the process of black-hole formation and evaporation cannot be described by an S matrix. In general, the initial situation will not be a pure quantum state either because of the evaporation of black holes at earlier times. What one has therefore is an operator, which will be called the superscattering operator $\mathcal{S}$, that maps density operators describing the initial situation to density operators describing the final situation. By the superposition principle this mapping must be linear. Thus if one regards the initial and final density operators ρ_1 and ρ_2 as second-rank tensors or matrices ρ_{1AB} and ρ_{2CD} on the initial and final Hilbert spaces, respectively, the superscattering operator will be a 4-index tensor $\mathcal{S}_{ABCD}$ such that

$$\rho_{2CD} = \sum\sum \mathcal{S}_{CDAB}\rho_{1AB}. \tag{5.1}$$

When the initial situation is a pure quantum state ξ_A the initial density operator will be

$$\rho_{1AB} = \xi_A\bar{\xi}_B. \tag{5.2}$$

If the initial state is such as to have a very small probability of forming a black hole, the final situation will also be a pure quantum state ζ_C which is related to the initial state by the S matrix:

$$\zeta_C = \sum S_{CA}\xi_A. \tag{5.3}$$

The final density operator will be

$$\rho_{2CD} = \zeta_C\bar{\zeta}_D. \tag{5.4}$$

Thus the components of the $\mathcal{S}$ operator on these states can be expressed as the product of two S matrices:

$$\mathcal{S}_{CDAB} = \tfrac{1}{2}(S_{CA}S_{BD}^{-1} + S_{AD}^{-1}S_{CB}). \tag{5.5}$$

However, for initial states that have a significant probability of forming a black hole, there is no S matrix and so one cannot represent $\mathcal{S}$ in the form (5.5).

Consider, for example, the scattering of two gravitons. In this case the initial situation is a pure quantum state and, if the energy is low, the final situation will be also a nearly pure state. This can be recognized by computing the entropy of the final situation which can be defined as

$$S_2 = -\sum\sum \rho_{2CD}\ln(\rho_{2CD}). \tag{5.6}$$

In this expression the logarithm is to be understood as the inverse of the exponential of a matrix. It can be computed by transforming to a basis in which ρ_{2CD} is diagonal. For energies for which there is a low probability of forming a black hole, the entropy S_2 will be nearly zero. However, as the center-of-mass energy of the gravitons is increased to the Planck mass, there will be a significant probability of a black hole forming and evaporating and the entropy S_2 will be nonzero.

The tensor $\mathcal{S}_{CDAB}$ is Hermitian in the first and second pairs of indices. Any density matrix has unit trace because, in a basis in which it is diagonal, the diagonal entries are the probabilities of being in the different states of the basis. Since ρ_{2CD} must have unit trace for any initial density matrix ρ_{1AB},

$$\sum \mathcal{S}_{CCAB} = \delta_{AB}. \tag{5.7}$$

One can regard this as saying that, starting from any initial state, the probabilities of ending up in different final states must sum to unity. The corresponding relation

$$\sum \mathcal{S}_{CDAA} = \delta_{CD} \tag{5.8}$$

would imply that for any given final state, the probabilities of it arising from different initial states should sum to unity. Two arguments will be given for Eq. (5.8). The first is a thermodynamic argument based on the impossibility of constructing perpetual-motion machines. The second is based on CPT invariance.

Because the mass measured from infinity is conserved, the superscattering operator $\mathcal{S}$ will connect only initial and final states with the same energy. Thus (5.7) will hold when the initial and final state indices are restricted to states with some given energy E. Similarly, if (5.8) holds, it should also hold when restricted to initial and final states of energy E. For convenience, in order to make the number of states finite, consider states between energy E and $E + \Delta E$ contained in a very large box with perfectly reflecting walls. Define ψ_{CD} to be $\sum \mathcal{S}_{CDAA}$, where the summation is over the finite number of states specified above. Suppose that

$$\psi_{CD} \neq \delta_{CD}. \tag{5.9}$$

By (5.7) restricted to the same states, $\sum\psi_{CC} = N$, where N is the number of states. By transforming to a basis in which ψ_{CD} is diagonal, one can see that (5.9) would imply that there was some state ξ_C such that

$$\sum\sum\psi_{CD}\xi_C\bar{\xi}_D=\sum\sum\sum \mathcal{S}_{CDAA}\xi_C\bar{\xi}_D>1. \tag{5.10}$$

This would imply that the sum of the probabilities of arriving at the final state ξ_C from all the different possible initial states was greater than unity. If one now left the energy E in the box for a very long time, the system would evolve to various different configurations. For most of the time the box would contain particles in approximately thermal distribution. Occasionally, a large number of particles would get together in a small region and would create a black hole which would then evaporate again. To a good approximation one could regard the time development of the density matrix of the system as being given by successive applications of the $\mathcal{S}$ operator restricted to the finite number of states. On the normal assumptions of thermal equilibrium and ergodicity one would expect that after a long time the probability of finding the system in any given state would be N^{-1} and the entropy would be $\ln N$. However, if (5.10) held, the probability of the system being in the state ξ_C would be greater than N^{-1} and so the entropy would be less than $\ln N$. One could therefore extract useful energy and run a perpetual-motion machine by periodically allowing the system to relax to entropy $\ln N$. If one assumes that this is impossible, (5.8) must hold.

The second argument for Eq. (5.8) is based on CPT invariance. Because the Einstein equations are separately invariant under C, P, and T, pure quantum gravity will also be invariant under these operations if the boundary conditions at hidden surfaces are similarly invariant. The matter fields are not necessarily locally invariant under C, P, and T separately, but they are locally invariant under CPT because their Lagrangian density is a scalar under local proper Lorentz transformations. Thus the quantum theory of coupled gravitational and matter fields will be invariant under CPT provided that the boundary conditions at hidden surfaces are invariant under CPT. That the boundary conditions at hidden surfaces should be invariant under CPT would seem a very reasonable assumption. In fact, the assumption of CPT for quantum gravity and the assumption that one cannot build a perpetual-motion machine are equivalent in that each of them implies the other. With CPT invariance, Eq. (5.8) follows from (5.7). Because black holes can form when there was no black hole present beforehand, CPT implies that they must also be able to evaporate completely; they cannot stabilize at the Planck mass, as has been suggested by some authors. CPT invariance also implies that for an observer at infinity there is no operational distinction between a black hole and a white hole: The formation and evaporation of a black hole can be regarded equally well in the reverse direction of time as the formation and evaporation of a white hole.[23] An observer who falls into a hole will always think that it is a black hole but he will not be able to communicate his measurements to an observer at infinity.

ACKNOWLEDGMENTS

The author is very grateful for helpful discussions with a number of colleagues, in particular, B. J. Carr, W. Israel, D. N. Page, R. Penrose, and R. M. Wald.

*Work supported in part by the National Science Foundation under Grant No. MPS75-01398 at the California Institute of Technology.

† Sherman Fairchild Distinguished Scholar at the California Institute of Technology.

[1] P. J. Roll, R. Krotkov, and R. H. Dicke, Ann. Phys. (N.Y.) 26, 442 (1964).

[2] V. B. Braginsky and V. I. Panov, Zh. Eksp. Teor. Fiz. 61, 873 (1971) [Sov. Phys.—JETP 34, 464 (1972)].

[3] R. Penrose, Phys. Rev. Lett. 14, 57 (1965).

[4] S. W. Hawking, Proc. R. Soc. London A294, 511 (1966).

[5] S. W. Hawking, Proc. R. Soc. London A295, 460 (1966).

[6] S. W. Hawking, Proc. R. Soc. London A300, 187 (1967).

[7] S. W. Hawking and R. Penrose, Proc. R. Soc. London A314, 529 (1970).

[8] E. M. Lifshitz and I. M. Khalatnikov, Adv. Phys. 12, 185 (1963).

[9] E. M. Lifshitz and I. M. Khalatnikov, Phys. Rev. Lett. 24, 76 (1970).

[10] C. Brans and R. H. Dicke, Phys. Rev. 124, 925 (1961).

[11] A. Trautman, Colloques, CNRS Report No. 220, 1973 (unpublished).

[12] S. W. Hawking, Nature 248, 30 (1974).

[13] S. W. Hawking, Commun. Math. Phys. 43, 199 (1975).

[14] G. 't Hooft and M. Veltman, Ann. Inst. H. Poincaré 20A, 69 (1974).

[15] S. Deser and P. van Nieuwenhuizen, Phys. Rev. D 10, 401 (1974); 10, 411 (1974).

[16] R. M. Wald, Commun. Math. Phys. 45, 9 (1975).

[17] L. Parker, Phys. Rev. D 12, 1519 (1975).

[18] C. B. Collins and S. W. Hawking, Astrophys. J. 180, 317 (1973).

[19] R. Omnès, Phys. Rep. 3C, 1 (1972).

[20] Ya. B. Zel'dovich and A. Starobinsky, Zh. Eksp. Teor. Fiz. 61, 2161 (1971) [Sov. Phys.—JETP 34, 1159 (1972)].

[21] W. G. Unruh, Phys. Rev. D 14, 870 (1976).

[22] D. G. Boulware, Phys. Rev. D 12, 350 (1975).

[23] S. W. Hawking, Phys. Rev. D 13, 191 (1976).