

Black hole horizons and complementarity

Youngjai Kiem

Joseph Henry Laboratories, Princeton University, Princeton, New Jersey 08544

Herman Verlinde

*Joseph Henry Laboratories, Princeton University, Princeton, New Jersey 08544
and Institute for Theoretical Physics, University of Amsterdam, Valckenierstraat 65, 1018 XE Amsterdam,
Noord-Holland, The Netherlands*

Erik Verlinde

*TH-Division, CERN, CH-1211 Geneva 23, Switzerland
and Institute for Theoretical Physics, University of Utrecht, P.O. Box 80.006, 3508 TA Utrecht,
Utrecht, The Netherlands*

(Received 2 June 1995)

We investigate the effect of gravitational back reaction on the black hole evaporation process. The standard derivation of Hawking radiation is reexamined and extended by including gravitational interactions between the infalling matter and the outgoing radiation. We find that these interactions lead to substantial effects. In particular, as seen by an outside observer, they lead to a fast growing uncertainty in the position of the infalling matter as it approaches the horizon. We argue that this result supports the idea of black hole complementarity, which states that, in the description of the black hole system appropriate to outside observers, the region behind the horizon does not establish itself as a classical region of space-time. We also give a new formulation of this complementarity principle, which does not make any specific reference to the location of the black hole horizon.

PACS number(s): 04.70.Dy, 04.62.+v

I. INTRODUCTION

Ever since the discovery of black hole evaporation [1] there has been a continuing debate on the relevance of the gravitational back reaction to the final quantum state of the radiation. The central question is whether back reaction effects could, even in principle, bring out the information about the initial quantum state of the matter that has formed the black hole. The answer to this question would be no, essentially by assumption, if one accepts that the state of the radiation is reliably computed using free propagation of quantum fields on a fixed classical background geometry, and that the gravitational effect of the quantum radiation is accurately described via an adiabatic change of the background geometry and the mass of the black hole. According to this scenario, which has in particular been advocated by Hawking,¹ strong gravitational effects take place too late or too far behind the horizon to be able to bring out the initial information.

An opposite point of view has been put forward by 't Hooft [3], who pointed out that from the perspective of the outside observer strong gravitational interactions take place near the horizon between the infalling matter and the outgoing virtual particles describing the Hawking radiation. He argued that this interaction could drastically change the standard semiclassical picture of the evaporation process, and in particular may give rise to

a complementarity between the physical world of the infalling observer and that of the outside observer. Indeed, for the infalling observer the horizon represents a smooth region of space-time, but the Hawking particles are not measurable to him. To the outside observer detecting the Hawking radiation, on the other hand, the horizon becomes a strongly interacting region, while the black hole behind it never seems to establish itself as a classical part of space-time. According to this alternative physical picture, which recently has also been advocated by several other authors [4–8], there is no longer any *a priori* reason why quantum coherence should be destroyed during the evaporation process.

To investigate the possibility of this second scenario, we will in this paper reexamine the derivation of Hawking radiation and present a new procedure for studying the onset of gravitational back reaction effects on the radiation spectrum. Specifically, we will study the propagation of a quantum state from an initial Cauchy surface Σ_{initial} located some finite distance outside the event horizon to a final Cauchy surface Σ_{final} located at a later time and much closer to the event horizon (see Fig. 1). Both surfaces are spacelike everywhere and can roughly be thought of as constant-time slices as seen by an outside observer. Eventually, we will be interested in the time evolution of the state on Σ_{final} , as the outside part of it starts to contain more and more of the outgoing thermal radiation.

Since both Cauchy surfaces are far away from the black hole singularity, it would at first sight appear to be sufficient to use free field theory in a fixed background to

¹For more recent explanations of this point of view, see [2].

relate the physical observations made on each of these surfaces. We will find, however, that this naive expectation is incorrect. Instead we will show that gravitational interactions, which take place between the modes as they propagate from the initial to final Cauchy surface, become increasingly significant as the time on Σ_{final} progresses. These interactions are associated with two types of collisions, namely, between the infalling and virtual outgoing particles near the horizon, and secondly between the outgoing virtual particles themselves. The interaction regions of these two types of virtual processes are schematically indicated in Fig. 1(a). In this paper we will mostly be concerned with the first type of interactions.

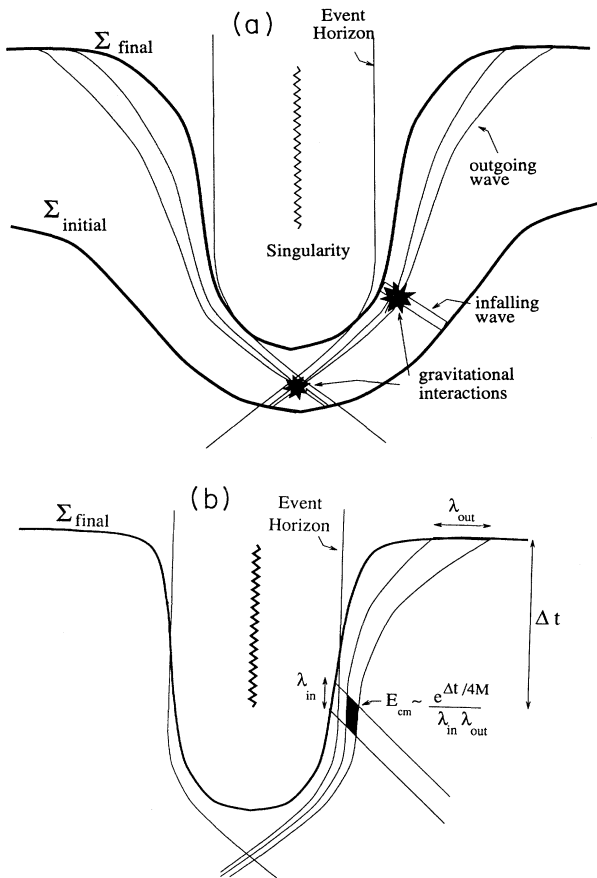


FIG. 1. (a) In this paper we will study the effect of back reaction on the propagation of quantum states in a black hole formation geometry. We will find that even for regular initial data on the initial Cauchy surface Σ_{initial} the virtual gravitational interaction between the in- and outgoing modes, as well as among the outgoing modes themselves, will become increasingly important with time on the final Cauchy surface Σ_{final} . (b) This figure shows two spacelike separated observables defined on Σ_{final} of typical wavelengths λ_{in} and λ_{out} . The field modes associated with these observables have collided in the past with a center of mass energy that grows as $\exp(\Delta t/8M)$. The proposed complementarity principle states that observables for which this collision energy exceeds some (possibly macroscopic) critical value do not simultaneously exist as mutually commuting operators.

Typically, the in- and outgoing particles involved in these processes propagate through the horizon at different angular directions. Hence, while the center of mass energies can be huge, the momentum transfer during these virtual collisions is typically small. In this special kinematical limit, there exists a rather large range of collision energies for which the quantum gravitational interaction between the particles is well controlled and can be described by means of semiclassical techniques [10].

This paper is organized as follows. To set up some notation, we briefly summarize in Sec. II the relevant formulas describing free field propagation in a fixed black hole background. In Sec. III we write the classical equations describing the gravitational interaction between in- and outgoing particles near the horizon. In Sec. IV we present a new procedure for including the quantum mechanical effects of the back reaction. Starting from the original formulas of Hawking, we will include a small quantum contribution to the infalling matter that creates the black hole, and compute its effect on the relation between the *in* and *out* modes. Finally, we will then consider the effect of these interactions on the calculation of the outgoing state.

But first we present a somewhat more concrete form of the complementarity principle put forward in [4–6]. Whether this principle is indeed dynamically realized depends on the yet unknown details of Planck scale physics, although the calculations of this paper provide some supporting evidence. An important advantage of our formulation is that it does not need to make any explicit reference to the location of the black hole horizon, and does not strictly rely on the assumed existence of a black hole S matrix.

A. The complementarity principle

Complementarity is a fundamental aspect of quantum mechanics, with familiar manifestations such as particle-wave duality and the uncertainty relations between momentum and position operators. It arises from the fact that a single quantum state occupies a finite volume of the classical phase space. This basic property of quantum mechanics also applies to a scalar field ϕ propagating on a dynamical black hole geometry. The back reaction of a classical ϕ configuration can change the background geometry, and thus quantum states are in principle supported on different classical geometries. However, we can usually ignore this fact, because we can in most situations safely truncate the Hilbert space to a subspace in which quantum gravitational interactions are guaranteed to be small, e.g., by restricting all modes of ϕ to frequencies sufficiently smaller than the Planck frequency.

Concretely, we can imagine introducing some position-dependent cutoff scale $\epsilon(x)$ on the Cauchy surface Σ . We can then represent the corresponding truncated Hilbert space on Σ by means of all states supported on field configurations with wavelengths larger than this cutoff

scale. For normal, regular Cauchy surfaces it is reasonable to expect that a constant cutoff ϵ of the order of the Planck length will still be sufficient to eliminate all strong coupling gravitational effects. However, in case the Cauchy surface contains different regions that are related via large relative boosts (such as Σ_{final} in Fig. 1), it is conceivable that a much stronger restriction of the free field Hilbert space may be needed to achieve a reliable semiclassical description.

To address this question, let us first formulate in a somewhat more precise way the criterion we would like to impose on the cutoff length scale $\epsilon(x)$. A first key point is that to a given cutoff $\epsilon(x)$ we can associate a corresponding size of stress-energy fluctuations. These stress-energy fluctuations are a necessary consequence of the presence of all modes of the second quantized scalar field up to the cutoff scale. Their typical size is determined by the behavior of the regulated expectation value $\langle T_{\mu\nu}(x)^2 \rangle_\epsilon$, and from free field theory we deduce that the quantum fluctuations of $T_{\mu\nu}$ typically grow like $\epsilon(x)^{-4}$ as the cutoff gets smaller.

Via the Einstein equations, this will result in correspondingly large quantum fluctuations in the local background geometry on the Cauchy surface Σ . For a given cutoff and Cauchy surface Σ , one can in principle calculate these by integrating the cumulative gravitational effect of the local stress-energy fluctuations. It is clear that when these geometry fluctuations become too large, relative to the cutoff scale, the semiclassical description of the Hilbert space as the space of states on a given Cauchy slice breaks down. We will call a cutoff violating this bound *supercritical*, and we will call it *subcritical* or *semiclassical* if the fluctuations of the geometry are controlled in this sense. It is clear from the above discussion that semiclassical cutoffs have a minimal size.

The question arises, however, how such a semiclassical truncation of the Hilbert space must be interpreted at a fundamental level? On the one hand, any short distance cutoff would appear to constitute an unacceptable violation of Lorentz invariance (= frame independence), since different observers will in general be inclined to truncate the Hilbert space in different ways. On the other hand, there is no correspondence principle that tells us that the Hilbert space must extend into this supercritical regime. On the contrary: if we were to assume it does extend into the supercritical regime, we would need to explain why there are no large space-time fluctuations at scales much larger than the Planck scale.

Thus it seems we are faced with a dilemma: apparently we must either give up strict Lorentz invariance, or find a way to deal with a Hilbert space supported on field configurations with arbitrarily high frequencies. A way out of this dilemma is provided by the black hole complementarity principle proposed in [3,6,5].

We now propose a new formulation of this principle, which we name *space-time complementarity*, because its formulation and possible consequences are in principle not restricted to the black hole context. The spirit in which this principle is meant is as a proposal for a reasonable effective description of some underlying, consistent theory of quantum gravity, such as the one provided by

string theory.² Clearly, however, a much more detailed knowledge of thus underlying fundamental theory is necessary to verify and quantify this effective description.

Space-time complementarity: A variable cutoff scale $\epsilon(x)$ on a Cauchy surface Σ provides a permissible semiclassical description of the second quantized Hilbert space only when the quantum fluctuations of the local background geometry induced by the corresponding stress-energy fluctuations do not exceed the cutoff scale itself. All critical cutoff scales that saturate this requirement provide complete, complementary descriptions of the Hilbert space.

The key step here is that, while different observers may under certain circumstances be inclined to use very different cutoffs, and thus very different bases of observables for doing measurements, the Hilbert space spanned by these different bases is assumed to be the same. The consequences of this assumption are particularly striking in a situation with two different observers whose reference frames are related by very large red- or blueshifts, such as the infalling and outside observer on a black hole background.

To illustrate this, let us consider the simultaneous measurement of an *outside* observable $\mathcal{O}_{\lambda_{\text{out}}}$, supported on free field configurations of typical wavelength λ_{out} , and an *inside* observable $\mathcal{O}_{\lambda_{\text{in}}}$ near or just behind the horizon, of typical wavelength λ_{in} . [See Fig. 1(b).] One would expect that such a simultaneous measurement should indeed be possible, since the two observables on Σ_{final} are spacelike separated.

To accurately compute transition amplitudes involving these operators, however, we must consider their past history. In the first instance, we can try to compute this past history using the free field propagation of the modes contained in these observables. We then discover that the past history of these observables in fact contains an ultrahigh energy collision very close to the horizon, with a center of mass energy that grows exponentially in the outgoing time. While it is true that this interaction takes place between virtual instead of real particle excitations, the absurd magnitude of the collision energy nevertheless indicates that the classical geometry is no longer the appropriate setting for considering the simultaneous measurement by these two observables.

In the following sections we will make this intuition more explicit, by showing that the relevant transition amplitudes or correlation functions indeed always involve very large stress-energy fluctuations, which collide near the horizon [see Fig. 1(b)]. The fact that this stress-energy is associated with virtual particles implies that it is of purely negative frequency in a local inertial frame. Although this means that the *in-in* expectation value of the stress-energy remains small, we will show that its

²L. Susskind has also advocated that a complementarity principle of the type formulated in this section may be realized in string theory. There indeed exist several indications that, compared to local field theory, string theory has drastically fewer degrees of freedom at short distances.

large quantum fluctuations still lead to nontrivial quantum gravitational effects that affect the calculation of the outgoing state.

Moreover, at a more fundamental level, this observation suggests that something could actually be wrong in the assumption that both these observables must be simultaneously present in the Hilbert space as two mutually commuting operators. Again, as before, there is no longer any correspondence principle that tells us that this has to be the case. Instead it seems to us that a strong case can be made for the opposite assumption, which provides the basis for the second formulation of the complementarity principle.

Space-time complementarity (kinematical). *Different microscopic observables that are spacelike separated on a Cauchy surface Σ , but have support on matter field configurations that, when propagated back in time, have collided with macroscopically large center of mass energies, are not simultaneously contained as commuting operators in the physical Hilbert space. Instead such operators are complementary.*

Let us end this section with a few short comments.

It is important to emphasize that in both formulations of the complementarity assumption no specific reference is made to the presence of the black hole horizon, nor to its location. The horizon region is therefore not considered differently from any other region of space-time.

Nevertheless, the complementarity assumption will have drastic consequences for the horizon region as seen by an observer who stays outside the black hole. This is particularly evident from the second formulation of the complementarity principle, since it immediately implies that, to an outside observer, the part of the Hilbert space that is associated with the region near or behind the horizon must be much, much smaller than would follow from free field theory. It is clear that this will have important consequences for the computation of quantities like the black hole entropy (cf. Ref. [11]).

While the kinematical and dynamical formulation are very similar in spirit, it is not immediately obvious that they are equivalent. To establish this equivalence, one would have to show that the above kinematical complementarity restriction is dynamically implied by the first principle. In other words, one would need to show that the simultaneous existence of operators like the in- and outside operators in Fig. 1(b) inevitably results in macroscopically large space-time fluctuations, thereby violating the restriction formulated in the first form of the complementarity principle. The aim of the following sections is to present evidence that this is indeed the case.

II. SCALAR FIELDS IN A BLACK HOLE GEOMETRY

In this section we consider the propagation of a free scalar field on the time-dependent geometry of a black hole that is being formed by gravitational collapse of a spherical body of matter. Here we will ignore back reaction. The metric outside the collapsing body is given by the Schwarzschild metric for a black hole with con-

stant mass M , while inside the matter distribution the metric is assumed to be regular, and not very different from Minkowski space. It is convenient to introduce the advanced and retarded time coordinates v and u which in the Schwarzschild region are defined by

$$u = t - r^*, \quad v = t + r^*, \quad (2.1)$$

with

$$r^* = r + 2M \ln \left(\frac{r}{2M} - 1 \right) - 2M,$$

where r and t represent the usual Schwarzschild radial and time coordinates. In terms of the Kruskal coordinates $x^+ = e^{v/4M}$ and $x^- = -e^{-u/4M}$ the metric outside the collapsing matter is given by

$$ds^2 = -\frac{32M^3}{r} e^{-r/2M} dx^+ dx^- + r^2 d^2\Omega, \quad (2.2)$$

where $d^2\Omega = d\theta^2 + \sin^2\theta d\varphi^2$ is the line element on the sphere.

Now consider the classical propagation of a scalar field ϕ on this geometry. Following [1] we want to determine the relation between a given outgoing wave and the corresponding initial wave that is obtained by propagating the former backwards in time through the collapsing matter. We may concentrate our discussion on the region close to the horizon, where the Klein-Gordon equation for the field ϕ takes the form

$$\left[\partial_u \partial_v - e^{(v-u)/4M} \left(-\frac{\Delta_\Omega}{r^2} + m^2 + \frac{2M}{r^3} \right) \right] \times r \phi(u, v, \Omega) = 0, \quad (2.3)$$

where Δ_Ω denotes the scalar Laplacian on the two-sphere. For our purpose we need to consider field configurations that have a finite frequency with respect to the Schwarzschild time t and a finite angular momentum.

Since an outgoing wave ϕ_{out} with a finite frequency at \mathcal{I}^+ oscillates extremely rapidly near the horizon (see Fig. 2), one may apply the geometric optics approximation to derive the form of the incoming wave ϕ_{in} in the asymptotic past [1]. In terms of the above wave equation, this procedure becomes exact in the region $r \rightarrow 2M$, where we have

$$e^{(v-u)/4M} \ll 1,$$

and thus the wave equation (2.3) simplifies to $\partial_u \partial_v \phi = 0$. In this way we find that near the horizon the field ϕ is decomposed into an incoming and an outgoing wave

$$\phi(u, v, \Omega) = \phi_{\text{in}}(v, \Omega) + \phi_{\text{out}}(u, \Omega). \quad (2.4)$$

By matching the in and out signals near the region where the horizon is formed in the initial stages of the gravitational collapse, one then finds that these in- and outgoing waves ϕ_{in} and ϕ_{out} are related via a simple reparametrization

$$\phi_{\text{in}}(v, \Omega) = \phi_{\text{out}}(u(v), {}^P\Omega), \quad (2.5)$$

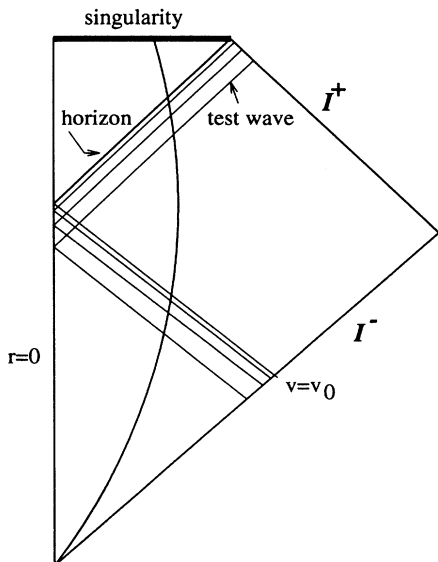


FIG. 2. The Penrose diagram for a black hole that is formed through spherical gravitational collapse. The additional lines near the horizon and $v = v_0$ indicate the light rays of a test wave that is used to determine the form of the outgoing state of the scalar field.

where $\Omega \equiv (\theta, \varphi)$ and ${}^P\Omega$ denotes the corresponding antipodal point on the two-sphere, i.e., ${}^P\Omega \equiv (\pi - \theta, \pi + \varphi)$. For large u , the reparametrization $u(v)$ takes the asymptotic form

$$u(v) = v_0 - 4M \ln \left(\frac{v_0 - v}{4M} \right) + \text{const}, \quad (2.6)$$

where v_0 is the limiting value for the null coordinate v describing the location of the incoming radial light ray that eventually coincides with the global event horizon. The constant on the right-hand side of (2.6) depends on the details of the gravitational collapse, and is of order M . In the following we will drop this constant, because it will not be important for our discussion. The first term v_0 is necessary to ensure that our equations are invariant under time translations which act as simultaneous shifts of u , v , and v_0 . We note that the coordinate relation (2.6) is noninvertible, since it maps the domain $v < v_0$ onto the complete range $-\infty < u < \infty$. The equations (2.5) and (2.6) play a central role in the derivation of Hawking radiation [1].

To describe the quantum physics near the horizon, the classical field variables $\phi_{\text{in}}(v, \Omega)$ and $\phi_{\text{out}}(u, \Omega)$ are replaced by second-quantized field operators. The standard free field canonical commutation relations for the incoming fields take the form

$$[\phi_{\text{in}}(v_1, \Omega_1), \partial_{v_2} \phi_{\text{in}}(v_2, \Omega_2)] = -2\pi i \delta(v_{12}) \delta^{(2)}(\Omega_{12}), \quad (2.7)$$

and a similar commutation relation holds for the outgoing fields. The initial Hilbert space on \mathcal{I}^- is generated by the creation operators obtained from ϕ_{in} :

$$\phi_{\text{in}}(v, \Omega) = \sum_{l,m} \int_{-\infty}^{\infty} \frac{d\omega}{\sqrt{2\pi\omega}} (a_{\omega lm}^\dagger e^{i\omega v} + a_{\omega lm} e^{-i\omega v}) \times Y_{lm}(\Omega). \quad (2.8)$$

For each given initial state $|\psi\rangle_{\text{in}}$, we would like to calculate the quantum state $|\psi\rangle_{\text{fin}}$ on a final Cauchy slice, which asymptotically approaches the Cauchy surface $\mathcal{I}^+ \cup \mathcal{H}^+$, formed by asymptotic future infinity and the event horizon.

We can expand the outgoing field on \mathcal{I}^+ as

$$\phi_{\text{out}}(u, \Omega) = \sum_{l,m} \int_{-\infty}^{\infty} \frac{d\omega}{\sqrt{2\pi\omega}} (b_{\omega lm} e^{i\omega u} + b_{\omega lm}^\dagger e^{-i\omega u}) \times Y_{lm}(\Omega) \quad (2.9)$$

to obtain the creation operators that generate the out Hilbert space. Ignoring the back reaction, these out modes can be expressed in terms of the in modes, by expanding the relation (2.6) describing the propagation backwards in time to \mathcal{I}^- in modes. One finds the linear relation

$$b_{\omega lm} = \int_0^\infty d\omega' (\alpha_{\omega\omega'} a_{\omega' lm} + \beta_{\omega\omega'} a_{\omega' lm}^\dagger), \quad (2.10)$$

$$b_{\omega lm}^\dagger = \int_0^\infty d\omega' (\alpha_{\omega\omega'}^* a_{\omega' lm}^\dagger + \beta_{\omega\omega'}^* a_{\omega' lm}).$$

Up to some irrelevant phase, the Bogolyubov coefficients have the asymptotic form

$$\alpha_{\omega\omega'} = e^{-i(\omega' - \omega)v_0} \frac{e^{2\pi M\omega} \Gamma(1 - i4M\omega)}{2\pi \sqrt{\omega(\omega' + i\epsilon)}}, \quad (2.11)$$

$$\beta_{\omega\omega'} = e^{+i(\omega' + \omega)v_0} \frac{e^{-2\pi M\omega} \Gamma(1 - i4M\omega)}{2\pi \sqrt{\omega(\omega' - i\epsilon)}}.$$

This relation between the asymptotic modes completely determines the asymptotic form of the outgoing state on future infinity \mathcal{I}^+ . It is described by a mixed state, since the outgoing fields on \mathcal{I}^+ cover only part of the final Cauchy surface. We refer to Hawking's original paper [1] for a more detailed discussion of the properties of this mixed state.

III. CLASSICAL BACK REACTION

One of the central assumptions in the standard derivation [1] of the outgoing radiation spectrum is that the incoming and outgoing fields can to a very good approximation be treated as free fields. In particular, it is assumed that the commutator between them

$$[\phi_{\text{in}}(v, \Omega_1), \phi_{\text{out}}(u, \Omega_2)] \quad (3.1)$$

vanishes for $v > v_0$. The underlying classical intuition is that the fields $\phi_{\text{in}}(v, \Omega)$ with $v > v_0$ will propagate without much disturbance into the region behind the black hole horizon, and thus become unobservable from the out

side. However, this intuition ignores the gravitational interactions between the in- and outgoing particles. Our aim in the following is to investigate the consequences of these interactions for the derivation of the final state.

First let us briefly recall why in principle one could expect that gravitational interactions can become important. An important feature of the mapping (2.5) is that an outgoing wave with a certain frequency ω translates into a in wave with infinitely many oscillations along $v = v_0$. For example, in the s -wave sector we have

$$e^{i\omega u(v)} = e^{i\omega v_0} \left(\frac{v_0 - v}{4M} \right)^{-4iM\omega} \theta(v_0 - v). \quad (3.2)$$

Hence a generic out wave, when propagated back in time, carries a very large outgoing stress-energy near the horizon. Although for the propagation of a nonsingular initial state this stress-energy manifests itself only in the form of *virtual* fluctuations, i.e., of purely negative frequency in a local inertial frame, it can in principle still lead to nontrivial gravitational effects. In order to investigate this point, we will begin with a description of these interactions at the classical level.

A. Classical dynamics of the horizon

As a preparation, let us study the effect of a spherical shell of matter with energy δM that falls into the black hole at some late advanced time v_1 . In particular, we want to know how this influences the advanced time v_0 at which the global event horizon forms. It is clear that due to the additional matter the Schwarzschild radius of the black hole increases by an amount $2\delta M$, and so after $v = v_1$ the global event horizon coincides with the new black hole horizon at $r = 2M + 2\delta M$. By tracing the corresponding light rays back to the origin $r = 0$ we discover that the global event horizon originates at a time $v_0 + \delta v_0$ that is slightly earlier than v_0 (see Fig. 3). Explicitly, we find

$$\delta v_0 = -4\delta M e^{-(v_1 - v_0)/4M}. \quad (3.3)$$

Even though this variation seems negligible, it leads to a significant effect on outgoing light rays. As illustrated in Fig. 3, a light ray that originally would have reached the outside observer at some retarded time u will as a result of the shift

$$v_0 \rightarrow v_0 + \delta v_0$$

arrive at a much later time $u + \delta u$. Using (2.6) one easily shows that

$$\delta u(u) = -4M \ln \left(1 + \frac{\delta v_0}{4M} e^{(u - v_0)/4M} \right). \quad (3.4)$$

Notice that even for a very small perturbation $\delta v_0 < 0$ the time delay $\delta u(u)$ becomes infinite at a finite time $u_{\text{lim}} - v_0 \sim -4M \ln(|\delta v_0|/M)$. The physical interpretation of this fact is that a light ray that is on its way to reach the asymptotic observer at some time $u > u_{\text{lim}}$ will as a result of the infalling shell cross the event hori-

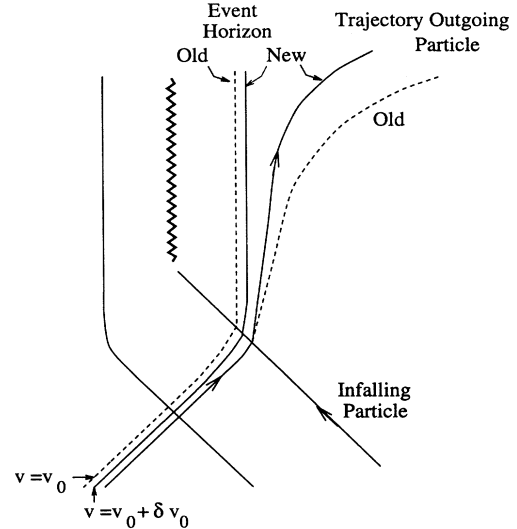


FIG. 3. An infalling shell of matter changes the position of the horizon by a small amount, but due to the redshift it has a large effect on the trajectories of outgoing light rays.

zon and be trapped inside the black hole horizon. It follows that an outgoing wave $\phi_{\text{out}}(u, \Omega)$ corresponding to a given incoming wave $\phi_{\text{in}}(v, \Omega)$ is transformed, as a result of the additional infalling matter, into

$$\phi_{\text{out}} \rightarrow \phi_{\text{out}}(u + \delta u(u), \Omega),$$

where $\delta u(u)$ is given above. This is quite a dramatic effect for a generic out wave.

Let us now turn to the question of how to incorporate the gravitational self-interactions of ϕ into the description of the wave propagation. The basic observation is that the presence of the scalar field ϕ also leads to small changes in the black hole mass M and in the position of the horizon due to incoming energy flux carried by the T_{vv} component of its stress-energy tensor. It is easy to show that the stress-energy $T_{vv} = -\frac{1}{2}(\partial_v \phi_{\text{in}})^2$ in a small interval between v_1 and $v_1 + \Delta v_1$ induces a change in the mass M equal to

$$\delta M = \int d^2\Omega \int_{v_1}^{v_1 + \Delta v_1} dv T_{vv}(v, \Omega). \quad (3.5)$$

Just as in the case of an infalling matter shell, this leads to a small correction δv_0 in the formation time v_0 . At first it may seem a good idea to describe this effect by substituting (3.5) into (3.3). However, an important difference from the previous situation is that the incoming stress-energy T_{vv} is not necessarily spherically symmetric, and therefore it is reasonable to expect that the shift δv_0 depends on the angular direction Ω . Indeed, within a certain linearized approximation of the Einstein equations one can derive, along the lines of [9], the following expression for δv_0 :

$$\delta v_0(\Omega_1) = 8 \int d^2\Omega_2 f(\Omega_1, \Omega_2) P_{\text{in}}(\Omega), \quad (3.6)$$

$$P_{\text{in}}(\Omega) = \int_{v_0}^{\infty} dv e^{(v_0-v)/4M} T_{vv}(v, \Omega),$$

where $f(\Omega, \Omega')$ satisfies

$$(\Delta_{\Omega} - 1)f(\Omega_1, \Omega_2) = -2\pi\delta^{(2)}(\Omega_{12}). \quad (3.7)$$

Here Δ_{Ω} denotes the Laplacian on the sphere. The function $f(\Omega_1, \Omega_2)$ describes the response of the position of the horizon at angular direction Ω_1 , due to a localized stress-energy influx from the direction Ω_2 . Notice that the expression (3.6) contains the same exponential factor as in (3.3), and thus for field configurations with a finite energy only represents a very small correction to v_0 . The derivation of the result (3.6) is summarized in the Appendix.

Let us make a short comment about the choice of the lower integration limit at $v = v_0$. In reality this lower limit is not sharply determined, because we should also take into account the gravitational back reaction due to infalling particles for $v < v_0$. As a practical way of dealing with this technical complication, we will in the following simply adopt as a *model* that all gravitational interactions are simply turned off below the critical line $v = v_0$. Classically, this indeed seems a reasonable procedure, since incoming particles at $v < v_0$ will not fall into the black hole (provided their energy is not too large) and will therefore not generate any shift in the critical time.

Now let us return to the problem of wave propagation on the black hole background. As reviewed in Sec. II, the outgoing fields are related to the in fields $\phi_{\text{in}}(v)$ for $v < v_0$ via the time evolution from \mathcal{I}^- to \mathcal{I}^+ . We now propose to incorporate the effect of the back reaction in the relation between ϕ_{in} and ϕ_{out} via the substitution $v_0 \rightarrow v_0 + \delta v_0(\Omega)$ in Eq. (2.6). In this way we obtain

$$\phi_{\text{out}}(u, \Omega) = \phi_{\text{in}}(v(u) + \delta v_0(\Omega), {}^P\Omega) \quad (3.8)$$

where

$$v(u) = v_0 - 4Me^{(v_0-u)/4M}, \quad (3.9)$$

and $\delta v_0(\Omega)$ is given above.

IV. QUANTUM MECHANICAL BACK REACTION

In this section we will study how these interactions can be incorporated at the quantum level. In particular, we are interested in how they affect the propagation of the ϕ quantum state. The idea will be to follow the original work of Hawking [1] as much as possible, except that we replace the relation (2.5) between the in- and outgoing waves by its corrected version (3.8). Hence the relation between the asymptotic in and out waves becomes non-linear.

In the following discussion, an important role will be

played by the in- and outgoing components of the stress-energy tensor. Recall that as quantum operators T_{vv} and T_{uu} not only measure the energy flux, but are also the generators of coordinate transformations in the v and u coordinates. For example, the commutation relation of T_{vv} with ϕ_{in} reads

$$\begin{aligned} [T_{vv}(v_1, \Omega_1), \phi_{\text{in}}(v_2, \Omega_2)] \\ = 2\pi i \delta^{(2)}(\Omega_{12}) \delta(v_{12}) \partial_v \phi_{\text{in}}(v, \Omega_2), \end{aligned} \quad (4.1)$$

and a similar relation holds between T_{uu} and the outgoing field ϕ_{out} .

A. The algebra of in and out fields

An immediate consequence of the gravitational back reaction is that it invalidates the assumption that the asymptotic in- and outgoing fields $\phi_{\text{in}}(v, \Omega)$ at $v > v_0$ and $\phi_{\text{out}}(u, \Omega)$ can be treated as independent, commuting variables. As we will now show, the interaction described above implies that this commutator is in fact replaced by a nontrivial and nonlocal ‘‘exchange algebra.’’

We will assume that the gravitational interaction can be incorporated in the relation between quantum operators ϕ_{in} and ϕ_{out} via the semiclassical procedure described above, by including the correction (3.6) to the critical time v_0 . The correspondence principle guarantees that, within a reasonable energy range, this is a valid approximation. The diffeomorphism between the asymptotic in and out waves thus depends on the quantum stress-energy tensor $T_{vv} = -\frac{1}{2}(\partial_v \phi_{\text{in}})^2$ contained in δv_0 . This implies that, as a quantum operator, the new critical line $v_0 + \delta v_0(\Omega)$ no longer commutes with the incoming fields. Using (3.6) and (4.1), we find

$$\begin{aligned} [\delta v_0(\Omega_1), \phi_{\text{in}}(v, \Omega_2)] \\ = -16\pi i f(\Omega_1, \Omega_2) e^{(v_0-v)/4M} \partial_v \phi_{\text{in}}(v, \Omega_2). \end{aligned} \quad (4.2)$$

It will be useful, therefore, to make the dependence of the outgoing variables on $\delta v_0(\Omega)$ explicit. To this end, we note that the relation (3.8) and (3.9) can formally be inverted as³

$$\begin{aligned} \phi_{\text{out}}(u, \Omega) = \exp[-e^{(u-v_0)/4M} \delta v_0(\Omega) \partial_u] \\ \times \phi_{\text{in}}(v(u), {}^P\Omega). \end{aligned} \quad (4.3)$$

When we combine this relation with (4.2), a straightfor-

³The fact that δv_0 is an operator-valued quantity in principle could introduce a problem with normal ordering at higher orders in this expansion. In the first instance, however, we will ignore this point and simply exponentiate the linearized interaction between the in- and outgoing modes. This procedure amounts to the ladder or eikonal approximation to linearized gravity, which, in the kinematical regime of interest, is known to provide the correct leading order result [10].

ward calculation shows that the incoming and outgoing fields satisfy the exchange algebra

$$\begin{aligned} & \phi_{\text{out}}(u, \Omega_1) \phi_{\text{in}}(v, \Omega_2) \\ &= \exp[-16\pi i f(\Omega_1, \Omega_2) e^{(u-v)/4M} \partial_v \partial_u] \\ & \quad \times \phi_{\text{in}}(v, \Omega_2) \phi_{\text{out}}(u, \Omega_1). \end{aligned} \quad (4.4)$$

This result is valid for v sufficiently later than v_0 and for Ω_1 not too close to Ω_2 .

The above formula (4.4) is closely related to 't Hooft's two-particle S matrix for Planckian scattering [10] in the limit of low momentum transfer. Note that it is symmetric in ϕ_{in} and ϕ_{out} , even though the starting point (3.8) seemed to be asymmetric. A possible way to understand this fact is that the exponentiation in (4.4) results from summing the contributions from multigraviton exchange in the eikonal approximation [10]. It is further important to note that the result (4.4) only represents the *onset* of the gravitational interaction between the infalling matter and the outgoing radiation. The result is valid in a limited regime, because when the center of mass energy between the in and out particles gets too large (or when the angular positions Ω_1 and Ω_2 come too close) nonlinear higher order effects will become dominant.

B. The field operators at the horizon

The fact that ϕ_{in} and ϕ_{out} do not commute is physically reasonable, because the infalling matter that is in the causal past of the operator ϕ_{out} in principle influences the geometry on which the outgoing field has propagated. However, the above result not only represents the effect of the infalling matter on the outgoing wave, but it also implies that there is a nontrivial back reaction effect on the infalling matter due to the presence of the outgoing fields. As the infalling wave approaches the horizon, causality dictates that the local field operator ϕ must commute with asymptotic operators ϕ_{out} that describe the outgoing radiation. From this we may conclude that the incoming field ϕ_{in} and the field at the horizon are not the same operator, but are related via a nontrivial evolution operator.

To distinguish ϕ_{in} from the field at the horizon, let us denote the latter by ϕ_{hor} . We will try to determine the proper definition of the field ϕ_{hor} in terms of ϕ_{in} by the condition that ϕ_{hor} , at least formally, satisfies

$$[\phi_{\text{hor}}(v, \Omega_1), \phi_{\text{out}}(u, \Omega_2)] = 0. \quad (4.5)$$

The idea for finding such operators ϕ_{hor} is to look for an *interaction representation* of the out fields of the form

$$\phi_{\text{out}}(u, \Omega) = \mathcal{U} \phi_{\text{in}}(v(u), \Omega) \mathcal{U}^{-1}, \quad v < v_0, \quad (4.6)$$

with $v(u)$ defined in (3.9) and where \mathcal{U} is some operator acting on the in Hilbert space representing the gravitational correction. Intuitively, we may think of \mathcal{U} as the time-ordered exponential of interaction Hamiltonian that describes the gravitational self-interaction of ϕ . Note

that the relation (4.6) manifestly respects the canonical commutation rules for ϕ_{out} and ϕ_{in} .

An equation of the form (4.6) would indeed immediately help us in finding the field ϕ_{hor} near the horizon, since one deduces from it that the operator $\mathcal{U} \phi_{\text{in}}(v, \Omega) \mathcal{U}^{-1}$ commutes with ϕ_{out} for $v > v_0$. This suggests that we should take

$$\phi_{\text{hor}}(v, \Omega) = \mathcal{U} \phi_{\text{in}}(v, \Omega) \mathcal{U}^{-1}, \quad v > v_0, \quad (4.7)$$

as the relation defining the fields at the horizon.

Before we determine the operator \mathcal{U} , we want to point out that by assuming the existence of relations of the form (4.6) and (4.7) we implicitly assume that all infalling modes that are supported at $v > v_0$ also fall into the black hole, and produce corresponding horizon modes. We will see momentarily that this assumption can only safely be made, within our approximation scheme, when we restrict ourselves to infalling modes at sufficiently late times $v > v_0 + \Delta v$.

A useful clue that will help us find the operator \mathcal{U} is the apparent symmetry of the algebra (4.4) between the incoming fields ϕ_{in} and outgoing fields ϕ_{out} . We can make this symmetry more manifest by introducing, besides the operator $P_{\text{in}}(\Omega)$ defined in (3.6), a similar expression

$$P_{\text{out}}(\Omega) = \int du e^{(u-v_0)/4M} T_{uu}(u, \Omega) \quad (4.8)$$

in terms of the outgoing fields. By using the commutator relation

$$\begin{aligned} & [P_{\text{out}}(\Omega_1), \phi_{\text{out}}(u, \Omega_2)] \\ &= -2\pi i \delta^{(2)}(\Omega_{12}) e^{(u-v_0)/4M} \partial_u \phi_{\text{out}}(u, \Omega_2) \end{aligned} \quad (4.9)$$

we can now replace the differential operator in the exponent of (4.3), and formally rewrite the basic relation (3.8) between ϕ_{in} and ϕ_{out} precisely as in (4.6), with the operator \mathcal{U} given by the expression

$$\mathcal{U} = \exp \left[i \int \int d\Omega_1 d\Omega_2 P_{\text{out}}(\Omega_1) f(\Omega_1, \Omega_2) P_{\text{in}}(\Omega_2) \right]. \quad (4.10)$$

Inserting this result into (4.7) gives the formal definition of the fields at the horizon.

It is instructive to work out the right-hand side of Eq. (4.7) using the commutation relation (4.1). A simple computation gives

$$\phi_{\text{hor}}(v, \Omega) = \phi_{\text{in}}(v - \Delta v(v, \Omega), \Omega), \quad (4.11)$$

where

$$\begin{aligned} \Delta v(v, \Omega) &= 4M \ln \left(1 + e^{(v_0-v)/4M} \right. \\ & \quad \left. \times \int d^2\Omega' f(\Omega, \Omega') P_{\text{out}}(\Omega') \right). \end{aligned} \quad (4.12)$$

A more direct derivation of Eq. (4.11) is given in the Ap-

pendix, where it is shown to represent the effect of metric fluctuations close to the horizon. As we see from this equation, the infalling matter fields close to the horizon are related to the original in fields via an operator-valued shift $v \rightarrow v - \Delta v$. The magnitude of this shift is determined by the *outgoing* stress-energy flux Eq. (4.8). It will become clear in the following that, for as long as our approximation scheme is valid, the quantum mechanical uncertainty in this quantity Δv will grow very fast.

A second important comment about (4.11) is that it indicates that our description of the gravitational interaction near the horizon can only be trusted as long as the argument of both ϕ_{hor} and ϕ_{in} is (sufficiently) larger than v_0 . In other words, the definition (4.11) of the infalling field at the horizon requires that

$$v > v_0 + \Delta v. \quad (4.13)$$

What this means is that for initial infalling fields $\phi_{\text{in}}(v, \Omega)$ closer to v_0 we can no longer say with certainty that these will reach the black hole horizon. A more careful analysis of the interactions in the region near $v = v_0$ will be required to determine the fate of these fields.

C. Back reaction of the Hawking state

We will now describe how these results can be used to investigate the effect of back reaction on the propagation of a quantum state. To this end, we return to the setup as described at the end of Sec. II, where we introduced oscillator bases for the asymptotic in and out Hilbert spaces. The out modes b_ω do not generate the complete final Hilbert space, however, and thus we will now also introduce a mode basis of the Hilbert space near the horizon \mathcal{H}^+ . A convenient definition of modes is

$$\begin{aligned} \phi_{\text{hor}}(v, \Omega) = & \sum_{l,m} \int_{-\infty}^{\infty} \frac{d\omega}{\sqrt{2\pi\omega}} \\ & \times (c_{\omega lm} e^{i\omega \tilde{u}(v)} + c_{\omega lm}^\dagger e^{-i\omega \tilde{u}(v)}) Y_{lm}(\Omega), \end{aligned} \quad (4.14)$$

where

$$\tilde{u}(v) = 4M \ln(v - v_0). \quad (4.15)$$

Combined together, the two sets of b and c modes generate the complete final Hilbert space, and thus the in modes can be expressed in terms of them. Using standard results as summarized in Sec. II together with the above interaction representation of the back reaction, one finds after a straightforward calculation (suppressing the l, m labels)

$$\begin{aligned} \mathcal{U}^{-1} a_\omega \mathcal{U} = & \int_0^\infty d\omega' [\alpha_{\omega'\omega}^* (b_{\omega'} - e^{-4\pi M \omega'} c_{\omega'}^\dagger) \\ & - \beta_{\omega'\omega} (b_{\omega'}^\dagger - e^{4\pi M \omega'} c_{\omega'})], \end{aligned} \quad (4.16)$$

$$\begin{aligned} \mathcal{U}^{-1} a_\omega^\dagger \mathcal{U} = & \int_0^\infty d\omega' [\alpha_{\omega'\omega} (b_{\omega'}^\dagger - e^{-4\pi M \omega'} c_{\omega'}) \\ & - \beta_{\omega'\omega}^* (b_{\omega'} - e^{4\pi M \omega'} c_{\omega'}^\dagger)], \end{aligned}$$

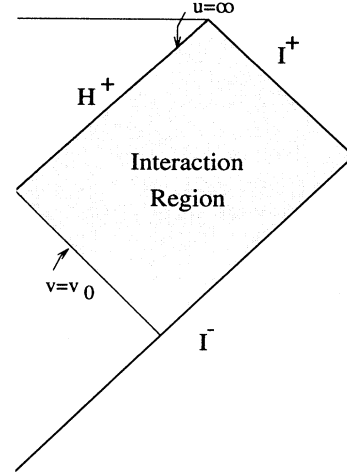


FIG. 4. In our model we assume that the interaction region between the in- and outgoing modes is bounded by the critical line $v = v_0$, the initial Cauchy surface at \mathcal{I}^- , and the final Cauchy surface near $\mathcal{H}^+ \cup \mathcal{I}^+$.

where $\alpha_{\omega\omega'}$ and $\beta_{\omega\omega'}$ are given in (2.11). The form of the outgoing state corresponding to the initial vacuum state is now easily found. The first expression for the annihilation mode a_ω shows that the final state $|\psi\rangle_{\text{final}}$ corresponding to the a vacuum $|0\rangle_{\text{in}}$ satisfies

$$(b_\omega - e^{-4\pi M \omega} c_\omega^\dagger) \mathcal{U}^{-1} |\psi\rangle_{\text{final}} = 0, \quad (4.17)$$

$$(b_\omega^\dagger - e^{4\pi M \omega} c_\omega) \mathcal{U}^{-1} |\psi\rangle_{\text{final}} = 0.$$

These equations can be readily solved, if we assume that $|\psi\rangle_{\text{final}}$ lies in the tensor product of the Fock spaces for the b and c modes. One finds

$$|\psi\rangle_{\text{final}} = \mathcal{U} |\psi\rangle_{\text{Hawking}}, \quad (4.18)$$

where \mathcal{U} is the gravitational high energy S matrix given in (4.10) and

$$\begin{aligned} |\psi\rangle_{\text{Hawking}} = & \frac{1}{N} \exp \left\{ \sum_{l,m} \int_0^\infty d\omega e^{-4\pi M \omega} b_{\omega lm}^\dagger c_{\omega l, -m}^\dagger \right\} \\ & \times |0\rangle_b \otimes |0\rangle_c, \end{aligned} \quad (4.19)$$

where N is a normalization constant.

Equation (4.18) for the gravitationally corrected final state is still rather formal. In the first place, we need to be more specific about the integration region in the (u, v) plane that is used in the definition (4.10) of \mathcal{U} . The idea of our approximation procedure is to only include the gravitational shift interaction between the in and out modes very close to the horizon, and as stated before we simply wish to ignore all interactions near and below $v = v_0$. Moreover, we have assumed in our description that the shift interaction extends all the way to $v = v_0$. The integration region in the (u, v) plane that we will use, therefore, is as indicated in Fig. 4.

Secondly, there are subtleties that arise from the operator nature of the stress-energy tensors contained in \mathcal{U} . In particular, we need to prescribe a specific operator

ordering. The appropriate prescription here seems to be to use time ordering, as is dictated via the identification of \mathcal{U} as the time-ordered exponential of the gravitational interaction Hamiltonian. Adopting this prescription, we write \mathcal{U} as

$$\begin{aligned} \mathcal{U} &= T \exp \left(i \int dt H_{\text{int}} \right) \\ &= T \exp \left[i \int \int_{v_0} du dv e^{(u-v)/4M} \right. \\ &\quad \left. \times \int \int d\Omega_1 d\Omega_2 T_{uu}(u, \Omega_1) f(\Omega_1, \Omega_2) T_{vv}(v, \Omega_2) \right], \end{aligned} \quad (4.20)$$

where the symbol T in front denotes time ordering. This formula summarizes the leading order gravitational interaction between the in- and outgoing fields at low momentum transfer, as obtained via the eikonal approximation.

As it stands, however, this expression still contains infinities that arise from the singular short distance expansion of T_{vv} and T_{uu} with themselves. In the following we will assume that these infinities are regularized with the help of some proper distance cutoff⁴ ϵ . We will comment further on this procedure and its physical meaning in Sec. IV E.

D. Fluctuations of the Hawking state

We would eventually like to get some insight into the size of the gravitational corrections we just calculated. To this end, we first consider the magnitude of the stress-energy fluctuations in the Hawking state, i.e., the final state, before including the back reaction. For this purpose, it will be convenient to choose as a basis of the final Hilbert space the coherent state basis for the b and c modes:

$$|\varphi\rangle = \exp \left[\int_0^\infty d\omega (\varphi_\omega b_\omega^\dagger + \varphi_{-\omega} c_\omega^\dagger) \right] |0\rangle_b \otimes |0\rangle_c. \quad (4.21)$$

These satisfy the relations $b_\omega|\varphi\rangle = \varphi_\omega|\varphi\rangle$ and $c_\omega|\varphi\rangle = \varphi_{-\omega}|\varphi\rangle$.

Using the expression (4.19) for the Hawking state, one easily computes that (to simplify the expressions we again suppress the angular dependence of the fields)

$$\langle \psi_{\text{Hawking}} | \varphi \rangle = \frac{1}{N} \exp[iS_0(\varphi)] \quad (4.22)$$

with

$$S_0(\varphi) = \int_0^\infty d\omega e^{-4\pi M\omega} \varphi_\omega \varphi_{-\omega}. \quad (4.23)$$

This result can in fact be rederived from a semiclassical saddle-point approximation. The overlap (4.22) is equal

to the transition element ${}_{\text{in}}\langle 0 | \varphi \rangle$, evaluated in the free scalar field theory on the black hole background. The transition element can be represented as a functional integral over all fluctuations of the scalar field ϕ with specific boundary conditions at \mathcal{I}^- and $\mathcal{I}^+ \cup \mathcal{H}^+$ (\mathcal{H}^+ denotes the event horizon) determined by the initial state ${}_{\text{in}}\langle 0 |$ and final coherent state $|\varphi\rangle$. Imposing vacuum boundary conditions at \mathcal{I}^- implies that ϕ has no positive energy modes, while at $\mathcal{I}^+ \cup \mathcal{H}^+$ the negative frequency part of ϕ is prescribed by the coherent state $|\varphi\rangle$.

The expression $S_0(\varphi)$ can be identified with the classical free field theory action of the saddle-point configuration ϕ_{cl} associated with the coherent state $|\varphi\rangle$:

$$S_0(\varphi) = S_{\text{free}}(\phi_{\text{cl}}) \quad (4.24)$$

with ϕ_{cl} given by the matrix element

$$\phi_{\text{cl}} = {}_{\text{in}}\langle 0 | \phi | \varphi \rangle. \quad (4.25)$$

We find that the relevant classical solution associated with the overlap (4.22) is given by

$$\begin{aligned} \phi_{\text{cl}}^{\text{in}}(v) &= \int_0^\infty d\omega \frac{1}{\sqrt{\omega}} \left[\varphi_\omega \left(\frac{v_0 - v}{4M} \right)_+^{-4iM\omega} \right. \\ &\quad \left. + \varphi_{-\omega} \left(\frac{v - v_0}{4M} \right)_+^{4iM\omega} \right], \end{aligned} \quad (4.26)$$

$$\phi_{\text{cl}}^{\text{out}}(u) = \int_0^\infty d\omega \frac{1}{\sqrt{\omega}} [\varphi_\omega e^{i\omega u} + e^{-4\pi M\omega} \varphi_{-\omega} e^{-i\omega u}],$$

where the subscript $+$ indicates that the corresponding function is defined to be analytic in the upper half complex v plane. These configurations should be thought of as the quantum fluctuations that are responsible for the production of the Hawking radiation.

The magnitude of the stress-energy tensor $T_{\mu\nu}(\phi_{\text{cl}})$ depends on the value of the parameters φ_ω , but also on the behavior of the modes $e^{i\omega u}$ and $(v - v_0)^{-4iM\omega}$ in the various regions of the black hole geometry. The components that potentially become large are T_{vv} and T_{uu} . We first consider the behavior of the T_{vv} component associated with the incoming field $\phi^{\text{in}}(v)$. We find

$$\begin{aligned} T_{vv}(\phi_{\text{cl}}) &= \frac{16M^2}{(v - v_0)^2} \int_0^\infty d\omega \left[T_\omega \left(\frac{v_0 - v}{4M} \right)_+^{-4iM\omega} \right. \\ &\quad \left. + T_{-\omega} \left(\frac{v - v_0}{4M} \right)_+^{4iM\omega} \right], \end{aligned} \quad (4.27)$$

where

$$\begin{aligned} T_\omega &= \int_0^\omega d\omega' \frac{1}{2} \sqrt{\omega'(\omega - \omega')} \varphi_{\omega'} \varphi_{\omega - \omega'} \\ &\quad + \int_0^\infty d\omega' e^{-4\pi M\omega'} \sqrt{\omega'(\omega + \omega')} \varphi_{-\omega'} \varphi_{\omega + \omega'} \end{aligned} \quad (4.28)$$

and a similar expression can be given for T_{-u} ($\omega > 0$).

We see that for finite nonvanishing values for the parameters φ_ω the stress-energy tensor is in general singular near $v = v_0$. The T_{uu} component, representing the

⁴This short distance cutoff must regulate the short distance singularities in the longitudinal (u, v) plane, as well as in the transverse Ω direction.

outgoing stress-energy flux, is related to T_{vv} by reflection off the $r = 0$ boundary. In the Kruskal coordinate $x^- = -e^{-u/4M}$ the outgoing stress-tensor is generically also singular near the horizon as soon as the φ_ω differ by a small amount from their exact expectation value. It is easy to convince oneself that such fluctuations are also really present. The average magnitude of φ_ω in the Hawking state $|\psi\rangle_{\text{Hawking}}$ is

$$\overline{\varphi_\omega \varphi_{\omega'}^*} = \langle n_\omega + 1 \rangle \delta(\omega - \omega') \quad (4.29)$$

and this indicates that the fluctuations of the (absolute value)² of φ_ω are comparable to the fluctuations in the particle number density observed in the out state. Thus for generic coherent out states the parameters φ_ω will indeed differ by a finite amount from their average value, and this will result in large, super-Planckian stress-energy fluctuations near the horizon.

E. Gravitational corrections and complementarity

We will now comment on the expressions (4.18)–(4.20) for the corrected final state. The following remarks will be mostly qualitative, as some of the relevant calculations are left for a future publication.

Let us first slowly turn on the back reaction. Using the expressions (4.18) and (4.20), we obtain the leading order correction to the final state Eq. (4.22), by adding to the classical free field action $S_{\text{free}}(\phi_{\text{cl}})$ an interaction term

$$\langle \psi_{\text{final}} | \varphi \rangle = \frac{1}{N} \exp[iS_0(\varphi) + iS_{\text{int}}(\varphi)] \quad (4.30)$$

with $S_{\text{int}} = \int dt H_{\text{int}}$ as given in Eq. (4.20). In leading order, this interaction term can be evaluated on the unperturbed classical field configuration ϕ_{cl} .

How large is this leading order correction? This turns out to be a somewhat subtle question. In principle, one can explicitly compute this correction term by inserting (4.27) and (4.28) in our expression (4.20) for S_{int} . One then finds, however, that the magnitude of the corrections critically depends on how one treats the end points of the integration over the u and v coordinates. If one defines the integration region, as indicated in Fig. 4, to have infinitely sharp boundaries at the horizon $u = \infty$ and the critical line $v = v_0$, one will find that the corrections to the final state on $\mathcal{H}^+ \cup \mathcal{I}^+$ in fact become very large. (This can easily be seen from the above expressions for T_{vv} .) However, if instead one cuts off the interactions at some finite distance from the horizon, as measured in some local coordinate system, one will find that the corrections are bounded, and, because of the exponentially growing redshift, will eventually get smaller with time for the final state on \mathcal{I}^+ .

A similar cutoff dependence arises if one tries to compute higher order corrections by further expanding the exponent in (4.20). In this case, additional short distance singularities arise because of the time-ordering prescription. In principle, in a finite consistent theory of gravity, such as perhaps string theory, these singularities should

be smoothed out by the short distance gravitational dynamics. However, given our lack of understanding of this short distance theory, it seems a reasonable procedure to represent its effect by introducing some cutoff scale $\epsilon(x)$.

To understand this procedure somewhat better, let us recall the discussion of Sec. IA on the space-time complementarity principle. There we also introduced a cutoff scale $\epsilon(x)$. The point of that discussion was that a reasonable gravitational cutoff must not only regulate the integrals of $T_{\mu\nu}$ but must also act at a more fundamental level and truncate the Hilbert space by eliminating all states with wavelengths smaller than the cutoff scale. Indeed, the short distance singularity in the operator product relation between two stress-tensor operators arises due to the contribution of intermediate states with arbitrarily short wavelengths. Regulating the singularity in the operator product expansions (OPE's) of $T_{\mu\nu}$ is therefore equivalent to throwing out these singular states.

What does this cutoff procedure imply for the calculation of the gravitational corrections to the Hawking spectrum? Suppose we take for ϵ some proper distance cutoff in the region very close to the horizon. This will indeed ensure that we can find a reasonable and controlled answer for the form of the final state on \mathcal{H}^+ . However, we immediately run into trouble if at the same time we want to calculate the form of the outgoing state on \mathcal{I}^+ , because practically all states on \mathcal{I}^+ correspond to singular states near the horizon that were thrown out by our cutoff procedure. A proper distance cutoff near the horizon is therefore not suitable for computing the outgoing spectrum (see, e.g., [12]).

Instead, to calculate the state on \mathcal{I}^+ , we are forced to choose a cutoff that allows the Hilbert space to contain very high frequency outgoing modes near the horizon, since the corresponding quantum fluctuations were used in the zeroth order calculation. The price one pays, however, is that these modes interact violently with all infalling particles, via the stress-energy fluctuations computed in Sec. IVD. This will in turn result in a fast growing quantum uncertainty in the geometry near the horizon, which is most clearly exhibited by inserting the expression for T_{uu} , as obtained from (4.27) and (4.28), into Eq. (4.11). If we nevertheless insist on producing a reliable semiclassical description of the entire final state on $\mathcal{H}^+ \cup \mathcal{I}^+$, we are forced to truncate the Hilbert space of the matter near the horizon so that it contains only very low frequency incoming modes. In this Hilbert space we can reasonably trust the calculations of the outgoing state, and we find that the corrections are finite. Qualitatively the outgoing state will look very much like the thermal state as derived by Hawking. However, it is clear that due to the necessary truncation of the Hilbert space the entropy associated with the black hole is drastically reduced compared to the conventional free field result.

Summarizing, we thus propose that the calculation of the outgoing state can be performed in a controlled fashion by working in a theory with such a truncated Hilbert space, with the idea that this procedure represents a reasonable effective description of some underlying consistent theory of quantum gravity. By regulating the calculation in this way, one produces a finite result for the

gravitationally corrected final state, which, however, critically depends on the choice of cutoff.

V. SUMMARY

We have presented a method for computing the gravitational corrections to the emission spectrum of an evaporating black hole. In this procedure we included gravitational interactions between the matter fields as they propagate from initial Cauchy slice near \mathcal{I}^- to a final slice near $\mathcal{H}^+ \cup \mathcal{I}^+$. Contrary to common expectations, we find that the gravitational interactions that play a role in this calculation are not suppressed, and can lead to potentially large physical effects.

It is clear that some parts of our approximation scheme can be improved. In particular, it should in principle be possible to compute exactly the leading order correction to the final state, which results from single graviton exchange between the in- and outgoing virtual particles. This will in particular clarify some issues of our calculation related to the interactions near $v = v_0$.

In going to higher orders, conventional field theoretical methods will become inadequate, because gravity is nonrenormalizable. Thus in any approach based on local field theory one is forced to introduce some effective description by introducing a cutoff. The physical picture that emerges from these considerations is described in Secs. IA and IVE.

ACKNOWLEDGMENTS

The research of E.V. and H.V. was partly supported by the Alfred P. Sloan Foundation, and by the Royal Dutch Academy of Sciences. The research of Y.K. and H.V. was supported by NSF Grant No. 90-21984 and the Packard Foundation. H.V. also acknowledges the support of NWO.

APPENDIX: DERIVATION OF (3.8) AND (4.11)

We start from the assumption that the back reaction effects are small, and therefore we allow ourselves to work in the weak field approximation. Weak gravitational fields are represented by small perturbations around the classical metric

$$ds^2 = ds_{\text{cl}}^2 + h_{\mu\nu} dx^\mu dx^\nu. \quad (\text{A1})$$

In the first approximation $h_{\mu\nu}$ satisfies the linearized Einstein equations with source equal to the stress-tensor $T_{\mu\nu}$ of the scalar field. Hence, $h_{\mu\nu}$ may be expressed as

$$h_{\mu\nu}(x) = 8G \int d^4y D_{\mu\nu}{}^{\lambda\sigma}(x, y) T_{\lambda\sigma}(y), \quad (\text{A2})$$

where $D_{\mu\nu}{}^{\lambda\sigma}(x, y)$ denotes the propagator for the graviton field on the classical black hole background. Next one should substitute this perturbation back into the scalar

wave equation to obtain the corrected equation that to this order includes all the gravitational self-interactions.

The calculation can be considerably simplified by focusing only on the interactions between the incoming and outgoing waves near the horizon, which turn out to be the most significant. It turns out to be convenient to work in the Kruskal parametrization in terms of x^\pm . We want to find an approximate solution of the linearized Einstein equations, which is valid close to the horizon and takes into account the back reaction effect of the $T_{\pm\pm}$ components of the stress-energy tensor. Thus we assume that T_{++} and T_{--} are separately conserved, i.e., $\partial_\mp T_{\pm\pm} = 0$, and that the other components of $T_{\mu\nu}$ are negligible. We further take the following ansatz for the metric near the black hole horizon:

$$ds^2 = -\frac{32M^3}{r} e^{-r/2M} [dx^+ + h_{--}(x^-, \Omega) dx^-] \times [dx^- + h_{++}(x^+, \Omega) dx^+] + r^2 d^2\Omega. \quad (\text{A3})$$

The linearized Einstein equations for this metric are

$$\kappa(\Delta_\Omega - 1)h_{\pm\pm} = T_{\pm\pm} \quad (\text{A4})$$

with $\kappa = 2^9 GM^4/e$. This equation integrates to

$$h_{\pm\pm}(x^\pm, \Omega) = \frac{1}{\kappa} \int d\tilde{\Omega} f(\Omega, \tilde{\Omega}) T_{\pm\pm}(x^\pm, \tilde{\Omega}). \quad (\text{A5})$$

Next we consider the wave equation for the scalar field ϕ in the background metric (A3). Just as in Sec. II, one finds that near the horizon it takes a simplified form:

$$\nabla_+ \nabla_- \phi = 0, \quad \nabla_\pm = \partial_\pm - h_{\pm\pm} \partial_\mp. \quad (\text{A6})$$

Using the fact that $[\nabla_+, \nabla_-] = 0$, one sees that the classical solutions again separate into a sum of an incoming and an outgoing part: $\phi = \phi_{\text{in}} + \phi_{\text{out}}$, with $\nabla_+ \phi_{\text{out}} = 0$ and $\nabla_- \phi_{\text{in}} = 0$. Solving these first order differential equations for ϕ_{in} and ϕ_{out} , one finds that the outgoing field ϕ_{out} must be of the form

$$\phi_{\text{out}}(x^-, x^+, \Omega) = \phi_{\text{out}} \left(x^- + \int_{x^+}^{\infty} dy^+ h_{++}(y^+, \Omega), \Omega \right). \quad (\text{A7})$$

Here we have chosen our integration limits such that for $x^+ \rightarrow \infty$ there is no shift in the argument. The matching with the in field occurs near the point $x_0^+ = e^{v_0/4M}$.⁵ Inserting the expression (A5) for h_{++} , taking $x^+ \rightarrow \infty$, and translating the result back to the u, v coordinates produces the result (3.8).

Similarly, the infalling matter is also sensitive to the metric fluctuations induced by the outgoing matter. The

⁵Note that we are indeed only considering the interactions that take place in the horizon region $x^+ > e^{v_0/4M}$. For values of x^+ near the critical time the dynamics is rather more complicated.

equation of motion for infalling waves reads

$$\partial_- \phi_{\text{in}}(x^+, x^-, \Omega) = h_{--}(x^+, \Omega) \partial_+ \phi_{\text{in}}(x^+, x^-, \Omega) \quad (\text{A8})$$

and can in a similar way be formally solved via

$$\phi_{\text{in}}(x^+, x^-, \Omega) = \phi_{\text{in}} \left(x^+ - \int_{x^-}^{\infty} dy^- h_{--}(x^+, \Omega), \Omega \right). \quad (\text{A9})$$

Substituting $x^- = 0$ and reexpressing the result in u, v coordinates leads to Eq. (4.11). Finally, we want to repeat that the solutions of the linearized Einstein equation and scalar wave equation that we just described are valid only in the close neighborhood of the black hole horizon. We have ignored all other components of $T_{\mu\nu}$ except $T_{\pm\pm}$, and therefore strictly speaking our analysis applies only to classical fields ϕ for which these other components are not too large.

-
- [1] S. W. Hawking, *Commun. Math. Phys.* **43**, 199 (1975); see also N. Birrell and P. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, Cambridge, England, 1982), and references therein.
- [2] S. Giddings, in *1994 Summer School in High Energy Physics and Cosmology*, Proceedings, Trieste, Italy (World Scientific, Singapore, 1995); A. Strominger, "Les Houches Lectures on Black Holes," presented at the Les Houches Summer School of Theoretical Physics, Les Houches, France, 1994, Report No. hep-th-950171 (unpublished).
- [3] G. 't Hooft, *Nucl. Phys.* **B335**, 138 (1990); "Unitarity of the Black Hole S Matrix," Utrecht Report No. THU-93/04 (unpublished).
- [4] C. R. Stephens, G. 't Hooft, and B. F. Whiting, *Class. Quantum Grav.* **11**, 621 (1994).
- [5] L. Susskind, *Phys. Rev. Lett.* **71**, 2367 (1993); *Phys. Rev. D* **49**, 6606 (1994); L. Susskind and J. Uglum, *ibid.* **50**, 2700 (1994); L. Susskind, L. Thorlacius, and J. Uglum, *ibid.* **48**, 3743 (1993).
- [6] E. Verlinde and H. Verlinde (unpublished).
- [7] F. Englert, S. Massar, and R. Parentani, *Class. Quantum Grav.* **11**, 2919 (1994).
- [8] K. Schoutens, E. Verlinde, and H. Verlinde, *Phys. Rev. D* **48**, 2690 (1993); in *String Theory, Gauge Theory and Quantum Gravity*, Proceedings of the Spring School, Trieste, Italy, 1993, edited by R. Dijkgraaf *et al.* (World Scientific, Singapore, 1994); S. R. Das and S. Mukherji, *Phys. Rev. D* **50**, 930 (1994); E. Keski-Vakkuri, G. Lifschytz, S. Mathur, and M. Ortiz, *ibid.* **51**, 1764 (1995).
- [9] T. Dray and G. 't Hooft, *Nucl. Phys.* **B253**, 173 (1985).
- [10] G. 't Hooft, *Phys. Lett. B* **198**, 61 (1987); *Nucl. Phys.* **B304**, 867 (1988); D. Amati, M. Ciafaloni, and G. Veneziano, *Int. J. Mod. Phys.* **3**, 1615 (1988); H. Verlinde and E. Verlinde, *Nucl. Phys.* **B371**, 246 (1992); *Class. Quantum Grav.* **10**, Suppl. 175 (1993).
- [11] G. 't Hooft, *Nucl. Phys.* **B256**, 727 (1985).
- [12] T. Jacobson, *Phys. Rev. D* **48**, 728 (1993).