

Superscattering matrix for two-dimensional black holes

S. W. Hawking

*Department of Applied Mathematics and Theoretical Physics, University of Cambridge,
Silver Street, Cambridge CB3 9EW, United Kingdom
and California Institute of Technology, Pasadena, California 91125*

(Received 28 February 1994)

A consistent Euclidean semiclassical calculation is given for the superscattering operator $\mathcal{S}$ in the Russo, Susskind, and Thorlacius (RST) model for states with a constant flux of energy. The $\mathcal{S}$ operator is *CPT* invariant. There is no loss of quantum coherence when the energy flux is less than a critical rate and complete loss when the energy flux is critical.

PACS number(s): 04.70.Dy

I. INTRODUCTION

In classical general relativity a lot of information is lost in the formation of a black hole. A black hole will settle down rapidly to a stationary state. The classical no hair theorems say that this state is characterized just by the values of conserved charges, like mass, angular momentum, and electric charge, that are coupled to gauge fields. In other words a black hole of a given mass, angular momentum, and electric charge, can be formed by the collapse of a very large number of different objects.

This loss of information was not a worry in the purely classical theory, because one could say the information was still inside the black hole, even if one could not get at it. But the situation changed when it was discovered that, according to quantum theory, black holes should radiate and slowly evaporate. This made the question about the information in a black hole much more pressing. If the black hole evaporated and disappeared completely, what would happen to the information? There were three possible ways the information might be preserved.

(1) The information might come out again at the end of the evaporation. The problem was that information requires energy to carry it and there would not be much energy left in the final stages.

(2) The information might come out continuously during the evaporation. The difficulty here was that the information would be carried by the in-falling matter, far beyond the apparent horizon. So, if it were to appear outside the horizon, it would violate causality. If one could send information faster than light like this, one could also send it back in time.

(3) The black hole might not evaporate completely, but might leave some long-lived remnant that could be said still to contain the information. But this would violate *CPT*, if black holes could form, but never disappear completely.

In this paper I want to take seriously the idea that information is lost. As I showed some time ago [1], this would imply a loss of quantum coherence. One could form a black hole from matter in a pure quantum state, but it would decay into radiation in a mixed quantum

state described by a density matrix. In ordinary quantum field theory the evolution is described by an S matrix. This can be thought of as a two index tensor on the Hilbert space that maps initial states to final states:

$$\psi_f^A = S^A_B \psi_i^B. \quad (1)$$

However, when one goes to quantum gravity, the possibility of forming real or virtual black holes means that the evolution is given by what I call a superscattering operator $\mathcal{S}$. This can be thought of as a four index tensor on Hilbert space that maps initial density matrices to final ones:

$$\rho_f^A_B = \mathcal{S}^A_{BC} \rho_i^C_D. \quad (2)$$

In general, the $\mathcal{S}$ operator will not be the product of an S matrix with its complex conjugate:

$$S^A_{BC} \neq S^A_C \bar{S}_B^D. \quad (3)$$

This proposal of a nonunitary evolution was greeted with outrage by most particle physicists. I was accused of violating quantum mechanics. That is not the case. One gets loss of quantum coherence and a mixed state whenever there is part of a system you do not measure. All I have done is point out you cannot measure the part of the quantum state that is inside the black hole.

The argument about whether quantum coherence is lost in black hole evaporation has dragged on inconclusively over the years because general relativity is not renormalizable. This has meant we have not been able to calculate what happens in the final stages of black hole evaporation. Maybe the answer lies in supergravity or superstring theory, but we do not know how to use them to do the calculation. However, in the last two years, there has been a revival of interest in two-dimensional black holes. These have the great advantage of being renormalizable, so it should be possible to use them to decide the issue. Those who felt a mission to defend quantum purity hoped that these two-dimensional models would show that information and quantum coherence were preserved. However, they have been disappointed with the results so far: all calculations to date have shown loss of information and quantum coherence. This

has left the unitary S matrix camp suggesting the breakdown of the large N approximation, but with no real argument.

Most of the calculations that have been done on two-dimensional black holes have been carried out in Lorentzian spacetime. They have assumed no horizons or singularities in the past. One can show [2] that this implies that there must be horizons or naked singularities in the future. Thus the calculations are manifestly not CPT invariant. This is reflected in the fact that the outgoing energy flux they predict is always below a certain critical level. However, the ingoing energy flux can have any value. Thus the superscattering operator given by these calculations will not be CPT invariant. To lose quantum coherence is bad enough, but to lose CPT as well sounds like carelessness.

I shall therefore outline a different approach to calculating the superscattering operator which is guaranteed to give CPT invariant results. It is the Euclidean path integral method, coupled with the no boundary condition [3]. In the cosmological case, the no boundary condition literally meant that spacetime had no boundary. In other words, the quantum state is defined by a path integral over all compact positive definite metrics. But in the particle scattering case, the appropriate quantum state is defined by a path integral over all positive definite metrics that have one or more asymptotically Euclidean regions but no other boundary. One can show that the asymptotic Green functions in this quantum state are CPT invariant [4]. It then follows that the superscattering operator is CPT invariant, in the sense of detailed balance [5]: the probability to go from the initial pure state A to a final state B is the same as the probability to go from the CPT conjugate of B to the CPT conjugate of A .

II. TWO-DIMENSIONAL BLACK HOLES

The starting point for two-dimensional black holes is the Callan-Giddings-Harvey-Strominger (CGHS) model [6]. This has a metric g and dilaton field ϕ coupled to N minimal scalar fields f_i :

$$L = \frac{1}{2\pi} \sqrt{-g} \left[e^{-2\phi} [R + 4(\nabla\phi)^2 + 4\lambda^2] - \frac{1}{2} \sum_{i=1}^N (\nabla f_i)^2 \right]. \quad (4)$$

The quantum effective action of the minimal scalars in the curved metric g can be evaluated exactly and added to the action of the classical CGHS model. One can then define new field variables χ and Ω in which the theory is a conformally invariant quantum field theory and the action looks rather like that of the Liouville model:

$$S = \frac{1}{\pi} \int d^2x \left\{ -\partial_+ \chi \partial_- \chi + \partial_+ \Omega \partial_- \Omega + \lambda^2 \exp \left[\frac{2}{\sqrt{\kappa}} (\chi - \Omega) \right] + \frac{1}{2} \sum_{i=1}^N \partial_+ f_i \partial_- f_i \right\}. \quad (5)$$

In making the theory a conformal field theory, one has to modify the action. One can change either the kinetic term, as Russo, Susskind, and Thorlacius (RST) do [7], or the nonderivative term, as Bilal and Callan [8] and de Alwis (BCdA) [9] do. Both modifications end up with the same Liouville-like theory, but the relation between the Liouville fields χ and Ω and the physical fields g and ϕ is different. The RST version has the great advantage that it admits the linear dilaton as a solution. The other, BCdA, version does not seem to have any natural ground state. I shall therefore work with the RST version.

In the RST model the Liouville fields are related to the physical fields by

$$\Omega = \frac{\sqrt{\kappa}}{2} \phi + \frac{e^{-2\phi}}{\sqrt{\kappa}}, \quad (6)$$

$$\chi = \sqrt{\kappa} \rho - \frac{\sqrt{\kappa}}{2} \phi + \frac{e^{-2\phi}}{\sqrt{\kappa}}, \quad (7)$$

where the metric is in the conformal gauge

$$ds^2 = -e^{2\rho} dx^+ dx^-. \quad (8)$$

The Liouville model looks to be almost linear, but this is deceptive. Both in the BCdA and RST versions, the field Ω is bounded below as a function of the physical field ϕ . It takes its minimum value Ω_c when $\phi = \phi_c$. There are two possible attitudes one can take to this. One, propounded by de Alwis [9], is that one should take the quantum theory to be defined by a path integral over the full range of χ and Ω from minus infinity to plus infinity. The other attitude, put forward by RST [7], is that the path integral should be restricted to the range of Ω that corresponds to real physical fields. I shall adopt this latter approach. The physical fields are what are important. The Liouville fields χ and Ω can be regarded just as mathematical constructions that help to solve the field equations, but are of no physical significance of themselves.

The field equations for the Liouville model are simple:

$$\partial_+ \partial_- (\chi - \Omega) = 0, \quad (9)$$

$$\partial_+ \partial_- (\chi + \Omega) = -\frac{2\lambda^2}{\sqrt{\kappa}} \exp \left[\frac{2}{\sqrt{\kappa}} (\chi - \Omega) \right]. \quad (10)$$

The first equation corresponds to the freedom to choose coordinates. One can therefore take

$$\chi - \Omega = 0 \quad (11)$$

at least locally. It is then easy to write down the general solution in this coordinate gauge:

$$\Omega = -\frac{\lambda^2}{\sqrt{\kappa}} x_+ x_- + F(x_+) + G(x_-). \quad (12)$$

To make this solution Euclidean, introduce a complex coordinate z in the plane with

$$z = x_+, \quad \bar{z} = -x_-. \quad (13)$$

This very simple theory becomes highly nontrivial if one imposes the requirement that the physical fields be nonsingular. One way to get a nonsingular solution is if

Ω is always above its minimum value Ω_c . The only solutions such as this that are also static and with only an asymptotically flat region can be written down in the $\chi = \Omega$ gauge:

$$\Omega = \frac{\lambda^2}{\sqrt{\kappa}} z\bar{z} + M + \frac{\sqrt{\kappa}}{2}. \quad (14)$$

Ω is a function only of r in polar coordinates, with a minimum at the origin. These solutions represent Euclidean black holes with the horizon at the origin. The polar angle θ corresponds to imaginary time. Because this is periodic, the black hole is in thermal equilibrium. There is a steady flow of energy in at infinity and a similar flow out.

Suppose now that one has a field configuration in which Ω actually reaches the critical value on some curve C . Then Ω will drop below the critical value in a neighborhood of C and the physical fields will be singular unless the gradient of Ω vanishes on C . Thus the necessary condition for the physical fields to be nonsingular is

$$\nabla\Omega|_C = 0. \quad (15)$$

I shall refer to this as the RST boundary condition. With this boundary condition, no signal can propagate through the curve C . One can therefore cut off the spacetime beyond C and consider only the region between C and asymptotically flat infinity. One can regard the curve C as being like the axis of symmetry in the dimensional reduction of a spherically symmetric spacetime to two dimensions. On the basis of this analogy, the appropriate boundary condition on the minimal scalars would be that their normal gradient should vanish on C .

If C is timelike, the normal direction will be spacelike. The field equation

$$\partial_+ \partial_- \Omega = -\frac{\lambda^2}{\sqrt{\kappa}} \exp \left[\frac{2}{\sqrt{\kappa}} (\chi - \Omega) \right] \quad (16)$$

will then imply that the second derivative of Ω in the normal direction will be positive. This means that $\Omega \geq \Omega_c$ in a neighborhood of C and the physical fields can be nonsingular. On the other hand, if C is spacelike, the normal direction will be timelike and the second derivative of Ω in the normal direction will be negative. Thus Ω will go below Ω_c and the physical fields will be singular. Thus the RST boundary condition cannot be imposed in Lorentzian spacetime if the boundary is spacelike.

By contrast, in the Euclidean regime the normal direction to C is always spacelike. This means one can always satisfy the boundary condition. Thus the Liouville model with the RST condition and the no boundary condition is a well defined quantum theory. One should therefore be able to calculate the superscattering operator and see whether it involves loss of quantum coherence. I shall give evidence that it does, in at least some situations.

III. THE QUANTUM RST MODEL

If one just had the Liouville model without any restriction on the range of Ω , the theory would be linear and the semiclassical approximation would be exact. This means

that the quantum theory would be determined completely by solutions of the quantum effective action, which will be the same as the original Liouville action. However, if the range of Ω is restricted, and the RST boundary condition imposed, the theory becomes effectively nonlinear. This means that the solutions of the Liouville theory are no longer the whole story. But one can still hope that they will give a first approximation.

As well as the field equations obtained by varying ϕ and ρ , there are constraint equations obtained by varying the components of the metric that are zero in the conformal gauge. In terms of the physical fields, the constraint equations can be written in a form like the Einstein equations:

$$\begin{aligned} & - \left[e^{-2\phi} + \frac{\kappa}{4} \right] (4\partial_+ \rho \partial_+ \phi - 2\partial_+^2 \phi) \\ & = \frac{1}{2} \sum_{i=1}^N \partial_+ f_i \partial_+ f_i - \kappa (\partial_+ \rho \partial_+ \rho - \partial_+^2 + t_+) . \end{aligned} \quad (17)$$

The term on the left that is the analog of the Einstein tensor is the trace-free part of the second covariant derivative of ϕ . The first term on the right is the classical energy-momentum tensor of the scalar fields. The remaining terms on the right can be regarded as the quantum induced energy-momentum tensor. It is nonlocal because it contains the functions $t_{\pm}(x_{\pm})$ that have to be determined by boundary conditions. Given a solution of the field equations, one can always find functions $t_{\pm}(x_{\pm})$ that will satisfy the constraint equations as the *definition* of the total energy-momentum tensor, classical plus quantum, of the scalar fields. In other words, given a solution of the ϕ and ρ field equations, one can read off from the constraint equations what the energy-momentum flow is.

In the $\chi = \Omega$ gauge there are a two-parameter family of solutions which have a Killing vector in the direction of the polar angle θ . The two parameters are the coefficient of the $\ln r$ term and the constant term in Ω . Imposing the RST boundary condition that $\nabla\Omega = 0$ where $\Omega = \Omega_c$ puts one relation between the two parameters. One therefore has a one-parameter family of static Euclidean solutions:

$$\begin{aligned} \Omega &= \frac{\lambda^2}{\sqrt{\kappa}} z\bar{z} - P \ln(\lambda^2 z\bar{z}) - P [\ln(P\sqrt{\kappa}) + 1] \\ & - \frac{\sqrt{\kappa}}{4} \left[\ln \left[\frac{\kappa}{4} \right] + 1 \right]. \end{aligned} \quad (18)$$

The linear dilaton is the member of this family with $P = \sqrt{\kappa}/4$. In it, Ω has a minimum of Ω_c on a circle around the origin. The total energy-momentum tensor of the scalar fields calculated in the manner I described is zero. Thus this solution represents the vacuum with no energy flux at infinity. However, for $0 \leq P < \sqrt{\kappa}/4$, there is a family of solutions with a constant positive flux of energy coming in at infinity and a similar flux going out. As the energy flux increases from zero, the circle at which $\Omega = \Omega_c$ shrinks in radius. When the flux reaches the critical level, the circle shrinks to zero and the solu-

tion becomes a Euclidean black hole with no boundary and with $\Omega > \Omega_c$ everywhere. There is a one parameter family of black hole solutions, but all have the same energy flux at infinity. There are no regular static Euclidean solutions with an energy flux greater than this value. That is what one might expect because, if one sent in energy at such a rate, one would create a black hole that would grow faster than it could radiate itself away. So one could not get a static solution.

IV. THE SUPERSCATTERING OPERATOR

One would expect these Euclidean solutions to give the dominant contributions to the superscattering operator for incoming states with a steady flux of energy. Consider first the solutions, such as the linear dilaton, that have a boundary at Ω_c . In these, the region inside the boundary circle is not regarded as part of the spacetime. This means there is no reason to identify the polar angle θ with any particular period. It is therefore natural to take θ to run the full range from minus infinity to infinity.

It is straightforward to calculate Green functions for the scalar fields in these spacetimes. The minimal scalars propagate in the flat background metric independent of the dilaton and the conformal factor. They are, however, affected by the global structure, such as whether θ is periodically identified, and by the presence of the boundary. On the basis that the boundary is like an axis of symmetry, it is natural to impose Neumann boundary conditions on the minimal scalars: the normal gradient of each field vanishes on the boundary. It is then straightforward to obtain the Green functions by the method of images.

The procedure to calculate the superscattering operator from the asymptotic Green functions was described in [5]. One first defines creation and annihilation operators for incoming and outgoing particles in terms of integrals of the field operators at infinity, e.g.,

$$a_i(\omega) = -\frac{i}{\sqrt{2\pi}} \int_{\Sigma} p_i(\omega, x) \vec{\nabla}_{\mu} f(x) d\Sigma^{\mu}, \quad (19)$$

where $a_i(\omega)$ is the annihilation operator for a particle in the incoming mode $p_i(\omega, x)$ of frequency ω , Σ is a time-like line near infinity, and the integral is carried out over real Lorentzian time. The superscattering operator is then the expectation value of a string of creation and annihilation operators for the initial and final states and their complex conjugates:

$$\mathcal{S}^{A_{BC}D} = \langle O_C O^A O_B O^D \rangle, \quad (20)$$

where O^D is a string of creation operators for the initial state $|\psi^D\rangle = O^D|0\rangle$ and O_C is a string of annihilation operators for the initial state $\langle\psi_C| = \langle 0|O_C$. Similarly O^A and O_B are the creation and annihilation operators for the final state.

One interprets the expectation values $\langle f(x_1), f(x_2), \dots, f(x_n) \rangle$ as the analytic continuation of the Euclidean Green function $G(x_1, x_2, \dots, x_n)$ keeping each of the differences $x_{i+1} - x_i$ as a vector with a small future-directed imaginary part. In other words, to

evaluate the superscattering matrix element $\mathcal{S}^{A_{BC}D}$ one does a multiple integral of the analytically continued Green function on the timelike line Σ near infinity with the contours of integration arranged in increasing order of imaginary time reading from left to right in the expectation value.

In ordinary nongravitational quantum field theory, the superscattering operator factorizes

$$\mathcal{S}^{A_{BC}D} = \langle O_C O^A \rangle \langle O_B O^D \rangle. \quad (21)$$

The second factor on the right is defined to be the S matrix and the first factor is its complex conjugate. There is no loss of quantum coherence and the S matrix is unitary. But in quantum gravity the superscattering matrix may not factorize, leading to loss of quantum coherence, as I shall show.

The static solutions with a boundary are conformal to half of flat space. This means that the Green functions for the scalar fields are the same as those for flat space with a reflecting wall. They will propagate a positive frequency wave function purely in the negative imaginary time direction, and a negative frequency wave function purely in the positive imaginary time direction. This means that there will be propagation only between the top two contours corresponding to O^D and O_B and between the bottom two contours corresponding to O^A and O_C . Thus the superscattering operator will factorize and there will be no loss of quantum coherence. One can adjust the solution so that the ingoing flux of the initial state is equal to the energy flux of the solution, calculated in the way I described. This gives a consistent semiclassical calculation of the superscattering operator for initial states with constant energy fluxes less than the critical flux.

The situation is very different, however, if one considers Green functions on a black hole background. In this case, the point at the origin of Euclidean space is part of the physical spacetime. In order for the metric to be regular there, the polar angle θ must be identified with a period of 2π . This means that the asymptotic Green functions will be periodic in imaginary time, with a period $\beta = 2\pi/\lambda$. The propagation of positive frequencies will still be in the negative imaginary time direction. But now there will be images of the contours of integration, periodically spaced in the imaginary time direction. This means that the propagation need not be just from the O^D contour to the O_B one and from the O^A contour to the O_C one. Instead, there can be a nonzero amplitude for the initial creation operators O^D to propagate to the initial annihilation operators O_C and for the final creation and annihilation operators to do the same. This means that the superscattering operator will not factorize, so there will be loss of quantum coherence.

The energy-momentum tensor for the scalar fields is

$$T_{\pm\pm} = \sum \partial_{\pm} f_i \partial_{\pm} f_i - \frac{N}{12} [\partial_{\pm} \rho \partial_{\pm} \rho - \partial_{\pm\pm}^2 + t_{\pm}(x_{\pm})]. \quad (22)$$

The functions $t_{\pm}(x_{\pm})$ should be chosen to make the second term on the right zero asymptotically. This will

mean that the energy flux at infinity is just given by the first term on the right, which is the classical energy-momentum tensor of the f_i fields, which are given by propagating the data for the initial and final states with the analytically continued Euclidean propagator. The total energy-momentum tensor will be regular on the Euclidean space except at the origin, where it will in general have a singularity. This singularity will, however, vanish if the classical energy-momentum tensors of the initial and final states both have exactly the critical energy flux. Thus in this case, and only in this case, one will have a consistent semiclassical calculation of the superscattering operator, but this time with loss of quantum coherence. Suppose that the initial state was a pure quantum state in one particular scalar field. Then the final state could be any state, in any combination of the scalar fields, that has the critical energy flux. But this is just the condition for a thermal spectrum. So the superscattering operator takes any initial pure state with the critical energy flux to a final mixed thermal quantum state. Thus there is loss of quantum coherence.

These semiclassical calculations give no loss of quantum coherence when the energy flux is below the critical value, and complete loss when it is critical. But one would expect that a full calculation would give a gradual transition. As the energy flux tends towards the critical value, the boundary at Ω_c in the semiclassical solutions becomes a curve of greater and greater acceleration. It would require only a small quantum fluctuation in Ω to make the transition to a black hole metric with no boundary. Thus one would expect a gradually increasing amount of quantum incoherence, as the energy flux approached the critical value.

In this paper I have been considering the superscatter-

ing operator for steady fluxes of energy. But what one really wants to consider are bursts with finite total energy, but with a maximum flux above the critical level. It seems that a Lorentzian semiclassical solution for this situation will have a boundary at Ω_c that is spacelike. Thus the RST boundary condition cannot be satisfied. This means that the theory is not well defined: in the absence of a boundary condition on part of the boundary, one cannot calculate Green functions and predict the results of scattering. This lack of predictability is usually minimized by assuming the metric is nonsingular in the past. But then there is necessarily a singularity in the future, at least one point of which must be naked. This is rather *ad hoc*, and unsatisfactory. It is also clearly not *CPT* invariant. To restore *CPT* symmetry, one should presumably consider metrics with singularities in both the past and the future [10]. But then the lack of a boundary condition at the past singularity, which would be naked, would mean that one was completely unable to predict the final state. The only way out would seem to be to impose boundary conditions in the Euclidean regime. There will still be singularities in the Lorentzian regime, but now the behavior of the Green functions at the singularities will be determined by analytic continuation from the Euclidean regime. Hopefully, this will give physically reasonable *CPT* invariant results for the scattering of bursts, probably with loss of quantum coherence. Further work on this is in progress.

ACKNOWLEDGMENTS

I am very grateful for discussions with James Grant, Justin Hayward, Andy Strominger, and Edward Teo. This work was partly carried out at Caltech.

-
- [1] S. W. Hawking, *Phys. Rev. D* **14**, 2460 (1976).
 - [2] S. W. Hawking and J. Hayward, *Phys. Rev. D* **49**, 5252 (1994).
 - [3] S. W. Hawking and J. B. Hartle, *Phys. Rev. D* **28**, 2960 (1983).
 - [4] S. W. Hawking, *Commun. Math. Phys.* **87**, 395 (1982).
 - [5] R. M. Wald, *Black Holes, Thermodynamics and Time Reversibility*, Quantum Gravity Vol. 2 (Clarendon, Oxford,

- 1981), p. 223.
- [6] C. G. Callan, S. B. Giddings, J. A. Harvey, and A. Strominger, *Phys. Rev. D* **45**, 1005 (1992).
- [7] J. G. Russo, L. Susskind, and L. Thorlacius, *Phys. Rev. D* **46**, 3444 (1992).
- [8] A. Bilal and C. G. Callan, *Nucl. Phys.* **B394**, 73 (1993).
- [9] S. P. de Alwis, *Phys. Rev. D* **46**, 5429 (1992).
- [10] A. Strominger, *Phys. Rev. D* **48**, 5769 (1993).