INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE CENTRATOM TRIESTE

SMR.762 - 44

SUMMER SCHOOL IN HIGH ENERGY PHYSICS AND COSMOLOGY

13 June - 29 July 1994

QUANTUM MECHANICS OF BLACK HOLES

STEVEN GIDDINGS
Dept. of Physics
University of California
Santa Barbara, CA
USA

Please note: These are preliminary notes intended for internal distribution only.

UCSBTH-94-xx
hep-th/yymmnnn
Draft -- do not distribute

Quantum Mechanics of Black Holes

Steven B. Giddings[†]

Department of Physics
University of California
Santa Barbara, CA 93106-9530

Abstract

Lectures at Trieste Summer School

[†] Email addresses: giddings@denali.physics.ucsb.edu, ~~steve@voodoo.bitnet.~~

1. Introduction

Hawking's 1974 discovery[1] that black holes evaporate ushered in a new era in black hole physics. In particular, this was the beginning of concrete applications of quantum mechanics in the context of black holes. But more importantly, the discovery of Hawking evaporation has raised a sharp problem whose resolution probably requires a better understanding of Planck scale physics, and which may therefore may serve as a guide (or at least a constraint) in our attempts to understand such physics. This problem is the information problem.

In brief, the information problem arises when one considers the Gedanken experiment of black hole formation through collapse of a carefully arranged *pure* quantum state $|\psi\rangle$, or in terms of quantum-mechanical density matrices, $\rho = |\psi\rangle\langle\psi|$. This black hole then evaporates, and according to Hawking's calculation the resulting outgoing state is approximately thermal, and is described as a *mixed* quantum state. The latter statement means that the density matrix is of the form $\sum_\alpha p_\alpha |\psi_\alpha\rangle\langle\psi_\alpha|$, for some normalized basis states $|\psi_\alpha\rangle$ and some real numbers p_α of which two or more are non-zero. Comparing pure and mixed states, we find that there is missing phase information in the latter. A measure of the missing information is the entropy, $S = -\mathrm{Tr}\rho\ln\rho = -\sum_\alpha p_\alpha \ln p_\alpha$. If Hawking's calculation can be trusted, this means that in the quantum theory of black holes pure states can evolve to mixed. This conflicts with the ordinary laws of quantum mechanics, which always preserve purity.

Hawking subsequently proposed [2] that quantum mechanics be modified to allow purity loss. However, as we'll see, inventing an alternative dynamics is problematical. This has lead people to consider other alternatives, namely that information either escapes a black hole or that it is left behind in a black hole remnant. Both of these possibilities also encounter difficulties, and as a result we have the black hole information problem.

In these lectures we'll develop these statements more fully, starting with a study of Hawking radiation. In the past few years an improved understanding of black hole evaporation has been obtained through study of two-dimensional models, and because of this and due to their greater simplicity we'll start by considering the Hawking effect in such a two-dimensional context.

2. Two-dimensional dilaton gravity

In 2d, formulating gravity with just a metric gives trivial dynamics; for example, the Einstein action is a topological invariant. Instead we consider theories with the addition of a scalar dilaton ϕ. A particular simple theory[3-5] has action

$$S = \frac{1}{2\pi}\int d^2x\sqrt{-g}\left[e^{-2\phi}(R+4(\nabla\phi)^2+4\lambda^2)-\tfrac{1}{2}(\nabla f)^2\right] , \tag{2.1}$$

where λ^2 is an analogue to the cosmological constant and f is a minimally coupled massless matter field that provides a source for gravity. Note that e^ϕ plays the role of the gravitational coupling, as its inverse square appears in front of the gravitational part of the action.

In two dimensions the general metric $ds^2 = g_{ab}dx^a dx^b$ can always locally be put into conformal gauge,

$$ds^2 = -e^{2\rho}dx^+dx^- , \tag{2.2}$$

with the convention $x^\pm = x^0 \pm x^1$. The equations resulting from the action (2.1) are most easily analyzed in this gauge. The matter equations are

$$\partial_+\partial_- f = 0 , \tag{2.3}$$

with general solution $f_i = f_+(x^+) + f_-(x^-)$. Next, the relation

$$\sqrt{-g}R = -2\Box\rho \tag{2.4}$$

allows rewriting of the gravitational part of the action,

$$S = \frac{1}{2\pi}\int d^2x\left\{2\nabla(\rho-\phi)\nabla e^{-2\phi}+4\lambda^2 e^{2(\rho-\phi)}\right\} . \tag{2.5}$$

The equation of motion for $\rho-\phi$ is therefore that of a free field, with solution

$$\rho-\phi = \frac{1}{2}\left(w_+(x^+)+w_-(x^-)\right) . \tag{2.6}$$

The equation for ϕ then easily gives

$$e^{-2\phi} = u_+ + u_- - \lambda^2\int^{x^+} e^{w_+}\int^{x^-} e^{w_-} \tag{2.7}$$

where $u_\pm(x^\pm)$, are also free fields. Finally, varying the action with respect to g^{++}, g^{--} gives the constraint equations,

$$\begin{aligned}\delta g^{++} : G_{++} &\equiv -e^{-2\phi}\left(4\partial_+\rho\partial_+\phi - 2\partial_+^2\phi\right) = \tfrac{1}{2}\partial_+ f\partial_+ f\\ \delta g^{--} : G_{--} &\equiv -e^{-2\phi}\left(4\partial_-\rho\partial_-\phi - 2\partial_-^2\phi\right) = \tfrac{1}{2}\partial_- f\partial_- f .\end{aligned} \tag{2.8}$$

These determine $u_\pm$ in terms of $f_\pm$ and $w_\pm$:

$$u_\pm = \frac{M}{2\lambda} - \tfrac{1}{2}\int e^{w_\pm}\int e^{-w_\pm}\partial_\pm f\partial_\pm f , \tag{2.9}$$

where M is an integration constant. In the following we will choose units so that $\lambda = 1$.

The theory is therefore completely soluble at the classical level. The unspecified functions $w_\pm$ result from the unfixed remaining gauge freedom; conformal gauge (2.2) is unchanged by a reparametrization $x^\pm = x^\pm(\sigma^\pm)$. This freedom may be used to set $w_+ + w_- = \sigma^+ - \sigma^-$, for example. In this gauge the general vacuum solution is

$$\begin{aligned}ds^2 &= -\frac{d\sigma^+d\sigma^-}{1+Me^{\sigma^- - \sigma^+}}\\ \phi &= -\tfrac{1}{2}\ln\left(M + e^{\sigma^+-\sigma^-}\right) .\end{aligned} \tag{2.10}$$

The case $M = 0$ corresponds to the ground state,

$$\begin{aligned}ds^2 &= -d\sigma^+d\sigma^-\\ \phi &= -\sigma .\end{aligned} \tag{2.11}$$

This is the present analogue of flat Minkowski space, and is called the linear dilaton vacuum. The solutions for $M > 0$ are asymptotically flat as $\sigma^+ - \sigma^- \to \infty$. At $\sigma^+ - \sigma^- \to -\infty$ they are apparently singular, but regularity is restored by the coordinate transformation

$$x^+ = e^{\sigma^+} , x^- = -e^{-\sigma^-} . \tag{2.12}$$

A true singularity appears at $x^+x^- = M$, and $x^\pm = 0$ is the horizon. The corresponding Penrose diagram is shown in Fig. 1; the solution is a black hole and M is its mass. Notice the important relation $e^{2\phi}|_{\rm horizon} = \frac{1}{M}$. For $M < 0$ the solution is a naked singularity.

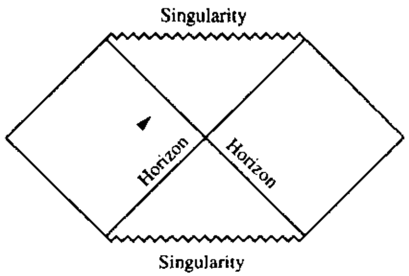

Fig. 1: Shown is the Penrose diagram for a vacuum two-dimensional dilatonic black hole.

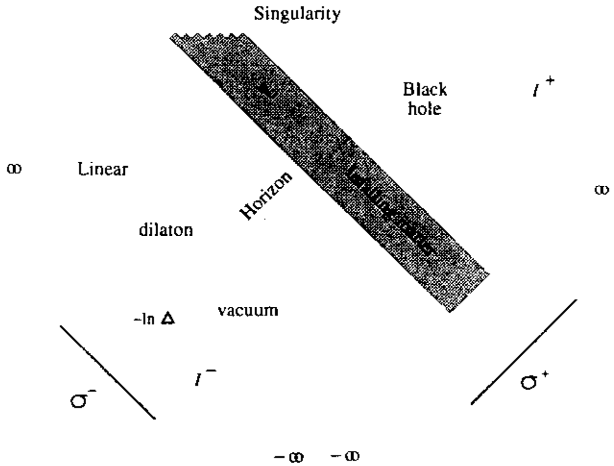

Fig. 2: The Penrose diagram for a collapsing black hole formed from a left-moving matter distribution.

Next consider sending infalling matter, $f = F(x^+)$, into the linear dilaton vacuum. This will form a black hole, as shown in Fig. 2. Before the matter infall the solution is given by (2.11). Afterwards it is found by using (2.6)-(2.9),

$$\begin{aligned} e^{-2\phi} &= M + e^{\sigma^+}\left(e^{-\sigma^-} - \Delta\right) \\ ds^2 &= -\frac{d\sigma^+ d\sigma^-}{1 + M\, e^{\sigma^- - \sigma^+} - \Delta\, e^{\sigma^-}} \end{aligned} \tag{2.13}$$

where one can easily show

$$\begin{aligned} M &= \int d\sigma^+ T_{++} \\ \Delta &= \int d\sigma^+ e^{-\sigma^+} T_{++} \end{aligned} \tag{2.14}$$

and

$$T_{++} = \tfrac{1}{2}(\partial_+ F)^2 \tag{2.15}$$

is the stress tensor. The coordinate transformation

$$\xi^- = -\ell n\left(e^{-\sigma^-} - \Delta\right)\ ,\ \xi^+ = \sigma^+ \tag{2.16}$$

returns the metric to the asymptotically flat form

$$ds^2 = -\frac{d\xi^+ d\xi^-}{1 + M\ e^{\xi^- - \xi^+}}\ . \tag{2.17}$$

3. Hawking radiation in two dimensions

Now that we have a collapsing black hole, we can study its Hawking radiation.[1] The quickest derivation arises by computing the expectation value of the matter stress tensor. Consider the stress tensor for right-movers; before the hole forms they are in their vacuum, and

$$\begin{aligned} \langle T_{--}\rangle &= \lim_{\hat\sigma^- \to \sigma^-} \langle 0|\frac{1}{2}\partial_- f(\hat\sigma^-)\partial_- f(\sigma^-)|0\rangle \\ &= \lim_{\hat\sigma^- \to \sigma^-} \frac{-1}{4(\hat\sigma^- - \sigma^-)^2}\ , \end{aligned} \tag{3.1}$$

where the second line uses the 2d Green function,

$$\langle 0|f(\hat\sigma)f(\sigma)|0\rangle = -\frac{1}{2}\left[\ln(\hat\sigma^+ - \sigma^+) + \ln(\hat\sigma^- - \sigma^-)\right]\ . \tag{3.2}$$

As usual, one removes the infinite vacuum energy by normal-ordering:

$$:T_{--}:_\sigma = T_{--} + \frac{1}{4}\frac{1}{(\hat\sigma^- - \sigma^-)^2}\ . \tag{3.3}$$

The formula (3.1) also holds at $\mathcal{I}^+$, but the flat coordinates are now $\xi^\pm$. Therefore to compare the stress tensor to that of the outgoing vacuum on $\mathcal{I}^+$, we should subtract the vacuum energy computed in the ξ coordinates,

$$:T_{--}:_\xi = T_{--} + \frac{1}{4(\hat\xi^- - \xi^-)^2}\ . \tag{3.4}$$

[1] For other references see [1,6,7].

The corresponding expectation value is

$$
\begin{aligned}
\langle : T_{--} :_\xi \rangle &= \lim_{\hat{\xi}^- \to \xi^-} \langle 0 | \frac{1}{2} \partial_-^{\hat{\xi}} f\left(\sigma^-(\hat{\xi}^-)\right) \partial_-^{\xi} f\left(\sigma^-(\xi^-)\right) |0\rangle + \frac{1}{4(\hat{\xi}^- - \xi^-)^2} \\
&= -\frac{1}{4} \lim_{\hat{\xi}^- \to \xi^-} \partial_-^{\hat{\xi}} \partial_-^{\xi} \ln\left(\sigma^-(\hat{\xi}^-) - \sigma^-(\xi^-)\right) + \frac{1}{4(\hat{\xi}^- - \xi^-)^2} .
\end{aligned} \tag{3.5}
$$

Next one expands $\sigma^-(\hat{\xi}^-)$ about ξ^-, and in a few lines finds

$$
\langle : T_{--} :_\xi \rangle = -\frac{1}{24} \left[\frac{\sigma^{-'''}}{\sigma^{-'}} - \frac{3}{2} \left(\frac{\sigma^{-''}}{\sigma^{-'}} \right)^2 \right] \tag{3.6}
$$

where prime denotes derivative with respect to ξ^-.

Using the relation (2.16) between the two coordinate systems gives the outgoing stress tensor from the black hole,

$$
\langle : T_{--} :_\xi \rangle = \frac{1}{48} \left[1 - \frac{1}{(1 + \Delta e^{\xi^-})^2} \right] . \tag{3.7}
$$

This exhibits transitory behavior on the scale $\xi^- \sim -\ln \Delta$, but as $\xi^- \to \infty$ it asymptotes to a constant value $1/48$. As will be seen shortly, this corresponds to the thermal Hawking flux at a temperature $T = 1/2\pi$.

A more detailed understanding of the Hawking radiation arises from quantizing the scalar field. Recall the basic steps of canonical quantization:

1. Find a complete orthonormal basis of solutions to the field equations.
2. Separate these solutions according to positive or negative frequency.
3. Expand the general field in terms of the basis functions with annihilation operators as coefficients of positive frequency and creation operators for negative frequncy.
4. Use the canonical commutation relations to determine the commutators of the field operators.
5. Define the vacuum as the state annihilated by the annihilation operators, and build the other states on it by acting with creation operators.

In curved spacetime general coordinate invariance implies that step two is ambiguous: what is positive frequency in one frame is not in another. Consequently the vacuum state is not uniquely defined. These two observations are at the heart of the description of particle creation in curved spacetime. This ambiguity was implicit in the different normal-ordering perscriptions in the above derivation.

Following these steps, begin by introducing a basis in the "in" region near $\mathcal{I}^-$; a convenient basis of right-moving modes with positive and negative frequency with respect to the time variable σ^0 are

$$
u_\omega = \frac{1}{\sqrt{2\omega}} e^{-i\omega\sigma^-} \ , \ u_\omega^* = \frac{1}{\sqrt{2\omega}} e^{i\omega\sigma^-} . \tag{3.8}
$$

The field f has expansion in terms of annihilation and creation operators

$$
f_- = \int_0^\infty d\omega \left[a_\omega u_\omega + a_\omega^\dagger u_\omega^* \right] \ (\text{in}) . \tag{3.9}
$$

The equations of motion imply existence of the conserved Klein-Gordon inner product,

$$
(f, g) = -i \int_\Sigma d\Sigma^\mu f \overleftrightarrow{\nabla}_\mu g^* \tag{3.10}
$$

for arbitrary Cauchy surface Σ. The modes (3.8) have been normalized so that

$$
(u_\omega, u_{\omega'}) = 2\pi\delta(\omega - \omega') = -(u_\omega^*, u_{\omega'}^*) \ , \ (u_\omega, u_{\omega'}^*) = 0 . \tag{3.11}
$$

These, together with the canonical commutation relations

$$
[f_-(x), \partial_0 f_-(x')]_{x^0 = x'^0} = \frac{1}{2} [f(x), \partial_0 f(x')]_{x^0 = x'^0} = \pi i \delta(x^1 - x'^1) \tag{3.12}
$$

imply that the operators a_ω satisfy the usual commutators,

$$
[a_\omega, a_{\omega'}^\dagger] = \delta(\omega - \omega') \ , \ [a_\omega, a_{\omega'}] = 0 \ , \ [a_\omega^\dagger, a_{\omega'}^\dagger] = 0 . \tag{3.13}
$$

Lastly, the in vacuum is defined by

$$
a_\omega |0\rangle_{in} = 0 . \tag{3.14}
$$

To describe states in the future modes are needed both the "out" region near $\mathcal{I}^+$ and near the singularity. The former are the obvious analogues to (3.8),

$$
v_\omega = \frac{1}{\sqrt{2\omega}} e^{-i\omega\xi^-} \ , \ v_\omega^* = \frac{1}{\sqrt{2\omega}} e^{i\omega\xi^-} . \tag{3.15}
$$

The latter are somewhat arbitrary as the region near the singularity is highly curved. A convenient coordinate near the singularity proves to be

$$
\hat{\xi}^- = \ln(\Delta^2 e^{\sigma^-} - \Delta) . \tag{3.16}
$$

and corresponding modes $\widehat{v}_\omega$ and $\widehat{v}^*_\omega$ are given by a formula analogous to (3.15). In terms of these modes, f is written

$$f_- = \int_0^\infty d\omega \left[b_\omega v_\omega + b^\dagger_\omega v^*_\omega + \widehat{b}_\omega \widehat{v}_\omega + \widehat{b}^\dagger_\omega \widehat{v}^*_\omega \right] \text{ (out + internal)} . \tag{3.17}$$

These modes are normalized as in (3.11), and the corresponding field operators obey commutators as in (3.13). The vacua $|0\rangle_{\text{out}}$ and $|0\rangle_{\text{internal}}$ are also defined analogously to (3.14).

The non-trivial relation (2.16) between the natural timelike coordinates in the in and out regions imply that a positive frequency solution in one region is a mixture of positive and negative frequency in another region. This mixing implies particle creation. For example, positive frequency out modes can be expressed in terms of the in modes,

$$v_\omega = \int_0^\infty d\omega' [\alpha_{\omega\omega'} u_{\omega'} + \beta_{\omega\omega'} u^*_{\omega'}] . \tag{3.18}$$

The Fourier coefficients $\alpha_{\omega\omega'}$, $\beta_{\omega\omega'}$ are called Bogoliubov coefficients, and they can be calculated by inverting the fourier transform,

$$\alpha_{\omega\omega'} = \frac{1}{2\pi}(v_\omega, u_{\omega'}) \quad , \quad \beta_{\omega\omega'} = -\frac{1}{2\pi}(v_\omega, u^*_{\omega'}) . \tag{3.19}$$

In the present model they can be given in closed form in terms of incomplete beta functions[7]

$$\begin{aligned} \alpha_{\omega\omega'} &= \frac{1}{2\pi}\sqrt{\frac{\omega'}{\omega}}\Delta^{i\omega} B(-i\omega + i\omega', 1 + i\omega) \\ \beta_{\omega\omega'} &= \frac{1}{2\pi}\sqrt{\frac{\omega'}{\omega}}\Delta^{i\omega} B(-i\omega - i\omega', 1 + i\omega) \ ; \end{aligned} \tag{3.20}$$

Although we will not use these formulas directly they are exhibited for completeness.

To investigate the thermal behavior at late times, $\xi^- \gg -\ln\Delta$, $\sigma^- \simeq -\ln\Delta$, we could examine the asymptotic behavior of (3.20), but a shortcut is to use the asymptotic form of (2.16) valid in this limit,

$$-e^{-\xi^-} = \Delta - e^{-\sigma^-} \simeq \Delta(\sigma^- + \ln\Delta) \equiv \tilde{\sigma}^- . \tag{3.21}$$

Note that this is the same as the relation between Rindler and Minkowski coordinates in the context of accelerated motion. Likewise one finds

$$e^{\hat{\xi}^-} \simeq \tilde{\sigma}^- . \tag{3.22}$$

Next, notice that functions that are positive frequncy in $\tilde{\sigma}^-$ are analytic in the lower half complex $\tilde{\sigma}^-$ plane. Therefore the functions

$$\begin{aligned} u_{1,\omega} &\propto (-\tilde{\sigma}^-)^{i\omega} = v_\omega + e^{-\pi\omega}\widehat{v}^*_\omega \\ u_{2,\omega} &\propto (\tilde{\sigma}^-)^{-i\omega} = \widehat{v}_\omega + e^{-\pi\omega} v^*_\omega \end{aligned} \tag{3.23}$$

are positive frequency. This means that the corresponding field operators $a_{1,\omega}$ and $a_{2,\omega}$ must annihilate the in vacuum. The inverse of the transformation (3.23) gives the relation between field operators,

$$\begin{aligned} a_{1\omega} &\propto b_\omega - e^{-\pi\omega}\hat{b}^\dagger_\omega \\ a_{2\omega} &\propto \hat{b}_\omega - e^{-\pi\omega} b^\dagger_\omega . \end{aligned} \tag{3.24}$$

In particular, the in vacuum obeys

$$\begin{aligned} 0 &= (a^\dagger_{1,\omega} a_{1,\omega} - a^\dagger_{2,\omega} a_{2,\omega})|0\rangle \\ &\propto (b^\dagger_\omega b_\omega - \hat{b}^\dagger_\omega \hat{b}_\omega)|0\rangle \\ &\propto (N_\omega - \hat{N}_\omega)|0\rangle , \end{aligned} \tag{3.25}$$

where N_ω, $\hat{N}_\omega$ are the number operators for the respective modes. This implies that

$$|0\rangle = \sum_{\{n_\omega\}} c(\{n_\omega\})\, |\widehat{\{n_\omega\}}\rangle |\{n_\omega\}\rangle \tag{3.26}$$

for some numbers $c(\{n_\omega\})$. These can be determined up to an overall constant from the equation $a_{1\omega}|0\rangle = 0$:

$$c(\{n_\omega\}) = c(\{0\}) \exp\left\{-\pi \int d\omega\omega n_\omega\right\} . \tag{3.27}$$

Thus the state takes the form

$$|0\rangle = c(\{0\}) \sum_{\{n_\omega\}} e^{-\pi\int d\omega\omega n_\omega} |\widehat{\{n_\omega\}}\rangle |\{n_\omega\}\rangle . \tag{3.28}$$

It is clear from this relation that the state inside the black hole is strongly correlated with the state outside the black hole. Observers outside the hole cannot measure the state inside, and so summarizes their experiments by the density matrix obtained by tracing over all possible internal states,

$$\rho^{\text{out}} = \text{Tr}_{\text{inside}}|0\rangle\langle 0| = |c(\{0\})|^2 \sum_{\{n_\omega\}} e^{-2\pi\int d\omega\omega n_\omega} |\{n_\omega\}\rangle\langle\{n_\omega\}| . \tag{3.29}$$

This is an exactly thermal density matrix with temperature $T = 1/2\pi$. The corresponding energy density is

$$\mathcal{E} = \int_0^\infty \frac{d\omega}{2\pi} \frac{\omega}{e^{\omega/T - 1}} = \frac{\pi}{12} T^2 = \frac{1}{48\pi} , \tag{3.30}$$

which agrees with (3.7) if we account for the unconventional normalization of the stress tensor,

$$\mathcal{E} = \frac{1}{2\pi} T_{00} . \tag{3.31}$$

Both the total entropy and energy of this density matrix are infinite, but that is simply because we have not yet included backreaction which causes the black hole to shrink as it evaporates.

Acknowledgements

This work was supported in part by DOE grant DOE-91ER40618 and by NSF PYI grant PHY-9157463.

References

[1] S.W. Hawking, "Particle creation by black holes," *Comm. Math. Phys.* **43** (1975) 199.

[2] S.W. Hawking, "The unpredictability of quantum gravity," *Comm. Math. Phys* **87** (1982) 395.

[3] C.G. Callan, S.B. Giddings, J.A. Harvey, and A. Strominger, "Evanescent black holes," *Phys. Rev.* **D45** (1992) R1005.

[4] J. Harvey and A. Strominger, "Quantum aspects of black holes," in *Recent Directions in Particle Theory*, proceedings of the Theoretical Advanced Study Institute, Boulder, CO, Jun 1992, J. Harvey and J. Polchinski, eds (World Scientific, 1993); hep-th/9209055.

[5] S.B. Giddings, "Toy models for black hole evaporation," in *String Quantum Gravity and Physics at the Planck Energy Scale*, proceedings of the International Workshop of Theoretical Physics, 6th Session, June 1992, Erice, Italy, N. Sanchez, ed., (World Scientific 1993); hep-th/9209113

[6] N.D. Birrell and P.C.W. Davies, *Quantum fields in curved space* (Cambridge U.P., 1982).

[7] S.B. Giddings and W.M. Nelson, "Quantum emission from two-dimensional black holes," *Phys. Rev.* **D46** (1992) 2486, hep-th/9204072.