Quantum Information on the Black Hole Horizon

Gerard 't Hooft

Institute for Theoretical Physics, University of Utrecht, Princetonplein 5, 3584 CC Utrecht, The Netherlands

Abstract. The scattering matrix approach to the black hole quantization problem is introduced and further elaborated. There appears to be a general concensus about the quantum degeneracy of black holes increasing exponentially with the horizon area. Attempts are described to reproduce the individual quantum states, in particular by exploiting an operator algebra that results from considerations of the gravitational back reaction.

1 Introduction

One would expect that the quantum mechanical properties of a black hole should follow naturally by applying large scale physics. Only the space-time region at the side of the observer, the "physical side" of the horizon, should be relevant. Indeed one can calculate accurately the quantum mechanical effects near a large black hole, as seen by an outside observer, by first studying what an infalling observer would experience, and then performing the appropriate general coordinate transformation. As is to be expected from quantum mechanical calculations, one finds "probabilities": chances that particles of certain types, with certain momenta, energies or other quantum numbers, emerge at certain places. It is when one wants to interpret these outcomes in terms of some Schrödinger equation for the black holes as a whole, that the first genuine problems emerge[1, 2]. In these lectures, the problems we encounter are exposed, and roads towards their resolution are explored.

2 The Space-time Metric of a Black Hole Under Formation

The space-time metric of a stationary, non-rotating and electrically neutral black hole is the *Schwarzschild metric*:

$$\mathrm{d}s^2 = -\Big(1-\frac{2M}{r}\Big)\mathrm{d}t^2 + \frac{\mathrm{d}r^2}{1-2M/r} + r^2\mathrm{d}\Omega^2\,, \tag{2.1}$$

where

$$\mathrm{d}\Omega^2 \equiv \mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\,. \tag{2.2}$$

Here, M stands for Gm_{BH}, where G is Newton's constant and m_{BH} is the black hole mass. Often we will employ the *Kruskal coordinates*[3], which we will write as (x, y, θ, ϕ), with

$$\left(\frac{r}{2M}-1\right)e^{r/2M} = xy\,; \qquad e^{t/2M} = x/y\,. \tag{2.3}$$

In terms of these coordinates we have (see Fig. 1)

$$\frac{dx}{x}+\frac{dy}{y} = \frac{dr}{2M(1-2M/r)}\,; \qquad \frac{dx}{x}-\frac{dy}{y} = \frac{dt}{2M}\,; \qquad ds^2 = \frac{32M^3}{r}e^{-r/2M}\,dx dy + r^2 d\Omega^2\,. \tag{2.4}$$

The apparent singularity at the horizon, $r = 2M$, has disappeared. The only true singularities are at the curves $xy = -1$, where $r = 0$. The region $\{x > 0,\ y > 0\}$ is the "outside region", the only region from which distant observers can obtain any information. The line $y = 0$, where $r = 2M$, is the "future event horizon"; the line $x = 0$, where also $r = 2M$, is the "past event horizon".

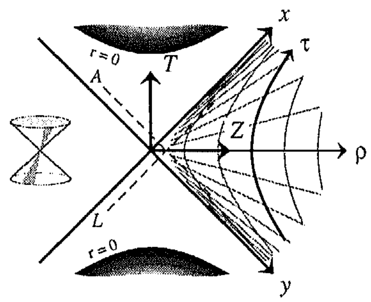

Fig. 1 Various coordinates used to describe the Schwarzschild metric. The local light cones are oriented everywhere as indicated at the left.

In the region $r \approx 2M$ one can write the metric as

$$ds^2 \approx \frac{16M^2}{e}dx dy + 4M^2 d\Omega^2 \tag{2.5}$$

and with the coordinate substitution

$$\frac{4M}{\sqrt{e}}x = Z+T\,, \qquad \frac{4M}{\sqrt{e}}y = Z-T\,, \qquad 2M(\theta-\tfrac{1}{2}\pi) = X\,, \qquad 2M\phi = Y\,, \tag{2.6}$$

at small X, Y, one finds that in terms of these coordinates space-time is approximately flat:

$$ds^2 \approx -dT^2 + dZ^2 + dX^2 + dY^2 . \tag{2.7}$$

The transformation

$$Z = \varrho \cosh \tau , \qquad T = \varrho \sinh \tau , \tag{2.8}$$

brings us back to the Schwarzschild coordinates (close to the horizon), apart from normalization factors:

$$t/2M = 2\tau , \qquad 8M(r - 2M) = \varrho^2 . \tag{2.9}$$

The description of a flat space-time (2.7) in terms of the coordinates (2.8) is called "Rindler space"[4]. We see that close to the horizon, the Schwarzschild coordinates r and t behave as Rindler space coordinates.

To see that black holes can actually be formed by ordinary matter we have to study time-dependent solutions. For details concerning construction of such solutions we refer to[5]. The Penrose diagram for a configuration with both ingoing and outgoing matter is shown in Fig. 2a. In a classical black hole, there is only ingoing matter. In a coordinate frame that is flat at some distance from the black hole, the configuration looks as in Fig. 2b.

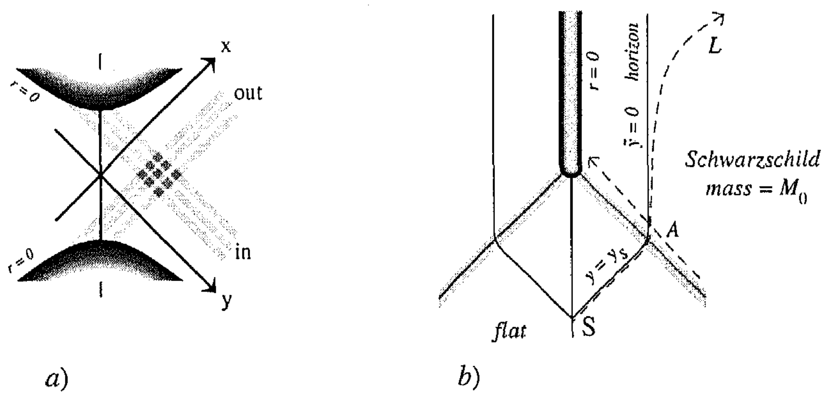

Fig. 2 a) Spherically symmetric configuration of matter radially moving inward and outward with the speed of light. b) Spherically symmetric black hole formed by radially inmoving lightlike matter.

3 Hawking Radiation[6, 7]

Consider the Minkowski coordinate frame $\{T, X, Y, Z\}$, or $\{T, \mathbf{X}\}$ for short, and a scalar field $\Phi(T, \mathbf{X})$. Let this field simply obey a Klein-Gordon equation,

$$(\partial^2 - m^2)\Phi = 0 . \tag{3.1}$$

The quantum theory is written in the Heisenberg representation, which means that the states $|\psi\rangle$ are space-time independent, but the fields are operators depending both on space and on time. Usually, a complete set of solutions of (3.1) is written in terms of the Fourier modes with respect to the Minkowski space coordinates, and one gets

$$\Phi(\mathbf{X}, T) = \int \frac{\mathrm{d}^3\mathbf{k}}{\sqrt{2k^0(\mathbf{k})(2\pi)^3}} \Big(a(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{X} - ik^0 T} + a^\dagger(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{X} + ik^0 T} \Big) , \tag{3.2}$$

$$\dot\Phi(\mathbf{X}, T) = \int \frac{-ik^0\mathrm{d}^3\mathbf{k}}{\sqrt{2k^0(\mathbf{k})(2\pi)^3}} \Big(a(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{X} - ik^0 T} - a^\dagger(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{X} + ik^0 T} \Big) . \tag{3.3}$$

Here $k^0(\mathbf{k}) = \sqrt{\mathbf{k}^2 + m^2}$, and the transformation from a and $a^\dagger$ to Φ and $\dot\Phi$ has been designed such that the following commutation rules are maintained:

$$[\Phi(\mathbf{X}, T),\, \Phi(\mathbf{X}', T)] = 0 , \qquad [\Phi(\mathbf{X}, T),\, \dot\Phi(\mathbf{X}', T)] = i\delta^3(\mathbf{X} - \mathbf{X}') , \tag{3.4}$$

$$[a(\mathbf{k}),\, a(\mathbf{k}')] = 0 , \qquad [a(\mathbf{k}),\, a^\dagger(\mathbf{k}')] = \delta^3(\mathbf{k} - \mathbf{k}') . \tag{3.5}$$

Not only do these commutation rules ensure that $a^\dagger$ and a act as creation and annihilation operators, but also the time dependence in (3.2) and (3.3) implies that the objects created and annihilated carry an amount of energy equal to k^0.

The operator H_M that generates boosts in the time coordinate T,

$$\frac{\partial\Phi}{\partial T} = -i[\Phi, H_M] , \tag{3.6}$$

is the Minkowski-Hamiltonian

$$H_M = \int \mathcal{H}_M(\mathbf{X})\mathrm{d}^3\mathbf{X} = \int \mathrm{d}^3\mathbf{k}\, k^0(\mathbf{k}) a^\dagger(\mathbf{k}) a(\mathbf{k}) . \tag{3.7}$$

We need first the transition to light-cone coordinates, and we define

$$a(\mathbf{k})\sqrt{k^0} = a_1(\tilde k, k^+)\sqrt{k^+} . \tag{3.8}$$

where $k^+ = \frac{1}{\sqrt2}(k^0 + k^3)$, and $\tilde k$ is the transverse component of $\mathbf{k}$. Since

$$\frac{\partial k^+}{\partial k^3}\Big|_{\tilde k} = \frac{1}{\sqrt2}\Big(1 + \frac{k^3}{\mu}\Big) = \frac{k^+}{k^0} , \qquad \mu \equiv \sqrt{\tilde k^2 + {k_3}^2 + m^2} \tag{3.9}$$

the new operators a_1, $a_1^\dagger$ are normalized by

$$[a_1(\tilde{k},k^+),\, a_1^\dagger(\tilde{k}',k^{+\prime})] = \delta^2(\tilde{k}-\tilde{k}')\delta(k^+-k^{+\prime})\,. \qquad (3.10)$$

To obtain operators a_2 that transform neatly under time boosts in Rindler space (i.e., Lorentz boosts in flat space-time), we define them as Fourier transforms with respect to $\ln(k^+)$:

$$a_2(\tilde{k},\omega) = (2\pi)^{-1/2}\int_0^\infty \frac{\mathrm{d}k^+}{\sqrt{k^+}} a_1(\tilde{k},k^+) e^{i\omega \ln\left(\frac{k^+\sqrt{2}}{\mu}\right)}\,, \qquad (3.11)$$

of which the inverse is:

$$a_1(\tilde{k},k^+)\sqrt{k^+} = (2\pi)^{-1/2}\int_{-\infty}^\infty \mathrm{d}\omega a_2(\tilde{k},\omega) e^{-i\omega \ln\left(\frac{k^+\sqrt{2}}{\mu}\right)}\,. \qquad (3.12)$$

With the normalization factors chosen in (3.11, 12), the operators a_2 and $a_2^\dagger$ again obey

$$[a_2(\tilde{k},\omega),\, a_2^\dagger(\tilde{k}',\omega')] = \delta^2(\tilde{k}-\tilde{k}')\delta(\omega-\omega')\,, \qquad (3.13)$$

Let the Minkowski-Hamiltonian H_M obey Eq. (3.7). The Rindler-Hamiltonian, H_R is then defined to be

$$H_R = H_R^I - H_R^{II}\,,\; H_R^I = \int_{\varrho>0} \mathrm{d}^3\mathbf{X}\,\varrho\mathcal{H}_M\,;\; H_R^{II} = \int_{\varrho<0} \mathrm{d}^3\mathbf{X}\,|\varrho|\mathcal{H}_M \qquad (3.14)$$

(note that the Rindler Hamiltonian H_R is dimensionless). If the region $\varrho > 0$ (see Fig. 1) is called quadrant I and $\varrho < 0$ quadrant II, we see that all observables made of fields in quadrant *II* commute with H_R^I and *vice versa*. One finds that

$$H_R = \int_{-\infty}^\infty \mathrm{d}\omega\,\omega\, a_2^\dagger(\tilde{k},\omega)a_2(\tilde{k},\omega)\,; \qquad (3.15)$$
$$H_R^I = \int_0^\infty \mathrm{d}\omega\,\omega\, a_I^\dagger(\tilde{k},\omega)a_I(\tilde{k},\omega)\,; \qquad H_R^{II} = \int_0^\infty a_{II}^\dagger(\tilde{k},\omega)a_{II}(\tilde{k},\omega)\,,$$

where a_I and a_{II} are related to a_2 and $a_2^\dagger$ as follows:

$$\begin{pmatrix} a_I(\tilde{k},\omega) \\ a_{II}(\tilde{k},\omega) \\ a_I^\dagger(-\tilde{k},\omega) \\ a_{II}^\dagger(-\tilde{k},\omega) \end{pmatrix} = \frac{1}{\sqrt{1-e^{-2\pi\omega}}} \begin{pmatrix} 1 & 0 & 0 & e^{-\pi\omega} \\ 0 & 1 & e^{-\pi\omega} & 0 \\ 0 & e^{-\pi\omega} & 1 & 0 \\ e^{-\pi\omega} & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_2(\tilde{k},\omega) \\ a_2(\tilde{k},-\omega) \\ a_2^\dagger(-\tilde{k},\omega) \\ a_2^\dagger(-\tilde{k},-\omega) \end{pmatrix}. \qquad (3.16)$$

They obey

$$[a_I(\tilde{k},\omega),a_I^\dagger(\tilde{k}',\omega')] = [a_{II}(\tilde{k},\omega),a_{II}^\dagger(\tilde{k}',\omega')] = \delta(\omega-\omega')\delta^2(\tilde{k},\tilde{k}')\,;$$
$$[a_I,\,a_{II}] = [a_I,\,a_{II}^\dagger] = 0\,. \tag{3.17}$$

Thus we observe that the corresponding Hilbert space is separable into two factor spaces: $\mathcal{H}=\mathcal{H}_I\otimes\mathcal{H}_{II}$. The space $\mathcal{H}_I$ is described by the Hamiltonian H_R^I and $\mathcal{H}_{II}$ is described by the Hamiltonian $-H_R^{II}$.

The Rindler- or Boulware vacuum state $|0,0\rangle$ is defined by

$$a_I|0,0\rangle = a_{II}|0,0\rangle = 0\,. \tag{3.18}$$

This is not the same as the vacuum experienced by a freely falling observer, who is said to experience the Minkowski- or Hawking vacuum $|\Omega\rangle$, which obeys

$$a_2(\tilde{k},\omega)|\Omega\rangle = 0\,. \tag{3.19}$$

It is not difficult to express this state in terms of the basis generated by a_I and a_{II}:

$$a_I(\tilde{k},\omega)|\Omega\rangle = e^{-\pi\omega}a_{II}^\dagger(-\tilde{k},\omega)|\Omega\rangle\,, \tag{3.20}$$
$$a_{II}(\tilde{k},\omega)|\Omega\rangle = e^{-\pi\omega}a_I^\dagger(-\tilde{k},\omega)|\Omega\rangle\,,$$

so that

$$|\Omega\rangle = \prod_{\tilde{k},\omega}\sqrt{1-e^{-2\pi\omega}}\sum_{n=0}^{\infty}|n\rangle_I|n\rangle_{II}e^{-\pi n\omega}\,, \tag{3.21}$$

where the square root is added for normalization.

Notice that

$$H_R|\Omega\rangle = (H_R^I-H_R^{II})|\Omega\rangle = 0\,, \tag{3.22}$$

which confirms that $|\Omega\rangle$ is Lorentz invariant; remember that H_R is the generator of Lorentz boosts.

If one does not have the means to observe any features at $\varrho<0$ this implies that one only has at one's disposal operators $\mathcal{O}$ composed of the fields in region I, that is, the operators a_I and $a_I^\dagger$. These act only in the factor space $\mathcal{H}_I$ but are proportional to the identity operator in $\mathcal{H}_{II}$:

$$\mathcal{O}\Big(|\psi\rangle_I|\psi'\rangle_{II}\Big) = |\psi'\rangle_{II}\Big(\mathcal{O}|\psi\rangle_I\Big)\,. \tag{3.23}$$

Let us limit ourselves momentarily to a single point $(\tilde{k},\omega)$. There the expectation value for such an operator in the state $|\Omega\rangle$ is

$$\langle\Omega|\mathcal{O}|\Omega\rangle = (1-e^{-2\pi\omega})\sum_{n_1,n_2}{}_{II}\langle n_1|\,{}_I\langle n_1|\mathcal{O}|n_2\rangle_I\,|n_2\rangle_{II}e^{-\pi\omega(n_1+n_2)}$$
$$= \sum_{n\geq 0}{}_I\langle n|\mathcal{O}|n\rangle_I e^{-2\pi n\omega}(1-e^{-2\pi\omega}) = \mathrm{Tr}\big(\mathcal{O}\,\varrho_\Omega\big)\,, \tag{3.24}$$

where ϱ_Ω is the *density matrix* $Ce^{-\beta H_I}$ corresponding to a thermal state at the temperature[6] $T = 1/2\pi$. Note that in Rindler space time, energy and temperature are dimensionless. If we scale with the appropriate factor $4M$ as in Eq. (2.9) we find the *Hawking temperature*,

$$T_H = 1/8\pi M = 1/8\pi G m_{\rm BH} . \qquad (3.25)$$

This result is highly independent of the way the black hole was formed. In case the collapse took place in a less symmetric way, or at various steps and intervals, one still finds that an observer falling in the black hole should observe the Hawking vacuum state there, and this necessarily leads to the density matrix ϱ_Ω. In particular, one could assume that the collapsing matter was in a *pure* quantum state, and even in that case, the outgoing radiation appears to be mixed according to the matrix ϱ_Ω. The question to be asked is how literally this result is to be taken. One could conclude

i) that black holes must be fundamentally different from other objects in nature. They do not obey a single Schrödinger equation (which after all would allow pure states to evolve only into pure states), but instead obey probabilistic equations of motion that are not purely quantum mechanical[2].

According to this view, a more basic theory at the Planck scale would show no quantum mechanical features of the familiar kind. Alternatively, perhaps,

ii) black holes do obey a Schrödinger equation, but this equation requires knowledge of all inaccessible observables behind the horizon, so that a black hole forms an infinitely degenerate state. In this case the black hole can never decay completely[7], but it decays into stable, infinitely degenerate, final states with masses of the order of the Planck mass, called *remnants.* Thirdly, however, one may suspect that

iii) the density matrix derivation depended on certain hidden assumptions of a statistical nature, such that the answer may be correct in a statistical sense, but more precise treatments may yield a purely quantum mechanical description of a black hole that nevertheless has only a finite degeneracy. This is the scattering matrix assumption, which we will further investigate from section 7 onwards.

Thus, one expects the system as a whole to react just as any other physical system does: when it absorbs infalling material, or even just infalling radiation, it should react some way or other, and enter into a state that is orthogonal to what it would have evolved into if the infalling material had been in a different mode or totally absent. This is just the experimentally observed fact that all known evolution laws in small-scale physics can be written in terms of a *unitary* evolution matrix. It is hard to understand how the world at the scale of ordinary atomic and elementary particle physics could behave quantum mechanically and evolve in a unitary way, if quantum mechanics were not at the basis of the laws of dynamics at the smallest distance scales.

It appears that the derivation of the density matrix ϱ_Ω in Eq. (3.24) cannot be exactly right, since it implies that infalling material of whatever variety should not affect the outgoing radiation at all (linearized quantum field theory was used). This would violate unitarity.

The density matrix ϱ_Ω has to be replaced by a pure state.

4 Black Hole Entropy, and its Interpretation in Terms of Quantum States

The fact that the radiation emitted, as described by Eq. (3.25), is *thermal*, opens up the possibility to approach this phenomenon from a thermodynamical point of view. Taking $m_{\rm BH}$ to be the energy and $T = T_H$ the temperature, one readily derives the *entropy* S:

$$T\mathrm{d}S = \mathrm{d}m_{\rm BH}\,; \qquad \mathrm{d}S = 8\pi G m_{\rm BH}\mathrm{d}m_{\rm BH}\,; \qquad S = 4\pi G m^2_{\rm BH} + C\,, \tag{4.1}$$

where C is an unknown integration constant, to be referred to as the "entropy normalization constant".

It is important to note that the expression obtained for the entropy S, apart from the integration constant, is always equal to $\frac{1}{4}A/G$, where A is the *area* of the horizon, a finding that will be very much at the center of our discussions.

Connecting the entropy to the *density of quantum mechanical states*[8], must be done with considerable care, since there will be two kinds of divergences: at the horizon and at spacelike infinity. In fact, one may very well question the mere *existence* of such quantum levels. This, however, is the key assumption of this paper: not only is the quantum mechanics of black holes meaningful, it can also be *derived*, and the constant C in Eq. (4.1) is finite and of order one (apart from subdominant terms). In order to enable us to judge the relation between the entropy just derived, and the density of quantum states, we now present a direct argument concerning the density of states, an argument that will also show that any infinities at the *horizon* must be absorbed in C, but the "infrared" infinities arising from spacelike infinity should be excluded; the latter represent the radiation field far from the black hole.

The spectral density of a black hole can be derived from its Hawking temperature by applying time reversal invariance[9]. We have to our disposal both the *emission rate* (the Hawking radiation intensity), and the *capture probability*, or the effective cross section of the black hole for infalling matter.

The cross section σ is approximately determined by the radius r^+ of the horizon:

$$\sigma = 2\pi r_+^2 = 8\pi M^2\,, \tag{4.2}$$

and slightly more for objects moving in slowly. The emission probability $W\mathrm{d}t$ for a given particle type, in a given quantum state, in a large volume $V = L^3$ is:

$$W\mathrm{d}t = \frac{\sigma(\mathbf{k})v}{V}\, e^{-E/T}\mathrm{d}t\,, \tag{4.3}$$

where **k** is the wave number characterizing the quantum state, v is the particle velocity, and E is its momentum.

Now we *assume* that the process is also governed by a Schrödinger equation. This means that there exist quantum mechanical transition amplitudes,

$$\begin{aligned} \mathcal{T}_{\text{in}} &= {}_{\text{BH}}\langle M+GE| \;\; |M\rangle_{\text{BH}}|E\rangle_{\text{in}}\,, \\ \text{and} \qquad \mathcal{T}_{\text{out}} &= {}_{\text{BH}}\langle M|\,{}_{\text{out}}\langle E| \;\; |M+GE\rangle_{\text{BH}}\,, \end{aligned} \tag{4.4}$$

where the states $|M\rangle_{\text{BH}}$ represent black hole states with mass M/G, and the other states are energy eigenstates of particles in the volume V. In terms of these amplitudes, using the so-called Fermi Golden Rule, the cross section and the emission probabilities can be written as

$$\sigma = |\mathcal{T}_{\text{in}}|^2 \varrho(M+GE)/v\,, \tag{4.5}$$

$$W = |\mathcal{T}_{\text{out}}|^2 \varrho(M)\frac{1}{V}\,, \tag{4.6}$$

where $\varrho(M)$ stands for the level density of a black hole with mass M. The factor v^{-1} in Eq. (4.5) is a kinematical factor, and the factor V^{-1} in W arises from the normalization of the wave function.

Now, time reversal invariance relates $\mathcal{T}_{\text{in}}$ to $\mathcal{T}_{\text{out}}$ (through complex conjugation). To be precise, all one needs is PCT invariance, since the parity transformation P and charge conjugation C have no effect on our calculation of σ. Dividing the expressions (4.5) and (4.6), and using (4.3), one finds:

$$\frac{\varrho(M+GE)}{\varrho(M)} = e^{E/T} = e^{8\pi ME} \tag{4.7}$$

(naturally, we assume the energy E to be small compared to the black hole mass M, so that the E^2 terms are relatively insignificant). This is easy to integrate:

$$\varrho(M) = e^{4\pi M^2/G + C} = e^S\,. \tag{4.8}$$

For large enough black holes, Eq. (4.8) may be rewritten as

$$\varrho(M) = 2^{A/A_0}\,, \tag{4.9}$$

where A is the horizon area and A_0 is a fundamental unit of area,

$$A_0 = 4G\ln 2\,. \tag{4.10}$$

This suggests a spin-like degree of freedom on all surface elements of size A_0.

As stated earlier, the importance of this derivation is the fact that the expressions used as starting points are the *actual* Hawking emission rate and the *actual* black hole absorption cross section. This implies that, if in more detailed considerations divergences are found near the horizon, these divergences should not be used as arguments to adjust the relation between entropy and level density by large renormalization factors. Furthermore, the Golden Rule argument can be used only to deal with one emitted particle at the time. Hence, we should not take the outside volume V so large that the dominant emission mode contains very many particles. Therefore, any divergences found when the outside volume is taken to infinity should be subtracted.

Extension to the more general Kerr-Newman solutions is straightforward.

5 The Brick Wall Model[9]

In this Section, we now present a model in which only *low energy* quantum fluctuations of the fields are taken into account. We apply quantum field theory up to some point r_1 close to the horizon: $r_1 = r_+ + h,\ h > 0$. For simplicity we only consider scalar fields $\Phi_i(r,\theta,\phi,t)$, whose only interaction is the gravitational one with the metric; generalization towards spinor, vector or even perturbative gravitational field excitations will be straightforward. To simplify things, we just represent all those by giving the fields Φ_i a multiplicity N, so $i = 1\dots, N$. At $r = r_1$ we impose a boundary condition:

$$\Phi_i(r,\theta,\phi,t) = 0 \qquad \text{if} \qquad r \le r_1\,. \tag{5.1}$$

The quanta of the fields will be given a temperature $T = T_H$. The question one may ask is: which value should one assign to the cutoff parameter h, such that the entropy of this system precisely takes the value (4.1), so that the density of quantum states corresponds to (4.8)? We will need an infrared cutoff in the form of a box with radius L:

$$\Phi_i(r,\theta,\phi,t) = 0 \qquad \text{if} \qquad r \ge L\,. \tag{5.2}$$

To determine the thermodynamic properties of this system, one must compute the energy levels $E(n,\ell,\ell_3)$ of the bosons Φ_i. The Lagrange density $\mathcal{L}$ in the metric (2.1) is given by

$$2\mathcal{L}(\mathrm{x},t) \;=\; \Big(1-\frac{2M}{r}\Big)^{-1}(\partial_t\Phi_i)^2 \;-\; \Big(1-\frac{2M}{r}\Big)(\partial_r\Phi_i)^2 \;-\; r^{-2}(\partial_\Omega\Phi_i)^2 \;-\; m_i^2\Phi_i^2\,. \tag{5.3}$$

In the approximation

$$m_i^2 \ll 2M/\beta^2 h\,, \qquad L \gg 2M\,, \tag{5.4}$$

the main contributions to free energy at a temperature $T = \beta^{-1}$ is found to be[9, 5]

$$F \approx -\frac{2\pi^3 N}{45h}\Big(\frac{2M}{\beta}\Big)^4 - \frac{2}{9\pi}L^3 N\int_m^\infty \frac{\mathrm{d}E(E^2-m^2)^{3/2}}{e^{\beta E}-1}\,. \tag{5.5}$$

The second part is the usual contribution from the vacuum surrounding the black hole at great distances, and as argued before, should be discarded. The first part is an intrinsic contribution from the horizon, and it is seen to diverge linearly as $h \downarrow 0$.

The contribution of the horizon to the total energy U and the entropy S are

$$U = \frac{\partial}{\partial\beta}(\beta F) = \frac{2\pi^3}{15h}\Big(\frac{2M}{\beta}\Big)^4 N\,, \tag{5.6}$$

$$S = \beta(U - F) = \frac{8\pi^3}{45h} 2M\Big(\frac{2M}{\beta}\Big)^3 N\,. \tag{5.7}$$

When this is adjusted to the Hawking value, Eq. (4.1), with $\beta = 1/T_H = 8\pi M$, we find that the cutoff parameter h must be chosen to be

$$h = \frac{NG}{720\pi M}\,. \tag{5.8}$$

The total energy U of the thermally excited particles is given by

$$GU = \tfrac{3}{8}M\,, \tag{5.9}$$

independently of N. Alternatively, one could have tuned the energy U to be equal to $m_{\rm BH}$, which would yield the same order of magnitude for h, but adjusting the physical degrees of freedom, *i.e.*, the entropy S, appears to us more sensible. Clearly, it makes little sense to allow $h\overset{?}{\rightarrow}0$, since then both the entropy and the energy would diverge.

We refer to the cutoff near the horizon as a "brick wall". The physical distance between the brick wall and the horizon is

$$\int_{r=2M}^{r=2M+h} \mathrm{d}s = \int_{2M}^{2M+h} \frac{\mathrm{d}r}{\sqrt{1-2M/r}} = 2\sqrt{2Mh} = \sqrt{\frac{NG}{90\pi}}\,, \tag{5.10}$$

which is independent of the mass $m_{\rm BH}$ of the black hole. The brick wall should be a property of any horizon of arbitrary size. If N is not too large, the brick wall thickness is of the order of the Planck length.

The brick wall model, with the values of β and h fixed according to Equations (3.25) and (4.1), actually reproduces the thermodynamic properties of a black hole quite nicely, and could have served as a realistic model for a black hole that fully obeys Schrödinger's equation and preserves quantum coherence, except for the fact that it also preserves all symmetries of the underlying quantum field theory; it could generate chemical potentials for the various globally conserved quantum numbers. Thus, not only the temperature must be constrained to keep the Hawking value, but also the chemical potentials are constrained to be zero. In principle, this is easy to realize, simply by introducing symmetry breaking effects in the brick wall boundary condition, but probably one would then be pushing this model too far; anyway, its most important deficiency is that we completely gave up invariance under general coordinate transformations near the horizon.

The most important lesson to be learned from the brick wall model is that Hawking radiation can indeed be seen to be compatible with quantum mechanical purity, if only one could introduce a cutoff at the Planck scale.

6 The Aichelburg-Sexl Metric near a Black Hole

The gravitational effect of an infalling particle in the Schwarzschild metric can be understood when we transform to a locally flat space-time, Eqs. (2.6). Consider the coordinate frames of Section 2. As Schwarzschild time t, or equivalently, Rindler time τ, evolves, the infalling particle is Lorentz boosted, as we see in Eq. (2.8). In terms of the flat coordinates, therefore, the energy of the particles increases exponentially, and thus it quickly reaches values where gravitational effects can no longer be ignored. These effects are easy to calculate in the approximation that the source particle moves with the speed of light[10, 11].

For simplicity, consider the case that the surrounding space-time is completely flat. In the rest frame we can approximate the metric as

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \frac{2\mu}{r}\mathrm{d}t^2 + \frac{2\mu}{r}\mathrm{d}r^2 \,, \tag{6.1}$$
$$r \equiv \sqrt{x_1^2 + x_2^2 + x_3^2} \,, \qquad \mathrm{d}x^2 \equiv -\mathrm{d}t^2 + \mathrm{d}r^2 + r^2\mathrm{d}\Omega^2 \,.$$

where $\mu = Gm$, and m is the mass of the source particle. This we rewrite as

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \frac{2\mu}{r}(u\cdot\mathrm{d}x)^2 + \frac{2\mu}{r}\mathrm{d}r^2 \,, \qquad r = \sqrt{x^2 + (u\cdot x)^2} \,, \tag{6.2}$$

where

$$u = (0,0,0,i) \,; \qquad u^2 = -1 \,. \tag{6.3}$$

In these expressions, we have neglected all effects that are of higher order in the particle mass μ, since we choose μ to be small. The particle's Schwarzschild radius r_+ is very small, and the Lorentz boost to be considered next will only further reduce the particle's dimensions.

The advantage of the notation chosen in Eq. (6.2) is of course that now the Lorentz boost is straightforward. In the boosted frame we can take

$$m\,u^\mu \Rightarrow (0,0,p,ip) = p^\mu \,, \qquad Gp = \frac{\mu v}{\sqrt{1-v^2/c^2}} \gg \mu \,. \tag{6.4}$$

In the limit $\mu \Rightarrow 0$, p fixed, one has $r \Rightarrow |x\cdot u|$. It will turn out to be useful to compare the metric then obtained with the flat space-time metric in two coordinate frames $y^\mu_{(\pm)}$, defined as

$$y^\mu_{(\pm)} = x^\mu \pm 2\mu\,u^\mu \log r \,. \tag{6.5}$$

We have:

$$\mathrm{d}y^2_{(\pm)} = \mathrm{d}x^2 \pm \frac{4\mu}{r}(u\cdot \mathrm{d}x)\,\mathrm{d}r - 4\mu^2\,\frac{\mathrm{d}r^2}{r^2}\,; \tag{6.6}$$

$$\mathrm{d}s^2 - \mathrm{d}y^2_{(\pm)} = \frac{2\mu}{r}\,\mathrm{d}\big[r \mp (u\cdot x)\big]^2 + 4\mu^2(\mathrm{d}\log r)^2\,. \tag{6.7}$$

Now consider the limit (6.4). We keep p fixed but let μ tend to zero. We now claim that when $(p\cdot x) > 0$, the metric $\mathrm{d}s^2$ approaches the flat metric $\mathrm{d}y^2_{(+)}$, whereas when $(p\cdot x) < 0$, we have $\mathrm{d}s^2 \Rightarrow \mathrm{d}y^2_{(-)}$, and at the plane defined by $(p\cdot x) = 0$ these two flat space-times are glued together according to

$$y^\mu_{(+)} = y^\mu_{(-)} + 4\mu\,u^\mu \log|\tilde{x}|\,, \tag{6.8}$$

where $\tilde{x} = (X, Y, 0, 0)$, the transverse part of the coordinates y^μ.

Verifying the flatness of space-time away from the plane $(p\cdot x) = 0$, is easy, but to ascertain the connection formula (6.8), is a bit more delicate. One can show[5] that

$$\begin{aligned}
\mathrm{d}s^2 \quad &\to \mathrm{d}y^2_{(+)} \qquad \text{if} \quad (p\cdot x) \gtrsim 0\,, && (6.9\mathrm{A})\\
\mathrm{d}s^2 \quad &\to \mathrm{d}y^2_{(-)} \qquad \text{if} \quad (p\cdot x) \lesssim 0\,, && (6.9\mathrm{B})\\
y_{(+)} &= y_{(-)} + 4\mu\,u^\mu \log r \quad \text{in the region} \quad (p\cdot x) \approx 0\,, && (6.9\mathrm{C})
\end{aligned}$$

which is equivalent to Eq. (6.8). This defines the Aichelburg-Sexl metric[10].

The effect a fast moving particle has on the surrounding space-time, is visualized in Fig. 3. In terms of light cone coordinates, we have the connection formula

$$\begin{aligned}
x^+{}_{(+)} - x^+{}_{(-)} &= 4Gp^+ \log|\tilde{x}| = 4\sqrt{2}\,Gp \log|\tilde{x}|\,;\\
x^-{}_{(+)} - x^-{}_{(-)} &= 0\,.
\end{aligned} \tag{6.10}$$

Here, $\tilde{x}$ is the transverse distance from the source particle, which is moving (highly relativistically) along the line $x^- = \tilde{x} = 0$. The r.h.s. of Eq. (6.10) is a Green function, $-\sqrt{2}f(\tilde{x})p$, satisfying the equation

$$\tilde{\partial}^2 f(\tilde{x}) = -8\pi G\,\delta^2(\tilde{x})\,, \tag{6.11}$$

where the sign is chosen such that f is large for small values of $\tilde{x}$.

This result can be generalized to the case of a particle moving into a finite size black hole. More details can be found in refs[11, 5]. For most purposes, however, it is sufficient to look at Rindler space, which corresponds to an infinite size black hole. There, the fast particle produces the Aichelburg-Sexl metric as described above.

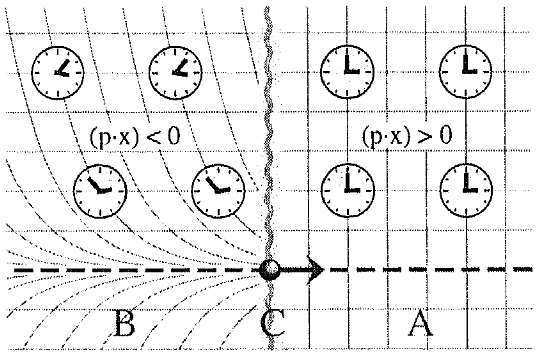

Fig. 3 Snapshot of the gravitational shock wave caused by a highly relativistic particle. If we have a rectangular grid and synchronized clocks before the particle passed by (region A), then, behind the particle (region B), the grid will be deformed, and the clocks desynchronized. The shift is proportional to the logarithm of the transverse distance.

7 Constructing the S-Matrix from the Gravitational Back Reaction

The postulate that scattering of particles against a black hole can be described by a quantum mechanical scattering matrix is an *assumption* that cannot be proved from the principles of quantum field theory and general relativity alone. Indeed, it may well be at variance with these theories, if the latter would be extrapolated to beyond the Planck scale. The S-matrix Ansatz applied here may be seen as a new physical principle, perhaps comparable to Max Planck's new postulate in his 1900 paper, that energies are quantized. The S-matrix Ansatz reads as follows[12]:

> *All physical interaction processes that begin and end with free, stable particles moving far apart in an asymptotically flat space-time, therefore also all those that involve the creation and subsequent evaporation of a black hole, can be described by one scattering matrix S relating the asymptotic outgoing states $|\mathrm{out}\rangle$ to the ingoing states $|\mathrm{in}\rangle$.*

In essence, the Ansatz will be used in the following way[13, 12]: consider one state $|\mathrm{in}_0\rangle$ and one state $|\mathrm{out}_0\rangle$, with a possible black hole in their connecting history. We assume some value for the transition amplitude $\langle\mathrm{out}_0|\mathrm{in}_0\rangle = \mathcal{N}$. This means that we replaced the out-state produced by the Hartle-Hawking!vacuum, which actually was a quantum mechanical mixture of states, by one arbitrary choice

$|\text{out}_0\rangle$. Then, using all the physical laws that we know and trust, we compute neighboring S-matrix elements, $\langle \text{out}_0 + \delta_{\text{out}} | \text{in}_0 + \delta_{\text{in}} \rangle$.

If there were no interactions, the effects from δ_{in} onto the out-states would not have been discernable. All amplitudes would have to be equal, and the scattering matrix thus obtained could never be unitary. Since in the calculations of Section 5, interactions between the Φ particles were ignored, those calculations were not good enough to give us our S-matrix. In this section, we will take only one type of interaction into account: the gravitational shift computed in the previous section. Thus, we only consider particles moving in and out in the longitudinal direction, with hyper-relativistic speeds when they are near the horizon. Far away from the horizon, as soon as $r - 2M = \mathcal{O}(2M)$, they will be allowed to go slower, indeed, out-moving particles may turn around to fall back in again. What has to be done in order to accommodate for such possibilities, is to define the S-matrix to consist of three ingredients:

$$S = S_{\text{out}} S_{\text{hor}} S_{\text{in}} , \tag{7.1}$$

where S_{in} relates the asymptotic in-states to wave packets moving inwards very near the horizon, S_{out} connects wave packets moving outwards very near the horizon to the asymptotic out-states, and S_{hor} is the really important part telling us how particles moving inwards very near the horizon affect the outgoing particles very near the horizon. S_{in} and S_{out} follow unambiguously from known laws of low-energy physics, and require little discussion.

In the limit $M \to \infty$, the region near the horizon can be described as a Rindler space. The angles θ and ϕ are replaced by flat transverse coordinates, and we rescale the momentum p accordingly. We recover the shift (6.10), determined by the Green function f of Eq. (6.11).

To begin our construction of the S-matrix, let us take δ_{in} to be one extra particle going in with momentum δp^-_{in}, at the transverse position $\tilde{x}'$. Since we use the conventions of Section 4, the value of δp^-_{in} is negative. The outgoing particles, at points $\tilde{x}$ near to the point $\tilde{x}'$, are shifted inwards, so that δx^-_{out} is negative, and

$$x^-_{\text{out}} \to x^-_{\text{out}} + \delta x^-_{\text{out}}(\tilde{x}) \,, \qquad \delta x^-_{\text{out}}(\tilde{x}) = f(\tilde{x} - \tilde{x}')\delta p^-_{\text{in}} \,, \tag{7.2}$$

where f obeys Eq. (6.11), or, if from now on $8\pi G = 1$,

$$\tilde{\partial}^2 f(\tilde{x}) = -\delta^2(\tilde{x}) \,. \tag{7.3}$$

We now temporarily suppress the superscripts $\{^\pm\}$, since the subscripts in and out suffice, and later we want to reintroduce $\{^\pm\}$ with different sign conventions. Any outgoing particle has a wave packet ψ, oscillating as $e^{ip_{\text{out}}x_{\text{out}}}$. With the shift δx_{out}, this wave turns into

$$e^{ip_{\text{out}}x_{\text{out}} - ip_{\text{out}}\delta x_{\text{out}}} = e^{-i\int \mathrm{d}^2\tilde{x}\left[\delta x_{\text{out}}(\tilde{x})\hat{P}_{\text{out}}(\tilde{x})\right]} \psi \,, \tag{7.4}$$

where $\hat{P}_{\text{out}}(\tilde{x})$ is the operator that generates a shift at transverse position $\tilde{x}$. It is also the total momentum density of the outgoing particles at transverse position $\tilde{x}$.

Now combining this with Eq. (7.2), we see that

$$\psi \Rightarrow \psi' = e^{-i\int \mathrm{d}^2\tilde{x}\left[\delta p_{\text{in}} f(\tilde{x}-\tilde{x}')\hat{P}_{\text{out}}(\tilde{x})\right]}\,\psi\,. \tag{7.5}$$

Repeating this many times, adding (or removing) different ingoing particles in the in-state, with momenta adding all up to $P_{\text{in}}(\tilde{x}')$ at the transverse position $\tilde{x}'$, we see that the total effect is:

$$\psi' = e^{-i\int \mathrm{d}^2\tilde{x}\mathrm{d}^2\tilde{x}'\left[P_{\text{in}}(\tilde{x}')f(\tilde{x}-\tilde{x}')\hat{P}_{\text{out}}(\tilde{x})\right]}\,\psi\,. \tag{7.6}$$

Notice the complete symmetry between in- and outgoing particles. $P_{\text{in}}(\tilde{x}')$ refers to all momenta of particles going in during a certain epoch where we have control over the ingoing particles. $\hat{P}_{\text{out}}(\tilde{x})$ refers to all particles seen going out during a similar epoch of observations. Before or after these two epochs we do not have the opportunity to observe or control. The states there are kept fixed as much as is possible. Of course, both P_{in} and $\hat{P}_{\text{out}}$ are operators; from now on we omit the hat $(\hat{\ })$.

Noting that, according to the result of Section 4, the total number of quantum states should be finite, we have reasons to believe that, by adding or subtracting a sufficient number of particles, we can generate *all* in-states from $|\text{in}_0\rangle$, and for the out-states it is even more natural to have $P_{\text{out}}(\tilde{x})$ refer to *all* outgoing particles. It is suggested to describe the in- and out states exclusively by giving the functions $P_{\text{in}}(\tilde{x})$ and $P_{\text{out}}(\tilde{x})$. One then obtains

$$\langle\{P_{\text{out}}(\tilde{x})\}|\{P_{\text{in}}(\tilde{x})\}\rangle = \mathcal{N}\exp\left[-i\int \mathrm{d}^2\tilde{x}\,\mathrm{d}^2\tilde{x}' P_{\text{in}}(\tilde{x}')f(\tilde{x}-\tilde{x}')P_{\text{out}}(\tilde{x})\right], \tag{7.7}$$

where $\mathcal{N}$ is a common normalization factor. The magnitude of this factor is fixed by requiring S to be unitary; its phase cannot be determined, but in most cases it will be a freely adjustable parameter anyway, since our amplitude tends to violate global conservation laws.

This scattering matrix is indeed unitary, if one imposes the inner product

$$\langle\{P_{\text{in}}(\tilde{x})\}|\{P_{\text{in}}{}'(\tilde{x})\}\rangle = \mathcal{N}'\prod_{\tilde{x}}\delta\left(P_{\text{in}}(\tilde{x}) - P_{\text{in}}{}'(\tilde{x})\right), \tag{7.8}$$

for the in-states, with again some normalization parameter $\mathcal{N}'$, and we impose a similar inner product rule for the out-states.

We should hasten to add, that the S-matrix (7.7) cannot be the ultimate result of our theory, since the states $|\{P_{\substack{\text{in}\\\text{out}}}(\tilde{x})\}\rangle$ with the inner product (7.8) form a *continuum* of states, and this is not the result we want. What this really means is that we still expect some cut-off mechanism when $|\tilde{x}-\tilde{x}'|$ approaches the Planck length. Indeed, if $|\tilde{x}-\tilde{x}'|$ approaches the Planck length, our present result is invalid, since then the *transverse* components of the momenta also produce shifts, and those have not been taken into account. If, however, we limit ourselves to a "coarse grained" description, specifying only features that are large compared to the Planck length, and if it could indeed be accepted that restricting oneself to the gravitational interaction forces only (and of those only the longitudinal ones), is reasonable, then (7.7) seems to be a reasonable approximation to the S-matrix that we are looking for. In view of this, let us first further analyze what this S-matrix implies.

8 Functional Operator Algebra on the Horizon

Consider the Hilbert space of in-states $|\{P_{\rm in}(\tilde x)\}\rangle$ with inner product (7.8), and define an operator $U_{\rm in}(\tilde x)$ that is canonically conjugated to $P_{\rm in}(\tilde x)$:

$$[P_{\rm in}(\tilde x),\, U_{\rm in}(\tilde x')] \;=\; -i\delta^2(\tilde x-\tilde x')\,, \tag{8.1}$$

$$[P_{\rm in}(\tilde x),\, P_{\rm in}(\tilde x')] \;=\; [U_{\rm in}(\tilde x),\, U_{\rm in}(\tilde x')] \;=\; 0 \tag{8.2}$$

(We regard all these operators as acting on in-states). The eigenstates of $U_{\rm in}(\tilde x)$ are the functional Fourier transforms of the eigenstates $|\{P_{\rm in}(\tilde x)\}\rangle$ of the operators $P_{\rm in}$:

$$|\{U_{\rm in}(\tilde x)\}\rangle \;=\; \mathcal{N}''\int \mathcal{D}P_{\rm in} e^{-i\int {\rm d}\tilde x\, P_{\rm in}(\tilde x)U_{\rm in}(\tilde x)}|\{P_{\rm in}(\tilde x)\}\rangle\,, \tag{8.3}$$

where $\mathcal{N}''$ is again a normalization factor.

Writing this as

$$\langle\{U_{\rm in}(\tilde x)\}|\{P_{\rm in}(\tilde x)\}\rangle \;=\; \mathcal{N}'''e^{i\int {\rm d}\tilde x\, P_{\rm in}(\tilde x)U_{\rm in}(\tilde x)}\,, \tag{8.4}$$

we find that the states $|\{U_{\rm in}(\tilde x)\}\rangle$ can be expressed in terms of the states $|\{P_{\rm out}(\tilde x)\}\rangle$, by using Equ. (7.7).

We find:

$$U_{\rm in}(\tilde x') \;=\; -\int {\rm d}\tilde x\, f(\tilde x-\tilde x')P_{\rm out}(\tilde x)\,, \tag{8.5}$$

and similarly:

$$U_{\rm out}(\tilde x') \;=\; \int {\rm d}\tilde x\, f(\tilde x-\tilde x')P_{\rm in}(\tilde x)\,, \tag{8.6}$$

where $U_{\rm out}$ is the operator canonically conjugated to $P_{\rm out}$, since in addition to Eqs. (8.1) and (8.2) we have for the out-states:

$$[P_{\rm out}(\tilde x),\, U_{\rm out}(\tilde x')] \;=\; -i\delta^2(\tilde x-\tilde x')\,, \tag{8.7}$$

$$[P_{\rm out}(\tilde x),\, P_{\rm out}(\tilde x')] \;=\; [U_{\rm out}(\tilde x),\, U_{\rm out}(\tilde x')] \;=\; 0 \tag{8.8}$$

By virtue of the fact that Eqs (8.5) and (8.6) relate operators on in-states to operators on out-states, we say that these generate the S-matrix. Rewriting the equations as

$$\tilde{\partial}^2 U_{\rm in}(\tilde{x}) = P_{\rm out}(\tilde{x}) , \qquad \tilde{\partial}^2 U_{\rm out}(\tilde{x}) = -P_{\rm in}(\tilde{x}) , \tag{8.9}$$

underlines the local nature of these equations with respect to the transverse coordinates $\tilde{x}$. Also:

$$\langle\{U_{\rm out}(\tilde{x})\}|\{U_{\rm in}(\tilde{x})\}\rangle = \mathcal{N}'''' \exp\big[-i\int {\rm d}^2\tilde{x}\,\tilde{\partial}U_{\rm out}(\tilde{x})\cdot\tilde{\partial}U_{\rm in}(\tilde{x})\big] . \tag{8.10}$$

Because of its local nature, this equation may be suspected to be more elementary than Eq. (7.7), which was derived earlier. Combining (8.10) with (8.4) and the analogous inner product between the $U_{\rm out}$ and $P_{\rm out}$ eigenstates, we rewrite Eq. (7.7) as

$$\langle\{P_{\rm out}(\tilde{x})\}|\{P_{\rm in}(\tilde{x})\}\rangle = \mathcal{N}\int \mathcal{D}U_{\rm in}(\tilde{x})\int \mathcal{D}U_{\rm out}(\tilde{x}) \tag{8.11}$$

$$\exp\big[\; i\int {\rm d}^2\tilde{x}\{-\tilde{\partial}U_{\rm out}(\tilde{x})\cdot\tilde{\partial}U_{\rm in}(\tilde{x}) + P_{\rm in}(\tilde{x})U_{\rm in}(\tilde{x}) - P_{\rm out}(\tilde{x})U_{\rm out}(\tilde{x})\}\big] ,$$

where $\mathcal{N}$ is again a different but universal normalization factor (henceforth, we write such factors simply as $\mathcal{N}$.)

Imagine now that both the in- and the out-state can be completely composed of a finite number, $N = N_{\rm in} + N_{\rm out}$, of particles. Let us denote the momenta of the ingoing particles as ${p^{-,i}}_{\rm in}$, $i = 1,\dots,N_{\rm in}$, entering at transverse coordinates $\tilde{x}^i$, and those of the outgoing particles, at transverse coordinates $\tilde{x}^j$, as $-{p^{+,j}}_{\rm out}$, $j = N_{\rm in}+1,\dots,N$. The reason for the minus sign here, is that now the total momentum going into the horizon can be seen as the sum of all 4-vectors p^μ of the in- and outgoing particles, as it is usually done in field theory. The operators $U_{\rm in \atop out}$ are put in a Lorentz vector x^μ without sign changes:

$$x^+ = U_{\rm in} , \qquad x^- = U_{\rm out} . \tag{8.12}$$

Substituting

$$P_{\rm in}(\tilde{x}) = \sum_{i=1}^{N_{\rm in}} p_{\rm in}^{-,i}\delta^2(\tilde{x}-\tilde{x}^i) , \qquad P_{\rm out}(\tilde{x}) = -\sum_{j=N_{\rm in}+1}^{N} p_{\rm out}^{+,j}\delta^2(\tilde{x}-\tilde{x}^j) , \tag{8.13}$$

one obtains

$$\langle{\rm out}|{\rm in}\rangle = \mathcal{N}\int\mathcal{D}x^+(\tilde{x})\int\mathcal{D}x^-(\tilde{x}) \tag{8.14}$$

$$\exp\big[i\int {\rm d}^2\tilde{x}\{-\tfrac{1}{2}\tilde{\partial}x^\mu(\tilde{x})\tilde{\partial}x^\mu(\tilde{x})\} + i\sum_{i=1}^{N} p^{\mu,i}x^\mu(\tilde{x}^i)\big] .$$

Here, the transverse components of x^μ are not functionally integrated over; they are the transverse coordinates. The factor $\frac{1}{2}$ compensates for double counting. The contribution of the transverse components of x^μ to the integrand must be subtracted, which corresponds to a renormalization of $\mathcal{N}$.

It is, however, more realistic to put the external particles in wave functions that are eigenstates of momenta only. Therefore, we must convolute this expression by transverse wave functions $e^{i\tilde{p}^i\cdot\tilde{x}^i}$, where the transverse components of the momenta, $\tilde{p}^i$, must be kept small compared to the Planck energy (otherwise, it would have been illegal to ignore the transverse gravitational shifts.) We then obtain

$$\langle\text{out}|\text{in}\rangle = \mathcal{N}\Big(\prod_i\int \mathrm{d}^2\tilde{x}^i\Big)\int \mathcal{D}x^+(\tilde{x})\int\mathcal{D}x^-(\tilde{x})$$

$$\exp\Big[\ i\int \mathrm{d}^2\tilde{x}\{-\tfrac{1}{2}\tilde{\partial}x^\mu(\tilde{x})\cdot\tilde{\partial}x^\mu(\tilde{x})\}+i\sum_{i=1}^{N}p^{\mu,i}x^\mu(\tilde{x}^i)\Big]\,, \qquad (8.15)$$

where now the effects of the wave functions are included in the contributions of the external momenta $p^{\mu,i}$ to the 'vertex insertions'. Thus, in contrast to Eq. (8.14), $p^{\mu,i}$ here have transverse components.

It is here that the striking resemblence to string amplitudes should be pointed out. We have the string integrand (for closed strings), as well as the integration over moduli space, which here is formed by the points $\tilde{x}^i$ where the particles cross the horizon. The fact that the action is linearized is understandable, since all transverse dimensions have been kept large compared to the longitudinal ones. What is more surprising is the value of the string constant: it is equal to i, in units where $8\pi G = 1$.

The way in which here the black hole horizon is identified with a string worldsheet is sketched in Fig. 4. At $t\to-\infty$ we have ingoing closed "strings". Arriving at the horizon these strings exchange a string, whose world sheet wraps around the horizon exactly once. The edges of the holes left behind are the outgoing closed strings.

At this point, let us once again focus on the nature of the Hilbert space of in- and outgoing particles. Suppose that, for simplicity, we discretize the transverse coordinates $\tilde{x}$. The functional integrals then become finite-dimensional. What distinguishes this space from the usual Fock space is now, that at every point $\tilde{x}$ *exactly one* "particle" is allowed. The only way to mimick the usual Fock space is to assume that every elementary point particle must be given a *different* value for its transverse coordinate $\tilde{x}$. This constraint may be considered to be negligible, if the $\tilde{x}$ are sufficiently fine-grained, but it is somewhat puzzling how to maintain this constraint in an infinite-volume limit. Apparently, unlike ordinary Fock space, a state with two or more particles at the *same* transverse position $\tilde{x}$, with momenta $p^{\mu,1},\ \dots\ ,\ p^{\mu,k}$, is indistinguishable from the state with just a single particle there, whose momentum is $\sum_{i=1}^{k}p^{\mu,i}$. This may seem to be odd, but it should be noted that this situation is identical to what one has in string theory, where the integrand for a many-particle amplitude is identical to the

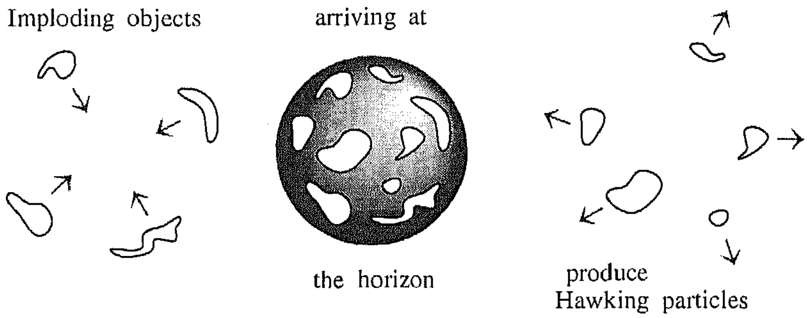

Fig. 4 The horizon as a string world sheet. Three snapshots of a collision event with a black hole intermediate state.

amplitude for fewer particles, when two or more of the vertex insertions happen to coincide in the string world sheet.

The operators $x^\mu(\tilde{x})$ may be regarded as an "average position" operator for *all* particles ever entering or leaving the horizon at that point. This may (partly) explain why this information never "disappears" behind the horizon: there are always sufficient numbers of particles to be seen outside. This is in fact guaranteed by our brick wall model: the number of particles at distance greater than h from the horizon are sufficient to represent all "information" concerning the state of the black hole.

There is a special interpretation for the commutation rule

$$[U_{\rm out}(\tilde{x}),\, U_{\rm in}(\tilde{x}')] \;=\; \int {\rm d}^2\tilde{x}'' f(\tilde{x}-\tilde{x}'')[P_{\rm in}(\tilde{x}''),\, U_{\rm in}(\tilde{x}')] \;=\; -if(\tilde{x}-\tilde{x}')\,. \tag{8.16}$$

We could decide to interpret $-U_{\rm out}(\tilde{x})$ as indicating the position of the horizon with respect to the particles seen to emerge from the black hole, and similarly, $-U_{\rm in}(\tilde{x})$ as the *time reverse* of this: the position of the past horizon with respect to the ingoing particles. Eq. (8.16) implies an uncertainty relation for these two quantities. For ordinary black holes, $U_{\rm out}(\tilde{x})$ is usually precisely defined, as it is determined by the momentum distribution of the ingoing particles that actually formed the black hole. $U_{\rm in}(\tilde{x})$ is the horizon of the time-reversed, or "white hole". In our picture, the white hole is the object formed by the Hawking particles if we follow these backwards in time. This is usually spread quantum mechanically over a large range of values. In our view, white holes are nothing but quantum superpositions of black holes. They relate to black holes just like the momentum and the position of a quantum particle are related to each other.

9 The Transverse Gravitational Force; A Discrete Spectrum of States[12]

So-far, the transversal component of the gravitational force has not been taken into account. One may suspect that this is the reason why our algebra is still represented by a continuum of states.

Unfortunately, as we will show, including the transverse gravitational force is difficult. We here only give an indication as to how one could proceed along these lines, so as to further improve our theory.

Let us recapitulate our algebra. From Section 8:

$$[P_{\rm in}(\tilde x),\, P_{\rm in}(\tilde y)] \;=\; 0 \;=\; [P_{\rm out}(\tilde x),\, P_{\rm out}(\tilde y)]\;; \tag{9.1}$$

$$[U_{\rm in}(\tilde x),\, U_{\rm in}(\tilde y)] \;=\; 0 \;=\; [U_{\rm out}(\tilde x),\, U_{\rm out}(\tilde y)]\;; \tag{9.2}$$

$$[P_{\rm in}(\tilde x),\, U_{\rm in}(\tilde y)] \;=\; -i\delta^2(\tilde x - \tilde y\;) \;=\; [P_{\rm out}(\tilde x),\, U_{\rm out}(\tilde y)]\;; \tag{9.3}$$

$$P_{\rm out}(\tilde x) \;=\; \tilde\partial^2 U_{\rm in}(\tilde x)\;;\quad P_{\rm in}(\tilde x) \;=\; -\tilde\partial^2 U_{\rm out}(\tilde x)\;; \tag{9.4}$$

$$[U_{\rm in}(\tilde x),\, U_{\rm out}(\tilde y)] = if(\tilde x - \tilde y)\;;\quad [P_{\rm in}(\tilde x),\, P_{\rm out}(\tilde y)] \;=\; -i\tilde\partial^2\delta^2(\tilde x - \tilde y)\,. \tag{9.5}$$

As our starting point we again use Eqs. (9.1) – (9.5), but assume these to be valid only when the functions $U_{\rm in}(\tilde x)$ and $U_{\rm out}(\tilde x)$ are slowly varying. For later convenience, we rename the transverse coordinates on the horizon as $(\sigma^1,\, \sigma^2)$, and now define a 2-surface $x^\mu(\tilde\sigma)$ embedded in 4-space:

$$x^+ \;=\; U_{\rm in}\;,\qquad x^- \;=\; U_{\rm out}\;,\qquad \tilde x \;=\; \tilde\sigma\;. \tag{9.6}$$

The orientation of the surface is given by the tensor

$$W^{\mu\nu}(\tilde\sigma) \;=\; -W^{\nu\mu}(\tilde\sigma) \;=\; \varepsilon^{ab}\,\frac{\partial x^\mu}{\partial\sigma^a}\,\frac{\partial x^\nu}{\partial\sigma^b}\,. \tag{9.7}$$

We have

$$\frac{\partial\tilde x^a}{\partial\sigma^b} \;=\; \delta^a_b\,. \tag{9.8}$$

Now first consider the case that $x^\pm$ are slowly varying. This implies

$$\begin{aligned} W^{12} \;&=\; 1\;; \qquad & W^{1\pm} \;&=\; \frac{\partial x^\pm}{\partial\sigma^2}\;; \\ W^{2\pm} \;&=\; -\frac{\partial x^\pm}{\partial\sigma^1}\;; \qquad & W^{+-} \;&=\; \mathcal{O}(\partial_\sigma x^\pm)^2\;. \end{aligned} \tag{9.9}$$

Commutation rules follow from Eqs. (9.2) and (9.5):

$$\begin{aligned} [W^{1+}(\tilde\sigma),\, W^{2-}(\tilde\sigma')] \;&=\; [W^{2+}(\tilde\sigma),\, W^{1-}(\tilde\sigma')] \;=\; i\partial_1\partial_2 f(\tilde\sigma - \tilde\sigma')\,; \\ [W^{1+}(\tilde\sigma),\, W^{1-}(\tilde\sigma')] \;&=\; -i\frac{\partial^2}{\partial{\sigma^2}^2}f(\tilde\sigma - \tilde\sigma')\;; \\ [W^{2+}(\tilde\sigma),\, W^{2-}(\tilde\sigma')] \;&=\; -i\frac{\partial^2}{\partial{\sigma^1}^2}f(\tilde\sigma - \tilde\sigma')\;. \end{aligned} \tag{9.10}$$

As a special case, we have

$$[W^{\mu+}(\tilde\sigma),\, W^{\mu-}(\tilde\sigma')] \;=\; -i\tilde\partial^2 f(\tilde\sigma-\tilde\sigma') \;=\; i\,\delta^2(\tilde\sigma-\tilde\sigma')\,, \tag{9.11}$$

where the index μ is summed over. It is this equation that we can reformulate in a manifestly Lorentz covariant form. One then may hope that not only the longitudinal, but also the shifts in all other directions will have been accommodated for. Since according to Eq. (9.9), W^{12} is the dominating component of the tensor $W^{\mu\nu}$, one may rewrite the right hand side of Eq. (9.11) as

$$\varepsilon^{+-12}W_{12}(\tilde\sigma)\delta^2(\tilde\sigma-\tilde\sigma') \;\approx\; \tfrac{1}{2}\varepsilon^{+-\mu\nu}W_{\mu\nu}(\tilde\sigma)\delta^2(\tilde\sigma-\tilde\sigma')\,, \tag{9.12}$$

with $\varepsilon^{+-12} \;=\; i\varepsilon^{3412} \;=\; i$. The covariant generalization is then:

$$[W^{\mu\alpha}(\tilde\sigma),\, W^{\mu\beta}(\tilde\sigma')] \;=\; \tfrac{1}{2}\delta^2(\tilde\sigma-\tilde\sigma')\varepsilon^{\alpha\beta\mu\nu}W^{\mu\nu}(\tilde\sigma)\,. \tag{9.13}$$

This equation, as well as (9.11), is invariant under all continuous reparametrizations of the $\tilde\sigma$ coordinates (note that $W^{\mu\nu}$, as defined by Eq. (9.7), transforms as a density.)

It is tempting to assume Eq. (9.13) to have a wider range of validity than the non-covariant Eqs. (9.1) – (9.10). After all, Lorentz invariance guarantees that Eq. (9.13) continues to hold when the derivatives of $x^{\pm}(\tilde x)$ are arbitrarily large. Unfortunately, the equations (9.11) do not form a closed algebra, since at the left hand side the index μ is still summed over. One can, however, limit oneself to the self-dual part:

$$K^{\mu\nu} \;=\; i(W^{\mu\nu} + \tfrac{1}{2}\varepsilon^{\mu\nu\kappa\lambda}W^{\kappa\lambda})\,, \tag{9.14}$$

which has only three independent components:

$$K_1 \;=\; i(W^{23}+W^{14})\;;\quad K_2 \;=\; i(W^{31}+W^{24})\;;\quad K_3 \;=\; i(W^{12}+W^{34})\,, \tag{9.15}$$

and, indeed, their algebra closes. From Eq. (9.14) we derive:

$$[K_a(\tilde\sigma),\, K_b(\tilde\sigma')] \;=\; i\varepsilon_{abc}K_c(\tilde\sigma)\delta^2(\tilde\sigma-\tilde\sigma')\,. \tag{9.16}$$

The operators $K_a(\tilde\sigma)$ are distributions. In order to construct representations of the algebra (9.16), we introduce test functions $f(\tilde\sigma)$, $g(\tilde\sigma)$, and write

$$L_a^{(f)} \quad \stackrel{\text{def}}{=} \quad \int K_a(\tilde\sigma)f(\tilde\sigma)\mathrm{d}^2\tilde\sigma\,, \tag{9.17}$$

$$\left[L_a^{(f)},\, L_b^{(g)}\right] \;=\; i\varepsilon_{abc}L_c^{(fg)}\,. \tag{9.18}$$

Restricting oneself to test functions f with $f^2=f$, which only take the values 0 or 1, we find that the operators $L_a^{(f)}$ obey the commutation rules of the angular momenta:

$$[L_a^{(f)}, L_b^{(f)}] = i\varepsilon_{abc}L_c^{(f)} . \tag{9.19}$$

Notice that, since these operators L_a are obtained by integrating K_a over the region(s) where $f = 1$, and because the definition of K_a can be traced back to Eq. (9.7), one can rewrite $L_a^{(f)}$ as a contour integral:

$$L_1^{(f)} = i \oint_{\delta f} (x^2 \mathrm{d}x^3 + x^1 \mathrm{d}x^4) , \qquad \text{etc.} , \tag{9.20}$$

where δf stands for the boundary of the support of f.

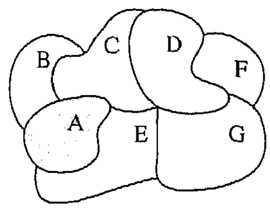

Fig. 5 Domains on the horizon corresponding to a representation of the algebra (9.16).

Suppose now that we have a set of test functions f which are equal to 1 on domains A or B, etc., and zero elsewhere. The domains form a lattice (of our choice) on the horizon, see Fig. 5. In each domain we have a set of three operators L_a that commute as angular momentum operators. The states could be formed out of the $|\ell, m\rangle$ eigenstates of $\mathbf{L}^2$ and L_3. If we combine domains to form some larger domain, the corresponding angular momentum operators must be added to form the new $\mathbf{L}$ operators, by the use of Clebsh-Gordan coefficients. Actually, if any of the ℓ values is larger than the minimal value $\frac{1}{2}$ (or perhaps, in some cases, 1), one can imagine splitting the corresponding domain into smaller ones with each the ℓ value $\frac{1}{2}$. Thus, one may end up with a lattice where on each site one has $m = \pm\frac{1}{2}$. It would not make much sense to maintain domains which have $\mathbf{L} = 0$, because the vanishing of the integrals (9.20) would imply that these regions have no spatial extent.

At first sight, this looks like a complete resolution of our problems. If each domain could be attributed an area equal to $4G \ln 2$ (see Eq. (4.10)), we exactly reproduce Eq. (4.9) for the level density. Unfortunately, life is not so simple. In Eq. (9.15), $W^{i4} = iW^{i0}$ are antihermitean operators, but W^{12}, W^{23} and W^{31} are hermitean. Therefore, the hermitean conjugates of K_a, and those of L_a, are the *anti*self-dual parts of $W^{\mu\nu}$. The L_a operators are not hermitean, and therefore the ℓ and m quantum numbers need not be subjected to the usual constraints of being half-integer, nor to obey the usual inequalities $|m| \leq \ell$.

10 Black Hole Complementarity

Let us return to the argument at the end of Section 9 concerning the notion of causality[14]. It has often been raised as a point of criticism against our scattering matrix Ansatz, see Fig. 6. An observer A passes through an horizon, while also an onserver B detects Hawking radiation. If we were allowed to treat them as living in a space-time that is fixed by external conditions, these two observers could be considered to be spacelike separated, and therefore one could conclude that their measurement operators commute. Hilbert space can be factored into a space of states whose properties can be detected by A, and another space of states whose properties can be detected by B, and possible further factors that can be seen neither by A nor by B. If, however, this space were considered to be the horizon of a black hole, one would expect the states seen by A to be related to the states seen by B through an S-matrix, and hence no longer independent. For the black hole physicist, there is no contradiction. Any measurement made by B, implies the introduction of states obtained from the Hartle-Hawking!vacuum by acting on it with operators that create or remove particles seen by B, which for A would by outrageously energetic. These particles would cause gravitational shifts that seriously affect the ingoing objects, including the fragile detectors used by A. Thus, these observations cannot be independent. What is new here, even for any possible flat space-time observer, is that trans-Planckian particles are involved (with this term we mean particles whose energies are far beyond the Planck value). In short: the metric is *not* determined solely by external circumstances, but also by the particles under consideration.

Complementarity here stands for the idea that states in Hilbert space near a black hole will appear to be profoundly different when the ingoing observer compares them with the outside one. General coordinate transformations fail to relate the experiences of the Hawking observer to the ones of the ongoing observer. Nevertheless, we talk about the same Hilbert space of states. In the limit of the infinite size black hole, this implies that a mapping should exist between the states living in a flat background metric, and black hole states that have all their information mapped onto the horizon. Unfortunately, a completely coherent physical picture clarifying this situation is still lacking.

Apparently, new phenomena strongly affect the conventional form of quantum mechanical Hilbert space when trans-Planckian particles enter the scene. With trans-Planckian particles around, spacelike separated operators may no longer commute with each other.

11 Outlook

Even though the philosophy, adhered to in this paper, is completely straightforward, and should not present fundamental conceptual problems, it nevertheless turned out to be extremely difficult to implement it completely. The effects of

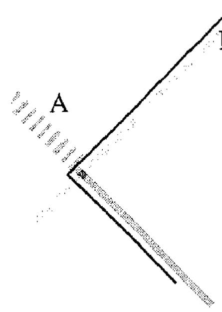

Fig. 6 The ingoing observer (A), and the Hawking observer (B).

transverse gravitational shifts were hard to implement, since these shifts do not commute with the longitudinal ones (because of their $\tilde{x}$-dependence). We have not mentioned another difficulty: the mass shell conditions for the in- and outgoing particles. We took these to be essentially massless, but most particles have a lower bound for their masses. Transverse momenta and masses, however, cause outgoing particles to fall back in again. The difficulty connected to this is the fact that, close to the horizon, ingoing and outgoing states will be difficult to distinguish. Presumably, the splitting of S according to $S = S_{\text{out}} S_{\text{hor}} S_{\text{in}}$ (Eq. 7.1), must be further refined.

The resemblance to string theory in our final results may suggest that one should readdress the black hole using string theory. Some caution however is called for. It is well-known that string theory requires a 10 or 26 dimensional target space, if tachyons and other unphysical features are to be avoided, but such arguments do not directly apply to our present aproach: unitarity and causality look very different, as is manifest from the observation that our string constant is purely imaginary. Secondly, by considering the "information content" of the states in our Hilbert space, we infer that a cut-off at the Planck scale is required that turns our world into a discrete one at that scale. This is quite unlike the starting points of string theory. Convergence of the various approaches may well be envisioned, but it is conceivable that two-dimensional conformal quantum field theory is no more (or less) relevant here than it is in certain statistical models such as the Ising model. We should keep in mind that QCD is also a theory that shows stringlike behavior, but clearly lives in four space-time dimensions, so that apparently the formal unitarity arguments are not applicable here.

The observation of black hole–white hole complementarity (Section 8) suggests an interesting relationship between the *black hole horizon* and the *white hole singularity*, and vice versa. After all, a white hole singularity would develop as soon as one allows Hawking particles to produce a gravitational field, as one would be tempted to do when contemplating time reversal invariance. Indeed, the point S in Fig. 2b, is not truly a point, but gets the extended shape of a caustic when ingoing matter is deprived of its spherical symmetry. The operator $U_{\rm out}(\tilde{x})$ could be regarded as the one describing this caustic. When ingoing matter is allowed to enter during sufficiently large time intervals, this caustic becomes a true fractal. At the same time, this fractal may be relavant for the description of the singularity in the time-reversed black hole. A duality relationship between the black hole singularity and the horizon has been proposed in the framework of string theory.

Bibliography

[1] S.W. Hawking, Commun. Math. Phys. **43** (1975) 199; J.B. Hartle and S.W. Hawking, Phys. Rev. **D13** (1976) 2188.

[2] S.W. Hawking, Phys. Rev. **D14** (1976) 2460; *id.* Commun. Math. Phys. **87** (1982) 395; S.W. Hawking and R. Laflamme, Phys. Lett. **B209** (1988) 39; D.N. Page, Phys. Rev. Lett. **44** (1980) 301;
R. Haag, H. Narnhofer and U. Stein, Commun. Math. Phys. **94** (1984) 219;
R. Sorkin, Phys. Rev. Lett. 56 (1986) 1885, Phys. Rev. D **34** (1986) 373;
P. Mitra, *Black hole Entropy*, Invited talk delivered at XVIII IAGRG Conference, Madras, February 1996, hep-th/9603184.

[3] S.W. Hawking and G.F.R. Ellis, *The Large Scale Structure of Space-time*, Cambridge: Cambridge Univ. Press, 1973.

[4] W. Rindler, Am.J. Phys. **34** (1966) 1174.

[5] G. 't Hooft, Int. J. Mod. Phys. **A11** (1996) 4623.

[6] R.M. Wald, Commun. Math. Phys. **45** (1975); *id.*, Phys. Rev. D **20** (1979) 1271; W.G. Unruh and R.M. Wald, Phys. Rev. D **29** (1984) 1047; W.G. Unruh, Phys. Rev. D **14** (1976) 870;
S.W. Hawking and G. Gibbons, Phys. Rev. D **15** (1977) 2738.

[7] G. 't Hooft, Acta Physica Polonica **B19** (1988) 187-202.

[8] J.D. Bekenstein, Nuovo Cim. Lett. **4** (1972) 737; *id.*, Phys. Rev. D **7** (1973) 2333, D **9** (1974) 3292.

[9] G. 't Hooft, Nucl. Phys. **B256** (1985) 727-745.

[10] P.C. Aichelburg and R.U. Sexl, Gen. Rel. and Gravitation **2** (1971) 303; W.B. Bonner, Commun. Math. Phys. **13** (1969) 163.

[11] T. Dray and G. 't Hooft, Nucl. Phys. **B253** (1985) 173.

[12] G. 't Hooft, Proc. of the 4th seminar on Quantum Gravity, May 25-29, 1987, Moscow, USSR. Eds. M.A. Markov, V.A. Berezin and V.P. Frolov, World Scientific 1988, p. 551; *id.*, "Black holes as clues to the problem of quantizing gravity", Lecture notes for the CCAST/WL Meeting on *fields, strings and quantum gravity*, Beijing, June 1989. (Gordon and Breach, London); *id.* Nucl. Phys. **B335** (1990) 138; Physica Scripta T **36** (1991) 247. *id.* "Scattering matrix for a quantized black hole", In book: *Black Hole Physics*, V. De Sabbata and Z. Zhang (eds.). 1992 Kluwer Academic Publishers, The Netherlands, p. 381; *id.* "S-Matrix theory for black holes", In Proceedings of a NATO Advanced Study Institute on *New Symmetry Principles in Quantum Field Theory*, held July 16-27, 1991 in Cargèse, France. Eds. J. Fröhlich, G. 't Hooft, A. Jaffe, G. Mack, P.K. Mitter and R. Stora NATO ASI Series, 1992 Plenum Press, New York, p. 275; *id.* "Black Holes, Hawking Radiation and the Information Paradox", Proceedings of *Trends in Astroparticle Physics*, Stockholm, Sweden, Sept. 22-25, 1994, THU-94/20.

[13] G. 't Hooft, Physica Scripta, **T15** (1987) 143-150.

[14] C.R. Stephens, G. 't Hooft and B.F. Whiting, Class. Quantum Grav. **11** (1994) 621.

Part VII

Panel Discussion

Panel Discussion: The Definitive Proofs of the Existence of Black Holes

Werner Collmar (Garching), Norbert Straumann (Zürich), Sandip K. Chakrabarti (Calcutta), Gerard 't Hooft (Utrecht), Edward Seidel (Potsdam), Werner Israel (Victoria)

It is our intention to give the non-expert reader a book at hand which enables him or her to recognize whether there are black holes around or not. In the various lectures of our school, the lecturers tried to address this question from theoretical as well as from observational points of view. In the panel discussion some of our lecturers were asked to sum up the present state of knowledge in the form of relatively short statements. They were explicitely requested to answer the fundamental question "Do black holes exist or not?". We hope that you will enjoy the output of our panel as much as we did.

The Editors

Werner Collmar:

As an astrophysical observer I want to summarize briefly (my knowledge on) the observational hints/evidence which we have for the existence of Black Holes (BHs) in outer space.

The most conclusive observations would be if we could directly image the astrophysical sites suspected to host a BH, like the center of active galaxies or of galactic binaries, for example. Active galactic nuclei (AGNs) are believed to contain massive BHs at their center. The Schwarzschild radius R_S of a $10^8\,M_\odot$ BH is 3×10^8 km or $\sim$2 AU (astronomical units). If a surrounding accretion disk extends out to $\sim$500 R_S, its angular size at a distance of 10 Mpc is $\sim 2\times10^{-4}$ arcsec, which is well below the angular resolution of the Hubble Space Telescope. Since no resolved direct imaging is possible, the observational evidence has to come from 'indirect' means.

In AGNs large luminosities are generated in small volumes, like the total luminosity of our galaxy would be generated within our solar system, and would be switched on and off on timescales of days. This extraordinary fact is not explainable with nuclear fusion as the source of energy. However, it is most naturally explained by the release of gravitational energy close to a massive BH, which is the most efficient process we know.

The effects of the gravitational force of BHs on the surrounding stars or gas can be studied. From velocity and velocity-dispersion measurements near the center of galaxies, the enclosed mass can be estimated. The most promising cases for BHs are found in our 'Milky Way' with $\sim 2.45\times10^6\,M_\odot$ within 0.015 pc [2] and in NGC 4258 from VLBI (Very Long Baseline Interferometry maser observations with a central mass of $3.6\times10^7\, M_\odot$ within a volume of 0.13 pc [8]. Massive star clusters in such small volumes would not be stable for long times and therefore have to collapse to BHs. X-ray observations have revealed an asymmetric and redshifted iron K_α line in the Seyfert galaxy MCG-6-30-15 [10]. The shape and the redshift are most easily explained if the line is generated in a plasma which is located within the strong gravitational field close to a massive BH.

If a BH is a member of a binary system, its mass can be estimated (via the mass function) by the dynamics of the system. Several compact sources have lower limits on their masses which are significantly larger than $3\, M_\odot$, the upper limit on the stability of a neutron star. The most promising stellar candidate to date is the X-ray nova V404 Cyg with an estimated mass for the compact object in the range between 10 and 15 $M_\odot$ [9].

Different spectral signatures in compact binaries have been proposed to be evidence for a BH as compact object instead of a neutron star, like spectral bumps at γ-ray energies or power-law spectra up to ~1 MeV instead of spectra cutting off exponentially at hard X-rays. Because these interpretations are model dependent, their support for the evidence of BHs is less compelling than the mass estimates, for example.

In the future new instruments from the radio to the γ-ray band will come into operations with improved sensitivity and improved spectral and angular resolution. With these instruments we shall have better diagnostic tools to further investigate such promising topics like this X-ray line in MCG-60-30-15, for example. To my mind however, even these improved instruments will not provide a 'single key' observation definitely proving the existence of BHs. I rather believe that — as time progresses — the overall evidence for BHs will rise asymptotically to one.

Norbert Straumann:

The evidence of black holes in some X-ray binary systems and for supermassive black holes in galactic centers is still indirect, but has become overwhelming during the past few years. There is so far, however, very little evidence that these collapsed objects are described by the Kerr metric.

In my brief remarks I would like to tell those of you, who had no closer look at the observational situation, what I consider in both categories to be at present the very best candidates. I will indicate only one prospect of observing specific signatures associated with the Kerr metric. Others will have to say more on this.

Independent of the remarkable recent developments, I had never any serious doubts that black holes exist in great numbers in the astronomical universe. This is simply because 'cold' self-gravitating matter can only exist in a small mass range below a few solar masses. There is, on the other hand, absolutely no reason for all the many massive stars (or associations of stars) to get rid of their mass in the course of their evolution in some violent processes (supernovae), in order to be able to settle down in a cold final state (white dwarf, neutron star, quark star (?), ...).

The value of the largest possible mass, $M_{\rm max}$, of a neutron star plays a decisive role in the observational identification of stellar-mass black holes. In view of the large uncertainties of the equation of state, it is important of having reliable limits of $M_{\rm max}$. This is possible on the basis of general assumptions. For instance, if one accepts a fluid dynamic description of matter in a neutron star, causality implies that the velocity of sound should not exceed the velocity of light. (It is true that the sound velocity is only a phase velocity, but it provides the characteristics of the hydrodynamic equations for acceptable relativistic formulations.) For non-rotating stars, an upper bound of about $3.2 M_\odot$ is obtained. On the basis of even weaker assumptions (microscopic stability, $\mathrm{d}p/\mathrm{d}\rho \geq 0$) one finds $M_{\rm max} < 5 M_\odot$ instead. Rotating neutron stars can have somewhat larger masses, see [6].

In binary X-ray systems the mass functions f can be measured with good accuracy. Since this provides a rigorous lower limit for the mass of the compact companion, the case for a black hole is extremely strong if f is bigger than $6 M_\odot$, say. Now, this is the case for the *X-ray nova V404 Cygni*, which has the mass function

$$f = (6.08 \pm 0.06)\ M_\odot\,.$$

Other good cases have been mentioned in some of the lectures, but this is, as far as I can see, the most secure stellar-mass black hole candidate we have at the moment.

As far as supermassive black holes are concerned, I would like to emphasize that gas-dynamical evidence is not very strong in general, because gas can easily be pushed around. Therefore, the many examples of central gas and dust disks perpendicular to jets, like M87, do not establish the existence of black holes.

Thanks to the recent work of Genzel and coworkers [2, 3, 7], we now know that a dark mass of about $2.6 \times 10^6 M_\odot$ must reside within about a light week in Sgr A*. Its density is thus greater than $2 \times 10^{12} M_\odot/\mathrm{pc}^3$. There exist no stable configurations of normal stars, stellar remnants of sub-stellar entities at that density. This concentration of dark matter *in the center of the Milky Way* is now the best case for a supermassive black hole.

Another, almost equally good case is provided by NGC 4258. The following observations show that the central mass must be a black hole (or something even more exotic). The peculiar spiral galaxy NGC 4258 (distance ~ 6.5 Mpc) has in its core a disk extending from about 0.2 to 0.13 pc in radius. This was discovered with a very precise mapping of gas motions via the 1.3 cm maser-emission of H_2O with VLBI (Very Long Baseline Interferometry). The angular resolution of the array was better than 0.5 milliarc seconds. (This is 100 times better than the resolution of the HST (Hubble Space Telescope).) The rotational velocity distribution in the disk follows an exact Keplerian law around a compact dark mass, and the velocity of the inner edge is 1080 km/s. From this one infers a dark mass concentration of $3.6 \times 10^7\ M_\odot$. As in the case of the Milky Way, there are no long-lived star clusters with these extremal properties.

Finally, I would like to point out the possibility to study the relativistic region of a black hole with X-ray astronomy. Recent ASCA (Advanced Satellite for Cosmology and Astrophysics) observations of MCG-6-30-15 and other Seyfert 1 galaxies have revealed that the 6.4 keV fluorescent iron line is very broad and provides some evidence that the emitting region is orbiting close to a black hole. Gravitational effects are apparently skewing the line shape [10, 4]. With the greatly improved spectral resolution of XMM (X-ray Multi-mirror Mission), this might become an important tool to obtain specific information on the gravitational field in the relativistic X-ray emitting region. J. Wilms et al. have made detailed studies, see their contribution in Chap. 5, see also B.C. Bromley et al. [1].

Sandip K. Chakrabarti:

The problems at hand are (a) whether black holes exist, (b) whether the compact objects we call 'black holes' are actually those predicted by solutions of Einstein's equations and (c) what is the best way to identify black holes.

Since black holes are necessarily black (save Hawking radiations which are unobservable with present technology), their detections must be indirect. However, some of the evidences are more 'circumstantial' than others. For instance, the origin, acceleration and collimation of jets, fast time variabilities, stellar velocity dispersions etc. are far from convincing proofs. Measurements of the mass of the central objects such as in M87 ($4 \times 10^9 M_\odot$) by HST spectroscopy or in NGC4258 ($4 \times 10^7 M_\odot$) by water maser emission or in our Galactic Center ($2.5 \times 10^6 M_\odot$) by proper motion study of stars produce mass concentrations of $2.0 \times 10^7 M_\odot/pc^3$, $2.5 \times 10^9 M_\odot/pc^3$, and $6.5 \times 10^9 M_\odot/pc^3$ respectively. Similarly, the measurements of mass function of binary systems indicated a mass function of V404 Cyg to be $f(M) = (6.08 \pm 0.06)\ M_\odot$, a few other candidates have $f(M) \gtrsim 3.0\, M_\odot$. These objects may be strong candidates for black holes but the nagging issue still remains: Are they black holes as predicted by Einstein's equations or just some compact object with hitherto unknown equation of state? I.e., do these objects have horizons with all the associated properties?

In order to prove the existence of the horizons one must look for detailed spectral signatures, since radiations forming the spectra come out of infalling matter which respects the inner boundary condition (IBC) on the horizon. As is known, the IBC selects completely different hydrodynamic and radiative hydrodynamic branches of the global solutions of the governing equations. Unfortunately, till today not all the equations could be written down with certainty in black hole environment, what to talk about solving them. However, some of the predictions of the existing advective models are sufficiently robust and do not depend on the detailed models. For instance, the hard/soft transitions can be seen even in neutron stars, but almost *constancy of spectral slopes* with change of luminosity by factors of hundreds and particularly the photon index of -2.5 in weak power-law tail of soft states are *not seen* in neutron stars. Difference of neutron stars and black holes cannot be made in terms of total luminosity (even if through black hole horizons the entire energy can be advected and in principle makes the flow non-luminous) since any number of physical processes (such as outflows) in neutron stars may invalidate the argument completely.

With the recognition that we must rely on spectral signatures for a complete stock taking of the black holes in the universe, it has become easier to identify black holes. For instance, Cyg X-1, which is the first suspected candidate for a black hole does not satisfy the mass function criteria at all, its $f(M)$ is only $(0.24 \pm 0.01)\ M_\odot$. But it shows the desired spectral slopes in both hard and soft candidates. Similarly, V404 Cygni and A0620-00 have large mass functions and also behave understandably in quiescence states as predicted by advective disk models. Supermassive black holes are difficult to be identified using spectral signatures, because their change of state would take thousands of years.

Gerard 't Hooft:

For an elementary particle physicist, the question of the existence of black holes has several aspects.

- From a philosophical point of view, we will never be able to prove the existence of anything with ultimate rigour. We cannot even prove that we exist ourselves. Clearly, what we are trying to do is to provide proofs "beyond all reasonable doubts".
- In quantum particle physics, it is of importance to know whether some object can exist in principle, even if its actual presence somewhere cannot be shown. A particle that can exist in principle, will represent an element in Hilbert space, and as such it will give important contributions to the Schrödinger equation. We call such particles *virtual* particles, and we want to know about them.
- This is why we want to know whether black holes with masses between 1 mg and 1 $M_\odot$ can exist, even if no astrophysical mechanism for their production was known. Furthermore, the importance of the existence, in the above sense, of these solutions of Einstein's equations is that they appear to be correct, acceptable solutions. If they did *not* exist, even in principle, this would imply a significant deficiency in our understanding of Nature. Not only the physics of the fundamental interactions would have to be revised, but even the laws at distance scales larger than centimeters, whereas these have been checked by many precision experiments.
- It is conceivable, however, that tiny black holes will turn out to be indistinguishable from more conventional forms of matter: the spectrum of "black hole states" might blend naturally into the spectrum of "elementary particles" (loosely speaking, black holes *are* elementary particles) and *vice versa*. In this case, large black holes (heavier than, say, a few milligrams) will still be so significantly different from other forms of matter that in practice confusion will be unlikely.
- For me, *astronomical* black holes represent the extreme other end of the scale. I am impressed by the evidence produced by the astrophysicists, and I believe

that they have come extremely close to proving the existence of black holes.

- There is the question of the existence of *primordial* black holes, in particular the ones with masses in the range of planetoids. These should decay, and quantum theory predicts that once their masses have decreased to become that of a small asteroid, they turn themselves into radiation energy through a violent explosion. Whether these really exist in our universe (not just virtually) and whether these explosions can be or have already been observed, remains extremely dubious, at best.
- I would like to know from the astrophysicists whether three-body interactions in globular star clusters can produce black holes, which should come shooting out of the center, singlets as well as doublets (P. Hut, private communication). I would also like to know what the *smallest mass* is for a black hole that can be produced via conventional astrophysical processes.

Edward Seidel:

Evidence for the existence of black holes, once considered fantasy by many, is mounting these days on a monthly basis. Most astronomers now accept black holes as a standard part of their observed universe, while only a decade ago many were rather skeptical. As standard astronomical observations improve, more and more black hole candidates are found, while existing candidates are even more firmly thought to be actual black holes.

However, while such evidence is now very strong indeed, it is generally circumstantial. We still await direct and incontrovertible proof of the existence of black holes, i.e., a detectable signal emitted by a black hole that unambiguously identifies it. *Gravitational waves emitted by black hole interactions* should provide that "smoking gun" signal that not only proves that black holes exist, but that will also reveal essential properties of the black holes. In particular, the so-called ringing or quasinormal modes of the black hole are damped wavetrains that excited rather generically from a perturbed black hole, and the wavelength and damping time of these modes depending only on its mass and spin (and charge, which is probably not relevant). In all numerical studies of black hole formation, perturbation, and even collision, these modes have been strongly excited, and if detected they should uniquely identify the source as a black hole.

As gravitational wave observatories such as LIGO, VIRGO, and GEO600, are under construction around the world, the gravitational wave signals expected from black hole interactions is gaining increasing attention. Black hole mergers from binary coalescence is now considered one of the most promising sources of gravitational waves, and according to a recent study by Flanagan and Hughes [5], such systems may well provide the first gravitational waves to be seen after the detectors go online in a few years. In what follows I summarize the main points of their very detailed analysis.

There are three phases of black hole coalescence: (a) inspiral, (b) merger, and (c) ringdown. The inspiral is the adiabatic orbit phase where the holes slowly spiral together, before they are around $r = 6M$ apart, where M is the total mass of the system. At this time, the merger phase begins, where dynamic instabilities drive the holes towards each other in near free fall. This is a most violent stage of evolution for which numerical relativity will be essential to compute waveforms. Finally, a single, distorted black hole results, which will quickly settle in to the ringdown phase, where normal modes are emitted. According to Flanagan and Hughes, taking account of the sensitivity and bandwidth of the first generation detectors coming online by the year 2001, low mass coalescence (< 30 solar masses) should be seen via their inspiral waves, while high mass systems (between 100 and 700 solar masses) should be seen via their ringdown waves. Both types should be visible out to about 200 Mpc, and the event rate for the low mass systems should be of order several per year. By using numerical relativity calculations as templates for the merger phase, one can enhance the detection rate, and very significantly, one can analyze the signals to better understand the sources.

This is a very exciting time for black hole physics: gravitational wave astronomy is about to be born, and fortunately numerical relativity is on the track of simulations that will be needed to enhance and understand the upcoming observations. The next decade should provide many proofs of the existence of black holes, and a plethora of information about them as well.

Werner Israel:

Twenty years ago I spent a very pleasant sabbatical year at the ... Institute in ... . Although this story is true, I have blanked out the names because its significance is generic – it could have happened at any of dozens of institutions at that time. Shortly after my arrival there was a coffee party, and after some warm words of welcome, the Director of the Institute remarked, "Werner is going to be with us for a year. We should all talk to him and try to cure him of these silly notions he has about the possibility of black holes."

I was very well treated and enjoyed the most cordial personal relations with my hosts throughout that year, but in this one respect at least I'm afraid I proved something of a disappointment to them. The cure failed. As my wife will, I'm sure, readily confirm, I am the most unreasonably mulish of people, also fairly deaf and impervious to foreign languages.

I recall going to lunch one day with about a dozen members of the Institute – staff and graduate assistants – and I decided to take a poll, asking each in turn whether he believed that black holes are possible. There was a string of firm denials, except for one assistant, from whom I got the interesting response: "Not usually. But when I'm applying for a U.S. research grant, *then* I believe – temporarily".

I am still in touch with a few of the colleagues who gathered around the table that day. Not one has changed his views.

With the foreshortening that memory lends to past events, it now seems that in December 1967, John Wheeler declared, "Let there be black holes". and lo! there were black holes. But this is not how it really happened. Comprehension of this idea took years, acceptance much longer. I remember Felix Pirani's words in 1967: "The Schwarzschild singularity at $r = 2m$ is almost certainly a fraud – but we don't yet know what kind of fraud".

There was an incident in Canada in 1979 when the Winnipeg vice squad, raiding a video store, seized copies of (among others) the Walt Disney movie, "The Black Hole", on suspicion of obscenity – probably the only time a Disney film has ever suffered this indignity.

Of course, the disbelief of my Institute colleagues had nothing to do with observations. They were put off by the inherent absurdity of the *concept* – the inevitability of singularities, space turning into time, and so on. At that time, observational progress was lethargic..– It took ten years before the first serious black hole candidate Cygnus X-1, optically identified in 1972 by Tom Bolton of Toronto and by Webster and Murdin in England was joined by a second, LMC X-3, identified by a team at the Dominion Astrophysical Observatory, Victoria, B.C.

But, as we have heard at this School, in the last few years the observational arena has been transformed by the Hubble space telescope and sophisticated ground-based techniques. We now have compelling evidence for the presence of compact dark objects in galactic nuclei with masses ranging from millions to billions times that of the sun.

The most convincing and extraordinary case is NGC 4258. Here one is observing (by interferometric techniques in the microwave range) maser radiation from water molecules in a dusty torus orbiting just 0.3 light-year from the centre, indicating the presence of an invisible mass of 36 million solar masses within this radius.

In our own galaxy, the recent work of Eckart and Genzel, who measured proper motions of three dozen stars within 1/30 light-year from the centre, point convincingly to a central dark mass of 2 million suns.

At the 1996 "Texas" symposium in Chicago, there were even reports of the first direct observational evidence for general-relativistic strong-field effects and for event horizons. The Japanese X-ray satellite ASCA has found peculiar frequency shifts in the X-rays from galactic nuclei, which (it is claimed) can only be caused by strong gravity. Richard Mushotzky of the Goddard Space Institute reported on observations of X-ray novae during intervals when accretion is slow. Those thought (on the basis of mass) to contain black holes are dimmer than ones with neutron stars, presumably reflecting the difference between the hard surface of a neutron star and an event horizon.

Despite all of these developments, many who rejected black holes in the 1970s as conceptually absurd remain hard-core skeptics today. The moral I should like to draw is this. Nothing in science is 100 % certain. But a time arrives when the balance of probabilities has become so overwhelming that the burden of proof must pass to the opposition, and further attempts at rational argument are pointless. We *still* have flat-earth societies; 87 % of U.S. parents want creationism taught in schools; and, as for me, I still check my horoscope in the newspaper from time to time. Today, if one encounters a skeptic for whom non-existence of black holes is a matter of religious faith, perhaps it is kinder (and certainly less trouble) not to get involved in an argument.

Bibliography

[1] B.C. Bromley, K. Chen, W.A. Miller, ApJ **475** (1997) 57

[2] A. Eckart and R. Genzel, Nature **383** (1996) 415

[3] A. Eckart and R. Genzel, MNRAS **284** (1997) 576

[4] A.C. Fabian, K. Nandra, C.S. Reynolds, W.N. Brandt, C. Otani, Y. Tanaka, H. Inoue, K. Iwasawa, Los Alamos eprint archive astro-ph/9507061

[5] E.E. Flanagan and S.A. Hughes, Phys.Rev. **D57** (1998) 4535

[6] J.L. Friedman in: *Black Holes and Relativistic Stars. A Symposium in Honor of S. Chandrasekhar*, Ed. R. Wald (University of Chicago Press, Chicago 1998)

[7] R. Genzel, A. Eckart, T. Ott and F. Eisenhauer, MNRAS **291** (1997) 219

[8] L.J. Greenhill et al., ApJ **440** (1995) 619

[9] T. Shahbaz et al., MNRAS **271** (1994) 10

[10] Y. Tanaka et al., *Nature* **375** (1995) 659