RECENT DEVELOPMENTS IN QUANTUM FIELD THEORY
J. Ambjørn, B.J. Durhuus & J.L. Petersen (editors)

BLACK HOLES AND QUANTUM MECHANICS

G. 't HOOFT

Institute for Theoretical Physics, University of Utrecht, P.O. Box 80.006, 3508 TA Utrecht, The Netherlands

ABSTRACT
The equivalence principle in general relativity may have a non-standard form when quantum effects are considered. A theory that may produce the complete spectrum of black holes is outlined.

1. INTRODUCTION

Many attempts are being made to formulate the laws of physics at the Planck length scale. Canonical quantization procedures when applied to Einstein's Lagrangian, reveal fundamental and seemingly uncontrollable space-time fluctuations at distance scales less than the Planck length[1]. It was natural that investigators turned their attention to sophisticated models of gravitation and matter in which the infinities in these fluctuations might cancel out such as "supergravity"[2], "string"[3] and now also "superstring"[4] theories. As yet these theories seem to give relatively little insight in the structure of space-time itself at the Planck length scale.

Various suggestions have been made that space-time might show no structure at all beyond the Planck length[5]. But what does a discrete - and curved - space-time look like? Which constraints should we impose on such numerous models to select out the physically viable? Which, if any, of our familiar concepts - continuity of space-time being just one of them - can still be used?

Gravitating systems are fundamentally unstable against collapse. Classically this is not a great problem: only for very large systems the gravitational force is stronger than the counter forces produced by matter. But in a quantum theory, with huge oscillations near the Planck length the possibility of gravitational collapse cannot be ignored. What we propose is that a healthy theory should not only take into account collapsed chunks of matter but must more likely contain them as essential ingredients. Perhaps all particles can in some sense be viewed upon as smaller or larger black holes.

Unavoidably our theory must exhibit a "smallest possible length scale": the smallest possible structure is a particle whose Schwarzschild radius coincides with its Compton wave length. We now notice a situation that reminds one of the familiar "bootstrap" idea; all particles much lighter than the Planck mass are likely to be described reasonably accurately by some Lagrange field theory. All particles much heavier than the Planck mass are black holes with fairly large radii. Their behavior also should follow from field equations - the same Lagrange field theory - with these larger length scales. It is this form of "duality" that interests us: it gives us the impression that quantum gravity should be a completely understandable, finite, problem.

But how do the quantum properties of black holes follow from Lagrange field theory? One comfortable result was derived by Hawking[6]: due to vacuum fluctuations near the horizon all heavy black holes must emit particles spontaneously, with a thermal spectrum corresponding to a temperature $T = 1/8\pi M$, where M is the mass of the black hole in natural units. Apparently like most other fundamental particles, black holes are unstable and decay into lighter objects.

This result is extremely powerful since it suggests that no additive conservation law can be exact, with the exception of electric charge conservation, because no chemical potential for quantities such as baryon- or lepton-number can be accepted. But unfortunately the obtained expressions only produce emission probabilities, not the quantum mechanical amplitudes. The quantum states are represented by density matrices. So it seems that the information produced by this argument is only statistical in nature. Suppose we had a precisely defined Lagrange field theory. Could we then not do better than this?

If the black hole were an ordinary soliton the answer would have been "yes". We would have been able to do calculations such as the ones by Rubakov and Callan[7] on magnetic monopoles. But black holes are not ordinary solitons and some fundamental and tantalizing difficulties prevent us from applying conventional laws of quantum mechanics.

Hawking had derived his result by relying heavily on the equivalence principle of general relativity: states in Hilbert space were assumed to be well-defined in any coordinate system and their inner products were all assumed to be coordinate independent. The difficulty is then that "states" seem to disappear into the horizon of the black hole and in spite of them being all orthogonal to each other they become fundamentally unobservable. One line of thought, as proposed by Hawking, is that pure quantum mechanics is no longer valid at Planck length scales: pure states may undergo transitions towards mixed quantummechanical states: the eigenvalues of the density matrix may no longer be constants of motion. This is an extremely important conclusion because it seems to be practically unavoidable whereas it also seems to imply the breakdown of quantum mechanics as we know it at the Planck scale[8].

But how sure are we of the equivalence principle for states in Hilbert space? Could it not be that a coordinate transformation has more subtle effects on Hilbert space if the corresponding observers from a certain moment on can no longer communicate with each other? What if one observer falls right into the system studied by another observer? What is the probability interpretation of a wave function if an observer has a finite chance to become killed by a space-like singularity?

Of course what we need foremost is a mathematically unique prescription for obtaining the laws of physics for every imaginable system. This "theory" should as much as possible reproduce all known results of ordinary quantum mechanics on the one hand and general relativity on the other. We will be quite content if this "theory" is first formulated in a coordinate-invariant way and then allows us to construct a Hamiltonian suitable to describe anything seen by any observer. But this construction might be dependent on the observer and in particular his "horizon". It could even be that the "probabilities" experienced by one observer are not the same as those of another. All is well if the two "classical limits" are as they should be.

We will now make the assumption that the black hole quantum properties[9] somehow follow from Lagrange quantum field theory at the same length scale. We are very well aware of the risk that this may be wrong. Still, we like to know how far one can get. Regrettably, the results to be reported in this paper will be extremely modest.

We will start by making a simplification that caused some confusion for some readers of my previous publication: we first concentrate on the steady state black hole: every now and then something falls in and something else comes out. *Nowhere a distinction is made between "primordial" black holes and black holes that have been formed by collapse*. It has been argued that Hawking's derivation in particular holds for collapsed black holes and not necessarily for ones eternally in equilibrium. However if we succeed to describe infalling things in a satisfactory way then one might expect that inclusion of the entire collapse (and the entire evaporation in the end) can naturally be incorporated at a later stage. Our main concern at present will be time scales of order $M \log M$ in Planck units, which is much shorter than the black hole's history. As we will see, understanding in- and outgoing things at this scale will be difficult enough, and indeed Hawking's radiation can very well be understood at this time scale.

2. KRUSKAL COORDINATES. BLACK HOLE AT EQUILIBRIUM

In the absence of matter, the metric of a black hole is

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\Omega^2 \; . \tag{2.1}$$

The Kruskal coordinates u, v are defined by

$$uv = \left(1 - \frac{r}{2M}\right)e^{r/2M} , \tag{2.2}$$

$$v/u = -e^{t/2M} , \tag{2.3}$$

and then we have

$$ds^2 = - \frac{32M^3}{r} e^{-r/2M} dudv + r^2 d\Omega^2 , \tag{2.4}$$

which is now entirely regular at $r > o$. However (2.2) and (2.3) admit two solutions at every (r,t): we have two universes connected by a "wormhole". The Schwarzschild region, I, is $r > o$, $u < o$. The other regions are indicated in Fig. 1.

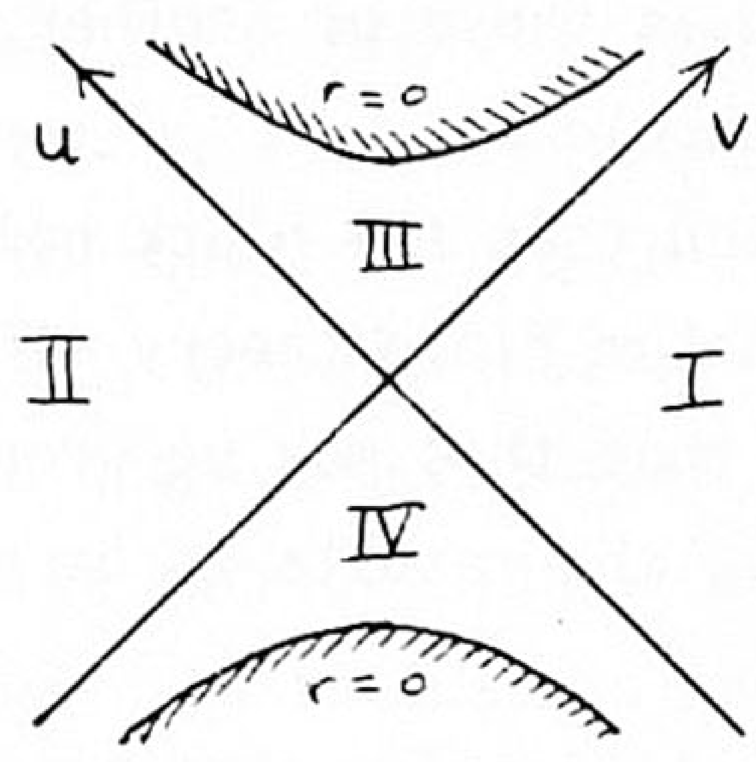

Figure 1

Now the classical picture of a black hole formed by collapse only shows regions I and III, the others being shielded by the imploding matter which accumulates at the past horizon (the u-axis). Similarly, an evaporating black hole (sometimes called a "white hole") only has regions I and IV. In both cases it is convenient to extend analytically the particle content in regions III or IV towards region II, and a black hole in equilibrium is perhaps best described by the entire system I-II-III-IV.

The equivalence theorem should now relate the Hilbert space as needed by an observer in the wormhole ("Kruskal observer") to the one needed to describe the "physical" world I as experienced by an outside observer ("Schwarzschild observer"). Imagine a limited number of soft particles that can be described by the Kruskal observer using standard physics. With "soft" we mean that the energies of these particles are so small that gravitational effects on the metric can be neglected. We have then a reasonable description of an important part of the Hilbert space for the wormhole observer. The evolution of this system is described by an Hamiltonian

$$H = \int \mathcal{H}(\vec{x})d\vec{x} > o , \tag{2.5}$$

with one ground state

$$H|o\rangle_k = o \tag{2.6}$$

where k stands for Kruskal. Due to curvature this vacuum is not exactly but only approximately conserved. H describes the evolution in the time coordinate $\tau = u+v$

Now the outside observer uses t as his time coordinate, and a generator of a boost in t produces

$$\delta v = \frac{v}{2M}\delta t\ , \tag{2.7}$$

$$\delta u = -\frac{u}{2M}\delta t\ , \tag{2.8}$$

so the generator of this boost is

$$h = \frac{1}{2M}\int d\vec{x}\ \rho H(\vec{x}) \quad ; \quad \rho = v-u\ . \tag{2.9}$$

We split $h = H_I - H_{II}$:

$$H_I = \frac{1}{2M}\int \rho H(\vec{x})d\vec{x}\ \theta(\rho) \quad ; \quad H_{II} = \frac{1}{2M}\int |\rho|\ H(\vec{x})d\vec{x}\ \theta(-\rho)\ . \tag{2.10}$$

We have

$$[H_I, H_{II}] = o\ , \tag{2.11}$$

and we can write the eigenstates of H_I and H_{II} as $|n,m\rangle$ with

$$H_I|n,m\rangle = n|n,m\rangle \quad ; \quad H_{II}|n,m\rangle = m|n,m\rangle\ . \tag{2.12}$$

Extensive but straightforward calculations show that the "Kruskal vacuum" $|o\rangle_k$ does not coincide with the "Schwarzschild vacuum" $|o,o\rangle$, but instead, we have

$$|o\rangle_k = C\sum_n |n,n\rangle e^{-4\pi Mn}\ , \tag{2.13}$$

where C is a normalization factor. Note that we do have

$$h|o\rangle_k = o\ , \tag{2.14}$$

which is due to Lorentz-invariance of $|o\rangle_k$.

If we consider the equivalence theorem in its usual form and consider all those particles that are trapped into region IV as lost and therefore unobservable then without any doubt the correct prescription for describing the observations of observers in I is to average over the unseen particles. Let $\mathcal{O}$ be an operator built from a field $\phi(\vec{x},t)$ with $\vec{x}$ in region I, then

$$[\mathcal{O},H_{II}] = 0 \quad , \tag{2.15}$$

$$\mathcal{O}|n,m\rangle = \sum_k \mathcal{O}_{nk}|k,m\rangle \ , \tag{2.16}$$

and

$$\langle\mathcal{O}\rangle = {}_k\langle o|\mathcal{O}|o\rangle_k = C^2 \sum_{n,n'} e^{-4\pi M(n+n')} \langle n',n'|\mathcal{O}|n,n\rangle = C^2 \sum_n e^{-8\pi Mn} \mathcal{O}_{nn} \ . \tag{2.17}$$

We recognize a Boltzmann factor $e^{-\beta n}$ with $\beta = 8\pi M$, corresponding to a temperature

$$T = 1/8\pi M \ . \tag{2.18}$$

This is Hawking's result in a nutshell. Black holes radiate and the temperature of their thermal radiation is given by (2.18). The only way in which the horizon entered in this calculation is where it acts as a shutter making part of Hilbert space invisible.

As stated in the introduction this result would imply that black holes are profoundly different from elementary particles: they turn pure quantum mechanical states into mixed, thermal, states. Our only hope for a more complete quantum mechanical picture where black holes also show pure transitions, that in principle allow for some effective Hamiltonian is to reformulate the equivalence principle. Let us assume that the location of the horizon has a more profound effect on the interpretation that one should give to a wave function.

A pair of horizons (the u- and the v-axis in Fig. 1) always separate regions where a boost in t goes in opposite directions with respect to a regular time coordinate such as u+v. As before[10] we speculate that these regions act directly as the spaces of bra states and ket states, respectively. Any "state" as described by a Kruskal observer actually looks like the product of a bra and a ket state to the Schwarzschild observer. More precisely, it looks like an element of his density matrix, ρ:

$$|n,m\rangle \rightarrow |n\rangle\langle m| = \rho \ . \tag{2.19}$$

Just like any density matrix its evolution is given by the commutator with H_I:

$$\frac{d}{dt}\rho_{nm} = -ih|n,m\rangle = -i(n-m)|n,m\rangle = -i[H_I,|n\rangle\langle m|] = -i[H_I,\rho] \, . \qquad (2.20)$$

Now the Kruskal vacuum $|o\rangle_k$ corresponds to the density matrix

$$\rho_{nn'} = C|n\rangle e^{-4\pi Mn}\langle n|\delta_{nn'} \, , \qquad (2.21)$$

which is a thermal state at temperature

$$T = 1/4\pi M \, , \qquad (2.22)$$

twice the usual result. The usual result would require not ρ from eq. (2.19) but $\rho\rho^\dagger$ to be the density matrix, from which of course (2.18) follows.

As long as we consider *stationary black holes with only soft particles* our mapping (2.19) is perfectly acceptable. The Hamiltonian (2.5) may ad libitum be extended to include any kind of interactions including those of curious observers. In the two classical limits we reproduce quantum mechanics and general relativity as required.

The only possible way to settle the question which of the procedures is correct and which of the temperatures (2.18) or (2.22) describe a black hole's radiation spectrum, is to include the effects of "hard" particles. This is also a necessary requirement for understanding the effects of implosion and explosion of black holes. Hard particles are particles whose rest masses may be small, but whose energies are so large that their gravitational effects may not be ignored.

3. HARD PARTICLES

The black holes considered in the previous section were only exactly time-translation-invariant if they were covered by a Kruskal vacuum $|o\rangle_k$. This is because translations in t correspond to Lorentz-transformations at the origin of the Kruskal coordinate frame and only a vacuum can be Lorentz-invariant. Naturally, $|o\rangle_k$ corresponds to a Schwarzschild density matrix ρ which is diagonal in the energy-representation.

Any other state will undergo boosts in t as if the Kruskal observer continuously applies Lorentz-transformations to his state, and eventually any "Soft" particle will turn into a hard particle. This is why hard particles, particles with enormously large Lorentz γ factors are unavoidable if we want to understand how a system evolves over time scales only slightly larger than $\mathcal{O}(M \log M)$. Hard particles alter their surrounding space-time metric. Some basic features of their effects on space-time are now well-known.

A hard particle in Minkowsky space produces a gravitational shock wave[11], sometimes called "impulsive wave", not unlike Cerenkov radiation. Before and behind this shock wave space-time is flat, but the way in which these flat regions are connected at the location of the shock wave produces delta-distributed curvature. Writing

$$u = t-z$$
$$v = t+z \tag{3.1}$$

we find that a particle moving in the positive z direction with momentum p, at $\tilde{y} = o$, produces a shock wave on the v axis where the two half-spaces are connected after a shift

$$\delta v = -4p \, \ell n(\tilde{y}^2) \ . \tag{3.2}$$

Here $\tilde{y}$ is the transverse coordinate. See Fig. 2.

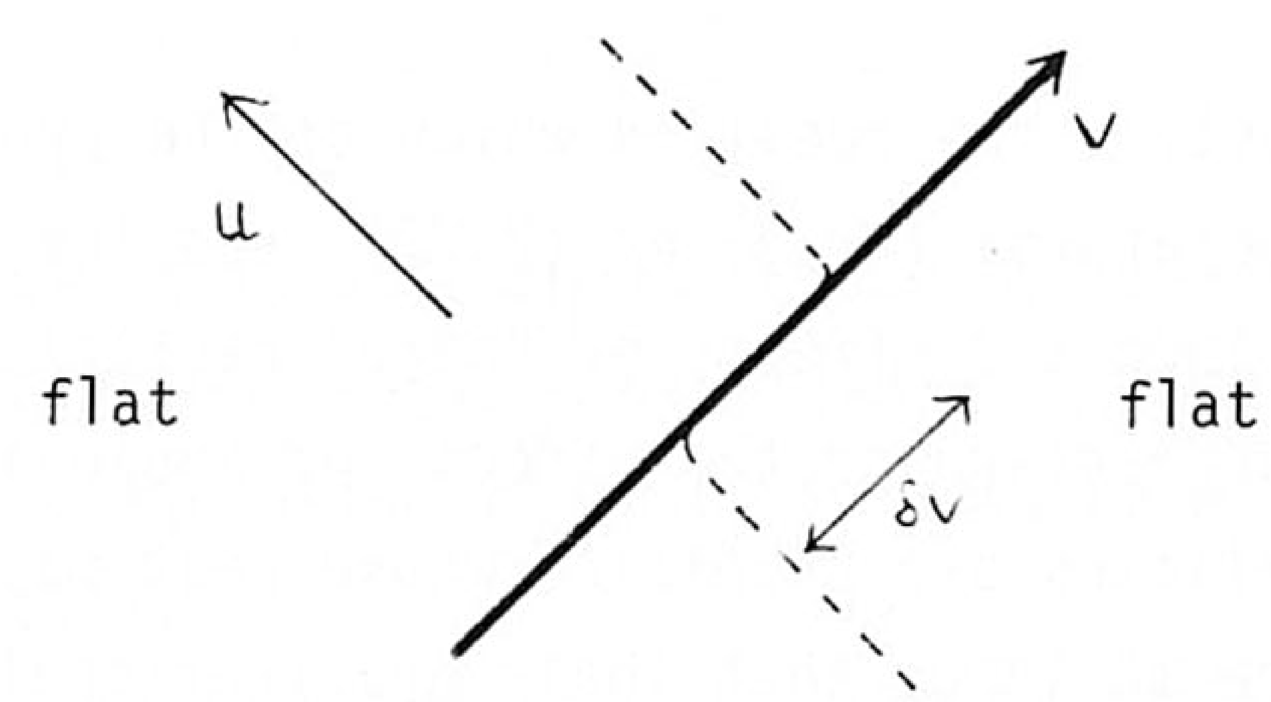

Figure 2

A way to picture this is to choose $g_{\mu\nu} = \eta_{\mu\nu}$ everywhere except at u=o, where all geodesics make a jump δv from past to future.

For us it is interesting to consider now a hard particle on one of the black hole's horizons. It was found that again a displacement of a form similar to (3.2) solves Einstein's equations. In Kruskal's coordinates u,v a hard particle with momentum p again produces a shift δv, with

$$\delta v(\vec{\Omega}) = pf(\vec{\Omega},\vec{\Omega}') \ , \tag{3.3}$$

where Ω' is the angle where the particle goes through the horizon and p its momentum. f is given by

$$\Delta f - f = -2\pi\kappa \ \delta(\theta) \ , \tag{3.4}$$

where θ is the angle between Ω and Ω'; Δ the angular Laplacian and κ a dimensionless numerical constant. The solution to (3.4),

$$f = \kappa \sum_{\ell} \frac{\ell+\frac{1}{2}}{\ell(\ell+1)+1} P_{\ell}(\cos\theta) , \tag{3.5}$$

can be seen to be positive for all θ.

Because of the shift, the causal structure of space-time is slightly changed. The Penrose diagram for a hard particle coming in along the past horizon is given in Fig. 3.

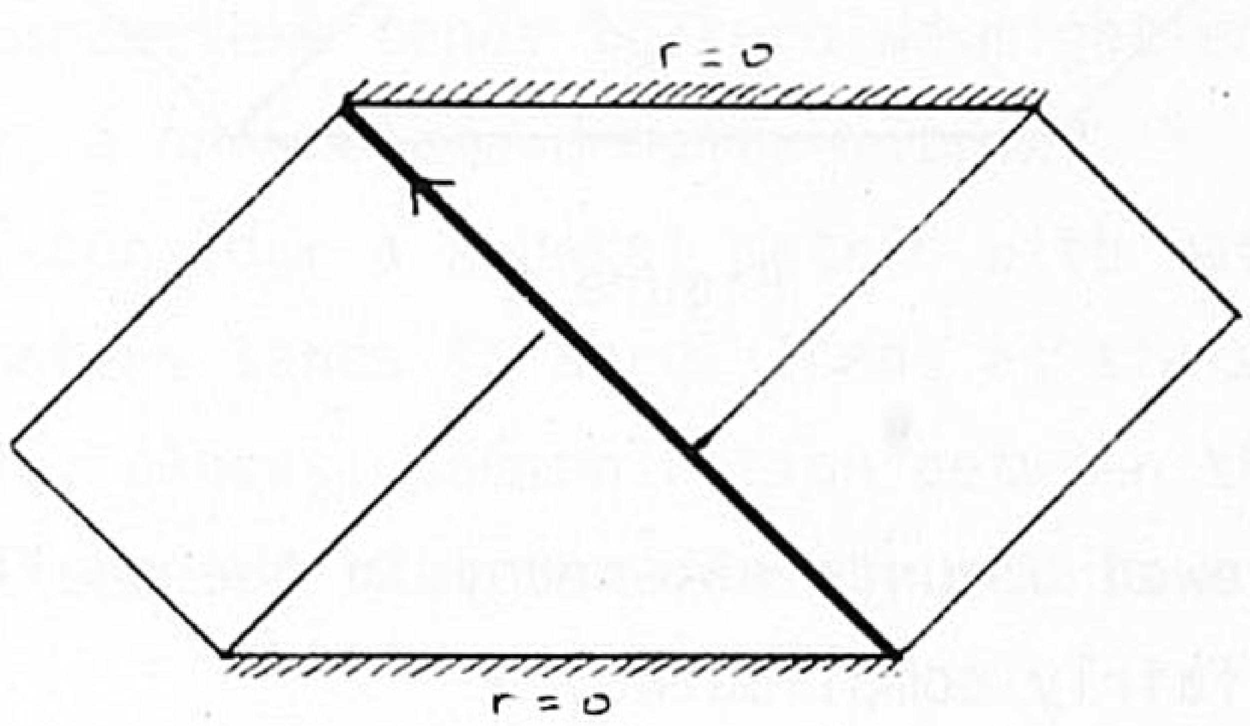

Figure 3

In Fig. 3 the geodesics are defined to go straight through the shock wave but enter into a more or less badly curved metric.

When two hard particles meet each other from opposite directions the curvature due to the resulting gravitational radiation is not easy to describe. We do need some description of this situation and therefore we introduced a simplification by imposing spherical symmetry. Hard particles are now replaced by spherically symmetric hard shells of matter entering or leaving the black hole. We guessed correctly that then Einstein's equations are also solved by connecting shifted Schwarzschild solutions with different mass parameters. The space-time structure of Fig. 4 results.

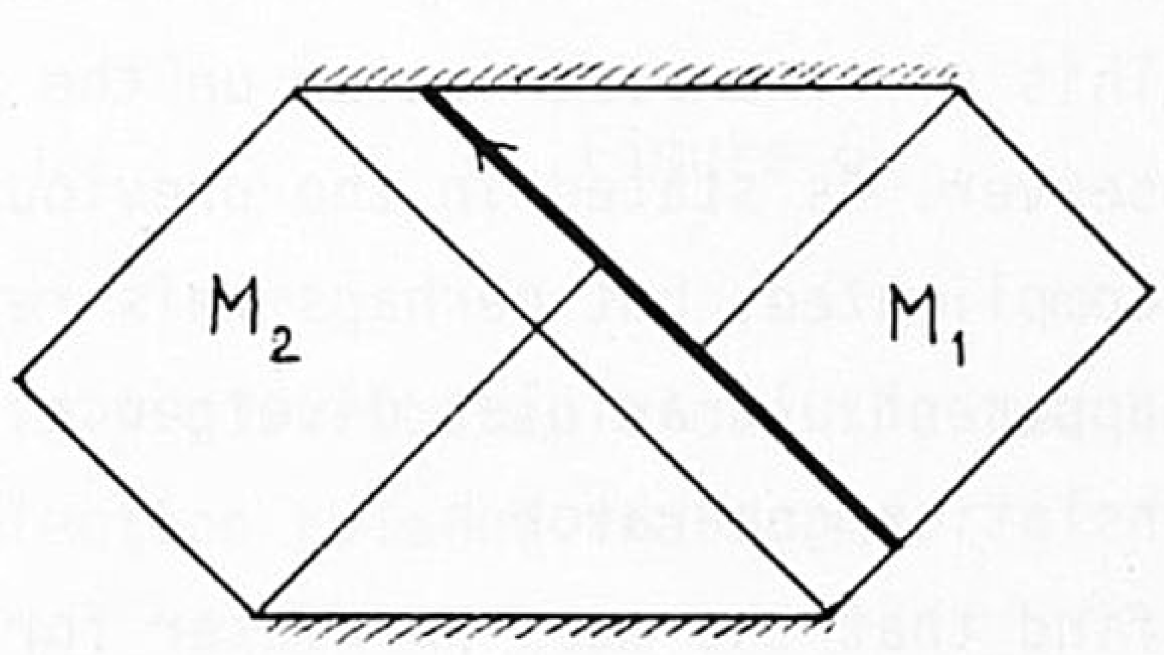

Figure 4

In Fig. 4 matter hits the future singularity at some distance from the past-horizon. In that case $M_1 > M_2$, if we require that the energy content of the shell of matter be positive.

This solution allows us now to combine various shells of ingoing and outgoing matter. One gets the Penrose diagram of Fig. 5.

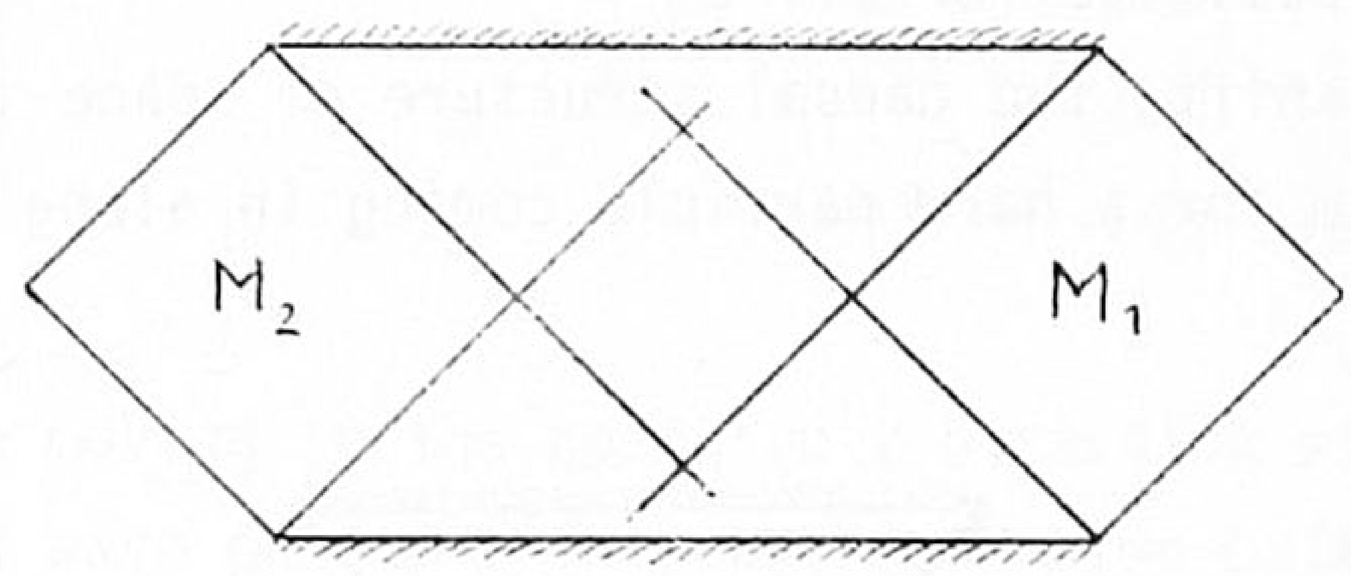

Figure 5

The algebra of the allowed amounts of energy in the shells and the resulting mass parameters M_i is fairly complicated.

An interesting limiting case occurs if one of the internal mass parameters tends to zero. If we require all shell-energies to be positive then such a zero mass region must always connect the future- with the past singularity by an r=o line. This r=o line is the origin of a polar coordinate representation of a flat space and one easily convinces oneself that then no longer any wormhole exists that connects us with another space. Bra- and ket-space are clearly disconnected and indeed we will argue that such a no-bra-space may perhaps be a way to describe a pure state for the Schwarzschild observer.

4. OFF-DIAGONAL DENSITY MATRIX AND PURE STATES - A SPECULATION

It is now reasonable to assume that for a complete description of the Hilbert space for a Schwarzschild observer we need all configurations with hard particles seen by the Kruskal observer. A restriction must be that the metric cannot be distorted so much that any of the Schwarzschild mass parameters become negative. This gives a restriction on the amount of matter acceptable to the Kruskal observer. As stated in the previous sector, the algebra of these requirements is complicated, but perhaps this restriction will be sufficient to cut off an apparent ultraviolet divergence in the spectrum of the Schwarzschild time translation generator h.

In general we will find that the mass-parameter for the black hole in universe I, M_1, differs from M_2 in universe II. We speculate that this could be a direct representation of an off-diagonal density matrix

$$\sum |M_1,\ldots\rangle \langle M_2,\ldots| \ , \tag{4.1}$$

where the sum is over the soft particle states.

Similarly, if one of the masses tends to zero, we get the matrix

$$\sum |M_1,...\rangle \langle \mu \to o,...| \quad (4.2)$$

Now if we are indeed allowed to speculate that light black holes behave more and more like ordinary elementary particles then in the vanishing mass limit black holes may occupy a much smaller number of quantum levels than the heavy ones. Thus, in (4.2) only "a few" bra states contribute. Therefore, perhaps, if one of the mass parameters tends to zero we might end up with a "pure state", or more precisely, a one-column density matrix.

Alternatively, consider a Kruskal metric with matter such that one of the center mass parameters tends to zero. Then, as stated, the wormhole connecting bra- and ket-space, closes. Communication between the two worlds become negligible and we might expect that the density matrix will tend to factorize:

$$\sum |M_1,...\rangle \langle M_2,...| \to (\sum|M_1,...\rangle)\ (\sum\langle M_2,...|)\ . \quad (4.3)$$

It becomes the product of two pure states.

We see that these various considerations converge to a description of pure state black holes: there must be exactly enough matter inside the Kruskal frame such that the wormhole disappears and space-time only keeps one asymptotic region (Fig. 6).

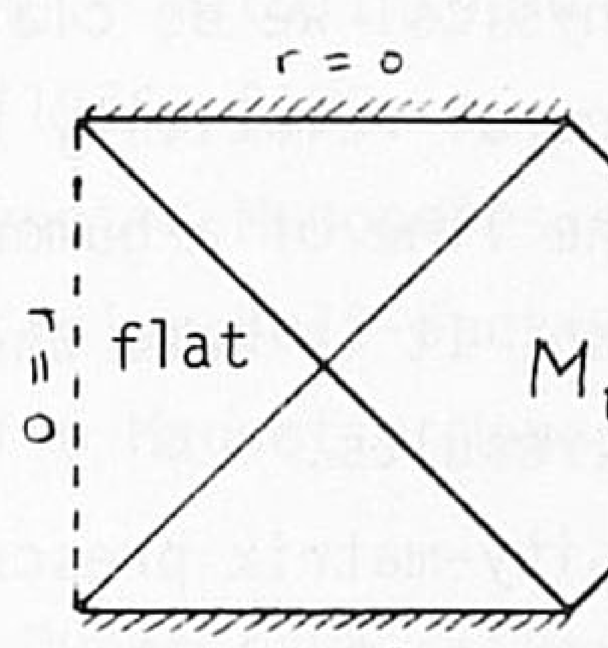

Figure 6

If all matter is mainly distributed along the two horizons this condition corresponds to a selection rule of the form:

$$P^{in}_{tot}\ .\ P^{out}_{tot} = C.M^2\ . \quad (4.4)$$

Even if one does not wish to go along with our density matrix formulation of the equivalence principle, condition (4.4) with Fig. 6 could be an interesting description of the pure state black hole, and it would be important to be able to derive a radiation temperature directly from this picture. Unfortunately our description now lacks any symmetry under time boosts and therefore there seems to be no easy way to describe the near stationary case of a slowly evaporating black hole.

A premature attempt to improve this situation is to realize that any eigenstate of the Hamiltonian with given total energy M must be purely periodic in (Schwarzschild-)time t. Hence only Fourier transforms of states satisfying (4.4), with frequency M, should be allowed:

$$\begin{cases} |\psi\rangle = \int dt\, e^{-iMt} \left| P^{in}_{tot}\, e^{t} ,\, P^{out}_{tot}\, e^{-t} \right\rangle ; \\ P^{in}_{tot} \cdot P^{out}_{tot} = M^2 . \end{cases} \tag{4.5}$$

One of the problems we still have to face is the apparent divergence of this time-integral.

5. DISCUSSION

A major objection against our density matrix theory for a black hole has been put forward by many critics. It usually amounts to saying that the standard calculation yielding Hawking's temperature $1/8\pi M$ is impeccable and only requires known laws of physics.

Now this is absolutely true if the usual equivalence principle is considered to be a known law of physics. We do claim that the equivalence principle has been used - without general relativity there would be no computable Hawking effect. Somewhere in the line of arguments it was necessary to apply transformations across a horizon. It is here where - perhaps - a different procedure might give different results.

If we were to adopt the density-matrix prescription (and even the author himself is far from certain that it should be adopted) then we can imagine where the usual derivation fails. To see the radiation one has to wait long compared to $M \log M$ after the collapse took place. The only stabil matter-metric configuration during such a long time is obtained if from the start the collapsing object were in a mixed state. If the collapsing object started out as a purestate we were forced to use the states of Fig. 6, satisfying (4.5), to describe it. Even before collapse we would have been forced to postulate outgoing matter at the past-horizon.

We suspect that the selection rule (4.4) should be used in the description of the evolution of our pure state beyond times of order $M \log M$ but were unable to implement it.

REFERENCES

1) B.S. DeWitt, "Quantum Theory of Gravity", Phys. Rev. 162 (1967) 1195, 1239.
G. 't Hooft and M. Veltman, "One Loop Divergences in the Theory of Gravitation", Ann. Inst. H. Poincaré 20 (1974) 69.
2) P. van Nieuwenhuizen, "Supergravity", Phys. Rep. 68C (1981) 189.
3) J.H. Schwarz, "Dual Resonance Theory", Phys. Rep. 8c (1973) 269.
S. Mandelstam, "Dual Resonance Models", Phys. Rep. 13c (1974) 259.
J. Scherk, Rev. Mod. Phys. 47 (1975) 123.
4) J.H. Schwarz, "Superstring Theory", Phys. Rep. 89 (1982) 223.
M.B. Green and J.H. Schwarz, "Anomaly Cancellations in Supersymmetric D=10 Gauge Theory and Superstring Theory", Phys. Lett. 149B (1984) 117.
D.J. Gross et al., "Heterotic String", Phys. Rev. Lett. 54 (1985) 502.
5) T. Regge, "General Relativity without coordinates", Nuovo Cim. 19 (1961) 558.
G. 't Hooft, "Quantum Gravity: a Fundamental Problem and Some Radical Ideas", in "Recent Developments in Gravitation", Cargèse 1978, ed. M. Lévy and S. Deser, Plenum Press, New York and London 1979, p. 323.
T.D. Lee, "Difference Equations as the Basis of Fundamental Physical Theories", Columbia Preprint CU-TP-297 (1984); Proceedings of the International School of Subnuclear Physics, Erice 1983.
6) S.W. Hawking, "Particle Creation by Black Holes", Commun. Math. Phys. 43 (1975) 199.
J.B. Hartle and S.W. Hawking, "Path Integral Derivation of Black Hole Radiance", Phys. Rev. D13 (1976) 2188.
7) V. Rubakov, "Superheavy Magnetic Monopoles and the Decay of the Proton", JETP Lett. 33 (1981) 644; "Adler-Bell-Jackiw Anomaly and Fermion Number breaking in the Presence of a Magnetic Monopole", Nucl. Phys. B203 (1982) 311.
C.G. Callan, "Disappearing Dyons", Phys. Rev. D25 (1982) 2141; "Dyon-Fermion Dynamics", Phys. Rev. D26 (1982) 2058; "Monopole Catalysis of Baryon Decay", Nucl. Phys. B212 (1983) 391.
8) S.W. Hawking, "Breakdown of Predictability in gravitational Collapse", Phys. Rev. D14 (1976) 2460.
9) G. 't Hooft, "On the Quantum Structure of a Black Hole", Nucl. Phys. B (1985), to be published.
10) G. 't Hooft, "Ambiguity of the Equivalence Principle and Hawking's Temperature", J. Geom. and Phys. 1 (1984) 45.
11) T. Dray and G. 't Hooft, "The Gravitational Shock Wave of a Massless Particle", Nucl. Phys. B253 (1985) 173.