QUANTUM GRAVITY AND BLACK HOLES

G. 't Hooft

Institute for Theoretical Physics
Princetonplein 5, P.O. Box 80.006
3508 TA Utrecht, The Netherlands

Abstract

After a sketchy resumé of the General Theory of Relativity the problem of quantizing this theory is addressed. The perturbative problem seems to be hardly different from gauge theory models in particle physics, but then some fundamental divergences and instabilities are pointed out. It is argued that "black holes" should be more closely studied in attempting to solve the problems that have arisen. The author then presents a discussion of his views on black holes that has also appeared elsewhere.

1. Curved space-time

There are many excellent treatizes on the subject of General Relativity[1] and the beginning student of this subject is highly advised to consult one or several of those. Only for pedagogical reasons and for the sake of establishing notations we will give here a short resumé of this beautiful theory, in as far as it is needed for our further considerations. A reader who is familiar with the principles of general relativity may decide to skip sections 1 and 2.

For representing points in space-time one uses four coordinates x^μ ($\mu = 0,1,2,3$). Contrary to the case in more conventional physics one will not be able in general to distinguish rectangular from curved coordinates, so any (continuous) representation of points by four numbers is acceptable.

For a pair of neighbouring points x^μ and x'^μ with $x'^\mu = x^\mu + dx^\mu$ (dx^μ being infinitesimal) one defines a distance ds by

$$ds^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu \tag{1.1}$$

where summation over repeated indices is assumed, and $g_{\mu\nu}$ is a symmetric 4×4 matrix that may depend on x^μ:

$$g_{\mu\nu} = g_{\mu\nu}(x) \ . \tag{1.2}$$

In order to make contact with special relativity one assumes that the matrix $g_{\mu\nu}$ has one negative and three positive eigenvalues. Usually (but not always) we will have

$$g_{oo} < 0 \tag{1.3}$$

and identify x^o with time t.

For integrations over space and time an infinitesimal volume element is needed, besides the notion of distance. For this one chooses

$$dV = \sqrt{-g}\, dx^0\, dx^1\, dx^2\, dx^3 \;, \qquad (1.4)$$

where g is the determinant:

$$g = \det(g_{\mu\nu}) \;. \qquad (1.5)$$

In flat space:

$$g_{\mu\nu} = \begin{pmatrix} -1 & & & \emptyset \\ & 1 & & \\ & & 1 & \\ \emptyset & & & 1 \end{pmatrix} \quad ; \quad g = -1 \;. \qquad (1.6)$$

We now need the notion of a scalar field $\varphi(x)$. It might represent any measurable quantity at x but we often use it to describe a spinless particle. The Lagrangian for a free spinless particle in any "curved" space-time may be written as

$$\mathcal{L} = -\sqrt{-g}\left\{g^{\mu\nu}\partial_\mu\phi^*\partial_\nu\phi + m^2\phi^*\phi\right\} \;; \qquad (1.7)$$

where

$$g^{\mu\nu} = (g_{\mu\nu})^{-1} \;. \qquad (1.8)$$

The Euler-Lagrange equation for this field is obtained from the extremal principle

$$\delta \int \mathcal{L}\, d^4x = 0 \;, \qquad (1.9)$$

which should hold for any infinitesimal variation $\varphi(x) \to \varphi(x) + \delta\varphi(x)$. Consequently, the Klein-Gordon equation is now written as

$$\partial_\mu\left[g^{\mu\nu}\sqrt{-g}\,\partial_\nu\phi\right] = m^2\sqrt{-g}\,\phi \;. \qquad (1.10)$$

86

The fundamental theme underlying these equations is invariance or covariance under coordinate transformations of the form

$$x^\mu \Rightarrow u^\mu(x) \quad ; \quad x^\mu = x^\mu(u) \ , \tag{1.11}$$

where the mapping $x \leftrightarrow u$ may be any differentiable function. In the new coordinates the scalar field is $\tilde{\phi}(u)$, with

$$\tilde{\phi}(u) = \phi(x(u)) \ . \tag{1.12}$$

The gradient of a scalar field is a vector field. Its transformation properties under (1.11) can be read off from

$$\phi(x+dx) = \phi(x) + \phi_\mu dx^\mu \ , \tag{1.13}$$

$$\tilde{\phi}(u+du) = \tilde{\phi}(u) + \tilde{\phi}_\mu du^\mu \ , \tag{1.14}$$

$$dx^\mu = \frac{\partial x^\mu}{\partial u^\nu} du^\nu \ , \tag{1.15}$$

so that

$$\tilde{\phi}_\mu(u) = \frac{\partial x^\lambda}{\partial u^\mu} \phi_\lambda(x(u)) \ . \tag{1.16}$$

A "co-vector field" is now any quantity $A_\mu(x)$ that transforms this way,

$$\tilde{A}_\mu(u) = \frac{\partial x^\lambda}{\partial u^\mu} A_\lambda(x(u)) \ . \tag{1.17}$$

A "contra-vector field" $B^\mu(x)$ is defined to transform in such a way that for any vector field $A_\mu(x)$, the quantity $B^\mu(x)A_\mu(x)$ transforms as a scalar. Since the inverse of the matrix $\partial x^\lambda/\partial u^\mu$ is $\partial u^\mu/\partial x^\lambda$ (under the usual definitions of partial integration), we find that B^μ must transform as

$$\tilde{B}^{\mu}(u) = \frac{\partial u^{\mu}}{\partial x^{\lambda}} B^{\lambda}(x(u)) \ . \tag{1.18}$$

Note that co-vectors and contra-vectors are distinguished by placing the index either below or above the field symbol.

Tensor fields $A_{\mu\nu}$ are defined to transform just as products of vector fields:

$$\tilde{A}_{\mu\nu}(u) = \frac{\partial x^{\lambda}}{\partial u^{\mu}} \frac{\partial x^{\kappa}}{\partial u^{\nu}} A_{\lambda\kappa}(x(u)) \ , \tag{1.19}$$

and similarly for fields $B^{\mu\nu}$ and mixed tensors $F_{\mu\nu\ldots}{}^{\alpha\beta\ldots}$.

Now a gradient of a scalar field transforms as a vector. A gradient $\partial_{\mu}A_{\nu}$ of a vector field A_{ν} is *not* a proper tensor. Of course, *products* $A_{\mu}A_{\nu}$ are. In order to obtain correction terms to a gradient such that it does transform as a tensor one introduces the "connection field" $\Gamma^{\alpha}{}_{\beta\gamma}(x)$. In spite of the special choice of upper and lower indices, Γ does not transform as a tensor. Instead, we require

$$\tilde{\Gamma}^{\alpha}{}_{\beta\gamma}(u) = \frac{\partial u^{\alpha}}{\partial x^{\mu}} \left[\frac{\partial x^{\lambda}}{\partial u^{\beta}} \frac{\partial x^{\kappa}}{\partial u^{\gamma}} \Gamma^{\mu}{}_{\lambda\kappa}(x) + \frac{\partial^2 x^{\mu}}{\partial u^{\beta} \partial u^{\gamma}} \right] . \tag{1.20}$$

One may convince oneself that (1.20) is consistent with repeated transformations, and that now

$$D_{\mu}A_{\nu} \overset{=}{\text{def}} \partial_{\mu}A_{\nu} - \Gamma^{\alpha}{}_{\mu\nu}A_{\alpha} \ , \tag{1.21}$$

does transform as a true tensor.

Similarly we have the true tensor

$$D_{\mu}B^{\nu} \overset{=}{\text{def}} \partial_{\mu}B^{\nu} + \Gamma^{\nu}{}_{\mu\alpha}B^{\alpha} \ , \tag{1.22}$$

for a contravector field B^μ. The symbol D_μ is called "covariant derivative". So far, the quantity $\Gamma^\alpha_{\beta\gamma}$ is an arbitrary field. We usually restrict ourselves to the case that it is symmetric in the lower indices β,γ.

Now in spaces with a metric tensor $g_{\mu\nu}$ there is *another* way to define covariant derivatives. Consider any vector field $A_\mu(x)$. How do we produce a tensor $D_\mu A_\nu(x)$?

1. Consider first one point $x = x_{(1)}$.
2. Choose special coordinates u at x close to $x_{(1)}$, such that

$$ds^2 = du^{1^2} + du^{2^2} + du^{3^2} - du^{0^2} + \mathcal{O}\left(du^2\ (x-x_{(1)})^2\right) , \qquad (1.23)$$

   or

$$\tilde{g}_{\mu\nu}(u(x_{(1)})) = \eta_{\mu\nu}$$
$$\partial_\alpha \tilde{g}_{\mu\nu}(u(x_{(1)})) = 0 . \qquad (1.24)$$

3. In these coordinates we define that, at $u = u(x_{(1)})$, $D_\mu A_\nu(u) = \partial_\mu A_\nu(u)$.
4. Transform to any other coordinate system by the transformation rule of true tensors.
5. Do the same for all other points x.

One can convince oneself that this procedure is uniqe. With this definition one always gets

$$D_\mu g_{\alpha\beta} = 0 \quad ; \quad D_\mu g^{\alpha\beta} = 0 \qquad (1.25)$$

($g^{\alpha\beta}$ is defined to be the inverse of the matrix $g_{\alpha\beta}$). Now the previous definition with a Γ field would give

$$D_\mu g_{\alpha\beta} = \partial_\mu g_{\alpha\beta} - \Gamma^\lambda{}_{\mu\alpha} g_{\lambda\beta} - \Gamma^\kappa{}_{\mu\beta} g_{\alpha\kappa} . \qquad (1.26)$$

This vanishes if we choose

$$\Gamma^{\alpha}{}_{\beta\gamma} = \tfrac{1}{2} g^{\alpha\mu}\left(\partial_\beta g_{\mu\gamma} + \partial_\gamma g_{\mu\beta} - \partial_\mu g_{\beta\gamma}\right) . \tag{1.27}$$

Thus, a metric tensor $g_{\mu\nu}$ generates uniquely a symmetric connection field $\Gamma^{\alpha}{}_{\beta\gamma}$.

The Riemann curvature tensor is now defined by looking at the commutator of covariant derivatives:

$$D_\mu D_\nu A_\alpha - D_\nu D_\mu A_\alpha = - R^{\lambda}{}_{\alpha\mu\nu} A_\lambda , \tag{1.28}$$

for all fields A_α. Writing it out in terms of the connection field we get

$$R^{\alpha}{}_{\beta\gamma\delta} = \partial_\gamma \Gamma^{\alpha}{}_{\beta\delta} - \partial_\delta \Gamma^{\alpha}{}_{\beta\gamma} + \Gamma^{\alpha}{}_{\gamma\sigma}\Gamma^{\sigma}{}_{\beta\delta} - \Gamma^{\alpha}{}_{\delta\sigma}\Gamma^{\sigma}{}_{\beta\gamma} . \tag{1.29}$$

We find the following properties:

1. $R^{\alpha}{}_{\beta\gamma\delta}$ is a tensor. This is obvious in eq. (1.28) but not if (1.29) is used as a definition.

2. $$R^{\alpha}{}_{\beta\gamma\delta} = - R^{\alpha}{}_{\beta\delta\gamma} \tag{1.30}$$

3. $$R^{\alpha}{}_{\beta\gamma\delta} + R^{\alpha}{}_{\gamma\delta\beta} + R^{\alpha}{}_{\delta\beta\gamma} = 0 \tag{1.31}$$

4. $$D_\mu R^{\alpha}{}_{\beta\gamma\delta} + D_\gamma R^{\alpha}{}_{\beta\delta\mu} + D_\delta R^{\alpha}{}_{\beta\mu\gamma} = 0 \tag{1.32}$$

5. For a contravector field B^μ we have

$$D_\mu D_\nu B^\alpha - D_\nu D_\mu B^\alpha = R^{\alpha}{}_{\lambda\mu\nu} B^\lambda \tag{1.33}$$

6. If $R^{\alpha}{}_{\beta\gamma\delta}(x)$ is zero in some region then space-time is flat there: there exist coordinates x^μ such that

$$g_{\mu\nu}(x) = \begin{pmatrix} -1 & & & \emptyset \\ & 1 & & \\ & & 1 & \\ \emptyset & & & 1 \end{pmatrix} . \tag{1.34}$$

The Ricci tensor is

$$R_{\mu\nu} = R^{\alpha}{}_{\mu\alpha\nu} = R_{\nu\mu} \tag{1.35}$$

and the scalar curvature

$$R = g^{\mu\nu}R_{\mu\nu} . \tag{1.36}$$

The Ricci tensor satisfies the Bianchi identity

$$g^{\alpha\beta}D_{\alpha}R_{\beta\nu} - \tfrac{1}{2}D_{\nu}R = 0 . \tag{1.37}$$

Defining

$$G_{\mu\nu} = R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu} \tag{1.38}$$

we have

$$g^{\alpha\beta}D_{\alpha}G_{\beta\nu} = 0 . \tag{1.39}$$

2. Einstein's equation

Usually, matter is introduced in the theory by writing down the stress-energy-momentum tensor $T_{\mu\nu}$. We take

$$T_{oo} = \rho = \text{energy (mass-)density.} \tag{2.1}$$

$$-T_{oi} = -T_{io} = P_i = \text{energy flow} = \text{momentum density.} \tag{2.2}$$

$$T_{ij} = t_{ij} = \text{stress tensor.} \tag{2.3}$$

$$\tfrac{1}{3}t_{ii} = p = \text{pressure.} \tag{2.4}$$

The stress tensor t_{ij} describes the force $d\vec{F}$ on a surface element $d\vec{S}$:

$$d\vec{F} = t.d\vec{S} \ . \tag{2.5}$$

We have conservation of energy and momentum:

$$\partial_\mu T_{\mu\nu} = 0 \ . \tag{2.6}$$

If we ignore the effect of a gravitational field on the energy-momentum then we also have

$$g^{\alpha\beta} D_\alpha T_{\beta\nu} = 0 \ . \tag{2.7}$$

The idea is now that (2.7) describes matter only, whereas (2.6) would include the energy and momentum of the gravitational field.

Einstein first guessed

$$R_{\mu\nu} \propto T_{\mu\nu} \ ? \ , \tag{2.8}$$

but it turned out that then (2.7), in combination with (1.37) would give

$$D_\mu R = 0 \quad ; \quad D_\mu T_{\lambda\lambda} = 0 \ ? \tag{2.9}$$

which cannot be true. He then found the correct equation:

$$G_{\mu\nu} = R_{\mu\nu} - \tfrac{1}{2} R g_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{2.10}$$

where G is the gravitational constant:

$$G = 6.672 \ . \ 10^{-8} \ \mathrm{cm}^3 \ \mathrm{g}^{-1} \ \mathrm{sec}^{-2} \ . \tag{2.11}$$

This is the only equation of this sort that is consistent with both (1.39) and (2.7). Eq. (2.10) can be cast into a Lagrange form. Consider

$$\mathcal{L}_1 = \sqrt{-g}\, R \ . \tag{2.12}$$

And consider all variations

$$g_{\mu\nu} \to g_{\mu\nu} + \delta g_{\mu\nu} \ . \tag{2.13}$$

One finds that

$$\delta \int \sqrt{-g}\, R\, d^4x = - \int \sqrt{-g}\, G^{\mu\nu} \, \delta g_{\mu\nu} \, d^4x \ . \tag{2.14}$$

This vanishes if

$$G_{\mu\nu} = 0 \tag{2.15}$$

which is Einstein's equation in the absence of matter.

If we consider the Lagrangian

$$\mathcal{L} = \sqrt{-g}\left(\frac{R}{16\pi G} + \mathcal{L}^{\text{matter}}\right) \tag{2.16}$$

with for instance

$$\mathcal{L}^{\text{matter}} = - g^{\mu\nu}\ \partial_\mu \phi^* \partial_\nu \phi - m^2 \phi^* \varphi \tag{2.17}$$

then one finds, for variations $g_{\mu\nu} \rightarrow g_{\mu\nu} + \delta g_{\mu\nu}$,

$$\delta\left(\int \sqrt{-g}\ \mathcal{L}^{\text{matter}}\ d^4x \right) = \tfrac{1}{2} \int T^{\mu\nu} \sqrt{-g}\ \delta g_{\mu\nu}\ . \tag{2.18}$$

So that

$$\delta \int \mathcal{L}\ d^4x = \int \sqrt{-g}\ \delta g_{\mu\nu} \left(- \frac{G^{\mu\nu}}{16\pi G} + \tfrac{1}{2} T^{\mu\nu} \right) d^4x\ . \tag{2.19}$$

Requiring now that $\delta \int \mathcal{L}\ d^4x$ vanish for all choices of $\delta g_{\mu\nu}$ corresponds to imposing Einstein's equation (2.10).

Having a Lagrangian we can now apply "standard" Hamilton-Lagrange theory to this system, and consider the quantization of gravity.

3. The temporal gauge

General coordinate transformations, eqs. (1.11) - (1.19), must be seen as gauge transformations. In some sense, quantum gravity is a gauge field theory. Like in a gauge theory, we must choose a gauge-fixing procedure. Writing

$$u^\mu = x^\mu + \eta^\mu \tag{3.1}$$

with η^μ infinitesimal, we find that a gauge transformation for the metric tensor can be expressed as

$$g_{\mu\nu} \to g_{\mu\nu} + D_\mu \eta_\nu + D_\nu \eta_\mu \ , \tag{3.2}$$

to be compared with

$$A_\mu \to A_\mu + D_\mu \Lambda \tag{3.3}$$

in a vector gauge theory. Because of this invariance, the physical Hilbert space is "smaller" than what would be suggested by the number of field variables. Conceptually, the best way to characterize the Hilbert space is by choosing the equivalent of the so-called "temporal gauge",

$$A_o = 0 \tag{3.4}$$

in vector theories, namely

$$g_{oo} = 1 \quad ; \quad g_{oi} = 0 \qquad (i=1,2,3) \ . \tag{3.5}$$

In this gauge we have

$$\Gamma^o{}_{oo} = \Gamma^o{}_{oi} = \Gamma^i{}_{oo} = 0 \ . \tag{3.6}$$

Now $\Gamma^i{}_{oo}$ can be identified with the leading components of the gravitational field. Apparently we have coordinates such that it vanishes: we have freely falling coordinates.

As long as we work in perturbation theory, (3.5) can always be imposed. The first part can be realized by choosing an η in (3.2) with a certain condition on $\partial_o\eta_o$. This will essentially fix η_o. The other conditions determine

$$\partial_o\eta_i + \partial_i\eta_o \tag{3.7}$$

and since η_o was already fixed, these conditions determine η_i. But clearly, there are free integration constants.

This means that, after having realized (3.5) by some choice of η, we will still have a subclass of coordinate transformations that will leave (3.5) unaffected. Since now we also have (3.6), the infinitesimal η that generate this class of residual gauge transformations satisfy

$$\partial_o\eta_o = 0 \qquad \rightarrow \qquad \eta_o(\vec{x},t) = \eta_o(\vec{x}) \tag{3.8}$$

and

$$D_o\eta_i + D_i\eta_o = 0 \tag{3.9}$$

which implies

$$\partial_o\eta^i + g^{ij}\partial_j\eta_o = 0\ . \tag{3.10}$$

So that

$$\eta^i(\vec{x},t) = \eta^i(\vec{x}) - \int g^{ij}\partial_j\eta_o(\vec{x})dt\ . \tag{3.11}$$

The $\eta_o(\vec{x})$ and $\eta^i(\vec{x})$ generate local but time-independent symmetry transformations. Due to Noether's theorem they must be associated with conservation laws. These are found by subjecting the Lagrangian in which

(3.5) was substituted, to infinitesimal gauge transformations not satisfying (3.8) - (3.11). The only terms that survive are those due to the fact that δg_{oo} and δg_{oi} were omitted, since otherwise $\delta \int \mathcal{L}\, d^4x$ would vanish identically. So we can also look at the terms generated by δg_{oo} and δg_{oi} only. They give

$$- \delta \int \mathcal{L}\, d^4x = \int \sqrt{-g}\, d^4x \left(T^{oo}-2G^{oo}\right) \partial_o \eta_o$$

$$+ \int \sqrt{-g}\, d^4x \left(T^o{}_k-2G^o{}_k\right)\left(\partial_o \eta^k+g^{kj}\partial_j \eta_o\right) = 0 \,, \tag{3.12}$$

where we have put $16\pi G = 1$ (this convention may seem to be an obvious choice, but differs from another obvious choice by a factor 16π. Apparently "large" numbers such as 16π can grow easily in quantum gravity).

From (3.12) we find the following conservation laws:

$$\partial_o\left(\sqrt{-g}\left(T^o{}_k-2G^o{}_k\right)\right) = 0 \,, \tag{3.13}$$

and

$$\partial_o\left(\sqrt{-g}\; T^{oo}-2G^{oo}\right) + \partial_k\left(\sqrt{-g}\; T^{ok}-2G^{ok}\right)\Big) = 0 \,. \tag{3.14}$$

The oo and oi components of Einstein's equations follow, if we put these "conserved currents" equal to zero. The other (ij) components follow from the Euler-Lagrange equations directly.

Let us compare this situation with an ordinary symmetry such as translation invariance. The associated conservation law is the conservation of momentum. Apparently we must require that the momentum itself is zero: the wave function $|\psi>$ is invariant under translations. In quantum gravity, the wave function $|\psi>$ must be invariant under the transformations generated by the η satisfying (3.8) - (3.11).

The importance of this is the following. Remember that the temporal gauge corresponds to freely falling coordinates. This would be quite cumbersome for describing e.g. the planetary system, or a black hole, because such coordinates would become very chaotic or even singular after a small lapse of time. Fortunately at any moment t we can choose to perform a gauge rotation of the subclass that leaves the gauge condition (3.5) intact. The wave function $|\psi>$ will not change. Thus at every t we can decide to "unwind" the coordinates.

A problem on the other hand is that η_o generates time translations. Apparently the Hamiltonian in this representation should vanish also. But this is not so. The correct attitude is to consider only those η that vanish at spatial infinity. Then $|\psi>$ needs not be invariant under time translations that also affect time at infinity. In short: time is defined by clocks located at spatial infinity. It is crucial to observe that apparently the definition of time involves the boundary conditions at spatial infinity. Without explicitly fixing the boundary conditions at spatial infinity there would not exist such a thing as a quantum-mechanical Hamiltonian.

4. Perturbation expansion and the Wick rotation

To set up a perturbative formalism for quantum gravity is deceptively simple[2)]. We define

$$g_{\mu\nu} = \eta_{\mu\nu} + \lambda_{\mu\nu} \tag{4.1}$$

with $\eta_{\mu\nu}$ = diag (−1,1,1,1) and $\lambda_{\mu\nu}$ is infinitesimal. Expanded up to second order in λ the Lagrangian is

$$\mathcal{L} = -\tfrac{1}{4}\left(\partial_\mu\lambda_{\alpha\beta}\right)^2 + \tfrac{1}{8}\left(\partial_\mu\lambda_{\alpha\alpha}\right)^2 + \tfrac{1}{2}\Lambda_\mu^2 - \tfrac{1}{2}\lambda_{\mu\nu}T_{\mu\nu} + \mathcal{L}^{\mathrm{matter}} \tag{4.2}$$

where

$$\Lambda_\mu = \partial_\alpha\lambda_{\alpha\mu} - \tfrac{1}{2}\partial_\mu\lambda_{\alpha\alpha} \ . \tag{4.3}$$

In the temporal gauge we require

$$\lambda_{o\mu} = 0 \ , \tag{4.4}$$

but of course other gauge choices, such as requiring Λ_μ to vanish, are also possible.

Let us Fourier transform in the three space coordinates and take as our dynamical variables $\lambda_{ij}(\vec{k},t)$. In $\vec{k}$ space the quadratic part of the Lagrangian diagonalizes, so that we can concentrate on one particular $\vec{k}$. Let us rotate $\vec{k}$ such that

$$\vec{k} = (0,0,k_3) \ . \tag{4.5}$$

Let us choose the beginning of the alphabet for indices that take the values 1,2 only (transverse indices), and define the traceless, transverse part $\tilde{\lambda}_{ab}$ of λ_{ij}:

$$\lambda_{ab} = \tilde{\lambda}_{ab} + \lambda\delta_{ab} \qquad (a,b=1,2) \tag{4.6}$$

$$\tilde{\lambda}_{aa} = 0 \quad ; \quad \lambda = \tfrac{1}{2}\lambda_{aa} \ . \tag{4.7}$$

We find that the Lagrangian then splits in three parts

$$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 \ , \tag{4.8}$$

with

$$\mathcal{L}_1 = -\tfrac{1}{4}k_3^2\,\tilde{\lambda}_{ab}\tilde{\lambda}_{ab} + \tfrac{1}{4}\dot{\tilde{\lambda}}_{ab}\dot{\tilde{\lambda}}_{ab} - \tfrac{1}{2}\tilde{T}_{ab}\tilde{\lambda}_{ab} \ , \tag{4.9}$$

where $\tilde{T}_{ab}$ is the transverse, traceless part of T_{ij};

$$\mathcal{L}_2 = \tfrac{1}{2}(\dot{\lambda}_{a3})^2 - \dot{\lambda}_{a3}\,\frac{i}{k_3}\,T_{oa} \ ; \tag{4.10}$$

$$\mathcal{L}_3 = \tfrac{1}{8}k_3^2\lambda^2 - \tfrac{1}{8}\dot{\lambda}^2 - \tfrac{1}{4}T_{oa}\lambda - \dot{\lambda}\dot{\lambda}_{33} - \tfrac{1}{2}\dot{\lambda}_{33}\dot{T}_{oo}/k_3^2 \ . \tag{4.11}$$

We make use of energy-momentum conservation:

$$ik\,T_{3\mu} = \dot{T}_{o\mu} \ . \tag{4.12}$$

Now $\mathcal{L}_1$ looks just like any Lagrangian for a driven harmonic oscillator and indeed $\tilde{\lambda}_{ab}$ which has two independent degrees of freedom, is an ordinary dynamical variable. It describes gravitons.

$\mathcal{L}_2$ implies $\partial_o\Pi_a = 0$, with, classically:

$$\Pi_a = \dot{\lambda}_{a3} - \frac{i}{k_3}\,T_{oa} \ . \tag{4.13}$$

As discussed in the previous section, we must impose $\Pi_a = 0$. Then $\mathcal{L}_2$ generates a Hamiltonian

$$H_2 = \tfrac{1}{2}\left|\Pi_a - \frac{i}{k_3}\,T_{oa}\right|^2 = \frac{1}{2k_3^2}\,T_{oa}^2 \ . \tag{4.14}$$

In $\mathcal{L}_3$ something new happens. The $\dot{\lambda}_{33}$ occurs only linearly. It acts as a Lagrange multiplier. This ensures:

$$\frac{d}{dt}\left(\lambda+T_{oo}/k_3^2\right) = 0 \ . \tag{4.15}$$

Again, we require more restrictively:

$$\lambda = - T_{oo}/k_3^2 \tag{4.16}$$

so that $\mathcal{L}_3$ does not contain any independent dynamical variables. It contributes to the Hamiltonian

$$H_3 = -\frac{1}{8k_3^2} T_{oo}^2 + \frac{1}{8k_3^2} T_{o3}^2 - \frac{1}{4k_3^2} T_{aa} T_{oo} \ , \tag{4.17}$$

of which the first part generates Newton's law.

We went through this exercise in perturbation theory in order to make a point. It is often useful in quantum field theory to perform the so-called Wick-rotation: $k_4 = ik_o$; $x_4 = it$. The functional integral then becomes

$$\int Dg \ \exp\left(i \int \mathcal{L} \ d^3x \ dt\right) \rightarrow \int Dg \ \exp \int \mathcal{L} \ d^4x \ . \tag{4.18}$$

In conventional quantum field theories the kinetic parts of the Lagrangian are all negative:

$$\mathcal{L} = - \tfrac{1}{2}(\partial_\mu \varphi)^2 - \frac{m^2}{2} \varphi^2 \ \ldots \tag{4.19}$$

so that the functional integral (4.18) can be well-defined. Now in our case $\dot{\lambda}_{33}$ acted as a Lagrange multiplier, α, securing some constraint,

$f = 0$. In Minkowski space one has

$$\int d\alpha\ e^{i\alpha f} \rightarrow \delta(f)\ . \tag{4.20}$$

After a Wick rotation one deals with Euclidean space, in which usually all fields and coordinates, including f, are real. But in order to obtain $\delta(f)$ as the result of a functional integration, we have to keep the i in (4.20). This means that in Euclidean space all Lagrange multipliers must be imaginary! Therefore, somewhat surprisingly, in the temporal gauge $\lambda_{o\mu} = 0$ (which becomes $\lambda_{4\mu} = 0$ in Euclidean space), we must require the field λ_{33} to be imaginary and not real. Actually, as in all integrals, there is some freedom in choosing the integration contours in the space of field variables. We could write

$$\alpha f = \tfrac{1}{4}(\alpha+f)^2 - \tfrac{1}{4}(\alpha-f)^2\ . \tag{4.21}$$

The functional integral will be well defined if α is imaginary and f real, but it converges better if $\alpha-f$ is chosen real, and $\alpha+f$ imaginary. So we see from (4.11) that it will be even better to choose

$$\sum_\mu \lambda_{\mu\mu}$$

to be imaginary, and all components of λ orthogonal to that to be real. In fact we found this requirement to hold in all gauge choices, not only the temporal gauge. We can understand this as follows. Write

$$g_{\mu\nu} = e^{\varphi}\ \hat{g}_{\mu\nu} \tag{4.22}$$

with

$$\det(\hat{g}_{\mu\nu}) = 1\ . \tag{4.23}$$

The Lagrangian then corresponds to

$$\int \sqrt{g}\, R = \int e^{\varphi}\Big(\hat{R} + \tfrac{3}{2}(\partial_\mu\varphi)^2 + \text{matter}\Big) \qquad (4.24)$$

$$= \int e^{\varphi}\Big[\tfrac{3}{2}(\partial_\mu\varphi)^2 - \tfrac{1}{12}\Big(\partial_\mu\hat{g}_{\alpha\beta}+\partial_\alpha\hat{g}_{\beta\mu}+\partial_\beta\hat{g}_{\mu\alpha}\Big)^2\Big] , \qquad (4.25)$$

where $\hat{R}$ is the curvature generated by $\hat{g}$ only, and $(\partial_\mu\varphi)^2$ stands for

$$\hat{g}^{\mu\nu}\ \partial_\mu\varphi\partial_\nu\varphi \ , \qquad (4.26)$$

etc.

In the perturbative regime we can neglect the exponent in front, and we notice that in order for the functional integral to be properly defined, φ must be imaginary and $\hat{g}$ real.

Naturally, we ask the question what requirement to use in the non-perturbative regime where the exponent e^{φ} cannot be ignored. We write

$$e^{\varphi}(\partial_\mu\varphi)^2 \propto (\partial_\mu e^{\frac{1}{2}\varphi})^2 \ . \qquad (4.27)$$

We see that the functional integral is well-defined if variations $\delta e^{\frac{1}{2}\varphi}$ of $e^{\frac{1}{2}\varphi}$ satisfy

$$|\mathrm{Re}(\delta e^{\frac{1}{2}\varphi})| \leqslant |\mathrm{Im}(\delta e^{\frac{1}{2}\varphi})| \qquad (4.28)$$

whereas we must also insist on

$$\mathrm{Re}(e^{\varphi}) \geqslant 0 \qquad (4.29)$$

to make the terms with $\hat{g}$ converge.

It is just barely possible to meet these requirements:

$$e^{\varphi} = 1 + ia \ , \ a \text{ real} \ . \qquad (4.30)$$

In the presence of scalar and vector fields S and A_μ, the matter-Lagrangian contributes

$$- \tfrac{1}{4}F_{\mu\nu}F_{\mu\nu} - \tfrac{1}{2}e^{\varphi}(\partial_\mu S)^2 - \tfrac{1}{2}e^{2\varphi}\left(m^2S^2 + \frac{\lambda}{12}\, S^4\right) . \tag{4.31}$$

We see that the kinetic parts are well behaved, but with (4.30), the mass term and scalar coupling terms diverge dangerously. The problem will not show up in perturbation theory, but will force us to choose φ dependent integration contours in S space.

With the choice (4.30) the integration over the conformal factor φ is just barely convergent. We suspect that this rather bad behaviour of the functional integral is related to an intrinsic instability in the theory of quantum gravity: collapse of matter into a black hole.

Another fundamental difficulty in quantum gravity was known for a long time: the theory is not renormalizable[2)]. Renormalization counter-terms with higher derivatives such as $(R_{\alpha\beta\gamma\delta})^2$, etc. would be necessary, which would add up eventually to non-local effects and violate causality. Now this problem seems to occur only at length scales in the order of the Planck scale, and its resolution will probably require drastic changes in the theory such as the "superstring" approach. But gravitational collapse may take place at length scales much larger than the Planck scale. Therefore we might be able to attack this problem using conventional quantum field theory only. This is the philosophy of our present research.

5. Black hole

The philosophy used as a starting point in quantizing gravity is that a Hilbert space exists, with a Hamiltonian, to be constructed for instance in the temporal gauge, where time is defined by clocks located at space-like infinity. Now one consequence of standard general relativity is well-known: if enough matter is accumulated in a small enough volume, then it collapses under the action of the gravitational force into a configuration called a "black hole". The black hole is the only naturally stable end product of this process. However, it is only stable in terms of the unquantized theory. If we set up a Hilbert space according to the above prescriptions, we find that its dimensionality is growing perpetually, even if one restricts oneself to states with total energy within certain bounds[3)].

This situation is unsatisfactory for various reasons. In the case of a very large black hole the problem does not look very serious: all regions of space-time visible to the outside observer are regular, so if we "solve" our field theories in those regions, we can predict the observations by using standard coordinate transformations. As first discovered by Hawking, these coordinate transformations produce a non-trivial background of radiation: the Hawking radiation emitted by a black hole[4)]. The fact that Hilbert-space is "expanding" is related to the ideal randomness of this thermal radiation: it corresponds to black body radiation with a temperature

$$T = \lambda / 8\pi M \ , \qquad (5.1)$$

where λ is probably 1, but other values are not entirely excluded[5)]. It has been deduced from thermodynamical arguments that this radiation gives black holes an entropy proportional to the area of the horizon:

$$S = 4\pi M^2/\lambda \tag{5.2}$$

But black-hole solutions could have sizes as small as or smaller than elementary particles. In fact, in some sense, elementary particles are black holes themselves. Is the internal Hilbert space of these particles bounded or infinite? If infinite then all axioms of quantum mechanics may fall apart.

Of course we are unable to prove that the axioms of quantum mechanics hold for black holes, but this author chose for the theory that quantum mechanics is valid, but the rules for general coordinate transformations might have to be adapted.

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.

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

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\Omega^2 \ . \tag{5.3}$$

The Kruskal coordinates u, v are defined by

$$uv = \left(1 - \frac{r}{2M}\right)e^{r/2M} \ , \tag{5.4}$$

$$v/u = -e^{t/2M} \ , \tag{5.5}$$

and then we have

$$ds^2 = -\frac{32M^3}{r} e^{-r/2M} \, du dv + r^2 d\Omega^2 \ , \tag{5.6}$$

which is now entirely regular at $r > o$. However (5.4) and (5.5) admit two solutions at every (r,t): we have two universes connected by a

---

* We now use units in which $G = 1$; see the remarks on units in section 3.

"whormhole". The Schwarzschild region, I, is $v > 0$, $u < 0$. The other regions are indicated in Fig. 1.

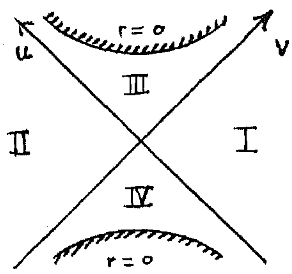


FIGURE I

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.

Are black holes fundamental particles or solitons? It would be tempting to view upon these objects as being the "magnetic monopoles" of gravity theory, extended solutions of localized equations. There are two important differences however. One is that, unlike magnetic monopoles, black holes can be arbitrarily small and their total mass reduces when they shrink. The other is that in the case of monopoles, a Callan-Rubakov procedure of quantization is possible[8]. But when we do quantum field theory near a black hole, we find that an ingoing wave reflects back. In the Schwarzschild coordinates, the speed of light,

$$\frac{dr}{dt} = 1 - \frac{2M}{r} , \tag{5.7}$$

tends to zero at the horizon (the points $r = 2M$). Wave packets accumulate there and never return: an infinite amount of information can be stored "indefinitely" at this horizon.

6. The brick wall model

It is not inconceivable that gravitational interactions will remove the infinities described in the previous section. If we look at wave packets that arrived at the region $r \simeq 2M$ long ago and those that will leave that region in the late future, then in the Kruskal frame we will see one wave packet squeezed onto the u axis and one onto the v-axis. One wave packet is Lorentz-boosted to tremendous energies with respect to the other. Because of these enormous relative energies one expects non-negligible gravitational interactions between the two. This will happen typically when they both come closer to the horizon than a distance comparable to the Planck length. Apparently then, at some distance h from the horizon the particles are no longer described by free wave packets.

This is how we got motivated to look at the following simplistic model[3)]. It should be regarded as an excercise rather than a theory. Let us assume that all wave functions vanish within some fixed distance h from the horizon:

$$\varphi(x) = 0 \quad \text{if} \quad x \leqslant 2M + h \ , \tag{6.1}$$

where M is the black hole mass. For simplicity we take $\varphi(x)$ to be a scalar wave function for a light ($m \ll 1 \ll M$) spinless particle. Later we will give them a multiplicity Z as a first attempt to mimic more closely the real world.

In view of a freely falling observer, condition (6.1) corresponds to a uniformly accelerated mirror which in fact will create its own energy-momentum tensor due to excitation of the vacuum. As in sect. 1 we stress that this presence of matter and energy may be observer dependent, but above all this model should be seen as an elementary exercise, rather than an attempt to describe physical black holes accurately.

Let the metric of a Schwarzschild black hole be given by

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2\, d\Omega^2 \ . \qquad (6.2)$$

Furthermore, we need an "infrared cutoff" in the form of a large box with radius L:

$$\varphi(x) = 0 \quad \text{if} \quad x = L \ . \qquad (6.3)$$

The quantum numbers are l, l_3 and n, standing for total angular momentum, its z-component and the radial excitations. The energy levels $E(l,l_3,n)$ can then be found from the wave equation

$$\left(1 - \frac{2M}{r}\right)^{-1} E^2\varphi + \frac{1}{r^2}\, \partial_r r(r-2M)\partial_r\varphi - \left(\frac{l(l+1)}{r^2} + m^2\right)\varphi = 0 \ . \qquad (6.4)$$

As long as $M \gg 1$ (in Planck units) we can rely on a WKB approximation. Defining a radial wave number $k(r,l,E)$ by

$$k^2 = \frac{r^2}{r(r-2M)}\left(\left(1 - \frac{2M}{r}\right)^{-1} E^2 - r^{-2}l(l+1) - m^2\right) , \qquad (6.5)$$

as long as the r.h.s. is non-negative, and $k = 0$ otherwise, the number of radial modes n is given by

$$\pi n = \int_{2M+h}^{L} dr\, k(r,l,E) \ . \qquad (6.6)$$

The total number N of wave solutions with energy not exceeding E is then given by

$$m^2 \ll 2M/\beta^2 h \ , \qquad L \gg 2M \ , \tag{6.11}$$

we find that the main contributions are

$$F \simeq -\frac{2\pi^3}{45h}\left(\frac{2M}{\beta}\right)^4 - \frac{2}{9\pi} L^3 \int_m^\infty \frac{dE(E^2-m^2)^{3/2}}{e^{\beta E}-1} \ . \tag{6.12}$$

The second part is the usual contribution from the vacuum surrounding the system at large distances and is of little relevance here. The first part is an intrinsic contribution from the horizon and it is seen to diverge linearly as $h \to 0$.

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

$$U = \frac{\partial}{\partial\beta}\,(\beta F) = \frac{2\pi^3}{15h}\left(\frac{2M}{\beta}\right)^4 Z \ , \tag{6.13}$$

$$S = \beta(U-\Gamma) = \frac{8\pi^3}{45h}\, 2M\left(\frac{2M}{\beta}\right)^3 Z \ . \tag{6.14}$$

We added a factor Z denoting the total number of particle types.

Let us now adjust the parameters of our model such that the total entropy is

$$S = 4\lambda^{-1} M^2 \ , \tag{6.15}$$

as in eq. (5.2), and the inverse temperature is

$$\beta = 8\pi\lambda^{-1} M \ . \tag{6.16}$$

This is seen to correspond to

$$h = \frac{Z\lambda^4}{720\pi M} \ . \tag{3.17}$$

Note also that the total energy is

$$\pi N = \int (2l+1)dl\ \pi n \stackrel{\text{def}}{=} g(E)$$

$$= \int_{2M+h}^{L} dr\left(1 - \frac{2M}{r}\right)^{-1} \int (2l+1)dl \sqrt{E^2 - \left(1 - \frac{2M}{r}\right)\left(m^2 + \frac{l(l+1)}{r^2}\right)}\ , \tag{6.7}$$

where the l-integration goes over those values of l for which the argument of the square root is positive.

What we have counted in (6.7) is the number of classical eigenmodes of a scalar field in the vicinity of a black hole. We now wish to find the thermodynamic properties of this system such as specific heat etc. Every wave solution may be occupied by any integer number of quanta. Thus we get for the free energy F at some inverse temperature β,

$$e^{-\beta F} = \sum e^{-\beta E} = \prod_{n,l,l_3} \frac{1}{1-e^{-\beta E}}\ , \tag{6.8}$$

or

$$\beta F = \sum_{N} \log\left(1-e^{-\beta E}\right)\ : \tag{6.9}$$

and, using (6.7),

$$\pi\beta F = \int dg(E)\ \log\left(1-e^{-\beta E}\right)$$

$$= -\int_{0}^{\infty} dE\ \frac{\beta g(E)}{e^{\beta E}-1}$$

$$= -\beta \int_{0}^{\infty} dE \int_{2M+h} dr\left(1 - \frac{2M}{r}\right)^{-1} \int \left(2l+1\right)dl$$

$$\times\ (e^{\beta E}-1)^{-1} \sqrt{E^2 - \left(1 - \frac{2M}{r}\right)\left(m^2 + \frac{l(l+1)}{r^2}\right)}\ . \tag{6.10}$$

Again the integral is taken only over those values for which the square root exists. In the approximation

$$U = \frac{3}{8} M , \tag{6.18}$$

independent of Z, and indeed a sizeable fraction of the total mass M of the black hole! We see that it does not make much sense to let h decrease much below the critical value (6.17) because then more than the black hole mass would be concentrated *at our side* of the horizon.

Eq. (6.17) suggests that the distance of the "brick wall" from the horizon depends on M, but this is merely a coordinate artifact. The invariant distance is

$$\int_{r=2M}^{r=2M+h} ds = \int \frac{dr}{\sqrt{1-2M/r}} = 2\sqrt{2}\ \mathrm{Mh} = \sqrt{\frac{Z\lambda^4}{90\pi}} \ . \tag{6.19}$$

Thus, the brick wall may be seen as a property of the horizon independent of the size of the black hole.

The conclusion of this section is that not only the infinity of the modes near the horizon should be cut-off, but also the value for the cut-off parameter is determined by nature, and a property of the horizon only. The model described here should be a reasonable description of a black hole aslong as the particles near the horizon are kept at a temperature as given by (6.16) and all chemical potentials are kept close to zero. The reader is invited to investigate further properties of the model such as the average time spent by one particle near the horizon, etc.

The model automatically preserves quantum coherence completely, but it is also unsatisfactory: there might be several conserved quantum numbers, such as baryon number*. What is wrong, clearly, is that we

* One may postpone this difficulty by inserting explicitly baryon number violating interactions near the horizon.

abandoned the principle of invariance under coordinate transformations at the horizon. The question that we should really address is how to keep not only the quantum coherence but also general invariance, while dropping all global conservation laws.

7. The equivalence theorem and Hawking radiation

Consider now the mapping from Schwarzschild coordinates to Kruskal coordinates and back. 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 a Hamiltonian

$$H = \int \mathcal{H}(x)dx > 0 \ , \qquad (7.1)$$

with one ground state

$$H|0>_k = 0 \qquad (7.2)$$

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 theoutside observer uses t as his time coordinate, and a generator of a boost in t produces

$$\delta v = \frac{v}{2M}\,\delta t \ , \qquad (7.3)$$

$$\delta u = -\frac{u}{2M}\,\delta t \ , \qquad (7.4)$$

so the generator of this boost is

$$h = \frac{1}{2M} \int dx\ \rho H(x) \quad ; \quad \rho = v-u \ . \tag{7.5}$$

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

$$H_I = \frac{1}{2M} \int \rho H(\vec{x}) d\vec{x}\ \theta(\rho) \quad ; \quad H_{II} = \frac{1}{2M} \int |\rho|\ H(\vec{x}) d\vec{x}\ \theta(-\rho) \ . \tag{7.6}$$

We have

$$[H_I, H_{II}] = 0 \ , \tag{7.7}$$

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

$$H_I |n,m\rangle = n|n,m\rangle \quad ; \quad H_{II} |n,m\rangle = m|n,m\rangle \ . \tag{7.8}$$

Extensive but straightforward calculations show that the "Kruskal vacuum" $|0\rangle_k$ does not coincide with the "Schwarzschild vacuum $|0,0\rangle$, but instead, we have

$$|0\rangle_k = C \sum_n |n,n\rangle e^{-4\pi Mn} \ , \tag{7.9}$$

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

$$h|0\rangle_k = 0 \ , \tag{7.10}$$

which is due to Lorentz-invariance of $|0\rangle_k$.

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

$$[\mathcal{O}, H_{II}] = 0 \ , \tag{7.11}$$

$$\Phi|n,m\rangle = \sum_k \Phi_{nk}|k,m\rangle \ , \tag{7.12}$$

and

$$\langle\Phi\rangle = {}_k\langle 0|\Phi|0\rangle_k = C^2 \sum_{n,n'} e^{-4\pi M(n+n')} \langle n',n'|\Phi|n,n\rangle = C^2 \sum_n e^{-8\pi Mn} \Phi_{nn} \ . \tag{7.13}$$

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

$$T = 1/8\pi M \ . \tag{7.14}$$

This is Hawking's result in a nutshell. Black holes radiate and the temperature of their thermal radiation is given by (7.14). 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 section 5 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[5] 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:

$$|n,m\rangle \rightarrow |n\rangle\langle m| = \rho \ . \qquad (7.15)$$

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

$$\frac{d}{dt}\rho_{nm} = -ih|n,m\rangle = -i(n-m)|n,m\rangle = -i[H_I, |n\rangle\langle m|] = -i[H_I,\rho] \ . \qquad (7.16)$$

Now the Kruskal vacuum $|0\rangle_k$ corresponds to the density matrix

$$\rho_{nn'} = C|n\rangle e^{-4\pi Mn}\langle n|\delta_{nn'} \ , \qquad (7.17)$$

which is a thermal state at temperature

$$T = 1/4\pi M \ , \qquad (7.18)$$

twice the usual result. The usual result would require not ρ from eq. (7.15) but $\rho\rho^\dagger$ to be the density matrix, from which of course (7.14) follows.

As long as we consider *stationary black holes with only soft particles* our mapping (7.15) is perfectly acceptable. The Hamiltonian (7.1) 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 (7.14) or (7.18) 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.

8. Hard particles

The black holes considered in the previous section were only exactly time-translation-invariant if they were covered by a Kruskal vacuum $|0>_k$. 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, $|0>_k$ 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 states, 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 Minkowski space produces a gravitational shock wave[6)], 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

$$\begin{aligned} u &= t-z \\ v &= t+z \end{aligned} \qquad (8.1)$$

we find that a particle moving in the positive z direction with momentum p, at $\tilde{y} = 0$, produces a shock wave on the v axis where the two half-spaces are connected after a shift

$$\delta v = -4p \; \ell n(\tilde{y}^2) \; . \tag{8.2}$$

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

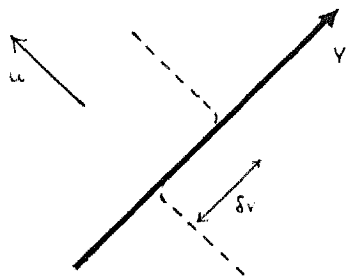


FIGURE 2

A way to picture this is to choose $g_{\mu\nu} = \eta_{\mu\nu}$ everywhere except at u=0, where all geodesics make a jump δv from past to future. See Fig. 2.

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 (8.2) solves Einstein's equations. In Kruskal's coordinates u,v a hard particle with momentum p again produces a shift δv, with

$$\delta v(\vec{\Omega}) = p f(\vec{\Omega},\vec{\Omega}') \; , \tag{8.3}$$

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

$$\Delta f - f = -2\pi\kappa \; \delta(\theta) \; , \tag{8.4}$$

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

$$f = \kappa \sum_{\ell} \frac{\ell+\frac{1}{2}}{\ell(\ell+1)+1} P_{\ell}(\cos\theta) \; , \tag{8.5}$$

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.

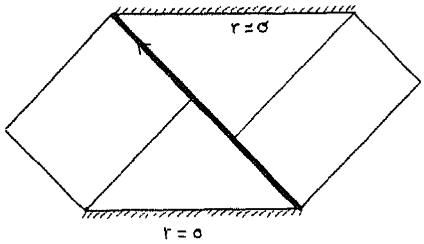


FIGURE 3

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[7]. The space-time structure of Fig. 4 results.

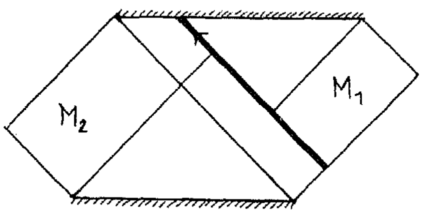


FIGURE 4

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.

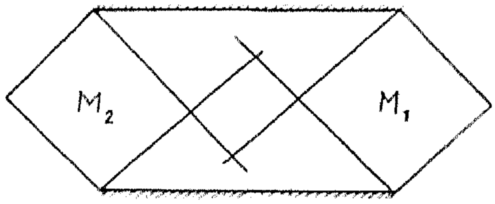


FIGURE 5

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=0 line. This r=0 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.

9. Purification of black hole states and the effect on the metric

We can now understand qualitatively some aspects of the mapping from Schwarzschild coordinates to Kruskal coordinates and back.

Suppose we have in Kruskal space an eigenstate of the Hamiltonian (7.1), say the lowest, the vacuum. The general coordinate transformation that maps this state onto Schwarzschild coordinates, where (7.5) is the Hamiltonian, produces some density matrix ρ. In our picture, a likely candidate is (7.17), with temperature (7.18), but this is not so crucial for the following. Our mapping is a1 ↮ b1 in Fig. 6.

Now let us consider another state in Kruskal space, such that in Schwarzschild space most of the ingoing particles disappear. We are searching then for solutions of equations of the form

$$a_s |\psi> = 0 \tag{9.1}$$

where a_s are annihilation operators in Schwarzschild space, whereas before we had

$$a_k |\psi> = 0 \tag{9.2}$$

where a_k are the particle-annihilation operators in Kruskal space. The relation between a_k and a_s is

$$a_s \sqrt{e^{\pi\omega} - e^{-\pi\omega}} = a_k(\omega) e^{\pi\omega/2} + a_k^{\dagger}(-\omega) e^{-\pi\omega/2} \tag{9.3}$$

$$b_s \sqrt{e^{\pi\omega} - e^{-\pi\omega}} = a_k(-\omega) e^{\pi\omega/2} + a_k^{\dagger}(\omega) e^{-\pi\omega/2} \tag{9.4}$$

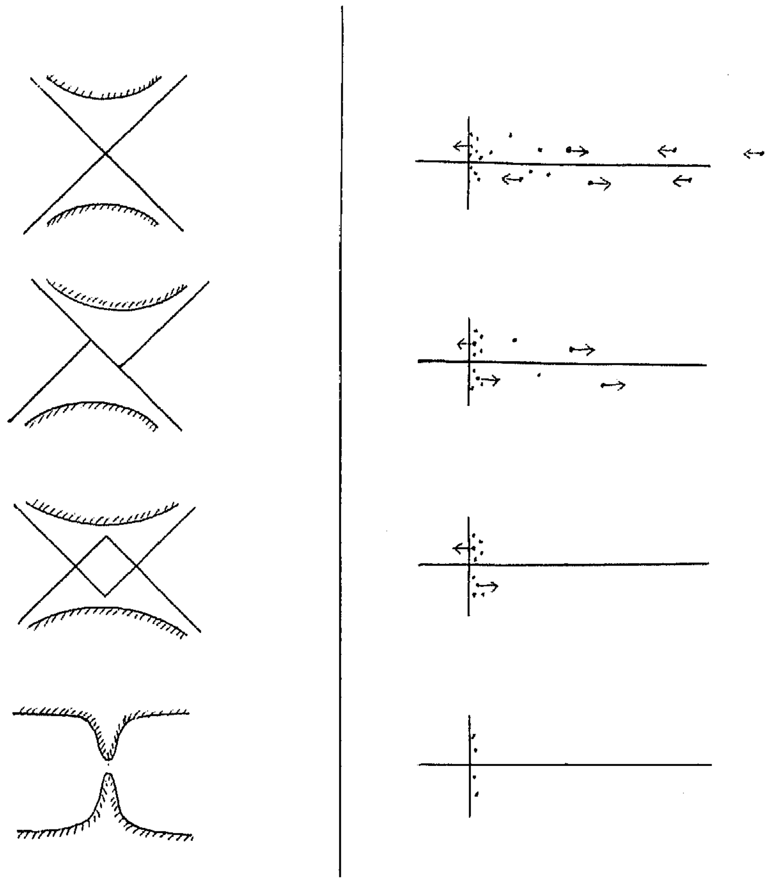
where ω is the energy eigenvalue of an ingoing or outgoing wave in Schwarzschild space. b_s are annihilation operators in section II of the Schwarzschild world. We may or may not require $b_s |\psi> \to 0$.

1)

2)

3)

4)



a) b)

FIGURE 6

Mapping from Kruskal coordinates (a) to Schwarzschild coordinates (b).

Clearly (9.1) and (9.2) cannot simultaneously be satisfied. So, imposing (9.1) instead of (9.2) implies admitting particles in the Kruskal world. If all a_s would annihilate $|\psi>$ then an infinite amount of matter would appear in the Kruskal world. That is more than we can handle. So rather we decide to annihilate all particles further from the horizon than some distance h in Schwarzschild space. The particles in

Kruskal space then have a finite total energy-momentum, with a corresponding effect on the metric. This is pictured in Fig. 6 (a2, b2).

We can do the same thing with the outgoing particles in Schwarzschild space, thus producing a black hole which is pure in a region that extends outside the distance h from the horizon. We see in Fig. 6 (3) what effect this has on the metric in Kruskal space.

Finally we try to reduce h to as small a size as possible. What is expected, though we were unable to prove, is that there is a limit for h that should not be surpassed because then the singularities of the Kruskal metric would touch each other: the singularity at $r \to 0$ would become space-like. In Fig. 6 (4) we are close to the limit. In that case the central region approaches the flat metric. If the central region becomes flat the wormhole to universe II is squeezed off: there will be no coupling anymore between bra- and ket-states. We suspect that precisely in this limit the "density matrix" describing the Schwarzschild world becomes that of a pure quantum mechanical state, that is

$$\rho = |\psi\rangle \langle\psi| \ , \qquad (9.5)$$

so that it has one eigenvalue 1 and the others are zero.

We believe that the scenario sketched in this section can replace the "brick wall model" to obtain a more natural description of a black hole in a pure state.

10. Conclusion

The aim of our investigation was to apply quantum field theory at moderately large distance scales in sufficiently smooth but non-trivial space-times of deduce the quantum properties of black holes. This aim was not achieved. Although the black hole is classically a stationary configuration, it is a run-away solution in the quantum-mechanical sense: the dimensionality of Hilbert space seems to explode indefinitely. On the one hand this is of course a disappointment. A more detailed theory of black holes could suggest a link between them and the elementary constituents of some "unified field theory". On the other hand this situation is extremely interesting, because it strongly indicates that even at large distance scales quantum field theory may have to be amended. Our "dissident" ideas about the density matrix at the horizon might be an example of such a change, but in any case they are not sufficient.

Curiously, some very fundamental notions in quantum mechanics may have to be reconsidered. In General Relativity a Hamiltonian only exists if time is defined by "clocks" at spatial infinity, as we saw in section 3. But which definitions should we use when an observer falls into a black hole? He should take his clocks with him.

It is quite likely that black holes can be described in many different ways, which are essentially equivalent. This is because we can put anything we like in space II without observing any difference in space I. This resembles the situation in a gauge theory where also the dynamical variables can be described in many different (but "gauge equivalent") ways. Perhaps black holes are huge gauge particles.

References

1) R. Adler, M. Bazin, M. Schiffer, "Introduction to General Relativity", McGraw-Hill 1965.

P.A.M. Dirac, "General Relativity", Wiley Interscience 1975.

S. Weinberg, "gravitation and Cosmology: Principles and Applications of the General Theory of Relativity", J. Wiley & Sons.

S.W. Hawking and G.F.R. Ellis, "The large scale Structure of Space-time", Cambridge Univ. Press 1975.

S. Chandrasekhar, "The Mathematical Theory of Black Holes", Clarendon Press, Oxford Univ. Press.

R.M. Wald, "General Relativity", 1984, Univ. of Chicago Press.

2) B.S. DeWitt, "Quantum Theory of Gravity I, II, III", Phys. Rev. 160 (1967) 1113; 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.

3) G. 't Hooft, "On the quantum structure of a black hole", Nucl. Phys. B256 (1985) 727.

4) S.W. Hawking, "Particle creation by black holes", Comm. Math. Phys. 43 (1975) 199.

J.B. Harkle and S.W. Hawking, "Path integral derivation of black hole radiance", Phys. Rev. D13 (1976) 2188.

W.G. Unruh, "Notes on black-hole evaporation", Phys. Rev. D14 (1976) 870.

5) G. 't Hooft, "Ambiguity of the equivalence principle and Hawking's temperature", J. Geom. Phys. 1 (1984) 45.

6) T. Dray and G. 't Hooft, "The gravitational shock wave of a massless particle", Nucl. Phys. B253 (1985) 173.

7) T. Dray and G.'t Hooft, "The effect of spherical shells of matter on the Schwarzschild black hole", Comm. Math. Phys. 99 (1985) 613.

8) 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.