Black Holes and the Foundations of Quantum Mechanics

Gerard 't Hooft

UNIVERSITY OF UTRECHT

Abstract

A new relativity principle in quantum mechanics may have observable consequences in the Hawking radiation of a Schwarzschild black hole. The original formulation of this phenomenon is questioned.

1. Introduction

Understanding the quantum mechanical properties of black holes may well be crucial if one wishes to devise a credible model of elementary particles in the region of the Planck length scale. Concerning a black hole with Schwarzschild radius much greater than the Planck length one expects to be able to learn much from considerations of ordinary quantum field theory in the modestly curved background of its metric. Indeed until shortly it seemed that one result is well established [1]: due to vacuum fluctuations around the Schwarzschild horizon all heavy black holes must emit particles of all sorts, and this emission has a thermal spectrum corresponding to a temperature $T = 1/(8\pi M)$, where M is the mass of the black hole, in Planck units. The importance of this result is that it suggests well-determined decay properties of all "fundamental" particles heavier than the Planck mass.

Unfortunately, two difficulties were standing in the way of further progress. One is that this new piece of information is clearly only statistical in nature. Purely thermal radiation can only be expected from sufficiently large systems so it is reasonable to assume that only very large black holes behave this way on the average. For not too large black holes we would like to have the quantummechanical amplitudes, not only the probabilities. Can we also say something about these amplitudes by applying quantum field theory in a curved space-time? The standard picture appears to deny such a possibility.

Secondly, in an attempt to find an improved formalism, this author recently proposed an alternative model [2] for the radiating black hole. Although not yet all consequences of this model have been elaborated, it has one striking feature: black holes radiate, as before, but here the temperature is $1/(4\pi M)$, not $1/(8\pi M)$. For this reason, most investigators dismissed this model. Hawkings value for the temperature, $1/(8\pi M)$, seemed to be an inescapable consequence of known laws of physics. So how can one possibly believe a model that gives a conflicting value? Now this would have been the end of just another untenable theory if the author had not further analysed the cause of the discrepancy. The flaw, he now claims, is not in the logic of the new theory but in the logic of the old one, in as far as inescapability of Hawkings result is claimed. In fact, the new model suggests a new alley to produce a completely quantum mechanical, not just statistical, description of large black holes.

We claim that the very foundation of quantum mechanics and relativity theory are touched upon.

Since the time the author first proposed the alternative model several years have past and by now he is convinced that it is presumably as naive or worse than the standard picture. So we do not desire to defend the factor 2 at all cost, but there is this curious possibility that the probability interpretation of wave functions needs to be reconsidered with extreme care when a black hole horizon is present.

In the next section we present the most essential steps in the derivation of Hawkings familiar result, and then in section 3 describe briefly our alternative. In section 4 we explain our attitude concerning the difference between the old and new models as a new "relativistic" effect. We then give in section 5 an extensive discussion of the objections against our views that are most commonly heard.

2. The Conventional Theory

The original Schwarzschild metric is defined in a region $r \geqslant 2M$ and is independent of its time coordinate t. As is well known there is a natural extension of this metric showing that $r = 2M$ is not a singular point. One writes, for instance [3],

$$uv = \left(1 - \frac{r}{2M}\right) e^{r/2M}, \tag{2.1}$$

$$\frac{v}{u} = -e^{t/2M}, \tag{2.2}$$

the so-called Kruskal coordinates. Defining

$$\rho = v - u; \tau = v + u, \tag{2.3}$$

we find that the Schwarzschild region, I,

$$\rho \geqslant |\tau|, \tag{2.4}$$

is solidly connected to another region, II;

$$\rho \leqslant -|\tau| \tag{2.5}$$

via regions III ($\tau > |\rho|$) and IV ($\tau < -|\rho|$). The metric in II is the mirror image of region I, a statement that remains true also for the Kerr, Reissner-Nordstrom and Kerr-Newman solutions. Regions II and IV do not exist or carry a different metric when the imploding object that give birth to the black hole is taken into account. If we nevertheless use these regions of the metric (2.1), (2.2) then they should be considered as nothing but convenient analytic extensions of the metric, to which we can also extrapolate the fields of the outgoing particles.

Consider the generator H of a translation in the Kruskal time coordinate τ. Since this is not a symmetry of the metric this hamiltonian is not exactly conserved but as long as the curvature of space-time is not too strong it does describe the evolution of a quantum field in this metric. We can write

$$H = \int \mathcal{H}(\mathbf{x})\mathbf{dx} \geqslant o \tag{2.6}$$

and a single ground state $|o\rangle_k$ with

$$H|o\rangle_k = o \tag{2.7}$$

is to be expected (after normal ordering of $\mathcal{H}$). Due to curvature this vacuum is not exactly but approximately conserved.

We wish however to describe the time evolution of the

Schwarzschild system as seen by an outside observer. The generator of a translation in t turns out to be

$$h = H_{\rm I} - H_{\rm II}, \tag{2.8}$$

with

$$H_{\rm I} = \frac{1}{2M}\int_{\rho=o}^{\infty} \rho\, \mathcal{H} d^3x;\; H_{\rm II} = \frac{1}{2M}\int_{\rho=-\infty}^{o} |\,\rho\, |\mathcal{H} d^3x, \tag{2.9}$$

and

$$[H_{\rm I}, H_{\rm II}] = o. \tag{2.10}$$

We can write the eigenstates of $H_{\rm I}$ and $H_{\rm II}$ as $|n,m>$ with

$$H_{\rm I}\,|\,n,m\rangle = n\,|\,n,m\rangle;\; H_{\rm II}\,|\,n,m\rangle = m\,|\,n,m\rangle. \tag{2.11}$$

A straightforward calculation now shows that the "Kruskal vacuum" $|\,o\rangle_K$, does not coincide with the "Schwarzschild vacuum" $|\,o,o\rangle$, but instead we have

$$|\,o\rangle_K = C\sum_n |\,n,n\rangle e^{-4\pi Mn}, \tag{2.12}$$

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

$$h\,|\,o\rangle_K = o, \tag{2.13}$$

which is due to t-invariance of $|\,o\rangle_K$.

In the view of the establishment the way to proceed is now as follows. Consider any observable operator O, built from the field operators ϕ in space I. Necessarily we have

$$[O, H_{\rm II}] = o;\; \langle n',m'|\,O\,|n,m\rangle = O_{nn'}\delta_{mm'}. \tag{2.14}$$

We find

$$\begin{aligned} O_K &= {}_K\langle o\,|\,O\,|\,o\rangle_K \\ &= C^2\sum_{n,n'} e^{-4\pi M(n+n')}\langle n',n'\,|\,O\,|\,n,n\rangle \\ &= C^2\sum_n e^{-8\pi Mn}\,O_{nn}. \end{aligned} \tag{2.15}$$

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

$$T = 1/(8\pi M). \tag{2.16}$$

Hence, the Kruskal vacuum $|\,o\rangle_K$ corresponds to a mixed state as seen by observers in I with temperature as given by (2.16); the black hole radiates. The reason why we have a mixed state is clear: all particles in space II are fundamentally unobservable and any quantum mechanical superposition of states in I and II corresponds to mixing of states in I only.

The above constitutes only the main lines of the argument. It can be strengthened by confirming that indeed $| o\rangle_k$ is the state that will emerge when matter collapses, and by checking that eq. (2.15) is quite generally valid also for non-scalar fields.

Thus, the logic that was used to obtain this result seems to be impeccable and unavoidable. Yet it could be wrong. In any case, the conclusion that collapsing matter in a pure state would produce a black hole in a mixed state would imply that abundant "mixture" of quantum mechanical states would occur at the Planck scale of length, time and energy. It seems an impossible task then to recover any kind of unified particle theory at that scale such that ordinary "pure" quantum mechanics would emerge at our presently available energies. More importantly, this conclusion seems to contradict the pure field theoretic arguments that *were* used as a starting point for the calculations.

3. The Alternative Model and Its Temperature

In our alternative model the information that went into space II is not entirely lost. We use the same Kruskal metric as before. Again we assume that at equilibrium the particle content of the black hole is described by the Kruskal vacuum $| o\rangle_k$ (or else a state that contains only a negligible amount of particles as seen by the Kruskal observer). All we do is give a different interpretation of the basis $| n,m\rangle$ in terms of the states in the Hilbert space for an outside observer. We noticed that the time evolution is

$$| n,m\rangle_t = e^{-int+imt} \, | n,m\rangle_o, \tag{3.1}$$

which is exactly that of a density matrix ρ_{nm}, or equivalently the product $| n\rangle\langle m |$ of bras and kets both defined *only* in the Hilbert space spanned by observables in space I. The fact that the metric in space II is the exact mirror image of space I (in black holes where all matter sources have been removed) implies that the Hamiltonian $H_{\rm I}$ has the same spectrum as $H_{\rm II}$. This, as stated before, remains true for rotating and charged black holes. In our alternative theory we use this fact to postulate that states $| o\rangle_k$ or $| n,m\rangle$ in Kruskal space described not states but density matrix elements for an outside observer in Schwarzschild space.

That this is a profound departure from the previously described picture becomes apparent when the temperature of a radiating black hole is computed. Again our starting point is the Kruskal vacuum $| o\rangle_K$. We now rewrite eq. (2.12) as

$$| o\rangle_K = C \sum_n | n\rangle \langle n | e^{-4\pi n} \tag{3.2}$$

Clearly this is a density matrix. It is diagonal and therefore stationary in the time t. We recognize the Boltzmann factor

$$e^{-4\pi n} = e^{-\beta n};\ T = 1/\beta = 1/(4\pi M). \tag{3.3}$$

The temperature is twice the usual value, eq. (2.17).

This different result should be easily detectable experimentally whenever a light black hole (light in astronomical sense) would be found.

Non-diagonal density matrices where the black holes in the bras and kets have *different* masses can also be constructed but then the intertial observer sees matter close to the horizon [2,4]. The back-reaction of this matter causes the mass-difference.

In the conventional model the inclusion of the back-reaction of Hawking's radiation required to describe a shrinking black hole is problematic, though interesting attempts are being made [5].

We note that the density matrix implicit in the standard theory (see eq. 2.15) is the square of ours (eq. 3.2). One may therefore summarize our proposal by emphasizing that we postulate a *linear* mapping between states $| \psi\rangle_K$ and density matrices ρ in Schwarzschild space I. In a quadratic mapping all states in II would be averaged over, which is not what we want.

4. A Reconciliation: Relativity Revisited

Should we dismiss the above model as being impossible? As long as we restrict ourselves to those systems to which the postulate can be applied (a black hole emitting and absorbing radiation close to its stationary equilibrium point) there seems to be no contradiction whatsoever. So we asked the question whether a principle can be formulated that is ubiquitously valid so that the theory can be made complete. To explain the principle which we found, we first remind the reader of the fundamental notions of relativity.

Relativity tells us that different observers might interpret their measurements of certain phenomena quite differently. When one observer sees an electric field, another might see a superposition of an electric and a magnetic field. When one observer states that the chronological order of two events *A* and *B* was (*A*,*B*), another might get the conflicting result (*B*,*A*). In short, rather than requiring physical observations to be invariant under Lorentz transformations, it might be necessary to require only *co*variance. What this really means, in our view, is the following:—there is no such thing as Lorentz-invariance of physically observable notions such as electromagnetic fields, or the chronological order of events. After all, the whole universe *does* have a preferred coordinate frame (moving with the same velocity as the average of the stars). What relativity means is that the laws of physics as experienced by observers who move with a certain velocity *can be derived* from the laws of physics as seen by a stationary one. But in the derivation transformations of all sorts may be necessary. This is called covariance, as opposed to invariance. All transformation rules are in principle acceptable. In fact, the further we let our phantasy go, the more likely we are to discover new forms of covariance, and with that new laws of nature.

General relativity fits well in this picture. We can derive the laws of physics as seen by an observer in a gravitational field if we know what is observed by a freely falling observer. But we should not necessarily require the observations to be identical. A transformation may be necessary. What we require in the case at hand is a further transformation in Hilbert space which is of a quite unusual type: pure states for an observer in free fall are linearly mapped onto elements of the density matrix for an accelerated observer. Wherever $\partial t/\partial\tau$ is negative, for certain values of $\mathbf{x}$, we have that an observable $\phi(\mathbf{x},\tau)$ for the freely falling observer acts on ket-space, but for the accelerated one on bra-space.

All this implies that if the density matrix is used to compute probabilities, as it should, then one observer may experience probabilities differently from others. As before, not the physics itself is invariant under general coordinate transformation, but by means of a transformation we can deduce the laws of physics in one frame from those in another.

One can verify that this covariance principle is sufficient to ensure conventional general invariance in the classical (i.e. non-quantum mechanical) limit and conventional quantum mechanics remains true locally (as long as horizons stay far away). In short we claim that there is no contradiction with the known laws of physics.

The situation in our theory could be viewed as being not fundamentally different from that in the famous Stern-Gerlach experiment. There the question which event actually occurred (which slit was passed by the electron) cannot be answered unambiguously. Here, the question with which probability an event occurred depends on the observer. One might of course add that, anyway, the infalling observer will not be able to communicate his observations to those who stay outside. Worse even, he and his entire laboratory are absorbed and will form part of the Hilbert space studied by the outside observer!

The conclusion of this section is that the difference in the predictions of the two theories lies in the fact that we require different transformation laws relating the observed results from the different observers. The difference only occurs when a horizon is present and nowhere else. The crucial factor is time reversal. It could be that time-reversal has more far-reaching effects on the laws of physics than previously taken for granted.

5. Objections Raised to the Theory, and the Answers

Anyone is invited to shoot down our alternative theory, and as stated in the Introduction, many attempts were made. We here list the objections most often heard together with the replies one could give. Although the latter are still far from perfect one might conclude that the objections are not insuperable.

OBJECTION: One might object against the use of the regions II and IV as described by the Kruskal metric: one should instead consider a black hole that has been formed by collapse of matter and in that case there is no mirror image space II at all. The bulk of matter that went in, no matter how long ago, removes it.

REPLY: Our alternative model, as formulated, is only suitable for describing a black hole close to its stationary equilibrium. During collapse it was not at all close to equilibrium. The model still has to be extended to describe large amounts of matter going in and/or out. Progress has already been made [2,4] but the problem is not yet completely understood. Note however that in any case we are dealing with a quantum theory here, in which some form of PCT invariance is to be expected. It would be unreasonable to treat region IV ($\tau < -|\rho|$) and region III ($\tau > |\rho|$) entirely differently. The standard picture, used in the objections, assumes that the Penrose diagram of the complete history of a blackhole can always be given. We however assume that this is only so for black holes in certain mixed states. A "pure" black hole always evolves. Note that the black hole is not stable, it decays, so that if one selects a pure state or a differently mixed state at a given time, one might inevitably obtain a *superposition* of metrics both in the far past and in the far future. Only one mixed configuration is stationary in time and we use the full Kruskal metric to describe it.

OBJECTION: Consider now a dust cloud as large as our galaxy. Take as an initial condition that all particles move slowly inward. An observer will still feel quite comfortable while he moves through the horizon. One can "compute" what will be seen by the observer who stays outside using nothing but standard physics. One gets $T = 1/(8\pi M)$.

REPLY: That is a low value indeed. But let's assume the outside observer can measure it. Even though standard physics was used in the calculation, one assumption was made whether one likes it or not. It is the assumption that the outside observer has the same Hilbert space to his disposal as the infalling one and that the probabilities he measures are the same as those of the other. But his Hilbert space is smaller because there are things he cannot see and that are irrelevant to him. Thus a further transformation law might be necessary. We may have covariance, rather than invariant probabilities. The infalling observer cannot communicate his results to the outside one and therefore there are no contradictions in our alternative model.

OBJECTION: Both space I and space II were used to describe the Hilbert space seen by the outsider. That corresponds to identifying points (u,v) with $(-u,-v)$ of the Kruskal coordinates. If we continue to Euclidean space then such an identification would indeed divide the periodicity β by two and the temperature doubles. But, such an identification would lead to a conical singularity at the origin of Kruskal space, and an extremely singular curvature there, whereas no appreciable amount of matter could be present there that could be responsible for such an enormous curvature.

REPLY: In the *classical* (i.e. non-quantum mechanical) *limit* points (u,v) and $(-u,-v)$ might indeed be considered identified. It doesn't really matter much because, remember, regions II and IV were analytic continuations. And in the classical limit no one can check whether matter was or wasn't present at $u = v = o$. But in our description of the quantum mechanics where we used the state $|\, o\rangle_K$, those points are definitely not identified. This state has no conical singularity at the origin so no singular amount of matter is needed there. As for Euclidean space, this should always be regarded as a mathematical tool, not as a physical reality.

OBJECTION: In the more detailed descriptions of the alternative models in refs [2] it is claimed that gravitational interactions between infalling matter and outgoing Hawking radiation diverges and that this is a reason to reconsider the complete formulation. Yet the standard arguments show that, in Kruskal coordinates, the metric tensor $T_{\mu\nu}$ is basically regular at the horizon so fear of large gravitational effects there is totally unjustified.

REPLY: This is only true in the description of the black hole in its mixed state, as generated by $|\, o\rangle_K$. If we try to describe a *pure state*, either by producing a density matrix with only one column, or with eigenvalues 1 and 0 only, then the corresponding state in Kruskal space would have infinite amounts of matter at the horizon, leading indeed to uncontrollable gravitational forces at the origin. We assume that in our model transitions between pure and mixed states do not take place and therefore "pure" black holes exist also.

OBJECTION: The alternative model suggests a complete Hamiltonian with a discrete spectrum for a black hole. Where is it?

REPLY: It is being worked on but no promises are made. We do not pretend to solve the world.

OBJECTION: In eq. (3.2) the right hand side is not properly normalized if the left hand side is (normalization is left unaffected in the standard theory). Indeed any definition of an inner product between "bra" states in space II and "het" states in space I seems to be unnatural or ad hoc as seen by a Kruskal observer.

REPLY: This may perhaps be the most severe objection to which we have no completely satisfactory answer. Of course it could be that a simple adaption of the normalization constant when a mapping is performed is the logical answer, and we note that on both sides of the equation the norm does not vary with time. But it is indeed strange that the "Identity matrix" **1** for a Schwarzschild observer corresponds to a highly excited state in Kruskal space. In another variant of our model it is this identity matrix which is mapped onto the vacuum state $|o\rangle_k$. In that case the temperature of a black hole may appear to be infinite!

6. Conclusion

There is no way to reject a priori a fundamental requirement to further transform the data from the infalling observer before they are interpreted by an outside observer. We have made no attempt (yet) to *prove* that such a requirement *must* be made. The temperature of a black hole could be twice the usual value, and as a consequence the final explosion of a mini-black-hole would have a little more than twice the energy as computed in the conventional way [6].

As indicated in the title of this paper the arguments presented here may affect the interpretative basis of quantum mechanics itself. Whether or not one accepts the views presented here may depend on the way in which one adheres to the dogmas of this theory.

Acknowledgment

The author acknowledges a fruitful correspondence with R.M. Wald.

Note Added in Proof

More recent attempts by this author and others to construct coherently quantum mechanical models for black holes usually favor the conventional picture, not our alternative one, so that Hawking's original value for the temperature emerges. Region II may refer to some other area *outside* the black hole horizon. But the matter is not settled, and we would still like to emphasize the possibility of a philosophy as sketched in this paper.

References

[1] S.W. Hawking, "Particle creation by black holes", Comm. Math. Phys. **43** (1975) 199.
J.B. Hartle 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.
R.M. Wald, "On particle creation by black holes", Comm. Math. Phys. **45** (1975) 9.
R.M. Wald, "Entropy and black hole theromodynamics", Phys. Rev. **D20** (1979) 1271.

[2] G. 't Hooft, "Ambiguity of the equivalence principle and Hawking's temperature". J. Geometry and Phys. **1** (1984) 45.
G. 't Hooft, "On the quantum structure of a black hole", Nucl. Phys. **B256** (1985) 727.

[3] S.W. Hawking and G.F.R. Ellis, "The large scale structure of space-time", Cambridge Univ. Press (1973).

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

[5] R. Balbinot, "Hawking radiation and the back reaction—a first approach", Class. Quantum Gravity **1** (1984) 573.

[6] D.N. Page and S.W. Hawking, "Gamma rays from primordial black holes", Astrophys. Journal **206** (1976) 1.