SOME DIFFICULTIES IN RECONCILING GRAVITY AND QUANTUM FIELD THEORY

and Some Recent Proposals

Luis Álvarez-Gaumé
Theory Division
CERN, CH-1211 Geneva 23
Switzerland

General Relativity is peculiar with respect to other classical theories in that it predicts its own limits. This is the subject of the Singularity Theorems [1]. During the sixties and seventies there was a good deal of activity in trying to determine the general properties of gravitational collapse and black hole formation, and this led to a number of beautiful and deep results on the conditions for the appearance of singularities under gravitational collapse, the formation of horizons, the Cosmic Censorship hypothesis, the no-hair theorems and black hole thermodynamics among others (for details and references see for instance [2]). The subject of this talk is to present a cursory overview of the conceptual problems one finds when trying to quantize gravity as a consequence of the existence of black holes and horizons in General Relativity, and some of the solutions which have been proposed in the context of String Theory or String inspired models of gravity. The organization of this talk is as follows. First we will quickly review some of the relevant properties of black holes. Second we will analyze the quantization of fields around black holes, the Hawking radiation process, black hole evaporation, Hawking’s superscattering operator $S$ and Hawking’s radical proposals that complete black hole evaporation violates the standard unitary evolution in Quantum Mechanics in the sense that in the presence of an evaporating black hole pure states evolve into mixed states. Third we will present some simplified two-dimensional models (two-dimensional dilaton gravity) where explicit quantum computations can be carried out [3] to a certain extent. This is one of the few examples where the reduction to two-dimensions keeps
most of the conceptual problems intact but leads to an enormous technical simplification. We will present some of the recent suggestions on how to avoid some of the more unpalatable conclusions in Hawking's analysis, and finally we will present at a more fundamental level the exact two-dimensional black hole solution found by E. Witten [4] in String Theory which may provide an arena where many of the issues and controversies in the subject might eventually be settled.

A useful way of portraying black hole space-times is in terms of Penrose diagrams [2], which provide a clear picture of the causal structure of a given space-time. If $I^+$ represent future null infinity in a given asymptotically flat space-time, we can construct the chronological past of $I^+$, $J^-(I^+)$ as the set of points of the space-time which can be joined to $I^+$ by time-like or null curves. When the chronological past $J^-(I^+)$ does not coincide with the full space-time we say that we have a black hole region. The boundary of this region is the horizon, and under certain assumptions embodied in the singularity theorems, the horizon contains inside a space-time singularity. The metric of a spherically symmetric static black hole is the Schwarzschild metric

$$ds^2 = (1 - \frac{2M}{r})dt^2 - \frac{dr^2}{1 - \frac{2M}{r}} - r^2(d\theta^2 + \sin^2\theta d\phi^2).$$

(1)

$r = 2M$ represents the position of the horizon and the curvature singularity is located at $r = 0$. In the presence of singularities our classical theories (and also quantum mechanics) need extra boundary conditions, and a spectre of unavoidable uncertainty creeps into the subject. Through the analysis of black hole evolution and the behavior of fields in the presence of black holes, it was possible to formulate the laws of Black Hole Thermodynamics (for details and references see [2]). In particular it was possible to assign an entropy to the horizon, which in the case of a spherically symmetric black hole is simply the area of its horizon. For (1) this yields

$$S = \frac{1}{4} \text{Area of Horizon} = 4\pi M^2.$$  

(2)

This implies naively that the number of states one can associate to a black hole grows like $exp(const.M^2)$. It is difficult to understand the origin of these states and what is their correct dynamics. This problem appears more acutely when one considers quantum fields in the presence of black holes. In principle if the mass of the hole is much larger than the Planck scale, one would expect that semiclassical arguments can be applied and that one should be able to treat gravity classically. After all, in Planck units the space-time curvature near the
horizon of a massive black hole is small. This naive expectation has been under attack for some time [5], and some workers argue now that some of the puzzles posed by quantization around a black hole do require detailed knowledge of Quantum Gravity. It is indeed quite debatable to what extent we can truly decouple quantum gravitational effects from matter in the presence of strong gravitational fields.

For large mass black holes one can carry out the quantization of a field in its background. In 1974 Hawking [6] showed that black holes are not black after all, and that they radiate thermally at a temperature proportional to the surface gravity of the horizon. This process of Hawking emission makes the horizon decrease, but if one includes the contribution to the entropy of the radiation, the total entropy of the hole plus radiation system increases. To give a flavor of how Hawking's result is obtained, we consider the quantization of a massless free scalar field in the presence of a collapsing body which eventually forms a black hole (more details and references can be found for example in the book [7]). Since the field is massless we can consider initial conditions on null surfaces. In particular in the past we chose initial conditions on $I^-$, past null infinity. On $I^-$ we can expand the solutions to the Klein-Gordon equation in terms of positive and negative frequency solutions with respect to past asymptotic time. The coefficients of these solutions are the creation and annihilation operators for the modes. Thus the field takes the form

$$\phi \sim \sum_j p_j a_j + p_j^* a_j^\dagger.$$  \hspace{1cm} (3)

In the asymptotic future the complete Cauchy surface contains future null infinity $I^+$ together with the horizon (a null surface as well). Thus we can choose a basis of solutions to the Klein-Gordon equation ($f_i, q_i$) where $f_i$ has support only near $I^+$ while the $q_i$'s have only support near the black hole horizon and the black hole interior but not at future null infinity. Each of these solutions can be divided into positive and negative frequency solutions (there is certainly an ambiguity in the choice of the solutions with support near the hole, but this does not change the results about the Hawking radiation), and as usual we assign annihilation operators to the positive frequency solutions and creation operators to the negative frequency ones. The field $\phi$ expanded in this basis takes the form

$$\phi \sim \sum_i f_i b_i + f_i^* b_i^\dagger + \sum_i q_i c_i + q_i^* c_i^\dagger.$$  \hspace{1cm} (4)

Out of the oscillators we can construct three Fock spaces $\mathcal{H}_{in}, \mathcal{H}_{out}, \mathcal{H}_{bh}$. The incoming state is taken in $\mathcal{H}_{in}$, and it will evolve into a state in the tensor product $\mathcal{H}_{out} \otimes \mathcal{H}_{bh}$. If for
simplicity we send in the vacuum state of the in Fock space, in terms of density matrices \( \rho_{in} = |0_{in}\rangle\langle 0_{in}| \), the claim is that the state seen by an asymptotic observer out in \( \mathcal{I}^+ \) is a thermal state with density matrix \( \rho_{out} = \exp(-\beta H_{out}) \), where \( H_{out} \) is the out-hamiltonian in \( \mathcal{I}^+ \), \( \beta = 2\pi/\kappa \), and \( \kappa \) is the surface gravity of the hole. For Schwarzschild this gives a temperature \( T = 1/8\pi M \). A pure state seems to evolve into a mixed state. The hole continues to radiate and heat up. If one envisages the scenario where the hole completely disappears into radiation, at the end we are left with flat space and radiation at infinity, and it seems that effectively the initial state (a pure state) has evolved into a mixed state thus violating the unitary evolution of quantum mechanics. Hawking took the radical point of view that this feature is ubiquitous when black holes are taken into consideration in quantum gravity [8], and that density matrices are evolved by the superscattering operator \( S \):

\[
\rho_{out} = S\rho_{in}
\]

\[
(rho_{out})_{ab} = \sum_{cd} S_{abcd}(\rho_{in})_{cd},
\]

where \( S \neq S \otimes S^\dagger \), i.e., the superscattering operator does not factorize into the tensor product of an \( S \)-matrix and its adjoint. The conservation of probability \( Tr\rho = 1 \) implies that

\[
\sum_A S_{Aacd} = \delta_{CD}
\]

\[
\sum_C S_{ABCC} = \delta_{AB}.
\]

If this were true it would imply a fundamental violation of the principles of Quantum Mechanics. One should certainly beware of extracting this kind of information from semiclassical arguments. The back reaction effects on the geometry are not probably taken into account properly. In [9] one finds a rather simple argument which makes it clear why one should beware of taking semiclassical arguments too far. Imagine an elementary point particle of a given mass, and imagine we build an s-wave wave packet concentrated on a thin spherical shell moving towards \( r = 0 \). We can compute the expectation value of the energy-momentum tensor in this configuration, yielding a distribution of energy and momentum which can be described as an infalling spherical shell. If we couple this energy-momentum tensor to the Einstein equations, it will lead to gravitational collapse, and the formation of a black hole, which in turn will Hawking radiate, leaving at the end flat space with thermal radiation at infinity. This therefore gives a scenario where we evolve a pure state into a mixed state.
However since point particles do not self-gravitate if one could compute all $\hbar$ corrections to this process the outcome would presumably be that the incoming wave-packet reflects through the origin and becomes an outgoing s-wave wave packet, without leading to any lose of quantum coherence.

It is clear that the formation and evaporation of a black hole is a complex process which leads certainly to a practical loss of information, as with all macroscopic phenomena. The point of dispute is the suggestion that when gravitational effects are taken into account, some information about the initial quantum state will be irretrievably lost behind the event horizon, that Hawking's radiation is purely thermal and carries little or no information about the initial quantum state. If the black hole evaporates completely the information would be lost in violation of the rules of quantum mechanics. Soon after this suggestion was made by Hawking, there were other scenarios proposed which would not lead to such radical conclusions (among the proposers are G. 'tHooft and D. Page [5]). Recently some of the ideas in [5] have been sharpened specially within the context of dilaton gravity [3],[9], where the concept of a "stretched horizon" is introduced. The stretched horizon can absorb, thermalize and re-emit information and it is a surface hovering at about a Planck unit of distance above the geometrical horizon. In [9] a number of postulates are formulated in order to deal with these issues. Although there is no proof that these postulates hold in any known theory incorporating gravity at the quantum level, they are reasonably appealing for us to reproduce them here:

Postulate 1. The process of formation and evaporation of a black hole, as viewed from a distant observer can be described entirely within the context of the standard Quantum Theory. There exist a unitary S-matrix describing the evolution from infalling matter to outgoing Hawking-like radiation.

Postulate 2. Outside the stretched horizon of a massive black hole, physics can be described to a good approximation by a set of semiclassical field equations.

Postulate 3. To a distant observer, a black hole appears to be a quantum system with discrete energy levels. The dimension of the subspace of states describing a black hole of mass $M$ is the exponent of the Beckenstein entropy.

Various gedanken experiments (see for instance ref. [9]) suggest that the notion of a stretched horizon is probably the only way we have at present to model the unknown quantum gravitational effects. The simplest way to evaluate the validity of these postulates,
and to build up some confidence in them is to consider simplified models where the decoupling of gravity can be controlled as far as possible. One could envisage the coupling of a free massless field to the Schwarzschild or Reissner-Nordstrom black holes truncated to $s$-wave modes only. The Reissner-Nordstrom metric is given by

$$ds^2 = (1 - \frac{2M}{r} + \frac{Q^2}{r^2})dt^2 - \frac{dr^2}{1 - \frac{2M}{r} + \frac{Q^2}{r^2}} - r^2d\Omega^2,$$

where for $Q = M$ we have the extremal black hole with vanishing Hawking temperature. The problem with this approach is that both the Schwarzschild and Reissner-Nordstrom black holes have throats that become infinitely deep and infinitely narrow, and therefore there is no consistent truncation to an effective two-dimensional field theory including only $s$-waves. A better example is the extremal magnetically charged black hole in dilaton gravity coupled to electromagnetism [10,11]. In this solution there are two asymptotic regions, one the standard one in spherically symmetric black holes, and a second one down the throat which asymptotically looks like $S^2 \times M_{1,1}$, where $M_{1,1}$ is two-dimensional Minkowski space, down the throat therefore it is possible to make a consistent truncation to $s$-waves only and this leads effectively to a two-dimensional theory of dilaton gravity. These configurations which look like a horn stuck in space-time have been called "cornucopions" [12]. This can be considered as inspiration to the dilaton black-hole studied in [3], where one begins directly in two-dimensions by considering dilaton gravity coupled to massless scalar fields. The action takes the form

$$S[f_i, \phi, g] = \frac{1}{2\pi} \int d^2 z \sqrt{-g} e^{-2\phi}(R + 4(\nabla\phi)^2 + 4\lambda^2)$$

$$- \frac{1}{4\pi} \int d^2 z \sqrt{-g} \sum_{i=1}^{N} (\nabla f_i)^2.$$  

This theory is exactly solvable classically. In particular we have the linear dilaton vacuum solution (LDV)

$$ds^2 = dt^2 - dx^2 \quad \phi = -\lambda x,$$

and the two-dimensional black hole metric

$$ds^2 = \frac{dz^+dz^-}{\lambda^4 - \lambda^2 x^+x^-},$$

$$e^{-2\phi} = \frac{M_0}{\lambda} - \lambda^2 x^+x^-.$$  

$M_0$ is the black hole mass, the singularities appear at $x^+x^- = M_0/\lambda^3$ and $x^- = 0, x^+ = 0$.
are the future and past horizons respectively. Furthermore it is possible to construct exact solutions containing collapsing matter leading to the formation of a black hole. The simplest one was studied in [3] where we have an incoming shock wave of conformal matter. In this case the space-time is obtained by gluing the LDV and the black hole solution along the shock wave line. This is probably the simplest possible scenario where by reducing the number of dimensions one keeps the conceptual difficulties intact but reduces drastically the technical complications. In order to handle the back-reaction effects of the Hawking radiation one should compute the quantum effective action and solve the corresponding equations of motion. Even for this simplified model this is an unsolved problem. However the authors of [3] considered the quantum effective action in the limit of a large number of matter fields. This unfortunately works well far away from the singularity, but there is a critical line

$$\phi = \frac{1}{2} \ln \frac{12}{N},$$

where the approximation breaks down [3]. This is the strong coupling region from the point of view of string theory if one remembers that the string coupling is related to the dilaton value by $g_{str} = e^{2\phi}$. Within the context of string theory this means therefore that the original action is just the weak coupling approximation to the full effective action. From a different point of view, the action (7) is a field theory containing scalars; in two-dimensions scalar fields have dimension zero, and therefore one should expect that upon quantization there will be an infinite number of counterterms generated. The theory is barely renormalizable technically. Although some proposals have been suggested to cure this problem in the context of field theory, most of them run into difficulties of other types. It seems that the hypotheses associated with the stretched horizon is at present the more reasonable way of dealing with the problem. Within this framework, the resolution of Hawking’s paradox takes the following form. Consider a black hole which evaporates completely. We can pay attention to different Cauchy surfaces $\Sigma_1, \Sigma_2, \Sigma_3$. The first one is a Cauchy surface before the black hole forms, and therefore it contains no horizons. The second one is chosen after the black hole forms and it extends inside the horizon, and the third one is chosen after the black hole has evaporated completely. $\Sigma_2$ can be divided into two regions $\Sigma_{out}$ and $\Sigma_{bh}$, representing respectively the pieces outside and inside the horizon. If we evolve a state $|\psi(\Sigma_1)\rangle$ the final state will be $|\psi(\Sigma_3)\rangle$, and in the intermediate surface we have $|\psi(\Sigma_{bh})\rangle \otimes |\psi(\Sigma_{out})\rangle$. To avoid having an information loss paradox, all information in the initial state must be obliterated before the state crosses the horizon, in other words the state $|\psi(\Sigma_{bh})\rangle$ must be independent of the initial state of infalling matter. The state inside the black hole is suspiciously unphysical because
it there is no observable outside we could use to give us information about this state. If this state is the same independently on the incoming state, it is clear that the final state after tracing over the black hole Hilbert space will still be pure. This proposal is unfortunately difficult to assess within the context of a concrete model where the approximations can be controlled.

Finally, there is yet another way of looking at the black hole problem wholly within string theory. In [4] E. Witten proposed and exact solution to string theory which at the classical level looks like the two-dimensional black hole of dilaton gravity. Since string theory is supposed to be a consistent theory of quantum gravity, many of the previous questions can in principle be formulated (and answered) within this context. Unfortunately, there are rather hard technical difficulties in obtaining the quantum solution to Witten's model in the language of string field theory, and for the time being we have little information about what answers string theory has in store for us. The basic difficulty is the study of conformal field theories defined in terms of non-compact current algebras.

To summarize, there is a good deal of activity in trying to understand the riddles posed by quantization in the presence of horizons and singularities, and although the complete solution will involve almost certainly some knowledge of quantum gravity, there are encouraging signs that some substantial progress can be achieved before the end of the century. As should hopefully be clear from the previous arguments one of the basic problems is the proper understanding of when we can consider gravity as classical, and deals only with matter fields quantum mechanically. In Quantum Electrodynamics we have a far more detailed understanding of this problem. Using coherent states we can obtain a semiclassical picture of electromagnetic interactions and develop a systematic semiclassical expansion. In gravity this is not so clear. Even if we knew how to write down a quantum theory including gravity and matter fields (a mini-superspace approximation for instance), we do not know on what states we should take expectation values to obtain a reasonable semiclassical expansion where the first term would be equivalent to the coupling of the Einstein equations with the expectation value of the energy-momentum tensor for some matter fields in a gravitational background. In particular, a related problem which appears automatically if one assumes that such an expansion is valid is the problem with the cosmological constant. We do need to understand properly the decoupling of quantum gravitational effects. It would be nice if this happened in our lifetimes.
REFERENCES


