Recent developments in the information loss paradox

Éanna Flanagan, Cornell
eef3-at-cornell.edu

In 1976 Stephen Hawking noted that the formation and evaporation of black holes as described in the semiclassical approximation appear to transform pure states into mixed states, in conflict with unitarity (Hawking 1976). This apparent paradox engendered a lively debate during the subsequent decades, with no generally accepted resolution. A poll taking during a 1993 conference showed 32% of the participants for information loss, with the remainder in favor of unitarity or other resolutions of the paradox (Page 1994). If such a poll were repeated today, however, the result would probably be markedly different — evidence in favor of the preservation of information has emerged from string theory, and more recently from Euclidean quantum gravity (Hawking 2004,2005) and loop quantum gravity (Ashtekar & Bojowald 2005a,b). There is thus an emerging consensus that black hole evaporation is unitary. However, the flaws in the original semiclassical arguments have not yet been clearly resolved; this aspect of the subject remains a puzzle.

The standard semiclassical argument for information loss is as follows (see, eg, Preskill 1992, Strominger 1994). Consider a solution of the semiclassical Einstein equations that describes the formation and evaporation of a black hole with initial mass $M \gg 1$ in Planck units. It is argued that the domain of validity of such a solution consists of all points in whose causal pasts the local curvature is everywhere sub-Planckian; this belief is supported by detailed calculations in two-dimensional models (Bose, Parker & Peleg 1996, Strominger 1994). One can in principle excise from the spacetime all points outside this domain of validity. The resulting spacetime is generally believed to have the following features (based again in part on explicit calculations in two-dimensional models; see eg. Strominger 1994): (i) It contains an apparent horizon whose area shrinks down to the Planck scale due to Hawking emission. (ii) It can be foliated by spacelike slices whose extrinsic curvatures are everywhere sub-Planckian such that one of the slices, Σ, contains most of the outgoing Hawking radiation (mass $\sim M$), yet intersects the infalling matter inside the black hole in a region of low curvature (Lowe et. al. 1995).

Next, one assumes that the black hole evaporates completely, and that after the evaporation spacetime can again be described using the semiclassical approximation. The initial conditions for this final semiclassical spacetime region are determined in part by quantum gravity. One assumes that this region can be foliated by spacelike slices, the first of which, Σ', can be chosen to coincide with Σ outside of some closed two-surface S. Thus, $\Sigma = \Sigma_{\rm in} \cup \Sigma_{\rm out}$ and $\Sigma' = \Sigma'_{\rm in} \cup \Sigma_{\rm out}$, where $\Sigma_{\rm out}$ is the region outside S, $\Sigma_{\rm in}$ is the portion of Σ inside S, which intersects the infalling matter inside the black hole, and $\Sigma'_{\rm in}$ is the portion of Σ' inside S, which contains the end-products of the last stages of evaporation. The construction can be specialized so that the mass contained within S is of order the Planck mass (since the semiclassical description of the evaporation is accurate down to the Planck scale), and so that the radius R of S is of order the time taken for the black hole to decay completely after it reaches the Planck scale (the "remnant lifetime").

If the quantum state on Σ' were pure, then the entanglement entropies $S_{\rm out} = -{\rm tr}(\hat\rho_{\rm out} \ln \hat\rho_{\rm out})$ and $S'_{\rm in} = -{\rm tr}(\hat\rho'_{\rm in} \ln \hat\rho'_{\rm in})$ of the density matrices on $\Sigma'_{\rm in}$ and $\Sigma_{\rm out}$ would agree. If the Hawking radiation is approximated to be in a thermal state, its entropy is of order $S_{\rm out} \sim M^2$ (Page

1976). One might worry that this approximation is not reliable and that subtle correlations between Hawking quanta could make the state on $\Sigma_{\rm out}$ very nearly pure, giving $S_{\rm out} \ll M^2$. However since the state on Σ is pure, $S_{\rm out} = S_{\rm in}$, so a nearly pure state on $\Sigma_{\rm out}$ would require a nearly pure state on $\Sigma_{\rm in}$. This would contradict the linearity of quantum mechanical evolution from initial states of the collapsing matter to states on Σ (Preskill 1992). Therefore the estimate $S_{\rm out} \sim M^2$ seems unavoidable. Finally, if the lifetime R is assumed to be of order the Planck time, the region $\Sigma'_{\rm in}$ is too small and contains too little energy to allow significant correlations, $S'_{\rm in} \lesssim 1$. [Preskill (1992) shows that $S'_{\rm in} \sim M^2$ would require a lifetime of order $R \sim M^4$.] Therefore $S'_{\rm in} \ll S_{\rm out}$ and the state on Σ' cannot be pure.

This semiclassical argument indicates that the information defined by $I = \ln \dim \mathcal{H} + {\rm tr}(\hat{\rho} \ln \hat{\rho})$ decreases, where $\mathcal{H}$ is the Hilbert space. However, as pointed out by Page (1994), it does not necessarily imply that information is lost in the colloquial sense that knowledge of the final (mixed) state on Σ' is insufficient to compute the initial (pure) state: pure-to-mixed evolutions can be described by invertible superscattering matrices. Nevertheless, if one accepts that I decreases, then information loss in the stronger, colloquial sense seems inevitable, since the state on $\Sigma_{\rm out}$ as computed in the semiclassical approximation depends only very weakly on the initial quantum state of the infalling matter.

Many different resolutions to the information loss paradox have been proposed; see eg. Thorlacius (2004) or Page (1994) for detailed reviews. Some novel recent suggestions are those of Horowitz & Maldacena (2004) and Gambini et. al. (2004). In the remainder of this article I will focus on evidence from various approaches to quantum gravity that black hole evaporation is unitary.

The evidence from string theory is well-known; see the articles by Gary Horowitz in issues 12 and 18 of Matters of Gravity. Briefly, the AdS/CFT duality of string theory implies that the formation and evaporation of small black holes in anti-deSitter space can described completely (albeit indirectly) by a manifestly unitary conformal field theory. String theory also provides an explanation for the failure of the semiclassical argument summarized above. Namely, the low energy effective theory describing evolution in the foliation used in the semiclassical argument is *non-local* even in low-curvature regimes: it contains states of strings stretched over macroscopic distances (Lowe et. al. 1995, Lowe & Thorlacius 1999). This non-locality allows external observers to measure the information coming out in the Hawking radiation, while infalling observers see nothing unusual happening while the infalling matter crosses the horizon. The idea of such low-energy non-locality is anathema to many relativists (see, eg. Jacobson et. al. 2005), but is supported by computations of commutators of spacelike separated operators in string theory (Lowe et. al. 1995). Thus, the key assumption that semiclassical general relativity is a good approximation at sub-Planckian curvatures is invalid in string theory.

Turn now to the Euclidean path integral approach to quantum gravity. For many years, Stephen Hawking has been the most prominent proponent of the view that black holes destroy information. His dramatic and much-publicized reversal on this issue last year at the GR17 conference (Hawking 2004) was described in Brien Nolan's article in issue 24 of Matters of Gravity. Hawking considers scattering processes in anti-deSitter space that classically would be expected to produce a black hole. He argues in this context that a path integral computation of correlation functions of operators at infinity will be dominated by a sum over two topological sectors, one corresponding to no black hole being present and one corresponding to a black hole being present. (Since observers making measurements at infinity cannot be sure if a black hole formed or not, the corresponding amplitudes should be summed.) The contribution from the topologically trivial sector (no black hole), which he had previously neglected, is sufficient to restore unitarity (Hawking 2005). Hawking's arguments so far are schematic and will likely be followed up by more detailed computations. In any case, it is clear that his arguments do not yet explain where the standard semiclassical argument breaks down, if it does.

Turn lastly to loop quantum gravity. Here, the first key result relevant to information loss is that the singularity inside Schwarzschild black holes is resolved (Ashtekar & Bojowald 2005b). The analysis is similar to earlier analyses of the resolution of cosmological singularities (Bojowald 2001, Ashtekar et. al. 2003). In the Schwarzschild case the theory that is quantized is a a finite-dimensional, minisuperspace model rather than the full, infinite-dimensional theory. However, the particular quantization chosen (which is unconventional) is motivated by considerations of the loop quantum gravity program on the full phase space. With this quantization, the Hamiltonian constraint reduces to a finite difference equation, which can be thought of as describing a discrete time evolution. For appropriate solutions of the difference equation, operators describing the geometry are well defined at the singularity, where they diverge classically. While one might object to the truncation of the phase space it is plausible that similar results might hold in the full theory.

Ashtekar & Varadarajan (2005) also analyze the more complicated model of the two-dimensional CGHS black hole (Callan, Giddings, Harvey & Strominger 1992), which incorporates Hawking radiation and backreaction. The key results here are (i) the singularity is again resolved, and (ii) there exists a region of the spacetime surrounding what would be the singularity where quantum gravitational effects are strong and a classical geometry does not exist. Before this region, and again after it, a semiclassical approximation is valid. In particular there are no baby universes, and no superpositions of macroscopically different geometries at late times. Ashtekar & Bojowald (2005a) argue that these conclusions should also be valid in four-dimensions If this is the case, then the fact that the singularity is resolved should imply that black-hole formation and evaporation is unitary, since there are no singularities in the quantum theory.

How is this loop quantum gravity scenario consistent with the standard semiclassical argument for information loss? The answer to this question is not yet completely clear. Ashtekar and Bojowald (2005a) speculate that the amount of mass that passes through the singularity (the mass contained inside the surface S discussed above) could be much larger than the Planck mass. This would require the semiclassical solution that describes the evaporation outside the black hole to break down before the Planck regime is reached, or to have properties other than those usually assumed (for example, an apparent horizon area of order the Planck area at the same retarded time as a Bondi mass much greater than the Planck mass; Ashtekar 2005). This option cannot be ruled out since a complete and detailed computation of the semiclassical solution has not yet been carried out in four dimensions; however it seems unlikely given the two-dimensional computations that have been performed. If it is true that the mass that passes through the singularity is much larger than the Planck mass, then the semiclassical argument for information loss can be evaded.

One of the appealing features of the loop quantum gravity computations, in my view, is that they are sufficiently explicit, direct and local that it should be possible with additional computations to pin down where and why the semiclassical argument fails. By contrast, the Euclidean quantum gravity approach restricts itself to computing asymptotic observables;

it is difficult using only these observables to try to understand why the local semiclassical arguments fail. Similarly, in string theory the most detailed understanding of the evaporation process is in terms of the dual description in the boundary conformal field theory; it is not easy to translate this into a detailed local understanding of the evaporation process (although see Lowe & Thorlacius 1999).

The most satisfying resolution of the information loss paradox, in my view, would be an explanation of why the semiclassical theory breaks down earlier than naively expected. Ideally this explanation would use only minimal assumptions about quantum gravity, and would be applicable to all three of the approaches to quantum gravity discussed here. The recent loop quantum gravity computations might be a step in this direction.

References:

Ashtekar, A., 2005, private communication.
Ashtekar, A. & Bojowald, M., 2005a, Class. Quant. Grav. **22**, 3349 (also gr-qc/0504029).
Ashtekar, A. & Bojowald, M., 2005b, *Quantum geometry and the Schwarzschild singularity*, gr-qc/0509075.
Ashtekar, A., Bojowald, M. & Lewandowski, L., 2003, Adv. Theor. Math. Phys. **7**, 233.
Ashtekar, A. & Varadarajan, M., 2005, in preparation.
Bojowald, M., 2001, Phys. Rev. D **64**, 084018
Bose, S., Parker, L. & Peleg, Y., 1996, Phys. Rev. D **53**, 7089.
Callan, C. G., Giddings, S. B., Harvey, J. A. & Strominger, A., 1992, Phys. Rev. D **45**, R1005.
Gambini, R., Porto, R. & Pullin, J., 2004, Phys. Rev. Lett. **93**, 240401.
Hawking, S.W., 1976, Phys. Rev. D **14**, 2460.
Hawking, S.W., 2004, presentation at 17th International Conference on General Relativity and Gravitation, Dublin, Ireland; transcript available at http://pancake.uchicago.edu/˜carroll/hawkingdublin.txt.
Hawking, S.W., 2005, *Information Loss in Black Holes*, hep-th/0507171.
Horowitz, G. & Maldacena, J., 2004, JHEP **008**, 0402 (also hep-th/0310281).
Jacobson, T., Marolf, D. & Rovelli, C., 2005, *Black hole entropy: inside or out?*, hep-th/0501103.
Lowe, D.A., Polchinski, J., Susskind, L., Thorlacius, L. & Uglum, J., 1995, Phys. Rev. D **52**, 6997 (1995).
Lowe, D.A. & Thorlacius, L., 1999, Phys. Rev. D **60**, 104012.
Page, D.N., 1976, Phys. Rev. D **14**, 3260.
Page, D.N., 1994, *Black Hole Information*, in “Proceedings of the 15th Canadian Conference on General Relativity and Relativistic Astrophysics”, Eds. R.B. Mann and R.G. McLenaghan, World Scientific, Singapore (also hep-th/9305040).
Preskill, J., 1993, *Do black hole destroy information?*, in “International symposium on black holes, membranes, wormholes and superstrings”, World Scientific, River Edge, NJ (also hep-th/9209058).
Strominger, A. 1994, *Les Houches lectures on black holes*, hep-th/9501071.
Thorlacius, L. 2004, *Black Holes and the Holographic Principle*, hep-th/0404098.