Confusions and questions about the information paradox

September 17, 2009

1 Steps in the Hawking argument

Here is a list of the steps that Hawking would use to argue that black holes pose a problem for quantum mechanics. Thus to resolve his problem, we have to point out which step is incorrect, and give a concrete reason why.

- Start with a low density shell of matter of large radius. The shell is in a pure state $|\psi_M>$, and collapses towards the center under its own gravity. The rest of spacetime is empty.

- Follow this evolution using a set of 'nice slices'; i.e. slices that satisfy the conditions N1-N5. The shell moves to smaller radii, and particle pairs (b, c) get created.

- The black hole mass must decrease because of the radiation, so we should make the black hole geometry change slowly to reflect this fact. But because the radiation happens slowly, for the emission of any given quantum we can just use the metric for a black hole of mass $M(t)$, in the region where the pair production is taking place.

- The state of the created pairs is such that there is a large entanglement entropy S between the radiated quanta $\{b\}$ and the $(M, \{c\})$ quanta in the hole.

- When the evaporation has proceeded to the point where the black hole mass is say $100 l_p$, we stop our discussion, Quantum gravity has not been strong anywhere. There are only three options:

  (A) The hole completely evaporates, in which case quantum theory is violated because we cannot assign any state to the $\{b\}$ quanta

  (B) A remnant of mass $\sim m_p$ remains. In this case we have great trouble with the fact that there is an infinite degeneracy of remnants which have bounded size and bounded mass

  (C) Information leaks out in small correlations among the radiated quanta, just like what happens when we burn a piece of paper. But this cannot happen if there are small corrections to the radiation process; we need corrections of order unity. This means that the evolution of quantum fields at the horizon will *not* be given by the expected evolution in gently curved space. Thus we would be saying the the black hole geometry is incorrect, and we do not form a traditional black hole in the theory.

2 Attempts to resolve the problem

Here we list some common attempts to resolve the paradox (given in italics), and below each, we state why the argument does not help.

2.1 Avoiding black holes

As seen from infinity, a collapsing shell never appears to fall through the horizon. So a black hole never forms, and there is no problem.

It is true that classically the shell is never seen to fall through the horizon. But the Hawking paradox is created by quantum modes that straddle the horizon; one half comes out and the other falls in, leading to an entangled state. Thus it does not help to argue that in classical evolution in Schwarzschild coordinates the shell does not fall in.

There is no problem for an observer sitting at infinity, because in his Schwarzschild coordinates he sees an infinite temperature at the horizon, which will cause all infalling matter to burn up and radiate its information out.

It is not all all clear how to complete such an argument. The Schwarzschild coordinates become bad at the horizon, and one consequence of this is the diverging redshift observed for all objects near the horizon. If we are to make any arguments in this Schwarzschild frame, then we will have to be accurate enough to do all the physics correctly in this bad coordinate frame. The inside of the hole is not covered by these coordinates, so it is not clear how that part of the geometry is going to be covered; quantum wavefunctions can straddle both sides of the horizon. In the Kruskal coordinate frame the infalling matter does not burn up; it just passes smoothly through the horizon. Thus it is unclear how one can reconcile to a different answer in the Schwarzschild frame. One would expect that the large fluctuations seen at the horizon in the Schwarzscild frame would have to cancel out in their net effects, though it would be hard to see this explicitly in most cases.

The black hole horizon region will be different in subtle ways depending on what fell into the black hole. Thus the Hawking evolution will be corrected in a subtle way, and information will come out.

People looked for small corrections to the Schwarzschild solution. This was called looking for 'hair' on the black hole. But no hair was found: if one solves the linearized wave equation for perturbations on the black hole background, then one finds that there are no nonsingular solutions.

2.2 The equivalence principle

Suppose I am falling into a black hole. By the equivalence principle I will see nothing at the horizon. Thus there cannot be any information about the black hole state at the horizon.

This is a completely circular argument. The equivalence principle does not say that we will fall freely through the black hole horizon. For example, suppose we are falling towards the earth. At first we will seem to be following the equivalence principle because we move along a geodesic. But when we reach $r = 4000$ miles, we will hit something solid, and bounce off. This does not violate the equivalence principle because every body has a size, and when we reach that size the geodesic motion under gravity will have to change. In a black hole we do not know a priori what the 'size' is. If it happens to be horizon radius, then we will *not* fall through a vacuum region at the horizon.

2.3 Thermality

The spectrum of Hawking radiation is thermal. Thermal states have no information, so information must be lost.

There is a lot of confusion in the use of the word 'thermal'. Do we mean that the spectrum is the planck black body spectrum? Black holes have significant 'grey body factors', so the spectrum is actually far from planckian. The actual spectrum of the radiation is completely irrelevant; all that is important is that the state of the produced pairs is entangled between the inside and outside of the hole. Do we mean that a thermal state is 'mixed', described only by a density matrix? The state on the complete Cauchy slice is a pure state, and we get a mixed state only when we trace over the inside of the hole. This would happen in *any* system: if we trace over on part, the other part becomes a mixed state. In black holes the problem comes when the inside of the hole disappears, leaving the entire state on a complete Cauchy slice to be a mixed state; this never happens in usual quantum theory.

2.4 Small corrections to Hawking radiation

When we burn a piece of paper, the information is encoded in subtle correlations between the emitted photons, so it is hard to see. Perhaps small corrections to Hawking's leading order computation introduce similar correlations, an information comes out in the Hawking radiation.

We have seen by an explicit computation that entanglement between the inside and outside of the hole keeps increasing, unless we have order unity corrections to the evolution of modes at the horizon. But such an order unity correction implies that we do not have a traditional black hole horizon.

Instanton effects should be included, and these can resolve the problem.

Certainly instanton effects can be considered. But an instanton is just a convenient way of summarizing in a Euclidean computation a set of effects in the Lorentzian section. So even if can achieve something with instantons, to solve the paradox we will have to show where the Lorentzian argument of Hawking breaks down. Secondly, is we write instanton effects as $\sim e^{-S}$, where S is the action of the Euclidean black hole, then this is a very small effect; on the other hand we have seen that we need an order *unity* effect on the created quanta to solve the problem. So it is not clear that just invoking the term 'instantons' can help avoid the paradox; we will have to do more.

2.5 AdS/CFT and the information paradox

We can make a black hole in AdS. Since AdS is dual to a CFT, and the CFT is unitary, there cannot be any information loss, and so there is no information paradox to solve in string theory.

This is again a completely circular argument. AdS/CFT duality is arguably one of the most interesting insights to emerge from string theory. It is also a very useful tool in understanding black hole behavior. But we cannot simply invoke this duality to bypass the information paradox. Since this is a very common confusion among students of string theory, we present it as the following discussion:

Student: I dont see why I should worry about Hawking's paradox. Now that we know that gravity is dual to a CFT, and the CFT is unitary, there can cannot be any information loss, and so there is no problem.

Hawking believer: That is an entirely circular argument, as I can easily show. Suppose I say: Quantum mechanics is unitary, so there can be no information loss. Would I have resolved Hawking's paradox?

Student: No, that would be silly. Hawking agrees that quantum mechanics is valid in all laboratory situations. All he argues is that once we make a black hole, then quantum mechanics is violated. So we cannot use our tests of quantum mechanics in the everyday world to argue that there will be no problem when black holes form.

Hawking believer: Good, that is correct. So now me let me ask the same question about AdS/CFT. You have computed the spectrum , 2-point functions, 3 point functions etc. and found agreement between the CFT and gravity descriptions. I understand that you have numerous such computations. But these processes do not involve black hole formation, and so do not address Hawking's argument. Is that correct?

Student: Yes, that is correct. But we also have a black hole solution, called AdS-Schwarzschild, which is similar to the standard Schwarzschild metric in its essential respects.

Hawking believer: Excellent. So I will now apply the Hawking theorem, proved in the above sections, and *prove* that normal assumptions about locality gives mixed states/remnants. Since your black hole has an 'information free horizon' just like the Schwarzschild hole, my arguments go through in exactly the same way. Thus you have three choices: (a) You can tell me why local Hamiltonian evolution breaks down under the niceness conditions N (b) you can agree to mixed states arising from pure states, which violates quantum theory; in that case you lose AdS/CFT and string theory as well, since these are built on a foundation of usual quantum theory (c) You can agree to have remnants in your theory, and explain why they do not cause the problems that people feared. Now which will it be?

Student: I don't know ... I see that you have forced me into a corner by using the Hawking theorem, and I will have to work as hard to solve it in my AdS case as I would have had to in the usual asymptotically flat case. So let me try to evade the problem by trying a different argument. I will use the CFT to *define* my gravity theory. Then I will get a gravity theory that has the expected weak field behavior, and I will never violate quantum mechanics, and I can never get information loss.

Hawking believer: Excellent. With this definition of your gravity theory, you will by construction never have the 'mixed state' possibility in Hawking's theorem. So now tell me: (a) Will you claim that traditional black holes do not form in this gravity theory (b) The black hole horizon forms, but the niceness conditions N do not give locality; in this case you should be sure to tell me how this happens and what niceness conditions you will add to recover conditions for the solar system limit (c) Do neither of the above but say that the theory has long lived remnants.

Student: Well ... I always assumed that I could have a normal black hole horizon, usual notions of niceness conditions, and still get all the information out in Hawking radiation so there are no remnants. But I see now that the Hawking theorem forbids exactly this possibility. I dont know how I can say anything about the options you list without studying the black hole formation/evaporation process in detail in either the CFT or the gravity theory.

Hawking believer: Exactly. You are welcome to do your analysis in either the CFT or the gravity theory, but at the end you must show me what happens when a black hole forms and evaporates in the gravity description.

Student: I see now that to solve Hawking's paradox I will have to understand the interior structure of the black hole. I cannot get by with any abstract arguments like 'AdS/CFT removes the paradox.'

Hawking believer: Exactly; in fact abstract arguments in general cannot distinguish between whether locality broke down and information came out in the Hawking radiation or if information leaked out from a long lived remnant. Solving the information paradox implies that you tell us which happens, and if you want the information to come out in the radiation, to show explicitly the process by which 'solar system physics' broke down while the niceness conditions N were still valid.

2.6 Violation of classical intuition

We observe black holes in the sky, in the sense that we see accretion into them in quasars etc. Would all this be violated if we changed our picture of the black hole in a basic way?

There are two issues here. One is the region over which the changes happen; if this is confined to within the horizon radius, then there are no observable effects outside. The other is the time scale involved in our experiment. There are two very different time scales in the black hole: the crossing timescale and the Hawking evaporation timescale. It is plausible that all microstates of the black hole will behave just like the traditional black hole over the crossing timescale, but will differ from each other (and emit Hawking radiation differently) over the Hawking evaporation timescale.