BLACK HOLES AND INFORMATION LOSS*

T. Banks

Department of Physics and Astronomy

Rutgers University

Piscataway, NJ 08855-0849

In this lecture, the author's point of view on the problem of Hawking Evaporation of Black Holes is explained. A possible resolution of the information loss paradox is proposed which is fully in accord with the rules of quantum mechanics. Black hole formation and evaporation leaves over a remnant which looks pointlike to an external observer with low resolving power, but actually contains a new infinite asymptotic region of space. Information can be lost to this new region without violating the rules of quantum mechanics. However, the thermodynamic nature of black holes can only be understood by studying the results of measurements that probe extremely small (sub-Planck scale) distances and times near the horizon. Susskind's description of these measurements in terms of string theory may provide an understanding of the Bekenstein-Hawking (BH) entropy in terms of the states of *stranded strings* that cross the horizon. The extreme nonlocality of string theory when viewed at short time scales allows one to evade all causality arguments which pretend to prove that the information encoded in the BH entropy can only be accessed by the external observer in times much longer than the black hole evaporation time. The present author believes however that the information lost in black hole evaporation is generically larger than the BH entropy and that the remaining information is causally separated from the external world in the expanding horn of a black hole remnant or *cornucopion*. The possible observational signatures of such cornucopions are briefly discussed.

* Supported in part by the Department of Energy under grant No. DE-FG05-90ER40559.

1. Introduction—The Facts in the Case

The subject that I am going to talk about in these lectures, the Hawking evaporation of black holes, has been with us for on the order of 19 years now. Although the last few years have seen an upsurge of interest and activity in the subject, it remains a frustrating field which stubbornly refuses to yield a satisfying resolution of its paradoxes. The number of very good physicists who have expressed fairly definitive opinions about the resolution of the Hawking puzzle is smaller than the number of definitive opinions they have expressed. The frustration is compounded by the fact that there is no hope for experimental resolution of the confusion. I would guess that there is still a sizable body of physicists and astronomers who remain unconvinced of the observational evidence for the existence of black holes. Given that they exist, the probability that we are ever going to examine a black hole close up seems very small. Even if we could examine one close up, the probability that we would happen to observe it at a time when it was emitting substantial amounts of Hawking radiation is nil. And most frustrating of all, even those physicists who insist that black hole radiation is not "really" incoherent agree that it looks like thermal radiation for all practical purposes.

The title of this section implies a certain degree of objectivity. In a field such as this, true objectivity is impossible. I will therefore be presenting "the facts" in a way which emphasizes that part of the data that supports my present opinions. It is best then to get those opinions out on the table, so that you can judge for yourself what they are worth and how much they are distorting my presentation of the facts. In brief: I believe that the Hawking evaporation of black holes terminates in stable remnants. An angular slice of the geometry of those remnants is shown in Fig. 1.

The remnant is a small "hole in space" attached to a semi-infinite horn which has a spherical cross section of small radius. These static geometries have a unique quantum ground state, but they are the remains of evolving geometries possessing a horizon which moved off to an infinite spacelike distance. The Penrose diagram of the full spacetime of one of these remnants is shown in Fig. 2.

There can be many different quantum states propagating behind the horizon of these remnants, but the external geometry and ADM mass of all of these states is the same.

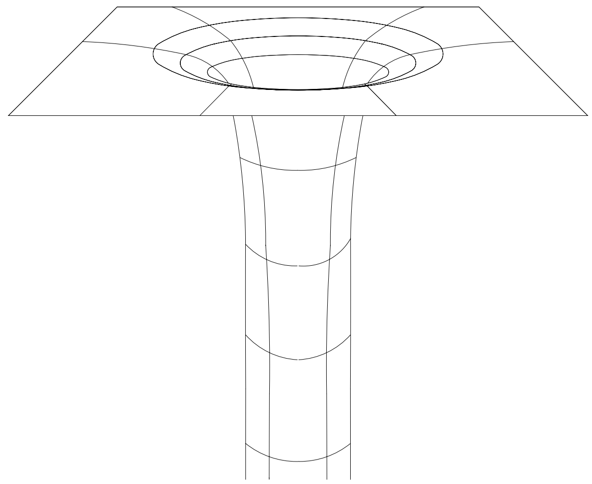

Fig. 1: An angular slice of the static geometry of a black hole remnant.

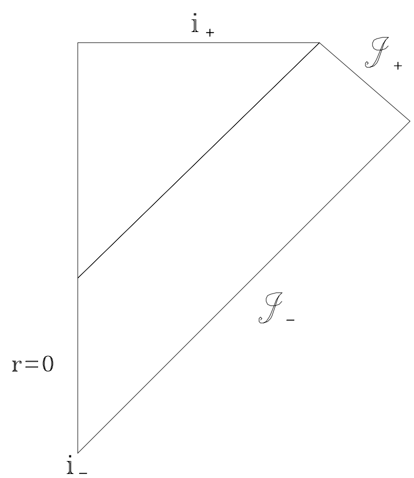

Fig. 2: The Penrose diagram of a black hole remnant.

These states form the repository for all of the information "lost" in Hawking evaporation. The full spacetime has (in the approximation in which quantum fluctuations of the gravitational field can be treated perturbatively) a unitary quantum mechanical evolution, but the past and future contain different numbers of causally separated spatial asymptotic regions. The S-matrix for transitions from the initial asymptotic region back to itself is not unitary. This picture by itself may seem like a resolution of the information loss puzzle, and for a while, I believed that it was. However, neither it nor any other discussion of information that is truly lost to the external observer can account for the BH entropy. This is a true thermodynamic entropy which describes interactions between the black hole and the external observer. As such, it must be associated with degrees of freedom which are causally connected to, and can interact with, the external observer. By assigning the entropy to correlations with degrees of freedom which are out of causal contact with the external observer, or in a topologically disconnected universe, one gives up the possibility of explaining its thermodynamic nature. I will argue in more detail below that the location of the BH degrees of freedom must be in an extremely tiny region in the vicinity of the horizon of the black hole. In discussing them, one is inevitably led into extremely short distance physics. I believe that Susskind's string model of these degrees of freedom, while still in a primitive stage of development, may lead to an ultimate explanation of the BH entropy. If this is the case, then the information represented by the BH entropy is not lost to the external observer.

My picture of Hawking evaporation then includes in some manner all of the current theoretical prejudices about the subject. I believe in remnants of black hole evaporation, but all information stored in them is causally inaccessible to the external observer. The S-matrix for a single asymptotic spatial region is not unitary once a black hole has formed. However, the information which absence is quantified by the BH entropy is not irretrievably lost. It is located on the horizon of the black hole, probably in the form of *stranded strings* (see below for a definition), and will be radiated into the original asymptotic region as the black hole evaporates.

1.1. Some Classical Facts

For the author at least, the only way to get intuition about what is going on in general relativity is to work in synchronous gauge. This is a name for any one of a collection of coordinate systems in which "time is time," and general relativity is a theory of the dynamics of spatial geometry. A synchronous gauge is chosen by picking a spacelike hypersurface and defining time to be the geodesic distance orthogonal to this hypersurface.

The time evolution of the geometry of the spatial sections of a neutral black hole is shown in Fig. 3.

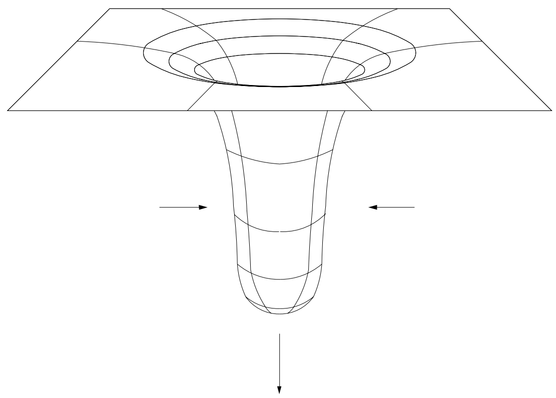

Fig. 3: The spatial geometry inside a Schwarzschild black hole evolves into an infinite, and infinitely thin, horn.

Note that the geometry has the "hornlike" shape of Fig. 1, but the horn is dynamic. It stretches along its length and shrinks transversely, reaching infinite length and zero width in finite proper time. This is the celebrated Schwarzschild singularity. It has two aspects: the width goes to zero, and the infinite length is achieved in finite time. Both of these correspond to curvature singularities.

From these pictures, it is easy to understand why the black hole has a horizon. In Einstein's general relativity, space is allowed to stretch faster than light can cross it. A light beam sent out from some object may never be able to get back to its point of origin because in the time that it has travelled, the space has expanded (and continues expanding). It is clear then that if we look at any point a finite distance inside the horizon on some time slice, the space between it and any point outside the horizon will expand so rapidly that light will not be able to travel between them. Note that even if the singularity were absent, and the transverse size of the horn remained finite, we could still have superluminal expansion of the interior geometry and would be forced to conclude that the system had a horizon.

The synchronous view of the black hole interior suggests a possible scenario for its nonsingular evolution. Namely, all that is necessary is to find a mechanism for stabilizing the horn against transverse collapse, and for slowing down the expansion in the radial direction so that the rate never becomes infinite. We will see below that this is precisely what happens for extremal charged black holes. One thing is clear about any such hypothetical mechanism for stabilization of the black hole. If we believe (as we should) that the Schwarzschild solution for a large mass black hole is valid down to times at which the transverse dimension has shrunk to microscopic size, then the radial extent of the geometry will be very large. In fact, as we shall see, the example of extremal charged black holes suggests that perhaps the geometry keeps on growing in the radial direction, even when the transverse collapse is averted. We are thus led to a picture of black hole final states as large one-dimensional protrusions on the geometry of space, which connect onto the space of the asymptotic observer through essentially pointlike openings. From many points of view, the external observer will regard these as point particles, but we will argue that this is a mistake when it comes to quantum mechanics.

We now turn to another well-known classical feature of black holes, the discrepancy between the description of black hole physics given by infalling observers, and the description given by those observers, like the observer at infinity, who are supported in the gravitational field and remain outside the horizon. When the black hole mass is large, the infalling observer experiences nothing in particular as she falls through the horizon. To the asymptotic observer on the other hand, the horizon is a very peculiar place. Nothing seems to fall through it. This can be ascribed to the behavior of the g_{00} component of the metric in Schwarzschild coordinates. Near the horizon, the metric has the form

$$-\frac{x}{2\ GM}dt^2 + {\frac{x}{2\ GM}}^{-1}dx^2 + (2\ GM)^2 d\Omega^2 \tag{1.1}$$

where $x = r - 2\ GM$. The proper time per unit Schwarzschild coordinate time goes to zero as the horizon is approached. This is a Lorentz contraction. A supported observer near the horizon is accelerating like mad to prevent himself from falling through. The instantaneous boost relating his frame to that of the infalling observer is extremely large, of order $e^{\frac{t}{4M}}$, where t is the Schwarzschild time coordinate. For a large mass black hole, even moderate energies will be boosted far above the Planck energy by such a transformation. Thus the Schwarzschild observer sees the infalling observer as highly Lorentz contracted. From his point of view, the geodesic observer's clocks run very slowly and the structure in her machinery has very little extent in the radial direction.

Now suppose that the asymptotic observer has a measuring apparatus of limited accuracy, which cannot measure arbitrarily small intervals of space or time. Very quickly, the infalling observer reaches a point at which the full extent of her apparatus in the radial direction is apparently squeezed into less than a minimally measurable distance from the horizon. All normal processes in the infalling observer's frame are slowed down so much that the external observer cannot discern anything changing with time over the minimal interval between ticks of his clock. Thus, a classical observer with limited powers of observation quickly loses all information about what is going on in the infalling observer's frame. The traditional "covariant" view of relativity is that this is a consequence of a bad choice of coordinates, but in the past few years, a new paradigm has developed which takes the Schwarzschild observer's point of view as the basis of the treatment of the astrophysics of black holes, and their interactions with the external environment. The Membrane Paradigm [1], as it is called, claims that all interactions of the black hole with the external world can be treated correctly by a model in which the black hole behaves *as if* there were a physical membrane with charge, current, and energy densities on it hanging on a timelike surface just above the horizon of the hole. The classical aspects of this picture can be *derived* from general relativity, but it has also been applied to the treatment of Hawking radiation. We shall see that such a picture for Hawking radiation seems to contradict the axioms of quantum field theory, but that it may be derivable from the dynamics of strings.

There are two important features of this classical picture that we will want to remember later when we discuss Susskind's conjectures about the nature of the BH entropy. The first is that the Schwarzschild observer's attempts to understand what is happening to his infalling colleague as she approaches the horizon require him to contemplate measurements of arbitrarily small length and time intervals. Thus, a proper theory of these measurements requires us to understand physics at the shortest distances. Secondly, although the Schwarzschild observer's picture is an "incorrect" picture of what is "really happening" to the infalling observer, it is a perfectly sensible account of everything that the Schwarzschild observer can actually measure. In the end, I believe that this will be the sort of situation that we will recover for the quantum theory of black holes, at least close to the semiclassical limit of large mass. Our picture of "what is really going on" in Hawking evaporation will be the formation of a remnant and the disappearance of particles behind a receding horizon. However, the asymptotic observer will be able to account for much of what he observes in terms of a gas of "stranded strings" glued to the membrane on the *stretched horizon* [2]. I will suggest below that there is probably some real information loss to the asymptotic observer, but that the thermodynamic entropy of Bekenstein and Hawking represents information that can in principle (though certainly not in practice) be retrieved by him.

1.2. Quantum Facts

Here I will briefly review the salient facts about the theory of Hawking radiation. I assume that the listener/reader is already familiar with this material, and I will only emphasize some important facts that are not usually presented. Hawking's calculation is carried out in the framework of an approximation to quantum gravity called *quantum field theory in curved spacetime*. One imagines the formation of a black hole by a classical matter distribution falling in from infinity in an initially flat spacetime. The first quantum correction to this classical process consists of quantizing the linear fluctuations of quantum fields around this classical solution. Hawking computed the S-matrix for this linear field theory. What does this involve? The classical geometry has a perfectly well-defined past asymptotic region. The future, however, consists of two causally disconnected asymptotic regions (the original one, and a region "down the horn" in the synchronous gauge picture), one of which becomes singular a finite time in the future. Hawking's idea was to treat the singular region "as if" it had a well-defined set of asymptotic states. We can then consider *inclusive cross sections* in which the external observer measures only what is causally accessible to him, summing over the unknown final states on the other side of the horizon. As usual, this will lead to a density matrix description of the final state that he measures.

The expansion parameter for this semiclassical approximation is the Planck mass divided by the mass of the black hole. In the limit of large mass, the spacetime curvature is small (order $\frac{1}{(GM)^2}$) everywhere except for the singularity.[1] Furthermore, the singularity is a large timelike geodesic distance (of order GM) from any point on the horizon.

According to Hawking's calculation [3], the expression for the outgoing density matrix is

$$\rho = e^{-\frac{16\pi^2 GMH}{h}}, \tag{1.2}$$

where h is Planck's constant. The black hole is thus seen to behave like a black body with temperature $T_H = \frac{h}{16\pi^2 GM}$. Since its energy is M, the first and second laws of thermodynamics give it an entropy $S_{BH} = \frac{8\pi^2 GM^2}{h}$ with the usual ambiguity of an additive constant. Note that this entropy is proportional to the area of the horizon $S_{BH} = \frac{1}{4}AM_P^2$. According to the Stefan-Boltzmann Law, and taking into account that the area of the horizon is $\sim M^2$, the black hole should lose energy at a rate $\frac{dM}{dt} = -\frac{1}{M^2}$, giving it a lifetime of order M^3. All of this is for a neutral nonrotating black hole. In general, the black hole temperature depends on its mass, angular momentum, and charge. In particular, for near extremal charged black holes, the lifetime is linear in the deviation from extremality.

The paradox of all this is that the black hole seems to decay into incoherent radiation. Below, we will review the argument that suggests that in standard quantum field theory, the decay of the hole proceeds incoherently until a time when its energy content is very small compared to the amount of information that it has yet to liberate. This appears to lead to a choice between three alternative scenarios for the climax of the radiation process, all of which appear to lead to paradoxes. We will enumerate and discuss them below.

The other key question raised by this discussion is the origin of the BH entropy $S_{BH} = \frac{1}{4}AM_P^2$. It does not seem to come from a counting of states. Below, I will discuss Susskind's proposal for calculating the BH entropy in string theory.

[1] In particular, the curvature is small near the horizon.

2. The Threefold Way

2.1. Subtle Correlations and Causality

At first sight, the most conservative approach to the problem of information loss is that which goes under the name of "subtle correlations." According to this dogma, the S-matrix for black hole formation and evaporation is unitary in the Hilbert space of the original asymptotic observer. The apparent loss of coherence exhibited in Hawking's calculation is ascribed to the inadequacy of his semiclassical approximation. The standard analogy is to the heating of a lump of coal: Suppose that we encode the information in the *Encyclopedia Brittanica* in Morse code and send it out as flashes of laser light that are directed at a large lump of coal. The laser flashes are absorbed by the coal, which heats up. All of the information in the *Encyclopedia* is now contained in the coal. Of course, the heated coal emits infrared radiation and eventually cools down. After it has cooled, the information is stored in the heat that it radiated, but for all practical purposes, this radiation is thermal and the information has been lost.

In this situation, we understand what is going on. The radiation from a cooling lump of coal is not really thermal; it is in a pure, albeit very complicated, quantum state. The useful information originally stored in the pulsed laser beam is now encoded in correlations between photons which were emitted from the coal at very different times and which are therefore very far from each other in space. This nonlocally stored information is of no practical use, and for local measurements, the pure state is equivalent to a mixed state. Is this all that is going on in the Hawking calculation?

There is a very strong argument that this is not the case, at least not within the conventional formalism of quantum field theory. The semiclassical picture of Hawking evaporation is valid for most of the evaporation of a large black hole. In particular, if we have an enormous black hole which has evaporated away 99% of its mass, leaving behind a hole which is still large, the Hawking calculation will be accurate to the past of the asymptotically null spacelike slice labelled 99 in the Penrose diagram of Fig. 4.

We can examine the state of that portion of the world which is behind the horizon. In the semiclassical approximation, this will be calculable and correlated with the state of

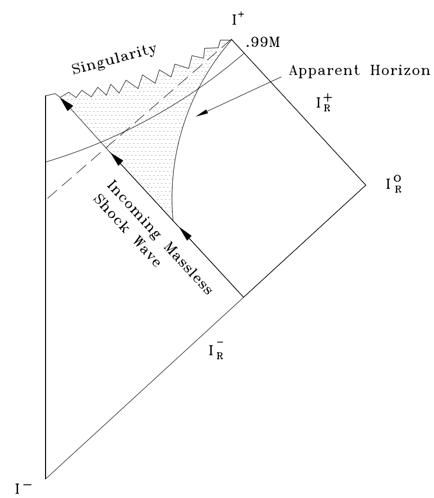

Fig. 4: Penrose diagram illustrating the argument that information cannot get out of a black hole until times longer than the Hawking lifetime.

the outside world. Observers falling through the horizon, but still not at the singularity, do not notice any particularly violent events. The interior state will be impure, with large entropy, determined by the correlations (in the initial state) between objects that fall into the black hole and those that do not. If we assume the usual rules of quantum mechanics and locality, this will also be the entropy of the state seen by an external observer. There do not seem to be any principles which would prevent this entropy of entanglement from being of the order of the BH entropy of the initial hole, $\sim \frac{M^2}{M_P^2}$. Thus, within the domain of validity of the semiclassical approximation, we can establish the existence of a time slice on which the energy of the external system is small, $\sim 0.01M$, but its entropy huge, $\sim M^2$. It seems to me that the only way to avoid the conclusions of this argument is to give up the assumption that the semiclassical approximation to quantum field theory correctly describes physics in regions of spacetime where the curvature is small. Precisely such a retreat from the conventional wisdom is proposed by the critics of this argument, though there seems to be no agreement among them as to the correct replacement for semiclassical quantum field theory. I think it is fair to say that the above argument establishes that the "subtle correlation" approach is far from conservative. It implies a radical rethinking of our approach to the quantum dynamics of spacetime.

The only one of these radical approaches that the author of the present lectures can make even a pretense of understanding, is that of Susskind [4]. Susskind's criticism of the above argument goes right to the heart of quantum field theory. Indeed, in making the above argument, we accepted without comment the assumption that the Hilbert space of the whole system is a tensor product of states inside the horizon and states outside. This is the fundamental principle that allows us to conclude that information cannot be radiated to the outside on the basis of the examination of the states inside the horizon. This assumption appears to be valid in quantum field theory and in naive cutoff schemes like lattice theories of gravity. Susskind makes the point that it is in no way valid in string theory. If we cut space into two pieces, there will always be strings which straddle the boundary. It is to these "stranded" string states that Susskind looks for the origin of black hole entropy and the resolution of the paradoxes of Hawking evaporation. We will present a brief description of his work in Sec. 6.

2.2. The Remains of the Data

Let us for the moment accept the argument given above and examine its consequences. The fact that the external world has large entropy on the time slice 99 is already in contradiction with the lump of coal analogy. When a lump of coal has radiated most of the energy it absorbed from the laser beam and its glow begins to fade, it has also radiated away most of the information in the *Encyclopedia*. By considering the amount of time it will take for such a low energy, approximately pointlike object to emit the enormous amount of information that it contains, one comes to the conclusion that the remnants of black holes of mass larger than 10,000 tons will have lifetimes longer than the age of the universe—they are essentially stable. Thus, within the context of quantum field theory, the unitary S-matrix scenario seems to reduce to the remnant scenario.

I have emphasized that Susskind's string theoretic criticism of the basic tensor product structure of the space of states in quantum field theory vitiates the force of these arguments. What I find hard to believe is that these criticisms significantly affect the description of the evolution of a large black hole inside its horizon, long before the singularity is encountered. Thus I believe that the picture of long, hornlike geometries connected onto our space by tiny holes must be valid in string theory as well as in ordinary quantum field theory. As we will see, consideration of such structures leads one to the notion of stable remnants with infinite numbers of degenerate states, even if one ignores the above arguments.

We turn now to an exposition of the apparent phenomenological catastrophe caused by the hypothesis of stable or quasi-stable black hole remnants. Hawking's calculation of the evolution of an evaporating black hole appears to be valid until the mass of the hole is of the order of the Planck mass. Its Schwarzschild radius is then of the order of the Planck length, and it appears pointlike to all but the most well-equipped external observers. The remnant scenario thus appears to require the existence of a new class of "particles," all of which have masses of the order of the Planck mass. On the other hand, there must be a distinct remnant particle for each possible state of matter that can collapse to form a black hole, and since remnants can be formed starting with black holes of arbitrarily large mass, there must be an essentially infinite number of different remnant species. Even a few species of stable Planck mass particles might cause difficulties for cosmology if they are produced with reasonable probability after inflation, but an infinite number of species is a complete disaster. Schwinger's calculation of the pair production of charged particles in a background electric field shows that the probability depends only on the mass and charge of the particle (for fixed field strength). An infinite number of degenerate charged species would give an infinite cross section for this process. These estimates of remnant production are completely wrong, because black hole remnants do not behave like elementary particles even though they look pointlike to an external observer. This argument will be taken up in the next section.

2.3. Information Loss

First, however, we must review the third major scenario for the endpoint of Hawking evaporation, that proposed by Hawking himself. Eschewing both the "information is returned in subtle correlations" approach and the idea of remnants (essentially for the reasons outlined above), Hawking instead proposed that an evaporating black hole simply disappeared, taking with it the information that was lost to the external observer in the collapse process. For very small black holes, the entire process of formation and evaporation occupies a small region of space time. Since it is (in Hawking's view) a completely local phenomenon, it should happen all the time in the form of virtual processes even when sufficient energy for real black hole formation is unavailable. Again due to the locality of the process, we should be able to construct a coarse-grained, or effective, theory, describing the effect of these virtual information-destroying processes on large scale physics. Hawking indeed proposed a formalism for computing such corrections. Since pure states can now evolve into mixed states, the effective theory must now map density matrices to density matrices in a way which does not preserve purity.

In ordinary quantum mechanics, the initial and final density matrices in a scattering experiment are related by

$$(\rho_{out})_A^B = S_A{}^C (\rho_{in})_C{}^D S^{\dagger}_D{}^B \ , \tag{2.1}$$

where S is the scattering matrix. Hawking proposed instead a general linear relation[2]

$$(\rho_{out})_A^B = (\$)_{AC}^{BD} (\rho_{in})_D^C . \tag{2.2}$$

If the \$ matrix factorizes into the product of an S-matrix and its inverse, then we have unitary evolution, preserving purity. If it does not so factorize, then purity is lost. Hawking proposed [6] that the true \$ matrix of the world had a small, nonfactorizable term of the form

$$\delta(\$)_{AB}^{CD} = C_{ij} (\frac{1}{M_P^p} \int O_i \int O_j \ , \tag{2.3}$$

where the O_i are operators of high dimension, as indicated by the powers of the Planck mass.

Unfortunately, these apparently small corrections to the \$ matrix are not small at all. To see this [7], we will have to make a small extension of Hawking's \$ matrix formalism and discuss local time evolution equations for the density matrix. That this must be possible follows from the assumption that processes of virtual formation and evaporation

[2] Nonlinear density matrix evolution equations lead to nonlocal phenomena which Polchinski [5] has dubbed "Everett Phones." EPR correlations can be used to send messages in such theories. It is not clear to exactly what extent this is ruled out by experiment, but we will not discuss nonlinear density matrix evolution in this review.

of small black holes are, in Hawking's picture, confined to a small space time region. With sufficient coarse graining, we must be able to incorporate their effect in a set of local evolution equations. We will see that even if we take the coarse-graining scale to be a nuclear time scale, the "small" terms in the evolution equation analogous to Eq. (2.3) are far from negligible. Let us begin by writing the most general linear coarse-grained evolution equation for the density matrix, which conserves probability [8]:

$$\dot{\rho} = i[H,\rho] + C_{\alpha\beta}[O_\alpha,[O_\beta,\rho]], \tag{2.4}$$

where H is hermitian and $C_{\alpha\beta}$ is a real matrix. The O_α run over a complete set of hermitian operators. To preserve the positivity of ρ, we must impose a condition on the relative sizes of the symmetric and antisymmetric parts of C. We *might* also want to impose a condition guaranteeing that entropy always increases [7].

When C is a symmetric matrix, it is possible to make a simple model which produces Eq. (2.4). It is simply quantum mechanics coupled to random sources via a Hamiltonian

$$H_R = H + J_\alpha(t)O_\alpha, \tag{2.5}$$

where the J's have white noise correlation functions

$$< J_\alpha(t)J_\beta(s) > = \ C_{\alpha\beta}\delta(t-s). \tag{2.6}$$

This interpretation makes most of the important features of Eq. (2.4) obvious. In particular, although the evolution equation is time translation invariant, it does not conserve energy. Time translation invariance guarantees only the conservation of the average energy. In a random system, there will generically be energy fluctuations, and the moments of the energy will not be preserved. Similarly, space translation invariance of Eq. (2.5) does not guarantee conservation of momentum.

The extent of this violation of the conservation laws depends on the extent to which the operators O_α and the correlation function of the sources are local. We have assumed that we are working at a scale for which the time correlation of the sources is local. Hawking's proposals lead one to expect all the nonlocality in the new terms in the equation to be at the Planck scale. As shown in Ref. [7], this leads to disaster. The inverse powers of the Planck length are cancelled by matrix elements of the local operators between states of very low and very high energy. In a flash, the vacuum is converted into a mixed state whose dominant components have very high energy. To make the violations of purity and of energy conservation small, we have to smear the operators over long distance scales, which leads to violations of locality. I believe that these arguments show that time-dependent information loss is not a viable proposal.

3. Horned Particles as the Endpoint of Hawking Evaporation

3.1. Near Extremal Charged Black Holes

The Hawking temperature of Reissner-Nordstrom black holes in Einstein-Maxwell gravity vanishes in the extremal limit $Q = M$. This is easily understood in terms of the geometry of the extremal black hole. For $Q = M$, the Killing vector which is timelike at infinity is everywhere timelike, the singularity is a timelike curve, and there is no horizon at any finite distance. Although quantum field theory on this background requires some kind of boundary condition on the singularity, it has a time-independent Hamiltonian. Quantum fields propagating on the background are in pure states.[3] It has therefore seemed plausible to many researchers that extremal charged black holes are the endpoint of Hawking evaporation for the case where a black hole manages to retain its charge.

The extremal Reissner-Nordstrom solution has a geometry similar to that of Fig. 1. (Actually, this is the geometry that comes out of dilaton-gravity which is the low energy approximation to string theory.) It is an infinite horn connected on to the rest of space by a small opening. While it looks like a point to a distant observer, it really contains an infinite volume of space behind it. Clearly, it resembles what the Schwarzschild geometry might have become if the transverse collapse had been averted, but the geometry had continued to grow in the radial direction. This geometry provides the clue to understanding what happens to information in black hole collapse. The infinite reaches of the horn of this black

[3] I am ignoring problems caused by the Cauchy horizon. There is a region of the spacetime, including the whole asymptotic region, which is causally separated from the Cauchy horizon.

hole suggest to us where the information that the external observer has lost might have gone.

We now come to the important question of a name for almost pointlike (from the point of view of an external observer) black holes with large internal geometries. In the context of the physics of near-extremal charged black holes, where this circle of ideas arose, I proposed that they be called *horned particles*, or *cornucopions*, "to celebrate both the shape of their internal geometry and the wealth of information hidden inside them." I will use the names *horned particles, cornucopions, and remnants* interchangeably from now on.

3.2. Relaxation to Extremality

If we throw some neutral matter into an extremal black hole, it ceases to be extremal, develops a horizon, and begins to Hawking radiate. If not too much matter is thrown in, the horizon appears far down the horn of the black hole, and its evolution reduces to a problem in two-dimensional physics. One can formulate a closed set of equations which describe this evolution. For the models that come directly from the Einstein Lagrangian or low-energy string theory, the solutions are singular, but one can change the Lagrangian in a manner that eliminates the singularities. These nonsingular equations were studied numerically by O'Loughlin and Lowe [9]. It turns out that they are very similar to the equations for the Reissner-Nordstrom geometry and Ref. [9] also carried out a numerical study of evaporation of Reissner-Nordstrom black holes. This case had been treated previously by Strominger and Trivedi [10], using approximate analytical techniques. The evolution is completely nonsingular[4] and leads to the following qualitative picture (Fig. 5).

When matter is incident on an extremal black hole, an apparent horizon is formed and the black hole begins to radiate. The apparent horizon recedes from the external observer, eventually leaving an infinite static geometry identical to the exterior of the extremal black hole solution. The full spacetime, however, has a horizon, and one can verify explicitly that the state of the field theory behind the horizon depends on the nature of the initial infalling matter. Thus, within these models, we have a consistent remnant scenario for

[4] There is, of course, a timelike singularity in the Reissner-Nordstrom case, but this does not affect the qualitative nature of the evolution.

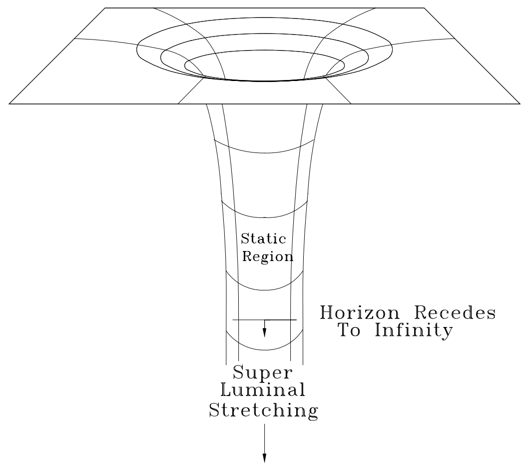

Fig. 5: Evolution of the spatial geometry of a nonextremal black hole back to an extremal remnant.

black hole evaporation, in which information is lost to the external observer. There is a different type of remnant for each kind of initial pulse that forms the black hole (and thus an infinite number of different remnants altogether), but they are all indistinguishable from the point of view of the external observer. The information that distinguishes between them is causally disconnected from him. On any finite time slice, the system still has an apparent horizon that is accessible to the external observer, and it may be possible for him to associate the information with states on the horizon. Asymptotically, he cannot do so. The apparent horizon has gone off to infinity, a different infinity than his own initial asymptotic region. His S-matrix will not be unitary.

3.3. Production of Remnants from the Vacuum

I now come to the discussion of the famous "problem" of the remnant scenario, the spectre of infinite production cross sections. The crucial point turns out to be that on any finite time slice in a synchronous gauge, a cornucopion is *not* a static geometry. It has an apparent horizon, behind which the spatial geometry is undergoing superluminal expansion. The infinite number of degenerate states of the cornucopion are the states of quantum fields lying behind this horizon.

The model for all discussions of production of nontrivial classical field configurations from the vacuum of a quantum field theory is a classic paper by Affleck and Manton [11]. These authors studied pair production of 't Hooft-Polyakov monopoles in a constant magnetic field, and showed that in the small field limit the amplitude reduced to Schwinger's formula for the corresponding amplitude for elementary monopoles. The essential idea of the calculation is that for sufficiently small accelerations, the only degree of freedom of a soliton that one should be able to excite is the noncompact collective coordinate for its motion through space. Thus, for the motion of a soliton in a weak external field, one can construct an approximate solution in which the soliton simply follows a curved world line. The idea of Affleck and Manton was that one can compute the pair production cross section for solitons in the same field by analytically continuing the solution for a pair of particles to Euclidean time. The uniformly accelerated hyperbolic Minkowski motion continues to a circle in Euclidean time, and one computes the production amplitude to be the exponential of the action of this Euclidean instanton.

We do not really understand the Euclidean continuation of quantum gravity, but at the semiclassical level, there is a long tradition of seemingly sensible calculations [12] which simply take over the idea of analytically continuing Minkowski geometries. Special provisions must be made for situations in which the vector field, which we use to define time at infinity, changes signature in some finite portion of the spacetime. The analytic continuation is performed by taking this vector field to be imaginary, but this only produces an Euclidean manifold in the region where the vector field is timelike. On the surface on which the change takes place, the vector field is null. In the analytically continued spacetime, this null surface is replaced by a point, and the portion of the Minkowski manifold beyond this null surface is discarded. Now consider the pair production of the degenerate states of cornucopions. Unlike solitons, these are not globally static configurations. They have, to a good approximation, a timelike Killing vector outside the apparent horizon,[5] but it becomes null on the apparent horizon. Following the rules of Euclidean gravity as we know them, the Euclidean continuation of a cornucopion trajectory does not contain the portion of space where the degenerate states of the cornucopion live. Thus, the tunneling

[5] This approximation becomes better and better as time goes on and the horizon recedes.

process cannot produce these states [13]. Heuristically, the infinite degeneracy of cornucopion states is localized in a region behind the horizon, causally disconnected from the external observer. The cornucopion stationary state of the external observer's quantum field theory contains only the static part of the geometry, in front of the horizon. When the external observer attempts to create pairs of these states, with an external field, he can at best create the static geometry, and this does not contain the degenerate states.

3.4. Remnants of the Imagination?

If black hole remnants exist, will we ever be able to find them? And if we do, will we ever be able to tell what they are? Unfortunately, the answer to the first question is probably no. Cornucopions are by hypothesis stable. Although we have no very good idea about the processes which might produce them in the very early universe, it is unlikely that the production process is so finely tuned that it can produce a density of remnants that is neither much larger nor much smaller than the density of ordinary matter in the universe. In the former case, cornucopions would be ruled out by observational astronomy; in the latter case, we would never be able to find them. We can save a model of the production process that produces too many cornucopions by invoking inflation, but in that case there are likely to be no black hole remnants in our portion of the universe. We would have to wait for macroscopic black holes to finish their Hawking evaporation before we could get our hands on real remnants.

But suppose we did so. Would we then be able to tell what the remnants were by their properties, or would they just behave like ordinary elementary particles? Classically, I believe the answer to this question is no. Small perturbations of the cornucopion change its internal structure and produce horizons. If we now imagine quantizing the classical cornucopion solution, it would formally have statistics, like an elementary particle. However, any realistic experiment in which we attempted to measure these statistics would be doomed to failure. The procedure of scattering the cornucopions to measure their statistics would invariably cause the formation of an internal horizon and the emission of Hawking radiation that would change the internal state in an uncontrollable way. The statistical phase would be unmeasurable, and the cornucopions would behave like classical distinguishable particles. A double-slit experiment for cornucopions would resemble Fig. 6.

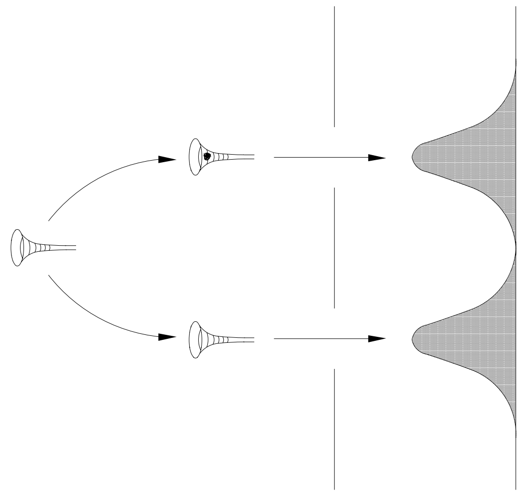

Fig. 6: A double-slit experiment for cornucopions.

There is a way in which the above argument could fail. Quantum mechanically, there might be a threshold energy for horizon formation. Given their infinite internal geometry, it is inevitable that the spectrum of excitations of cornucopions will have a continuum, but the continuum might be separated from the ground state by a gap. Experiments in which the cornucopions were moved with sufficiently small velocity would not excite the continuum, and in these experiments they would behave like quantum particles. Determination of the size of the gap in the cornucopion excitation spectrum seems to me to be one of the most interesting open problems in the study of black hole remnants. Unfortunately, it is a very hard problem and probably requires the study of strongly coupled string theory.

Finally, one should note that the pessimistic assessment of the probability of finding black hole remnants in the first paragraph of this section might be wrong. Perhaps the differences between them and elementary particles are sufficiently great that their production probability in an inflationary universe is much larger than we presently imagine.

4. Unitarity, Information Loss, and CPT

The general picture that arises from the discussion of the previous section is that black hole collapse and evaporation leads to the formation of a new asymptotic region of space which is causally disconnected from the old asymptotic region. The spacetimes that we have discussed can be foliated by spacelike hypersurfaces, and within the semiclassical approximation for the geometry, there will be no breakdown of unitarity. Local physics will continue to obey the rules of quantum mechanics.

It is clear, however, that the S-matrix measured by observers in the original asymptotic region will not be unitary. The cornucopion spacetime has a horizon buried deep within the horn of the cornucopion. On any spacelike slice, the state of the system outside the horizon, which is all that can ever be probed by the original asymptotic observer, is correlated with that inside the horizon. The external density matrix is not pure, and the purely external scattering process must be described by a Hawking $\$$ matrix.

The cornucopion scenario thus unifies the idea of information loss in the observable universe with the idea of black hole remnants, retaining the merits of both while discarding their difficulties. In particular, Hawking's original claim that formation and evaporation of "small virtual black holes" would lead to a reformulation of the fundamental microphysical laws to accomodate time-dependent information loss is discarded, and with it the problems unveiled in Ref. [7]. Virtual quantum fluctuations of geometry that do not change the topological properties of space will be describable as distortions of the classical background over small volumes, which subside after a short time. No new asymptotic regions are created in such fluctuations, and it should be possible to "integrate them out" and construct a local effective Lagrangian for long distance physics. This local Lagrangian will obey the usual rules of quantum mechanics. Small, topology-changing processes can also be integrated out, and they too lead to a picture in which the information content of a single connected component of the universe does not change with time. Thus, to paraphrase J. A. Wheeler, we have *Information Loss Without Information Loss*. The rules of quantum mechanics are locally preserved, but the global S-matrix will not be unitary unless we take into account contributions from causally disconnected asymptotic regions.

Hawking has emphasized that any remnant scenario for black hole evaporation implies a violation of CPT in the observable universe. This is indeed the case. Suppose that it is possible to scatter a pair of elementary particles with sufficiently large energy and momentum transfer that they form a black hole. The resulting hole will Hawking radiate, and asymptotically settle down into a cornucopion state in which a new asymptotic region of space is formed. The CPT inverse of this process involves a conspiracy of the matter in this infinite asymptotic region, which causes the "internal universe" to collapse and spew out precisely two particles into the region of space in which we live. On the other side of the horizon in the meantime, particles must be sent in from infinity to converge on the point where the cornucopion throat is sitting and be absorbed by it as inverse Hawking radiation. Such initial conditions, involving as they do a rather special conspiracy between causally separated points in two infinite universes, are clearly of measure zero in the space of all possible initial conditions. CPT is violated in the same way that it usually is in macroscopic systems: the inverses of processes which increase coarse-grained entropy require very special initial conditions, which can never be realized in practice. Once cornucopions are formed, there is essentially zero probability that they will spontaneously dissipate. Similarly, there is zero probability that two cornucopions of opposite charge will annihilate when the mouths of their horns meet. Even if their charges are opposite, the probability that their internal states are exactly CPT conjugates of each other is zero. This is not just a question of the huge number of possible degenerate states available in the large interior of the cornucopion. The interior state of the cornucopion is time dependent and involves an expanding geometry. The CPT conjugate state is one with a contracting geometry and very special initial conditions. It will not be realized in a cornucopion which formed from the collapse of matter in the external world.

Indeed, the annihilation process for cornucopions should be the time reverse of the production process that we discussed above. That process, as we saw above, was both very improbable and resulted in production of cornucopions of very small volume. Thus, in order to annihilate large-volume, expanding cornucopions, we must first force them to evolve classically into the configuration which is produced at the end of the tunneling process described in Ref. [13]. This is very unlikely to happen. Furthermore, causality prevents it from happening as a consequence of simply moving the mouths of the cornucopions so that they coincide in the external space. Cornucopion production, whether through the improbable process of pair production in an external field followed by classical expansion of the internal geometry, or through gravitational collapse, is an effectively irreversible process from the point of view of a single causally connected sector of the universe.

In passing, we note that the above discussion is the strongest argument in favor of the existence of neutral black hole remnants. We can certainly imagine bringing together the mouths of two oppositely charged cornucopions, obtaining a neutral object. The above arguments suggest that that neutral object will have a large and complicated internal geometry, which will not disappear. It will be a neutral cornucopion. This is even more apparent if we imagine neutralizing the charge of a cornucopion by dropping an elementary particle of opposite charge into it. This microscopic perturbation cannot destroy the complicated internal geometry.

4.1. Entropy

No discussion of information loss and unitarity in black hole physics is complete without mention of the BH entropy formula. It is common to describe this entropy as a measure of the number of internal states of a black hole. In the context of the cornucopion scenario, this cannot be the case. In this picture, a black hole has an enormous internal geometry and an essentially infinite number of internal states.

A black hole of mass M could be in any one of an infinite number of states depending on its past history. The information theoretic entropy of the external world would depend on which state the hole was in. However, most of the information contained in the correlations between the internal and external states of the black hole can have no impact on the future interaction of the black hole with the external world because it is causally disconnected from the external portion of spacetime. Thus, this measure of entropy can have no bearing on the *thermodynamic* entropy of the hole, which describes how it exchanges energy with the outside world. A system can come into thermal equilibrium only with a collection of states with which it is in causal contact and constantly exchanging energy.

The thermodynamic entropy of a large black hole can only be related to states located very close to the horizon. If we look at states localized at any finite distance from the horizon, with a resolution coarse enough so that the semiclassical picture of the geometry is valid, then these states are expanding away from the horizon at superluminal velocities. They are out of causal contact with the outside world, and in the cornucopion scenario they will remain so forever. Only states "infinitesimally close" to the horizon can be in causal contact with the outside. Here, "infinitesimally close" probably means within a length scale l which is small enough so that experiments probing physics on the scale l cannot be described by the semiclassical approximation. This probably means that l is of the order of the Planck length, or the slightly larger fundamental length of string theory. It is thus plausible that the number of states of a black hole that might be in causal contact with the external world is proportional to the area of the horizon, and it is perhaps natural that the proper units (i.e., those in which the proportionality constant is of order one) of area are Planck units.[6]

The fact that the information theoretic entropy of a black hole must, in the remnant scenario, be thought of either as infinite (the logarithm of the number of possible final states of the remnant) or as dependent on the black hole's entire previous history (the entropy of entanglement of the internal state produced in a particular process of formation of the hole with the external world) has been among the many arguments levelled against the remnant scenario. I believe that the above paragraph shows clearly that neither of these quantities is a relevant measure of the thermodynamic entropy of the hole. Further, it is clear that if the black hole has a thermodynamic entropy, it should be proportional to the area of its horizon. A microphysical demonstration that the vicinity of the horizon really contains the number of states indicated by the BH entropy formula would seem to depend on knowledge of physics at very short distances. Here we make contact with the point of view of 't Hooft [14] who has long insisted that the divergence of the entropy of entanglement of the state of a quantum field theory outside the horizon with that of the inside was the key to understanding the BH entropy formula. 't Hooft believes that

[6] In string theory, the string length is the more natural unit. This discrepancy will have important consequences below.

this demonstrates that the understanding of black hole entropy and of the information paradox is a problem of short distance physics, and will guide us in the construction of the fundamental theory of small-scale geometry. While I do not agree with his assertion that unitarity of the S-matrix for the original asymptotic region is a necessary ingredient in the construction of the theory, I do agree that the thermodynamic nature of black holes can only be understood in terms of short-distance physics.

I have emphasized throughout this talk that our discussion of the evaporation of black holes depends only minimally on Planck scale physics and not at all on quantum gravity. We have seen that it is possible to make the idea of information loss to an asymptotic observer consistent with the basic rules of unitary quantum field theory in a self-consistently generated classical gravitational field. The microphysical derivation of the BH formula is the one aspect of this subject that still seems to hint at the need for a more fundamental theory. I believe that I have outlined a way in which this formula could be compatible with the idea of remnants, but the derivation of the formula itself seems to be outside the domain of reliability of the semiclassical field theoretic considerations that have been our guide up to this point. Furthermore, this discussion makes it clear that the remnant scenario (nor, I believe any scenario which leads to effective information loss) cannot account for the information corresponding to the BH entropy. In the next section I will describe one attempt that has been made to understand the BH entropy formula from a microscopic point of view.

5. Stranded Strings and Black Hole Entropy

The cornucopion scenario for black hole evaporation is a logically complete (though technically incomplete) description of this process. By contrast, the material I am about to discuss is very much "work in progress." Its authors (primarily 't Hooft, Susskind, and their collaborators) do not pretend to have achieved a logically complete understanding of the process of Hawking evaporation. I am in the unfortunate situation of understanding only imperfectly what these authors *have* accomplished. I enter into a public discussion of their work with great trepidation and apologize in advance for my inevitable distortion of their points of view. All errors in the paragraphs that follow are my own.

One of the fundamental ingredients in 't Hooft's work on black holes is the BH entropy formula. One starts from the assumption that black holes behave like thermal objects with Hawking temperature and BH entropy, and attempts to make this fact consistent with the postulate that the S-matrix measured by the initial asymptotic observer is unitary. A special role for Schwarzschild coordinates is implicit in these assumptions. The black hole is static, and thermodynamic considerations make sense, only in these coordinates.

One of the most revealing of the calculations performed by 't Hooft is that of the entropy of states of quantum fields in a small region of a fixed Schwarzschild time slice, very near the horizon. 't Hooft argues that in order for a quantum field theoretic calculation of the entropy and energy of the black hole to give an answer of the order of the classical values (i.e., M for the energy and the BH entropy), one must cut off the quantum field theory by insisting that there are no states at a proper distance closer to the horizon than $\sqrt{\frac{N}{90\pi}}M_P^{-1}$ where N is the number of field degrees of freedom. But this removes many of the modes of the field which are involved in Hawking's calculation of the outgoing density matrix. Modes which can escape to infinity with energies of the order of the Hawking temperature have extremely high energy near the horizon and do not satisfy 't Hooft's brick wall boundary condition.

Susskind's point of departure is the 't Hooft entropy calculation described above. He brings in ideas from the Membrane Paradigm, which has had many successful applications to the astrophysics of black holes. The basic idea of the Membrane Paradigm is that, from the point of view of an external observer, things that fall through the horizon of a black hole may equally well be imagined to reside on a membrane suspended above the horizon. Susskind, Thorlacius, and Uglum [2] proposed to take the membrane or "stretched horizon" to lie only a few fundamental lengths from the horizon (i.e., the stretched horizon is very analogous to 't Hooft's brick wall), and postulated that the BH entropy can be understood in terms of degrees of freedom living on this membrane. They further suggested that the process of Hawking evaporation could be completely understood in terms of interactions between the external world and these degrees of freedom, and that the resulting S-matrix would be unitary.

't Hooft's calculation shows that this suggestion makes no sense in quantum field theory. The number of Schwarzschild degrees of freedom per unit area of the membrane will be infinite in field theory, and will be much larger than $e^{\frac{M^2}{M_P^2}}$ for any cutoff that keeps all the degrees of freedom necessary to describe Hawking radiation. We have discussed previously an even more serious (since it does not make reference to very high energies or special coordinate systems) field theoretical argument that information cannot get out of a black hole with the Hawking radiation. This argument invoked the tensor product structure of field theoretic Hilbert spaces, causality, and the smoothness of the spacetime metric on a certain spacelike hypersurface. In field theory, if we tried to associate degrees of freedom with a stretched horizon, the Hilbert space would (by locality) be a tensor product of an inside space, an outside space, and a horizon space. We could make conclusions about information carried by the infalling particles by examining the inside space alone, and use the tensor product structure to make negative conclusions about the possibility of that information being carried by the horizon.

I would like to present two more arguments (due to Susskind) that the membrane picture cannot work in quantum field theory. To do so, we consider a black hole of extremely large mass, or what is the same thing, Rindler space. The spacetime curvature is completely negligible for our considerations. We know that Rindler space is simply a wedge of ordinary Minkowski space. We want to consider observations made by a Rindler observer following a timelike trajectory whose proper distance from the Rindler horizon is a few multiples of the fundamental length.

Now let us consider throwing a wave packet of particles at the horizon. The wave packet is assumed to be localized in the transverse direction at infinity. How will such a packet appear to our Rindler observer hovering just above the horizon, according to the rules of quantum field theory? To some extent, this is a difficult question because it involves scattering cross sections at very high energies and low momentum transfers. This regime is not well understood even in weakly coupled field theories. We are interested in the question of whether the cross section grows with energy, and in how much information is contained in the growing cross section. A rigorous answer to the question of growth of the cross section is not known, but what we do know fairly rigorously, both from field theoretic studies in models with spin-one gauge bosons and hadron phenomenology,[7] is that the growing part of the cross section is more or less universal. In parton-model language, the quantum numbers that distinguish hadrons from one another are carried by the valence partons, which carry finite fractions of the longitudinal momentum of the hadron in the infinite-momentum frame.[8] These behave like normal particles with transverse wave functions which do not grow with energy, but remain localized in the transverse plane. Their wave functions also Lorentz contract in a normal manner and are thin pancakes in the infinite momentum frame. The bulk of the high energy cross section is due to *wee partons*, which have a $\frac{dx}{x}$ distribution in longitudinal momentum and logarithmically spreading wave functions in the transverse plane. In hadron physics, the wee parton distribution is universal and contains no information about which hadrons are scattering (except perhaps whether they are baryons or mesons). We emphasize that the existence of growing cross sections and wee partons is a conjecture in field theory. If they do not exist, then our argument is even simpler. The important point is that they do not carry (much) information even if they exist.

The implications of this behavior for our Rindler observer are clear.[9] He will find that the information in an infalling particle wave packet remains localized in the transverse plane. Furthermore, if he restricts his attention to the region outside a stretched horizon which is some fixed distance from the horizon, he will soon find that he cannot measure anything about the infalling state at all. The flat pancake which carries all the information falls below the stretched horizon in a time that is short compared to the Hawking

[7] Here we are extrapolating the behavior of hadrons at energies $\sim 1\ TeV$ to energies many orders of magnitude above the Planck scale.

[8] A rather appropriate frame for studying super-Planck energy collisions near the horizon.

[9] Actually, the statements made above apply to an inertial observer who happens to have the same velocity as the Rindler observer at the moment he makes his measurements. If the measurements are sufficiently localized in space and in infinite momentum frame time, there is probably not very much difference between these two observers in quantum field theory. The distinction might be more important in string theory, where we do not really know how to solve the theory in a noninertial reference frame. One extremely important aspect of the situation that cannot be understood in terms of inertial frames is the difference between the experiences of the inertial and accelerated observer.

evaporation time of the black hole. Thus, according to the rules of quantum field theory, it is impossible for a Rindler observer to imagine that the information carried by infalling particles gets smoothly spread over the stretched horizon, to be emitted later as isotropic Hawking radiation.

Susskind points out that in string theory the situation is quite different. Growing cross sections are built into string theory, since it is Regge behaved and the intercept of the graviton trajectory is two. A very picturesque way of understanding this has been developed by Karliner, Klebanov, and Susskind (KKS) [15] who did Monte Carlo simulations of the distribution of strings predicted by the wave functions of free bosonic string theory in light cone gauge. Consider a small box of fixed size in the transverse plane, and a string whose center of mass is in that box. KKS ask how much of the actual length of the string is in that box. The answer depends on a cutoff that they imposed on longitudinal momentum. Remembering that low longitudinal momentum means large light cone energy, we see that this is a cutoff on the time resolution of the observer looking at the string. Remember also that in light cone gauge string theory, a cutoff on longitudinal momentum is a spatial world sheet cutoff, a cutoff on the number of modes of the string. KKS find that as the cutoff is taken to infinity, the proportion of the string that is in the box goes to zero. This is a symptom of the logarithmic spread of the string in the transverse plane. The region (measured in string units) in which the string is confined grows logarithmically with the cutoff.

To apply this to the black hole problem, note that for a Schwarzschild observer supported near the horizon, watching a string fall into a black hole at a time T (long) after the black hole is formed, the Lorentz boost between the observer's frame and the infalling string frame corresponds to a time dilation[10] $e^{\frac{TM_P^2}{4M}}$. Thus, in order to see what is going on in the string frame, the observer needs an exponentially fine time resolution, corresponding to a spread of the string over an area $\frac{TM_P^2}{4MM_S^2}$. Thus, in the Hawking evaporation time $T \sim \frac{M^3}{M_P^4}$, the string has spread over an area $\frac{M^2}{M_P^2 M_S^2}$. There is an important factor of the

[10] Here, we are treating the Schwarzschild frame as a highly boosted inertial frame because we do not know how to do unitary string quantum mechanics in noninertial frames. This is a weak point in Susskind's arguments that deserves investigation.

string coupling in these formulae corresponding to the fact that the natural length scale for string fluctuations is larger than the Planck length by a factor of $\frac{1}{g}$. This ensures that the string has spread over an area larger than the horizon, or rather, that it has been able to cover the horizon many times. In higher dimensions, the string spreading, evaporation time, and horizon area scale differently and the string has spread over regions bigger than the horizon by powers of the black hole mass.[11]

Another important difference between string theory and field theory is that the information which distinguishes between states is carried not by *valence partons*, which are localized excitations, but by vertex operators which are conformal fields smoothly spread over the fluctuating string. (Examples are the Kac-Moody currents which carry gauge quantum numbers.) The well-known inability of string theory to reproduce localized form factors for hadrons is a symptom of this effect. This implies that the localized Schwarzschild observer will not be able to conclude that the information carried by an infalling string is localized near its center of mass. A collection of observers spread over the horizon would be necessary to extract the information (and they would then probably completely change the state of the string). Thus, there is no contradiction with the claim that the information can be emitted as isotropic thermal radiation.

In a similar vein, the Regge behavior of string scattering amplitudes and their growing cross sections imply that strings do not Lorentz contract to sizes smaller than the string length. Thus, if we put a stretched horizon within a string length of the horizon, strings will never "fall through it" from the point of view of a Schwarzschild observer. Thus in string theory, as opposed to quantum field theory, the picture of information carried by a stretched horizon which interacts with the outside world is not ruled out.

[11] In all of these considerations, it is important to take the string coupling to be extremely weak. The region in which our description of string scattering is reliable is $\frac{g^2}{M_S^2}E < b$, the impact parameter. In order for the energy to be high enough for string spreading to cover the horizon, we need $ln(\frac{E}{M_S}) = \frac{g^4 M^2}{M_S^2}$ However, we must also be discussing experiments with impact parameters much less than the Schwarzschild radius of the black hole. Thus, we must have $\frac{M}{M_S} \gg 1$ and $g \ll \sqrt{\frac{M_S}{M}}$. This inequality is not satisfied by the string coupling in the real world. Thus Susskind's estimates are really valid only in an imaginary world with very weak coupling. I thank E. Martinec for a discussion of this point.

In assessing Susskind's claims about string theory and black holes, it is important to understand that he is not proposing to change the picture of physics as seen by an infalling observer near a large mass black hole. The spacetime geometry is smooth and so is the observer's coordinate system, and no string theoretic corrections to a semiclassical field theory description are important. This is not the case for the Schwarzschild observer. We have recalled that in classical general relativity the Schwarzschild observer can access all of the information about objects that long ago fell into the black hole, but only by making observations on extremely short distance and time scales. She can attribute all interactions of the black hole with the exterior to a membrane on the stretched horizon, but unless she takes her stretched horizon extremely close to the true horizon and measures time with extreme precision, she can only discuss average properties of the hole. It is to describe these extremely precise measurements of short times and distance scales that string theory, or some other theory of the small scale structure of the world, becomes necessary. Susskind's string theory of black hole thermodynamics is, like the classical Membrane Paradigm, a description of the physics that is tied to a particular coordinate frame (and those related to it by smooth coordinate transformations).

We now come to what most researchers consider the most important argument against a unitary S-matrix, the analysis of the spacelike surface we called 99 in Sec. 2. The refutation of this argument goes right to the heart of the difference between quantum field theory and string theory. The essence of the argument is that the state of the system on this spacelike hypersurface lives in a Hilbert space which is a tensor product of an "inside" space and an "outside" space. If we wish, we can add a "membrane" space intermediate between the two, but this changes nothing. We claim that we know from the prior history of the system that the inside state is correlated with the outside. Then, from knowledge that the inside state has not changed very much ("nothing much happens to the infalling observer as he falls through the horizon of a large black hole"), we conclude that the outside density matrix has a large entropy. When we try to rerun this argument in string theory, we run into an immediate snag. Consider ordinary light-cone gauge string field theory, and let us try to divide space into two along one of the transverse directions. We want to consider the entire Hilbert space as a product of an $x_1 > 0$ and and $x_1 < 0$ Hilbert space.

We cannot. The full Hilbert space contains states which are created from the vacuum by creation operators for strings which straddle the boundary. These do not belong in either the plus or minus Hilbert space. The tensor product structure of the Hilbert space of any quantum field theory is completely unrecognizable here. For a black hole spacetime, these nonfield-theoretic states will be strings "stranded on the stretched horizon."

String field theory in a half space is in fact a complicated interacting theory. In addition to states of strings completely within the half space, it contains the stranded states described above. And even when the string coupling vanishes, the half space theory contains complicated interactions in which a closed string is annihilated and replaced by an arbitrary number of stranded strings. These are required to reproduce the dynamics of free strings moving in the full space, which cross the boundary (Fig. 7).

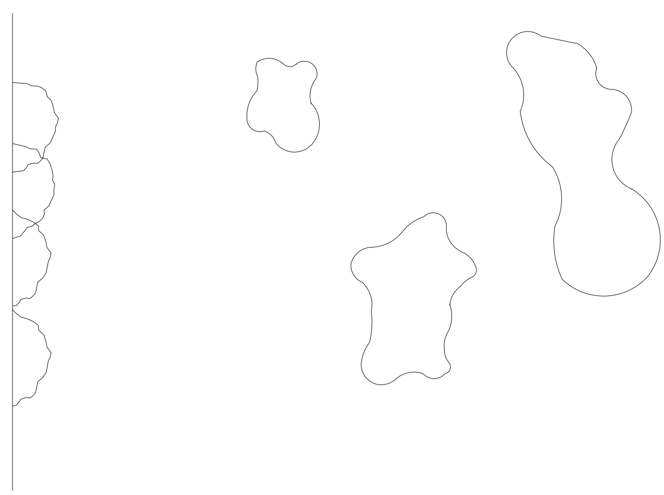

Fig. 7: Stranded strings violate the locality postulate of field theory but can only be seen with very short time resolution.

It is important to note that the wave functionals of even low energy states of the string have nonzero amplitudes for finding large strings in them. Thus, the creation operator for a photon state with center-of-mass coordinate arbitrarily far from the boundary has a nonzero piece which creates stranded strings. It is only when we agree not to study the photon state on short time scales that we are able to ignore this. The large strings which appear with finite amplitude in the photon wavefunctional all have large world sheet wave numbers. When we consider averages over light-cone time intervals much larger than the string scale, we can construct an effective theory in which string degrees of freedom with large world sheet wave numbers are integrated out. The only large strings which appear in this effective theory are large smooth strings. Neither the photon wavefunctional nor that of any low energy excitation of the string has any significant amplitude for such large smooth strings. Thus, the stranded string states in the full string field theory Hilbert space can be ignored if we are making low energy observations of low energy states.

The situation is quite different for observations with very short time resolution. Consider a wave packet of photons travelling in the negative x_1 direction, the center of which is at some large negative x_1. The string wavefunctional for this photon has finite amplitudes for string configurations which wander arbitrarily far from the center of the wave packet, even into the region where $x_1 > 0$. These are configurations which contain large contributions from very high world sheet wave numbers, and they oscillate with very high frequency. However, measurements with sufficiently fine time resolution in the $x_1 > 0$ region should be able to discern them. In principle, these measurements could determine the state of a photon whose center of mass was arbitrarily far away.

Susskind thus argues that much of the information (indeed, he claims all the information) that in field theory is lost behind the horizon of the black hole, is encoded in string theory in the state of the stranded strings on the stretched horizon. It would be a remarkable consistency check of these ideas if one could calculate the entropy of stranded strings in a black hole background and find that it agreed with the BH entropy. This calculation has not yet been done, but even without attempting it we can see that there is a puzzle about the detailed form of the BH entropy formula in any theory in which the fundamental length is larger than the Planck length by a factor of the dimensionless coupling that controls semiclassical expansions.

In particular, in string theory, the fundamental unit of length is the string length $l_S \sim \sqrt{\alpha'}$, related to the Planck length by the dimensionless string coupling, $g_S l_S \sim l_P$. Thus, in terms of the natural units of string theory, the BH entropy is an effect of order $\frac{1}{g_S^2}$. Any straightforward calculation of the quantum entropy of strings in a black hole background will, to leading order, be independent of g_S. Indeed, this would seem to be a property of any quantum theory. In the semiclassical expansion, entropy arises by tracing over states which are small fluctuations about a classical background. To leading order, their description is independent of the coupling, and one would expect the expansion of the entropy to begin at order g_S^0.

Susskind proposes to avoid this apparent difficulty by the following heuristic, though rather convincing, line of argument. Imagine doing string theory in a light cone frame moving parallel to the horizon. In this gauge, we can, as above, try to make sense of local quantities in string theory. In particular, Susskind proposes to take over from local field theory the notion that an observer, in this frame localized a distance d from the horizon, effectively sees a bath of strings at a d dependent temperature, approaching the Hagedorn temperature as $d \downarrow l_S$. If this is the case, then the string partition function at genus one will blow up in a black hole background. It has long been believed that in the finite temperature case, this divergence is the signal of a phase transition to a new phase of string theory. In particular, Atick and Witten [16] have argued that this phase transition is of first order. Formally, the precursor of the transition is a tachyonic mass for certain winding modes of strings around the compact time dimension of a finite temperature cylinder. Atick and Witten argue that string interactions stabilize this instability, leading to a finite condensate of winding modes which contributes to the free energy at order $\frac{1}{g_S^2}$. Susskind proposes to identify the black hole analog of the genus zero entropy of Atick and Witten with the BH entropy.

It is beyond the reach of present stringy technology to calculate the Atick-Witten entropy at genus zero, much less Susskind's profound generalization of it. Nonetheless, taken as a whole, Susskind's proposal is, in this author's opinion, the first serious attempt to explain the BH entropy in terms of a quantum mechanical sum over states. His whole picture, if viewed as the quantum mechanical analog of the Membrane Paradigm, is consistent with everything we know about black holes and strings. In particular, it does not contradict the use of effective field theory for the description of the vicinity of the horizon of a large black hole as viewed by an infalling observer. Classical general relativity tells us that a Schwarzschild observer with fine enough resolving power is able to retrieve all information about things that fell into the black hole in "the remote past."[12] It is only for the quantum mechanical analysis of these super-Planckian measurements that one needs to resort to string theory.

Viewed in this manner, the BH entropy formula hints at a very interesting conclusion. In classical physics, we can throw an infinite amount of information into a black hole, and the Schwarzschild observer can measure it. The BH formula suggests that this infinity is cut off by quantum effects. Bekenstein has tried to argue [17] that this cutoff reflects a fundamental quantum gravitational restriction on the amount of information that can be contained inside a volume bounded by an area A. As far as I know, all such arguments implicitly assume that the volume enclosed is finite, and certainly do not take into account geometries with horizons which recede off to infinity. The cornucopion scenario is in direct contradiction with Bekenstein's bounds, if they are taken to refer to the entire entropy of entanglement of the external world with the world behind the black hole horizon. If instead they are taken to represent only the entropy of states near the horizon that are in causal contact with the external world, the contradiction is removed. I believe that this is all that is necessary to prevent violations of the second law of thermodynamics in the presence of black holes.

Instead, in view of Susskind's picture of the origin of the BH formula, I would ascribe the finiteness of BH entropy to a limitation on the information-carrying capacity of stranded strings on the stretched horizon. Consider a string state whose center of mass is thrown into the black hole. In free string theory, it leaves behind stranded strings on the horizon. Once interactions are taken into account, these stranded strings can break, combine with other strings, and in this way, "lose contact" with the original state which deposited them. Information about the infalling state is now truly lost to the stranded strings and thus to the external observer. The horizon is thus a semipermeable membrane for information—that information represented by the BH entropy is information about the state of the stranded strings on the stretched horizon, but it is not complete information about what fell into the black hole. The idea that complete information is in principle

[12] As seen by an infalling observer. Of course, all these measurements are made on a single Schwarzschild time slice.

accessible is an artifact of classical physics, corresponding to the divergence of the BH entropy in the $g_S \to 0$ limit.

This lecture has been only a brief summary of Susskind's ideas. I have stressed primarily those points where I felt that a clearer explanation was necessary than could be found in the literature. For more details, I refer the reader to Susskind's original papers [4].

6. Conclusions

I believe that by combining the notion of cornucopions with Susskind's ideas about string theory and BH entropy, we have for the first time the outline of a sensible story about the problem of information loss in black hole evaporation. These two seemingly contradictory aspects of the description are the analog of the infalling and supported observers' points of view in classical general relativity. The cornucopion scenario tells us what is "really going on" behind the horizon. It cannot easily account for the BH entropy, which is a notion relevant to the observations of the external observer. Susskind's interpretation of the BH entropy is a concrete realization of the heuristic treatment of this entropy by the Membrane Paradigm. It is remarkable that we have to give up the rules of quantum field theory in order to find a consistent realization of these semiclassical ideas. It is even more remarkable that string theory seems to provide the required generalization of field theory. If these arguments are verified, 't Hooft's bold claim, that the resolution of the paradoxes of Hawking radiation would lead us to the correct theory of quantum gravity, will be vindicated.

Taken by itself, Susskind's picture could be advanced (and Susskind so advances it) as proof that *all* of the information in the black hole is returned to the external observer. As indicated above, I do not think that this is necessarily correct. All of the recent work on remnants has led to a picture in which black hole formation and evaporation terminates, at least for black holes of sufficiently large magnetic charge, in a semiclassical spacetime in which the number of different causally disconnected asymptotic regions is different in the future than in the past. Basically, this is the picture of a black hole that comes out of classical general relativity, with only the singular behavior of that theory removed. I believe that perturbative string theory on such a spacetime would lead to the prediction that the S-matrix for the original asymptotic region is nonunitary (basically, because even though string theory is nonlocal, it satisfies the cluster property of the S-matrix).

7. Acknowledgments

Conversations about black hole physics with G. 't Hooft, E. Verlinde and H. Verlinde, S. Das, and E. Martinec are gratefully acknowledged. Much of the content of this lecture was written while the author was in residence at the Institute for Theoretical Physics in Santa Barbara during the spring semesters of 1993 and 1994. I would like to thank J. Langer, M. Stone, and the Institute staff for providing me with a wonderful place to get some work done. Conversations with M. Srednicki, S. Giddings, and especially A. Strominger were invaluable to me, and without them these lectures would be much less clear than they are. I would also like to take this opportunity to thank my collaborators, A. Dabolkhar, M. Douglas, and especially M. O'Loughlin for hours of enjoyable debate about the subject of black holes and information loss. My colleague S. Shenker deserves his usual acknowledgment for penetrating insights that have opened up whole new areas of investigation for me. This work was completed at the Aspen Center for Physics in the summer of 1994 and the Laboratoire de Physique Theorique of the Ecole Normale Superieure in Paris in the fall. I would like to thank all of the members and staff of these institutions for their hospitality. Finally, I would like to thank L. Susskind for many conversations about physics, and for continually challenging us all with his original insights.

References

[1] D. A. MacDonald, R. H. Price, and K. S. Thorne, *Black Holes: The Membrane Paradigm* (Yale University Press, New Haven, 1986), p. 367.

[2] L. Susskind, L. Thorlacius, and R. Uglum, Phys. Rev. D **48**, 3743 (1993).

[3] S. Hawking, Comm. Math. Phys. **43**, 199 (1975).

[4] L. Susskind, Phys. Rev. D **50**, 2700 (1994), hep-th 9309145; *ibid.* D **49**, 6606 (1994); Phys. Rev. Lett. **71**, 2367 (1993).

[5] J. Polchinski, Phys. Rev. Lett. **66**, 397 (1991).

[6] S. Hawking, Phys. Rev. D **14**, 2460 (1976).

[7] T. Banks, M. Peskin, and L. Susskind, Nucl. Phys. B **244**, 135 (1984).

[8] G. Lindblad, Comm. Math. Phys. **48**, 119 (1976).

[9] M. O'Loughlin and D. Lowe, Phys. Rev. D **48**, 3735 (1993).

[10] A. Strominger and S. Trivedi, Phys. Rev. D **48**, 5778 (1993).

[11] I. Affleck and N. Manton, Nucl. Phys. B **194**, 38 (1982).

[12] S. Hawking and G. Gibbons, Phys. Rev. D **15**, 2752 (1977); D. Gross, M. Perry, and L. Yaffe, Phys. Rev. D **25**, 330 (1982); S. Coleman and F. Delucia, Phys. Rev. D **21**, 2133 (1980).

[13] T. Banks, M. O'Loughlin, and A. Strominger, Phys. Rev. D **47**, 4476 (1993).

[14] G. 't Hooft, Nucl. Phys. B **256**, 727 (1985).

[15] M. Karliner, I. Klebanov, and L. Susskind, Int. J. Mod. Phys. A **3**, 1981 (1988).

[16] J. Atick and E. Witten, Nucl. Phys. B **310**, 291 (1988).

[17] J. Bekenstein, Phys. Rev. D **7**, 2333 (1973); Phys. Rev. D **9**, 3292 (1974); Phys. Rev. D **49**, 1912 (1994).