TOPIC IX

THE INFALL PROBLEM

What happens to an object that falls onto the horizon?

We will call this question the 'infall problem'. It is sometimes confused with the information puzzle, but it is a very different question. In some sense it is actually the *opposite* question:

(i) We have seen that the information paradox arises if the horizon is a vacuum region. Thus most attempts to resolve the paradox look for some way of modifying the physics at the horizon. This modification should somehow encode into the horizon the information of the matter which initially made the hole, in order that the radiation from the hole carry out this information. Thus each black hole microstate should be different in some way at the horizon. More precisely, the radiation quanta emitted with energies $E \sim T$ should differ by order unity between different microstates.

(ii) The infall problem on the other hand arises from a very strong intuitive desire to preserve one feature of the classical black hole: namely, that an infalling observer feel nothing as he crosses the horizon. If thus desire is to be satisfied, the different microstates will have to behave in the *same* way for the purposes of this infall.

It is clear that there is a tension between (i) and (ii). We will soon list some of the approaches that have been taken to deal with this tension. But before we proceed, we note that the problems (i) and (ii) are not problems on the same footing:

(i) The information paradox is a very deep issue, that *must* be resolved; otherwise general relativity would be in conflict with quantum theory.

(ii) The infall problem arises from an aesthetic desire that some feature of our classical intuition be preserved at the quantum level. But there is no essential *requirement* on our theory that it preserve this intuition.

In particular, we have seen that fuzzballs resolve the information paradox in string theory by altering the structure at the horizon. While we will discuss the infall problem with fuzzballs below, this discussion has no impact the resolution of the information paradox itself.

Lecture notes 1

Different approaches to the problem

Let us now summarize some conjectures that have been made to deal with the tension noted above between the information paradox and the infall problem:

(a) *Complementarity:* We have seen that the quantum physics of black holes leads to problems if the horizon is a vacuum region. 't Hooft made a bold suggestion: the correct theory of quantum gravity must be such that the degrees of freedom of the hole actually reside at the horizon. Since the horizon is a 2-dimensional surface rather than a 3-dimensional volume, this conjecture says that gravity is 'holographic': its degrees of freedom in a region are not 'one per unit volume' but 'one per unit boundary area'.

If the data of the hole is at its surface, then information can be emitted from this surface. But what about the intuition that an infalling object see empty space as it falls through the horizon? 't Hooft suggested the idea of *complementarity*: this smooth infall is an equivalent description of the black hole dynamics in a different set of variables.

Thus there are really three different conjectures here:

(i) Some quantum gravity effect makes the information of the hole reside at the horizon, and this information is carried out by Hawking radiation.

(ii) There is an alternative set of variables in which an infalling particle to behave as if it is falling freely through a vacuum horizon.

't Hooft suggested the following mechanism for violating semiclassical physics at the horizon. Infalling particles shift the location of the horizon. But Hawking radiation quanta emerge from this horizon, so the effect of this shift is imprinted on the outgoing radiation. This transfers infomation from the infalling matter to the outgoing radiation, giving a unitary S-matrix for the entire process of black hole formation and evaporation.

Unfortunately, this argument turns out to be too naive; as we will note below, shifting the location of the horizon does not actually imprint any significant information on the emitted quantum. Thus it is not clear how the requirement (i) is to be obtained. It is also not clear how (ii) is to be realized.

In spite of these difficulties, 't Hooft's proposals had a significant impact on the general thinking about black holes, and set the stage for other developments to follow.

Susskind and his collaborators developed the idea of complementarity in more detail. They proposed that the picture of black hole dynamics depends on the observer:

(i) For the purpose of an observer who stays outside the hole, the physics of the black hole can be described using a surface placed just outside the horizon; this surface is called the stretched horizon. All infalling matter is absorbed by this surface, thermalized, and returned to infinity as radiation; thus there is no information loss.

(ii) For the purpose of an observer who falls into the hole, the horizon appears to be a vacuum region; thus this infalling observer carries his information onto the hole.

(iii) The descriptions (i) and (ii) are compatible because a person who falls through the horizon cannot communicate what he sees to the outside.

The difficulties with this proposal are much the same as the ones with 't Hooft's proposal. The no-hair theorem had not been broken at the time of this proposal, so it is not clear what effect would lead to information reflecting of the horizon, as required by (i). The requirement (ii) suggested that there was no 'real' structure at the horizon with an invariant meaning, since in at least one description the horizon was an exact vacuum. It was also not clear what physics would lead to (iii); the standard formulation of quantum theory does not give different realities for observers at different places, even if these observers are unable to communicate with each other in the future.

But as with 't Hooft's proposal, the impact of Susskind's ideas was large, because they focused attention on the tension between the requirements of information retrieval and smooth infall.

To distinguish the above line of thinking from what we will discuss below, we will call it *traditional complementarity.*

(c) *Fuzzball complementarity:* With fuzzballs, we have a very different situation. We have broken the no-hair theorem, and found nontrivial structure at the horizon. In fact we have no region interior to the horizon, since spacetime ends just outside the place where the horizon would have formed. The structure at this horizon is 'real' in the sense that it is a covariant gravitational solution like any other gravitational solution: the horizon region is not a vacuum in any coordinates.

How then can we have any notion of free infall? In [], it was argued that one can still have an *approximate* notion of smooth infall. The idea is as follows:

(i) The fuzzball surface radiates low energy quanta, with energies $E \sim T$, in place of the Hawking radiation from the hole. The state of these quanta depends on the choice of fuzzball state, and so this radiation carries out the information of the black hole microstate.

(ii) But now imagine a quantum with $E \gg T$ impinging on the fuzzball surface. This quantum cannot 'go through' this surface, since there is no interior to go to; the spacetime has ended at this surface. But the energy E carried by the quantum can deform this surface. Let this deformation be characterized by a set of frequencies

$$\{\nu\}^f = \{\nu_1^f, \nu_2^f \dots\} \tag{1.1}$$

(iii) Now consider the semiclassical black hole. The set of quantum dynamical processes associated to this hole will be described by some frequencies; let us call them

$$\{\nu\}^{bh} = \{\nu_1^{bh}, \nu_2^{bh}, \dots\} \tag{1.2}$$

(iv) Now suppose that for a generic fuzzball, we have

$$\{\nu_1^f, \nu_2^f \dots\} \approx \{\nu_1^{bh}, \nu_2^{bh}, \dots\} \tag{1.3}$$

Then the dynamics of the fuzzball surface will reproduce the physics of the traditional hole to a good approximation in the limit $E \gg T$. This is the idea of fuzzball complementarity.

Lecture notes 2

The conjecture of fuzzball complementarity

In fig.**??** we again reproduce our schematic picture of a fuzzball microstate. The spacetime ends in a quantum mess a little outside the place where a horizon would have been in the traditional picture of the black hole.

Suppose an object falls onto the surface of this fuzzball. Since spacetime ends before the horizon is reached, the object cannot pass smoothly through the horizon into the black hole 'interior'. Should we not say that the object hits a barrier at the horizon, and gets destroyed?

Indeed, several people working on fuzzballs did argue that exactly this should happen. Bena, for example has explained the structure of the extremal microstate as follows: "If you fall down the throat of such a microstate, you will crack you head and die"[]. While this looks like a natural statement, our goal here is to ask if there is another possibility. That is, is there some way that we can have the microstate structure of fuzzballs, but yet preserve some approximation to our classical intuition which suggests smooth free fall through the horizon?

To see how such an alternative possibility can arise, let us begin with an analogy: the example of AdS/CFT duality.

2.1 Infall onto a stack D3 branes

If fig.**??**(a) we depict a stack of N D3 branes, all placed at the same location, parallel to each other. We assume that

$$N \gg 1 \tag{2.1}$$

We consider an object that falls onto this stack. For concreteness, we take this object to be a closed string. This object has internal dynamics, for example the closed string has oscillation frequencies

$$\{\nu^{closed}\} = \{\nu_1^c, \nu_2^c, \dots\} \tag{2.2}$$

and is in a state

$$|\psi\rangle = \sum_j C_j^{closed} |E_j^{closed}\rangle \tag{2.3}$$

where $E_j^{closed} = \nu_j^c$ is the energy of the quantum state with frequencies ν_j^c.

When the closed string reaches the stack of D-branes, then its energy gets converted to a collection of open strings. These open strings spread out along the surface of the D3 branes, so the energy that was contained in the closed string now moves out in transverse directions, instead of continuing along the initial direction of infall of the closed string.

One would be tempted to say at this point that the initial closed string has been 'destroyed' upon reaching the stack of D3 branes. After all, it has been split into a large number of open strings, which move off in different directions on the branes. But we know that the dynamics of the D3 branes has an alternative description, depicted in fig.**??**(b). There are no D3 branes visible in this description; instead we have a smooth spacetime geometry. The infalling closed string loop passes smoothly through the 'neck' region of this geometry into an AdS region, continuing to move in its original direction. In this description, it certainly does *not* look as if the closed string has been 'destroyed' in its interaction with the D3 branes.

We would like to understand this phenomenon in more detail, since the physics here will be one of the ingredients that we would like to borrow, though in a somewhat modified form, in arriving at the conjecture of fuzzball complementarity.

The state of the closed string has been converted into a complicated wavefunctional describing the state of a large number of open strings. Let the states in this open string description be $|E_j^{open}\rangle$. The closed string was falling towards the D3 branes with the state (2.3). Upon reaching the D3 branes, its state gets converted to an open string state as

$$\sum_j C_j^{closed}|E_j^{closed}\rangle \;\rightarrow\; \sum_j C_j^{open}|E_j^{open}\rangle \tag{2.4}$$

In general such a change of state would lead a a change of dynamics, and the closed string would cease to exist as a well defined object. But suppose we have

$$E_j^{closed} \approx E_j^{open} \tag{2.5}$$

$$C_j^{closed} \approx C_j^{open} \tag{2.6}$$

then the dynamics of the infalling closed string is, to a first approximation, *not* altered. The states E_j^{closed} have been given a new representation as states E_j^{open} in a new Hilbert space, but the energy levels and their weights have not been altered (eq.(2.5),(2.6)). The internal dynamics of any object, however complicated, is captured by the energy levels and their weights, so all that has happened is that the dynamics has been mapped isomorphically to a new description. In particular, if the closed string was replaced by a person, the person would feel no 'pain' when he falls onto the stack of D3 branes: he would just feel the gentle changes of metric as the spacetime changes from flat space to the 'neck' region and then to an AdS geometry.

2.2 Fuzzball complementarity

Let us now see if we can reproduce some version of the AdS/CFT example for an object that falls onto the surface of a fuzzball.

Let the infalling object again be a closed string in the state (2.3). When this closed string reaches near the fuzzball surface, its energy causes an excitation of this surface. This excitation is analogous, in the AdS/CFT example, to the excitation of open strings on the D3 branes. We can therefore write the absorption of the closed string on the fuzzball surface in a way similar to eq. (2.4)

$$\sum_j C_j^{closed}|E_j^{closed}\rangle \;\rightarrow\; \sum_j C_j^F|F_j\rangle \tag{2.7}$$

where the fuzzball state $|F_j\rangle$ has an energy E_j^F. Suppose we again have

$$E_j^{closed} \approx E_j^F \tag{2.8}$$

$$C_j^{closed} \approx C_j^F \tag{2.9}$$

Then the dynamics of the infalling closed string would be mirrored by the dynamics of the excitations of the fuzzball surface. We would have 'holographically encoded' the state of the infalling closed string onto the fuzzball surface. If the closed string was replaced by a person, this person would feel no 'pain' as he falls onto the fuzzball surface; the state on the RHS of (2.7) would describe his evolution almost as well as the state on the LHS of (2.7).

This is the outline of the idea of fuzzball complementarity. We will now probe it in more detail, finding in the process the various limitations on how such a possibility can be implemented.

2.3 The conditions for fuzzball complementarity

We do not of course know if the conjecture of fuzzball complementarity is true. But assuming for the moment that it is, we note several conditions which have to be satisfied for the fuzzball surface to be able to mimic free infall through the relation (**??**).

The need for fuzzballs

The fuzzball paradigm differs from the traditional picture of the black hole by having 'real' degrees of freedom at the horizon. In other words, the state near $r \approx 2M$ is not the vacuum, but something that depends on the choice of black hole microstate. If we did not have fuzzballs, the infalling object would indeed just pass smoothly though $r = 2M$ into the black hole interior, and there would be no states $|F_i\rangle$ that we have in (2.7).

2.3.1 The approximation $E \gg T$

We have postulated the map (2.7) for the transition from the states of the infalling object to the states of fuzzballs. Thus postulate was made in line with the map (2.4) for the AdS/CFT case. But there is an important difference between the two cases, which goes to the heart of the information paradox.

In the AdS/CFT case the closed string falls through a vacuum spacetime: the flat space region changes to a 'neck' and then to an AdS region, and the changes of the metric are the only effects that the closed string feels. Suppose the size of the closed string is much less that the curvature radius of the AdS. Then these changes of curvature give only mild tidal forces on the closed string, and it is only these mild tidal effects that prevent the $\approx$ sign from being an equality.

In the fuzzball case, however, there is a more basic limitation on how accurate the $\approx$ sign can be in []. Each fuzzball microstate is different, so they cannot all respond to the infalling object in exactly the same way. Thus the map (2.7) cannot be exactly the same for different choices of the initial state of the hole. Clearly (2.7) must be an approximation which ignores the differences between these microstates. What is this approximation?

The information of the fuzzball is carried out by the quanta radiated by the fuzzball, just the way it would be for any normal body. These quanta have energies of order $E \sim T$. Thus when we examine the dynamics at energies $E \sim T$, each fuzzball must look different; in fact the overall state of the radiated quanta from one fuzzball state $|F_i\rangle$ must be orthogonal to the overall state of the radiated quanta from a different fuzzball state $|F_j\rangle$, where we have taken $\langle F_j|F_i\rangle = 0$.

A universal behavior for the fuzzballs, can however emerge if we look at the dynamics in a domain $E \gg T$. Let us now see how such a universal behavior may arise, and how it can approximately mimic infall into the traditional black hole geometry for infalling quanta with $E \gg T$.

2.3.2 A toy model for fuzzball complementarity

We have seen above the the idea of fuzzball complementarity is motivated by the notion of gauge gravity duality. Thus we will start be making a toy model for AdS/CFT duality. We will then extend this toy model to one which exhibits fuzzball complementarity. Thus it should be noted that the model we will arrive at is a toy model which illustrates the idea behind the conjecture; the actual dynamics of fuzzballs is likely to be much more complicated.

We proceed in the following steps:

(i) In AdS/CFT duality, when we let the gravity coupling be at the value of interest, then the CFT is a strongly coupled field theory. But we have computed the radiation from a *weakly* coupled field theory, and observed that it agrees with the radiation from a black hole. While this weakly coupled CFT cannot reproduce all the physics of black holes, let us use it for the moment to make our toy model of AdS/CFT.

In this model we have box of volume V, in which we place a graviton h_{ij}. The black hole will be replaced by an effective string; the oscillations modes of this string give the field theory. The left moving excitations are created by $\hat{a}^\dagger_{n,L}$ and the right moving ones by $\hat{a}^\dagger_{n,R}$. There is a coupling that annihilates the graviton and creates a pair of vibrations

$$H_{int} \sim \sum_{m,n} \hat{A}_h \hat{a}^\dagger_{m,L} \hat{a}^\dagger_{n,R} \tag{2.10}$$

but we assume that there are no interaction between the modes on the string themselves. Thus we have replaced the actual strongly coupled CFT by a free field theory; this is the simplification assumed for our toy model.

(ii) We are interested in the transition from the graviton state to the state of vibrations on the string. Let the state of the graviton be $|h\rangle$ and its energy be E_0. The crucial point is that the states on the string form an almost continuous band, which results in a 'fermi golden rule' absorption of the graviton onto the string. The fact that we have a dense set of level of the string will be very important in what follows. To understand its significance, first consider the case where the modes on the string are replaced by a system with just *one* state $|s\rangle$, also with energy E_0. Let the amplitude for transition per unit time from $|h\rangle$ to $|s\rangle$ be R. We can choose the phase of $|s\rangle$ to make R real. The Hamiltonian of this 2-state system has the form

$$\hat{H} = \begin{pmatrix} E_0 & R \\ R & E_0 \end{pmatrix} \tag{2.11}$$

The eigenvalues and eigenvectors are

$$(E_0 + R): \quad \frac{1}{\sqrt{2}}(1,1), \qquad\qquad (E_0 - R): \quad \frac{1}{\sqrt{2}}(1,-1) \tag{2.12}$$

If we start in the graviton state $|h\rangle = (1,0)$ at $t = 0$, the the subsequent evolution is

$$|\psi(t)\rangle = \frac{1}{2}(1,1)e^{-i(E_0+R)t} + \frac{1}{2}(1,-1)e^{-i(E_0-R)t} = e^{-iE_0 t}\left(\cos(Rt)|h\rangle + \sin(Rt)|s\rangle\right) \tag{2.13}$$

We see that the amplitude oscillates between the two states $|h\rangle$ and $|s\rangle$.

By contrast, if $|s\rangle$ is replaced by a *band* of states, then the amplitude flows from $|h\rangle$ to the band, but it does not flow back from the band to $|h\rangle$. We have already seen the computation for absorption into a band in section (**??**); now let us depict this physics pictorially.

(iii) Let the states of the band be $|s_k\rangle$, with energies

$$E_k = E_0 + k\Delta, \qquad -\infty < k < \infty \tag{2.14}$$

We can choose the phases of the $|s_k\rangle$ so that the amplitude of transition per unit time R_k from $|h\rangle$ to $|s_k\rangle$ is real. The transition will be predominantly to states with $E_k \approx E_0$, and in this region we assume that $R_k \rightarrow$ for each k; this assumption is only for convenience and is not necessary.

For convenience, let us shift all energies by a constant, so that $E_0 = 0$; this is equivalent to removing a phase factor $e^{-iE_0 t}$ from all states. We start in the state $|h\rangle$ at $t = 0$. The transition to the state $|s_k\rangle$ generates the amplitude in this state with phase $\phi_k = 0$, since we have taken R real. This is depicted in fig.**??**(a), where we show $\phi_k \approx 0$ for each k at small times t.

But as more time passes, the amplitude in the state $|s_k\rangle$ evolves due to its energy, as $\sim Exp[-iE_k t]$. The states with $E_k > E_0 = 0$ get a factor $e^{-i\phi_k}$ with phase $\phi_k > 0$, while the states with $E_k < E_0 = 0$ get a factor $e^{-i\phi_k}$ with phase $\phi_k < 0$. This is depicted in fig.**??**(b), where we draw the phase ϕ_k as a function of k at time $t > 0$.

(iv) The above describes the absorption on the initial graviton $|h\rangle$ into the degrees of freedom of the CFT. Let us now consider the gravity dual of this absorption. This dual picture is depicted in fig.**??**, where the graviton enters into the AdS region, passing through the throat of the geometry.

Why should we think of the CFT process as an absorption? In the 2-state system (2.11) the amplitude moved from $|h\rangle$ to $|s\rangle$ and then back to $|h\rangle$. In the case of the band (2.14) each state $|s_k\rangle$ has the amplitude $R^* = R$ to transition back to the state $|h\rangle$. But the evolution $Exp[-iE_k t]$ gives the states $|s_k\rangle$ different phases, and their their contributions to $|h\rangle$ come with these same phases. This leads to a phase cancellation, and therefore the amplitude does not flow back from the band $|s_k\rangle$ to $|h\rangle$.

As t increases, the phases evolve further, as shown in fig.**??**(c). The phase cancellation becomes even stronger, and so it is 'even more difficult' for amplitude to flow back from the band $|s_k\rangle$ to $|h\rangle$. In the gravity picture, this corresponds to the quantum moving further away from the 'neck' from where it could have transitioned to the state $\frac{1}{2}\rangle$; in other words, the quantum moves deeper into the throat.

This is the analogue of AdS/CFT duality in our toy model where we have assumed a free CFT. In the CFT we have a transition

$$|h\rangle \to \sum_k C_k |s_k\rangle \tag{2.15}$$

The dynamics is thus in the evolution of the set $\{C_k\}$

$$C_k \to C_k e^{-iE_k t} \tag{2.16}$$

This evolution has a simple dual description in terms of the progression of the graviton wavefunction down the throat of a dual geometry. The direction down the throat did not exist as a dimension in the CFT itself, but emerged as an aspect of the evolution of the set $\{C_k\}$.

(v) Our goal is to see if a similar effective direction can emerge to describe infall into the interior of the horizon. But before we do that, let us note that the above motion down the AdS throat can be reversed, to make the graviton move *up the throat.* We can start with the CFT state

$$|\psi\rangle = \sum_k C_k |s_k\rangle, \qquad C_k = e^{iE_k t_0} \tag{2.17}$$

These phases are depicted in fig.**??**(d). The evolution (2.16) brings these phases back into alignment after time t_0, at which point the wavefunction can transition back to the state $|h\rangle$. The evolution of the CFT state for $0 < t < t_0$ describes, in the gravity dual, a graviton moving up the AdS throat to the neck, from where it can exit to the flat space region outside. Thus we see that the evolution of CFT states describes both up and down motions in the gravity dual, in a completely symmetric fashion.

(vi) Now suppose that at dome location $r = r_0$ down the AdS throat we have a black hole. What happens when our infalling graviton reaches this location?

In the traditional picture of the hole, the particle just passes into the black hole interior. Such a situation would be very puzzling from the viewpoint of the dual CFT. This is because on thinks of the time t of the CFT as the Schwarzschild time in the gravity dual. This time t goes to infinity as the graviton approaches the black hole horizon; to see the motion inside the horizon we have to switch to Kruskal coordinate time. So it would seem that the CFT captures only the part of the graviton trajectory that lies outside the hole. If the hole indeed has an interior, then the CFT description is incomplete; some other degrees of freedom besides the CFT would be required to describe the full gravity theory.

But with the fuzzball paradigm we do not have this problem. The AdS throat ends in a 'cap'; there is no horizon. Thus the infalling graviton does not pass into a new region that is not captured by the CFT. Thus it is consistent to say that the infalling graviton goes no further than the place where the horizon would have been in the traditional geometry, and that the full motion up and at the cap is captured by the dual CFT.

(vii) We are however interested in exploring if there can emerge some *effective* description where the graviton does appear to continue its infall through a horizon. Note that the cap does not have a unique geometry; on the contrary, there are a large number $Exp[S_{bek}]$ different states possible at the cap. Thus there are a large number of gravitational degrees at the end of the throat, and the infalling graviton will excite these degrees of freedom. We wish to see if we can find an analogue of the AdS/CFT magic using these degrees of freedom. More precisely, the analogy we seek is the following:

(a) In the AdS/CFT evolution described in (i)-(vi) above, the graviton state $|h\rangle$ interacted with a dense band of sates $|s_k\rangle$. This band absorbs the graviton; we had seen that if we just have one state $s\rangle$ then the amplitude oscillates between $|h\rangle$ and $|s\rangle$, but with a band $|s_k\rangle$ we get absorption into the band. In a similar manner, the gravitational states at the cap form a dense set; the level density is high because the entropy S_{bek} is a large number. When the infalling graviton reaches the cap, it will be absorbed into the band of fuzzball states $|F_k\rangle$ localized near the cap.

(b) In the AdS/CFT case we had the transition

$$|h\rangle \to \sum_k C_k|s_k\rangle \tag{2.18}$$

There was no throat direction in the CFT as such, but the evolution of the set $\{C_k\}$ could be mapped to infall into an effective throat. In the fuzzball case, the initial state of the fuzzball $|F_i\rangle$ gets excited upon the infall of the graviton as

$$|F_i\rangle \to \sum_k C_{ik}|F'_k\rangle \tag{2.19}$$

Here the $|F_i\rangle$ on the LHS is a fuzzball of mass M, while the $|F_k\rangle$ on the RHS are fuzzballs of mass $M+E$, where E is the energy of the infalling graviton. The number N_i of fuzzball states at the initial mass M and the number N_f at the final mass $M+E$ are related by

$$\frac{N_f}{N_i} = \frac{Exp[S_{bek}(M+E)]}{Exp[S_{bek}(M)]} \approx \frac{Exp[S_{bek}(M)+\Delta S]}{Exp[S_{bek}(M)]} = e^{\Delta S} \approx e^{\frac{E}{T}} \tag{2.20}$$

Thus for an infalling quantum with $E \gg T$

$$\frac{N_f}{N_i} \gg 1 \tag{2.21}$$

Thus the situation is very similar to the AdS/CFT case, where we had comparatively few graviton states $|h\rangle$ and a much larger number of states $|s_k\rangle$. (In fact we had taken just one graviton state $|h\rangle$ in our toy model.)

(c) In the AdS/CFT case the evolution was contained in the functions $\{C_k(t)\}$. But we could recast this evolution by introducing an emergent direction along a throat, and letting the graviton progress down this throat. In the fuzzball case the evolution in the cap region is contained in the functions $\{C_{ik}(t)\}$. We would now like to see if we can recast this evolution as an effective motion along an emergent direction that describes infall into a black hole interior $r < r_h$, where r_h is the horizon.

Note that there is one novel feature that we must now have that we did not have in the AdS/CFT case. Infall into a horizon is a one-way process; we cannot come out, at least to leading order in the semiclassical approximation. By contrast we had seen in (v) above that in the AdS/CFT case we could equally well both up and down the emergent throat direction. Let us now see how we can modify our model to get an emergent direction where it is easy to travel in, but hard to travel back out.

(viii) Consider the toy model depicted in fig.??. The energy levels are arranged along a hierarchy, labelled by an index $n = 0, 1, 2, \ldots$ increasing along the horizontal axis. At $n = 0$, we have just one energy level, which represents the infalling graviton state $|h\rangle$. The states at $n = 1$ are a subset of the fuzzball states; these states form a closely spaced band, and the amplitude moves from $n = 0$ to the band at $n = 1$ in the manner of the AdS/CFT model described in (i)-(v) above.

Lecture notes 3

The firewall argument

In 2012, an interesting argument was made about the nature of infall into black holes. This argument came to be known as the 'firewall problem'. Let us discuss this argument.

We already know from the Hawking theorem that if the region around the horizon is the vacuum, then there is a problem with the growing entanglement between emitted quanta and the remaining hole. We have also seen that the problem is removed if we have a fuzzball structure which replaces the traditional horizon; the hole then radiates from this surface like a normal body. For the purposes of the firewall argument, however, we do not need any particular structure or dynamics of the hole; all we need to assume is that some mechanism causes information to be radiated from the surface of the hole, just the way it would be radiated from a normal body. Then, the argument says, *an infalling observer will necessarily have a destructive impact with the surface of such a hole; i.e., it cannot pass through the horizon region 'without drama'.*

The destructive impact is supposed to arise from the interaction between the infalling observer and the Hawking radiation emitted by the hole. Recall that the Hawking radiation quanta are very low energy: a solar mass black hole will emit quanta with wavelength $\sim 3\,\text{km}$. Thus the surprising part of the argument is that the infalling observer will find himself interacting with quanta of higher and higher energies as he approaches the horizon, with the energy of these quanta becoming of order the planck energy when the observer reaches within planck distance of the horizon.

If such an argument were true, then it would rule out the conjecture of fuzzball complementarity, which suggests that observers falling freely from afar feel nothing novel as they reach the horizon. But as we will see below, the firewall argument starts with an extra assumption:

Assumption: *No novel physical phenomena will arise outside the surface of the hole; i.e., the physics outside this surface will be normal low energy physics for all processes. In particular, the response of the black hole surface to an infalling object is causal; i.e., the surface does not distort until null rays from the infalling object reach it.*

While such an assumption at first looks quite natural, we will see that it conflicts with the expected behavior of black holes. More precisely, if we make this assumption, then information cannot come out of the hole without violating causality. Thus it is not clear if the firewall argument can be valid in any known theory of gravity.

In spite of this problem, the firewall argument is very useful, since it puts constraints on conjectures like the conjecture of fuzzball complementarity; one gets conditions on when and where the tunneling behavior discussed in section **??** must start. For this reason, we will describe the firewall argument below, and spend some time in studying how its assumptions conflict with causality.

3.1 The firewall claim

Consider the following three postulates about the bahvior of a black hole:

Postulate 1: The process of formation and evaporation of a black hole, as viewed by a distant observer, can be described entirely within the context of standard quantum theory. In particular, there exists a unitary S-matrix which describes the evolution from infalling matter to outgoing Hawking-like radiation.

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

Postulate 3: A freely falling observer experiences nothing out of the ordinary when crossing the horizon.

The firewall argument says that after the time that the evaporation of the hole has proceeded past its halfway point, *all three of these postulates cannot be true simultaneously.*

Let us see what these postulates say:

(i) Postulate 1 is just the requirement that the black hole radiate like a normal body. In particular, the entanglement between the radiation and the hole should not keep increasing; after the halfway point, it should start to decrease, in line with the arguments of Page [**?**] which we had noted in section [].

(ii) Postulate 2 is the new assumption that we referred to above. Let the black hole horizon be at the location $r = r_h$. Consider a surface that is one planck length outside this horizon; i.e., at the location

$$r_s = r_h + l_p \tag{3.1}$$

We call this surface the 'stretched horizon'. For our present purposes it represents the boundary of the black hole: we assume that any novel physics associated with the black hole ends at a distance $\sim l_p$ outside the horizon. Thus for $r > r_s$, all physics is normal low energy physics; in particular, since the curvature here is low, there are no large effects arising from quantum gravity.

(iii) Postulate 3 is just the requirement of complementarity: an infalling observer should detect nothing abnormal as he crosses the horizon.

Thus the firewall argument says that if we assume (i) that some phenomenon resolves the information problem, by making the black hole radiate like a normal body, and (ii) that no novel phenomena arise outside the hole, then we cannot have (iii); i.e., an infalling observer will not be able to fall 'without drama' through the horizon of the hole.

3.2 What the firewall argument is *not*

The firewall argument, when first proposed, caused a fair amount of confusion. In particular, many people confused it with Hawking's original information paradox of 1975. Thus let us begin by first clarifying what the argument does *not* try to say:

(A) Some people thought that the firewall argument was aimed at proving that there was nontrivial structure at the horizon. But this is not true, since the need for nontrivial structure at the horizon is the content of Hawking's original argument of 1975. More precisely, Hawking had argued that

H: If the horizon is a vacuum region, then there is a problem with monotonically growing entanglement between the radiation and the hole.

This is exactly equivalent to the statement:

H': If we do not want this monotonically growing entanglement, then the horizon cannot be a vacuum region; i.e., there must be some nontrivial structure at its location.

One might think that the required horizon structure could be a small perturbation to the vacuum horizon, but we have ruled this out by the small correction theorem; this theorem had the corollary:

C: If we wish to prevent the monotonically growing entanglement at the horizon, then the corrections to low energy physics at the horizon *must be order unity*

In short, we already know that we need order unity corrections at the horizon; thus this cannot be what the firewall argument is claiming.

(B) Some people thought that the firewall argument was giving a new derivation of Hawking's original argument. But this is not the case either; the argument uses the same model of entangled bits that was used to prove a rigorous form of the Hawking argument in the small corrections theorem.

(C) Some people thought that the firewall argument gives a construction of a firewall at the horizon. But this is not the case; the argument does not concern itself with any details of quantum gravity. Rather, the argument says that *if* someone has a theory in which there are degrees of freedom at the horizon emitting radiation, then an infalling object will be burnt by this radiation. In particular, we have seen that fuzzballs emit radiation from degrees of freedom at the location of the horizon. Thus the question becomes: will a fuzzball behave like a firewall? If the firewall argument were true, then an observer falling onto a fuzzball would get burnt before reaching the surface of the fuzzball. As we will see however, the firewall assumption has a conflicting set of assumptions, and fuzzballs are not expected to obey these assumptions.

Now let us turn to what the argument was actually trying to do. As noted above, the Hawking theorem tells us that we need order unity corrections at the horizon of we are to resolve the problem of monotonically growing entanglement. But what is the nature of these order unity corrections? For a black hole of radius r_h, the Hawking quanta also have a wavelength $\lambda \sim r_h$. Thus all that is required by the small corrections theorem is that there be a change of order unity in the state of modes with $\lambda \sim r_h$. The firewall argument seeks to extend this conclusion by arguing that an infalling observer will feel the 'heat' from quanta of all wavelengths $l_p \lesssim \lambda \lesssim r_h$. The price to be paid for this extension is the additional assumption mentioned above, which we will need to examine in more detail.

3.3 The intuition behind the firewall argument

The basic idea behind the firewall argument can be seen from fig.**??**. In fig.**??**(a) we depict the black hole as an object whose physics we do not know. This region with unknown physics is bounded by the stretched horizon, which is drawn as a solid line. The hole radiates quanta from the region $r \le r_s$ which escape out to infinity.

Outside the stretched horizon, postulate 2 tells us that the physics is just that of quantum fields on curved space; the effects of quantum gravity are negligible. The curvature of this space does have one important effect: the redshift at a radius r is given by

$$R = (1 - \frac{2M}{r})^{\frac{1}{2}} \tag{3.2}$$

Now consider the quanta of Hawking radiation. When these quanta are near infinity, they have a wavelength $\lambda_\infty \sim r_h \sim M$. But now imagine following these quanta back to a location r closer to the horizon. Due to the redshift, the wavelength at radius r is

$$\lambda(r) \sim \lambda_\infty R \sim (1 - \frac{2M}{r})^{\frac{1}{2}} M \tag{3.3}$$

Thus the quanta become blueshifted to higher energies as we approach the horizon, and at the stretched horizon $r_s = 2M + l_p$, they have a wavelength

$$\lambda_s \sim l_p \tag{3.4}$$

An infalling object will interact with these high energy quanta as he falls closer and closer to the stretched horizon. He will therefore get burnt by these high energy quanta, and cannot pass smoothly through the horizon into a black hole interior.

One might ask: why does a similar burning not happen with Hawking's original picture of black hole radiation? This picture is depicted in fig.**??**(b). The Hawking quanta again have wavelength $\lambda \sim M$ at infinity. But as we move to the region

$$r - 2M \lesssim M \tag{3.5}$$

there are *no* radiation quanta. Rather, we just have a region which is locally the vacuum. We can say that near the horizon, the outgoing quanta pair up with their infalling partners to 'cancel' out. This is of course just a restatement of the fact that the region around the horizon is the vacuum, and the vacuum modes transition to 'real quanta' only in the region $r - 2M \gtrsim M$.

Thus the firewall argument is making the following observation. In Hawking's picture of pair creation, the region near the horizon is the vacuum, and the radiation quanta materialize as 'real' particles only far away from the horizon. But this process of radiation creates the problem of monotonically growing entanglement. Suppose we did *not* want to have this monotonically growing entanglement; instead we wanted the entanglement to start decreasing after the half-way point of evaporation, the way it would for a normal body. Suppose in addition we ask that all novel black hole dynamics be confined to within the stretched horizon r_s. Then the surface of this stretched horizon would act like the surface of a hot body, emitting quanta to the region $r > r_s$. Such quanta are 'real' quanta from the moment they are radiated, and travel out to infinity just like any normal particle would. The energy of these quanta is very high near $r = r_s$; these high energy quanta then climb out of the gravitational field of the hole and reach infinity with wavelength $\lambda \sim M$. An infalling observer will thus encounter high energy outgoing quanta as he come near the horizon, and consequently get burnt.

3.4 The argument in detail

Let us now describe the firewall argument in detail. There are many ways of presenting the argument; we will choose one that aligns most closely with the notation and language we have used in studying the information paradox and complementarity.

The goal of the firewall argument is to rule out the possibility of complementarity. That is, if in one description the black hole radiates information from its surface, then there cannot be an alternate description where the horizon looks like a vacuum region.

We proceed in the following steps:

(a) Postulate 1 requires that radiation from the hole be like that from a normal body. Thus the entanglement of the hole with its radiation should start going down after the halfway point, in line with the arguments of Page. If, on the other hand, we find that the entanglement keeps going up, even after the halfway point, then we will have a contradiction with postulate 1.

(b) Hawking's computation of radiation found that the entanglement keep going up monotonically. So let us see if we have the conditions assumed in Hawking's computation in our present situation.

(c) By postulate 3, there should be a spacetime patch around the horizon where the quantum state is that of the vacuum. We depict this in fig.**??**(a). We consider a mode b outside the horizon and a mode c inside the horizon; these are the same modes that in the Hawking computation, evolved into entangled particle pairs. To make the quantum state in our patch the local vacuum, the states in these modes b, c must be entangled. We use the same toy model of entangled bits that we used in the Hawking computation, letting the state in these modes be

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle_{b_{N+1}}|0\rangle_{c_{N+1}} + |1\rangle_{b_{N+1}}|1\rangle_{c_{N+1}}\right) \tag{3.6}$$

We have added a subscript $N+1$ to the b, c modes to signify that N steps of emission have already occurred. We assume that N is large enough that we are past the halfway point of evaporation, so that we need the entanglement to decrease rather than increase at step $N+1$.

(d) By postulate 2, the physics outside the stretched horizon is just normal semiclassical physics. The mode b_{N+1} is outside the stretched horizon, and so will evolve to evolve towards infinity just the way it did in Hawking's computation. When b_{N+1} reaches near infinity, the particles in it will be normal local particles.

(e) The entanglement of b_{N+1} with c_{N+1} was given by $S_{ent} = \ln 2$. We do not care about the future evolution of the mode c_{N+1}; whatever be the dynamics inside the hole, it does not affect the physics outside the stretched horizon by postulate 2. Thus when b_{N+1} reaches infinity, the entanglement of the radiation with the hole has gone *up* by $\ln 2$, instead of going down. This is in contradiction with the requirement of postulate 1, noted in step (a) above.

(f) Thus we see that postulates 1, 2 and 3 cannot all be true. Suppose we assume postulate 1 (that there is some mechanism that resolves the Hawking entanglement problem) and also postulate 3 (that all novel physics is confined within the stretched horizon). Then we cannot have postulate 2; i..e, we cannot have any description of the physics where the state near the hole is the vacuum. In other words, we cannot have complementarity.

But if the state of the modes b_{N+1}, c_{N+1} is not the vacuum state (3.6), then it contains particles in these modes. We can consider these modes when their wavelengths are very small. For any wavelength λ, we can consider the mode b_{N+1} as a wavepacket localized at a distance $\sim \lambda$ from the horizon, and the mode c_{N+1} will be at the same distance inside the horizon. We can choose any wavelength $\lambda \gtrsim l_p$, since the stretched horizon is at a distance l_p from the horizon, and outside this stretched horizon we have assumed the normal semiclassical behavior of modes by postulate 2. We will therefore have a nonzero particle number in all modes near the horizon, with

$$l_p \lesssim \lambda \lesssim M \tag{3.7}$$

where the upper limit gives the wavelength of the Hawking particles when they have left the vicinity of the horizon. An infallling observer will encounter with very high energy particles as he comes close to the horizon. Since physics here is assumed to be normal semiclassical physics, we conclude that he will get 'burnt' by these high energy particles. This is the firewall argument.

3.5 The flaw in the firewall argument

Let us now see the problem with the firewall argument. We will see that while the argument itself is technically correct, the difficulty lies with the starting assumptions. We proceed in the following steps:

(i) Start with a black hole of radius $r_h = 2M$. The stretched horizon is at $r_s = 2M + l_p$.

(ii) Consider a spherical shell of mass δM falling onto the surface of this hole. We let this shell be composed of massless quanta moving radially inwards. Then this shell moves in at the speed of light. As a consequence, no signal from this shell reaches the stretched horizon until the shell itself actually hits the stretched horizon. This infalling shell is depicted in fig.**??**(a).

(iii) By postulate 2 of the firewall argument, the physics in the region $r > r_s$ is normal semiclassical physics. Thus nothing unusual happens as the shell moves from $r = \infty$ to $r = r_s$; we just get the classical geometry created by the infall of the shell δM onto the hole of mass M. In particular, since causality is obeyed in the classical geometry, the stretched horizon cannot not move out to $r > r_s$ until after the shell reaches $r = r_s$.

(iv) Consider the situation where the shell of mass δM is just about to reach the the stretched horizon of our black hole of mass M. The total mass in the region $r \le r_s$ is now

$$M_T = M + \delta M \tag{3.8}$$

The horizon radius for this mass is

$$r_T = 2M + 2\delta M \tag{3.9}$$

Recall that $r_s = 2M + l_p$. We can easily ensure that

$$r_T > r_s \tag{3.10}$$

by taking

$$\delta M \gg m_p \tag{3.11}$$

Thus we have the following situation: we have a mass M_T which is well inside its Schwarzshild radius r_T. In fact the mass is confined inside a radius r_s, which is close to the horizon radius of the starting hole of mass M, while the new horizon radius corresponds to a mass $M + \delta M$. This is depicted in fig.**??**(b)

(v) Now consider the region

$$r_s < r < r_T \tag{3.12}$$

In this region the metric is given by usual Einstein gravity. This is because the original stretched horizon could not have responded in any way (since it has not yet been touched by the infalling shell), and the physics in the region $r > r_s$ is assumed to be given by normal semiclassical physics. Thus in this region the light cones point inwards, just the way they would in the interior of any horizon.

(vi) We now see that the information in the shell δM is causally trapped. If we want this information to get out of the location r_s to the outside region $r > r_T$ then we need to make this information move outside the light cones; i.e., we need the information carriers to move faster than the speed of light.

But if we postulate that information can travel outside the light cone, then there was no information puzzle in the first place.

But most theories of quantum gravity respect causality, so they do *not* have superluminal motion. In any theory that respects causality, the information in the shell δM can never emerge till this shell stops being trapped inside a horizon. Note that if we have causality, then the stretched horizon cannot even continue to stay at the location $r = r_s$; once the light cones tilt inwards as in fig.**??**(b), the structure at the stretched horizon has to fall monotonically to smaller r values. In fact since the original hole and the shell are all now inside the horizon r_T, we have the same behavior that we had for matter in a classical black hole, where everything inside the horizon ends up at the singularity.

(vii)

3.6 How fuzzball complementarity bypasses the firewall argument

The firewall argument says that we cannot have complementarity; i.e., we cannot have the sense of free infall if we preserve postulates 1 and 2. Yet we have argued in section **??** that it is possible to get information out of the hole, and yet have *fuzzball complementarity*, where observers who fall in freely with energies $E \gg T$ see approximately free infall. Does the firewall argument rule out the conjecture of fuzzball complementarity? We will see that the answer is no. We proceed in the folloiwng steps:

(i) We start by observing that the spirit of fuzzball compelmentarity is very different from the spirit of the complementarity that the firewall argument was addressing. The firewall argument addresses 'traditional complementarity', where in some description the region around the horizon is the vacuum. In paricular the argument focuses of the field modes that will emerge to make Hawking radiation quanta; thus we call these $E \sim T$ modes. With fuzzballs, we already know that the region around the horizon is not the vacuum. Thus what we seek to do is get an effective dynamics of the fuzzball in the limit

$$E \gg T \tag{3.13}$$

and ask if *in this approximation* the dynamics of the fuzzball can mimic free infall. The modes $E \sim T$ are *not* in the vacuum state, since they are the ones that will carry the information of the fuzzball state out to unfinity. The firewall argument has no condition like (3.13), and so does not address the idea of fuzzball complementarity.

(ii) One may still ask if the firewall argument can be extended in some way to rule out the idea of fuzzball complementarity. In other words, how does the limit (3.13) relate to the steps in the firewall argument?

Consider a particle with $E \gg T$ falling towards the horizon. We have seen above that the difdficulty with the firwall argument is postulate 2, which says that no novel effects happen before an infalling particle reaches the stretched horizon. By contrast, the usual classical horizon extends out to meet the incoming particle; thus this horizon meets the infalling particle before the particle reaches the stretched horizon. In fuzzball complementarity, we conjecture that tuneling into fuzzballs happens just *before* the particle would have fallen through its horizon; this is how we evade the causaility problems that arise if the particle does get trapped behind a horizon. Thus we violate postulate 2 of the argument.

(iii) In more detail, we can make the following estimates. An infalling particle of energy E will lead to an increase in entropy by

$$\Delta S_{bek} \approx \frac{E}{T} \tag{3.14}$$

We assume that the entropy S is reflected in the surface area A of the fuzzball, through the usual relation $S_{bek} = \frac{A}{4G}$. Then the increase (3.14) implies an increase in the area

$$\Delta A \approx \frac{4GE}{T} \tag{3.15}$$

For a rough approximation, let us assume that the deformation of the horizon is of the shape of a hemisphere with area ΔA. Suppose we are dealing with a Schwarzschild hole in D spacetime dimensions. The radius of this hemisphere is

$$s_h \sim (\Delta A)^{\frac{1}{D-2}} \sim \left(\frac{GE}{T}\right)^{\frac{1}{D-2}} \sim \left(\frac{E}{T}\right)^{\frac{1}{D-2}} l_p \tag{3.16}$$

where l_p is the planck length in D spacetime dimensions, defined through $G = l_p^{D-2}$. Thus the transition to fuzzballs happens at a distance $\sim s_h$ from the original horizon.

Now we note that the temperature of the radiation drops with the distance from the horizon. At a distance R from the horizon, this temperature is

$$T \sim \frac{1}{s_h} \tag{3.17}$$

The firewall argument relies of the fact that an infalling object will burn up by interaction with this raditon before new effects (that lead to complementarity) can start. But as E increases, we see that s_h increases and T drops. Consider a quantum that is sent towards the hole, starting with energy E at infinity and falling to a distance s from the horizon. Let $\bar{s}$ be the value of s at which the probaibility of interacton with the radiation becomes order unity. We find []

$$\bar{s} \sim \left(\frac{E}{T}\right)^{\frac{1}{2(D-2)}} l_p \tag{3.18}$$

Assuming that $D > 2$, we see that for $E \gg T$,

$$s_h \gg \bar{s} \tag{3.19}$$

Thus the transition to fuzzballs happens before significant interaction with the radiation can occur. We therefore see that the firewall argument cannot rule out the possibility of fuzzball complementarity.

(iv) To put all this in another language, the firewall argument and the conjecture of fuzzball complementarity are focusing on quite different degrees of freedom. The firewall argument is concerned with the $E \sim T$ modes that will emerge as Hawking, and focuses on how these modes are entangled with other modes in the problem. Fuzzball complementarity, on the other hand, is concerned with the dynamics of *new* degrees of freedom that are created when an energetic quantum falls onto the hole. The number of these new degrees of freedom is given by (3.14). Since the entropy is the logarithm of the number of states, the number of states N_f after the addition of energy E is related to the number N_i before the infall by

$$\frac{N_f}{N_i} = \frac{Exp[S_{bek}(M+E)]}{Exp[S_{bek}(M)]} \approx \frac{Exp[S_{bek}(M)+\Delta S_{bek}]}{Exp[S_{bek}(M)]} = e^{\Delta Sbek} \approx e^{\frac{E}{T}} \qquad (3.20)$$

where M is mass of the hole before infall and M_E is the mass after. If $E \gg T$,, we see that

$$\frac{N_f}{N_i} \gg 1 \qquad (3.21)$$

so that most of the states of the hole after infall are *new* states that were not accessible to the system before the infall. These new states are not entangled with anything outside the fuzzball, and so are not subject to the statements about entanglement used in teh firewall argument. But it is the dynamics of these new states that is conjectured to mimic 'approximately free infall' for the infalling quantum, so we see that the degrees of freedom involved in the firewall argument and the degrees of freedom responsible for fuzzball complementarity are quite different.

Lecture notes 4

Unruh radiation

In the mid seventies, an interesting effect was discovered by Fulling [], Davies [] and Unruh []. Take flat spacetime, and consider a quantum field ϕ on this spacetime. Let ϕ be in the vacuum state; i.e., there are no particles of the field ϕ in our state. Thus

$$\hat{a}_{\vec{k}}|0\rangle_M = 0 \tag{4.1}$$

where the subscript M indicates that $|0\rangle_M$ is the vacuum of Minkowski space.

Now consider a detector that can detect particles of the field ϕ. If this detector is kept at rest, it will detect no particles. It will also detect no particles if it is moving with constant velocity, since the vacuum $|0\rangle_M$ is a Lorentz invariant state. But suppose we let the detector move with an *acceleration*. Then the detector detects particles of ϕ; in fact it detects a thermal bath of such particles, with a temperature T that is proportional to the acceleration a. This phenomenon is called the Unruh effect.

Why does the accelerating detector see particles? At a qualitative level, the physics is similar to that of *bremsstrahlung* – the radiation of photons from an accelerating electron. The electron couples to the photon field A, and so creates in its vicinity a distortion δA photon field vacuum. For the electron at rest let the deformation be $\delta A_{(0)}$. No photons are radiated by this deformation; after all, we cannot extract any energy from the electron which is an elementary particle. The same holds for an electron moving with a constant velocity v, since we can go to a frame where this electron will appear at rest. Let the deformation in this case be $\delta A_{(v)}$.

Now suppose the electron was at rest, and was then given a kick so that it changed its velocity to v. The deformation of the vacuum $\delta A_{(0)}$ cannot change immediately to $\delta A_{(v)}$. The deformation we have, $\delta A_{(0)}$, can be written as the new required deformation $\delta A_{(v)}$ plus some additional photon excitations. These additional excitations escape, giving the bremsstrahlung radiation from the accelerated electron.

A similar situation holds with our particle detector. Since this detector must measure particles of the ϕ field, it couples to this field and creates a distortion $\delta\phi$ of ϕ field in its vicinity. If the detector is kept at rest, then the state of the detector ψ_D and the deformation of the field $\delta\phi_{(0)}$ settle to an equilibrium state where we say that no particles are being detected. The same situation holds if the detector is in uniform motion, but the vacuum deformation $\delta\phi_{(v)}$ is different from $\delta\phi_{(0)}$. If the detector accelerates, then the deformation $\delta\phi_{(0)}$ cannot change to $\delta\phi_{(v)}$ immediately, and the difference shows up as quanta of the ϕ field. The new feature this time – not present in the bremsstrahlung discussion – is that the detector is not an elementary particle. Instead, it can interact with a ϕ quantum and get excited to a new internal state; we say that this excitation corresponds to the detection of a ϕ particle by the detector. In particular, the detector can get excited by the ϕ quanta that arise from the difference between $\delta\phi_{(0)}$ and $\delta\phi_{(v)}$. Thus we say that an accelerating detector sees ϕ particles even though we did not start with any ϕ particles in our state $|0\rangle_M$.

While this phenomenon looks straightforward enough, the surprise is in the details of the detection rate. If the detector has a constant acceleration a, then the ϕ quanta that it sees form a *thermal* distribution, with temperature

$$T = \frac{a}{2\pi} \tag{4.2}$$

Moreover, this temperature can be related to the temperature we would attribute to the Hawking radiation from a black hole, if we assumed that the hole radiated thermally from a surface placed at the horizon. This relation between the Unruh effect and black hole physics has led to an effort to relate the two effects, and this effort in turn has led to deep puzzles about black holes. As we will see, one of the lessons from the fuzzball paradigm is that these two effects are *not* as closely related as had been imagined. Let us now look at these issues in more detail.

4.1 The Rindler vacuum

Consider the Minkowski metric

$$ds^2 = -dT^2 + dX^2 + dY^2 + dZ^2 \tag{4.3}$$

We will let the acceleration be in the X direction; thus we just consider the T, X plane in what follows. It is useful to recall Rindler coordinates which cover the 'right wedge' of the T, X Minkowski spacetime

$$T = r\sinh t, \quad X = r\cosh t \tag{4.4}$$

in which the metric becomes

$$ds^2 = -r^2 dt^2 + dr^2 \tag{4.5}$$

The trajectories $r = r_0 = const$ are depicted in fig.**??**; these are paths of constant acceleration. To see this, note the proper distance along such a trajectory is

$$d\tau = r_0 dt \tag{4.6}$$

The proper velocity has components

$$U^T = \frac{dT}{d\tau} = \frac{dT}{dt}\frac{dt}{d\tau} = \cosh t, \quad U^X = \frac{dX}{d\tau} = \sinh t \tag{4.7}$$

The components of the acceleration are

$$a^T = \frac{dU^T}{d\tau} = \frac{1}{r_0}\sinh t, \quad a^X = \frac{dU^X}{d\tau} = \frac{1}{r_0}\cosh t \tag{4.8}$$

Thus

$$a^\mu a_\mu = -(a^T)^2 + (a^X)^2 = \frac{1}{r_0^2} \tag{4.9}$$

Thus the magnitude of the acceleration is

$$a = \frac{1}{r_0} \tag{4.10}$$

Thus the acceleration is higher for smaller r_0.

The Unruh computation showed that a detector moving along the trajectory $r = r_0$ picked up excitations as if it was immersed in a thermal bath with temperature

$$T = \frac{a}{2\pi} = \frac{1}{2\pi r_0} \tag{4.11}$$

Thus the temperature becomes higher as we approach $r = 0$. Let us now see how these observations relate to the black hole.

4.2 The near horizon region of the black hole

The metric of the Schwarzschild black hole is given by (**??**). We have already seen in section **??** that in the region

$$r = 2M + \epsilon, \quad 0 < \epsilon \ll M \tag{4.12}$$

the metric of this hole looks like the metric of Minkowski space expressed in Rindler coordinates

$$ds^2 = -r^2dt^2 + dr^2 + dy_1^2 + dy_2^2 \tag{4.13}$$

where

$$r = \tag{4.14}$$

Bibliography