Lecture notes 1

Quantum field theory

1.1 Quantum field theory

What is the nature of the vacuum?

We have noted that in quantum theory the vacuum will fluctuate, and these fluctuations will take the form of particle pairs that appear and disappear. Since it is these pairs that lead to Hawking radiation and the consequent paradox, it is vital to understand the quantum vacuum. To do this, we will review the structure of quantum field theory – the theory that emerges when we put together quantum mechanics and special relativity. We will find however that there are some deep unresolved difficulties with the picture of the vacuum in quantum field theory, and these difficulties become even more severe when incorporate gravity into the theory.

1.1.1 A theory of many particles

Let us start with nonrelativistic quantum mechanics. Consider a particle in one dimension x. The particle is described by a wavefunction $\psi(x)$. Suppose we have a device which can perform a measurement to check if this particle is between x_1 and x_1+dx. The probability that the device finds the particle in this interval is $|\psi(x_1)|^2dx$. The overall probability for the particle to be somewhere is unity

$$\int_{-\infty}^{\infty}|\psi(x)|^2dx = 1 \tag{1.1}$$

Suppose we take two measuring devices, one of which checks for the particle in the interval (x_1, x_1+dx) and the other checks for the particle in the interval (x_2, x_2+dx). One of these devices may find a particle, or neither may find a particle, but it cannot be the case that *both* find a particle. The reason is simple: there is only one particle overall. If the measurement collapses the wavefunction to the location x_1, then this collapsed wavefunction satisfies (1.1) around the location x_1, leaving no probability for the particle to be at x_2.

We can immediately see however that this feature must fail in a relativistic theory. Suppose the first device makes its measurement at position x_1 and time $t=0$. A second device at position x_2 does not immediately find out that such a measurement has been made; after all, information cannot travel faster than the speed of light. Thus if it is possible for the second device to detect the particle, then it must be possible to make this detection at $t=0$ regardless of whether the first device detected a particle or not. We conclude that there must be a nonzero probability that both devices detect particles if their separation is a spacelike interval.

Where does the second particle come from? Fortunately in a relativistic theory this is not a problem; the process of measurement requires some energy E, and this energy can create a particle of mass m if $E \geq mc^2$. Thus a detector may either pick up a particle which was already present, or create a new one in the process of detection.

An immediate consequence of this observation is that relativistic quantum theory must necessarily be a theory of many particles. But if we have more than one particle, then we are led to the strange way in which many-particle states are counted in quantum theory. Let us now turn to this statistics of particles.

1.1.2 Counting boson and fermion states

Let us start with classical physics. Suppose we have two balls, and we toss them randomly into a set of two bins. There are 4 ways that the balls could land in the bins:

(1) Both balls in bin 1
(2) Both balls in bin 2
(3) The first ball in bin 1 and the second ball in bin 2
(4) The second ball in bin 1 and the first ball in bin 2

Assuming that each outcome has the same probability, we find that the probability for the balls to end up in the same bin in $2/4 = 1/2$. This way of counting, appropriate to classical particles, is called Maxwell-Boltzmann statistics.

In quantum theory, it turns out that we need to count in a somewhat different way. Take two identical particles, and again toss them into two bins. If these particles are bosons, then we get the possibilities:

(1) Both particles in bin 1
(2) Both particles in bin 2
(3) One particle in bin 1 and one particle in bin 2

Thus the probability that both bosons are in the same bin is $2/3$ which is larger than the value $1/2$ we had with classical particles. We say that 'bosons try to cluster together in the same state'. But how did we lose possibility (4), where the two particles would have been interchanged between the bins? The traditional argument is: "The particles are identical, so we cannot tell if they have been interchanged. And if we cannot tell that they have been interchanged, then we should not count their interchanged configuration as a new state."

While this conclusion is correct, and agrees perfectly with observations, the argument itself may not sound very convincing. We may not be able to tell the difference between the particles, but would the particles behave differently just because of this inability on our part? It might seem that this odd way

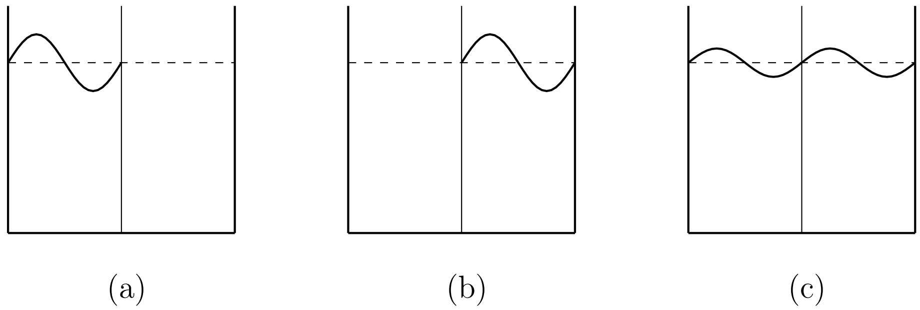

Figure 1.1: caption ...

of counting of particle states must have its origin in the mysteries of quantum theory. But as we will now see, many everyday classical systems exhibit the *same* counting as the example of bosons above.

Consider a tank of water, as shown in fig.**??**(a). The surface of this water can have ripples, which are the objects we will count. To mimic the above example of balls thrown into bins, we discretize our system as follows. We divide the tank into a left and a right half, to make two 'bins'. In each bin we can have ripples, but we allow the ripples to have discrete heights: $1\,\mathrm{mm}, 2\,\mathrm{mm}, \ldots$. A ripple with height $1\,\mathrm{mm}$ corresponds to one particle in the bin, a ripple with height $2\,\mathrm{mm}$ corresponds to two particles and so on.

Now let us count all states which have a total of two ripples. We find the possibilities:

(1) A ripple with height $2\,\mathrm{mm}$ in the left bin.
(2) A ripple with height $2\,\mathrm{mm}$ in the right bin.
(3) A ripple with height $1\,\mathrm{mm}$ in the left bin and a ripple with height $1\,\mathrm{mm}$ in the right bin.

We see that the counting is just like the one for bosons, though there is nothing quantum about this problem! A little reflection shows how this example of ripples differers from the example of classical balls in bins. The balls were *objects* that could be put in different places. But the ripples are not really 'objects'; they are just deformations of an underlying medium – the water in the tank. Thus it makes no sense to "put the ripple in the left tank into the right tank and the ripple in the right tank into the left tank". As a consequence we end up with only the three possibilities (a)-(c) that are analogous to the possibilities that we used to count the states of bosons in our bins.

But this suggests that the bosons should also be thought of an ripples on some medium; if we do that, then the way we count their states would be completely natural. The question of course is: what is this medium? To answer this, we postulate the existence of a 'quantum field ϕ'; excitations of this field will give our bosons, and the theory of excitations of this quantum field is called 'quantum field theory'.

1.1.3 A model for the field

The photon is also a boson – it has spin equal to 1, an integer – and we need a field for describing photons as well. This field will be called the electromagnetic field, and if we follow the discussion above, photons will be excitations of this field.

At this point the reader may get worried by this picture. Photons are particles of light, and we are asking light to move as waves on a medium. But doesn't this sound just like the theory of ether? In that theory, light was described by ripples in a hypothetical medium called ether. That picture was in conflict with Einstein's theory of relativity, and was disproved by the Michelson-Morley experiment. Can we model photons as excitations of a field, and still have a theory where the speed of light is the same in all uniformly moving frames?

The ether theory was in conflict with relativity because we were trying to keep our traditional idea of a fixed universal time. In this traditional approach one would ask: in which frame is the ether at rest? This rest frame would be a preferred frame, since in this frame the velocity of light would be the same in all directions. If we go to a frame moving with some velocity v compared to this rest frame, then the ether would appear to be moving with velocity $-v$ in the new frame. Then light waves would travel at speeds $c+v$ along the direction of this flow, and with speeds $c-v$ in the direction opposite to the flow. Checking the speed of light in different directions will exhibit this asymmetry. This asymmetry is what the Michelson-Morley experiment looked for, and failed to find.

But this difficulty with having a medium only arises if we keep a frame-independent notion of time. As we will now see, it is in fact quite easy to make a simple mechanical model for the field which will be consistent with relativity. Consider a set of pointlike masses m, placed in a row with spacing a. The masses are joined by springs with spring constant k and relaxed length a.

The mass at any position can be displaced in the x direction by a small amount. We will use the symbol ϕ to denote this displacement, and use the equilibrium location x of that particular point mass to say which mass we are displacing. Thus the displacement will be given by a function $\phi(x)$, where at this stage x can take only discrete values $x = na$. We now imagine a limit where a is small, so we have an almost continuous distribution of mass points. We then get an almost continuous function $\phi(x)$, which we will call our scalar field. What is the evolution of this field $\phi(x,t)$?

It is not hard to guess this evolution. This chain of masses and springs sustains waves, which are described by the equation

$$\frac{\partial^2\phi}{\partial t^2} - c^2\frac{\partial^2\phi}{\partial x^2} = 0 \tag{1.2}$$

where

$$c = \tag{1.3}$$

This equation has solutions of the form $\phi = f(x-ct)+g(x+ct)$, which describe waves moving to the left and right with speed c.

If we consider the change of variables familiar from special relativity

$$\begin{aligned} x' &= \cosh\beta x - \sinh\beta t \\ t' &= -\sinh\beta x + \cosh\beta x \end{aligned} \tag{1.4}$$

then we find that (1.2) becomes

$$\frac{\partial^2\phi}{\partial t'^2} - c^2\frac{\partial^2\phi}{\partial x'^2} = 0 \tag{1.5}$$

so we see that the speed of the waves is again $\pm c$. Thus the behavior of our waves respects the postulate of special relativity.

1.1.4 Quantizing the field

1.1.5 A problem

We seem to have obtained a theory which respects the Lorentz invariance. But one step in the procedure was the taking of the limit (**??**) where the wavelengths being studied were taken to be much longer than the lattice spacing a. But what happens if we look at the theory at wavelengths of $\lambda \sim a$? The lattice of atoms did define a preferred rest frame – the frame in which all point masses n are at rest when the field is unexcited. If we go to a frame moving with respect to this rest frame, then the masses will appear to be in motion. Further, the spacing between the masses would be length contracted to a value smaller than a. So we still seem to have a violation of the principle of relativity, which asks that all uniformly moving frames behave identically.

The natural way out would be to take the limit $a \to 0$, so that we don't see the lattice spacing at all. And this is in fact what is done in quantum field theory. But it leaves us with a very serious problem, whose solution we still do not know.

We had seen that the theory of point masses joined by springs was a theory of a collection of harmonic oscillators. Even if we don't excite the field, each oscillator has a ground state energy $\frac{1}{2}\hbar\omega$. Adding over the frequencies(**??**), we find

$$E_{ground} = \tag{1.6}$$

1.2 The outline of quantum field theory

We can now state in qualitative terms how quantum field theory describes particles:

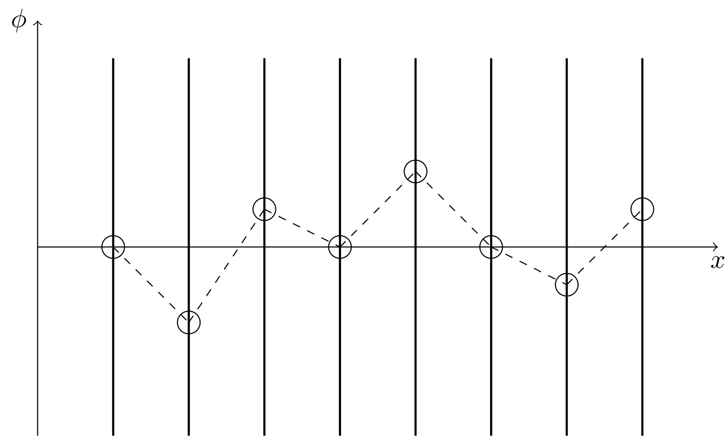

Figure 1.2: caption ...

(a) The degrees of freedom are encoded in a set of coupled harmonic oscillators. There is an oscillator variable at each point of space. Each such oscillator is coupled to its neighbors; thus the overall theory is described by a local Lagrangian.

(b) We can change variables so that a set of coupled oscillators looks like a set of decoupled oscillators. If each such decoupled oscillator is placed in its vacuum state, then we get the lowest energy state of our entire system. This state will be identified with the vacuum.

(c) Now let us consider the excitations of these coupled oscillators. Large excitations can be treated to a good approximation by classical mechanics, and we get the field equations (**??**). These equations are analogous to the equations of classical electrodynamics where the scalar field ϕ is replaced by a vector field A_μ.

(d) For small excitations, on the other hand, we should consider the quantum theory of our oscillators. The excitations of a quantum oscillator are discrete, and can be labelled by an integer $n = 0, 1, 2, \dots$. If we have a single excitations $n = 1$, then we say that we have one particle. If $n = 2$, then we have two particles, and so on. Thus the theory automatically describes multiparticle states. We also have the possibility of creating and annihilating particles, since we can increase or decrease the number of excitations of a harmonic oscillator by adding suitable interactions in the theory.

Note that if we excite the oscillators in a given region of space, then the coupling of these oscillators to other oscillators causes the excitation to spread over the space x. This spread gives the normal dispersion of a wavefunction over space, encoded in the Schrodinger equations describing a single particle. The eigenstates of the Hamiltonian are wavefunctions spread over all of space, and correspond to particles with definite momentum rather than definite position.

We will now carry write out this quantum theory of coupled oscillators in more detail. But we can already see how general relativity will impact the above picture. In general relativity, mass curves spacetime. When matter moves around, for example when making a black hole, the the spacelike slices of the geometry deform. Thus the distances between the points of space change, which in turn changes the couplings between the oscillators at these points.

There is one effect of this change which will be important for us. Suppose we start with a state which is the vacuum on the initial slice of spacetime. After matter moves and the spacelike slice deforms, the state we have will not in general be the vacuum state of the coupled oscillators that we get after the deformation. We can write this new state as the vacuum plus some particle excitations, since any state of the theory is given by some choice of excitations levels for the oscillators. This is the phenomenon of 'particle creation in curved space'. In particular, this is the phenomenon which gives Hawking's pair creation of particles near a black hole horizon, and leads to the information puzzle.

1.3 Diagonalizing the coupled oscillators

The wave equation (1.2) has solutions that moves at the speed of light, so it describes massless particles. For greater generality, we write the equation including a mass term

$$\frac{\partial^2\phi}{\partial t^2} - \frac{\partial^2\phi}{\partial x^2} - m^2\phi = 0 \tag{1.7}$$

where we have now chosen units where $c = 1$. This equation follows from an action

$$S = \int dt\ L \tag{1.8}$$

The Lagrangian is

$$\hat{\mathcal{L}} = \int dx\ \mathcal{L} \tag{1.9}$$

where

$$\mathcal{L} = \frac{1}{2}(\partial_t\phi)^2 - \frac{1}{2}(\partial_x\phi)^2 - \frac{1}{2}m^2\phi^2 \tag{1.10}$$

is the Lagrangian *density*.

We expand ϕ in Fourier modes

$$\phi(t,x) = \sum_{n=-\infty}^{\infty} \phi_n(t)e^{ik_n x} \tag{1.11}$$

It is convenient to think of the space x as a periodic box of length L. This gives a discrete set of values for k_n

$$k_n = \frac{2\pi n}{L} \tag{1.12}$$

Since ϕ is real we have

$$\phi_n = \phi^*_{-n} \tag{1.13}$$

Writing

$$\phi_n = \phi_n^R + i\phi_n^I \tag{1.14}$$

we get

$$\phi_n^R = \phi_{-n}^R, \quad \phi_n^I = -\phi_{-n}^I \tag{1.15}$$

Let us now compute the Lagrangian $\mathcal{L}$. A little algebra gives

$$L = \sum_{n=1}^{\infty}[L(\dot{\phi}_n^R)^2 - L\omega_n^2(\phi_n^R)^2] + \sum_{n=1}^{\infty}[L(\dot{\phi}_n^I)^2 - L\omega_n^2(\phi_n^I)^2] + \frac{L}{2}[(\dot{\phi}_0)^2 - \omega_0^2(\phi_0)^2] \tag{1.16}$$

where

$$\omega_n^2 = k_n^2 + m^2 \tag{1.17}$$

We see that the free scalar field is equivalent to a collection of harmonic oscillators. We can thus quantize the scalar field by just quantizing these harmonic oscillators.

1.4 The harmonic oscillator

Let us first recall the quantization of a single harmonic oscillator. Let the position variable be q. The Lagrangian is

$$\mathcal{L} = \frac{1}{2}m\dot{q}^2 - \frac{1}{2}kq^2 \tag{1.18}$$

This gives

$$p = m\dot{q}, \quad H = \frac{p^2}{2m} + \frac{k}{2}q^2 \tag{1.19}$$

We write

$$\omega = \sqrt{\frac{k}{m}} \tag{1.20}$$

Upon quantization the Hamiltonian becomes

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{k}{2}\hat{q}^2 \tag{1.21}$$

We can expand this as

$$\hat{H} = (\sqrt{\frac{k}{2}}\hat{q} - i\frac{\hat{p}}{\sqrt{2m}})(\sqrt{\frac{k}{2}}\hat{q} + i\frac{\hat{p}}{\sqrt{2m}}) + \frac{1}{2}\omega \tag{1.22}$$

Define

$$\hat{a}^\dagger = \frac{1}{\sqrt{\omega}}(\sqrt{\frac{k}{2}}\hat{q} - i\frac{\hat{p}}{\sqrt{2m}}), \quad \hat{a} = \frac{1}{\sqrt{\omega}}(\sqrt{\frac{k}{2}}\hat{q} + i\frac{\hat{p}}{\sqrt{2m}}) \tag{1.23}$$

This gives

$$[\hat{a}, \hat{a}^\dagger] = 1 \tag{1.24}$$

and

$$\hat{H} = \left(\hat{a}^\dagger \hat{a} + \frac{1}{2})\right) \tag{1.25}$$

Note that the original position variable can be written as

$$\hat{q} = \sqrt{\frac{\omega}{2k}}(\hat{a} + \hat{a}^\dagger) \tag{1.26}$$

1.5 Quantization of the scalar field

Let us now apply this quantization of a harmonic oscillator to the full set of oscillators describing the scalar field $\hat{\phi}$.

Each Fourier component $\phi_n^R, \phi_n^I, (n > 0)$ is the 'position' coordinate of a harmonic oscillator. This oscillator has

$$m = 2L, \quad k = 2L\omega_n^2, \quad \omega = \sqrt{\frac{k}{m}} = \omega_n \tag{1.27}$$

Thus we have

$$\hat{\phi}_n^R = \frac{1}{2\sqrt{L\omega_n}}(\hat{a}_n^R + (\hat{a}_n^R)^\dagger) \tag{1.28}$$

$$\hat{\phi}_n^I = \frac{1}{2\sqrt{L\omega_n}}(\hat{a}_n^I + (\hat{a}_n^I)^\dagger) \tag{1.29}$$

For the zero mode we have

$$m = L, \quad k = L\omega_0^2, \quad \omega = \omega_0 \tag{1.30}$$

$$\hat{\phi}_0 = \sqrt{\frac{1}{2L\omega_0}}(\hat{a}_0 + \hat{a}_0^\dagger) \tag{1.31}$$

Using (1.15) we can write the classical field as

$$\begin{aligned} \phi(t) &= \sum_n (\phi_n^R(t) + i\phi_n^I(t))e^{ik_n x} \\ &= \sum_{n>0} 2\phi_n^R(t)\cos(k_n x) - \sum_{n>0} 2\phi_n^I(t)\sin(k_n x) + \phi_0(t) \end{aligned} \tag{1.32}$$

Upon quantizing we should just convert the 'position variables' $q(t)$ of the harmonic oscillators to 'position operators' $\hat{q}$

$$\hat{\phi} = \sum_{n>0} 2\tilde{\phi}_n^R \cos(k_n x) - \sum_{n>0} 2\hat{\phi}_n^I \sin(k_n x) + \hat{\phi}_0 \tag{1.33}$$

In terms of creation and annihilation operators we get

$$\hat{\phi}(t) = \sum_{n>0} \frac{1}{\sqrt{L\omega_n}}(\hat{a}_n^R + (\hat{a}_n^R)^\dagger)\cos(k_n x) - \sum_{n>0} \frac{1}{\sqrt{L\omega_n}}(\hat{a}_n^I + (\tilde{a}_n^I)^\dagger)\sin(k_n x) + \sqrt{\frac{1}{2L\omega_0}}(\hat{a}_0 + (\hat{a}_0)^\dagger) \tag{1.34}$$

We thus see that the operator $(\hat{a}_n^R)^\dagger$ is a creation operator associated to the wavefunction $\cos(k_n x)$ and $(\hat{a}_n^I)^\dagger$ is associated to $\sin(k_n x)$. We would like to create and annihilate particles of definite momentum $e^{ik_n x}$. For $n > 0$ define

$$\hat{a}_n \equiv \frac{1}{\sqrt{2}}[\hat{a}_n^R + i\hat{a}_n^I], \quad n > 0 \tag{1.35}$$

For $n < 0$ we can similarly define

$$\hat{a}_n \equiv \frac{1}{\sqrt{2}}[\hat{a}_n^R + i\hat{a}_n^I], \quad n < 0 \tag{1.36}$$

where the operators $\hat{a}_n^R, \hat{a}_n^I$ for negative n are given through (1.15). We then find

$$\hat{a}_n^\dagger = \frac{1}{\sqrt{2}}[(\hat{a}_n^R)^\dagger - i(\hat{a}_n^I)^\dagger], \quad n \neq 0 \tag{1.37}$$

We find the commutation relations

$$[\hat{a}_m, \hat{a}_n^\dagger] = \delta_{m,n} \tag{1.38}$$

The zero mode operators stays as before. Writing

$$\cos(k_n x) = \frac{1}{2}[e^{ik_n x} + e^{-ik_n x}], \quad \sin(k_n x) = \frac{1}{2i}[e^{ik_n x} - e^{-ik_n x}] \tag{1.39}$$

we get

$$\hat{\phi} = \sum_{n=-\infty}^{\infty} \frac{1}{\sqrt{L}} \frac{1}{\sqrt{2\omega_n}}[\hat{a}_n e^{ik_n x} + \hat{a}_n \partial e^{-ik_n x}] \tag{1.40}$$

This form is not very symmetrical between x, t since the operators depend on x but not on t. We move to the Heisenberg picture where we have

$$\hat{O}(t) = e^{i\hat{H}t}\tilde{O}e^{-i\hat{H}t} \tag{1.41}$$

For a harmonic oscillator

$$\hat{H} = \omega\tilde{a}^\dagger\hat{a} \tag{1.42}$$

Note that

$$[\hat{a}, \hat{H}] = \omega\hat{a}, \quad [\hat{a}^\dagger, \hat{H}] = -\omega\hat{a}^\dagger \tag{1.43}$$

We can thus move $\hat{a}$ across $e^{-i\tilde{H}t}$ in

$$\hat{a}(t) = e^{i\tilde{H}t}\hat{a}e^{-i\tilde{H}t} \tag{1.44}$$

getting

$$\hat{a}(t) = \hat{a}e^{-i\omega t} \tag{1.45}$$

$$\hat{a}^\dagger(t) = \hat{a}^\dagger e^{i\omega t} \tag{1.46}$$

Then in the Heisenberg picture the operator $\hat{\phi}$ becomes

$$\hat{\phi}(x,t) = \sum_{n=-\infty}^{\infty} \frac{1}{\sqrt{V}} \frac{1}{\sqrt{2\omega_n}} [\hat{a}_n e^{ik_n x - i\omega_n t} + \hat{a}_n^\dagger e^{-ik_n x + i\omega_n t}] \tag{1.47}$$

We now see a covariant dot product between the position vector (t, x) and the momentum (ω_n, k_n).

1.6 Quantum fields in curved space

We can easily extend our above discussion to 3+1 dimensional spacetime. The field operator becomes

$$\hat{\phi} = \sum_{\vec{k}} \frac{1}{\sqrt{V}} \frac{1}{\sqrt{2\omega}} \left(\hat{a}_{\vec{k}} e^{i\vec{k}\cdot\vec{x} - i\omega t} + \hat{a}_{\vec{k}}^\dagger e^{-i\vec{k}\cdot\vec{x} + i\omega t} \right) \tag{1.48}$$

where V is the volume of the spatial box where we have taken the field to live, and $\omega = \sqrt{|\vec{k}|^2 + m^2}$ for a field with mass m. The vacuum is the state annihilated by all the $\hat{a}$

$$\hat{a}_{\vec{k}}|0\rangle = 0 \tag{1.49}$$

and the $\hat{a}_{\vec{k}}^\dagger$ create particles.

In *curved* spacetime, on the other hand, there is no canonical definition of particles. We can choose any coordinate t for time, and decompose the field into positive and negative frequency modes with respect to this time t. Let the positive frequency modes be called $f(x)$; then their complex conjugates give negative frequency modes $f^*(x)$. The field operator can be expanded as

$$\hat{\phi}(x) = \sum_n \left(\hat{a}_n f_n(x) + \hat{a}_n^\dagger f_n^*(x) \right) \tag{1.50}$$

Then we can define a vacuum state as one that is annihilated by all the annihilation operators

$$\hat{a}_n|0\rangle_a = 0 \tag{1.51}$$

The creation operators generate particles; for example a 1-particle state would be

$$|\psi\rangle = \hat{a}_n^\dagger|0\rangle_a \tag{1.52}$$

We have added the subscript a to the vacuum state to indicate that the vacuum is defined with respect to the operators $\hat{a}_n$. But since there is no unique choice of the time coordinate t, we can choose a different one $\tilde{t}$. We will then have a different set of positive and negative frequency modes, and an expansion

$$\hat{\phi}(x) = \sum_n \left(\hat{b}_n h_n(x) + \hat{b}_n^\dagger h_n^*(x) \right) \tag{1.53}$$

Now the vacuum would be defined as

$$\hat{b}_n |0\rangle_b = 0 \tag{1.54}$$

and the $\hat{b}_n^\dagger$ would create particles.

Given this situation, which is the 'correct' definition of particles? The situation can be summarized as follows:

(i) If spacetime is flat, then we should use the Minkowski time t to define modes as in (**??**), and this defines our particles. We can of course go to a frame that is uniformly moving with respect to our initial frame, and then we have a new Minkowski time t'. But it turns out that this change does not affect the definition of the vacuum. The different operators $\hat{a}_{\vec{k}}$ get relabelled because the momentum changes $\vec{k} \to \vec{k}'$, but annihilations operators remain annihilation operators and so the definition of the vacuum (1.51) is unaffected.

(ii) In general spacetime is not flat. Suppose that at a location x the curvature length scale is l. Then over distances much less than l the space can be treated as essentially flat. Thus for wavelengths $\lambda \sim |\vec{k}|^{-1} \ll l$, local particles defined the same way as in (a). But for $\lambda \gtrsim l$, we have no unique definition of particle.

(iii) Even though the definition of particle is ambiguous in general, there is a well defined meaning to $\langle T_{\mu\nu}(x)\rangle$, the expectation value of the stress-energy tensor at a point in spacetime. This quantity is independent of the definition we choose for particles, and in fact can be computed by doing a path integral over the field ϕ without choosing any definition of particle.

The underlying assumption in defining the operator $T_{\mu\nu}(x)$ is that its expectation value should vanish in the vacuum state of flat spacetime. We can then carry the same definition of $T_{\mu\nu}(x)$ to curved space, where we find a nonzero value in general. The assumption $\langle T_{\mu\nu}(x)\rangle = 0$ in the flat space vacuum appears to be a natural one. But hidden in this assumption is the cosmological constant problem: why must we subtract the energy of vacuum fluctuations in exactly such a way that this expectation value vanishes? There is as yet no clear answer to this question.

1.6.1 A conserved inner product

In nonrelativistic quantum mechanics, the inner product

$$(\chi, \psi) = \int d^3\, x \chi^*(x) \psi(x) \tag{1.55}$$

is conserved for two solutions χ, ψ of the Schrodinger equation. This inner product is positive definite, since (ψ, ψ) is the integral of the probability density $|\psi|^2$. For the correctly normalized wavefunction, (ψ, ψ) equals unity at all times, since the particle has a probability 1 to be somewhere in the whole space at any time t.

In quantum field theory, particles can be created and destroyed, since they are just excitations of an underlying field. Nevertheless, there is a conserved inner product for for the functions f, f^* appearing in the expansion of $\hat{\phi}$. This inner product is however, not positive definite; it is positive on functions like the f that multiply the annihilation operator $\hat{a}$, and is negative in functions f^* which multiply that $\hat{a}^\dagger$. This inner product will be very useful to us in relating different expansions of the field $\hat{\phi}$.

1.6.2 Obtaining the inner product

Consider two functions f_1, f_2 which satisfy the waveequation

$$\Box f = \frac{1}{\sqrt{-g}} \partial_\mu \left(g^{\mu\nu} \sqrt{-g} f_{,\nu} \right) = 0 \tag{1.56}$$

Thus $\Box f_1^* = 0$ and $\Box f_2 = 0$. Consider the product $f_1^* \Box f_2 = 0$, and integrate over the spacetime slab between two hypersurfaces $\mathcal{S}_{lower}$ and $\mathcal{S}_{upper}$

$$\begin{aligned} 0 &= \int d^4x \sqrt{-g} f_1^* \Box f_2 \\ &= \int d^4x \sqrt{-g} f_1^* \frac{1}{\sqrt{-g}} \partial_\mu \left(g^{\mu\nu} \sqrt{-g} f_{2,\nu} \right) \\ &= \int d^4x f_1^* \partial_\mu \left(g^{\mu\nu} \sqrt{-g} f_{2,\nu} \right) \end{aligned} \tag{1.57}$$

We can now integrate The ∂_μ derivative by parts. We assume that the functions die off at large spatial distances, so the only boundary terms are $\tilde{B}_{upper}$ and $\tilde{B}_{lower}$ from the hypersurfaces $\mathcal{S}_{upper}$ and $\mathcal{S}_{lower}$

$$0 = -\int d^4x f^*_{1,\mu} g^{\mu\nu} \sqrt{-g} f_{2,\nu} + \tilde{B}_{upper} - \tilde{B}_{lower} \tag{1.58}$$

The boundary terms can be written in a covariant form

$$B = \int d^3\xi d\Sigma_\mu f_1^* g^{\mu\nu} \sqrt{-g} f_{2,\nu} \tag{1.59}$$

where $\xi^i, i = 1, 2, 3$ are coordinates in the 3-dimensional spatial hypersurface. Alternatively, we can take the intrinsic metric h_{ij} of the hypersurface, and write

$$B = \int d^3\xi \sqrt{h} f_1^* \partial_n f_2 \tag{1.60}$$

where ∂_n is the derivative in the direction along the unit normal to the hypersurface.

The bulk integral in (1.58) is symmetric under the interchange $f_1^* \leftrightarrow f_2$. Thus we can cancel it by taking

$$\begin{aligned} 0 &= \int d^4x\sqrt{-g} f_1^* \Box f_2 - \int d^4x \sqrt{-g} f_2 \Box f_1^* \\ &= B_{upper} - B_{lower} \end{aligned} \tag{1.61}$$

where

$$B = \int d^3\xi \, (f_1^* \partial_n f_2 - f_2 \partial_n f_1^*) \tag{1.62}$$

Since The difference of B on the upper and lower hypersurfaces vanishes, we find that B is conserved along the evolution for any two solutions f_1, f_2 of the waveequation. If we choose $f_1 = f_2$, then we get a purely imaginary quantity. It is conventional to add a factor i to make this real; thus we define the inner product

$$(f_1, f_2) = i \int d^3\xi \sqrt{h} \, (f_1^* \partial_n f_2 - f_2 \partial_n f_1^*) \tag{1.63}$$

1.6.3 The inner product on a null hypersurface

The conserved inner product (1.63) was defined on a spacelike hypersurface. We will however sometimes find it convenient to consider a *null hypersurface*. For example we will encounter a metric of the form

$$ds^2 = -F(r) du dv + r^2 d\Omega^2 \tag{1.64}$$

where $r = r(u, v)$. The hypersurface defined by $u = constant$ is null. It is spanned by angular directions θ, ϕ which are spacelike, but also by the direction v which is null. Along such a hypersurface the volume element $d^3\xi\sqrt{h}$ vanishes. Also, the normal to the hypersurface is along the direction v which is null, so we cannot compute the derivative along a unit normal. These two difficulties cancel each other out, however, and we can take the limit where our spacelike hypersurface becomes null. It is easier however to derive the conserved inner product on null surfaces by starting again from the waveequation, which is what we do now.

For the metric (1.64) the waveequation is

$$\Box f = \frac{1}{\sqrt{-g}} \left(\partial_u \left(g^{uv} \sqrt{-g} f_{,v} \right) + \partial_v \left(g^{uv} \sqrt{-g} f_{,u} \right) + \partial_a \left(g^{ab} \sqrt{-g} f_{,b} \right) \right) = 0 \tag{1.65}$$

where a, b run over the angular directions. Let us integrate $f_1^* \Box f_2$ over the slab between two null hypersurfaces $u = u_{lower}$ and $u = u_{upper}$. We assume that the f_i fall off at large v, so the only boundary terms in an integration by parts are from these two hypersurfaces. We have

$$\begin{aligned} 0 &= \int d^4x\sqrt{-g}\,(f_1^*\Box f_2 - f_2\Box f_1^*) \\ &= \int dudvd\theta d\phi\, f_1^*[\partial_u\left(g^{uv}\sqrt{-g}f_{2,v}\right) + \partial_v\left(g^{vu}\sqrt{-g}f_{2,u}\right) + \partial_a\left(g^{ab}\sqrt{-g}f_{2,b}\right)] \\ &- \int dudvd\theta d\phi\, f_2[\partial_u\left(g^{uv}\sqrt{-g}f^*_{1,v}\right) + \partial_v\left(g^{vu}\sqrt{-g}f^*_{1,u}\right) + \partial_a\left(g^{ab}\sqrt{-g}f^*_{1,b}\right)] \end{aligned} \tag{1.66}$$

The only boundary terms come from integration by parts on ∂_u, and so we get

$$0 = B_{upper} - B_{lower} \tag{1.67}$$

where for any surface

$$B = \int dvd\theta d\phi\,\left(f_1^* g^{uv}\sqrt{-g}f_{2,v} - f_2 g^{uv}\sqrt{-g}f^*_{1,v}\right) \tag{1.68}$$

For the metric (1.64), we have

$$g^{uv} = -\frac{2}{F}, \quad \sqrt{-g} = \frac{1}{2}Fr^2\sin\theta \tag{1.69}$$

Putting in the factor i that we added in (1.63), we find the inner product computed on a $u = constant$ surface

$$(f_1, f_2) = i\int dvd\Omega\, r^2\,(f_1^*\partial_v f_2 - f_2\partial_v f_1^*) \tag{1.70}$$

TOPIC III

HAWKING'S PUZZLE: A SECOND PASS

Now that we have some understanding of general relativity and quantum field theory, let us return to the Hawking puzzle. We will see that the puzzle arises in the following steps:

(a) The horizon is a place where particles trajectories get 'separated'. Particles trying to fly out from a point just outside the horizon do manage to escape to infinity. But particles trying to fly out from a point just inside the horizon cannot escape; they end up falling to the center of the black hole. Thus trajectories that start off very close to each other get 'pulled apart' if they are on opposite sides of the horizon.

(b) This pulling apart of trajectories converts a vacuum state to a state which contains particle pairs; one member of the pair being outside the horizon and one inside.

(c) The important point is that these particles are in an 'entangled state'. It is this entanglement that creates a problem near the endpoint of black hole evaporation, as we will see.

Thus the Hawking puzzle arises from very simple and robust features of general relativity and quantum field theory. General relativity shows us that the horizon is a place where trajectories separate. The nature of the quantum vacuum then implies that this vacuum will be unstable; pairs of particles will be created, one inside the hole and one outside. The puzzle arises from the nature of the quantum state of these particles: the two particles are *entangled*, instead of having separate wavefuntions of their own. If the hole evaporats away, then the puzzle becomes one about the state of the particles left outside as Hawing radiation: what are these particles entangled with?

Later we will make a third pass at the puzzle, using some results from quantum information theory to show that the steps above are robust against any small modifications of the physics. This will remove a large category of proposed resolutions of the puzzle, that seek a way out by invoking sutble quantum gravity correctons to the leading semiclassical physics used in steps (i)-(iii) above.

Lecture notes 2

The essential physics behind pair creation

Let us begin by describing the basic physics that leads to pair creation at the horizon of a black hole. We will first see how geodesics on the two sides of the horizon diverge away from each other. We will then make a toy model for the quantum vacuum, and see how this divergence of trajectories leads to the creation of entangled particle pairs.

2.1 The divergence of trajectories at the horizon

Recall the Schwarzschild metric (**??**)

$$ds^2 = -(1 - \frac{2M}{r})dt^2 + (1 - \frac{2M}{r})^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{2.1}$$

We wish to consider particles that are trying to escape from the hole. The particles that can escape most easily are massless particles, moving outwards radially at the speed of light. For such trajectories the only nonzero displacements are dt, dr and we must have $ds^2 = 0$. There are three cases to consider:

(i) Suppose the particle starts a little outside the horizon, at $r = 2M + \epsilon$. Then we have from $ds^2 = 0$

$$\frac{dr}{dt} = \pm\left(1 - \frac{2M}{r}\right) \tag{2.2}$$

To have the particle go outwards, we take the positive sign. Let us ask for the time it takes for this particle to escape to a location r_f that is away from the horizon

$$\int_{2M+\epsilon}^{r_f} \frac{dr}{1 - \frac{2M}{r}} = \int_0^T dt \tag{2.3}$$

This gives

$$T \approx 2M \log\frac{1}{\epsilon} \tag{2.4}$$

Thus the particle does ultimately escape, but the time to escape becomes large as ϵ goes to zero.

(ii) Suppose the particle starts *at* the horizon $r = 2M$. Then the only way to get $ds^2 = 0$ is to take $dr = 0$, which means the particle stays at the horizon.

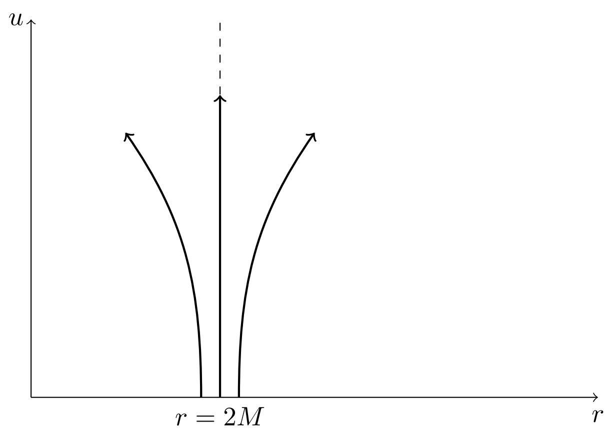

Figure 2.1: caption ...

(c) Suppose the particle starts a little *inside* the horizon, at a position $r = 2M - \epsilon$ We have

$$\frac{dr}{dt} = \pm\left(1 - \frac{2M}{r}\right) = \mp\left(\frac{2M}{r} - 1\right) \tag{2.5}$$

Let us compute the time for the particle to reach a position r_f that is away from the horizon $0 < r_f < 2M$. This needs the negative sign in the above relation, and we find

$$\int_{2M+\epsilon}^{r_f} \frac{dr}{\frac{2M}{r} - 1} = \int_0^T dt \tag{2.6}$$

This gives

$$T \approx 2M \log \frac{1}{\epsilon} \tag{2.7}$$

Thus the particle escapes the vicinity of the horizon, but again the time to escape becomes large as ϵ goes to zero.

The situation is schematically shown in fig.**??**. We see that a small region straddling the horizon gets stretched to a *large region* after we wait for a sufficiently long time. In fact we can start with an arbitrarily small region

$$|r - 2M| < \epsilon \tag{2.8}$$

and see that after a time

$$t \sim \tag{2.9}$$

the region will stretch to a size $\sim M$ which describes the length scale of the Schwarzschild geometry.

This persistent stretching at the horizon is what will lead to the evaporation of the hole. While spatial slices stretch and comtract in any process in general relativity, such deformations are usually quite limited. that is, if the length scale in the metric is $\sim L$, then regions of size $\sim L$ may stretch to a size αL with α being a factor of order unity. Further, the deformations of the spatial slices will typically stop after a while when the metric settles down to a new startionary state. But in the presence of a horizon, we see that the stretching can be arbitrarily large, and does not stop as long as the horizon exists.

Our next question is: what is the consequence of this stretching?

We have seen in section **??** that the quantum fields on spacetime can be described by a set of coupled harmonic oscillators. When a slice stretches, the distance between neighboring points increases. This makes the coupling between the corresponding oscillators weaker. This change of coupling can convert a vacuum state of the oscillators to a state that contains pairs of excitations. But excitations of the oscillators describing the quantum field correspond to particles. Thus we will find that the stretching of slices seen above will lead to the creation of particle pairs from the vacuum.

We will study this phenomenon of pair creation in three steps:

(i) We will first consider a single harmonic oscillator. We start in the ground state of this oscillator. At some time $t = 0$, we change the frequency of this oscillator. We then see how the state of the oscillator contains pairs of excitations above the vacuum.

(ii) We then make a toy model of the situation we will encounter with the black hole. We consider two oscillators, one on each side of the horizon. These oscillators will be coupled to each other, the way neighboring oscillators are coupled in quantum field theory, and we will let the initial state of the system be the ground state of the coupled oscillator pair. We have seen above that geodesics on the two sides of the horizon separate away from each other. We will model this effect by removing the coupling between the oscillators at some time $t = 0$. We will find that the two oscillators will now have pairs of excitations, and the overall states will be *entangled* between the two oscillators. This state has all the features of the full quantum problem that will be relevant to the information paradox, so this is a useful toy model.

(c) Finally we will set up the problem of Hawking radiation in the black hole metric, using the 1+1 dimensional case for simplicity. We will present details of the computations in the appendix.

2.2 Pair creation for a single harmonic oscillator

We will solve this simple case in two ways: first, in the Schrodinger picture, where the physics is a little clearer, and second, in the Heisenberg picture, which makes the computations simpler in the general case that we will find in field theory.

2.2.1 Sudden frequency change: Schrodinger picture

Consider the Lagrangian of a harmonic oscillator with frequency ω.

$$\mathcal{L} = \frac{1}{2}\dot{\phi}^2 - \frac{1}{2}\omega^2\phi^2 \tag{2.10}$$

where we have called the position variable ϕ, in line with the fact this this variable will be analogous to the value of the scalar field in the full problem. We can write ϕ and its conjugate momentum π in terms of creation and annihilation operators

$$\begin{aligned} \hat{\phi} &= \frac{1}{\sqrt{2\omega}}(\hat{a} + \hat{a}^\dagger) \\ \hat{\pi} &= = -i\sqrt{\frac{\omega}{2}}(\hat{a} - \hat{a}^\dagger) \end{aligned} \tag{2.11}$$

The vacuum state $|0\rangle_a$ of the oscillator is defined by

$$\hat{a}|0\rangle_a = 0 \tag{2.12}$$

We assume that for times $t < 0$ the oscillator is in its vacuum state $|0\rangle_a$. At time $t = 0$ the frequency of the oscillator is suddenly changed from ω to a different value $\tilde{\omega}$. The state of the oscillator, however, cannot change immediately. This is because the state evolves according to the Schrodinger equation

$$i\frac{\partial\psi}{\partial t} = \hat{H}\psi \tag{2.13}$$

so that for infinitesimal ϵ we have

$$\psi|_{t=\epsilon} - \psi|_{t-\epsilon} = -i(2\epsilon)\hat{H}\psi \tag{2.14}$$

For our $\hat{H}$ and $\psi = |0\rangle_a$, the state $\hat{H}\psi$ has finite norm. Thus in the limit $\epsilon \to 0$ we find that ψ does not suddenly change at $t = 0$ when we change the frequency of the oscillator.

The operators $\hat{\phi}, \hat{\pi} = -i\partial/\partial\phi$ do not change when we change the Hamiltonian. But their expressions in terms of the new creation and annihilation operators does change. Let the creation and annihilation operators for the oscillator with frequency $\tilde{\omega}$ be $\hat{b}, \hat{b}^\dagger$. Then we have

$$\begin{aligned} \phi &= \frac{1}{\sqrt{2\tilde{\omega}}}(\hat{b} + \hat{b}^\dagger) \\ \pi &= -i\sqrt{\frac{\tilde{\omega}}{2}}(\hat{b} - \hat{b}^\dagger) \end{aligned} \tag{2.15}$$

Comparing with (2.11), we find after a little algebra

$$\hat{a} = \alpha\hat{b} + \beta\hat{b}^\dagger \tag{2.16}$$

with

$$\alpha = \frac{1}{2}\left(\sqrt{\frac{\omega}{\tilde{\omega}}} + \sqrt{\frac{\tilde{\omega}}{\omega}}\right) \tag{2.17}$$

$$\beta = \frac{1}{2}\left(\sqrt{\frac{\omega}{\tilde{\omega}}} - \sqrt{\frac{\tilde{\omega}}{\omega}}\right) \tag{2.18}$$

Our chosen state at $t = 0$ was $|0\rangle_a$; this state was defined by the relation (2.12). We can write this relation as

$$\left(\alpha\hat{b} + \beta\hat{b}^\dagger\right)|0\rangle_a = 0 \tag{2.19}$$

If however we take our creation and annihilation operators as $\hat{b}, \hat{b}^\dagger$, the we would define a vacuum state $|0\rangle_b$ through the relation

$$\hat{b}^\dagger|0\rangle_b = 0 \tag{2.20}$$

We see that our initial state $|0\rangle_a$ is not the same state $|0\rangle_b$ which would be called the vacuum for the oscilaltor at times $t > 0$. We noted above that the state itself does not suddenly change at $t = 0$ when we change the frequency of the oscillator from ω to $\tilde{\omega}$. Thus when we cross the time $t = 0$ we should write our state $|0\rangle_a$ as a linear combination of basis states for the oscillator with frequency $\tilde{\omega}$; the subsequent evolution of the state will be through the phase e^{-iE_nt} for basis states with different energies E_n.

The states $|n\rangle_b$ with $n = 0, 1, 2, \ldots$ form a complete basis of the Hilbert space, where $|n\rangle_b$ is the nth excited level of the oscillator with frequency $\tilde{\omega}$. Thus we can write $|0\rangle_a$ as a linear sum over this basis. It turns out the the linear sum is given by an elegant ansatz of the form

$$|0\rangle_a = Ce^{\frac{1}{2}\gamma b^\dagger b^\dagger}|0\rangle_b \tag{2.21}$$

where C is a normalization constant. To find the coefficient γ appearing in this expansion, we note the relation We have

$$be^{\frac{1}{2}\gamma b^\dagger b^\dagger}|0\rangle_b = \gamma b^\dagger e^{\frac{1}{2}\gamma b^\dagger b^\dagger}|0\rangle_b \tag{2.22}$$

This relation can be checked by expanding the exponentials in series, and using the commutation relation $[\hat{b}, \hat{b}^\dagger] = 1$. Then we find

$$\gamma = -\frac{\beta}{\alpha} = -\frac{\tilde{\omega} - \omega}{\tilde{\omega} + \omega} \tag{2.23}$$

$$C = \tag{2.24}$$

2.2.2 Analyzing the relation (2.21)

2.2.3 The Heisenberg picture for a single oscillator

Let us now redo the problem of pair creation for a single oscillator in the Heisenberg picture; this is the picture that will is more convenient to use for quantum fields on curved spacetime. The Heisenberg operator $\hat{\phi}(t)$ is defined in terms of the operator $\hat{\phi}$ as

$$\hat{\phi}(t) = e^{i\hat{H}t}\hat{\phi}e^{-i\hat{H}t} \tag{2.25}$$

We can expand $\hat{\phi}(t)$ in terms of creation and annihilation operators

$$\hat{\phi}(t) = f(t)\hat{a} + f^*(t)\hat{a}^\dagger \tag{2.26}$$

We have put the time-dependence in the functions f, f^*, while the operators $\hat{a}, \hat{a}^\dagger$ are time indpendent. The state $|\psi\rangle$ does not evolve in the Heisenberg picture, and the action of $\hat{a}, \hat{a}^\dagger$ on a state will give rise to other states which will also not evolve with t.

The operator $\hat{\phi}(t)$ should satisfy the field equation (2.10) for the harmonic oscillator; this implies that f, f^* should satisfy this equation as well. For an oscillator with fixed frequency ω, a basis of solutions is

$$e^{-i\omega t}, \quad e^{i\omega t} \tag{2.27}$$

To find the linear combinations that must appear in (2.26), we compute a 2-point correlator: we start with the vacuum $|0\rangle_a$, created an excitation by applying $\hat{\phi}(t_1)$, and then annihilate this excitation to return to the vacuum by applying $\hat{\phi}(t_2)$:

$$\begin{aligned} {}_a\langle 0|\hat{\phi}(t_2)\hat{\phi}(t_1)|0\rangle_a &= {}_a\langle 0|\left(f(t_2)\hat{a} + f^*(t)2)\hat{a}^\dagger\right)\left(f(t_1)\hat{a} + f^*(t_1)\hat{a}^\dagger\right)|0\rangle_a \\ &= f(t_2)f^*(t_1) \end{aligned} \tag{2.28}$$

We require that the state that propagates begtween t_1 and t_2 have positive energy; i.e., it should evolve as e^{-iEt} with $E \geq 0$. We will call such functions positive frequency. We find that we must have

$$\hat{\phi}(t) = Ce^{-i\omega t}\hat{a} + C^* e^{i\omega t}\hat{a}^\dagger \tag{2.29}$$

which gives

$$f(t_2)f^*(t_1) = |C|^2 e^{-i\omega(t_2 - t_1)} \tag{2.30}$$

To fix C, we note the commutation relation between $\hat{\phi}(t)$ and $\hat{\pi} = d/dt\hat{\phi}(t)$:

$$[\hat{\phi}(t), \hat{\pi}(t)] = i \tag{2.31}$$

This gives

$$\hat{\phi}(t) = \frac{1}{\sqrt{2\omega}}e^{-i\omega t}\hat{a} + \frac{1}{\sqrt{2\omega}}e^{i\omega t}\hat{a}^\dagger \tag{2.32}$$

The value of C can be alternatively chacterized as a normalization of f. Define the inner product between two functions of t

$$(h_1, h_2) = i\left(h_1^* \dot{h}_2 - \dot{h}_1^* h_2\right) \tag{2.33}$$

If h_1, h_2 satisfy the equation (2.10) for the oscillator, then this inner product is independent of t, since it is proprtional to the Wronskian of two solutions. We find that

$$(f, f) = 1, \quad (f^*, f^*) = -1, \quad (f, f^*) = 0 \tag{2.34}$$

Thus the inner product is not positive definite: the positive frequency function f has positive norm while the negative frequency function f^* has negative norm.

2.2.4 Pair creation in the Heisenberg picture

Let us return to our problem where the frequency of then oscillator is ω for $t < 0$ and $\tilde{\omega}$ for $t > 0$. We again write $\hat{\phi}t)$ in the form (2.26). The equation for f is now

$$\begin{aligned} \ddot{f} + \omega^2 f &= 0, \quad t < 0 \\ \ddot{f} + \tilde{\omega}^2 f &= 0, \quad t > 0 \end{aligned} \tag{2.35}$$

For times $t < 0$ it is natural to think of states as being defined over the vacuum $|0\rangle_a$. Thus we use modes f that agree with (2.32) at $t < 0$

$$\begin{aligned} f &= \frac{1}{\sqrt{2\omega}} e^{-i\omega t}, \quad t < 0 \\ f &= a_1 e^{-i\tilde{\omega} t} + a_2 e^{i\tilde{\omega} t}, \quad t > 0 \end{aligned} \tag{2.36}$$

where the constants a_1, a_2 are determined by the requirement of continuity at $t = 0$

$$f(0^-) = f(0^+), \quad \dot{f}(0^-) = \dot{f}(0^+) \tag{2.37}$$

For times $t > 0$ it is natural to think of states as being defined over a vacuum $|0\rangle_b$, which is defined using the creation and annihilation operators $\hat{b}, \hat{b}^\dagger$ for an oscillator with frequency $\tilde{\omega}$. Calling the relevant function g instead of f, we have

$$\begin{aligned} g &= a_3 e^{-i\omega t} + a_4 e^{i\omega t}, \quad t < 0 \\ g &= \frac{1}{\sqrt{2\tilde{\omega}}} e^{-i\tilde{\omega} t}, \quad t > 0 \end{aligned} \tag{2.38}$$

where the constants a_3, a_4 are again determined by the requirement of continuity at $t = 0$

$$g(0^-) = g(0^+), \quad \dot{g}(0^-) = \dot{g}(0^+) \tag{2.39}$$

We now have two ways to write the field operator $\hat{\phi}(t)$:

$$\hat{\phi}(t) = f(t)\hat{a} + f^*(t)\hat{a}^\dagger \tag{2.40}$$

$$\hat{\phi}(t) = g(t)\hat{b} + g^*(t)\hat{b}^\dagger \tag{2.41}$$

Thus we get the relation

$$f(t)\hat{a} + f^*(t)\hat{a}^\dagger = g(t)\hat{b} + g^*(t)\hat{b}^\dagger \tag{2.42}$$

Our goal is to relate the operators $\hat{a}, \hat{a}^\dagger$ to the operators $\hat{b}, \hat{b}^\dagger$. Let us take the inner product (2.33) of both sides with f; i.e, we compute $(f, \cdot)$ on both sides. We get

$$\hat{a} = (f, g)\hat{b} + (f, g^*)\hat{b}^\dagger \tag{2.43}$$

Since f, f^*, g, g^* are all solutions of the same second order differential equation, the inner product is again independent of t. In general we would need to find the explicit form of f, g by finding the constants $a_1, \dots a_4$, but in the present case of 'sudden change of frequency' we have a shortcut: we can compute the inner products at $t = 0$. The inner product involves only the value and first derivative of the functions, and these quantities are the same at $t = 0^-$ and $t = 0^+$. Thus we can use the expression for f for $t < 0$ (where it is simpler) and the expression for g for $t > 0$ (where it is simpler). We find

$$\begin{aligned} (f, g) &= \frac{(\tilde{\omega} - \omega)}{\sqrt{2\omega}\sqrt{2\omega}} \equiv \alpha \\ (f, g^*) &= -\frac{(\tilde{\omega} + \omega)}{\sqrt{2\omega}\sqrt{2\omega}} \equiv \beta \end{aligned} \tag{2.44}$$

so that

$$\hat{a} = \alpha\hat{b} + \beta\hat{b}^\dagger \tag{2.45}$$

We see that we get the same relation as (**??**). The rest of the computation as before, leading to the relation (**??**).

2.3 Toy model of Hawking evaporation

Consider again the depiction of the black hole in fig.**??**. We have a horizon, and geodesics on different sides of this horizon diverge away from each other.

To see the effect of this divergence, consider a scalar field ϕ living on this spacetime. In fig.**??** we depict wavemodes f_L and f_R on the left and right sides of the horizon respectively. An occupation number $n_L = 0$ for f_L means that there is no particle in the mode f_L, an occupation number $n_L = 1$ means there is one particle in mode f_L, and so on. Similarly we can populate mode f_R with the possibilities $n_R = 0, 1, 2, \dots$.

Given our picture of quantum field theory, we can represent the mode f_L as a harmonic oscillator, and n_L is the excitation level of this oscillator. Similarly, the mode f_R corresponds to a second harmonic oscillator, with excitation level n_R. Let each oscillator have frequency ω.

In fig.**??** we depict these oscillators on two time slices. On the 'earlier' time slice the modes f_L, f_R are close to each other, and their corresponding oscillators should be coupled. At late times, the modes are far from each other, and the corresponding oscillators should be almost decoupled. Let the variable describing the left oscillator be ϕ_L and the variable describing the right oscillator be ϕ_R. We make a toy model of the physics by letting the oscillator be coupled for $t < 0$ and decoupled for $t > 0$. Thus the Lagrangian is

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}\dot{\phi}_L^2 + \frac{1}{2}\dot{\phi}_R^2 - \frac{1}{2}\omega^2\phi_L^2 - \frac{1}{2}\omega^2\phi_R^2 - a\phi_R\phi_L, \quad t < 0 \\ &= \frac{1}{2}\dot{\phi}_L^2 + \frac{1}{2}\dot{\phi}_R^2 - \frac{1}{2}\omega^2\phi_L^2 - \frac{1}{2}\omega^2\phi_R^2, \quad t > 0 \end{aligned} \tag{2.46}$$

2.3.1 The state for $t < 0$

We can decouple these two oscillators by going to a new basis

$$\phi_1 = \frac{1}{\sqrt{2}}(\phi_L + \phi_R), \qquad \phi_2 = \frac{1}{\sqrt{2}}(\phi_L - \phi_R) \tag{2.47}$$

which gives two uncoupled oscillators with frequencies

$$\phi_1 : \quad \omega_1 = \sqrt{\omega^2 + a}, \quad \omega_2 = \sqrt{\omega^2 - a} \tag{2.48}$$

The oscillator with variable ϕ_1 has creation and annihilation operators $\hat{a}_1, \hat{a}_1^\dagger$ and The oscillator with variable ϕ_1 has creation and annihilation operators $\hat{a}_2, \hat{a}_2^\dagger$.

We wish match our notation as closely as possible to the notation we have used for decoupling oscillators in quantum field theory. Instead of an infinite line of points where the field ϕ was defined, we now just have two points. In place of the functions $f(t, x)$ at timne t on this line x, we now have a function of t defined on two points. We write functions on this 2-point space using a 2-component vector (a, b), with a corresponding to ϕ_L and b corresponding to ϕ_R. We define two component functions

$$f_1 = \frac{1}{\sqrt{2\omega_1}}e^{-i\omega_1 t}\frac{1}{\sqrt{2}}(1, 1), \qquad f_2 = \frac{1}{\sqrt{2\omega_2}}e^{-i\omega_2 t}\frac{1}{\sqrt{2}}(1, -1) \tag{2.49}$$

The inner product between modes f, g is

$$(f, g) = i[f^* \cdot \partial_t g - \partial_t f^* \cdot g] \tag{2.50}$$

The above modes are normalized as

$$(f_i, f_j) = \delta_{ij}, \qquad (f_i^*, f_j^*) = -\delta_{ij}, \qquad (f_i^*, f_j) = 0 \tag{2.51}$$

Now consider the 'field operator'

$$\hat{\phi} = (\hat{\phi}_1, \hat{\phi}_2) \tag{2.52}$$

Since the oscillators have been decoupled in the ϕ_1, ϕ_2 basis, we can expand the field operator just the way we did for a single oscillator

$$\hat{\phi} = f_1\hat{a}_1 + f_1^*\hat{a}_1^\dagger + f_2\hat{a}_2 + f_2^*\hat{a}_2^\dagger \tag{2.53}$$

We start with the vacuum state for these coupled oscillators

$$\hat{a}_i^\dagger|0\rangle_a = 0, \quad i = 1, 2 \tag{2.54}$$

2.3.2 Evolution for $t > 0$

At the late time slice, the modes f_L, f_R are well separated, and the coupling between them is weak. We have modeled this by letting the oscillators corresponding to these modes be decoupled for $t > 0$. The analogue of the modes (2.49) is

$$g_1 = \frac{1}{\sqrt{2\omega}} e^{-i\omega t}(1,0), \qquad g_2 = \frac{1}{\sqrt{2\omega}} e^{-i\omega t}(0,1) \tag{2.55}$$

Note that

$$(g_i, g_j) = \delta_{ij}, \qquad (g_i^*, g_j^*) = -\delta_{ij}, \qquad (g_i^*, g_j) = 0 \tag{2.56}$$

The same field operator $\hat{\phi}$ can be written as

$$\hat{\phi} = g_1 \hat{b} + g_1^* \hat{b}^\dagger + g_2 \hat{c} + g_2^* \hat{c}^\dagger \tag{2.57}$$

Thus we have

$$f_1 \hat{a}_1 + f_1^* \hat{a}_1^\dagger + f_2 \hat{a}_2 + f_2^* \hat{a}_2^\dagger = g_1 \hat{b} + g_1^* \hat{b}^\dagger + g_2 \hat{c} + g_2^* \hat{c}^\dagger \tag{2.58}$$

2.3.3 Matching at $t = 0$

As we did in the case of the single oscillator in the Heisenberg picture, we wish to express the conditions (2.54) (which define our state $|0\rangle_a$) as conditions involving the oscillators $\hat{b}, \hat{b}^\dagger, \hat{c}, \hat{c}^\dagger$. This will then allow us to express the state $|0\rangle_a$ in terms of $\hat{b}^\dagger, \hat{c}^\dagger$ excitations.

We take the inner product $(g_1, \cdot)$ on both sides of (2.58). This gives

$$\begin{aligned}
\hat{b} &= (g_1, f_1)\hat{a}_1 + (g_1, f_1^*)\hat{a}_1^\dagger + (g_1, f_2)\hat{a}_2 + (g_1, f_2^*)\hat{a}_2^\dagger \\
&= \frac{\omega + \omega_1}{2\sqrt{2}\sqrt{\omega\omega_1}}\hat{a}_1 + \frac{\omega - \omega_1}{2\sqrt{2}\sqrt{\omega\omega_1}}\hat{a}_1^\dagger + \frac{\omega + \omega_2}{2\sqrt{2}\sqrt{\omega\omega_2}}\hat{a}_2 + \frac{\omega - \omega_2}{2\sqrt{2}\sqrt{\omega\omega_2}}\hat{a}_2^\dagger \\
\hat{c} &= (g_2, f_1)\hat{a}_1 + (g_2, f_1^*)\hat{a}_1^\dagger + (g_2, f_2)\hat{a}_2 + (g_2, f_2^*)\hat{a}_2^\dagger \\
&= \frac{\omega + \omega_1}{2\sqrt{2}\sqrt{\omega\omega_1}}\hat{a}_1 + \frac{\omega - \omega_1}{2\sqrt{2}\sqrt{\omega\omega_1}}\hat{a}_1^\dagger - \frac{\omega + \omega_2}{2\sqrt{2}\sqrt{\omega\omega_2}}\hat{a}_2 - \frac{\omega - \omega_2}{2\sqrt{2}\sqrt{\omega\omega_2}}\hat{a}_2^\dagger
\end{aligned} \tag{2.59}$$

While we can find the state $0\rangle_a$ in terms of $\hat{b}^\dagger, \hat{c}^\dagger$ for any value of the coupling a, the algebra is a little simpler for $a \ll \omega^2$. In this limit we have, keeping the leading order expression for each term

$$\omega_1 \approx \omega + \frac{a}{2\omega}, \quad \omega_2 \approx \omega - \frac{a}{2\omega} \tag{2.60}$$

Then we have for the operators and their conjugates

$$\begin{aligned}
\hat{b} &\approx \frac{1}{\sqrt{2}}\hat{a}_1 - \frac{a}{4\sqrt{2}\omega^2}\hat{a}_1^\dagger + \frac{1}{\sqrt{2}}\hat{a}_2 + \frac{a}{4\sqrt{2}\omega^2}\hat{a}_2^\dagger \\
\hat{c} &\approx \frac{1}{\sqrt{2}}\hat{a}_1 - \frac{a}{4\sqrt{2}\omega^2}\hat{a}_1^\dagger - \frac{1}{\sqrt{2}}\hat{a}_2 - \frac{a}{4\sqrt{2}\omega^2}\hat{a}_2^\dagger \\
\hat{b}^\dagger &\approx \frac{1}{\sqrt{2}}\hat{a}_1^\dagger - \frac{a}{4\sqrt{2}\omega^2}\hat{a}_1 + \frac{1}{\sqrt{2}}\hat{a}_2^\dagger + \frac{a}{4\sqrt{2}\omega^2}\hat{a}_2 \\
\hat{c}^\dagger &\approx \frac{1}{\sqrt{2}}\hat{a}_1^\dagger - \frac{a}{4\sqrt{2}\omega^2}\hat{a}_1 - \frac{1}{\sqrt{2}}\hat{a}_2^\dagger - \frac{a}{4\sqrt{2}\omega^2}\hat{a}_2
\end{aligned} \tag{2.61}$$

Now we note that the combination

$$\hat{b} + \frac{a}{4\omega^2}\hat{c}^\dagger \tag{2.62}$$

has only annihilation operators $\hat{a}_1, \hat{a}_2$. Thus

$$\left(\hat{b} + \frac{a}{4\omega^2}\hat{c}^\dagger\right)|0\rangle_a = 0 \tag{2.63}$$

which has the solution

$$|0\rangle_a = Ce^{-\frac{a}{4\omega^2}\hat{b}^\dagger\hat{c}^\dagger}|0\rangle_b \otimes |0\rangle_c \tag{2.64}$$

Thus we see that if we have two oscillators with the same frequency, weakly coupled together, and then we remove the coupling suddenly, then the ground state of the initial system becomes an entangled state of the two uncoupled oscillators.

2.4 The entangled nature of the state (2.64)

We can expand the exponential in (2.64) to find

$$|0\rangle_a = C\left[\,|0\rangle_b \otimes |0\rangle_c - \left(\frac{a}{4\omega^2}\right)|1\rangle_b \otimes |1\rangle_c + \left(\frac{a}{4\omega^2}\right)^2|2\rangle_b \otimes |2\rangle_c + \dots\right] \tag{2.65}$$

Let us note the physical picture of black hole evaporation captured by a state of this kind. Hawking evaporation is a quantum process. This if we consider a given time interval, then there are probabilities for different events to occur during that interval. There is some probability that no particle is emitted; this is captures by the first term on the RHS of (2.65) which has the factor $|0\rangle_b$. But this factor comes along with the factor $|0\rangle_c$, which tells us no particle falls into the interior of the hole either. This makes sense: the particles were created in pairs, so if no particle is emitted in mode b, then no particle in falls into the hole in the corresponding mode c. The probability for this eventuality – that of no particles being produced – is given by

$$P_{00} = |C|^2 \tag{2.66}$$

The second term on the RHS (2.65) says that 1 particle is emitted in mode b, and correspondingly, 1 particle falls into the hole in mode c. The probability for this occurrence is

$$P_{11} = |C|^2 \left(\frac{a}{4\omega^2}\right)^2 \tag{2.67}$$

The next term has two particles in each mode, and so on.

States of the form (2.65) are said to be *entangled* between the systems b and c: the choice of state for the system b is dependent on the choice of state of system c. By contrast, a 'factored' state would have no relation between the states in systems b and c. Examples of factored states are

$$|0\rangle_b \otimes |0\rangle_c, \quad |0\rangle_b \otimes |1\rangle_c, \quad \frac{1}{\sqrt{2}}\left(|0\rangle_b + |1\rangle_b\right) \otimes \frac{1}{\sqrt{2}}\left(|0\rangle_c + |1\rangle_c\right) \tag{2.68}$$

The essential issue with our entangled state (2.65) can be captured by taking a simpler entangled state which has just two terms in the sum

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle_b|0\rangle_c + |1\rangle_b|1\rangle_c\right) \tag{2.69}$$

We will work with this simpler state $\psi\rangle$ in the discussion below.

The widespread occurrence of entangled states

Entangled states are found very often in nature. Consider two electrons, labelled 1 ad 2. The states of each electron can be described by using a basis where the z spin is $\pm\frac{1}{2}$; let us call these states $|\pm\rangle_1, |\pm\rangle_2$.

Suppose the two electrons are in a singlet state; i.e., a state with overall spin 0

$$|\psi\rangle_{singlet} = \frac{1}{\sqrt{2}}\left(|+\rangle_1|-\rangle_2 - |-\rangle_1|+\rangle_2\right) \tag{2.70}$$

This is an entangled state since the spin of electron 1 is dependent on the spin of electron 2. Electron 1 does not have a definite state by itself: it is spin up if electron 2 has spin down and it has spin down if electron 2 has spin up.

More generally, entangled states are created by interactions. If a particle p is interacting with several other particles, then after some time the state of p will become entangled with the state of the reming system described by the other particles.

Thus in quantum theory, entangled states are more the norm than the exception. But as we will now see, the entanglement created by black hole evaporation leads to a serious problem for quantum theory.

Entanglement in the evaporating hole

We proceed in the following steps:

(i) Start with a ball of matter, in some state $\Psi\rangle_A$. This is the matter (which we had earlier termed 'A'), which will collapse and make the hole. While this matter A could have been entangled with some other system somewhere else, it is simplest to assume that there is no such entanglement; thus $\Psi\rangle_A$ is what is called a 'pure' state.

(ii) Now let A collapse to make a black hole. The hole is in a pure state characterized by A; let us call this state $\tilde{\Psi}\rangle_A$.

(iii) Now suppose a pair is created out of the vacuum by the Hawking process. Let us model the state of this pair by the simplified state (2.69):

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle_{b_1}|0\rangle_{c_1} + |1\rangle_{b_1}|1\rangle_{c_1}) \tag{2.71}$$

where we have added a subscript '1' to the modes b, c to indicate that these modes correspond to the first pair that will be emitted from the hole. Imagine dividing spacetime into two regions:

The outer region: The region outside the hole. For concreteness, let this be the region $r > 10M$, where we are away from any effects that might be particular to the region near the horizon.

The inner region: This is the region $r < 10M$, containing the hole and its vicinity.

We see that the state in the outer region is entangled with the state in the inner region, due to the entanglement in (2.69). The state of b_1, c_1 does not depend on the details of the state of the matter A which made the hole.

(iv) Now consider the emission of a second pair from the vacuum. The state of this pair is

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle_{b_2}|0\rangle_{c_2} + |1\rangle_{b_2}|1\rangle_{c_2}) \tag{2.72}$$

Thus the members of the second pair are entangled with each other, but they are not related to the initial matter of the hole A or the members of the first pair. The overall state in the outer region is now 'more' entangled with the state in the inner region; we will note how to define the amount of this entanglement in section **??**.

(v) After N steps of emission, the overall state is

$$\begin{aligned} |\Psi\rangle &= |\tilde{\Psi}\rangle_A \otimes \frac{1}{\sqrt{2}}\Big(|0\rangle_{b_1}|0\rangle_{c_1} + |1\rangle_{b_1}|1\rangle_{c_1}\Big) \\ &\otimes \frac{1}{\sqrt{2}}\Big(|0\rangle_{b_2}|0\rangle_{c_2} + |1\rangle_{b_2}|1\rangle_{c_2}\Big) \\ &\quad \cdots \\ &\otimes \frac{1}{\sqrt{2}}\Big(|0\rangle_{b_N}|0\rangle_{c_N} + |1\rangle_{b_N}|1\rangle_{c_N}\Big) \end{aligned} \tag{2.73}$$

The state in the outer region is now very heavily entangled with the state in the inner region. This is still not a problem however; any many-particle quantum system will typically be highly entangled with other systems that it is able to interacted with.

(vi) The problem comes from the curious nature of gravity: that it is attractive. This gives the potential energy a negative sign, and we have seen that the net energy of the initial matter A together with the infalling particles $\{c_k\}$ keeps going down with each emission. The black hole shrinks towards a situation where it has zero mass. The simplest assumption – and the one that Hawking made in 1975 – is that the black hole evaporates away to nothing, giving us the vacuum state at $r = 0$.

The mass M of the initial matter A is now in the Hawking radiation quanta $\{b_k\}$. So energy is conserved; which is satisfying. But the problem comes when we ask: what is the state of the quanta $\{b_k\}$?

The quanta $\{b_k\}$ had no definite state by themselves; their state was defined in conjunction with their partners $\{c_k\}$. Thus the state in the mode b_1 was $|0\rangle_{b_1}$ *if* the state in the mode c_1 was $|0\rangle_{c_1}$, and the state in the mode b_1 was $|1\rangle_{b_1}$ *if* the state in the mode c_1 was $|1\rangle_{c_1}$. But if the quanta in the mode c_1 have disappeared from the universe, then what is the state in the mode b_1?

This is the crux of the information paradox. If we cannot assign a state to the quanta in the modes $\{b_k\}$, then we have a violation of quantum theory, where the universe must be described by a state $\Psi\rangle(t)$ at all times. We started with matter A described by a definite state $|\Psi\rangle_A$, but at the end of the evaporation process we are left with quanta $\{b_k\}$ which cannot be assigned a state. Thus the process of black hole formation and evaporation cannot be captured by normal quantum evolution, which takes states to other states

$$|\Psi\rangle_f = e^{-i\hat{H}t}|\Psi\rangle_i \tag{2.74}$$

2.4.1 States vs density matrices

The above conclusion is startling, so let us take a moment to explore it in more detail.

Consider again the state (2.70) of two electrons in an overall singlet state

$$|\psi\rangle_{singlet} = \frac{1}{\sqrt{2}}\left(|+\rangle_1|-\rangle_2 - |-\rangle_1|+\rangle_2\right) \tag{2.75}$$

Neither electron has a state by itself; this is an entangled state of the two electrons. Now suppose electron 1 were to disappear from the universe. Could we assign any state to electron 2?

One might think that we could just delete all reference to the state of electron 1, getting the state

$$|\psi\rangle_{singlet} \to \frac{1}{\sqrt{2}}\left(|-\rangle_2 - |+\rangle_2\right) \tag{2.76}$$

for electron 2. This looks like a sensible state; it is an eigenstate of the Pauli matrix $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ with eigenvalue -1, so it represents a spin for electron 2 pointing in the direction $-\hat{x}$.

But there is a problem with this procedure. We could have chosen a slightly different basis for the states of electron 1

$$|+\rangle_1 \to |+\rangle_1, \qquad |-\rangle_1 \to e^{i\theta}|-\rangle_1 \tag{2.77}$$

The singlet state (2.75) would now be written as

$$|\psi\rangle_{singlet} = \frac{1}{\sqrt{2}}\left(|+\rangle_1|-\rangle_2 - e^{-i\theta}|-\rangle_1|+\rangle_2\right) \tag{2.78}$$

If we now delete all reference to electron 1, we find

$$|\psi\rangle_{singlet} \to \frac{1}{\sqrt{2}}\left(|-\rangle_2 - e^{-i\theta}|+\rangle_2\right) \tag{2.79}$$

This is *not* the same as the state (2.76). For example, for $\theta = \pi/2$, we get an eigenstate of the Pauli matrix $\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ with eigenvalue -1, so it represents a spin for electron 2 pointing in the direction $-\hat{y}$.

Thus we see that the structure of quantum mechanics does not allow us to delete all reference to one particle of an entangled pair, and obtain a unique state for the remaining particle.

If we cannot get a state for electron 2, is there *anything* that we can say about it? In the initial state (2.75) there was one thing we could say about electron 2 by itself: the *probability* for it to have spin $|+\rangle$

$$P_{\frac{1}{2}} = \left|\frac{1}{\sqrt{2}}\right|^2 = \frac{1}{2} \tag{2.80}$$

and the probability for it to have spin $|-\rangle$ was $P_{-\frac{1}{2}} = \frac{1}{2}$ as well. If we somehow 'delete' spin 1 from our universe, then the states (2.79) still have the the probabilities $P_{\pm\frac{1}{2}} = \frac{1}{2}$, regardless of the value of θ; i.e., regardless of how we chose the basis of states for electron 1. If we can specify only the probabilities, and not the actual state for an electron, then we say that the electron is described by a *density matrix* rather than a quantum state. We will see a detailed mathematical description of density matrices later.

Let us return to entangled state (2.73) of the black hole. If the hole evaporates away leaving only the quanta in modes $\{b_k\}$, then we cannot have a state for the $\{b_k\}$, but we can have *probabilities* for various possible states. Since each mode b_k can be in the states 0 or 1, there are 2^N possible states for the modes $\{b_k\}$, with each state labelled by a sequence of 0s and 1s; e.g. $0100010\ldots011$. From (2.73) we see that the probabilities for each of these states is the same

$$P_{000\ldots00} = \frac{1}{2^N}, \ldots P_{011\ldots01} = \frac{1}{2^N}, \ldots P_{111\ldots11} = \frac{1}{2^N} \tag{2.81}$$

This information about probabilities is described by the density matrix of the set $\{b\}$. The density matrix clearly has less information than a quantum state; any state of the form

$$|\Psi\rangle_b = \frac{1}{2^{\frac{N}{2}}}\left(e^{i\theta_1}|000\ldots 00\rangle + \ldots e^{i\theta_n}|011\ldots 01\rangle + \ldots e^{i\theta_{2^N}}|111\ldots 11\rangle\right) \quad (2.82)$$

has the probabilities (2.81). Thus we lose the phase information contained in the θ_k when we go from a state to a density matrix.

In the process of particle scattering, the usual quantum evolution (2.74) is described by an 'S-matrix'. This S-matrix is unitary

$$S^\dagger = S^{-1} \quad (2.83)$$

Hawking proposed that in a theory with gravity, there is no S-matrix, and thus no unitarity; all we can have is a 'Dollar matrix' $ which describes the evolution of density matrices. This of course is a radical change in the foundations of quantum theory, and caused much consternation. It also lead people to consider the idea of *remnants*.

2.4.2 Remnants, and their difficulties

Given the seriousness of the problem caused by the complete evaporation of the hole, one may try to say that the evaporation stops when the hole reaches planck scale. When the hole is much larger than planck scale, the semiclassical approximation would appear to be valid, and then the separation of geodesics at the horizon continues to give the creation of entangled particle pairs. But when the hole is planck size, then new quantum gravity effects may creep in, and somehow stop the further evaporation of the hole towards the vacuum.

The planck sized object that is left if the evaporation of the hole stops is called a remnant. This remnant must contain the initial matter A, as well as all the negative energy quanta in the modes $\{c_k\}$. With this situation, there is no problem with the quantum mechanics as such, since we have not lost the wavefunction of the matter A, and we have retained all members of the entangled sets $\{b\}, \{c\}$. The difficulties with remnants all stem from the question: how can such a small remnant hold all these quanta?

One might try to bypass this problem by postulating that all the quanta involved in $A, \{c\}$ get crushed to a unique state $|\psi\rangle_{remnant}$ inside the hole, and this state is characterized by a small spatial extent ($\sim l_p$) and small energy ($\sim m_p$). But this would be a violation of quantum mechanics. The full state $|\Psi\rangle$ (2.73) had a large entanglement between the exterior region and what is in the hole. More precisely, the state $|\Psi\rangle$ has the form

$$|\Psi\rangle = \frac{1}{2^{\frac{N}{2}}}\sum_{i=1}^{2^{\frac{N}{2}}} |\chi_i\rangle|\psi_i\rangle \quad (2.84)$$

where $|\chi_i\rangle$ are orthonormal states in the interior and $|\psi_i\rangle$ are orthonormal states in the exterior. Now it is true that we do not know the nature of the evolution in the interior when the hole reaches planck size. But under a unitary evolution (2.74), orthonormal states evolve to orthonormal states. Thus any evolution inside the hole can only produce a change of the form

$$\frac{1}{2^{\frac{N}{2}}}\sum_{i=1}^{2^{\frac{N}{2}}}|\chi_i\rangle|\psi_i \to \frac{1}{2^{\frac{N}{2}}}\sum_{i=1}^{2^{\frac{N}{2}}}|\chi'_i\rangle|\psi_i \tag{2.85}$$

where the $|\chi'_i\rangle$ again form a 2^N dimensional set of orthonormal states. Thus the remnant must have at least 2^N possible internal states. But N can be as large as we want: if we start with larger and larger holes and let them evaporate down to planck size, then we get larger and larger values of N in (2.73). Thus we should be able to have infinitely many states possible for planck sized remnants.

Lecture notes 3

Particle creation in black holes

3.1 The full structure of the classical black hole

We have seen that Hawking radiation arises because the vacuum around the horizon is *unstable* to the production of particle pairs. This instability, in turn arises from the fact that particle trajectories of the two sides of the horizon get 'pulled apart': a trajectory starting just outside the horizon can escape to infinity, while a trajectory starting just inside must fall to the center of the hole. The simplest trajectories are the radial trajectories of particles moving at the speed of light. Our goal is to write the metric in a form where we can study such trajectories easily, both outside and inside the horion. This form of the metric is achieved by going to Kruskal coordinates, which allow us to see the entire geometry of the classical black hole metric.

The physics of light rays is best seen in *null* coordinates. In flat 1+1 dimensional Minkowski space

$$ds^2 = -dt^2 + dx^2 \tag{3.1}$$

we can define the'null coordinates

$$u = t + x, \quad v = t - x \tag{3.2}$$

A lightlike trajectory moving left is given by $u = constant$ and one moving right is given by $v = constant$. Let us now look for similar coordinates for the black hole.

Our Schwarzschild metric is

$$ds^2 = -(1 - \frac{2M}{r})dt^2 + (1 - \frac{2M}{r})^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{3.3}$$

We write this as

$$ds^2 = (1 - \frac{2M}{r})\left[-dt^2 + \frac{dr^2}{(1-\frac{2M}{r})^2}\right] + r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{3.4}$$

It is useful to define a new coordinate r^* by

$$dr^* = \frac{dr}{(1 - \frac{2M}{r})} \tag{3.5}$$

which gives

$$r^* = r + 2M\ln(\frac{r}{2M} - 1) \tag{3.6}$$

where we have set the arbitrary additive constant to zero. Clearly this definition is singular at $r = 2M$, so let us restrict attention for the moment to the region outside the hole $r > 2M$. The metric is now

$$ds^2 = (1 - \frac{2M}{r}) \left[-dt^2 + dr^{*2} \right] + r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{3.7}$$

We can again define null coordinates

$$u = t + r^*, \quad v = t - r^* \tag{3.8}$$

Consider a null geodesic falling radially into the hole. Thus θ, ϕ are constant, and the worldline will be given by solving $ds^2 = 0$ in the (t, r^*) space. At infinity where the metric is flat the ingoing geodesic is $t + r = const.$. From (3.9) we see that taking into account the metric of the hole changes this to

$$t + r^* \equiv u = u_0 \tag{3.9}$$

Similarly, a radially outgoing null ray is given by

$$t - r^* \equiv v = v_0 \tag{3.10}$$

3.1.1 Coordinate ranges

Let us now look at the ranges of the coordinates u, v. For the radial coordinate, we see from (3.6) that the range $r = (2M, \infty)$ maps to $r^* = (-\infty, \infty)$.

Now consider the null geodesics (3.9) falling radially into the hole. By taking geodesics starting from a given r^* with different values of t we can cover the full range

$$-\infty < u_0 < \infty \tag{3.11}$$

Similarly, for radially outgoing geodesics, we can cover the full range

$$-\infty < v_0 < \infty \tag{3.12}$$

Suppose an observer starts outside the black hole, and decides tp fall into the hole. At some future time, he will pass through the horizon of the hole. The spacetime points where such observers cross the horizon are define the 'future horizon'. For an infa;;ing trajectory $u = u_0$, we have

$$v = t - r^* = u_0 - 2r^* \tag{3.13}$$

As the observer approaches the horizon, we see that $v \to \infty$. Thus the future horizon is given by the points

$$-\infty < u < \infty, \quad v = \infty \tag{3.14}$$

3.1.2 Defining Kruskal coordinates

From (3.14) we see that our coordinates (u, v) 'end' at the horizon. We have already noted, however, the the horizon is actually a regular place in the classical black hole, and we should be able to smoothly pass to a region interior to this horizon. To see the horizon as a normal place in our manifold, we would like to have coordinates that are smooth at the horizon. In particular, we need the horizon to appear at *finite* values of our coordinates, unlike (3.14). To achieve this, we write

$$U = e^{\alpha u}, \quad V = -e^{-\alpha v} \tag{3.15}$$

where we will choose the constant α later. Assuming $\alpha > 0$, we see that the region outside the horizon is

$$U > 0, \quad V < 0 \tag{3.16}$$

and the horizon itself is

$$0 < U < \infty, \qquad V = 0 \tag{3.17}$$

Thus we have brought the horizon to a finite position in our new coordinates U, V, and if the metric is smooth at $U = V = 0$ then we can continue the spacetime past the region (3.16).

Let us now see if the coordinates U, V can be made smooth at the horizon. From (3.15) we get

$$dU = \alpha e^{\alpha u} du, \qquad dV = \alpha e^{-\alpha v} dv \tag{3.18}$$

Thus the metric (3.9) becomes

$$ds^2 = -(1-\frac{2M}{r})\frac{e^{-\alpha(u-v)}}{\alpha^2}dUdV + r^2 d\Omega_2^2 = -\frac{(r-2M)}{r}\frac{e^{-\alpha(u-v)}}{\alpha^2}dUdV + r^2 d\Omega_2^2 \tag{3.19}$$

Now note that

$$e^{-\alpha(u-v)} = e^{-2\alpha r^*} = e^{-2\alpha[r+2M\ln(\frac{r}{2M}-1)]} = e^{-2\alpha r}\left(\frac{2M}{r-2M}\right)^{4\alpha M} \tag{3.20}$$

We now see that if we choose

$$\alpha = \frac{1}{4M} \tag{3.21}$$

then we cancel the factor $r - 2M$ in (3.19), getting

$$ds^2 = -\frac{32M^3}{r}e^{-\frac{r}{2M}} dUdV + r^2 d\Omega^2 \tag{3.22}$$

The metric is now written in coordinates U, V, θ, ϕ. The variable r should now be thought of as a function $r(U, V)$, given through the transcendental relation

$$UV = -(\frac{r}{2M} - 1)e^{\frac{r}{2M}} \tag{3.23}$$

Since we do not need the explicit form of this function for our analysis, we leave it as the symbol r in most of our expressions below.

3.2 The extended black hole spacetime

The Schwarzschild metric (3.3) described the region outside the horizon

$$-\infty < t < \infty, \qquad 2M < r < \infty \tag{3.24}$$

This region is covered by the Kruskal coordinates

$$\begin{aligned} U &= (\frac{r}{2M} - 1)^{\frac{1}{2}} e^{\frac{r}{4M}} e^{\frac{t}{4M}} \\ V &= -(\frac{r}{2M} - 1)^{\frac{1}{2}} e^{\frac{r}{4M}} e^{-\frac{t}{4M}} \end{aligned} \tag{3.25}$$

with

$$U > 0, \quad V < 0 \tag{3.26}$$

Let us now look at the spacetime we get when we continue to other values of U, V:

(i) First we note that the location $r = 0$ is a real singularity of the metric; the curvature diverges as $r \to 0$, and so we cannot remove the singularity here by a change of coordinates. From (3.23) we see that $r = 0$ corresponds to $UV = 1$. Thus the singularity lies on two surfaces

$$\begin{aligned} U > 0, \quad V > 0, \quad UV = 1: &\quad \text{future singularity} \\ U < 0, \quad V < 0, \quad UV = 1: &\quad \text{past singularity} \end{aligned} \tag{3.27}$$

We depict these surfaces as hyperbolae on the U, V plane in fig.**??**.

(ii) Since we cannot continue past the singularity at $UV = 1$, we will restrict ourselves to the region $UV < 1$. We now note that the metric (3.22) and the expression (3.23) for $r(U, V)$ are symmetric under the interchange $U \leftrightarrow V$. The region (3.26) described the exterior of the hole, stretching from the horizon $r = 2M$ to asymptotic infinity. We can get a second copy of such a region, from the range

$$U < 0, \quad V > 0 \tag{3.28}$$

Thus the wedge on the left in fig.**??** describes a second asymptotic infinity.

(iii) We have seen in (3.17) that a future horizon is located at

$$0 < U < \infty, \quad V = 0 \tag{3.29}$$

By the symmetry under $U \leftarrow V$, we get another future horizon at

$$0 < V < \infty, \quad U = 0 \tag{3.30}$$

Now consider the surfaces

$$-\infty < U < 0, \quad V = 0 \tag{3.31}$$

and

$$-\infty < V < 0, \quad U = 0 \tag{3.32}$$

We will call these surfaces the 'past horizons'. The future horizons were surfaces across which particles could not emerge from the hole to the exterior. For the past horizons, we see from fig.**??** that no particle can fall through these surfaces *into* the hole from the exterior.

The wedge bounded by the future singularity and the future horizons is a black hole: particles can fall into this wedge but not emerge out. The wedge bounded by the past singularity and past horizons is called a 'white hole': particles can emerge from this wedge, but not go in.

The spacetime obtained above is called the fully extension of the Schwarzschild metric. As we will see later, it is also called the 'extremal black hole'. We can introduce Schwarzschild like coordinates to cover each of the four wedges of this fully extended diagram. For example, for the future wedge, we define the coordinate r^* as

$$d\tilde{r}^* = -\frac{dr}{(\frac{2M}{r}) - 1} \tag{3.33}$$

which gives

$$\tilde{r}^* = r + 2M\ln(1 - \frac{r}{2M}) \tag{3.34}$$

Define

$$\tilde{u} = \tilde{r}^* + \tilde{t}, \quad \tilde{v} = \tilde{r}^* - \tilde{t} \tag{3.35}$$

and

$$U = e^{\frac{\tilde{u}}{4M}}, \quad V = \frac{\tilde{v}}{4M} \tag{3.36}$$

Then the metric (3.22) becomes

$$ds^2 = (\frac{2M}{r} - 1)d\tilde{t}^2 - \frac{dr^2}{\frac{2M}{r} - 1} + r^2 d\Omega^2 \tag{3.37}$$

so we get Schwarzschild the coordinates inside the horizon with $\tilde{t}$ now being spacelike and r being timelike.

This extended metric has several remarkable features. For one thing, we found a white hole along with the black hole. This is not too surprising, since the Schwarzschild metric (3.3) that we started with was symmetric under the reflection of the time direction $t \to -t$. If the full extension of the metric described a black hole from which nothing could come out, then we should also find a region into which nothing can go in.

The extended black hole metric describes a vacuum spacetime; i.e., there is no matter anywhere in the region $r > 0$. One can imagine that the actual spacetime will be quite different if we made the black hole by starting with a ball of matter and letting it collapse to $r = 0$. collapsing a ball of matter. For example, the white hole region has a singularity in the past. If we start with initial conditions that were nonsingular, then we would not have this past singularity, and perhaps no while hole region.

The other remarkable feature of the extended metric is the appearance of second asymptotic infinity – the left wedge. Again, this is not something we would expect if we made a black hole by collapsing a ball of matter. there is an infinite region outside the ball, and we can imagine that there may be puzzling features inside the horizon region $r < 2M$ after the ball collapses. But we do not expect to generate by such a collapse a region which contains *a whole new asymptotic infinity*.

We will see below that when a black hole is made by collapsing a ball of matter, we do not get the white hole region or the second asymptotic infinity. We do get a region inside the future horizon – the interior of the hole. The existence of this interior region is responsible for Hawking's puzzle with black hole evaporation.

It may therefore seem that the full extended black hole is a mathematical object with no physical relevance. But as we will see later, the information paradox can be posed for the extended black hole spacetime as well, and analyzing the paradox in this context will lead to important clues about what quantum gravity effects can do in spacetimes with horizons.

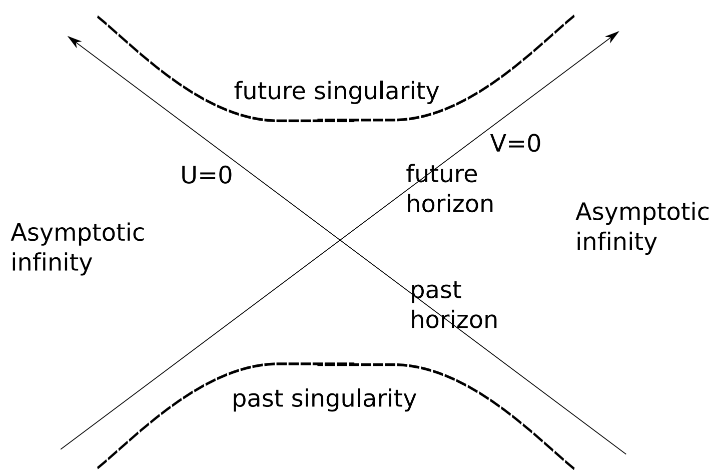

Figure 3.1: The fully extended Schwarzschild geometry

3.3 The Penrose diagram

Before we proceed to study the black hole made by collapse, let us pause to discuss a convenient way of depicting the various spacetime geomerries that we will encounter.

The U, V coordinates cover all of our extended black hole spacetime. These coordinates do not have a bounded range. Thus if we try to draw the U, V space on a sheet of paper, then we can cover only a finite subset of the whole spacetime, and in particular we cannot see the picture of how the 'points at infinity' border our spacetime.

We would like to bring these 'points at infinity' to a finite distance from the points in the interior of our spacetime. To do this, we make a *conformal rescaling* of the metric. Here the word 'conformal' means that at each point the metric is scaled by a number

$$g_{ab}(x) \to \Omega^2(x) g_{ab}(x) \tag{3.38}$$

The factor $\Omega^2(x)$ depends on the position x, and will be taken to be large near infinity; this is what will squeeze the infinite spacetime to a compact region. This squeezing will introduce a distortion in our depiction, which we will have to take into account when looking at the resulting figure. But one fact does not change: if the separation between two points was null – i.e., $ds^2 = 0$ – then it will remain null after the rescaling (3.38). We usually draw null lines as lines with slope ± 1, and this will remain the case after the rescaling. Since horizons is are null surfaces, they will appear as a line with slope ± 1 in our rescaled figure.

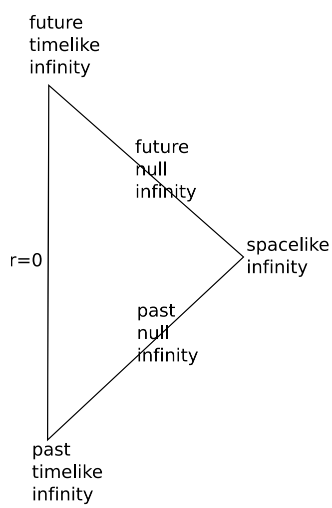

Figure 3.2: Penrose diagram of Minkowski space

Let us first carry out this rescaling process for Minkowski spacetime. We first write the spatial metric in polar coordinates

$$ds^2 = -dt^2 + dr^2 + r^2 d\Omega_2^2 \tag{3.39}$$

Define

$$U = t + r, \quad V = t - r \tag{3.40}$$

so that

$$ds^2 = -dUdV + r^2 d\Omega^2 \tag{3.41}$$

where now the coordinates are U, V, θ, ϕ, and

$$r = \frac{1}{2}(U - V) \tag{3.42}$$

Since $r > 0$, we find the alllowed range

$$-\infty < U < \infty, \quad -\infty < V < \infty, \quad U \geq V \tag{3.43}$$

We observe that the range of U, V is not finite. To make it finite, define

$$\tilde{U} = \tanh U, \quad \tilde{V} = \tanh V \tag{3.44}$$

Now we have

$$-1 < \tilde{U} < 1, \quad -1 < \tilde{V} < 1, \quad \tilde{U} \geq \tilde{V} \tag{3.45}$$

and the metric is

$$ds^2 = -[\frac{dU}{d\tilde{U}}\frac{dV}{d\tilde{V}}]d\tilde{U}\tilde{V} + r^2 d\Omega^2 \tag{3.46}$$

But

$$\frac{dU}{d\tilde{U}} = \operatorname{sech}^2 U = \frac{1}{1-\tilde{U}^2}, \quad \frac{dV}{d\tilde{V}} = \operatorname{sech}^2 V = \frac{1}{1-\tilde{V}^2} \tag{3.47}$$

so we have

$$ds^2 = \frac{1}{(1-\tilde{U}^2)(1-\tilde{V}^2)}[-d\tilde{U}d\tilde{V} + r^2(1-\tilde{U}^2)(1-\tilde{V}^2)d\Omega^2] \tag{3.48}$$

So far we have just rewritten Minkowski spacetime in new coordinates. Now make a conformal transformation, defining a *new* metric

$$g'_{ab} = (1-\tilde{U}^2)(1-\tilde{V}^2)g_{ab} \tag{3.49}$$

This new metric is

$$ds'^2 = -d\tilde{U}d\tilde{V} + r^2(1-\tilde{U}^2)(1-\tilde{V}^2)d\Omega^2 \tag{3.50}$$

Let us ignore the angular directions; since we have spherical symmetry there is no nontrivial structure in these directions, and we cannot depict more than two directions on our figure anyway. Thus we focus on the metric

$$ds'^2 = -d\tilde{U}d\tilde{V} \tag{3.51}$$

This metric, with the coordinate ranges (3.45), gives the spacetime picture in fig.3.2. Such a figure is called a *Penrose diagram* of the spacetime.

The null directions in this spacetime are $\tilde{U} = U_0$ and $\tilde{V} = V_0$; these are lines with slope ± 1.

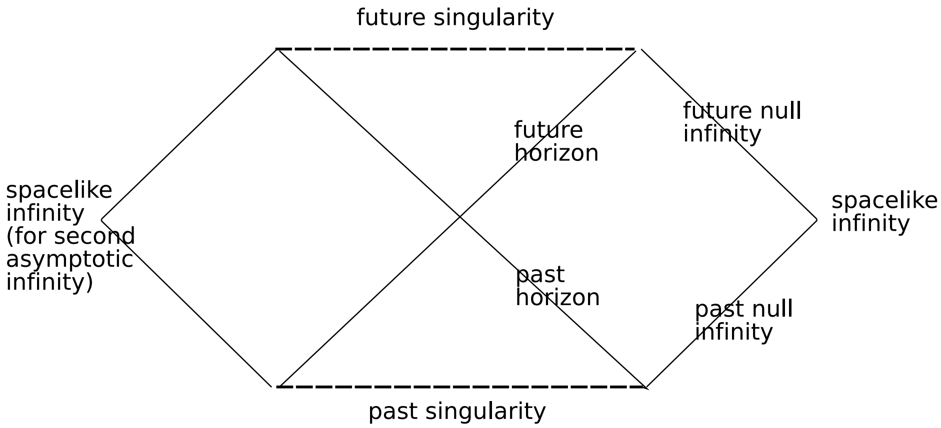

Figure 3.3: Penrose diagram for the 'eternal Schwarzschild hole'

Let us now find the Penrose diagram for the extended black hole metric (3.22)

$$ds^2 = -\frac{32M^3}{r} e^{-\frac{r}{2M}} dU dV + r^2 d\Omega^2 \tag{3.52}$$

We squeeze the range of U, V by the transformation (3.44). We then scale the metric as in (3.49). Dropping the angular directions as before, we find

$$ds'^2 = -\frac{32M^3}{r} e^{-\frac{r}{2M}} d\tilde{U} d\tilde{V} \tag{3.53}$$

Let us now look carefully at the ranges spanned by our coordinates. The spacetime again ends at $r = 0$; this time there is a singularity there instead of a 'simple origin of coordinates'. But $r = 0$ is now given by solving $UV = 1$ which is

$$\tanh^{-1}\tilde{U}\ \tanh^{-1}\tilde{V} = 1 \tag{3.54}$$

This is a curve in $\tilde{U}, \tilde{V}$ space, and points beyond this curve are not in the spacetime represented by the Penrose diagram, since they lie past the singularity. Thus our coordinates span the range

$$-1 < \tilde{U} < 1, \quad -1 < \tilde{V} < 1, \quad \tanh^{-1}\tilde{U}\ \tanh^{-1}\tilde{V} < 1 \tag{3.55}$$

We draw the resulting Penrose diagram in fig.3.3.

The singularity runs along a curve from $\tilde{U} = 0, \tilde{V} = 1$ to $\tilde{U} = 1, \tilde{V} = 0$. Note that the rescaling that we have perform to bring infinity to a finite place is not changed if we perform a further conformal rescaling at *interior* points of spacetime. Thus we can imagine a further rescaling which makes the singularity a straight line from $\tilde{U} = 0, \tilde{V} = 1$ to $\tilde{U} = 1, \tilde{V} = 0$; this is easier to draw, and is typically what is done in drawing Penrose diagrams. The essential property of the singularity we cannot change in the picture is that the singularity is *spacelike*; The constant r surface $r = 0$ is inside the horizon and so is spacelike instead of timelike.

3.4 The black hole formed by collapse

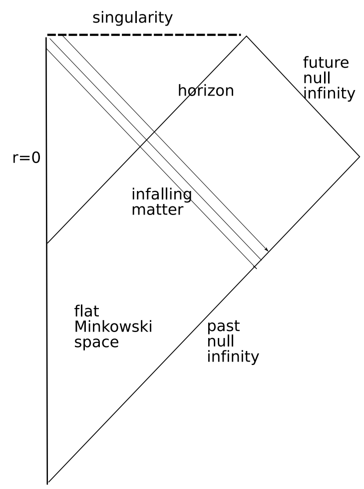

Figure 3.4: Penrose diagram of the black hole made by collapse of a shell

We have seen that the extended black hole spacetime obtained by extending the Schwarzschild metric has some strange features. We expect these features to be absent when we take a more realistic hole, which has been made by starting with a nonsingular distribution of matter and letting this collapse to create a horizon.

The simplest geometry for this matter is that of a spherically symmetric shell, with infinitesimal thickness. We then find the following:

(a) In Newtonian mechanics, the potential inside a spherical shell is $\Phi = constant$, so there is no gravitational field. A similar situation holds in general relativity, where a result called the Birkoff theorem tells us that metric inside a spherical shell is flat.

(b) In Newtonian mechanics the potential outside a spherically symmetric body of mass M is is the same as the potential of a point mass M at the center of the shell: $\Phi = -GM/r$. A similar situation holds in general relativity, where the Birkoff theorem tells us metric outside a spherical body of mass M is the Schwarzschild metric with mass M.

Thus we have the metric inside and outside the shell, and all we need to do is relate the choice of coordinates on the two sides by matching conditions at the shell.

These matching conditions need some work however. The shell falls in the metric which it crates by its own stress energy tensor, and so its trajectory is a nontrivial function of its mass. The problem is simplified if we take a *null* shell, which is composed of massless particles moving radially inwards. Since these massless particles must move at the speed of light, the trajectory of the shell can be written down immediately.

The metrics created by null shells are called Vaidya metrics. We will take a single infinitesimally thin null shell, and find the full spacetime geometry by matching coordinates on the two sides of this shell.

3.4.1 The metric inside the shell

Inside the shell we have flat Minkowski spacetime

$$ds_f^2 = -dt_f^2 + dr^2 + r^2 d\Omega^2 \tag{3.56}$$

The subscript f indicates the metric is 'flat'. We have also placed this subscript on the ti,e coordinate, to distinguish it from the time coordinate we will find outside the shell. The r coordinate does not need such a subscript, since we will take $ds^2 \to r^2 d\Omega^2$ both inside and outside the shell. Thus r has the geometric significance of being the proper radius of the angular sphere at any location, and in particular the value of r just inside the shell will agree with the value just outside the shell. The angular coordinates θ, ϕ can also be taken to be the same inside and outside, so we do not give them a subscript either.

We pass to null coordinate by writing

$$u_f = t_f + r, \quad v_f = t_f - r \tag{3.57}$$

The metric inside the shell then becomes

$$ds_f^2 = -du_f dv_f + r^2 d\Omega^2 \tag{3.58}$$

The inner surface of the shell describes a radial null trajectory in this metric. By choosing th origin of t_f appropriately, we can let this trajectory be

$$u_f = 0 \tag{3.59}$$

Note that

$$r = \frac{u_f - v_f}{2} = -\frac{v_f}{2} \tag{3.60}$$

3.4.2 The metric outside the shell

Outside the shell we have the metric corresponding to a mass M

$$ds_{BH}^2 = -\frac{32M^3}{r} e^{-\frac{r}{2M}} dU dV + r^2 d\Omega^2 \tag{3.61}$$

We have added a subscript K (for Kruskal) to U, V to differentiate them from the Minkowski coordinates u_f, v_f inside the shell. Outside the horizon, i.e. in the region $V < 0$, we have

$$U = e^{\frac{u}{4M}}, \quad V = -e^{-\frac{v}{4M}} \tag{3.62}$$

where

$$u = t_K + r^*, \quad v = t_K - r^* \tag{3.63}$$

and t_K goes over to the usual time coordinate at infinity.

The outer surface of the shell must follow an ingoing null geodesic in the metric (3.61). This geodesic will have the form $U = U_0$ for some constant U_0. Recall that $U = Exp[u/4M]$, where $u = t + r^*$. We can choose the origin of t to set $u = 0$ along the shell, which gives

$$U = 1 \tag{3.64}$$

3.4.3 Matching coordinate systems

Consider two infinitesimally separated points on the trajectory of the infalling shell. We have noted that the value of r is a geometric object, since it gives the proper radius of the angular sphere. Thus the separation dr between these two points must be the same on both sides of the shell.

On the inner surface of the shell, we have

$$r = \frac{u_f - v_f}{2} = -\frac{v_f}{2} \tag{3.65}$$

where we have used (3.59). Thus

$$dr = -\frac{1}{2} dv_f \tag{3.66}$$

On the outer surface of the shell, we use the relation (3.23) giving $r(U, V)$. Noting that $U = 1$ along the trajectory of the shell, we find

$$dV = -\frac{r}{4M^2} e^{\frac{r}{2M}} dr \tag{3.67}$$

We substitute the value of r from the inner side of the shell, since this is given by the simpler expression (3.65). We get

$$dr = -\frac{4M^2}{r} e^{-\frac{r}{2M}} dV = \frac{8M^2}{v_f} e^{\frac{v_f}{4M}} dV \tag{3.68}$$

Equating (3.66) and (3.68), we get

$$dv_f = -\frac{16M^2}{v_f} e^{\frac{v_f}{4M}} dV \tag{3.69}$$

This has the solution

$$V = \frac{(v_f + 4M)}{4M} e^{-\frac{v_f}{4M}} + C \tag{3.70}$$

The constant C is fixed by looking at the location of the horizon $r = 2M$. Inside the shell, we see from (3.60) that the horizon is $v_f = -4M$. Outside the shell we have seen that the horizon is at $V = 0$. Thus $C = 0$, and we have

$$V = \frac{(v_f + 4M)}{4M} e^{-\frac{v_f}{4M}} \tag{3.71}$$

We have seen that trajectories that start on opposite sides of the horizon diverge away from the horizon on opposite sides:

(a) The trajectories with $V < 0$ lie outside the horizon, and eventually escape to infinity where the natural coordinate is v (given by (3.62)).

$$v = -4M \ln(-V) = v_f - 4M \ln\left(-\frac{(v_f + 4M)}{4M}\right) \tag{3.72}$$

For trajectories that start very close to the horizon, we have $v_f + 4M \to 0$, and we observe a logarithmic relation between the Minkowski coordinate v_f and the coordinate V at infinity

$$v \approx -4M \ln\left(-\frac{(v_f + 4M)}{4M}\right) \tag{3.73}$$

(b) Trajectories with $V > 0$ lie inside the horizon, and eventually fall into the singularity. We will consider out full quantum state on a spacelike hypersurface where we catch these trajectories before they reach the singularity. We need to define modes for this region of the hypersurface, but the precise choice of these modes will not be relevant. As we have noted before, there is no unique definition of particle modes if we are not in a region where spacetime is essentially flat. The spacetime inside the hole cannot be approximated by a flat metric over distances of the order of the wavelength of Hawking quanta: the wavelength of these quanta is $\sim M$, and the curvature length scale of the hole is also $\sim M$. But the absence of a unique definition of particle in the black hole interior will not be relevant to us, since we focus on the particles that have been emitted by the hole. We are interested in how these emitted particles are entangles with what is inside the hole, and this entanglement itself is independent of how we define particles inside the hole.

To carry out or computations, we make a choice of particle modes in the black hole interior in a manner that is similar to the definition of particles in the exterior. In the exterior, the natural coordinate far from the hole was v, defined through $V = -Exp[-v/4M]$. In the interior we define

$$V = e^{\frac{\tilde{v}}{4M}} \tag{3.74}$$

Thus

$$\tilde{v} = 4M \ln(V) = -v_f + 4M \ln\left(\frac{(v_f + 4M)}{4M}\right) \tag{3.75}$$

For trajectories that start very close to the horizon, we get

$$\tilde{v} \approx 4M \ln\left(\frac{(v_f + 4M)}{4M}\right) \tag{3.76}$$

3.5 Wavemodes

Let us now look at the wavemodes that we will use to define particles in various regions of our black hole spacetime.

3.5.1 The eikonal approximation

Let us now see how to construct the wavemodes we need in the eikonal approximation. Consider the modes that lie outside the horizon, and will each infinity to give the emitted Hawking particles. In this region our metric (3.22) has the form

$$ds^2 = -F(r)dudv + r^2 d\Omega^2 \tag{3.77}$$

where

$$u = t + r^*, \quad v = t - r^* \tag{3.78}$$

and

$$\frac{dr^*}{dr} = \frac{1}{1 - \frac{2M}{r}} \tag{3.79}$$

Here the variables in the metric are u, v, θ, ϕ, and r is a function $r(u, v)$.

In general a solution of the waveequation $\Box\phi = 0$ will not be purely outgoing or purely ingoing; this is because the nontrivial dependence of the metric coefficients on r reflect each of these two kinds of waves to the other. But we can imagine that if the mode had a wavelength much shorter than the curvature scale $\sim M$ of the geometry, then it would travel straight out to infinity without significant reflection. Outgoing null rays have $u = constant$, so we could try an ansatz

$$h \sim Y_{lm}(\theta, \phi) e^{-ikv} \tag{3.80}$$

where the condition of short wavelengths is

$$k \gg \frac{1}{M} \tag{3.81}$$

We will focus on spherically symmetric modes, since most of the Hawking emission is in these modes. Then the angular harmonic is $Y_{0,0} = 1$. But this ansatz needs a little correction: we would like to ensure that the inner product between two modes $(h_k, h_{k'})$ is conserved as the mode evolves. This will introduce a slowly varying prefactor in our ansatz

$$h_k = A(r) e^{-ikv} \tag{3.82}$$

Let us find $A(r)$. The waveequation is

$$\frac{1}{\sqrt{-g}}\partial_\mu\left(g^{\mu\nu}\sqrt{-g}\phi_{,nu}\right)=0 \tag{3.83}$$

We have for the metric (3.77)

$$g^{uv}=-\frac{2}{F(r)}, \quad \sqrt{-g}=\frac{1}{2}F(r)r^2\sin\theta \tag{3.84}$$

Then (3.83) gives

$$\partial_u\left(r^2A(r)(-ik)e^{-ikv}\right)+\partial_u\left(r^2A(r)_{,v}e^{-ikv}\right)+\partial_v\left(r^2A(r)_{,u}e^{-ikv}\right)=0 \tag{3.85}$$

which yields

$$-ik\partial_u\left(r^2A(r)\right)+\partial_u\left(r^2A(r)_{,v}\right)-ikr^2A(r)_{,u}+\partial_v\left(r^2A(r)_{,u}\right)=0$$

For large k, we keep only the terms containing a factor k, which gives

$$\partial_u\left(r^2A(r)\right)+r^2A(r)_{,u}=0 \tag{3.86}$$

Note that for a function $Q(r)$, we have

$$\partial_uQ(r)=Q'(r)\frac{\partial r}{\partial u}=Q'(r)\frac{dr}{dr^*}\frac{\partial r^*}{u}=\frac{1}{2}(1-\frac{2M}{r})Q'(r) \tag{3.87}$$

where we have used $r^*=\frac{1}{2}(u-v)$. Then (3.86) gives

$$A'(r)+rA(r)=0 \tag{3.88}$$

which has the solution

$$A(r)=\frac{C'}{r}=\frac{C}{(4\pi r^2)^{\frac{1}{2}}} \tag{3.89}$$

where in the second step we have rewritten the result in a way that shows the physical origin of the prefactor A: the expansion of the angular spheres leads to a drop in the amplitude of the mode as it moves outwards. Thus our modes in the region outside the hole have the form

$$h_k=\frac{C}{(4\pi r^2)^{\frac{1}{2}}}e^{-ikv} \tag{3.90}$$

As we will see below, it turns out that the norm for these modes is conserved through the evolution, not only in the large k approximation but for all k. But it should be noted that these modes are only approximate solutions to the waveequation, and need an order unity correction when $k\sim 1/M$.

The generic quanta emitted by the hole have $k\sim 1/M$, rather than $k\gg 1/m$. Why then are we using the eikonal approximation? There are two reasons:

(i) The important property of the emitted quanta is the fat that they are entangles with quanta that fall into the hole. This nature of this entanglement will be captured by the modes used in our approximation, even if the overall magnitude of entanglement is off by a factor of order unity.

(ii) The temperature T of the emitted radiation is captured exactly by the modes in the eikonal approximation, even though the *rate of radiation* is off by a factor $\Gamma[k]$ of order unity. We will see that the factors $\Gamma[k]$ are greybody factors, present even in the radiation from a normal body and dependent on the shape and size of this body. The temperature on the other hand is a more robust quantity independent of the geometry of the body, and can be read off from the detailed balance between the body and its radiation.

3.5.2 Modes in the flat spacetime region

As mentioned before, we are interested in spherically symmetric solutions. We wish to focus on modes that are outgoing near the horizon. In a spherically symmetric geometry, there are, strictly speaking, no purely outgoing or ingoing modes: an ingoing mode reaches the origin $r = 0$ and then reflects back as an outgoing mode. But we are going to use an eikonal approximation, where we look at modes that have wavelength $\lambda \ll M$ near the horizon. and just follow wavefronts as they move away from the horizon. Thus we can look at short wavelength outgoing modes when they are near the horizon. We write

$$\bar{v}_f = -4M \tag{3.91}$$

Then $v_f = \bar{v}_f$ at the horizon. Let us define

$$f_{k_a} = \frac{1}{\sqrt{2\pi}\sqrt{2k_a}}\frac{1}{4M}e^{-ik_a(v_f-\bar{v}_f)} = \frac{1}{\sqrt{2\pi}\sqrt{2k_a}}\frac{1}{4M}e^{ik_a(r-t_f-4M)} \tag{3.92}$$

Here $k_a > 0$, and we see that we have a mode where the phase fronts move radially outwards at the speed of light. The factor $1/(4M)$ arises from m

$$\frac{1}{\sqrt{4\pi r^2}} \approx \frac{1}{4\pi(2M)^2} = \frac{1}{4M} \tag{3.93}$$

and is needed to normalize the spatial integral when we compute the inner products below. We find that

$$\begin{aligned}
(f_{k_a}, f_{k'_a}) &= i\int dr(4\pi r^2)\left(f^*_{k_a}\partial_{t_f}f_{k'_a} - f^*_{k'_a}\partial_{t_f}f_{k_a}\right)\\
&\approx i\int dr(4M)^2\left(f^*_{k_a}\partial_{t_f}f_{k'_a} - f^*_{k'_a}\partial_{t_f}f_{k_a}\right)\\
&\approx \frac{(k_a+k'_a)}{(2\pi)\sqrt{4k_1k'_a}}\int dr e^{i(k_a-k'_a)(r-2M)}e^{-i(k_a-k'_a)(t_f+2M)}\\
&= \frac{(k_a+k'_a)}{(2\pi)\sqrt{4k_1k'_a}}\left(2\pi\delta(k_a-k'_a)\right)e^{-i(k_a-k'_a)(t_f+2M)}\\
&= \delta(k_a-k'_a)
\end{aligned} \tag{3.94}$$

In the third step we have made the approximation of short wavelengths. The r integral runs over $(0, \infty)$, but we are looking at highly oscillating modes near $r = 2M$ and so expect the expect the integrals over our wavepackets to have their contribution from a small range of r around $r = 2M$. Thus we can formally extend the range of the r integral to an infinite range on both sides of the horizon; this gives the delta function. Proceeding similarly, we find the full set of inner products

$$(f_{k_a}, f_{k'_a}) = \delta(k_a - k'_a), \quad (f^*_{k_a}, f^*_{k'_a}) = -\delta(k_a - k'_a), \quad (f^*_{k_a}, f_{k'_a}) = 0 \tag{3.95}$$

3.5.3 Wavemodes at infinity

Near infinity the natural time coordinate is t. We define the wavemodes

$$g_{k_b} = \frac{1}{\sqrt{2\pi}\sqrt{2k_b}} \frac{1}{(4\pi r^2)^{\frac{1}{2}}} e^{-ik_b v} = \frac{1}{\sqrt{2\pi}\sqrt{2k_b}} \frac{1}{(4\pi r^2)^{\frac{1}{2}}} e^{ik_b(r^*-t)} \tag{3.96}$$

Suppose we are looking around a point $\bar{r} \gg M$. Then the logarithm in $r^* = r + 2M\log(\frac{r}{2M} - 1)$ oscillates slowly compared to the term r, since$\partial_r \log r = 1/r$. We therefore have

$$\begin{aligned}
(g_{k_b}, g_{k'_b}) &= i\int dr(4\pi r^2)\left(g^*_{k_b}\partial_{t_f} g_{k'_b} - g^*_{k'_b}\partial_{t_f} g_{k_b}\right) \\
&= \frac{(k_b + k'_b)}{(2\pi)\sqrt{4k_b k'_b}} \int dr e^{i(k_b - k'_b)(r + 2M\log(\frac{r}{2M} - 1)} e^{-i(k_b - k'_b)t} \\
&\approx \frac{(k_a + k'_a)}{(2\pi)\sqrt{4k_b k'_b}} \int dr e^{i(k_b - k'_b)(r + 2M\log(\frac{\bar{r}}{2M} - 1)} e^{-i(k_b - k'_b)t} \\
&\approx \delta(k_b - k'_b)
\end{aligned} \tag{3.97}$$

where we have again extended the range of the r integral to an infinite one on both sides of $\bar{r}$. Thus we have

$$(g_{k_b}, g_{k'_b}) = \delta(k_b - k'_b), \quad (g^*_{k_b}, g^*_{k'_b}) = -\delta(k_b - k'_b), \quad (g^*_{k_b}, g_{k'_b}) = 0 \tag{3.98}$$

We will consider these modes in the region near the horizon when we are computing the inner product with the modes f_{k_a}. Thus let us check that the modes g_{k_b} have the correct normalization when considered near the horizon. If we follow the lines $v = constant$ back from infinity to the vicinity of $r = 2M$, we find that they lie close to but outside the horizon. Let us use the Schwarzschild coordinates here, considering the hypersurface $t = constant$ as out spacelike hypersurface. The volume element is

$$d\Sigma = 4\pi r^2 \frac{dr}{(1 - \frac{2M}{r})^{\frac{1}{2}}} \tag{3.99}$$

The normal derivative involved in the computation of the inner product is

$$\partial_n = (g_{tt})^{-\frac{1}{2}}\partial_t = \frac{1}{(1 - \frac{2M}{r})^{\frac{1}{2}}}\partial_t \tag{3.100}$$

The inner product is

$$
\begin{aligned}
(g_{k_b}, g_{k'_b}) &= i\int \frac{dr}{1-\frac{2M}{r}}(4\pi r^2)\left(g^*_{k_b}\partial_t g_{k'_b} - g^*_{k'_b}\partial_t g_{k_b}\right) \\
&= \frac{(k_b+k'_b)}{(2\pi)\sqrt{4k_bk'_b}}\int \frac{dr}{1-\frac{2M}{r}} e^{i(k_a-k'_a)r^*} e^{-i(k_b-k'_b)t} \\
&= \frac{(k_b+k'_b)}{(2\pi)\sqrt{4k_bk'_b}}\int dr^* e^{i(k_a-k'_a)r^*} e^{-i(k_b-k'_b)t} \\
&= \delta(k_b-k'_b) \qquad (3.101)
\end{aligned}
$$

which agrees with (3.97).

3.5.4 Modes inside the horizon

Inside the horizon, we can introduce Schwarzschild type coordinates (3.37), but now $\tilde{t}$ is a spacelike direction and the direction of decreasing r is the forward timelike direction. We define

$$
h_{k_c} = \frac{1}{\sqrt{2\pi}\sqrt{2k_c}}\frac{1}{(4\pi r^2)^{\frac{1}{2}}} e^{-ik_c\tilde{v}} \qquad (3.102)
$$

To compute the inner product, we take a spacelike surface $r = r_c$, with

$$
0 < r_c < 2M \qquad (3.103)
$$

The spatial volume element is

$$
d\Sigma = 4\pi r^2 dt\left(\frac{2M}{r}-1\right)^{\frac{1}{2}} \qquad (3.104)
$$

The normal derivative involved in the computation of inner products is

$$
\partial_n = -(g_{rr})^{-\frac{1}{2}}\partial_r = -\left(\frac{2M}{r}-1\right)^{\frac{1}{2}}\partial_r \qquad (3.105)
$$

The inner product is

$$
\begin{aligned}
(h_{k_c}, h_{k'_c}) &= -i\int (4\pi r^2)dt\left(\frac{2M}{r}-1\right)\left(h^*_{k_c}\partial_r h_{k'_c} - h^*_{k'_c}\partial_r h_{k_c}\right) \\
&= \frac{(k_c+k'_c)}{(2\pi)\sqrt{4k_ck'_c}}\int dt\left(\frac{2M}{r}-1\right)\frac{dr^*}{dr} e^{i(k_c-k'_c)r^*} e^{-i(k_c-k'_c)t} \\
&= \frac{(k_c+k'_c)}{(2\pi)\sqrt{4k_ck'_c}}\int dt e^{i(k_c-k'_c)r^*_c} e^{-i(k_c-k'_c)t} \\
&= \delta(k_c-k'_c) \qquad (3.106)
\end{aligned}
$$

3.5.5 The mode expansions

In the flat space region we set

$$\hat{\phi} = \int dk_a \left(f_{k_a} \hat{a}_{k_a} + f^*_{k_a} \hat{a}^\dagger_{k_a} \right) \tag{3.107}$$

On the late time slice we have a part at infinity and a part inside the hole. Thus on this slice we write

$$\hat{\phi} = \int dk_b \left(h_{k_b} \hat{b}_{k_b} + h^*_{k_b} \hat{b}^\dagger_{k_b} \right) + \int dk_c \left(\tilde{h}_{k_c} \hat{c}_{k_c} + \tilde{h}^*_{k_c} \hat{c}^\dagger_{k_c} \right) \tag{3.108}$$

Equating the two expansions, we get

$$\begin{aligned} \int dk_a \left(f_{k_a} \hat{a}_{k_a} + f^*_{k_a} \hat{a}^\dagger_{k_a} \right) \quad &= \\ \int dk_b \left(h_{k_b} \hat{b}_{k_b} + h^*_{k_b} \hat{b}^\dagger_{k_b} \right) \quad &+ \quad \int dk_c \left(\tilde{h}_{k_c} \hat{c}_{k_c} + \tilde{h}^*_{k_c} \hat{c}^\dagger_{k_c} \right) \end{aligned} \tag{3.109}$$

Since the modes h_b and $h)c$ are localized on very different parts of the late time hypersurface, they have vanishing inner products

$$(h_{k_b}, \tilde{h}_{k_c}) = 0, \quad (h^*_{k_b}, \tilde{h}_{k_c}) = 0, \quad (h_{k_b}, \tilde{h}^*_{k_c}) = 0, \quad (h^*_{k_b}, \tilde{h}^*_{k_c}) = 0 \tag{3.110}$$

We can isolate the operator $\hat{b}_{k_b}$ by computing $(h_{k_b}, \cdot)$ on both sides of (3.109)

$$\hat{b}_{k_b} = \int dk_a \left((h_{k_b}, f_{k_a}) \hat{a}_{k_a} + (h_{k_b}, f^*_{k_a}) \hat{a}^\dagger_{k_a} \right) \tag{3.111}$$

Similarly, we find

$$\hat{c}_{k_c} = \int dk_a \left((\tilde{h}_{k_c}, f_{k_a}) \hat{a}_{k_a} + (\tilde{h}_{k_c}, f^*_{k_a}) \hat{a}^\dagger_{k_a} \right) \tag{3.112}$$

$$\hat{b}^\dagger_{k_b} = \int dk_a \left((h^*_{k_b}, f_{k_a}) \hat{a}_{k_a} + (h^*_{k_b}, f^*_{k_a}) \hat{a}^\dagger_{k_a} \right) \tag{3.113}$$

$$\hat{c}^\dagger_{k_c} = \int dk_a \left((\tilde{h}^*_{k_c}, f_{k_a}) \hat{a}_{k_a} + (\tilde{h}^*_{k_c}, f^*_{k_a}) \hat{a}^\dagger_{k_a} \right) \tag{3.114}$$

3.6 Computing the inner products

3.6.1 Computing $(h_{k_b}, f^*_{k_a})$

We have

$$f^*_{k_a} \quad = \quad \frac{1}{\sqrt{2\pi}\sqrt{2k_a}} \frac{1}{(4\pi r^2)^{\frac{1}{2}}} e^{ik_a(v_f+4M)} \tag{3.115}$$

$$\begin{aligned} h^*_{k_b} &= \frac{1}{\sqrt{2\pi}\sqrt{2k_b}}\frac{1}{(4\pi r^2)^{\frac{1}{2}}}e^{ik_b v} \\ &= \frac{1}{\sqrt{2\pi}\sqrt{2k_b}}\frac{1}{(4\pi r^2)^{\frac{1}{2}}}e^{ik_b[v_f - 4M\ln\left(-\frac{(v_f+4M)}{4M}\right)]} \end{aligned} \tag{3.116}$$

We have

$$\begin{aligned} (h_{k_b}, f^*_{k_a}) &= i\int dv_f d\Omega r^2 \,\left(h^*_{k_b}\partial_{v_f} f^*_{k_a} - f^*_{k_a}\partial_{v_f} h^*_{k_b}\right) \\ &= i\int (4\pi r^2) dv_f \,\left(h^*_{k_b}\partial_{v_f} f^*_{k_a} + (\partial_{v_f} f^*_{k_a}) h^*_{k_b}\right) \\ &= i\int (8\pi r^2) dv_f \,\left(h^*_{k_b}\partial_{v_f} f^*_{k_a}\right) \end{aligned} \tag{3.117}$$

We find

$$\begin{aligned} (h_{k_b}, f^*_{k_a}) &= i\int_{-\infty}^{-4M} dv_f \frac{2}{2\pi}\frac{1}{\sqrt{4k_a k_b}} e^{ik_b[v_f - 4M\ln\left(-\frac{(v_f+4M)}{4M}\right)]}(ik_a)e^{ik_a(v_f+4M)} \\ &= -\int_{-\infty}^{-4M} dv_f \frac{1}{2\pi}\sqrt{\frac{k_a}{k_b}} e^{ik_b[v_f - 4M\ln\left(-\frac{(v_f+4M)}{4M}\right)]} e^{ik_a(v_f+4M)} \end{aligned} \tag{3.118}$$

Let us define

$$X = -(v_f + 4M) \tag{3.119}$$

and write $k_b \equiv k$ Then we find

$$(h_k, f^*_{k_a}) = -\int_{X=0}^{\infty} dX \frac{1}{2\pi}\sqrt{\frac{k_a}{k}} e^{ik[-X-4M-4M\ln\left(\frac{X}{4M}\right)]} e^{-ik_a X} \tag{3.120}$$

3.6.2 Computing $(\tilde{h}^*_{k_c}, f^*_{k_a})$

We have

$$\begin{aligned} \tilde{h}_{k_c} &= \frac{1}{\sqrt{2\pi}\sqrt{2k_c}}\frac{1}{(4\pi r^2)^{\frac{1}{2}}}e^{-ik_c\tilde{v}} \\ &= \frac{1}{\sqrt{2\pi}\sqrt{2k_c}}\frac{1}{(4\pi r^2)^{\frac{1}{2}}}e^{-ik_c[-v_f + 4M\ln\left(\frac{(v_f+4M)}{4M}\right)]} \end{aligned} \tag{3.121}$$

$$\begin{aligned} (\tilde{h}^*_{k_c}, f^*_{k_a}) &= i\int dv_f d\Omega r^2 \,\left(\tilde{h}_{k_c}\partial_{v_f} f^*_{k_a} - f^*_{k_a}\partial_{v_f}\tilde{h}_{k_c}\right) \\ &= i\int (8\pi r^2) dv_f \,\left(h_{k_c}\partial_{v_f} f^*_{k_a}\right) \end{aligned} \tag{3.122}$$

We find

$$
\begin{aligned}
(\tilde{h}^*_{k_c}, f^*_{k_a}) &= i\int_{-4M}^{\infty} dv_f \frac{2}{2\pi}\frac{1}{\sqrt{4k_a k_c}} e^{-ik_c[-v_f+4M\ln\left(\frac{(v_f+4M)}{4M}\right)]}(ik_a)e^{ik_a(v_f+4M)} \\
&= -\int_{-4M}^{\infty} dv_f \frac{1}{2\pi}\sqrt{\frac{k_a}{k_c}} e^{-ik_c[-v_f+4M\ln\left(\frac{(v_f+4M)}{4M}\right)]} e^{ik_a(v_f+4M)}
\end{aligned}
\tag{3.123}
$$

Let us define

$$X = (v_f + 4M) \tag{3.124}$$

and write $k_c \equiv k$. Then we find

$$(\tilde{h}^*_k, f^*_{k_a}) = -\int_0^{\infty} dX \frac{1}{2\pi}\sqrt{\frac{k_a}{k}} e^{-ik[-X+4M+4M\ln\left(\frac{X}{4M}\right)]} e^{ik_a X}$$

3.6.3 Relating the two computations

We go to a new contour, where we have new values of X. Thus we get

$$
\begin{aligned}
(\tilde{h}_k, f^*_{k_a}) &= -\int_{X=0}^{\infty} dX \frac{1}{2\pi}\sqrt{\frac{k_a}{k}} e^{-ik[-X+4M+4M\ln\left(\frac{X}{4M}\right)]} e^{ik_a X} \\
&= -\int_{X=0}^{-\infty} dX \frac{1}{2\pi}\sqrt{\frac{k_a}{k}} e^{-ik[-X+4M+4M\ln\left(\frac{X}{4M}\right)]} e^{ik_a X} \\
&= -\int_{X'=0}^{\infty} (-dX') \frac{1}{2\pi}\sqrt{\frac{k_a}{k}} e^{-ik[X'+4M+4M\ln\left(\frac{X'}{4M}\right)+4M(i\pi)]} e^{-ik_a X'} \\
&= e^{4\pi Mk}\int_{X'=0}^{\infty} dX' \frac{1}{2\pi}\sqrt{\frac{k_a}{k}} e^{-ik[X'+4M+4M\ln\left(\frac{X'}{4M}\right)]} e^{-ik_a X'}
\end{aligned}
\tag{3.125}
$$

where in the last step we have written $X' = -X$.

Thus we find that

$$(\tilde{h}_k, f^*_{k_a}) = -e^{4\pi Mk}(h_k, f^*_{k_a}) \tag{3.126}$$

Thus the combination

$$\hat{b}_k + e^{-4\pi Mk}\hat{c}^\dagger_k \tag{3.127}$$

has only annihilation operators $\hat{a}^\dagger_{k_a}$ and no creation operators $\hat{a}^\dagger_{k_a}$. This this combination annihilates the vacuum $|0\rangle_a$

$$\left(\hat{b}_k + e^{-4\pi Mk}\hat{c}^\dagger_k\right)|0\rangle_a = 0 \tag{3.128}$$

3.7 Correlating wavepackets

The expression (**??**) shows that modes $\hat{b}_k$ near in finity are entangled with modes $\hat{c}_k$ inside the hole. We see a correlation in fourier space labelled by k: the excitation level of fourier mode k outside the hole is correlated with the excitation of the same mode k inside the hole. A fourier mode however describes a function that is delocaiized in space: for example, a wavefunction $Exp[ikr]$ extends over all values of r. By contrast, in our earlier heuristic picture of pair creation we had visualised particles that were rather loalized in positon space: when an electron emerged from the hole, a positron fell in, and if at a later time a positron emerged from the hole, then an electron would fall in at that later time. Thus we would like to rewrite our state (**??**) in a way that exhibits some locality, so that we can see particles being emitted with reasonably well locaized wavefunctions, and also see their infalling partners having reasonaby well localized wavefunctions at correponding positions inside the horizon.

We proceed as follows:

(i) We wish to make wavepackets that are reasonably localized in position space. This means that we will need to have a nonzero spread in the corresponding momentum k. Thus let us start with k space. Our wavefunctions have the form $\sim Exp[-ikv]$ with $k > 0$. Thus we take the interval $0 < k < \infty$ and divide this into intervals of length ϵ. Let the nth interval extend over

$$n\epsilon \le k \le (n+1)\epsilon, \qquad -0 <\le n < \infty \tag{3.129}$$

In this interval, we take a complete basis of functions of k:

$$\begin{aligned} \hat{H}_{n,s}(k) &= \frac{1}{\sqrt{\epsilon}} e^{\frac{isk}{\epsilon}}, \quad n\epsilon \le k \le (n+1)\epsilon \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{3.130}$$

where $-\infty < s < \infty$ is an integer. The functions $\hat{H}_{n,s}(k)$ are orthonormal

$$\int dk \hat{H}^*_{n,s}(k)\hat{H}_{n',s'}(k) = \delta_{n,n'}\delta_{s,s'} \tag{3.131}$$

and give a complete set of functions on the like $-\infty < k < \infty$. Our position space wavepackets will be defined as the fourier transforms of the $\hat{H}_{n,s}$. Since fourier transformation can be thought of as just a change of basis for the functions, the orthonormality of $\hat{H}_{n,s}$ implies that our position space wavepackets will be also be automatically orthonormal in the norm $\int dx|H(x)|^2$.

(ii) Define the position space wavepackets

$$H_{n,s}(v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk e^{-ikv} \hat{H}_{n,s} \tag{3.132}$$

Since $\hat{H}_{n,s}(k)$ is localized near wavenumbers $k \approx n\epsilon$, we see that $H_{n,s}(v)$ will be reasonably localized in momentum space: $\Delta k \sim \epsilon$. To see where $H_{n,s}(v)$ is localized in position space v, note that $\hat{H}_{n,s}(k)$ is oscillating in k as $e^{isk/\epsilon}$. Thus the k integral in (3.135) will damp out the integral unless this phase oscillation is cancelled by the oscillation from the factor e^{-ikv}. Thus the wavepacket will peak around

$$v_{n,s} \approx \frac{s}{\epsilon} \tag{3.133}$$

The spatial width of the wavepacket is

$$\Delta v \sim \frac{1}{\Delta k} \sim \frac{1}{\epsilon} \tag{3.134}$$

Thus we have wavepackets of length $\sim 1/\epsilon$, peaked at locations that are separated by a distance $\sim 1/\epsilon$.

Let us now compute the wavepackets (3.135) explicitly

$$\begin{aligned} H_{n,s}(v) &= \frac{1}{\sqrt{2\pi}} \int_{n\epsilon}^{(n+1)\epsilon} dk \hat{e}^{-ikv} \frac{1}{\sqrt{\epsilon}} e^{\frac{isk}{\epsilon}} \\ &= \frac{1}{\sqrt{2\pi}\sqrt{\epsilon}} \frac{e^{i(n+1)(s-v\epsilon)} - e^{in(s-v\epsilon)}}{i(\frac{s}{\epsilon} - v)} \\ &= \sqrt{\frac{\epsilon}{2\pi}}\; e^{i(n+\frac{1}{2})\epsilon(\frac{s}{\epsilon}-v)}\; \frac{\sin[\frac{1}{2}(s-v\epsilon)]}{[\frac{1}{2}(s-v\epsilon)]} \end{aligned} \tag{3.135}$$

Thus we have a wavefunction oscillating in v with a wavenumber $\approx (n+\frac{1}{2})\epsilon$, modulated by a facto of the form $\sin x/x$, which peaks at $x = 0$. This peak is at $v = s/\epsilon$, as anticipated above.

Note the reverse transform (3.135) is

$$\hat{H}_{n,s}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dv e^{ikv} H_{n,s} \tag{3.136}$$

(iii) Let us now turn to the expansion of the field operator $\hat{\phi}$:

$$\hat{\phi} = \int dk \left(\hat{b}_k h_k(v) + \hat{b}_k^\dagger h^*(v) \right) \tag{3.137}$$

The functions $h(v)$ are normalized not in the L^2 norm, but with the inner product (**??**). Thus these functions are obtained by starting with the functions

$$q_k(v) \equiv \frac{1}{\sqrt{2\pi}} e^{-ikv} \tag{3.138}$$

which are normalized as

$$\int dv q_k^*(v) q_{k'}(v) = \delta(k-k') \tag{3.139}$$

and adding a factor $1/\sqrt{2k}$

$$h_k(v) = \frac{1}{\sqrt{2k}} q_k(v) \tag{3.140}$$

We can similarly make mode functions based on the wavepackets $H_{n,s}(v)$ which are normalized in the L^2 norm (3.131). In fourier space these functions were peaked around $k = n\epsilon$ with a small width ϵ. Thus we can take

$$h_{n,s}(v) \approx \frac{1}{\sqrt{2n\epsilon}} H_{n,s}(v) \tag{3.141}$$

as mode functions that are orthonormal in the inner product (**??**). Note that we have used the same symbol h for these wavemodes, and will distinguish them from the $h_k(v)$ by the extra index they carry. We can write the field operator using these modes

$$\hat{\phi} = \sum_{n,s} \left(\hat{b}_{n,s} h_{n,s}(v) + \hat{b}^\dagger_{n,s} h^*_{n,s}(v) \right) \tag{3.142}$$

Equating (3.137) and (3.142)

$$\int dk \left(\hat{b}_k h_k(v) + \hat{b}^\dagger_k h^*(v) \right) = \sum_{n,s} \left(\hat{b}_{n,s} h_{n,s}(v) + \hat{b}^\dagger_{n,s} h^*_{n,s}(v) \right) \tag{3.143}$$

Taking $(h^*_k, \cdot)$ on both sides, and noting that $(h^*_k, h^*_{k'}) = -\delta(k - k')$, we have

$$\hat{b}^\dagger_k = -\sum_{n,s} \left(\hat{b}_{n,s} (h^*_k, h_{n,s}) + \hat{b}^\dagger_{n,s} (h^*_k, h^*_{n,s}) \right) \tag{3.144}$$

We have

$$\begin{aligned}
(h^*_k, h^*_{n,s}) &= i \int dv \left(h_k(v) \partial_v h^*_{n,s}(v) - (\partial_v h_k(v)) h^*_{n,s}(v) \right) \\
&= -2i \int dv (\partial_v h_k(v)) h^*_{n,s}(v) \\
&= -2i \int dv \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2k}} (-ik) e^{-ikv} \frac{1}{\sqrt{2n\epsilon}} H^*_{n,s}(v) \\
&= -\int dv \frac{1}{\sqrt{2\pi}} \sqrt{\frac{k}{n\epsilon}} e^{-ikv} H^*_{n,s}(v) \\
&= -\sqrt{\frac{k}{n\epsilon}} \hat{H}^*_{n,s}(k) \\
&\approx -\hat{H}^*_{n,s}(k)
\end{aligned} \tag{3.145}$$

where in the first step we have done an integration by parts on the first term in the inner product, and in the last step we have noted that $\hat{H}^*_{n,s}(k)$ is peaked sharply around $k \approx n\epsilon$. A similar computation gives $(h_k^*, h_{n,s}) \approx 0$, since $\hat{H}_{n,s}$ vanishes for wavenumbers $k < 0$. Thus we have

$$\hat{b}_k^\dagger \approx \sum_{n,s} \hat{H}_{n,s}^*(k)\hat{b}_{n,s}^\dagger \tag{3.146}$$

(iv) Now consider the operators $\hat{c}_k^\dagger$, Let us *define* a new basis of operators $\hat{c}_{n,s}^\dagger$ through

$$\hat{c}_k^\dagger = \sum_{n,s} \hat{H}_{n,s}(k)\hat{c}_{n,s}^\dagger \tag{3.147}$$

The purpose of this definition can be seen by looking at the exponent in (**??**)

$$\int_0^\infty dk e^{-4\pi Mk}\hat{b}_k^\dagger\hat{c}_k^\dagger = \int_0^\infty dk e^{-4\pi Mk}\left(\sum_{n,s}\hat{H}_{n,s}^*(k)\hat{b}_{n,s}^\dagger\right)\left(\sum_{n',s'}\hat{H}_{n',s'}(k)\hat{c}_{n',s'}^\dagger\right) \tag{3.148}$$

The product $\hat{H}_{n,s}^*(k)\hat{H}_{n',s'}(k)$ vanishes unless $n = n'$, and when $n = n'$ then it has support over a narrow range of k with $k \approx n\epsilon$. Thus the factor $e^{-4\pi Mk}$ is almost constant over this range of km and we can take it out of the integral:

$$\begin{aligned}\int_0^\infty dk e^{-4\pi Mk}\hat{b}_k^\dagger\hat{c}_k^\dagger &\approx e^{-4\pi Mn\epsilon}\int_0^\infty dk\left(\sum_{n,s}\hat{H}_{n,s}^*(k)\hat{b}_{n,s}^\dagger\right)\left(\sum_{n',s'}\hat{H}_{n',s'}(k)\hat{c}_{n',s'}^\dagger\right)\\ &= \sum_{n,s}\sum_{n',s'} e^{-4\pi Mn\epsilon}\hat{b}_{n,s}^\dagger\hat{c}_{n',s'}^\dagger\delta_{n,n'}\delta_{s,s'}\\ &= \sum_{n,s} e^{-4\pi Mn\epsilon}\hat{b}_{n,s}^\dagger\hat{c}_{n,s}^\dagger\end{aligned} \tag{3.149}$$

where we have used (3.131). Thus the state created by Hawking evaporation can be written as one where the excitations at infinity in the wavepackets $h_{n,s}$ are correlated with corresponding wavepackets in the interior of the hole.

(v) Finally we consider the nature of the wavepackets for $\hat{b}_{n,s}$ and $\hat{c}_{n.s}^\dagger$. We have written the wavepackets for $\hat{b}^\dagger$ as functions of v, but we can consider them on a spacelike hypersurface $t = t_0$. Since $v = t - r^*$, we see from (3.133) that the wavepacket for $\hat{b}_{n,s}^\dagger$ is localized around

$$r^* = t - \frac{s}{\epsilon} \tag{3.150}$$

Thus the modes for larger integers s are localized closer to the hole, indicating that they are emitted later.

Now consider the modes for $\hat{c}_{n,s}^\dagger$. The field operator inside the hole is expanded as

$$\hat{\phi} = \int dk\left(\hat{c}_k\tilde{h}_k(\tilde{v}) + \hat{c}_k^\dagger\tilde{h}_k^*(\tilde{v})\right) \tag{3.151}$$

With the change of basis (3.147) we have

$$\begin{aligned}\hat{\phi} &= \int dk \sum_{n,s} \Big(\hat{c}_{n,s} \hat{H}^*_{n,s}(k) \tilde{h}_k(\tilde{v}) + \hat{c}^\dagger_{n,s} \hat{H}_{n,s}(k) \tilde{h}^*_k(\tilde{v}) \Big) \\ &\equiv \sum_{n,s} \Big(\hat{c}_{n,s} \tilde{h}_{n,s}(v) + \hat{c}^\dagger_{n,s} \tilde{h}^*_{n,s}(v) \Big) \end{aligned} \tag{3.152}$$

where the mode functions are now

$$\begin{aligned}\tilde{h}_{n,s}(v) &= \int dk \hat{H}^*_{n,s}(k) \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2k}} e^{-ik\tilde{v}} \\ &= \int_{n\epsilon}^{(n+1)\epsilon} dk \frac{1}{\sqrt{\epsilon}} e^{\frac{isk}{\epsilon}} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2k}} e^{-ik\tilde{v}} \end{aligned} \tag{3.153}$$

The integral over k peaks at the location where the coefficient of k in the exponent vanishes:

$$\frac{s}{\epsilon} - \tilde{v} = \frac{s}{\epsilon} - \tilde{r}^* + \tilde{t} = 0 \tag{3.154}$$

We look at these modes on a surface $\tilde{r}^* = \tilde{r}^*_c$. Then we see that larger values of s correspond to later values of $\tilde{t}$.

Putting this together with the behavior of the $\hat{b}$ modes, we get the following picture. Larger values of s correspond to later emissions. The corresponding particles outside the hole are closer to the hole on a slice $t = t_0$, and the corresponding particles inside the hole are at later times $\tilde{t}$ on the slice $\tilde{r}^* = \tilde{r}^*_c$.

$$|0\rangle_a = \prod_{k>0,s} e^{-e^{-4\pi M k} \hat{B}^\dagger_{k,s} \hat{C}^\dagger_{k,s}} |0\rangle_b \otimes |0\rangle_c \tag{3.155}$$

Bibliography