Rep. Prog. Phys., Vol. 41, 1978. Printed in Great Britain

Thermodynamics of black holes

P C W DAVIES

Department of Mathematics, King's College, Strand, London WC2, UK

Abstract

In spite of the fundamental difficulties associated with the thermodynamics of self-gravitating systems, the objects known as black holes appear to conform to a very straightforward generalisation of standard laboratory thermodynamics. In this review the generalised theory is examined in detail. It is shown how familiar concepts such as temperature, entropy, specific heats, phase transitions and irreversibility apply to systems containing black holes, and some concrete results of the theory are presented. The thermodynamic connection is based on Hawking's celebrated application of quantum theory to black holes, and in this review the quantum aspects are described in detail from several standpoints, both heuristic and otherwise. The precise mechanism by which the black hole produces thermal radiation, its nature and origin, and the energetics of back-reaction on the hole are reviewed. The thermal states of quantum holes are also treated using the theory of thermal Green functions, and the entropy of the hole is shown to be related to the loss of information about the quantum states hidden behind the event horizon. Some related topics such as accelerated mirrors and observers in Minkowski space, super-radiance from rotating holes and the thermodynamics of general self-gravitating systems are also briefly discussed.

This review was received in January 1978.

0034-4885/78/0008–1313 $05.00

Contents

1. Gravity and thermodynamics

In laboratory physics, the topics of thermodynamics and gravitation lead a rather separate existence. Astronomers, however, have long had to contend with the conjunction of the two. In the broadest sense, thermodynamics regulates the *organisation* of activity in the universe, and gravity controls the *dynamics*, at least on the large scale. The interaction between these conceptually dissimilar aspects of fundamental physics is a grey area full of paradoxes, muddle and uncharted hazards.

One of the central difficulties about the thermodynamics of gravitating systems is the apparent absence of true *equilibrium*. This problem, which can lead directly to 'peculiar' effects, has long been known to astronomers. Stars are hot, self-gravitating balls of gas inside which the weight of the star is supported by its own internal kinetic or zero-point quantum pressure. Unlike ordinary laboratory thermodynamic systems, a star is made hotter, not by adding energy, but by removing it. If the Sun were to suddenly lose all its heat energy, it would rapidly shrink to a fraction of its present size and, after a period of oscillation, take up a new condition at a much higher temperature. The ever-present threat of gravitational shrinkage under its own weight makes every star potentially unstable against catastrophe. We can think of stars like the Sun as *metastable*—a temporary interlude between a distended cloud of gas and totally imploded matter.

1.1. Black holes

The key to understanding black holes, and especially their connection with thermodynamics, is to appreciate the meaning of the so-called *event horizon*. Consider a very compact and massive star. The strength of gravity at its surface can be increased either if the star shrinks, or if more mass is added. According to the general theory of relativity, gravity affects the properties of light, and this is manifested in the behaviour of light rays which leave the surface of the star travelling radially outwards. Because the light has to 'do work' to overcome the surface gravity and escape from the star, its energy, and hence frequency, will be somewhat diminished. This famous gravitational red shift has been measured in light leaving such relatively low-gravity objects as the Sun, and even the Earth. For more compact and massive objects the red shift can become enormous. Even on Newtonian grounds it is clear (as was pointed out by Pierre Laplace as long ago as 1798) that when the escape velocity from the surface of a star exceeds that of light, something odd must happen. According to both Newtonian gravity and relativity, this turn of events comes about for a spherical, uncharged star if the radius of the star shrinks below $2GM/c^2$, where M is the mass, c is the speed of light, and G is Newton's gravitational constant. This size is very small, being of the order of 1 km for the Sun and 1 cm for the Earth. Ordinary stars are enormously more distended than this, although neutron stars closely approach the critical size.

A straightforward calculation shows that as the star approaches the critical radius, the light from its surface becomes redshifted without limit, so that it can no longer be seen; it is black. This phenomenon can be viewed in a number of ways. One can think of the light struggling to escape against the intense gravity so that it takes longer and longer to travel to infinity. Care is needed here though; *locally* an inertial observer will always measure the light to be travelling at the usual speed, but far away from the star it *appears* to move slower. Perhaps it is better to think of time near the star's surface as being greatly dilated, so that the reduced frequency of the light can be regarded as a sort of slowing of events or clock rates. Either way, it is clear that as the star approaches the critical radius $2GM/c^2$ nothing at all can be seen of the events which occur on (and in) the star thereafter. The spherical surface $r=2GM/c^2$ therefore acts as an event *horizon*, separating the events which can be viewed from a great distance, from those which cannot, however long one waits.

Studies of stellar structure indicate that many stars may not be composed of sufficiently stiff material to prevent them shrinking to this critical radius, purely under the gravity of their own contents. If a star cannot withstand its own weight, it implodes catastrophically and travels through the critical radius in a very short time as measured in its own (falling) time scale. Of course, as viewed from infinity, this process becomes infinitely dilated, and as the star approaches the horizon it appears to be 'frozen'. At this stage, the red shift increases exponentially on an e-folding time that is typically 10^{-5} s. Within the briefest moment the star effectively disappears from the universe.

A useful picture of the event horizon can be obtained by imagining spherical wavefronts of light which are emitted radially outwards from different surfaces $r=$ constant. Those spheres travelling from $r>2GM/c^2$ gradually expand and eventually escape to infinity, but those emitted *inside* the critical radius actually *shrink* towards the centre, even though they are emitted in a direction away from the centre. Crudely speaking, the gravity there is so strong that it drags the light backwards. The event horizon is the spherical surface of light that just escapes to infinity after an infinite duration. From afar, it appears to hover, static, at $r=2GM/c^2$.

According to relativity, matter and information cannot propagate faster than light, so if light cannot escape from inside the horizon, neither can anything else. Thus, once the star has retreated through this surface it can never return to the outside universe, or signal its fate either. We can use the theory of relativity to conjecture the likely fate of the star, though this does not have any direct connection with the thermodynamic properties. As the surface of the star is composed of ordinary matter, it must move along a time-like trajectory, sandwiched between inward directed and outward directed light surfaces. As both these surfaces shrink in radius, it follows that the surface of the star must shrink also. No force, no pressure, however powerful, can cause it to remain static at a fixed radius $r<2GM/c^2$. The shrinkage is inexorable and rapid. Within a fraction of a second, star-time, the surface apparently disappears into a single point and the density of the star becomes infinite. This state of affairs cannot be taken too seriously, because the space–time curvature also rises without apparent limit, and space–time would become smashed, invalidating all our theories of physics. In some sense, not yet understood, the singularity represents a type of boundary or edge to space–time as at present conceived. Some very powerful theorems due mainly to Hawking and Penrose (Hawking and Ellis 1973) prove that under a very wide range of likely circumstances (not just spherical collapse) space–time singularities will form, although not all of the imploding star need crash into them.

The region inside the horizon, once the star has shrunk away to nothing, is empty and, from the exterior universe, black and inaccessible. It is therefore called a *black hole*. Astronomers widely believe that black holes will form as the natural end state of the evolution of massive stars, but so far there is no direct observational evidence of their existence.

The most general known solutions to Einstein's field equations of general relativity which contain black holes are the so-called Kerr–Newman family, which describe an axisymmetric, matter-free space–time representing a black hole which rotates and carries an electric charge. These solutions form a three-parameter set, labelled by the total mass–energy M, the angular momentum $J = |\boldsymbol{J}|$ and the electric charge e. For $J = 0$ the hole is spherical and, if $e = 0$ the solutions reduce to the case discussed above, originally discovered by Karl Schwarzschild in 1916, where the horizon is located at $2GM/c^2$ (often called the Schwarzschild radius for this reason). The shape of the horizon and the behaviour of light rays in the general case is more complicated, but the essential qualitative features are the same as for the Schwarzschild black hole.

Because the interior of the hole is invisible and inaccessible from the exterior universe, we cannot tell the difference between two holes with the same M, J and e. Whether the imploding star is made of antimatter, neutrinos, pions or green cheese, the end-state black hole will look the same. Only these three global parameters (which can all be measured by performing experiments far from the hole) have physical significance. Thus a given global, or macro, state (M, J, e) can be realised by an enormous number of internal microstates. This at once suggests that black holes have a very high entropy, and represent in some sense the maximum entropy, *equilibrium end state* of gravitational collapse.

It is instructive to consider the approach to equilibrium as the star implodes, to see how the information about the internal microstates is wiped out by the collapse. Only a qualitative understanding of this phase exists, and nothing like the equivalent of a Boltzmann equation or H theorem to describe the irreversible progress towards equilibrium.

Suppose the imploding star has a lot of microstructure—temperature gradients, an electric charge and current distribution, density perturbations away from exact spherical symmetry, etc. During the brief period in which it implodes towards the horizon, the distant observer sees the collapse rapidly slow to a halt and fade out as the red shift escalates exponentially. If, for example, one wishes to measure the temperature difference between two neighbouring patches of the surface using some kind of bolometer, then as the radiation shifts towards the red the temperature difference also diminishes as the two patches rapidly appear to approach zero temperature together. If one measures the surrounding pattern of electric fields to determine the charge difference between neighbouring regions then the effect of the growing space–time curvature is to bend the lines of force so that they appear to approach a purely radial configuration corresponding to a *uniform* charge distribution. In a sense, therefore, the effect of the collapse is to impose a type of *coarse-graining* as far as measurements from a distance are concerned. In the late stages of the collapse, as equilibrium is approached exponentially, all information about temperature and charge distributions disappears. Similarly, it can be shown that density perturbations, baryon and lepton number, and other parameters which might characterise the internal states of the star, become unmeasurable from outside the black hole. (For a summary of these so-called 'no-hair' theorems see Misner *et al* 1973, chap 33.)

In summary, the collapse of a star to a black hole is rather like the irreversible degeneration of information and organisation in a gas. In the latter case macroscopic information is destroyed by molecular collisions and it becomes inaccessible because it has been distributed among the microscopic degrees of freedom which we cannot resolve, whereas in the former case the information is destroyed by the gravitational field (i.e. the space–time structure) and it becomes inaccessible because of the event horizon. In both cases the physical appearance of the system changes from an ordered, structured state to a few-parameter disordered state which contains no *observable* memory of the initial system. Clearly, the bigger the hole, the more information it has wiped out and the greater the number of internal microstates which can produce it. Hence a measure of the entropy ought to be provided by the *size* of the hole.

1.2. The four laws of classical black hole 'thermodynamics'

Curious analogues between the behaviour of black holes and that of thermodynamic equilibrium systems were noted some time ago (for a review see Carter 1973). Crudely speaking, because gravity always attracts, there is a general tendency for self-gravitating systems to grow rather than shrink. In the black hole case, the inability for light to emerge from inside the event horizon precludes the escape of any material, so the horizon acts as a sort of asymmetric one-way surface: things can fall in and make the hole bigger but not come out and make it smaller. This is reminiscent of the second law of thermodynamics, in which there is an asymmetric tendency for a one-way increase in entropy. The size of the black hole is analogous to the entropy. It keeps on increasing.

This analogy is almost trivial for a spherical, electrically neutral (Schwarzschild) black hole. In the more general case of black holes that possess angular momentum J and electric charge e, the size of the black hole depends both on J and e in a rather complicated way. If the total surface area of the horizon is used as a measure of size then this is given by the formula (Smarr 1973):

$$A=4\pi\left[2M^2-e^2+2M^2\left(1-\frac{e^2}{M^2}-\frac{J^2}{M^4}\right)^{1/2}\right] \tag{1.1}$$

where $e^2<M^2$ and $J^2<M^4$ (throughout, units with $G=c=1$ will be used) so it is not clear at a glance whether a disturbance to the black hole which changes both e and J, as well as M, will always increase the total area.

In fact, as shown in a theorem by Hawking (1972) the horizon area cannot decrease in any process, even for these more general black holes, so long as locally negative energy (which gravitates repulsively) is not involved. One famous example due to Penrose (1969) concerns a method for extracting mass–energy from a rotating black hole. The mechanism consists of propelling a small body into the region just outside the event horizon where (due to a dragging effect on the space surrounding the black hole caused by its rotation) some particle trajectories possess negative energy relative to infinity. If the body is exploded into two fragments, one of which is placed on one of these negative-energy paths, and this part disappears down the hole, it will reduce the total mass M of the hole somewhat and the mass–energy thereby released by this sacrificed component appears in the remaining fragment which is ejected to infinity at high speed. During this energy transfer the black hole's rotation rate is diminished somewhat, so J also decreases. Inspection of (1.1) shows that when J decreases, the area *increases*, but when M decreases, the area decreases. The changes in M and J are therefore in competition, but a careful calculation shows that J always wins and the area increases.

Actually, if the class of all trajectories is studied, it is found that in general the area increases by an amount corresponding to a considerable fraction of the mass of the propelled body. However, the efficiency of energy extraction can be improved by approaching closely to a limiting class of trajectories for which the event horizon area remains constant. The limiting case is therefore reversible and corresponds to an *isentropic* change in thermodynamics. In practice, 100% efficiency (complete reversibility) would be impossible.

This strong analogy between event horizon area and entropy led to the use of the name 'second law' in connection with Hawking's area theorem, which is therefore written thus:

$$\mathrm{d}A \geqslant 0 \tag{1.2}$$

(equality corresponding to reversibility).

There are also analogues of the zeroth, first and third laws of thermodynamics. From (1.1) we can obtain:

$$\mathrm{d}M = (8\pi)^{-1}\kappa \, \mathrm{d}A + \Omega \, \mathrm{d}J + \Phi \, \mathrm{d}e \tag{1.3}$$

where $(8\pi)^{-1}\kappa \equiv \partial M/\partial A$, etc, which is really just an expression of mass–energy conservation and corresponds to the first law. If A plays the role of entropy then we see from the structure of (1.3) that κ plays the role of *temperature* ($\kappa \, \mathrm{d}A \sim T \, \mathrm{d}S$). The interesting thing is that κ can be shown to be *constant* across the event horizon surface. We thus have an expression of a 'zeroth' law, analogous to the thermodynamic one which says that in thermodynamic equilibrium there exists a common temperature parameter for the whole system. The quantity κ is known as the surface gravity of the black hole. Its precise definition need not concern us here, but its significance lies in the fact that it determines the e-folding time which controls the rate at which the collapsing star red shifts and approaches equilibrium. For a Schwarzschild hole, $\kappa = (4M)^{-1}$ and the constant 8π in (1.3) has been chosen to agree with this. The remaining terms in (1.3) simply describe the work done (energy extracted) from changes in angular momentum ($\Omega \, \mathrm{d}J$) and electric charge ($\Phi \, \mathrm{d}e$), and have a very obvious structure: Ω is the (magnitude of) angular velocity and Φ the electric potential at the event horizon.

Finally, there is the third law. It is straightforward to show that if J^2 or e^2 become large enough such that:

$$\frac{J^2}{M^4} + \frac{e^2}{M^2} = 1 \tag{1.4}$$

then κ vanishes (although A does not). This corresponds to absolute zero (though with finite entropy). A black hole with parameters given by (1.4) is known as an extreme Kerr–Newman black hole. It is the limiting case of an object which still possesses an event horizon. Should the left-hand side become even infinitesimally greater than one, then the horizon would disappear and we would be left with a *naked singularity*, i.e. the singularity would no longer be invisible inside a black hole but would be able to influence, and be observed by, the outside universe. This circumstance is considered so undesirable for physics that most physicists believe in the so-called cosmic censorship hypothesis due to Penrose (1969): naked singularities cannot form from gravitational collapse. Cosmic censorship implies the unattainability of 'absolute zero', $\kappa = 0$ (i.e. condition (1.4) for an extreme black hole), so it plays the role of the third law.

1.3. Entropy, information and event horizons

In spite of the compelling similarities between the four laws of thermodynamics and those of black holes, it would seem that we are unable to *identify* the laws for the following reason. A black hole is supposed to be *black*, i.e. have zero temperature. This clearly makes nonsense of any attempt to regard the surface gravity κ as a real temperature. Indeed, in the simplest case of a Schwarzschild black hole ($J=e=0$), we obtain from (1.1):

$$A=16\pi M^2 \tag{1.5}$$

$$\kappa\equiv 8\pi\frac{\partial M}{\partial A}=\frac{1}{4M}. \tag{1.6}$$

Moreover, if we attempt to employ the well-known thermodynamic relation (obtained from integrating $\mathrm{d}M=\kappa\ \mathrm{d}A=T\ \mathrm{d}S$):

$$\text{energy}=2\ \text{entropy}\times\text{temperature}\dagger \tag{1.7}$$

to the black hole, then as the energy M is finite, zero temperature would imply *infinite entropy*. In contrast, if we rewrite (1.5) in the form (1.7):

$$M=A\kappa/4\pi \tag{1.8}$$

we observe that the right-hand side is the product of two *finite* quantities.

How can we understand the notion of the entropy of a black hole and why should it be infinite? Great insight is provided by using the relation between entropy and *information* (see, for example, Davies 1974). If a system is highly ordered then its entropy is low. Such a system requires a great deal of information to describe it or, alternatively, we can say that it has a high information content. This leads to the identification:

$$\text{information}\leftrightarrow\text{negative entropy}.$$

When a system becomes disordered, its entropy goes up and its state requires less information to describe it. In thermodynamic equilibrium, only a very small number of parameters (e.g. overall temperature, volume, number of particle species) are needed; this is the state of maximum entropy and minimum information content.

The relevance of information to the discussion of black hole entropy concerns the nature of the event horizon. As discussed in §1.1 it might be expected on general grounds that a black hole would possess entropy—and rather a lot of it—by virtue of all the information it has swallowed up.

How can we estimate how much information has gone to make up a black hole of a given size? One crude estimate (Hawking 1976) is to count the number of internal degrees of freedom, assigning one bit of information to each. This means calculating how many particles go into the production of a black hole of, say, mass M. On the basis of classical physics this number can be arbitrarily large, because we can always choose particles of arbitrarily small mass: for example, zero rest mass particles (photons, neutrinos) with arbitrarily low energy. It therefore seems that the information content, and hence the entropy, associated with the black hole should indeed be unbounded, as we suspected.

It is tempting to regard the unbounded entropy as in some way connected with the inherent instability of self-gravitating matter against total collapse. Another system unstable against total collapse is the classical atom, wherein the orbiting electron can radiate unlimited entropy and spiral catastrophically and without limit indefinitely

† The factor 2 arises because A is proportional to the *square* of M.

close to a point nucleus. When quantum theory is applied to this system, the entropy becomes bounded and the atom requires a stable ground state.

If we take into account the quantum nature of matter we make the great discovery that the entropy of a black hole should also be finite. The reason is simple: we cannot choose particles of *arbitrarily* small mass to constitute a black hole because of the quantum relation between energy and wavelength:

$$E = h/\lambda.$$

At the very least the wavelength λ must be less than the size of the black hole if we are to regard the particle of energy E as being located inside the hole. Choosing $\lambda \simeq 2M$, the radius of a Schwarzschild black hole (in units $G=c=1$), leads to a minimum particle energy, or mass, of the order of h/M. Hence, the maximum number of such particles that go to make up a black hole of mass M is about M^2/h. An estimate of the entropy is then:

$$S = \xi k \left(\frac{M^2}{h}\right) \tag{1.9}$$

where k is Boltzmann's constant and ξ is a number of order unity, to be calculated from a proper, full theory of quantum black holes. It is obvious from (1.9) that the entropy S diverges in the classical limit $h \to 0$.

It is gratifying indeed that, using (1.5), we can write the entropy (1.9) in the form:

$$S = \left(\frac{\xi k}{16\pi h}\right) A \tag{1.10}$$

showing the entropy as proportional to the area of the event horizon, exactly as indicated on the basis of the four laws of black holes. This result was first suggested by Bekenstein (1973).

Finally, we deduce from (1.7):

$$T = \left(\frac{h}{2\xi k}\right) M^{-1} \tag{1.11}$$

for the *temperature* of the Schwarzschild black hole. On the basis of this formula, it is only in the classical limit $h \to 0$ that the hole is completely black ($T=0$). Otherwise, it appears to possess a temperature which, using (1.6), is:

$$T = \left(\frac{2h}{\xi k}\right) \kappa \tag{1.12}$$

which is indeed proportional to the surface gravity κ as we would have hoped.

What does this temperature mean physically? To assign a temperature to something implies that it can be in equilibrium with a surrounding heat bath at the same temperature. To be so, a black hole would have to emit heat energy at the same rate as it absorbed it. But the event horizon prevents any heat radiation from escaping from the hole—heat can flow in but not out.

If the concept of temperature makes sense then we must be able to associate with it a thermal equilibrium radiation spectrum. (There is an unimportant technical point here. The curved space round a black hole modifies the thermal equilibrium spectrum somewhat from the Planck form familiar from flat space–time.) The characteristic wavelength λ_0 of this radiation is:

$$\lambda_0 \simeq h/kT. \tag{1.13}$$

Using the value (1.11) for the black hole temperature T yields:

$$\lambda_0 \simeq 2M = \text{radius of black hole} \tag{1.14}$$

(assuming $\xi \simeq 1$). This result is not, of course, unexpected because we chose the entropy of the black hole in the first place with this wavelength in mind. Still, it shows that it is not meaningful to try and locate the origin of the heat radiation to be inside or outside the black hole, anymore than it is meaningful to discuss where in an atom a photon of light is created. That is, we do not have to actually say that there is heat radiation flowing *out of the black hole itself* to be able to associate with it a temperature given by (1.11). Nevertheless, we do have to assume that the hole is a source of heat radiation in some sense, i.e. it is not cold and black, but hot. In arriving at (1.11) we have hardly used any real physics, only some heuristic arguments about information and entropy. The precise nature of the mysterious heat radiation associated with black holes can only be determined by a proper quantum treatment.

2. Quantum black holes

The application of quantum theory to black holes by Hawking (1975) established the result that they should indeed emit thermal radiation with a temperature given by equation (1.11). His calculation fixes the value of ξ:

$$\xi = 8\pi^2. \tag{2.1}$$

Sometimes people find Hawking's result hard to understand because it depends on the details of quantum field theory in curved space–time. However, it is possible to gain considerable insight into the mechanism of the Hawking radiation process by considering a much more familiar system which nevertheless reproduces geometrically precisely the black hole situation. We shall first consider this simpler system, and then go on to discuss the black hole case in detail.

2.1. Radiation from a moving mirror

In quantum theory, we are used to the idea of quanta being created by sources; for example, photons being created by electrically charged currents. One way of visualising this is that the presence of the source disturbs the vacuum (no-quantum state) of the electromagnetic field and excites some of the modes. The mode excitations are the created photons.

Less familiar is the possibility of disturbing the vacuum (exciting modes) without having a source present at all. One way to do this is to contort the modes *geometrically*. In the case of a star imploding to form a black hole the dynamical gravitational field manifests itself through a changing geometry and the modes of the quantum field (e.g. electromagnetic field) are imbedded in this changing geometry and are thereby excited by the background motion.

A simpler example of a geometrical disturbance occurs when the field modes are constrained by reflecting boundaries (mirrors). In particular, if a mirror moves about, it will move the modes about too and thereby excite them, transferring energy to them in the form of quanta. Physically this will appear as particles created by the moving mirror, even though the mirror itself does not act as a *source* for the field. (Of course, a real mirror contains electric currents, although it may be overall electrically neutral.

However, the internal structure of the mirror is irrelevant to the discussion here because it is merely a device for reproducing on the field the effect of the gravity of a collapsing star.)

We shall now outline how a moving mirror can create radiation out of the vacuum (Fulling and Davies 1976, Davies and Fulling 1977). For simplicity we shall consider space to have one dimension (x) only and treat a massless scalar field $\phi(x, t)$ rather than the electromagnetic field. The wave equation is ($c=1$):

$$\frac{\partial^2\phi}{\partial x^2}-\frac{\partial^2\phi}{\partial t^2}=0. \tag{2.2}$$

The field modes under discussion are solutions of (2.2) which form a complete set of quantum states. In the absence of a mirror, these are conventionally chosen to have the form of complex exponential waves $\exp[-i\omega(t+x)]$ and $\exp[-i\omega(t-x)]$ for waves of frequency ω moving to the left and right, respectively.

If a mirror is inserted at the origin, $x=0$, then these left- and right-moving waves are no longer independent because the left-moving waves reflect from the mirror and become right-moving. To the right of the mirror the modes must now be taken as

$$\exp[-i\omega(t+x)]-\exp[-i\omega(t-x)] \tag{2.3}$$

which are chosen to vanish at the mirror ($x=0$). We shall call these modes ψ_ω to avoid writing this cumbersome expression repeatedly, and include in the definition a conventional normalisation factor $(4\pi\omega)^{-1/2}$.

The field ϕ is quantised by first expanding it in terms of the complete set of modes ψ_ω:

$$\phi(x, t)=\sum_\omega (a_\omega\psi_\omega+a_\omega{}^*\psi_\omega{}^*) \tag{2.4}$$

where the a_ω are expansion coefficients and * denotes complex conjugation, then quantising each mode (independently) as a simple harmonic oscillator. (In the case treated here Σ_ω denotes $\int_0^\infty d\omega$.) The amplitudes a_ω, $a_\omega{}^*$ then become quantum *operators*, which respectively annihilate (de-excite) or create (excite) quanta in the mode ψ_ω.

The vacuum (no-quantum) state, denoted by $|0\rangle$, is defined as the state which vanishes under the action of all the operators a_ω:

$$a_\omega|0\rangle=0 \qquad \text{for all } \omega. \tag{2.5}$$

The operator which represents the number of particles in the mode ω is $a_\omega{}^*\, a_\omega$. Thus the expected number of quanta of all frequencies contained in the vacuum state is, using (2.5):

$$\langle 0|\sum_\omega a_\omega{}^* a_\omega|0\rangle=0. \tag{2.6}$$

Now suppose that the mirror moves about (see figure 1). This cannot affect the left-moving waves $\exp[-i\omega(t+x)]$ to the right of the mirror because these are incoming from infinity. The reflected (right-moving) waves will, however, suffer a Doppler shift due to the motion of the mirror. This is a red shift if the mirror recedes to the left, and a blue shift if it approaches to the right. For a general mirror motion, the reflected waves will be very complicated, but it is still possible in this model to calculate them exactly as a functional of the mirror trajectory. We do not need them explicitly for this discussion, so the reflected waves will simply be denoted

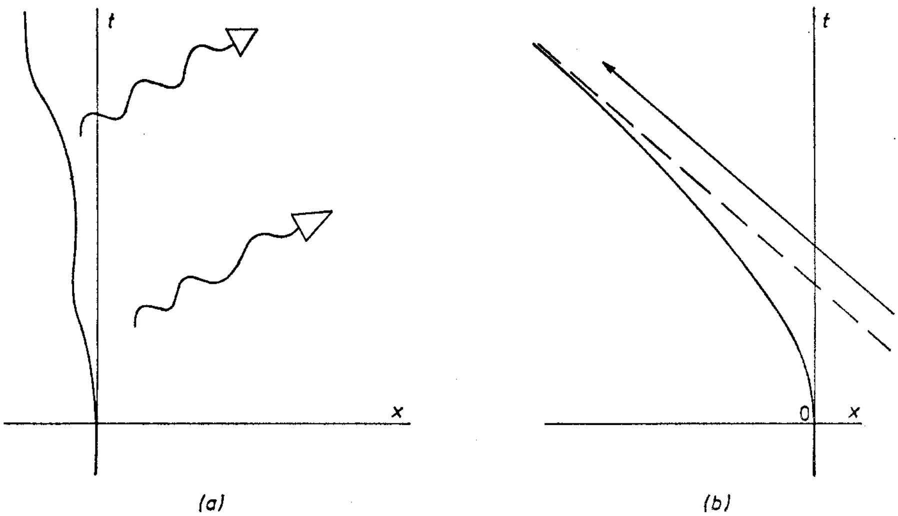

Figure 1. (a) Radiation from a moving mirror. The motion of the reflecting boundary disturbs the vacuum of the quantum field and causes the production of quanta which propagate away to the right as shown. (b) When a mirror accelerates away to the left along the asymptotically null (light-like) trajectory $x \to -t - A \exp(-2\kappa t) + B$ the radiation which is produced at late times has a Planck (thermal equilibrium) spectrum. The broken line is the null asymptote $v = B$. Incoming null rays that are later than this ray do not reflect from the mirror but disappear off to left-hand infinity.

$\exp[-i\omega f(t-x)]$ where f is a complicated, but calculable, function depending on the details of the mirror motion.

In the new situation the modes of the scalar field will be of the form:

$$\exp[-i\omega(t+x)] - \exp[-i\omega f(t-x)] \tag{2.7}$$

and these we shall call $\bar{\psi}_\omega$. The field $\phi(t, x)$ can be expanded as before:

$$\phi(t, x) = \sum_\omega (b_\omega \bar{\psi}_\omega + b_\omega^* \bar{\psi}_\omega^*) \tag{2.8}$$

where we now have a new set of amplitude operators denoted by b_ω.

Generally the b_ω will differ from the a_ω but they are related by a transformation of the form:

$$b_\omega = \sum_{\omega'} (\alpha_{\omega\omega'}^* a_{\omega'} - \beta_{\omega\omega'}^* a_{\omega'}^*) \tag{2.9}$$

known as a Bogolubov transformation. The coefficients α and β may be evaluated by expanding the modes $\bar{\psi}_\omega$ in terms of the modes ψ_ω:

$$\bar{\psi}_\omega = \sum_{\omega'} (\alpha_{\omega\omega'} \psi_{\omega'} + \beta_{\omega\omega'} \psi_{\omega'}^*). \tag{2.10}$$

As usual in field theory, the complex conjugate of the modes, such as ψ_ω^*, are considered to have negative frequency. Equation (2.10) therefore expresses the fact that mode functions of the type (2.7) cannot in general be expanded in purely positive-frequency exponentials (this is a well-known property of Fourier analysis). So long as $\beta_{\omega\omega'} \neq 0$, $\bar{\psi}_\omega$ will contain a superposition of positive- and negative-frequency ψ_ω modes and this will be the general situation for a non-trivial function $f(t-x)$ in (2.7).

We come now to the crucial point. Because $b_\omega \neq a_\omega$ in general, the state $|0\rangle$ which represents the vacuum (no-quanta) state for a static mirror will *not* be a vacuum state for the modes $\bar{\psi}_\omega$ (and operators b_ω) associated with the moving mirror. It will, in general, contain quanta. Indeed, the expected number of quanta in the mode τ is:

$$\langle 0|b_\tau{}^* b_\tau|0\rangle = \sum_{\omega'} |\beta_{\tau\omega'}|^2 \tag{2.11}$$

a result which follows using (2.5) and inverting (2.9). Thus if $\bar{\psi}_\tau$ contains any negative-frequency ψ_ω modes then (2.11) will be non-zero. Physically this has a clear meaning. Suppose that prior to some moment, say $t=0$, the mirror is static and there are no field quanta present. The quantum state is therefore $|0\rangle$ as defined by equation (2.5) in terms of the modes ψ_ω. The mirror then embarks on a period of acceleration and the modes become modified to $\bar{\psi}_\omega$. The state however remains as $|0\rangle$ (we are working with free fields in the Heisenberg picture). Equation (2.11) then tells us that the motion of the mirror has created quanta out of the vacuum and this radiation then proceeds to flow away to the right.

One of the significant features about the particle creation is that a receding mirror, which has the effect of weakening (red-shifting) any reflected *classical* radiation, is just as effective in the spontaneous production of particles as an approaching (blue-shifting) mirror. Hence, high red-shift trajectories are efficient for particle creation.

One particular such trajectory is of interest:

$$x \to -t - A\exp(-2\kappa t) + B \qquad \text{as } t \to \infty \tag{2.12}$$

where A, B and κ are just arbitrary real and positive constants. This trajectory, which is joined smoothly to the static configuration at $t=0$, is shown in figure 1(*b*). The mirror recedes to the left with ever-increasing acceleration, approaching the speed of light asymptotically. Notice that there is a latest ray which can reflect from the mirror out to the right. After this one, all incoming rays disappear off to the left.

The red shift from the surface of a mirror moving along the trajectory (2.12) increases exponentially with an e-folding time κ, so that any intrinsic luminosity of the mirror, or any reflected incoming radiation, will also fade out exponentially. However, the spontaneous particle creation does not. The outgoing (reflected) part of the modes $\bar{\psi}_\omega$ behaves like:

$$\exp\{i\omega A \exp[-\kappa(t-x)]\} \tag{2.13}$$

which is one of the few functions for which the transformation coefficients $\alpha_{\omega\omega'}$ and $\beta_{\omega\omega'}$ can be computed in terms of known functions (Davies and Fulling 1977). The result can then be used in (2.11) to calculate the number of particles created in each mode ω, which tells us the *spectrum* of the radiation. This turns out to be (we use units $\hbar = 1$):

$$\frac{1}{\exp(2\pi\omega/\kappa) - 1} \tag{2.14}$$

which is a Planck (thermal equilibrium) spectrum at a temperature $\kappa/2\pi$. Of course, if the mirror continues to accelerate indefinitely, then an infinite number of quanta (i.e. a steady flux of radiation) will be created in each mode.

2.2. *Radiation from black holes*

We are now ready to deal with the black hole problem. When radiation (e.g. light) passes through a gravitational field it suffers a gravitational frequency shift, an effect that can be measured even on Earth. If a beam of light could be shone through the centre of a static compact star, such as a neutron star, it would acquire a blue shift as it fell inwards and a red shift as it climbed away again on the other side. The net shift is zero. However, if the star implodes while the light is passing through it, the climb outwards becomes considerably harder, because the escape energy goes up. Hence the energy gained by the infall does not pay for the work needed to escape on the other side and there is a net red shift. Because a star undergoing catastrophic gravitational collapse shrinks appreciably during the time it takes for the light to traverse its interior, it is clear that this red shift will be very large. In fact, it increases exponentially with an *e*-folding time determined by the surface gravity, κ. For a star of solar mass this time is very short indeed—about 10^{-5} s. For a smaller mass it is proportionally greater.

The gravitational red shift imparted to the electromagnetic field by a collapsing star is a direct analogue of the Doppler red shift imparted by the receding mirror which moves along the trajectory (2.12). Both increase exponentially on a time scale controlled by the constant κ (in the mirror case this is just a parameter characterising the trajectory). Both have a latest ray that can pass into the system and out again. In the mirror case any later rays disappear off to the left, in the collapsing star these rays disappear down the black hole; they are trapped by the horizon and are drawn towards the singularity (see figure 2).

The effect on the field modes is almost identical. Around a spherical *static* star there will be incoming S wave (spherically symmetric) modes which start out at large radial distance like:

$$\exp\left[-\mathrm{i}\omega(t+r)\right]/r \tag{2.15}$$

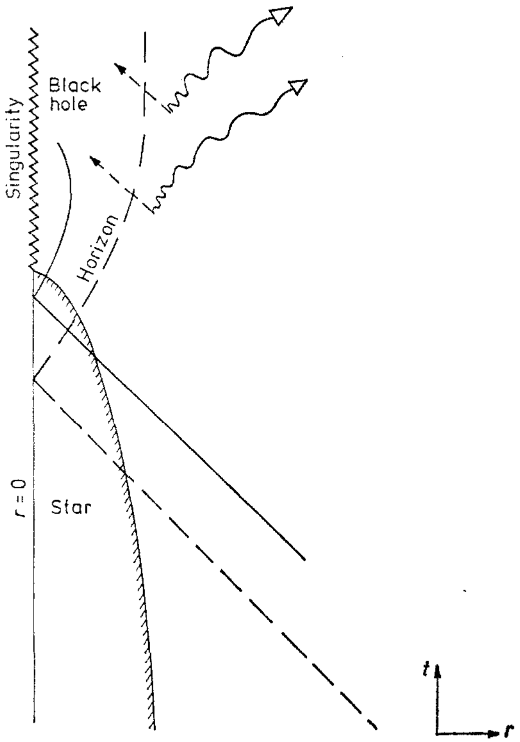

Figure 2.
Black hole evaporation. The star implodes to a singularity at $r=0$. Its gravity bends the nearby null rays so that there is a latest ray (broken line) that can just pass through the centre of the collapsing star and escape on the other side to infinity. This latest ray forms the event horizon. Later rays are trapped inside it and cannot escape from the black hole. From the standpoint of geometrical optics the situation outside the horizon is identical to the moving mirror system of figure 1(*b*) (they differ in appearance only because of a conformal transformation). The Hawking radiation produced is paid for by an ingoing negative energy flux (broken arrows).

pass through the star and emerge on the other side, at large r, like

$$\exp[-i\omega(t-r)]/r. \tag{2.16}$$

If the star collapses, it is straightforward to show that at late times, the outgoing modes (2.16) get red-shifted exponentially to:

$$\frac{1}{r}\exp\{i\omega A\exp[-\kappa(t-r)]\} \qquad (A \text{ constant}) \tag{2.17}$$

which (apart from the use of radial coordinates) is identical to the mirror-reflected modes (2.13) with κ the surface gravity (see §1.2).

The collapsing star therefore produces a flux of quanta travelling radially outwards for exactly the same reason as the moving mirror—a geometrical disruption of the vacuum—caused by the collapse of the star. Moreover, the geometrical optics of the two systems is identical and the calculation of the particle spectrum is nearly the same. The main difference is the fact that some of the outgoing radiation scatters gravitationally from the space curvature around the black hole and ends up going back down again. This is a frequency-dependent effect and so it alters the spectrum somewhat. However, the spectrum is still that of a black hole in *thermal equilibrium*, because there is a statistical balance between this back-scattering and that of incoming radiation being scattered out again. This is the reason that the thermal equilibrium spectrum associated with a black hole is not generally quite the same as the Planck spectrum. The spontaneous radiation flux therefore has the characteristics of a thermal spectrum with a temperature (see (2.14)):

$$T=\frac{\kappa}{2\pi k}=\frac{1}{8\pi kM} \tag{2.18}$$

precisely as expected from (1.12) with $\xi=8\pi^2$. This brilliant confirmation of the thermodynamic basis of black holes was demonstrated by Hawking (1974), and provides a most elegant connection between two previously unrelated topics in physics.

Although the outline described here refers to a massless scalar field, the basic mechanism applies to all quantum fields. In the real world, the black hole would produce photons, neutrinos and, presumably, gravitons. If the temperature is high enough ($kT\gtrsim mc^2$) then massive particle production will also occur. For a solar-mass black hole, the temperature is very low indeed—about 10^{-7} K. For a microscopic hole about the size of an atomic nucleus the temperature is $\sim 10^{12}$ K and electron–positron pairs would be produced. These mini-holes could have been formed from the aggregation of about 10^{14}–10^{15} g of dense primeval matter during the cosmological big bang (Hawking 1971, Carr 1977).

2.3. *The energy balance*

Compelling though it is, the prediction of a steady flux of particles from a black hole still leaves a number of mysteries. Where do the particles come from? Where does the energy come from? What happens in the end when the energy supply runs out?

In the example of the moving mirror it is natural to identify the surface of the mirror as the place where the radiation is created. There is no such privileged place for the black hole. It is most important to realise that the material of the imploding star itself need not be directly coupled to the quantum field to produce the Hawking effect. The particles are produced out of empty space by the gravitational disturbance caused by the star's implosion and *not* by the matter of the star acting as a *source*. Two bodies of the same mass (assuming for now $J=e=0$) made of totally dissimilar material would give rise to identical Hawking radiation.

As discussed in §1.3, it is not meaningful to locate the origin of the created particles, because their wavelength is comparable to the size of the black hole. However, we can ask for the energy density and energy flux of this radiation at different locations round the black hole, because these quantities *are* locally defined. To proceed covariantly, we should calculate the expectation value of the stress-energy-momentum tensor, or stress tensor for short, in the chosen quantum state (e.g. the state $|0\rangle$ corresponding to an initial vacuum). This is denoted $\langle T_{\mu\nu}\rangle$. An outline of the method of calculation will be given in §4.6. In the moving-mirror case it is easy to evaluate explicitly (Fulling and Davies 1976). The result shows that the energy of the radiation arises at the mirror surface and flows away to the right, remaining at constant density along the retarded null rays $t-x$.

Mathematical complexity has so far precluded a complete evaluation of $\langle T_{\mu\nu}\rangle$ for the black hole, but the general features are illustrated by a model *two-dimensional* black hole, for which $\langle T_{\mu\nu}\rangle$ has been calculated (for massless fields) in complete detail (see §4.6). The results (Davies 1976a, Davies *et al* 1976) show that around a massive, static body the gravitational field (or space curvature) induces a static vacuum stress, which is rather like the vacuum polarisation produced by the Coulomb field of a nucleus (and which contributes to the Lamb shift). This vacuum stress can be envisaged as a cloud of *negative* energy surrounding the body and falling off rapidly in density at large r.

Curiously, there is a laboratory analogue of this negative vacuum energy, which can actually be measured (Tabor and Winterton 1969). If two mirrors (conducting plates) are placed parallel and opposite each other there is an electromagnetic force of attraction between them even though the plates are electrically neutral. The force, which is independent of the electronic charge, arises because the presence of the plates disturbs the vacuum of the electromagnetic field and creates a static cloud of (uniform) negative-energy density in the space between them. The vacuum disturbance is really a geometrical effect due to the fact that the electromagnetic field modes are forced to become a discrete set of standing waves in the direction perpendicular to the plates. Thus, some long-wavelength modes are excluded because they cannot 'fit' into the space. Crudely speaking, each mode oscillator carries an unobservable zero-point vacuum energy $\frac{1}{2}\hbar\omega$, so the exclusion of these modes lowers the energy (to some negative value) between the plates. The phenomenon is known as the Casimir effect (Casimir 1948).

The presence of a cloud of negative vacuum energy is negligible around an ordinary star, but near a black hole it can become intense, and is comparable in magnitude to the energy of the Hawking radiation. When the star implodes, a hole appears in the centre of the negative-energy cloud, so this negative energy continuously streams into the black hole as the Hawking radiation streams away to infinity. Conservation of energy demands that the Hawking flux is paid for by this ingoing negative energy and the effect is to steadily reduce the mass–energy of the black hole at just the right rate to compensate for the thermal radiation appearing in the surroundings. In this way the mass of the hole supplies the energy of the Hawking radiation, *not* by this energy emerging from inside the hole (which is impossible because of the event horizon) but by receiving *negative* mass–energy from the *inflowing* vacuum stream. Thus the requirements of both causality and energy conservation are simultaneously satisfied.

As the mass of the hole is steadily reduced it shrinks in size. The violation of Hawking's area theorem (1.2) is, we recall, explained by the presence of negative energy. From the thermodynamic point of view there is also consistency: the entropy of the black hole (as measured by its area) diminishes, but the creation of high entropy thermal radiation in the surrounding environment more than offsets this decrease and saves the second law of thermodynamics.

One puzzling feature that remains concerns the experience of an observer who falls into the black hole (Hawking 1975). The Hawking radiation, as we have seen, is intimately connected with the existence of the event horizon and the exponential red shift associated with it. However, an observer who falls into a black hole does not experience the red shift or the horizon. Indeed, the horizon has no local significance (you cannot tell when you are crossing it—horizons are global constructs depending on the experiences of observers over the whole of future time). But freely falling observers do *not* experience a loss of information; they can continue to see what is going on inside the collapsing matter after it enters the black hole. All this implies that a freely falling observer should not encounter many created particles. Yet if the particles are really there, surely he will pass by them as he falls towards the horizon? More than this: because of the time dilation effect, the falling observer seems to a distant observer to take an *infinite* time to reach the horizon, even though in his own frame it takes a finite (and very short) time. Thus, as far as the distant observer is concerned, the falling observer should encounter *all* the particles emitted in that direction over the whole lifetime of the black hole.

The resolution of this paradox comes again from the appreciation that the particles cannot be localised inside the region in which the observer is falling. The wavelengths of the quanta are necessarily much larger than the observer's body, or his apparatus, so we should not expect him to attribute much significance to clicks in a counter, for instance. On the other hand, he can (in principle) measure the flux of energy which passes him as he falls, and it might be thought that the time-integrated energy of *all* the Hawking radiation would be very large. However, as far as energy is concerned there is no way the observer can distinguish the Hawking radiation flux from the other flux of energy which is observed due to his sweeping down through the cloud of static vacuum energy. Moreover the latter, being negative, subtracts from the Hawking flux contribution, so that the total energy flux which he encounters is actually rather small. In summary, a freely falling observer does not encounter very much quantum energy, even though the distant observer does.

Until now we have been deliberately vague about the total quantity of radiation emitted from the hole. If the back-reaction is ignored, the Hawking process is simply a steady flux which continues indefinitely. In practice, the energy loss causes a slow shrinkage of the hole, as has been described. Thus, hot black holes slowly *evaporate*. As they do so their mass diminishes and according to (2.18) the temperature then rises, at least in the Schwarzschild case. This in turn accelerates the evaporation rate and the process proceeds catastrophically towards an explosive end. In the last second the energy release is equivalent to many hydrogen bombs. It is conceivable that the intense burst of radiation, or secondary radiofrequency effects (Rees 1977), could be detected.

So long as the rate of shrinkage of the hole is much less than the frequency of the

Hawking radiation, the effect of the back-reaction on the radiation spectrum can be ignored—the black hole still behaves as if it is approximately in thermodynamic equilibrium. Eventually this approximation will fail. This happens when $dM/dt \simeq \lambda_0 \simeq M$. But dM/dt is given by the rate of energy loss, which according to Stefan's law is proportional to T^4 or M^{-4} for a black hole (ignoring the deviation from the Planck spectrum). The area of the hole is $16\pi M^2$, so the rate of mass loss by the hole is of the order of M^{-2}. Thus, back-reaction is important when $M^{-2} \simeq M$, or $M \simeq 1$ which in our units is the so-called Planck mass ($\simeq 10^{-5}$ g). When the hole shrinks to this mass its size is a mere 10^{-33} cm. The present derivation of the Hawking effect then breaks down and the details of the evaporation process might become strongly modified. However, we cannot really guess what might happen because the Planck dimensions considered also herald the onset of quantum gravity effects, in which higher-order quantum processes of the gravitational field *itself* start to become important. As we do not yet have a sensible theory of quantum gravity it is not known what the effect of quantum gravity, and hence the ultimate fate of the hole, will be. Perhaps the black hole just disappears altogether, taking the material of the star with it. This outcome would, of course, mean abandoning conservation of baryon number, etc, because we start out with a star of ordinary matter and end up with just thermal radiation.

2.4. *Super-radiance*

If the collapsing body is rotating it will produce a black hole with non-zero angular momentum J and angular velocity Ω. The system will no longer be spherically symmetric (though it will be axisymmetric at late times) so that the S-wave modes such as (2.15) and (2.17) no longer accurately illustrate the behaviour of the quantum field. The red shift imparted to the outgoing modes will depend on their orientation relative to the rotation axis—that is, on their quantum number m (the usual axial component of the field angular momentum). The effect of this when worked through the analysis is to modify the Planck factor (2.14) to:

$$\frac{1}{\exp\left[(\omega - m\Omega)/kT\right] - 1} \tag{2.19}$$

(this is quite apart from the other modification due to back-scattering of the waves from the space curvature mentioned on page 1327).

The quantity T which occurs in (2.19) is still related to the surface gravity by the formula:

$$T = \kappa/2\pi k$$

but κ is no longer simply $(4M)^{-1}$ as it is in the Schwarzschild case. Instead it has a more complicated form which is given in equation (3.4) in the next section.

It would not be correct to regard (2.19) as a thermal equilibrium spectrum. To see this, suppose that the rotating black hole were immersed in an isotropic heat bath at the temperature T, then the system would not be in equilibrium for the following reasons. Consider two quanta emitted with orbital angular momentum axial component parallel and antiparallel to the spin vector of the black hole respectively. The former will have $m > 0$ and the latter $m < 0$. Hence the magnitude of factor (2.19) will be greater for the quantum which carries away angular momentum parallel to the black hole's spin axis than the other which is antiparallel. This means that the probabilities of the emission of the two quanta are not the same. There is preferential emission of particles which tend to *reduce* the spin of the black hole. This implies that the hole will slowly lose angular momentum and spin down. It is not in equilibrium. Nevertheless, equilibrium could be established by immersing the black hole in a corotating bath of thermal radiation which reproduced the features of the spectrum (2.19).

The emission of quanta by a rotating black hole can actually be understood without recourse to the Hawking process involving a collapsing body, a feature which follows from inspection of (2.19). If we let $T \to 0$, then for a Schwarzschild ($\Omega = 0$) black hole it is clear that (2.19) vanishes, which simply says that the black hole does not radiate by the Hawking process. However, if $\Omega \neq 0$ then there exist frequencies ω for which $\omega - m\Omega < 0$, and then the limit $T \to 0$ does *not* cause (2.19) to vanish. The origin of this residue of non-thermal radiation, which was actually discovered before Hawking's radiation, can be traced to a wave analogue of the Penrose process for energy extraction by a particle from a rotating black hole, discussed on page 1318. An incoming classical wave can be amplified by passing near (but not into) the rotating hole and removes some of the rotational energy as a consequence. Quantum mechanically this is rather like a laser, with the rotating hole stimulating the emission of quanta. On the basis of detailed balancing arguments it then follows that there should also be spontaneous emission and this is just the residue of non-thermal radiation discussed above (Zel'-dovich 1970, Unruh 1974).

It is interesting to notice that this radiation removes both mass–energy and angular momentum from the black hole. However, the condition:

$$\omega < m\Omega \tag{2.20}$$

ensures that the emitted quanta always remove more angular momentum ($m\Omega$) than energy (ω). Thus the ratio J/M for the black hole will be reduced rather than increased by this process. This is a good illustration of cosmic censorship at work, for if J/M were to increase, it could approach the limiting value of one at which the black hole converts into a naked singularity.

If the black hole carries an electric charge and its temperature is great enough to permit massive particle production, then it will preferentially radiate charge of the opposite sign in an attempt to discharge itself. This is a form of charge super-radiance, analogous to the spin-down tendency of rotational super-radiance. An analysis has been given by Gibbons (1975). The super-radiant frequencies are those for which

$$\omega < m\Omega + e\Phi. \tag{2.21}$$

If we wish to consider a charged, rotating black hole in equilibrium with a surrounding heat bath then we must expect that, due to these super-radiant effects, the heat bath will not contain ordinary thermal radiation but will be corotating with the black hole and possess a compensating electric potential. The latter is, of course, unnecessary for black holes with masses $\gg 10^{15}$ g, because their temperatures are too low to produce charged particles.

3. Thermodynamic black hole processes

With the establishment of the quantum basis of the thermodynamic connection by Hawking, it is possible to bring black holes within the framework of ordinary thermodynamic theory (Davies 1977a, Hut 1977). The event horizon area *is* the entropy of the hole and the temperature is now seen to have its usual significance. Some of the features of black hole thermodynamics are rather strange, though.

3.1. The fundamental thermodynamic equations

It is convenient in this section to put Boltzmann's constant $k=1/8\pi$ and $\hbar\equiv h/2\pi=1$. The establishment of the value of ξ (see equation (2.1)) enables the precise numerical connection to be made between the event horizon area A and the entropy S. From equation (1.10) this is:

$$S=\tfrac{1}{4}A. \tag{3.1}$$

This relation remains true for the generic case (e, $J\neq 0$) so from equation (1.1) we can write down the fundamental thermodynamic equation for a black hole:

$$M^2=2S+\frac{1}{8S}(J^2+\tfrac{1}{4}e^2)+\tfrac{1}{2}e^2 \tag{3.2}$$

and use this to study black hole processes on the basis of equilibrium thermodynamics.

The first law, equation (1.3), now assumes the conventional form:

$$\mathrm{d}M=T\,\mathrm{d}S+\Omega\,\mathrm{d}J+\Phi\,\mathrm{d}e \tag{3.3}$$

where

$$T=\frac{\partial M}{\partial S}=\frac{1}{M}\left(1-\frac{J^2+\frac{1}{4}e^4}{16S^2}\right)=4\kappa \tag{3.4}$$

$$\Omega=\frac{\partial M}{\partial J}=\frac{J}{8MS} \tag{3.5}$$

$$\Phi=\frac{\partial M}{\partial e}=\frac{e(e^2+8S)}{16MS} \tag{3.6}$$

and κ is the surface gravity of a charged, rotating hole.

Euler's theorem allows (3.3) to be integrated to obtain the Gibbs–Duhem relation:

$$\tfrac{1}{2}M=TS+\Omega J+\tfrac{1}{2}\Phi e$$

which reduces to (1.7) when $J=e=0$.

Equation (3.2) can be inverted to give an expression for the entropy:

$$S=\tfrac{1}{4}M^2-\tfrac{1}{8}e^2+\tfrac{1}{4}M^2(1-e^2/M^2-J^2/M^4)^{1/2}. \tag{3.7}$$

In spite of the similarity between this set of equations and that for a conventional thermodynamic system, there are a number of features peculiar to the black hole case. One of these is that the entropy of a black hole cannot really be visualised as spread with a uniform density through it; black hole entropy is a global concept. Physically, this means that we cannot divide black hole systems into constituent subsystems which differ only by *scale*, as we can with, say, a laboratory gas. Mathematically, this limitation is connected with the appearance of M^2 rather than M on the left of equation (3.2): the total energy is *not* a homogeneous *first*-order function of 'extensive' parameters. If two Schwarzschild black holes of masses M_1 and M_2 are coalesced without loss of total energy, the final mass will be M_1+M_2 and the final entropy $\frac{1}{2}(M_1+M_2)^2$, which is always greater than the initial entropy $\frac{1}{2}M_1{}^2+\frac{1}{2}M_2{}^2$.

Black hole combination is therefore in itself an irreversible process: entropy considerations prevent a black hole from bifurcating.

From equation (3.7) it is clear that S is maximised by making $J = e = 0$. Thus the effects of super-radiance, discussed in §2.4, which tend to irreversibly reduce the values of J and e by radiating away angular momentum and charge are seen to be a direct consequence of the second law of thermodynamics.

3.2. Equilibrium and stability

If a black hole is to be in thermodynamic equilibrium at some temperature T, then it must be surrounded by a heat bath at the same temperature. Suppose the system is enclosed in a box of volume V with impermeable walls, containing a mixture of n Schwarzschild black holes and some thermal radiation. Neglecting the small volume of the black holes, the distortion of the Planck spectrum due to space curvature and surface effects associated with a box of finite size, we may write for the total entropy of the contents of the box:

$$S = \tfrac{1}{2}\, nM^2 + \tfrac{4}{3}\,(aV\, M_r{}^3)^{1/4} \qquad (3.8)$$

where M_r is the mass of radiation and a is the usual radiation constant. The total energy is constant, which we call E, so:

$$S = \tfrac{1}{2}\, nM^2 + A(E - nM)^{3/4} \qquad (3.9)$$

where $A \equiv \tfrac{4}{3}\,(aV)^{1/4}$.

The system will be in equilibrium when $\mathrm{d}S/\mathrm{d}M = 0$, which occurs for a black hole mass M which is a root of the equation:

$$f(M) \equiv nM^5 - EM^4 + \left(\frac{3A}{4}\right)^4 = 0 \qquad (3.10)$$

lying between 0 and E. The function $f(M)$ has only two turning points: at $M = 0$ and $M = (4/5)\, E/n$, so it possesses at most three roots. One of these must always be negative, because $f(M) \to -\infty$ as $M \to -\infty$, but $F(0) = (3A/4)^4 > 0$. This also shows that $F(0)$ is a maximum. Hence the turning point at $M = (4/5)\, E/n$ must be a minimum. The remaining two roots will therefore only exist if this minimum lies below the axis $f(M) = 0$, i.e. if $F(4E/5N) < 0$. So we can conclude that the contents of the box will only possess equilibrium states if $F(4E/5N) < 0$, or if:

$$V \leqslant \frac{4^4}{5^5}\frac{E^5}{an^4}. \qquad (3.11)$$

This curious result can be better understood by computing the *specific heat* of a Schwarzschild black hole from (3.4):

$$C \equiv T\left(\frac{\partial S}{\partial T}\right)_{J=e=0} = -\frac{M}{T} = -\frac{1}{T^2} = -M^2. \qquad (3.12)$$

That this is *negative* corresponds to the fact that the black hole heats up as it radiates energy ($T \propto M^{-1}$). It occasions no surprise, because negative specific heats are a familiar feature of self-gravitating systems (see §1.1).

Imagine a box with fixed volume V, containing a quantity of energy E in the form of thermal radiation. Suppose we wish to convert a fraction of this radiation energy into one (or more) black holes, in such a ratio that the final radiation temperature equals the black hole temperature and equilibrium prevails. First we make a very small black hole. Its temperature is very high—much higher than the radiation. Now we allow the hole to grow at the expense of the radiation. As the hole gets bigger, its temperature *falls*. At the same time, the surrounding heat bath, whose energy is being plundered to feed the hole, falls also in temperature. There is thus a competition, with the black hole attempting to grow and fall to the temperature of the heat bath, and the heat bath cooling to supply the mass of the hole.

There are two possible ways in which equilibrium might be achieved: a small hole (high temperature) whose formation does not deplete the heat bath by much, or a large hole (low temperature) where a large fraction of the heat bath has been removed, lowering its temperature appreciably. These two situations correspond to the two roots of equation (3.10). For a fixed volume, neither of these equilibria can be achieved unless the total energy E available is sufficiently great, because for smaller values of E the reduction in temperature of the heat bath is correspondingly greater to supply the energy of a black hole of given mass. As more and more energy is converted, the heat bath falls in temperature faster than the hole and they can never reach a common temperature. The limiting case is given by the equality sign in (3.11), where the two roots coincide.

The question now arises about *stability*. Do the roots of (3.10) correspond to entropy maxima or minima? As there cannot be two maxima without a minimum between them, or vice versa, it follows that one root must be a maximum and the other a minimum. Because $\mathrm{d}S/\mathrm{d}M<0$ at both extremes $M=0$, $M=E$, it is clear that the low-M black hole corresponds to stable equilibrium and the high-M black hole to unstable equilibrium.

If the latter situation were to occur, the system would achieve stability in one of two ways. The small hole could start to evaporate, heating up the surrounding radiation bath, but not fast enough to match its own rise in temperature. The evaporation would thus escalate and the black hole would eventually explode and presumably disappear, leaving just the radiation. Alternatively, the hole could start to grow and deplete the heat bath, but initially not fast enough to match its own drop in temperature. The growth would continue until the energy of the heat bath was low enough for its rate of fall in temperature to match that of the hole, and then the system would come into equilibrium in a situation corresponding to the other stable root of (3.10).

It is also clear from (3.11) that equilibrium is harder to achieve when several black holes are present. However, several black holes cannot be in *stable* equilibrium together. Because S is proportional to the *square* of M, it is always entropically more favourable for them to coalesce. If they do this without loss of energy, then their final mass will be nM, so their entropy will be $\frac{1}{2}n^2M^2$, compared to the initial entropy of $\frac{1}{2}nM^2$.

An alternative procedure for merging together black holes is to require that the final entropy be the same as the initial entropy (reversible change) and enquire what fraction of the mass of the system may be extracted as energy (Hawking 1973). For n Schwarzschild black holes of equal initial mass one obtains the fraction $1-n^{-1/2}$. Thus, for two black holes, the maximum extractable energy allowed by the second law of thermodynamics is about 29%. The prospect of, say, two solar-mass black holes divesting themselves in one go of 10^{54} erg—more than twenty times their total energy output from nuclear burning over the billions of years they spend as ordinary stars—is mind-boggling. Moreover, for a sufficiently large number of holes ($n\gg 1$) the fractional energy release can clearly be even greater.

In the general case of rotating charged black holes, still more energy may be released. If the initial holes differ only in the sign of the charge, and they are coalesced along their spin axes in a counter-rotating fashion, the final J and e will be zero. The maximum fractional energy extracted is then:

$$1-\frac{1}{\sqrt{2}}\,n^{-1/2}\left[1-\frac{e^2}{2M_i^2}+\left(1-\frac{e^2}{M_i^2}-\frac{J^2}{M_i^4}\right)^{1/2}\right]^{1/2} \qquad (3.13)$$

where M_i is the mass of one of the initial black holes. In the limiting case (1.4) of maximum e and J, (3.13) reduces to:

$$1-\frac{1}{\sqrt{2}}\,n^{-1/2}\left(1-\frac{e^2}{2M_i^2}\right)^{1/2}. \qquad (3.14)$$

To maximise expression (3.14) one should choose $J=0$, $e^2=M_i^2$ (non-rotating, supercharged hole). This yields the greatest possible fractional energy that black holes can release according to the second law: $1-\frac{1}{2}\,n^{-1/2}$. This is about 65% for two black holes. The other limit, $e=0$, $J=M_i^2$, yields only 50%.

It is instructive to compare the entropy of a black hole with that of ordinary matter of the same mass. For a non-relativistic gas, the entropy of N particles is of the order of Nk. If the particles are, say, of atomic mass ($\simeq 10^{-19}$ in these units) the entropy of a mass M is thus about $10^{19}\,M$, compared to $\frac{1}{2}\,M^2$ for the black hole. In the case of the Sun, $M\simeq 10^{33}\,\mathrm{g}\simeq 10^{38}$ units, so its entropy is about 10^{57}, which is nearly twenty powers of ten less than that of a black hole of the same mass. The two entropies will be about equal for a mass of 10^{19} units, or around 10^{14} g, which is the mass of a large asteroid. It is thus extremely entropically favourable for a solar-mass object to implode to form a black hole, but unfavourable for mini-holes of less than about 10^{14} g to form from non-relativistic matter. It is doubtful, however, if this is relevant to the formation of primordial mini-holes, as the cosmological fluid in the big bang was far from thermodynamic equilibrium and was also highly relativistic.

A black hole of mass 10^{14} g has a further curious significance. Its size is about one fermi (10^{-13} cm), or about the size of an atomic nucleus. If it is formed by the implosion of 10^{38} protons with their spins aligned, its angular momentum J will be equal to M^2, i.e. it will be an extreme Kerr black hole of limiting rotation rate.

There is a further significance to black holes of this mass. Page (1976) has carried out detailed computer calculations for the emission rate of gravitons, neutrinos and photons, which yield a total luminosity for a Schwarzschild black hole of

$$(3{\cdot}4\times 10^{46})\left(\frac{M}{1\ \mathrm{g}}\right)^{-2} \qquad \mathrm{erg\ s^{-1}}.$$

The *lifetime* of this hole against total evaporation is obtained by integrating this over time to obtain $10^{-26}\,(M/1\ \mathrm{g})^3$ s. For a mass of about 10^{14} g this works out at about 10^{10} yr, which is just the present age of the universe. These various coincidences are examples of the famous 'big number' coincidences between cosmological and atomic constants first noted by Eddington and Dirac (see, for example, Davies 1977b).

3.3. Phase transitions and the third law

Imagine a black hole at some temperature T in equilibrium with a surrounding heat bath. In general the hole will have angular momentum J and electric charge e.

Suppose there is a small, reversible transfer of energy between the hole and the environment which occurs in such a way that J and e remain unchanged. The specific heat corresponding to this transfer is easily calculated from (3.4) and (3.7):

$$C_{J,e} \equiv T\left(\frac{\partial S}{\partial T}\right)_{J,e} = \frac{8\,MS^3T}{J^2+\frac{1}{4}e^2-8T^2S^3} \tag{3.15}$$

which reduces to (3.12) in the limit $J=e=0$.

If we examine the opposite limit corresponding to condition (1.4), it follows from (3.7) and (3.4) that $T \to 0$ in this limit. Thus a black hole may be cooled down by rotating it or charging it up. In equation (3.15) we see that as $T \to 0$, $C_{J,e} \to 0$ through *positive* values. Recall, however, that the specific heat of a Schwarzschild black hole is *negative.* Consequently, at some values of J and e between these limits, $C_{J,e}$ changes sign. This occurs when the denominator of the right-hand side of (3.15) vanishes. At this point the specific heat passes from negative to positive values through an infinite discontinuity. This is a feature commonly associated with a phase transition of the second order. Just what *physical* difference there is in the black hole structure in the different phases is unclear. It is known (Bertin and Radicati 1976) that a similar type of phase transition occurs in Newtonian rotating fluids, for which the higher angular momentum phase represents a breakdown of axisymmetry.

The values of J^2 and e^2 at which the transition occurs are found to be αM^4 and βM^2 respectively, where α and β are positive roots of the equation:

$$\alpha^2+6\alpha+4\beta=3.$$

For an uncharged hole ($e=0$) this gives $J \simeq 0{\cdot}68\, M^2$ and for a non-rotating hole ($J=0$) it gives $e \simeq 0{\cdot}86\, M$. It is easy to show that these transitions occur when $\Omega \simeq 0{\cdot}23\ T$ and $\Phi = 3^{-1/2}$, respectively (Davies 1977a).

In the high J, e phase, there is a change in the stability characteristics of the black hole. It is possible for them to be in stable equilibrium with a (suitably rotating) surrounding heat bath of infinite volume. If the hole is non-rotating but carries a sufficiently large electric charge, its temperature can be made to approach zero as $e^2 \to M^2$. In this situation the hole cannot discharge itself by super-radiance, because in order to emit charged particles its temperature must be at least of the order of $T \simeq m_e c^2/k$ where m_e is the mass of the electron—the lightest charged particle. If its environment is cooled, it will slowly radiate mass, and hence increase the value of the ratio e^2/M^2 nearer to the extreme value of one. Moreover, as its specific heat is positive in this phase the hole cools down with its surroundings.

An intriguing question is: can it reach absolute zero, thereby violating the third law of thermodynamics, and open up the prospect that a small perturbation might cause the horizon to disappear altogether and expose a naked singularity? From equation (3.7) it is clear that the entropy approaches the finite value $M^2/8$ in the limit $e^2 \to M^2$. The fact that a zero-temperature black hole has a finite entropy rather than zero is easily understood in terms of statistical mechanics. In a laboratory thermodynamic system, the quantum state at absolute zero will be the lowest energy state of the system, and as such will be unique. The zero-temperature macrostate can therefore be realised by just one microstate, which yields an entropy of zero. In contrast, a black hole with $e^2 = M^2$ can clearly be realised by an enormous variety of internal microstates, so we expect it to possess finite entropy. It follows that black holes violate the third law as formulated by Planck. Moreover, the Nernst formulation of this law, that entropy differences between states that can be connected by an isothermal process must vanish as $T\to 0$, is also violated. For example, the quantity $(\partial S/\partial\Omega)_T\to -M^3$ as $T\to 0$. Hence the usual arguments about the unattainability of absolute zero do not seem to extend to black holes.

Nevertheless, it is not possible to argue from this that naked singularities *can* be produced by cooling down supercharged black holes. The law of cooling for such a hole is very complicated but approaches $T\propto t^{-1}$ as time $t\to\infty$. Hence, it will never attain absolute zero unaided in a finite time. On the other hand, the Reissner–Nordstrom solution of general relativity, on which this model of the charged black hole is based, is an idealisation (e.g. exact spherical symmetry, electrovac exterior) and it may be that a statistical fluctuation or some external perturbation could enable the limit (1.4) to be exceeded. All one can say is that there is no *thermodynamic* reason why not. There may, of course, be non-thermodynamic reasons why cosmic censorship cannot be violated. Certainly, simple dynamical mechanisms to increase J or e^2 beyond the black hole limit seem to fail for curious reasons. For example, if a massless subatomic particle with spin is shot into a black hole for which the limit (1.4) has already been attained, then it will impart a small additional angular momentum of about one unit. However, it will also increase M by a small amount. In order to 'fit' into the hole, a massless particle such as a neutrino, photon or graviton must have a minimum energy of M^{-1} so that its wavelength is less than M. The left-hand side of (1.4) then becomes:

$$\frac{(J+1)^2}{(M+1/M)^4}+\frac{e^2}{(M+1/M)^2}. \tag{3.16}$$

Using the fact that $J\gg 1$, $M\gg 1$, and that (1.4) is a first approximation, we may rewrite (3.16) as:

$$1+\frac{2(J+e^2)}{M^4}-\frac{4}{M^2}$$

which is always *less* than one, because $J<M^2$ and $e^2<M^2$. Thus, in absorbing the particle, the hole acquires insufficient angular momentum per unit mass to cool below the thermodynamic limit $T=0$. It actually heats up instead. Of course, this conclusion may be invalidated if there exist in nature massless fields of very large spin.

4. Two-dimensional models

The radial modes for a field in the vicinity of a black hole cannot be written down in terms of known functions. This shortcoming obscures many of the features of the thermal emission process because numerical, rather than analytic, techniques must be applied. Fortunately, most of the qualitative features of quantum black holes are present in the two-dimensional analogues, which do possess explicit solutions, and study of these models has provided great insight into the nature of the thermal radiation.

4.1. The relation between coordinates and quantum states

A spherically symmetric black hole with mass M and electric charge e is described by the space–time metric:

$$ds^2 = \left(1 - \frac{2M}{r} + \frac{e^2}{r^2}\right) dt^2 - \left(1 - \frac{2M}{r} + \frac{e^2}{r^2}\right)^{-1} dr^2 - r^2 (d\theta^2 + \sin^2\theta \, d\phi^2) \qquad (4.1)$$

where $M^2 > e^2$ and r is a radial coordinate chosen to make the surface area of a sphere of radius r equal to $4\pi r^2$, as in Minkowski space. The metric (4.1) is evidently singular at $r = r_\pm \equiv M \pm (M^2 - e^2)^{1/2}$. These are not singularities in the geometry itself but in the coordinate system (t, r), similar to those which occur to latitude and longitude on the surface of a sphere at the poles. In fact, the outer surface r_+ corresponds to the event horizon: notice that $r_+ \to 2M$ as $e \to 0$. This surface has global significance, but locally an inertial observer would find nothing unusual about the space–time geometry there. The inner surface r_- is another type of horizon inside the hole itself which need not concern us here.

The physics of the Hawking process depends in essence on the geometry of the (r, t) surface; the angular dependence, which enters through the final term of equation (4.1), is really only incidental. A natural two-dimensional model of this black hole is therefore obtained by suppressing the angular dependence (removing the final term) and treating one-half $(r > 0)$ of two-dimensional space–time with coordinates (t, r). If necessary, the origin of the spherical coordinates, $r = 0$, can be modelled by assuming a perfectly reflecting barrier at this point, or by reflecting the geometry in the origin. It is not possible to include the effects of rotation in this model as a black hole with $J \neq 0$ is not spherically symmetric.

A simple change of coordinates:

$$r^* = r + \frac{r_+^2}{r_+ - r_-} \ln (r - r_+) - \frac{r_-^2}{r_+ - r_-} \ln (r - r_-)$$

converts the two-dimensional residue of (4.1) into

$$ds^2 = \left(1 - \frac{2M}{r} + \frac{e^2}{r^2}\right) (dt^2 - dr^{*2}). \qquad (4.2)$$

For some purposes it is convenient to work with the so-called null coordinates u, v, defined by:

$$u = t - r^*$$

$$v = t + r^*$$

whence (4.2) reduces to

$$ds^2 = \left(1 - \frac{2M}{r} + \frac{e^2}{r^2}\right) du \, dv. \qquad (4.3)$$

The advantage of the forms (4.2) and (4.3) lies in the fact that they are *conformal* to Minkowski space. That is, a transformation of the form $ds^2 \to \Omega(t, r) \, ds^2$, for some space–time function Ω, converts the metric to the flat form $du \, dv$. The significance of conformal flatness is that the wave equation for a massless scalar field:

$$\Box \phi = 0 \qquad (4.4)$$

is *invariant* under conformal transformations. This property is also true for the massless spinor equation. Thus, in these coordinates, equation (4.4) is simply the same as the ordinary flat space form:

$$\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial r^{*2}}\right) \phi = \frac{\partial^2 \phi}{\partial u \, \partial v} = 0 \qquad (4.5)$$

and we may immediately write down a complete set of normalised mode solutions in the standard way:

$$\psi_\omega^{\rightleftharpoons}(t, r^*) = \frac{1}{2\sqrt{\pi\omega}} \exp(-i\omega u) \qquad \frac{1}{2\sqrt{\pi\omega}} \exp(-i\omega v). \tag{4.6}$$

These complex exponentials represent right ($\rightarrow$) and left ($\leftarrow$) moving waves, respectively. They may be used to construct a Fock space and vacuum state for the quantised theory in the fashion described in §2.1.

We now come to an important point. Equation (4.3) has the general form:

$$ds^2 = C(u, v)\, du\, dv. \tag{4.7}$$

There is a Fock space, and in particular a vacuum state, associated with the coordinates u and v through the choice of modes (4.6). However, under a *coordinate* transformation:

$$\begin{aligned} u &\to \bar{u}(u) \\ v &\to \bar{v}(v) \end{aligned} \tag{4.8}$$

the metric changes to:

$$ds^2 = \bar{C}(\bar{u}, \bar{v})\, d\bar{u}\, d\bar{v} \tag{4.9}$$

where $\bar{C}(\bar{u}, \bar{v}) = C(u, v)\,(du/d\bar{u})\,(dv/d\bar{v})$. But the transformed metric (4.9) is still conformally flat and possesses mode solutions:

$$\bar{\psi}_\omega^{\rightleftharpoons} = \frac{1}{2\sqrt{\pi\omega}} \exp(-i\omega\bar{u}) \qquad \frac{1}{2\sqrt{\pi\omega}} \exp(-i\omega\bar{v}) \tag{4.10}$$

exactly like (4.6) but in the new coordinates $\bar{u}$, $\bar{v}$. Indeed, for every choice of coordinates there exists a set of standard exponential modes, and hence a different set of states. In particular, each coordinate system possesses its own vacuum state.

When it comes to the physics of black hole thermal emission we have to face the question: what is the quantum state in which the field observables are to be evaluated as expectation values? Which coordinate system do we choose for our mode solutions? It is important to realise that the choice of coordinates in which to actually *calculate* is irrelevant—as required by general covariance. The significance of a particular choice is that the modes should reduce to simple exponential form like (4.6) or (4.10).

As always, the choice of quantum state depends on the physical circumstances of interest. For example, suppose we choose the modes (4.6) based on the standard coordinates u and v (or t and r) used in equation (4.3). Far from the black hole the metric (4.3) approaches that of Minkowski space and the coordinates t, r are the usual time and space coordinates of special relativity. Thus the modes (4.6) become standard exponential wave solutions of conventional quantum field theory. On the other hand, near the horizon, as $r \to r_+$, $r^* \to -\infty$ and $u \to \infty$ so that the ($\rightarrow$) modes oscillate infinitely fast. This indicates some serious pathology in the quantum mechanics of any states built out of these modes, and indeed it can be shown (see §4.6) that the expectation values in these states of various field observables, such as the local energy density measured by an inertial observer, actually diverge at r_+.

The implication of this bad behaviour is that quantum states based on (4.6) could not be realised for a black hole: back-reaction effects of the diverging field energy would in practice drastically modify the geometry near the horizon. On the other hand, these states could be realised in the region exterior to a static star whose radius is much larger than the horizon because the metric (4.3) would then only apply down as far as the surface of the star. Inside the star the modes would change to some new well-behaved functions. (If the star is not static then the outgoing modes will no longer be simple exponentials because of red-shift effects: see §2.2.) In many ways the vacuum state based on these modes would seem to be a natural generalisation to a static star of the Minkowski vacuum, to which it reduces at large r.

4.2. *The eternal black hole*

The unsuitability of the modes (4.6) for a black hole is directly related to the behaviour of the u or r^* coordinate at the horizon r_+. As explained, this is not a physical singularity in the geometry, a fact which may be seen by computing the space–time curvature for the geometry (4.3):

$$R=\frac{4M}{r^3}\left(1-\frac{3}{2}\frac{e^2}{Mr}\right) \tag{4.11}$$

which is clearly non-singular at $r=r_+$. The singularity is, therefore, a feature of the coordinate system only and a change of coordinates will remove it. Field modes based on new coordinates $\bar{u}$, $\bar{v}$ which are regular at r_+ will not then give rise to divergent quantities at r_+ in the quantum field theory.

The standard analytic extension of the space–time (4.3) into the interior of the black hole, i.e. across the horizon $r=r_+$, uses new coordinates $\bar{u}$, $\bar{v}$ defined in terms of u and v by the transformation:

$$\begin{aligned}\bar{u}&=\exp\left[-\left(\frac{r_+-r_-}{2r_+^2}\right)u\right]=\exp(-\kappa u)\\ \bar{v}&=\exp\left[\left(\frac{r_+-r_-}{2r_+^2}\right)v\right]=\exp(\kappa v)\end{aligned} \tag{4.12}$$

where κ is the surface gravity of the charged black hole, being the same quantity which is related to the Hawking temperature through equations (1.12) and (3.4). (The latter follows by inserting the definitions of $r_\pm$ in κ.) The appearance of an exponential in (4.12) is directly related to the exponential red shift which is the cause of the thermal radiation.

In the $\bar{u}$, $\bar{v}$ coordinates the metric (4.3) is transformed to:

$$ds^2=-\frac{4r_+^4}{r^2}\exp(-2\kappa r)(r_+-r_-)^{-2}(r-r_-)^{\alpha}\,d\bar{u}\,d\bar{v} \tag{4.13}$$

where $\alpha=(r_+^2+r_-^2)/r_+^2$. The metric (4.13) is clearly non-singular at $r=r_+$.

Having extended the space–time analytically inside the black hole, it is necessary to ask where it will end. If a collapsing star is encountered inside the hole, then the geometry will no longer remain that of (4.13) and we should have to match this metric onto some other metric which describes the interior of the star. However, the thermal properties of black holes seem to be completely independent of the details of the actual gravitational collapse, so great simplification may be achieved if the imploding star is omitted from the picture altogether, and the empty space–time (4.13), often called an eternal black hole, is used as the background on which to examine our quantum fields. In the next section this will be justified by comparing the energy flux emitted by a star which implodes into a black hole with that from an eternal black hole.

Whether or not there exist in the universe any eternal black holes, or whether all black holes form from imploding stars, is not known.

Although the geometry of an eternal black hole is identical to that of the vacuum region outside an imploding star, the *topology* is different. The reason for this is that, in the absence of the star, the space–time described by (4.13) must continue on without end through the interior of the hole, unless it terminates at a physical singularity. Such a singularity does indeed form a boundary to the space–time inside the black hole at $r=0$. It can be seen from (4.11) and (4.13) that not only does the metric become singular here but the curvature does also. We cannot continue the space–time beyond this point. However, it turns out that the singularity does not block off all the space–time inside the hole.

An observer in the exterior universe at fixed distance from the centre moves along a line of constant $r > r_+$. From equation (4.12) we see that:

$$\bar{u}\bar{v} = \exp(2\kappa r^*)$$

so that a line of constant r^* (or r) is a rectangular hyperbola in the $\bar{u}$, $\bar{v}$ plane. It follows that there exists *another* rectangular hyperbola, obtained by reflecting the first in the origin ($u \to -\bar{u}$, $\bar{v} \to -\bar{v}$). Consequently, there is a whole new region of space–time—a sort of mirror universe—with the same metric (4.3), existing on the 'other side' of the black hole, joined onto our universe through the inside of the hole (see figure 3). If the hole had formed from an imploding star instead, this mirror universe would not exist.

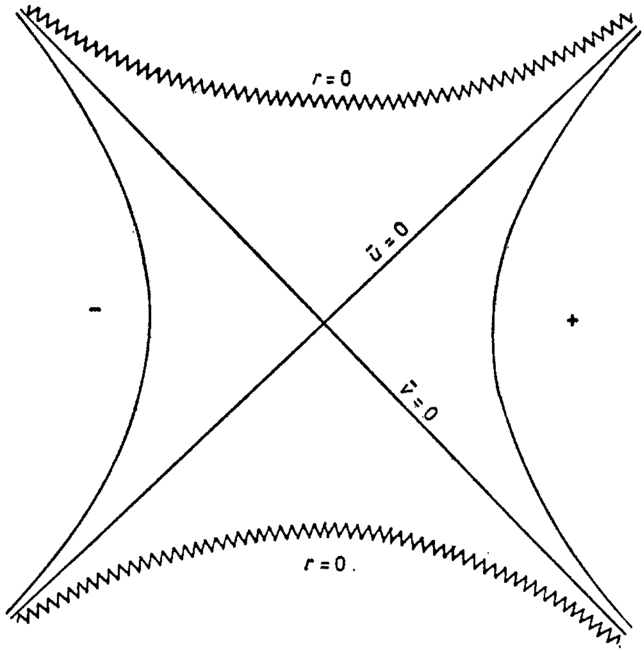

Figure 3. Eternal uncharged black hole. This space–time diagram shows the causal structure of the extended time-symmetric hole. The space–time is bounded above and below by the $r=0$ singularity, but there is a space-like throat joining our universe (wedge-shaped region marked +) to a mirror universe (marked −) through the interior of the hole. The null ray $\bar{u}=0$ is the event horizon for our universe. The hyperbolae represent the world lines of inertial observers far from the hole. In the charged case the space–time can also be extended vertically.

The mirror universe cannot be reached by an observer falling into the black hole from our universe, for it is space-like-separated from us. However, so long as the charge of the black hole is non-zero, the space–time can be continued in a time-like direction also, without hitting any singularities and opens out into yet another universe that way. In fact, the continuation can be repeated *ad infinitum* into an unlimited sequence of other universes. Moreover, this time-like extension can also exist inside a black hole formed from an imploding, electrically charged star. The question of whether an observer could really travel through a charged black hole into these other universes is another matter (Birrell and Davies 1978, Simpson and Penrose 1973, McNamara 1978).

There is a further crucial difference between an eternal black hole and the imploding star. Because the former is static and everywhere empty, it is time-symmetric. There will be a past horizon and past singularity which are the time reversals of the future event horizon and future singularity. Sometimes the time-reversed hole is called a white hole.

4.3. *Thermal Green functions*

We now have two different sets of coordinates which cover the eternal (maximally extended) black hole: the original 'Schwarzschild-like' coordinates u and v defined by (4.3) and which cover 'our' universe and the 'mirror' universe in two separate patches, and we have the so-called 'Kruskal-like' coordinates $\bar{u}$, $\bar{v}$ defined in terms of u and v through (4.12) which cover the whole space in one patch. Associated with each coordinate system is a complete set of modes, given by (4.6) for the u, v system and by (4.10) for the $\bar{u}$, $\bar{v}$ system.

The scalar field ϕ may be expanded in terms of either set:

$$\phi=\sum_{\omega}[a_\omega\bar{\psi}_\omega^{\rightarrow}+a_\omega^*(\bar{\psi}_\omega^{\rightarrow})^*]+[\leftarrow] \tag{4.14}$$

$$\phi=\sum_{\omega}[c_\omega{}^{(+)}\psi_\omega^{\rightarrow}+c_\omega^*({}^{(+)}\psi_\omega^{\rightarrow})^*+d_\omega{}^{(-)}\psi_\omega^{\rightarrow}+d_\omega^*({}^{(-)}\psi_\omega^{\rightarrow})^*]+[\leftarrow] \tag{4.15}$$

$\bar{\psi}$ referring to the set associated with the barred coordinates, (4.12). The symbol $[\leftarrow]$ denotes a similar expression for left-moving waves of the form $\exp(-i\omega v)$. In (4.15) ϕ has been expanded in terms of *two* sets of ψ_ω modes denoted by ${}^{(\pm)}\psi_\omega$. The $(+)$ set refer to those that cover the region outside the black hole in our universe, and the $(-)$ set refer to the mirror universe. Note that ${}^{(+)}\psi_\omega^{\rightleftarrows}$ vanish in the exterior mirror universe, and ${}^{(-)}\psi_\omega^{\rightleftarrows}$ vanish in our universe.

With each set of modes is associated a vacuum state. The first, which we denote by $|\bar{0}\rangle$, is defined by:

$$a_\omega|\bar{0}\rangle=0. \tag{4.16}$$

The second, denoted by $|0\rangle$, is defined by:

$$c_\omega|0\rangle=d_\omega|0\rangle=0. \tag{4.17}$$

These are not the same state. For example:

$$c_\omega|\bar{0}\rangle\neq 0.$$

The ψ_ω modes are related to the $\bar{\psi}_\omega$ modes through a transformation of the type (2.10) which mixes positive and negative frequencies. The operators a_ω, c_ω and d_ω will be related by a Bogolubov transformation of the type (2.9), so as $\beta_{\omega\omega'} \neq 0$ the vacuum state $|\bar{0}\rangle$ will contain quanta of the modes ψ_ω and vice versa.

It is a straightforward matter to calculate what quanta are present in the ψ_ω modes when the system is in the vacuum state $|\bar{0}\rangle$. This may be done directly by explicitly expanding one set of modes in terms of the other according to a relation like (2.10). However, the thermal properties of black holes are better displayed by exploiting quicker and more elegant routes.

First, note that the coordinates $\bar{u}$ and $\bar{v}$ do not reduce to ordinary Minkowski coordinates $t-r$, $t+r$ at large r (far from the black hole where space–time is flat). On the other hand the u, v coordinates do. Hence we conclude that the vacuum state $|\bar{0}\rangle$ is not the conventional vacuum state of ordinary flat space quantum field theory even in the region distant from the black hole. We must therefore expect that a black hole whose quantum state is $|\bar{0}\rangle$ will be regarded by an ordinary observer far from the hole as being immersed in a bath of quanta.

Perhaps the easiest way to explore the nature of this bath of quanta is to compare the so-called two-point, or Green, functions:

$$G(x'', x') = \langle 0 | \phi(x'')\phi(x') | 0 \rangle \tag{4.18}$$

$$\bar{G}(x'', x') = \langle \bar{0} | \phi(x'')\phi(x') | \bar{0} \rangle \tag{4.19}$$

computed in the two different vacuum states. A simple mode-sum calculation gives:

$$G(x'', x') = \frac{1}{4\pi} \ln \Delta u \, \Delta v \tag{4.20}$$

$$\bar{G}(x'', x') = \frac{1}{4\pi} \ln \Delta \bar{u} \, \Delta \bar{v} \tag{4.21}$$

for the respective vacua, where $\Delta u = u'' - u'$, etc. In arriving at (4.16) and (4.17) it has been necessary to discard an infinite constant. This is a manifestation of an infrared divergence which is always present in these two-point functions in two-dimensional massless scalar field theory. The presence of the infinite constant is not a serious problem as it disappears when G is differentiated to obtain observable quantities such as the stress tensor (see §4.6).

The relation between the two G (Gibbons and Perry 1976, Dowker 1977) is revealed by substituting for $\bar{u}$ and $\bar{v}$ in (4.21) from the transformation equations (4.12):

$$\bar{G}(x'', x') = \frac{1}{4\pi} \{\kappa(r^{*''} + r^{*'}) + \ln 2 \, [\cosh \kappa(t'' - t') - \cosh \kappa(r^{*''} - r^{*'})]\}. \tag{4.22}$$

It is seen that $\bar{G}$ is unchanged under the replacement $t'' \to t'' + 2\pi i n \kappa^{-1}$, n integer. That is, $\bar{G}$ is periodic in *imaginary* time, with period $2\pi/\kappa$:

$$\bar{G}(t'' + 2\pi i n \kappa^{-1}, r^{*''}; t', r^{*'}) = \bar{G}(t'', r^{*''}; t', r^{*'}). \tag{4.23}$$

This feature is characteristic of *thermal* Green functions. That is, if $\bar{G}$ is averaged over a grand canonical ensemble to give:

$$\bar{G}_\beta(x'', x') = \mathrm{Tr}\,[\exp(-\beta H)\phi(x'')\phi(x')]/\mathrm{Tr}\,\exp(-\beta H) \tag{4.24}$$

corresponding to thermal equilibrium at temperature $kT = \beta^{-1}$, where H is the Hamiltonian (defined with respect to the time coordinate t), one readily finds from the Heisenberg equations of motion:

$$\bar{G}_\beta(t'', r''; t', r') = \bar{G}_\beta(t', r'; t''+\mathrm{i}\beta, r'').$$

Observing the symmetry of (4.22) under interchange of primes, we can identify $\bar{G}$ as a thermal Green function at a temperature $T=\kappa/2\pi k$, which is precisely the Hawking temperature for the black hole (see equation (3.4) with $k=(8\pi)^{-1}$).

Next we note that the zero-temperature Green function ($\beta \to \infty$ or $\kappa \to 0$) is:

$$\bar{G}_\infty = \frac{1}{4\pi} \ln [(t''-t')^2 - (r^{*''} - r^{*'})^2]$$

$$= \frac{1}{4\pi} \ln \Delta u \, \Delta v \tag{4.25}$$

which, from (4.20), reveals that:

$$\bar{G}_\infty = G. \tag{4.26}$$

Using the identity:

$$\cosh x - \cosh y = \tfrac{1}{2}(y^2 - x^2) \prod_{n=1}^{\infty} \left(1 + \frac{x^2}{(2n\pi + \mathrm{i}y)^2}\right) \times \left(1 + \frac{x^2}{(2n\pi - \mathrm{i}y)^2}\right) (4n^2\pi^2 + y^2)^2 (2n\pi)^{-4}$$

and discarding another infinite constant, we readily obtain:

$$\ln [2(\cosh x - \cosh y)] = \sum_{n=-\infty}^{\infty} \ln [(y + 2\pi n\mathrm{i})^2 - x^2]$$

which enables (4.22) to be written, using (4.25):

$$\bar{G}(t'', r''; t', r') = \sum_{n=-\infty}^{\infty} \bar{G}_\infty(t'' + 2\pi n\mathrm{i}\, \kappa^{-1}, r''; t', r') \tag{4.27}$$

as an infinite series of zero-temperature images.

These properties of the Green function $\bar{G}$ reveal that the vacuum state $|\bar{0}\rangle$ behaves in the region outside the black hole, where the coordinates u, v (t, r) are more appropriate, like a bath of thermal equilibrium radiation at the Hawking temperature $\kappa/2\pi k$. In arriving at this conclusion we have nowhere used the implosion of a star, or the tracing of null rays through a collapsing object, as did Hawking in his original derivation. This illustrates the fact that the Hawking effect owes its existence primarily to the space–time structure associated with the black hole and not to the details of the hole's formation. Indeed, in the original ray-tracing argument, outlined in §2.2, the final answer turned out to be independent of the details of the imploding star in the late-time limit when the flux settles down to a thermal spectrum. This very basic nature of the Hawking effect accounts for the remarkably large number of different ways in which it has been derived.

Further information may be obtained about the $|\bar{0}\rangle$ vacuum by computing the expectation value of the stress tensor $T_{\mu\nu}$ of the quantum field in this state. As will be discussed in §4.6, this quantity is formally infinite, but the difference $\langle\bar{0}|T_{\mu\nu}|\bar{0}\rangle - \langle 0|T_{\mu\nu}|0\rangle$ is finite. The expectation value for $T_{\mu\nu}$ in any given vacuum state may be obtained by differentiating the corresponding two-point function G (see equation

(4.54)). For example, for the u–u component:

$$\langle \bar{0} | T_{uu} | \bar{0} \rangle - \langle 0 | T_{uu} | 0 \rangle = \lim_{x'', x' \to x} \partial_{u''} \partial_{u'} [\bar{G}(x'', x') - G(x'', x')]. \qquad (4.28)$$

From equation (4.27) we now see that $G(x'', x')$ is just the $n=0$ term in the expansion of $\bar{G}$. Therefore we may insert for $\bar{G}-G$ in (4.28) the summation (4.27) with the $n=0$ term removed. One readily obtains the result:

$$\kappa^2/48\pi \qquad (4.29)$$

which corresponds exactly to the energy density of a bath of thermal (Planckian) radiation at a temperature $\kappa/2\pi k$.

The energy flux $\langle T_{rt} \rangle$ can be calculated similarly. However, we merely need to note that the model is time-symmetric to deduce that the net flux vanishes. The temperature $\kappa/2\pi k$ is, of course, the temperature of the black hole as calculated by Hawking. Thus, we see that the $|\bar{0}\rangle$ vacuum corresponds physically to a black hole in *thermal equilibrium* with a surrounding heat bath. The hole does not evaporate: it emits and absorbs thermal radiation at an equal rate.

4.4. *Relating the vacuum states*

In the previous subsection it was suggested that the vacuum $|\bar{0}\rangle$ appeared as a *thermal state* as far as quanta of the modes ψ_ω were concerned. We shall now demonstrate this explicitly by calculating the unitary transformation from one vacuum state to the other:

$$|\bar{0}\rangle = U|0\rangle \qquad (4.30)$$

where U is a unitary operator to be found. This equation will enable us to compute such things as the expected number of quanta in each mode ψ_ω when the black hole is in the state $|\bar{0}\rangle$.

As already remarked, relation (4.30) could be deduced by expanding the modes ψ_ω in terms of $\bar{\psi}_\omega$ by brute force. However, we do not need to do this for the following reason. The state $|\bar{0}\rangle$ is the vacuum of the $\bar{\psi}_\omega$ modes, but it is also the vacuum of any linear superposition of these modes which does not mix up positive and negative frequencies (for then $\beta_{\omega\omega'}=0$). There is a very easy way of characterising such pure positive-frequency superpositions, namely that they are *analytic* functions of $\bar{u}$, bounded in the lower-half complex $\bar{u}$ plane. (This is clearly true for $\exp(-i\omega\bar{u})$.) Thus, finding U in (4.30) reduces to finding a suitable linear combination of the ψ_ω modes which is bounded analytic in the lower-half $\bar{u}$ plane. As it happens, this is almost trivial. Restricting for the moment to right-moving waves ($\rightarrow$) only, we note from (4.6) and (4.12) that, in $\bar{u}$ coordinates:

$$\begin{aligned} {}^{(+)}\psi_\omega{}^{\rightarrow} &= \frac{1}{2\sqrt{\pi\omega}} \exp(i\omega\kappa^{-1}\ln\bar{u}) && \bar{u}>0 \\ &= 0 && \bar{u}<0 \end{aligned} \qquad (4.31)$$

$$\begin{aligned} {}^{(-)}\psi_\omega{}^{\rightarrow} &= 0 && \bar{u}>0 \\ &= \frac{1}{2\sqrt{\pi\omega}} \exp[i\omega k^{-1}\ln(-\bar{u})] && \bar{u}<0 \end{aligned} \qquad (4.32)$$

$\bar{u}>0$ corresponding to 'our' universe and $\bar{u}<0$ to the 'mirror' universe. Although the discontinuity at $\bar{u}=0$ prevents the individual ${}^{(\pm)}\psi_\omega$ modes being analytic, the sum of the two would be analytic if it were not for the minus sign in the argument of the logarithm in ${}^{(-)}\psi_\omega$. However, factoring this out gives $\exp(\pm\pi\omega\kappa^{-1})$, so the linear combination:

$$ {}^{(+)}\psi_\omega^{\rightarrow} + \exp(-\pi\omega\kappa^{-1})\,{}^{(-)}\psi_\omega^{\rightarrow} \tag{4.33}$$

is analytic. (The sign of the exponent in (4.33) is determined by choosing the branch cut for the logarithm to lie in the upper-half $\bar{u}$ plane, so as not to spoil the required analyticity in the lower-half plane.)

We shall call modes like (4.33) $\Psi_\omega^{\rightarrow}$. Inserting the correct normalisation factors:

$$\Psi_\omega^{\rightarrow} = \cosh\phi_\omega\,{}^{(+)}\psi_\omega^{\rightarrow} + \sinh\phi_\omega\,{}^{(-)}\psi_\omega^{\rightarrow} \tag{4.34}$$

where $\tanh\phi_\omega = \exp(-\pi\omega\kappa^{-1})$. Because the modes Ψ_ω are analytic in the lower-half $\bar{u}$ plane, they will be a pure positive-frequency superposition of $\bar{\psi}_\omega$ modes and so possess the vacuum state $|\bar{0}\rangle$. Thus, the required superposition (4.34) of ψ_ω modes is very simple and does not even involve different frequencies ω, which makes it easy to relate the two vacuum states. This construction is due originally to Unruh (1976).

If ϕ is expanded in terms of Ψ_ω:

$$\phi = \sum_\omega [b_\omega \Psi_\omega^{\rightarrow} + b_\omega^*(\Psi_\omega^{\rightarrow})^*] + [\leftarrow] \tag{4.35}$$

then, as stated:

$$b_\omega|\bar{0}\rangle = 0. \tag{4.36}$$

It follows from (4.15), (4.34) and (4.35) that

$$b_\omega = \cosh\phi_\omega c_\omega - \sinh\phi_\omega\, d_\omega. \tag{4.37}$$

This simple transformation can also, with the help of the commutation relations for the c_ω and d_ω, be written:

$$b_\omega = \exp(-\mathrm{i}G)\, c_\omega \exp(\mathrm{i}G) \tag{4.38}$$

where

$$G = \sum_\omega \mathrm{i}\phi_\omega(c_\omega^* d_\omega^* - c_\omega d_\omega). \tag{4.39}$$

Now from (4.36) and (4.38):

$$\exp(\mathrm{i}G)\, b_\omega|\bar{0}\rangle = c_\omega \exp(\mathrm{i}G)|\bar{0}\rangle = 0. \tag{4.40}$$

But we know by definition that $c_\omega|0\rangle = 0$, so (4.40) provides the relation:

$$|0\rangle = \exp(\mathrm{i}G)|\bar{0}\rangle$$

which when inverted gives (4.30):

$$|\bar{0}\rangle = \exp(-\mathrm{i}G)|0\rangle. \tag{4.41}$$

If the exponential is expanded, the $c_\omega d_\omega$ operators will all give zero when acting on $|0\rangle$. Using this and the commutation rules the exponent can be rearranged to give:

$$\begin{aligned}|\bar{0}\rangle &= \exp\Big(\sum_\omega -\ln\cosh\phi_\omega + \tanh\phi_\omega c_\omega^* d_\omega^*\Big)\,|0\rangle \\ &= \prod_\omega (\cosh\phi_\omega)^{-1} \sum_{n=0}^{\infty} \exp(-n\pi\omega\kappa^{-1}) \frac{(c_\omega^*)^n}{\sqrt{n!}} \frac{(d_\omega^*)^n}{\sqrt{n!}}\,|0\rangle \\ &= \prod_\omega (\cosh\phi_\omega)^{-1} \sum_{n=0}^{\infty} \exp(-n\pi\omega\kappa^{-1})|n_\omega^{(+)}\rangle\,|n_\omega^{(-)}\rangle \end{aligned} \tag{4.42}$$

where the state $|n_\omega^{(+)}\rangle\ |n_\omega^{(-)}\rangle$ corresponds to n_ω quanta in the mode ψ_ω in our universe and another n_ω quanta in the corresponding mode in the mirror universe.

We now ask the question: what is the expectation value of an observable A as measured by an observer who remains in our universe, when the quantum state of the black hole is $|\bar{0}\rangle$? The operator $\hat{A}$ corresponding to such an observable, being restricted to our universe, leaves $|n_\omega^{(-)}\rangle$ unaffected, so this part of the state just factors out to give an overall factor of one in $\langle\bar{0}|\hat{A}|\bar{0}\rangle$. Suppose, for example, we choose $\hat{A}$ to be the number operator $N_\sigma \equiv c_\sigma^* c_\sigma$ for quanta in our universe in the mode $\omega=\sigma$. Then $N_\sigma|n_\sigma^{(+)}\rangle = n_\sigma|n_\sigma^{(+)}\rangle$ and we note that every $\omega\neq\sigma$ simply contributes a factor of one in the product Π_ω in the expectation value $\langle\bar{0}|N_\sigma|\bar{0}\rangle$. The only factor which is different from unity is at the σ frequency. Thus:

$$\langle\bar{0}|N_\sigma|\bar{0}\rangle = (\cosh\phi_\sigma)^{-2}\sum_{n=0}^{\infty} n_\sigma \exp(-2n\pi\sigma\kappa^{-1}) \tag{4.43}$$

where we have also used the orthogonality between states of different occupation number n to reduce the double summation to a single summation.

Noting that:

$$(\cosh\phi_\sigma)^{-2} = 1-\exp(-2\pi\sigma\kappa^{-1}) = \left[\sum_{n=0}^{\infty}\exp(-n\pi\sigma\kappa^{-1})\right]^{-1}$$

we can write (4.43) as:

$$\langle\bar{0}|N_\sigma|\bar{0}\rangle = \sum_{n=0}^{\infty} n_\sigma \exp(-\beta E_n)\left[\sum_{n=0}^{\infty}\exp(-\beta E_n)\right]^{-1} \tag{4.44}$$

where $\beta = 2\pi\kappa^{-1}$ and E_n is the energy of the nth mode.

In the general case (4.44) is replaced by:

$$\langle\bar{0}|A_\sigma|\bar{0}\rangle = \sum_{n=0}^{\infty} \alpha_\sigma \exp(-\beta E_n)\left[\sum_{n=0}^{\infty}\exp(-\beta E_n)\right]^{-1} \tag{4.45}$$

where α_σ is the σ-mode eigenvalue of $\hat{A}_\sigma$. Equation (4.45) has the form of a thermal average over the eigenvalues α_σ.

For the specific case of the number operator, the summations in (4.44) are immediately evaluated to give:

$$\langle\bar{0}|N_\sigma|\bar{0}\rangle = \frac{1}{\exp(\beta\sigma)-1} \tag{4.46}$$

which is a Planck thermal equilibrium spectrum at temperature:

$$T = (k\beta)^{-1} = \frac{\kappa}{2\pi k} \tag{4.47}$$

which is again the Hawking temperature for the black hole. We conclude that the vacuum state $|\bar{0}\rangle$ corresponds to a bath of thermal radiation at the Hawking temperature in the region exterior to the black hole.

This is not all. If we ask for the probability that the vacuum $|\bar{0}\rangle$ contains n_1 particles in ψ_ω mode 1, n_2 in mode 2, etc, we obtain from (4.42):

$$|\langle n_1, n_2, \ldots, n_\omega, \ldots|\bar{0}\rangle|^2 = \prod_\omega \exp(-2n_\omega\pi\omega\kappa^{-1})[1-\exp(-2\pi\omega\kappa^{-1})] \tag{4.48}$$

$$= \prod_\omega P(n_\omega) \tag{4.49}$$

where $P(n_\omega)$ is the probability that the mode ω contains n_ω particles. From the form of (4.49) it is clear that the probabilities for the excitation of different modes are all independent. This implies that the quanta are emitted at random and the state of the field outside the black hole is a mixture, which must in general be described by a (diagonal) density matrix $\Pi_\omega\ \delta_{n_\omega n_\omega'} P(n_\omega)$. This has a clear physical interpretation. An observer who is restricted to the exterior of the black hole can have no information about the state of field in the mirror universe, hidden from him by the event horizon. He must therefore trace over the modes ${}^{(-)}\psi_\omega$ and use a density matrix for the accessible modes ${}^{(+)}\psi_\omega$.

These results are based on work due to Israel (1976). They are further strengthened by an analysis by Unruh (1976) of the way in which a model particle detector would respond to the $|\bar{0}\rangle$ vacuum if constrained to move along the non-inertial path $r=$ constant (i.e. at a fixed radial distance) outside the hole. He finds that the detector absorbs energy in precisely the same way that it would if immersed in a bath of thermal radiation at the temperature $\kappa/2\pi k$.

All the results of this section have so far been restricted to the right-moving ($\rightarrow$) modes which in the exterior region in our universe describe particles which travel away from the vicinity of the hole out towards infinity. A distant inertial observer would interpret the vacuum state $|\bar{0}\rangle$ as a steady flux of thermal radiation with the Hawking temperature $\kappa/2\pi k$ being emitted by the hole. Identical consideration can be given to the left-moving ($\leftarrow$) modes, which will contribute another similar term to G in (4.39). The analysis of these modes proceeds in the same fashion as above and (inevitably, due to the inherent time symmetry of the space–time and the quantum state) leads to an equal and opposite, ingoing, flux of thermal radiation. Thus, there is no net transfer of energy between the black hole and its environment. The state $|\bar{0}\rangle$ therefore describes a black hole immersed in a bath of thermal radiation, in thermodynamic equilibrium at a temperature $\kappa/2\pi k$.

We could also describe the Hawking evaporation process by deliberately making the state time-asymmetric, i.e. choosing a vacuum state in which only the right-moving modes were of the form $\bar{\psi}_\omega$ but the left-moving modes were chosen to be of the type ${}^{(\pm)}\psi_\omega$. There would then be no flux of particles travelling towards the black hole from great distance in our universe—only an outgoing flux. This vacuum state was suggested by Unruh (1976) as a model of the Hawking process in which the effect of the imploding star is reproduced by choosing an appropriate vacuum state in the eternal black hole. Similarly, the state $|\bar{0}\rangle$ is a model of a black hole formed from an imploding star which is then placed in a heat bath at the same temperature. Although the topology of the space–time associated with an imploding star is different from that of the eternal black hole (there is no mirror universe), nevertheless the event horizon is still present and the behaviour of the quantum field in the exterior region is the same.

In dealing with the full four-dimensional case, most of the details of the present treatment go through with little change. The main difference is the presence of back-scattering—the outgoing quanta have a certain probability of being scattered by the gravitational field around the hole and being deflected back into it. There is an equal probability of ingoing quanta being reflected out again. This means that the thermal equilibrium nature of the $|\bar{0}\rangle$ state is unchanged, but as already mentioned the spectrum of the radiation is no longer Planckian: equation (4.46), for example, would contain a frequency-dependent transmission factor for the probability of the emitted quantum reaching infinity. There is also the possibility of including the effect of rotation of the hole. This enters into the results rather like a chemical potential.

The reason that the thermal properties of black holes formed by an imploding star are so well described by the simplified eternal black hole model, and that the details of the implosion do not enter into the final result, can be seen by inspecting the form of the modes in the two treatments. The modes of the field which propagate through the imploding body and out again emerge redshifted. During the early stages of the implosion, the form of the outgoing modes will depend on how fast the collapse of the star proceeds, the internal structure, and so on. However, as the star nears the event horizon, these details get washed out and the modes settle down to the simple form given by equation (2.17). This is precisely the radial equivalent of the modes $\bar{\psi}_\omega$, so it is no surprise that the vacuum state $|\bar{0}\rangle$ corresponding to $\bar{\psi}_\omega$ correctly describes the black hole end state which follows the late stages of the collapse. We can regard the implosion phase when the radiation spectrum is non-thermal as the approach to thermodynamic equilibrium and the asymptotic, static, end state represented by the black hole as the equilibrium condition.

4.5. *Accelerated observers in Minkowski space*

The association of sets of modes with coordinate systems discussed in the preceding subsection is not a connection with relevance only to curved space–time and black holes. Even in Minkowski space we can solve the field equation in any coordinate system we choose. If a consistent quantisation of the field is carried out in modes constructed from a coordinate system which is different from the usual (Cartesian) Minkowski coordinates then the usual vacuum state of conventional quantum field theory will *not* be the vacuum state of these other modes, i.e. the Minkowski vacuum will contain quanta of the non-standard modes. In particular, if the new coordinates are related to the Minkowski coordinates by a transformation of the type (4.12), with the Minkowski vacuum corresponding to $|\bar{0}\rangle$, then we would expect quanta of the other modes to be present in the form of a thermal bath of radiation. How can we understand the thermal nature of the ordinary vacuum state of quantum field theory in the absence of an evaporating black hole?

Still in two dimensions, let us call the Minkowski coordinates t, x and define null coordinates:

$$\begin{aligned} \bar{u} &= t - x \\ \bar{v} &= t + x. \end{aligned} \tag{4.50}$$

Then, in analogy with (4.12) we can define new coordinates η, ξ and $u = \eta - \xi$, $v = \eta + \xi$ through:

$$\begin{aligned} \bar{u} &= \exp(-u) \\ \bar{v} &= \exp(v). \end{aligned} \tag{4.51}$$

The modes $\bar{\psi}_\omega$ associated with (4.50) are the usual Minkowski plane wave modes, but those associated with (4.51), ψ_ω, will not be.

To understand the significance of the new coordinates, note from (4.51) that:

$$\bar{u}\bar{v} = t^2 - x^2 = \exp(2\xi) \tag{4.52}$$

so that a line of constant ξ is a rectangular hyperbola in Minkowski space, asymptotic to the null rays $x = \pm t$ which pass through the origin of the Minkowski coordinate system (see figure 4). It is well known that an observer who moves along such a world line experiences a uniform proper acceleration of magnitude $\exp(-\xi)$.

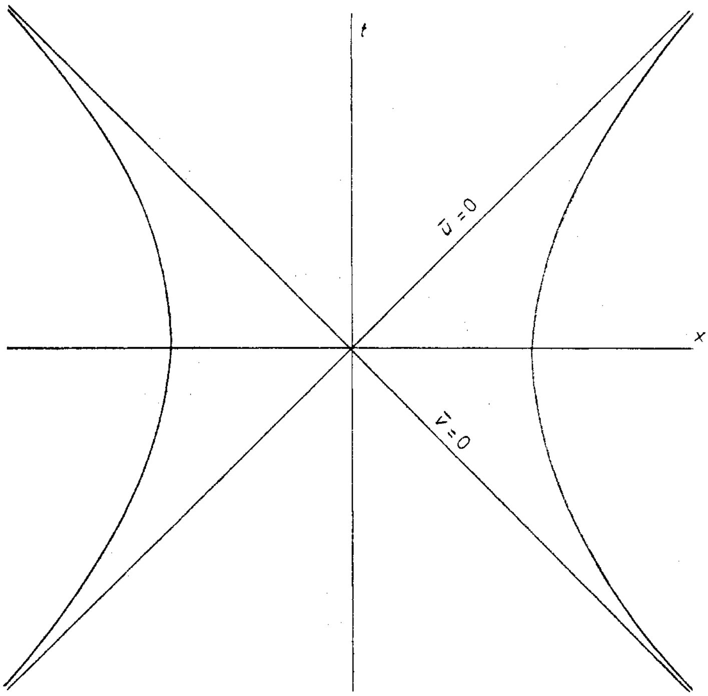

Figure 4. Accelerated observers. The causal structure of Minkowski space as viewed by uniformly accelerating observers (hyperbolae) is identical to the eternal black hole (see figure 3) without singularities. The null ray $\bar{u}=0$ is an event horizon for these observers in the right-hand wedge-shaped region.

Because the world line is asymptotic to the null rays $x= \pm t$, this observer will not be able to see events which occur in the region $x<t$. The future asymptote is an event horizon. Moreover (4.52) also describes a conjugate hyperbola to the *left* of the origin ($x<0$), representing an observer who accelerates to the left and for whom the null ray $x= -t$ is an event horizon. Although we are in Minkowski space, the causal structure as far as these accelerated observers are concerned is remarkably similar to an eternal black hole, with the regions $|x|>t$ corresponding to the exterior of the hole, that to the left of the origin being analogous to the 'mirror' universe located on the remote side of the black hole interior. Perhaps all this is not too surprising if we remember that, according to the equivalence principle, a uniform acceleration is indistinguishable locally from a static gravitational field.

The construction of modes and the connection between vacuum states proceeds in direct analogy to the black hole case. In particular, equation (4.42) is unchanged (when we put $\kappa=1$). The physical interpretation of the thermal radiation is somewhat different, however. Whereas the vacuum $|\bar{0}\rangle$ of a black hole appears to an *inertial* observer far from the hole as a bath of thermal radiation, the Minkowski vacuum seems like a thermal bath to an *accelerated* observer. Indeed, Unruh (1976) has shown that a simple model particle detector carried by such an observer would record the presence of radiation with a Planck spectrum. The fact that accelerated systems see a different physical situation from inertial systems comes as no surprise, because even classically the presence of inertial forces will cause 'peculiar' effects in non-inertial reference frames.

Unlike the black hole case, where the surrounding heat bath contains energy (see expression (4.29)) the Minkowski vacuum naturally has *zero* energy density. The quanta which are detected by the accelerated observer cannot be regarded as carrying energy or momentum in the usual sense. The energy which is absorbed by the Unruh detector must presumably be traced to the agency which is responsible for accelerating the detector.

4.6. *The vacuum energy*

In §4.3 we saw how the energy density difference between the $|\bar{0}\rangle$ and $|0\rangle$ states of a black hole behaves like that of a bath of thermal radiation at a local temperature $\kappa/2\pi k$. This thermal energy is, of course, superimposed on the background vacuum energy of the state $|0\rangle$. Because the space–time is curved, this will not be zero (see §2.3). In this subsection we shall outline how the vacuum energy density can actually be calculated.

The stress tensor *operator* for a massless scalar field in two dimensions is:

$$T_{\mu\nu}=\partial_\mu\phi\;\partial_\nu\phi-\tfrac{1}{2}g_{\mu\nu}g^{\sigma\rho}\;\partial_\sigma\phi\;\partial_\rho\phi. \tag{4.53}$$

The vacuum expectation value of $T_{\mu\nu}$ can be expressed as an infinite integral over the complete set of modes. Alternatively, the mode integral can be performed before the differentiation, so that $\langle 0|T_{\mu\nu}|0\rangle$ can be obtained by differentiation of the two-point function $G(x'', x')\equiv\langle 0|\phi(x'')\phi(x')|0\rangle$, after which the points x'', x' can be set equal to x.

For example, using (4.20) and noting that $g_{uv}=0$:

$$\langle 0|T_{uu}|0\rangle=\lim_{u'',\,u'\to u}\;\partial_{u''}\,\partial_{u'}G(x'', x')\propto\lim_{\Delta u\to 0}\,(\Delta u)^{-2} \tag{4.54}$$

which clearly diverges quadratically as the points x'', x' come together at x. This divergence is an expression of the infinite vacuum energy which arises from the sum of the $\frac{1}{2}\hbar\omega$ zero-point energy of all the field oscillators. In four dimensions the divergence is quartic and occurs even in Minkowski space. In special relativity this is not a problem because only energy differences are observable, so the infinite vacuum energy can be trivially subtracted away, but in general relativity energy is a source of gravity and the divergences must be handled in a more systematic way.

As in other branches of quantum field theory, the ultimate aim is to absorb the divergent quantities into renormalised physical constants and to argue that only the observed physical constants are relevant. In the case of quantum field theory in curved space–time this can only be done by generalising Einstein's equations to include higher-order terms. The divergences can then be absorbed into the coupling constants of these terms, as well as G, and λ—the cosmological constant.

The mathematical techniques which must be used to separate a finite residue from these renormalisable divergences vary, and are generally rather complicated so will not be described here. Instead, we need only use one result which emerges from all these techniques—the so-called conformal anomaly. The stress tensor operator (4.53) is formally traceless:

$$g^{\mu\nu}T_{\mu\nu}=0 \tag{4.55}$$

a feature closely associated with the conformal invariance of the field equation (4.4) and of $T_{\mu\nu}$. However, because the quantum expectation value $\langle 0|T_{\mu\nu}|0\rangle$ is divergent, the vanishing of the trace does not extend to the expectation value (we henceforth omit the 0 from the expectation value for convenience). Davies *et al* (1976) obtained

$$T \equiv g^{\mu\nu}\langle T_{\mu\nu}\rangle = \frac{1}{24\pi} R \neq 0. \tag{4.56}$$

The scalar curvature R is, in fact, the only geometrical scalar available in two dimensions. An identical result is true for a massless spinor field (Davies and Unruh 1977).

Armed with this information, we may compute the vacuum energy and the strength of the Hawking flux uniquely by integrating the covariant conservation equation:

$$\langle T^{\mu\nu}\rangle_{;\nu} = 0. \tag{4.57}$$

In null coordinates u, v this yields (Davies 1977c):

$$\partial_v \langle T_{uu}\rangle + \tfrac{1}{4} C\, \partial_u T = 0 \tag{4.58}$$

where C is, as usual, the conformal factor of the metric. Using the definition of the scalar curvature:

$$R = 4\left(C^{-2}\frac{\partial^2 C}{\partial u\, \partial v} - C^{-3}\frac{\partial C}{\partial u}\frac{\partial C}{\partial v}\right) \tag{4.59}$$

and interchanging the order of differentiation, enables (4.58) to be integrated:

$$\langle T_{uu}\rangle = -\frac{1}{12\pi} C^{1/2}\, \partial_u{}^2\, C^{-1/2} + f(u) \tag{4.60}$$

where $f(u)$ is an arbitrary function of u. To fix it, we note that the stress tensor in the static geometry around the black hole must be independent of time and fall to zero in the asymptotically flat region at large r, where $C \to 1$. This requires $f(u) = 0$. Moreover, by symmetry $\langle T_{uu}\rangle = \langle T_{vv}\rangle$.

Substituting for $C = 1 - 2M/r + e^2/r^2$ gives:

$$\langle T_{uu}\rangle = \langle T_{vv}\rangle = (24\pi)^{-1}\left(-\frac{M}{r^3} + \frac{3}{2}\frac{M^2}{r^4} + \frac{3}{2}\frac{e^2}{r^4} - \frac{3Me^2}{r^5} + \frac{e^4}{r^6}\right). \tag{4.61}$$

Equation (4.61) describes a static cloud of vacuum energy density surrounding the black hole. It is negative in the region $r > r_+$ and falls rapidly to zero in the asymptotic region $r \to \infty$. Near the horizon the u, v coordinate system has a coordinate singularity but we may transform (4.61) to the $\bar{u}$, $\bar{v}$ coordinates defined by (4.12), which are regular at $r = r_+$. This operation introduces an additional factor of $16M^2/\bar{u}^2$ in the expression (4.61) for $\langle T_{uu}\rangle$. Thus, at the horizon where $u = \infty$, $\bar{u} = 0$, this component of the stress tensor diverges quadratically in $\bar{u}$. It is for this reason that the vacuum state $|0\rangle$, used here for computing $\langle T_{\mu\nu}\rangle$, is regarded as unacceptable for the quantum state in the vicinity of a black hole (although it is still suitable for the vacuum region outside a static 'star', which has no horizon).

To obtain the stress tensor corresponding to the $|\bar{0}\rangle$ vacuum, we must add the expression (4.61) to that of the thermal radiation given by (4.29). That expression also diverges at $r = r_+$ in the $\bar{u}$, $\bar{v}$ system, but the two divergences actually cancel each other. Thus, a freely falling observer will only encounter a *finite* flux of energy when he crosses the horizon, in spite of the fact that during the brief interval of proper time required for him to reach the horizon, the black hole has emitted an infinite quantity of energy to infinity (neglecting back-reaction). This was discussed in §2.3.

Instead of adding (4.29) to (4.61) one can calculate the stress tensor for the $|\bar{0}\rangle$ vacuum directly from (4.60) by substituting for C from the 'Kruskal form' of the metric (4.13) rather than the 'Schwarzschild form' (4.2). Thus, the stress tensor calculation, based solely on the value for the conformal anomaly (4.56), provides a completely independent derivation of the Hawking effect and gives precisely the same value of the temperature (it cannot predict the spectrum, however). Alternatively the argument can be inverted and the Hawking effect used to fix the coefficient $1/24\pi$ of the conformal anomaly (Christensen and Fulling 1977).

5. Conclusions

The usual laws of thermodynamics cease to apply when gravitational fields are present. Specifically, if space–time horizons occur, then information and entropy may disappear from the observable regions of the universe. However, by extending the concepts of thermodynamics to include black hole regions, the laws may be generalised with remarkable ease to encompass these situations. The evidence for the generalised laws lies primarily in the result of Hawking that black holes emit radiation with a thermal equilibrium spectrum. This establishes the concept of a definite temperature for black holes. Although only based on a semiclassical approximation in which the quantum degrees of freedom of the gravitational field are ignored, the elegant connection with thermodynamics that Hawking's result provides is compelling evidence that it is a result of fundamental significance, and not merely an accident of the approximations used.

The weakest link in the reasoning rests with the identification of entropy with the event horizon area. Only heuristic arguments have been advanced to actually *equate* entropy with area rather than, say, some monotonic increasing function of it. Once the relationship is accepted, then the constant of proportionality ξ is determined from Hawking's result. Ideally, one would like a totally independent derivation of the entropy formula. The problem is that entropy is usually associated with the arrangements of microscopic degrees of freedom, but it is hard to see how the internal degrees of freedom of a black hole can be given a precise meaning in the absence of a quantum theory of the hole's gravitational field. Possibly a few internal degrees of freedom could be quantised (along the lines of quantum cosmology) and the evaporation regarded as a succession of transitions from excited internal states. In this way the entropy could be calculated by taking a thermal average over these states.

A related problem is that the black hole represents the thermodynamic equilibrium limit of gravitational collapse. In statistical mechanics, entropy can be defined for a general configuration of microstates, not just those corresponding to equilibrium. Similarly, it would seem that there must exist an entropy associated with gravitational fields which goes over to the black hole expression in the end-state limit. After all, when a star collapses, the ordinary entropy of the star does not suddenly convert to the black hole entropy. The collapse is a continuous process and the entropy of the star must gradually fade away, while that of the gravitational field grows. It would seem to be a major outstanding problem to understand how this limiting process occurs. Its solution would provide the equivalent for gravitational collapse of Boltzmann's H theorem for a box of gas.

One of the novel features of black hole entropy is its relative nature. Only for an external observer does the collapsing star lose information. It is possible to discuss how the increasing curvature inexorably coarse-grains the observations from a distance, steadily making information about the internal structure of the imploding star harder to obtain. At the same time the quantum radiation, at first highly non-equilibrium, gradually approaches the analogue of the Planck form. It ought to be possible to describe the thermodynamics of this transition phase in detail.

As a first step in dealing with non-equilibrium black hole problems, one can consider fluctuations about equilibrium and the irreversible dissipative processes associated with them. Investigations by Candelas and Sciama (1977) show that black holes obey the standard theory of non-equilibrium thermodynamics, including the fluctuation–dissipation theorem. This strongly suggests that the thermodynamic basis of self-gravitating systems may be fruitfully extended beyond the stationary end state of black holes to more general gravitational fields and quantum states.

Perhaps the most attractive feature of black holes is that they enlarge our notion of thermodynamics. In an area where both quantum theory and relativity are suspect, it is a remarkable thought that the laws of thermodynamics may remain intact and provide as strong a guide to producing the elusive theory of quantum gravity as they did for Planck and Einstein in producing quantum mechanics.

Acknowledgments

I should like to thank Drs P C Aichelburg, R Beig, L H Ford and C J Isham for helpful discussions in the course of preparing this review. I am grateful to the Institut für Theoretische Physik, Universität Wien for hospitality during part of this preparation. This work was supported in part by the Einstein Memorial Foundation (AR 77).

References

Bekenstein J D 1973 *Phys. Rev.* D **7** 2333
Bertin G and Radicati L A 1976 *Astrophys. J.* **206** 815
Birrell N D and Davies P C W 1978 *Nature* **272** 35
Candelas P and Sciama D W 1977 *Phys. Rev. Lett.* **38** 1372
Carr B 1977 *Mon. Not. R. Astron. Soc.* **181** 293
Carter B 1973 *Black Holes* ed DeWitt and DeWitt (London: Gordon and Breach)
Casimir H B G 1948 *Proc. Kon. Ned. Akad. Wetenschap.* **51** 793
Christensen S M and Fulling S A 1977 *Phys. Rev.* D **15** 2088
Davies P C W 1974 *The Physics of Time Asymmetry* (Surrey University Press/University of California Press)
—— 1976a *Proc. R. Soc.* A **351** 139
—— 1976b *Nature* **263** 377
—— 1977a *Proc. R. Soc.* A **353** 499
—— 1977b *Space and Time in the Modern Universe* (Cambridge: Cambridge University Press)
—— 1977c *Proc. R. Soc.* A **354** 529
—— 1977d *New Scientist* **75** 238
Davies P C W and Fulling S A 1977 *Proc. R. Soc.* A **356** 237
Davies P C W, Fulling S A and Unruh W G 1976 *Phys. Rev.* D **13** 2720
Davies P C W and Unruh W G 1977 *Proc. R. Soc.* A **356** 259
Dowker J S 1977 *Preprint* University of Manchester
Fulling S A and Davies P C W 1976 *Proc. R. Soc.* A **348** 393
Gibbons G W 1975 *Commun. Math. Phys.* **44** 245
Gibbons G W and Perry M 1976 *Phys. Rev. Lett.* **36** 985
Hawking S W 1971 *Mon. Not. R. Astron. Soc.* **152** 75
—— 1972 *Commun. Math. Phys.* **25** 152

—— 1973 *Black Holes* ed DeWitt and DeWitt (London: Gordon and Breach) pp1–56
—— 1974 *Nature* **248** 30
—— 1975 *Commun. Math. Phys.* **43** 199
—— 1976 *Phys. Rev.* D **13** 191
—— 1977 *Sci. Am.* **236** 34
Hawking S W and Ellis G F R 1973 *The Large Scale Structure of Space–Time* (Cambridge: Cambridge University Press)
Hut P 1977 *Mon. Not. R. Astron. Soc.* **180** 379
Israel W 1976 *Phys. Lett.* **57A** 107
McNamara J M 1978 *Proc. R. Soc.* A **358** 499
Misner C W, Thorne K S and Wheeler J A 1973 *Gravitation* (New York: Freeman)
Page D N 1976 *Phys. Rev.* D **13** 198
Penrose R 1969 *Nuovo Cim.* **1** *Special No* 252
Rees M J 1977 *Nature* **266** 333
Sciama D W 1976 *Vistas in Astronomy* **19** 385
Simpson M and Penrose R 1973 *Int. J. Theor. Phys.* **7** 183
Smarr L 1973 *Phys. Rev. Lett.* **30** 71
Tabor D and Winterton R H S 1969 *Proc. R. Soc.* A **312** 435
Unruh W G 1974 *Phys. Rev.* D **10** 3194
—— 1976 *Phys. Rev.* D **14** 870
Zel'dovich Ya B 1970 *Zh. Eksp. Teor. Fiz. Pis. Red.* **12** 443 (Engl. trans. *JETP Lett.* **12** 307)