Black holes and thermodynamics

R. Adamson

Department of Physics, University of Toronto, Toronto, Canada

Abstract

There exists a set of striking similarities between the laws governing the equilibrium mechanics of stationary black holes and the classical laws of thermodynamics. While the full development of this relationship requires the use of quantum field theory, the connection between thermodynamics and black hole mechanics is evident even in the framework of classical general relativity. This paper presents the set of parameters describing black holes which are the analogues of the properties describing thermodynamic systems. It then develops the black hole mechanical analogue of the zeroth, first, second and third laws of classical thermodynamics in general relativity. It ends with a discussion of the significance of the analogy, particularly as it relates to quantum theories of spacetime and gravity.

1 Introduction

An amazing feature of the physical world is that despite the staggeringly large number of individual particles making up a macroscopic system, each with its own momentum and position, such systems can usually be meaningfully described with only a small number of physical properties. Systems that are said to be in thermal equilibrium are especially simple and their behaviour is completely predictable if only a handful of properties are known about them - internal energy, entropy, temperature, volume, angular momentum, total mass and so on. When such systems interact with the outside environment, they do so according to four laws known collectively as the laws of thermodynamics. They are:

- Zeroth Law - The temperature of a system in thermodynamic equilibrium is constant throughout the system
- First Law - In any physical process in which an infinitesimal amount of work $\delta W$ is done on a system at temperature $T$, the entropy $S$ and energy $E$ of the system will change according to the equation
\[ \delta E = T \delta S + \delta W \]
- Second Law - In any physical process the entropy of a closed system can never decrease
- Third Law - No system can be cooled to absolute zero.[1]

One of the most surprising results of black hole research over the last forty years has been the realization that the mechanics of stationary black holes can be completely describe by a system of laws bearing a striking resemblance to the laws of thermodynamics. As it turns out, a stationary black hole can be completely described by only a small number of parameters, and
its interactions with its environment are completely governed by simple relationships between these parameters.

This paper will begin by giving a more concrete description of what a black hole is and by defining the parameters that can be used to describe it. It will then go on to derive the laws of black hole mechanics and point out the similarities between them and the laws of thermodynamics. Finally it will briefly discuss the significance of this analogy and the effect it has had on the search for a theory of quantum gravity.

1.1 Isometries and Killing vectors

Throughout this paper we will use the symmetry properties of black holes as reflected in the isometries of spacetime and the associated Killing vectors fields. Although this method is less intuitive than explicit metric solutions in a coordinate system, it is far more powerful and less computationally intensive. In this section we explain how the symmetries of spacetime are to be used in the rest of the paper.

If one has a metric defined on a spacetime manifold, then it is possible to map each point in the manifold onto another point. If such a mapping is differentiable then it is called a diffeomorphism. If a diffeomorphism leaves the metric of the manifold unchanged, then it is called an isometry. Examples of these kinds of mappings can easily be constructed in familiar manifolds. For example, the Euclidean plane has three isometries - translation along two different axes and rotation about some origin. Each of these transformations maps points on the plane to other points on the plane, but the metric tensor is unchanged, remaining always that of flat two-dimensional space (although its representation will, of course, depend on the coordinate system being used).

If an isometry exists in a spacetime, then we can define orbits in the spacetime such that if one were to travel along them one would not see a change in the local metric. The tangent vectors to these orbits are called Killing vectors and the field of all such vectors for a given isometry is called the Killing vector field.

In the language of Lie derivatives, a diffeomorphism is an isometry if and only if the corresponding Lie derivative of the metric vanishes. The standard relation between the Lie derivative of a tensor and the covariant derivative imposes a very important condition on the Killing vectors $\xi^a$ that generate the isometry [2]

$$\mathcal{L}_\xi g_{ab} = \xi^c \nabla_c g_{ab} + g_{cb} \nabla_a \xi^c + g_{ab} \nabla_b \xi^c = 0$$

$$\nabla_a \xi_b + \nabla_b \xi_a = 0$$

(2)

Since the covariant derivative of the metric is zero and multiplication by the metric in the second and third terms simply lowers the $c$ index. This relation is known as the Killing equation.

We restrict our interest in this paper to stationary black holes. Stationarity means that the spacetime in which the black hole lives must have an isometry in the time dimension with a corresponding timelike Killing vector field. In general it can have other isometries as well. We also restrict our attention to spacetimes in which there is a region of flat spacetime at spatial distances far from the black hole. This gives us a prefered family of observers who are capable of describing the black hole in terms familiar to those of us who are used to living in flat space - mass, energy, angular momentum and so on.
2 Black holes

A black hole is a region of spacetime such that [3]

$$\mathcal{B} = M - I^-(\mathbb{I}^+)$$  \hspace{1cm} (3)

Where \( \mathcal{B} \) is the black hole region, \( M \) is entire spacetime of the universe and the expression \( I^-(\mathbb{I}^+) \) can be read as the past of all the events that occur on the 'boundary' of the spacetime, at infinitely large distances and infinitely late times. This definition captures the essential feature of a black hole, namely that it represents a region of spacetime completely separate from the rest of the universe, a region from which no matter, energy or information can ever escape.

The black hole region represents a future endpoint which is disjoint from the region of large distances and times, so we can define a boundary in the spacetime inside of which the entire future of all observers lies inside the black hole. This boundary is called an event horizon and it is a closed, immersed submanifold of the spacetime manifold \( M \). It can most simply be thought of as a two dimensional boundary enclosing some volume of space at each moment in time.

From this definition and the rules of differential geometry a number of useful properties and theorems about black holes can be proved. The ones that are used in this paper are listed below without proof, but rigorous proofs can be found in Wald [2] or Hawking and Ellis[4]:

- All black holes must contain a spacetime singularity at which the curvature of spacetime is divergent.

- The event horizon of a black hole is a null hypersurface whose null generators are the orbits of a Killing vector field. The Killing vector field \( \chi^a \) is therefore orthogonal to the event horizon and \( \chi^a\chi_a = 0 \) on the event horizon.

- The most general black hole solution to the Einstein equation contains a black hole orthogonal to a Killing vector field composed of a timelike and a periodic spacelike part. In this case, the black hole event horizon is a null hypersurface orthogonal to a linear combination of two Killing vectors written as

$$\chi^a = \xi^a + \Omega\phi^a$$  \hspace{1cm} (4)

Where \( \xi^a \) is a timelike Killing vector field and \( \phi^a \) is a spacelike, periodic vector field. \( \Omega \) is a linearity constant.[5]

This statement is a consequence of the theorem that the most general stationary solution to Einstein’s equation containing a black hole is the Kerr solution.

The Killing field \( \chi^a \) becomes increasingly timelike as one moves away from the black hole.

The first item is commonplace knowledge, and the second one seems at least to be reasonable. A common textbook diagram of a black hole shows light cones 'tipping over' as they cross the event horizon, so that the future of an observer goes from being timelike to spacelike as the black hole boundary is crossed. It makes sense that the Killing vector pointing in the direction of time should go from having a negative norm outside the black hole to a positive norm inside.
the hole and can therefore be expected to be zero on the event horizon itself. To form a black hole event horizon, one just takes the set of all points where the Killing vector is null. These points form a hypersurface normal to the Killing vector at all points.

The last property is the most surprising one. With all the free variables contained in the tensors of the Einstein equation it is amazing to think that the sole requirement of stationarity reduces the space of all possible solutions down to a single two-parameter family. This result is reminiscent of thermodynamic systems, where a huge phase space of possible states is reduced to a state that can be describe with a handful of parameters, once the condition of thermal equilibrium is imposed.

The small number of free parameters in the general stationary black hole solution would seem to suggest a similarly simple mechanical description, and indeed it turns out that the mechanics of stationary black holes are governed by four very simple equations relating physically meaningful quantities. Two of these quantities are familiar from non-relativistic mechanics, namely mass and angular momentum. These will be discussed in the section on the first law of black hole mechanics. A third quantity, the surface area of the event horizon is a fairly intuitive geometric parameter. The fourth parameter is the less familiar notion of surface gravity which will be discussed in the next section. Using these four parameters alone, a complete description of the behaviour of stationary black holes can be achieved - a considerable simplification from the full-blown solution of Einstein’s equation!

3 Surface gravity and the zeroth law of black hole mechanics

3.1 Definition and physical interpretation of surface gravity

In this section we derive a parameter $\kappa$ of black holes called the 'surface gravity' and show that it is constant over the event horizon of any stationary black hole [6]. We begin by assuming a black hole with an event horizon that constitutes a null hypersurface orthogonal to the Killing vector field $\chi^a$. From the properties listed in the previous section, the Killing field has a norm of zero on the horizon, so on the horizon $\chi^a \chi_a = 0$. In particular, $\chi^a \chi_a$ is constant on the horizon, which means that its covariant derivative is normal to the horizon and therefore parallel to $\chi^a$.

We define a proportionality constant $\kappa$ via the equation

\[
\nabla^a (\chi^b \chi_b) = -2\kappa \chi^a \]

\[
\chi_b \nabla^a \chi^b + \chi^b \nabla^a \chi_b = -2\kappa \chi^a
\]

Where we have expanded the covariant derivative using the Leibniz rule and the Killing equation (2).

We assume for now that $\kappa$ is non-zero. Once the physical meaning of $\kappa$ has been made clear we can come back and examine what it would mean for $\kappa$ for be equal to zero.

Frobenius’s theorem [7] in differential geometry gives a criterion for whether or not a given vector field in a manifold $M$ is normal to some submanifold of $M$ with a well defined volume element. Although we have been somewhat loose with the definitions, it should be apparent
that the black hole event horizon constitutes such a submanifold, so Frobenius’s theorem gives
a constraint on how the Killing vector field normal to the event horizon must behave. This
constraint is that for any arbitrary vector field \( v^a \) in the tangent space of the event horizon, and
for the Killing field \( \chi^a \), the differential form \( \nabla_{[a} \chi_{b]} \) must be equal to \( \chi_{[a} v_{b]} \). We can eliminate \( v \) by expanding the following expression.

\[
\chi_{[a} \nabla_{b} \chi_{c]} = \chi_{a} \nabla_{b} \chi_{c} - \chi_{b} \nabla_{a} \chi_{c} + \chi_{b} \nabla_{c} \chi_{a} - \chi_{c} \nabla_{b} \chi_{a} + \chi_{c} \nabla_{a} \chi_{b} - \chi_{a} \nabla_{c} \chi_{b} \tag{6}
\]

\[
= \chi_{a} \chi_{b} v_{c} - \chi_{b} \chi_{a} v_{c} + \chi_{b} \chi_{c} v_{a} - \chi_{c} \chi_{b} v_{a} + \chi_{c} \chi_{a} v_{b} - \chi_{a} \chi_{c} v_{b}
\]

\[
= 0
\]

We can also apply the Killing equation to the first line of (6) to obtain

\[
\chi_{c} \nabla_{a} \chi_{b} = \chi_{b} \nabla_{a} \chi_{c} - \chi_{a} \nabla_{b} \chi_{c} \tag{7}
\]

\[
= -2 \chi_{[a} \nabla_{b]} \chi_{c}
\]

The factor of two appears because the antisymmetric brackets on the right side implicitly
contain a normalization factor of 1/2. If we now multiply both sides by \( \nabla^{a} \chi^{b} \), and expand the
antisymmetric brackets we get

\[
\chi_{c} (\nabla^{a} \chi^{b}) (\nabla_{a} \chi_{b}) = -(\nabla^{a} \chi^{b}) (\chi_{a} \nabla_{b} \chi_{c}) + (\nabla^{a} \chi^{b}) (\chi_{b} \nabla_{a} \chi_{c})
\]

\[
= -(\nabla^{a} \chi^{b}) (\chi_{a} \nabla_{b} \chi_{c}) - (\nabla^{b} \chi^{a}) (\chi_{b} \nabla_{a} \chi_{c})
\]

\[
= -2 (\chi_{a} \nabla^{a} \chi^{b}) (\nabla_{b} \chi_{c})
\]

\[
= -2 \kappa (\chi^{b} \nabla_{b} \chi_{c})
\]

\[
= -2 \kappa^{2} \chi_{c}
\]

\[
\kappa^{2} = -\frac{1}{2} (\nabla^{a} \chi^{b}) (\nabla_{a} \chi_{b})
\]

Where both the contravariant form of Killing’s equation (2) \( \nabla^{a} \chi^{b} = -\nabla^{b} \chi^{a} \) and the
definition of \( \kappa \) from (5) were used. Note that one can raise a summed index in one place and lower
it in another without changing the sum. \(^1\)

So we now have a method for calculating \( \kappa \) if the Killing vector field on the event horizon
is known, although we haven’t yet been able to ascribe any physical meaning to it. This will
require a few more calculations. We begin this by developing the following identity which will
be true everywhere, not just on the horizon.

\(^1\)This follows from writing out the same expressing with different indeces, performing the raising and lowering
with the metric and contracting over the indeces.
\begin{equation}
3(\chi^a \nabla^b \chi^c)(\chi_b \nabla_c \chi_a) = \frac{3}{36} (\chi^a \nabla^b \chi^c - \chi^a \nabla^c \chi^b + \chi^b \nabla^c \chi^a - \chi^b \nabla^a \chi^c + \chi^c \nabla^a \chi^b - \chi^c \nabla^b \chi^a) \times \\
(\chi_a \nabla_b \chi_c - \chi_a \nabla_c \chi_b + \chi_b \nabla_c \chi_a - \chi_b \nabla_a \chi_c + \chi_c \nabla_a \chi_b - \chi_c \nabla_b \chi_a)
= 6 \times \frac{1}{12} (\chi^a \nabla^b \chi^c \chi_a \nabla_b \chi_c + \chi^a \nabla^b \chi^c \chi_b \nabla_a \chi_c + \chi^a \nabla^b \chi^c \chi_c \nabla_a \chi_b + \chi^a \nabla^b \chi^c \chi_b \nabla_c \chi_a)
+ \chi^a \nabla^b \chi^c \chi_b \nabla_a \chi_c + \chi^a \nabla^b \chi^c \chi_c \nabla_a \chi_b + \chi^a \nabla^b \chi^c \chi_b \nabla_c \chi_a)
= \chi^a \chi_a (\nabla_b \chi^c)(\nabla_b \chi_c) - 2(\chi^a \nabla^b \chi^c)(\chi_b \nabla_a \chi_c)
\end{equation}

\begin{equation}
\frac{3(\chi^a \nabla^b \chi^c)(\chi_b \nabla_c \chi_a)}{\chi^a \chi_a} = -2\kappa^2 - 2(\chi^a \nabla^b \chi^c)(\chi_b \nabla_a \chi_c)
\end{equation}

Where we repeatedly used Killing’s equation (2) in the second last step, changed the ordering of dummy summation indeces and substituted the result of equation (8). Upon examining the behaviour of this expression on the horizon we notice that both the numerator and the denominator on the left hand side go to zero, the numerator by Frobenius’s theorem (6) and the denominator because \chi^a is null on the horizon. However, if we apply l’Hopital’s rule by taking the covariant derivative of the numerator and denominator we find that the covariant derivative of the numerator will vanish since it is quadratic in terms that go to zero on the horizon. On the other hand, the covariant derivative of the denominator is non zero by (5), so in the limit of approaching the event horizon the left hand side vanishes. Using this fact and equation (8), we have that on the event horizon,

\begin{equation}
\kappa^2 = -(\chi^b \nabla_b \chi^c)(\chi_a \nabla_a \chi_c)/(\chi^d \chi_d)
\end{equation}

If we now define a vector \(a^c = (\chi^b \nabla_b \chi^c)/(-\chi^a \chi_a)\) and a scalar \(V = (-\chi^d \chi_d)^{1/2}\) we see that this expression can be written as

\begin{equation}
\kappa = \lim(V a)
\end{equation}

where \(a\) is the length of \(a^c\) defined as \(a = (a^c a_c)^{1/2}\) and the limit is taken as one approaches the event horizon.

A clear physical meaning can be ascribed to this formula for \(\kappa\) [8]. There exists in curved spaces a well defined notion of what it means to locally ‘stay in one place’. If one moves in the local frame in such a way as to leave the local metric unchanged, then one can be said to be staying still. By definition, though, translation in spacetime which leaves the metric unchanged must be along the orbit of a Killing vector field, so ‘staying still’ means following the orbit of a Killing vector field. If we consider a black hole whose event horizon Killing vectors are purely timelike, we can develop a clear meaning for \(\kappa\).

In special relativity, a particle with a unit four-velocity \(u^a\) that satisfies the geodesic equation \(u^a \nabla_a u^b = 0\) is said to be undergoing force-free motion. If, however, \(u^a \nabla_a u^b = a^b \neq 0\), then the particle is said to be accelerating with an acceleration defined by \(a^b\). Classically, this acceleration is said to be due to a force \(f^c = ma^c\) [9]. If we look at the particular case of the velocity
corresponding to 'staying in place', then \( u^a \) will just be the normalized time-like Killing vector

\[
u^a = \frac{\xi^a}{(\xi^b\xi_b)^{1/2}}.\]

The corresponding acceleration is

\[
a^b = \frac{\xi^a \nabla_a \xi^b}{\xi^a \xi_a} = -\frac{\xi^a \nabla_a \xi^b}{V^2} \tag{12}\]

and the acceleration required to stay in place against the 'force' of gravity will be

\[
ma^b = m \frac{\xi^a \nabla_a \xi^b}{V^2} = m \frac{1}{2V^2} (\xi^a \nabla^b \xi_a + \xi_a \nabla^b \xi^a) = m \frac{1}{V} \nabla^b (\xi^a \xi_a)^{1/2} = m \frac{1}{V} \nabla^b V \tag{13}\]

Now suppose we have a particle close to the event horizon, but held in place by a long string held by an observer far from the black hole. The observer far from the hole will see the particle as having an energy

\[
E = m (-\partial_t)^a u_a = -\frac{m \xi^a \xi_a}{V} = mV \tag{14}\]

and the force this observer will have to exert on the string to hold it in place is

\[
F_{\infty} = (F^a F_a)^{1/2} = [(-\nabla^a E)(-\nabla_a E)]^{1/2} = m [\nabla^a V \nabla_a V]^{1/2} \tag{15}\]

On the other end of the string in the reference frame of the particle, the particle needs a force equal to \( ma \) to stay in place

\[
F = (a^b a_b)^{1/2} = m \frac{1}{V} [\nabla_a V \nabla^a V]^{1/2} \tag{16}\]

This shows that the force required to hold a particle in place is a factor of \( V \) smaller at large distances compared to what it is near the black hole. We can now interpret \( \kappa \) as exactly the value of the force \( F = Va \) required for a distant observer to hold a particle of unit mass on the event horizon in place. This is why \( \kappa \) is called the surface gravity of the black hole.\(^2\) Although the above interpretation of forces exerted at infinity holds only if the Killing vector field that is null on the event horizon is actually the timelike Killing field, the term 'surface gravity' is used for \( \kappa \) even in the more general case of a not purely timelike Killing field being normal to the event horizon.

\(^2\)Obviously the amount of force that an observer on the event horizon would have to exert in his own frame to keep from falling into the hole is infinite.
Another property of $\kappa$ that will be stated without proof [10] is that it represents a proportionality factor between the affinely parametrized null geodesics that generate the event horizon and the Killing vector field parametrization. If we label the affine parameter $\lambda$, then each component of the affinely parametrized null geodesic generators $k^a$ and the Killing vector $\chi^a$ is

$$\lambda k^a = \frac{1}{\kappa} \chi^a$$

(17)

This property is related to the fact that $\chi^a$ scales with the redshift factor, while the affine parameter does not. It is mentioned here because it will be required when we discuss the first law of black hole mechanics.

3.2 The zeroth law of black hole mechanics

The above derivation introduces the surface gravity $\kappa$ as a useful parameter in describing black holes with a clear physical meaning for an observer far from the hole. We now go on to show that the value of $\kappa$ must be constant over the event horizon of the black hole.[11]

We can rewrite equation (5) as $\chi^b \nabla_b \chi^a = \kappa \chi_a$ and multiply on both sides by $\chi^c \nabla_c$. Applying the Leibniz rule we get:

$$\chi_a \chi^c \nabla_c + \kappa \chi^c \nabla_c \chi_a = \chi^c \nabla_c (\chi_b \nabla_b \chi^a)$$

(18)

The last step uses an identity developed from the definition of the Riemann tensor and the Killing equation:

$$\nabla_a \nabla_b \chi^c - \nabla_b \nabla_a \chi^c = R_{abc}^d \chi_d$$

(19)

We can then take cyclic permutations and add and subtract in a manner similar to the derivation of differential equation for $R_{abc}^d$ to obtain

$$2 \nabla_b \nabla_c \chi_a = -R_{ba[c}^d \chi_d = -2R_{cab}^d \chi_d$$

(20)

We also make use of the identity from equation (7):

$$\chi_c \nabla_a \chi_b = -2 \chi_{[a} \nabla_b] \chi_c$$

(21)

Multiplying the right side of this expression by $\nabla_b \chi_a$ yields the first term right hand side of equation (18). Plugging the left side of (21) in for this we can rewrite (18) as
\[(\chi[d\nabla_c]\chi^b)\nabla_b\chi_a = -\frac{1}{2}(\chi^b\nabla_d\chi_c)\nabla_b\chi_a\] 
\[= -\frac{1}{2}\kappa\chi_a\nabla_d\chi_c
\]
\[= \kappa\chi[d\nabla_c]\chi_a\]

This is the same as the second term on the left hand side of (18). Canceling this out and rewriting we get,

\[\chi_a\chi[d\nabla_c]^{\kappa} = \chi^bR_{ab[c}^e\chi_d]\chi_e\] 

(23)

Now we can get a similar expression by multiplying (21) by \(\chi[d\nabla_c]\) to obtain

\[(\chi[d\nabla_c]\chi_c)\nabla_a\chi_b + \chi_c\chi[d\nabla_c]\nabla_a\chi_b = -2(\chi[d\nabla_c]\chi[a]\nabla_b]\chi_c - 2(\chi[d\nabla_c]\nabla[b]\chi_c)\chi_a)\]

(24)

Using (7) to twice expand the last part of the first term on the left into an antisymmetric bracket expression we can cancel the first term on the right. Rewriting the remaining double derivatives using (20) gives

\[-\chi_cR_{ab[e}^f\chi_d]\chi_f = 2\chi[aR_{b[c]}^ef\chi_d]\chi_f\]

(25)

If we now raise the \(c\) index on both sides, let \(c = e\) and contract, the left side of (27) will vanish by the symmetries of the Riemann tensor. If we then expand the second set of antisymmetric brackets on the right side, we get

\[-\chi[aR_{b[e}^f\chi_d^f\chi_f = \chi[aR_{b[c]}^ef\chi_d]\chi_f\]

(26)

We now note that the right hand side is somewhat similar to the right hand side of (25). Expanding both equations in the antisymmetric brackets and applying the symmetries of the Riemann tensor and the Killing equation, we find that they are, in fact, equal. We have therefore shown that the right hand side of equation (26) is equal to the left and side of (23)

\[\chi_a\chi[d\nabla_c]^{\kappa} = -\chi_a\chi[dR_c]^{\kappa}\chi_f\]

\[\chi[d\nabla_c]^{\kappa} = -\chi[dR_c]^{\kappa}\chi_f\]

(27)

At this point we have to rely on a statement made here without proof that the right hand side of this expression vanishes on the event horizon. This can be derived as a consequence of Raychaudhuri’s equation presented in a later section and the fact that the event horizon is orthogonal to the vectors that generate null geodesics [12]. With these it can be shown that the component \(\chi[dR_c]^{\kappa}\) is parallel to \(\chi_f\) and hence \(\chi[dR_c]^{\kappa}\chi_f = 0\) on the horizon where \(\chi^a\) is null. This statement applies the notion of the event horizon as the place at which time becomes “frozen” to the energy-stress-momentum tensor and hence, through Einstein’s equation, to \(R_{ab}\).
Admittedly, though, without the mathematical formalism to back it up, such a description is of limited usefulness. The full derivation can be found in Wald[1]. We now have

\[ \chi[\mathbf{d} \nabla_c \kappa] = \chi_d \nabla_c \kappa - \chi_c \nabla_d \kappa = 0 \quad (28) \]

The only way that this last expression can hold generally is if \( \nabla_c \) is parallel to \( \chi_d \) on the event horizon \(^3\), which means that it is normal to the event horizon. The gradient of \( \kappa \) on the surface of the event horizon is therefore zero, so that \( \kappa \) is constant over the whole horizon. This property of stationary black holes is known as the zeroth law of black hole mechanics. It is analogous to the zeroth law of thermodynamics which states that temperature is constant throughout a system in thermal equilibrium. As we develop the other laws we will see that this analogy between surface gravity and temperature is a very deep one and that in a way to be made clear later, they can actually be thought of as representing the same physical quantity.

### 3.3 The third law of black hole mechanics

When we introduced the parameter \( \kappa \), we demanded that it be non-zero on the surface of the black hole. Now that we understand this parameter to be the surface gravity of the hole, it seems intuitively obvious that it cannot be zero, otherwise the black hole would seem unattractive to an observer far from the hole. It seems even more unphysical to imagine \( \kappa \) being negative, for then the black hole would seem repulsive to a distant observer, going against all intuitive notions of how a black hole should behave. It seems reasonable, then, to require that \( \kappa \) be non-negative and that it only be allowed to approach zero in the limiting case of a very weak black hole. We can therefore introduce another law of black hole mechanics called the third law:

\[ \kappa \geq 0 \quad (29) \]

The law can also be proven by calculating \( \kappa \) explicitly in the most general stationary black hole metric, the Kerr metric. The non-negative nature of \( \kappa \) is then guaranteed by the requirement that the solution not have any closed timelike curves[13].

The corresponding thermodynamic law states that the temperature of a real thermodynamic system is always greater than absolute zero. Again, we see that there is a correspondence between temperature and surface gravity since the behaviour of the two properties seems to be governed by identical physical laws.

### 4 The area increase theorem

A parameter mentioned in section 1 that we have so far failed to discuss is the surface area of the black hole event horizon. Just as we were able to impose physical restrictions on the surface gravity of a black hole, we will also be able to restrict how the black hole area changes in physical processes. In order to do this we will first have to develop a few useful properties of the event horizon and its generators.

\(^3\)To see this set \( \xi^d = (1, 0, 0, 1) \), and write out the terms of the equation. It is obvious that \( \nabla_c \kappa \) will also be of the form \((a, 0, 0, a)\) for some constant \( a \).
Let $\eta$ be a null hypersurface (an event horizon, for instance). Pick a null geodesic generator of $\eta$ and let $\lambda$ denote the affine parameter of the geodesic and $k^a$ its tangent vector under this parametrization. We note that near the event horizon $k^a$ will locally be parallel to the Killing vector field $\chi^a$, but that the two vector fields will have different parametrizations. Recall that the affine parameter of a geodesic can be thought of as akin to a time parameter, even when the geodesic is a null geodesic.

The expansion $\theta$ of a point $p$ on the generator is defined by $\theta = \nabla_a k^a$. This quantity can be thought of as a spatial divergence of nearby geodesics as one moves along the geodesics. Most notably, since the hypersurface $\eta$ is defined as being normal to the null geodesic generators one can think of the expansion as characterizing the local change in an area element of $\eta$ as one moves it along the local null geodesics.

In general expansion is not the only parameter describing the evolution of geodesics as one moves along them. The most general such evolution can also involve a shear tensor $\sigma_{ab}$ describing the tendency of some geodesics to move further than others for a small change in $\lambda$ and a twist tensor $\omega_{ab}$ describing the rotation of an area element of the null hypersurface. Like the twist and shear tensors in solid bodies, these tensors are trace-free with the shear tensor being symmetric and the twist tensor antisymmetric. The overall behaviour of geodesics is described by the (0-2) tensor $B_{ab}$ defined as

$$B_{ab} = \frac{1}{3} \theta h_{ab} + \sigma_{ab} + \omega_{ab} = \nabla_b k_a$$

Where $h_{ab}$ is the projection of the metric onto its spatial part $h_{ab} = g_{ab} + k_a k_b$. Since $k_a$ is timelike and affinely parametrized, adding $k_a k_b$ to the metric will cancel the timelike part, leaving only a ‘spatial metric’ $h_{ab}$. To see how this tensor evolves as one moves along the geodesics, it can simply be dropped into the geodesic equation and expanded out

$$k^c \nabla_c B_{ab} = k^c \nabla_c \nabla_b k^a$$

$$= k^c \nabla_b \nabla_c k_a + R_{cda} \nabla^d k^a$$

$$= \nabla_b (k^c \nabla_c k_a) - (\nabla_b k^c) (\nabla_c k_a) + R_{cda} \nabla^d k^a$$

$$= -B^c_b B_{ac} + R_{cda} \nabla^d k^a$$

We have used the definition of the Riemann curvature tensor in the first step and in the last step the fact that since $k^a$ is the tangent vector of the geodesic, $k^c \nabla_c k_a$ is zero by the geodesic equation. We can now raise the $b$ index, set $a = b$ and take the trace.

$$k^c \nabla_c B^b_a = -B^{cb} B_{ac} + R^{b}_{cda} k^c k^d$$

$$k^c \nabla_c \theta = -\frac{d\theta}{d\lambda} = -\frac{1}{2} \theta^2 - \sigma_{ab} \sigma^{ab} + \omega_{ab} \omega^{ab} - R_{ab} k^a k^b$$

11
This final expression is known as the Raychaudhuri equation and it governs the expansion of null geodesic surfaces with time. Taking a closer look at the last term using the expression for the Ricci tensor contained in the Einstein equation, we have

$$R_{ab}k^a k^b = [8\pi T_{ab} + \frac{1}{2} R g_{ab}] k^a k^b = 8\pi [T_{ab} k^a k^b + \frac{1}{2} T]$$ (33)

Where we have used the fact that $k^a$ has an affine parametrization so that $k^a k^a = -1$ and the identity obtained by taking the trace of the Einstein equation, $R = -8\pi T$. Now there are special conditions that $T_{ab}$ is required to satisfy that will impose restrictions on $R_{ab}$. Most notably, since the quantity $T_{ab} k^a k^b$ is the local energy density as measured by an observer with 4-velocity $k^a$, it is expected that for time-like and null $k^a$ this quantity will be everywhere non-negative. Furthermore, although internal stresses in matter can make a negative contribution to $T$, for physically reasonable matter, this contribution should always be much smaller than the mass and momentum terms which contribute positively to $T_{ab} k^a k^b$. It therefore seems reasonable to assume further that $T_{ab} k^a k^b \geq -\frac{T}{2}$. Through the Einstein equation this imposes the restriction that $R_{ab} k^a k^b \geq 0$ which can be thought of as a requirement that gravity always be attractive, i.e. that the presence of matter always causes geodesics to converge rather than diverge. The terms $\sigma_{ab} \sigma^{ab}$ and $\omega_{ab} \omega^{ab}$ are also greater than or equal to zero. They are spatial tensors and so in Riemann normal coordinates the covariant and contravariant tensor will have the same value for a given set of indices, so the trace is simply a sum of squares. We therefore arrive at an important restriction on the expansion of geodesics, namely

$$\frac{d\theta}{d\lambda} \leq -\frac{1}{2} \theta^2$$ (34)

$$\int_{\theta_0}^{\theta} \frac{d\theta}{\theta^2} \leq - \int_{0}^{\lambda} \frac{1}{2} d\lambda$$

$$\frac{1}{\theta(\lambda)} \geq \frac{1}{\theta_0} + \frac{1}{2} \lambda$$

From the above we observe that if $\theta_0$ were negative, then as $\lambda$ increased, moving forward along the null geodesic, a point would be reached where the right side of the inequality would be zero so at some point the left side must go to zero if the inequality is to hold. This, however, implies an infinite expansion of nearby geodesics which seems highly unphysical and can be shown to contradict some of the basic properties of black holes. The only alternative is to have $\theta_0$ positive, in which case $\theta$ will remain non-negative for all $\lambda$. Since increasing $\lambda$ means moving forward in time to a time-like observer, this restriction means that the area of a black hole event horizon as viewed by a distant observer must never decrease with time.[14]

This statement is called the second of laws of black hole mechanics, and indeed, it bears a clear resemblance to the second law of thermodynamics which states that the entropy of a
closed system must never decrease with time. As was the case with temperature and surface gravity, at first glance this analogy seems extremely superficial. The above demonstration of area increase can be proved rigorously from the rules of Riemannian geometry and some physically reasonably assumptions about the mass-energy momentum tensor. The second law of thermodynamics, on the other hand, is only expected to hold probabilistically, with higher probability as the size of the system increases. Despite this, further developments in black hole mechanics have made clear that not only are these two quantities related, but that they are equivalent! This equivalence will be discussed further in the final section.

5 The first law of black hole mechanics

The area increase theorem imposes a constraint on how black holes can evolve in time, but it does not give a picture of how the event horizon area is related to the other properties of a stationary black hole, its mass, angular momentum and surface gravity. In what follows we will show that there is in fact a simple equation governing how a small change in one of these properties will cause changes in the others once the black hole has 'settled down' to an equilibrium state.[15]

If one imagines dropping a small amount of matter into a black hole, the local value of $T_{ab}$ near the black hole surface will change slightly by an amount $\delta T_{ab}$. The resulting change in the black hole area is governed by the Raychaudhuri equation (32). The change in $\theta^a$, $\sigma^a\sigma_{ab}$, and $\omega^a\omega_{ab}$ will only come about due to changes in the local curvature through Einstein’s equation. Such changes can be neglected to first order in $\delta T_{ab}$ and so the the Raychaudhuri equation simplifies to

$$\frac{d\theta}{d\lambda} = -8\pi\delta T_{ab}k^a k^b$$

(35)

As was briefly mentioned in the discussion of surface gravity, the relation between the killing vector field orthogonal to the event horizon and the affine parameter of the null geodesic generators of the event horizon is

$$k^a = \frac{1}{\kappa \lambda} \chi^a$$

(36)

$$= \frac{1}{\kappa \lambda} (\xi^a + \Omega \phi^a)$$

Where we have separated the Killing vector field normal to the event horizon into timelike and spacelike parts.

Evaluation of the effect of the change $\delta T_{ab}$ on the black hole once it has 'settled down' is now simply a matter of integrating both sides of the Raychaudhuri equation over the black hole event horizon surface and over all future values of $\lambda$. 

13
\[ \kappa \int d^2 S \int_0^\infty \lambda \frac{d\theta}{d\lambda} = -8\pi \int_0^\infty \int d^2 S \delta T_{ab} k^a k^b \]
\[ \kappa \left. \frac{d^2 S(\theta\lambda)}{d\lambda} \right|_0^\infty - \kappa \int d^2 S \int_0^\infty \theta d\lambda = -8\pi \left( \int_0^\infty d\lambda \int d^2 S \delta T_{ab} \xi^a k^b + \Omega \int_0^\infty dV \int d^2 S \delta T_{ab} \phi^a k^b \right) \]

An exact evaluation of these integrals would require considerable proficiency in Riemannian geometry, but the physical quantities that they represent are apparent even without any explicit calculations.\[16\]

The boundary term on the left hand side can be made equal to zero by taking the surface out to a region of spacetime where the expansion is negligible. Such an action is permitted by a version of Gauss’s law for curved spaces. The second term on the left is the integral of the expansion of each infinitesimal area element of the event horizon over the surface of the event horizon. This will simply be the infinitesimal change in event horizon surface area \( \delta A \) caused by the infinitesimal change in \( T_{ab} \).

On the right side we note that the action of the vector fields on \( T_{ab} \) is simply to project onto one of its components. Since \( \xi^a \) and \( k^a \) both point forward in time, the first integral will be an integral of the \( T_{00} \) component, which in special relativity is called the mass \( M \). The \( T_{ab} \phi^a k^b \) term is a projection onto the time-\( \phi \) component of \( T_{ab} \), which just the negative of the quantity that in special relativity is called the angular momentum \( J \). So for an observer far from the black hole where space is nearly flat, the interpretation given to (37) would be as

\[ \kappa \delta A = 8\pi (\delta M - \Omega \delta J) \] \[ (38) \]

This equation is known as the first law of black hole mechanics. Its name suggests an analogy with the first law of thermodynamics

\[ T \delta S = \delta E + \delta W \] \[ (39) \]

Indeed the analogy seems appropriate since it puts into correspondence event horizon area and entropy as well as surface gravity and temperature. The correspondence between mass and energy is no surprise given that these quantities are identical according to \( E = mc^2 \). The interpretation of \( \Omega \delta J \) turns out to be appropriate too. Recall that the energy of a rotating body is \( J\omega \) where \( J \) is the angular momentum and \( \omega \) the angular speed. Varying this, we get a differential energy of \( \omega \delta J \), so if the variable \( \Omega \) is interpreted as the angular speed of the black hole, then the term \( \Omega \delta J \) just represents the change in rotational energy as the matter is added to the black hole.

It is interesting to note that energy can actually be extracted from a black hole’s rotational degree of freedom in a thought-experiment called a Penrose process\[17\]. It is quite surprising that it is possible to transfer energy out of a black hole given that black holes are commonly thought of as letting nothing escape their gravity. The Penrose process is limited, though, in
how much energy it can extract. As energy is transferred out, the black hole 'slows down' and when it becomes completely static, no more energy can be extracted from it.\(^4\)

6 The last piece of the puzzle: Hawking radiation

Of the analogies that have been assigned so far between thermodynamic properties and black hole parameters, the one that appears least physically significant is the one between surface gravity \(\kappa\) and temperature. Penrose processes aside, classical black holes are objects that absorb all incoming matter and radiation and radiate nothing. Thermodynamically they would be described as perfect black bodies with a temperature of zero, so the analogy between temperature and \(\kappa\) would appear to be purely formal. This viewpoint was shattered, though, when Hawking showed that if quantum field theory is applied to black holes, then they can be described as having a non-zero temperature, and aside from proportionality constant, that that temperature \textit{is} \(\kappa\). While this result is quite simple, the derivation is complex and difficult to interpret\(^{18}\). The Hawking radiation process considering the behaviour of plane-waves over the entire history of a spacetime containing a black hole and comparing the waves at infinitely late times to the waves at infinitely early times. In this picture, a black hole behaves as a sort of amplifier for the quantum background radiation. No attempt will be made here to perform this derivation; we will simply quote the final result.

\[ T = \frac{\hbar \kappa}{2\pi c k} = \frac{\hbar c^3}{8\pi G M k} \]  

Where \(k\) is the Boltzmann constant and the usually suppressed terms \(c, G\) and \(\hbar\) have been restored for clarity. In the second line \(\kappa\) has been rewritten in terms of black hole mass for the specific example of the Schwarzschild metric. \(\kappa\) generally scales with inverse mass for all stationary black holes.\(^5\)

Unlike with a Penrose process, there is no theoretical upper limit to the fraction of the black hole’s energy that can be radiated via through Hawking radiation. Since energy conservation requires that the black hole mass be reduced to compensate for the radiated energy, it actually appears possible for a black hole to completely 'evaporate' in this way, leaving nothing behind. Since temperature is inversely proportional to mass, smaller black holes would be expected to evaporate more quickly and, paradoxically, to get hotter the more they radiate. It has been posited that this would lead to a runaway condition of increasing energy and a massive explosion.\(^6\)

\(^4\)In terms of specific metrics, this corresponds to a Kerr metric going to a Schwarzschild metric as the parameter \(a\) goes to zero

\(^5\)in the completely general Kerr metric \(\kappa = \frac{(M^2 - a^2 - e^2)^{1/2}}{2M[(M^2 - a^2 - e^2)^{1/2} - e]}\) where \(M\) is the mass, \(e\) is the charge and \(a\) is the angular momentum

\(^6\)It was this property of black holes that convinced the Reagan administration in the 1980s to fund research into making black holes in the laboratory for use in weapons! \[19\]
7 Summary and conclusions

We can now summarize the laws of black hole mechanics and the corresponding laws of thermodynamics:

<table>
<thead>
<tr>
<th>Black holes</th>
<th>Thermodynamics</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa = \text{constant}$</td>
<td>$T = \text{constant}$</td>
</tr>
<tr>
<td>$\delta M = \frac{1}{8\pi}\kappa \delta A + \Omega \delta J$</td>
<td>$\delta E = T \delta S + \delta W$</td>
</tr>
<tr>
<td>$\delta A \geq 0$</td>
<td>$\delta S \geq 0$</td>
</tr>
<tr>
<td>in any physical process</td>
<td>in any physical process</td>
</tr>
<tr>
<td>$\kappa \geq 0$</td>
<td>$T \geq 0$</td>
</tr>
</tbody>
</table>

Throughout this paper we have referred to these sets of equations as being analogous to each other but in fact the connection is deeper than mere analogy. The two sets of equations are actually identical! To see this, consider the process of Hawking radiation. In order that energy be conserved, a black hole that radiates energy must lose mass which in turn decreases its event horizon surface area, thereby violating the area increase theorem. The violation of the theorem is not unexpected because when quantum effects are taken into account the energy-momentum tensor may no longer satisfy $T_{ab} \geq 0$ on the event horizon. It would seem then that one can have Hawking radiation or the area increase theorem, but not both. Similarly, an object with finite entropy dropped into a black hole disappears without a trace, so the loss of matter to a black hole constitutes a net decrease in the entropy of the universe in violation of the second law of thermodynamics. Both of these problems can be resolved by considering that dropping an object into a black hole increases its surface area and Hawking radiation creates thermal particles outside the black hole with a finite entropy. This suggests a generalized second law of thermodynamics[20]:

$$S + A/4 \geq 0$$  (41)

Where $S$ is the entropy outside the black hole and $A$ is the surface area of the black hole event horizon. From this it can be seen that more than just an analogue of entropy, $A$ is the physical entropy of the black hole. The other parameters, $\kappa$ and $M$ can also be viewed not just as analogues of temperature and mass-energy, but as the same physical quantities. In this way, the two sides of the table converge into a single set of equations and can be summarized by saying that black holes obey the laws of thermodynamics, and the quantities of event horizon surface area, surface gravity and mass are identical to the thermodynamic properties of entropy, temperature and energy to within some scaling terms.

It has been stated by many authors that the connection outlined in this paper is a very deep one, and that it represents an important stepping stone the path to a robust quantum
theory of gravity. Some theories of quantum gravity such as quantum loop gravity even use the
thermodynamic equations as a starting point from which to develop a statistical mechanical
description of black holes[21]. Their microscopic description of spacetime is arranged to reduce
to the laws of black hole thermodynamics in the appropriate macroscopic limit in the same
way that quantum mechanics applied to large systems will yield the laws of classical thermo-
dynamics. There remain, however, some fundamental questions in trying to come up with a
microscopic, statistical picture of black holes. It is not obvious, for instance, where the entropy
in a black hole is ‘stored’. In quantum statistical mechanics there is a direct correspondence
between the entropy of a system and the number of quantum states of equal energy that the
system can be in. Assuming that there is an equivalent microscopic origin to black hole entropy,
it is unclear whether this entropy can best be thought of as lurking in the cloud of Hawking
radiation surrounding the black hole, as being on the event horizon or as being contained within
a few Planck wavelengths of the singularity itself [20]. Such issues act to both motivate and
direct studies into a full-blown quantum theory of gravity, and the first test that any such the-
ory should be expected to pass is that it simplify to the above equations in some macroscopic
limit. In this sense the laws described in this paper may well represent an important first step
towards a unified theory of nature.
References

[5] Wald, 1984. These properties are either derived or discussed in chapter 9 “Singularities” and chapter 12 “Black Holes”.
[6] Wald, 1984. This derivation is an expanded version of the one found in section 12.5 “Black holes and Thermodynamics”.
[9] Misner, C. W, Thorne, K. S., Wheeler, J. A.,”Gravitation” W.H. Freeman and Co, San Francisco 1970. Although this information is available in any good text on GR, the best one I found was in chapter 6 of “Gravitation”.
[10] Wald, 1984, section 12.5. The demonstration of this fact is straightforward, but would have required too much additional background to perform an effective derivation.
[12] Although the rigorous argument in favour of this is not too complicated, it would, again, require the introduction of several new concepts. See Wald, 1984 sections 9.2 and 12.5 for details.
[14] Wald, 1994. This is a simplified, less rigorous version of the argument given in chapter 6.
[15] Wald, 1994 section 6.2. This is a simpler, less rigorous derivation than the one in Wald, 1984, section 12.5.
[16] For a thorough discussion of the validity of this argument, see Misner et al., 461. For an explicit evaluation of these integrals see Wald, 1984, section 11.2 and for details on the volume element in curved spacetime see Wald, 1984 appendix B.
[17] Misner et al, chapter 33
[18] For a lengthy, thorough and completely incomprehensible derivation, see Hawking, S.W. “Particle Creation by Black Holes,” Commun. Math. Phys., 43, 199-220. For a slightly clearer explanation see Wald, 1984, Chapter 14. For the only description that made any sense to me, see http://superstringtheory.com/blackh/blackh3a.html

This was also emphasized by Lee Smolin in the Physics Dept colloquium he gave last semester.