2 The Laws of Black Hole Thermodynamics

In classical GR, black holes obey 'laws' that look analogous to the laws of thermodynamics. These are classical laws that follow from the Eintsein equations. Eventually, we will see that in quantum gravity, this is not just an analogy: these laws are the *ordinary* laws of thermodynamics, governing the microsopic UV degrees of freedom that make up black holes.

2.1 Quick review of the ordinary laws of thermodynamics

The first law of thermodynamics is conservation of energy,

$$\Delta E = \mathcal{Q} \tag{2.1}$$

where $\mathcal{Q}$ is the heat transferred to the system.[9] For quasistatic (reversible) changes from one equilibrium state to a nearby equilibrium state, $\delta\mathcal{Q} = TdS$ so the 1st law is

$$TdS = dE \ . \tag{2.2}$$

Often we will turn on a potential of some kind. For example, in the presence of an ordinary electric potential Φ, the 1st law becomes

$$TdS = dE - \Phi dQ \tag{2.3}$$

where Q is the total electric charge. If we also turn on an angular potential, then the 1st law is

$$TdS = dE - \Omega dJ - \Phi dQ \ . \tag{2.4}$$

The second law of thermodynamics is the statement that in any physical process, entropy cannot decrease:

$$\Delta S \geq 0 \ . \tag{2.5}$$

These laws can of course be derived (more or less) from statistical mechanics. In the microscopic statistical theory, the laws of thermodynamics are not exact, but are an

[9] Often the rhs is written $Q + W$ where $-W$ is the work done by the system. We'll set $W = 0$.

extremely good approximation in a system with many degrees of freedom.

2.2 The Reissner-Nordstrom Black Hole

The Riessner-Nordstrom solution is a charged black hole in asymptotically flat space. It will serve as an example many times in this course.

Consider the Einstein-Maxwell action (setting units $G_N = 1$),[10]

$$S = \frac{1}{16\pi}\int d^4x\sqrt{-g}\,(R - F_{\mu\nu}F^{\mu\nu}) \tag{2.6}$$

where $F_{\mu\nu} = \nabla_\mu A_\nu - \nabla_\nu A_\mu = \partial_\mu A_\nu - \partial_\nu A_\mu$. This describes gravity coupled to electromagnetism. The equations of motion derived from this action are

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R = 8\pi T_{\mu\nu} \tag{2.7}$$

$$\nabla_\mu F^{\mu\nu} = 0 \tag{2.8}$$

with the Maxwell stress tensor

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta S^{matter}}{\delta g^{\mu\nu}} = \frac{1}{4\pi}\left(-\frac{1}{4}g_{\mu\nu}F_{\alpha\beta}F^{\alpha\beta} + F_{\mu\gamma}F_\nu{}^\gamma\right) \ . \tag{2.9}$$

The Reissner-Nordstrom solution is

$$ds^2 = -f(r)dt^2 + \frac{dr^2}{f(r)} + r^2 d\Omega_2^2 \tag{2.10}$$

with

$$f(r) = 1 - \frac{2M}{r} + \frac{Q^2}{r^2} \ , \tag{2.11}$$

and an electromagnetic field

$$A_\mu dx^\mu = -\frac{Q}{r}dt, \qquad \text{so} \qquad F_{rt} = \frac{Q}{r^2} \ . \tag{2.12}$$

This component of the field strength is the electric field in the radial direction, so this is exactly the gauge field corresponding to a point source of charge Q at $r = 0$.

[10] See Carroll Chapter 6 for background material.The factors of 2 in that chapter are confusing; see appendix E of Wald for a consistent set of conventions similar to the ones we use here.

This is a static, spherically symmetric, charged black hole. There is nothing on the rhs of the Maxwell equation (2.8), so the charge is carried by the black hole itself; there are no charged particles anywhere. The parameter Q in the solution is the electric charge; this can be verified by the Gauss law,

$$Q_{electric} = \frac{1}{4\pi}\int_{\partial\Sigma} \star F = \frac{1}{4\pi} r^2 \int d\Omega_2 F_{rt} = Q \ . \tag{2.13}$$

This integral is over the boundary of a fixed-time slice Σ, ie a surface of constant t and constant $r \gg 1$.

Horizons and global structure

Write

$$f(r) = \frac{1}{r^2}(r - r_+)(r - r_-), \quad r_\pm = M \pm \sqrt{M^2 - Q^2} \ . \tag{2.14}$$

Then r_+ is the event horizon and r_- is the Cauchy horizon (also called the outer and inner horizon). The coordinates (2.10) break down at the event horizon, though the geometry and field strength are both smooth there. There is a curvature singularity at $r = 0$. See Carroll's textbook for a detailed discussion, and for the Penrose diagram of this black hole.

We will always consider the case $M > Q > 0$. If $|Q| > M$, then $r_+ < 0$, so the curvature singularity is not hidden behind a horizon. This is called a naked singularity, and there are two reasons we will ignore it: First, there is a great deal of evidence for the *cosmic censorship conjecture*, which says that reasonable initial states never lead to the creation of naked singularities.[11] Second, if there were a naked singularity, then physics outside the black hole depends on the UV (since the naked singularity can spit out visible very heavy particles), and we should not trust our effective theory anyway.

[11]There are also interesting violations of this conjecture in some situations, but in mild ways [Pretorious et al.]

2.3 The 1st law

Now we will check that this black hole obeys an equation analogous to (2.3), if we define an 'entropy' proportional to the area of the black hole horizon:

$$S \equiv \frac{1}{4\hbar G_N} \times \text{Area of horizon} \ . \tag{2.15}$$

We've temporarily restored the units in order to see that this is the area of the horizon in units of the Planck length $\ell_P = \sqrt{\hbar G_N}$. (Now we'll again set $G_N = \hbar = 1$.) For now this is just a definition but we will see later that there is a deep connection to actual entropy. The horizon has metric $ds^2 = r_+^2 d\Omega_2^2$, so the area is simply

$$A = 4\pi r_+^2 = 4\pi(M + \sqrt{M^2 - Q^2})^2 \ . \tag{2.16}$$

Varying the entropy gives

$$dS = \left(\frac{4\pi}{f'(r_+)}\right) dM - \left(\frac{4\pi Q}{f'(r_+) r_+}\right) dQ \ . \tag{2.17}$$

Rearranging, this can be written as the 1st law in the form

$$T dS = dM - \Phi dQ \tag{2.18}$$

with

$$T \equiv \frac{r_+ - r_-}{4\pi r_+^2} = \frac{\sqrt{M^2 - Q^2}}{2\pi(M + \sqrt{M^2 - Q^2})^2}, \quad \Phi = \frac{Q}{r_+} = \frac{Q}{M + \sqrt{M^2 - Q^2}} \ . \tag{2.19}$$

M, the mass of the black hole, is the total energy of this spacetime, so this makes sense. Φ also has a natural interpretation:

$$\Phi = -A_0|_{r=r_+} \ . \tag{2.20}$$

It is the electric potential of the horizon.

But, we have no good reason yet to call T the 'temperature' or S the 'entropy' (this will come later).

The 1st law relates two nearby equilibrium configurations. There are two ways we can think about it: *(i)* as a mathematical relation on the space of solutions to the equations, or *(ii)* dynamically, as what happens to the entropy if you throw some energy and charge into the black hole.

T is related to the *surface gravity* of the black hole

$$T = \frac{\kappa}{2\pi} \ , \tag{2.21}$$

which is defined physically as the acceleration due to gravity near the horizon (which goes to infinity) times the redshift factor (which goes to zero). If you stand far away from the black hole holding a fishing pole, and dangle an object on your fishing line near so it hovers near the horizon, then you will measure the tension in your fishing line to be κM_{object}. It can be shown that κ is constant everywhere on the horizon of a stationary black hole. This is analogous to the '0th law of thermodynamics': in equilibrium, temperature is constant.

If we restore units, then note that $S \propto \hbar$, so

$$T \propto \hbar \ . \tag{2.22}$$

Exercise: Thermodynamics of 3d Black Holes

Difficulty level: easy

Three-dimensional gravity has no true graviton, since a massless spin-2 particle has $\frac{1}{2}D(D-3) = 0$ local degrees of freedom. However, with a negative cosmological constant, there are non-trivial black hole solutions, found by Banados, Teitelboim, and Zanelli. The metric of the non-rotating BTZ black hole is

$$ds^2 = \ell^2 \left[-(r^2 - M^2)dt^2 + \frac{dr^2}{r^2 - M^2} + r^2 d\phi^2 \right] \ , \tag{2.23}$$

where ϕ is an angular coordinate, $\phi \sim \phi + 2\pi$. This is a black hole of mass M in a spacetime with cosmological constant $\Lambda = -\frac{1}{\ell^2}$.

(a) Compute the area of the black hole horizon to find the entropy.

(b) Vary the entropy, and compare to the 1st law $TdS = dM$ to find the temperature of the black hole.

(c) Put all the factors of G_N and $\hbar$ back into your formulas for S and T. S should be dimensionless and T should have units of energy (since by T we always mean $T \equiv k_B T_{thermodynamic}$). Does T have any dependence on G_N?

Exercise: Thermodynamics of rotating black holes.

Difficulty level: Straightforward, if you do the algebra on a computer

The Kerr metric is

$$ds^2 = -\frac{\Delta(r)}{\rho^2}(dt - a\sin^2\theta d\phi)^2 + \frac{\rho^2}{\Delta(r)}dr^2 + \rho^2 d\theta^2 + \frac{1}{\rho^2}\sin^2\theta(adt - (r^2+a^2)d\phi)^2 \ , \quad (2.24)$$

where

$$\Delta(r) = r^2 + a^2 - 2Mr \ , \qquad \rho^2 = r^2 + a^2\cos^2\theta \ , \qquad (2.25)$$

and $-M < a < M$. This describes a rotating black hole with mass M and angular momentum

$$J = aM \ . \qquad (2.26)$$

(a) Show that the entropy is

$$S = 2\pi M r_+ = 2\pi M(M + \sqrt{M^2 - a^2}) \ . \qquad (2.27)$$

(b) The first law of thermodynamics, in a situation with an angular potential Ω, takes the form

$$TdS = dM - \Omega dJ \ . \qquad (2.28)$$

Use this to find the temperature and angular potential of the Kerr black hole in terms of M, a. (*Hint:* The angular potential can also be defined as the angular velocity of the horizon: $\Omega = -\frac{g_{tt}}{g_{t\phi}}|_{r=r_+}$.)

2.4 The 2nd law

The second law of thermodynamics says that entropy cannot decrease: $\Delta S \geq 0$. This law does not require a quasistatic process; it is true in any physical process, including those that go far from equilibrium. (For example, if gas is confined to half a box, and we remove the partition.)

Hawking proved, directly from the Einstein equation, that in any physical process *the area of the event horizon can never decrease.* This parallels the second law of thermodynamics! This is a very surprising feature of these complicated nonlinear PDEs. We will not give the general proof; see Wald's textbook.

Exercise: Black hole collision

Difficulty level: 2 lines

The 2nd law also applies to multiple black holes. In this case the statement is that the total entropy – ie the sum of the areas of all black holes – must increase. Argue that if two uncharged, non-rotating black holes collide violently to make one bigger black hole, then at most 29% of their initial rest energy can be radiated in gravitational waves.

Exercise: Perturbative 2nd law

Difficulty level: Easy if familiar with particle motion on black holes

The 2nd law applies to the full nonlinear Einstein equation. In most cases, like a black hole collision, it is hopeless to actually solve the Einstein equations explicitly and check that it holds. But one special case where this can be done is for small perturbations of a black hole. In this exercise we will drop a charged, massive particle into a Reissner-Nordstrom black hole, and check that the entropy increases.[12]

Suppose we drop a particle of energy ϵ and charge q into a Reissner-Nordstrom black hole along a radial geodesic (to avoid adding angular momentum), with $\epsilon \ll M$ and

[12] Reference: MTW section 33.8.

$q \ll Q$. This will change the mass and charge of the black hole,

$$M \to M + \epsilon, \qquad Q \to Q + q \ . \tag{2.29}$$

Although initially there will be some fluctuations in the spacetime and ripples on the horizon from the particle that just passed through, these will quickly decay so that we have once again the Reissner-Nordstrom solution, now with the new energy and charge. Therefore, in this process the area of the black hole horizon changes according to the 1st law (2.18),

$$\delta S = \frac{1}{T}(\epsilon - \Phi q) \ . \tag{2.30}$$

(a) The infalling particle follows a trajectory $x^\mu(\tau)$ where τ is proper time. Its 4-momentum is

$$p^\mu = \frac{dx^\mu}{d\tau} \ . \tag{2.31}$$

In a spacetime with a time-translation Killing vector $\zeta^{(t)}$, the energy of a charged particle

$$\epsilon = -(p - qA) \cdot \zeta^{(t)} \ . \tag{2.32}$$

This is conserved along the path of the particle (which is not a geodesic, since it feels an electromagnetic force). For a charged particle on the Reissner-Nordstrom black hole, find ϵ in terms of $f(r)$ and the components of p^μ.

(b) Assume $Q > 0$. For one sign of q, the energy ϵ can be negative. Which sign?

If we drop a negative-energy particle into a black hole, the mass of the black hole decreases. Therefore it is possible to extract energy using this process. For uncharged but rotating black holes, a similar procedure can be used to extract energy in what is called the *Penrose process*. Particles far from the black hole cannot have negative energy, so negative-energy orbits are always confined to a region near the horizon. This region is called the *ergosphere*.

(c) Although we can decrease the energy of the charged black hole, we cannot decrease the entropy. To show this, we need to find the minimal energy of an orbit crossing the horizon. Assume the particle enters the horizon along a purely radial orbit,[13]

[13]This assumption is not necessary. In the general case, the particle can add angular momentum to the black hole, so we need to consider the charged, rotating Kerr-Newman spacetime. This is treated

$p^\theta = p^\phi = 0$. The proper time along the orbit is

$$d\tau^2 = -ds^2 = f(r)dt^2 - \frac{dr^2}{f(r)} \ . \tag{2.33}$$

Use this equation to write ϵ in terms of p^r, q, and $f(r)$.

(d) ϵ is conserved along the orbit, so you can evaluate it where the particle crosses the horizon, $r = r_+$. Show that the minimal value of ϵ is

$$\epsilon_{min} = qA_t(r = r_+) \ . \tag{2.34}$$

Conclude that the 2nd law of thermodynamics is obeyed.

(e) Reversible processes are those in which $\Delta S = 0$. How would you reversibly drop a charged particle into a Reissner-Nordstrom black hole? (*i.e.,* what charge would it have and how would you drop it?)

2.5 Higher curvature corrections

Everything in this section so far has assumed the gravity action is $\int \sqrt{-g}(R - 2\Lambda)$. As discussed in section 1.1, this is incomplete: there should be higher curvature corrections suppressed by the scale of new physics.

In a general theory of gravity including curvature corrections, the formula for the entropy also receives corrections,

$$S = \frac{\text{Area}}{4} \ + \ \text{higher curvature corrections} \ . \tag{2.35}$$

The more general formula is called the 'Wald entropy'. We will postpone the general discussion of the Wald entropy until later; for now suffice it to say that the 1st law still holds. The 2nd law, however, does not. There are known counterexamples involving black hole collisions.[14] To my knowledge this is not fully understood. A likely

in detail in [MTW section 33].

[14]See arXiv: [hep-th/9305016], [0705.1518], [1011.4988].

explanation is that this signals a breakdown of the effective field theory — *i.e.,* that when these violations occur we must include higher corrections or corrections from new physics in the UV.

2.6 A look ahead

We have seen that the classical Einstein equations lead to laws of black hole mechanics that are analogous to the laws of thermodynamics. In quantum gravity, it is not just an analogy.

Temperature

What we called 'T' is a true temperature: black holes radiate as blackbodies with temperature T. This is Hawking radiation. It does not rely on quantizing gravity itself — it is a feature of quantum field theory in curved space, which will be derived in the next couple lectures.

Generalized second law

The entropy S is also a real entropy. This means that the total entropy of a system is the ordinary entropy (of whatever gas is present, or a cup of tea, etc) *plus* the total entropy of all the black holes in the system. The *generalized second law* is the statement that the total entropy cannot decrease:

$$S_{tot} = S_{black\ holes} + S_{stuff} \ , \qquad \Delta S_{tot} \geq 0 \ . \tag{2.36}$$

If you throw a cup of hot tea into a black hole, then this entropy seems to vanish. This is puzzling, because if we didn't know about black hole entropy, we might conclude that the ordinary 2nd law (applied to the tea) had been violated by destroying entropy. However, the generalized second law guarantees that in this process the area of the horizon will increase, and this will (at least) make up for lost entropy of the tea.

Counting microstates

Finally, we know that in quantum mechanics, entropy is supposed to count the states of a system:

$$S(E) = \log\left(\#\text{ states with energy } E\right)\ . \tag{2.37}$$

For a supermassive astrophysical black hole like the one at the center of the Milky Way, this is an enormous number, or order $\left(\frac{10^6 km}{\ell_P}\right)^2 \sim 10^{88}$. For comparison, the entropy of all baryons in the observable universe is around 10^{82}, and the entropy of the CMB is about 10^{89}. So black holes must have enormous number of states!

Classical black holes have no microstates. They are completely specified by M, J, Q (this statement is called the *no hair theorem*). How, then, can they have entropy? The answer should be that in the UV completion of quantum gravity, black holes have many microstates. This is exactly what happens in certain examples in string theory, and in AdS/CFT, as we'll see later. Unlike Hawking radiation, to understand the microscopic origin of black hole entropy requires the UV completion of quantum gravity. Turning this around, this means that black hole entropy is a rare and important gift from nature: an infrared constraint on the ultraviolet completion, that we should take very seriously in trying to quantize gravity.