

MIT OpenCourseWare
<http://ocw.mit.edu>

8.821 String Theory
Fall 2008

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.

8.821 F2008 Lecture 22: Black Hole Mechanics *is* Thermodynamics

Lecturer: McGreevy

December 10, 2008

We will begin to “cleave some blazing swath” through the set of results of classical gravity. We will see that black hole mechanics is thermodynamics. Our next goal will be to answer the question: “Thermodynamics of *what?*” First, though, let’s tie up some loose ends from the last lecture.

1 Loose Ends

Last time we wrote down a confining solution and argued that it had a mass gap. There was the question of why this argument fails in AdS.

In the confining solution Φ_ω is normalizable so that

$$\omega^2 \int \phi^2 = \int (\partial_z \phi)^2 \tag{1}$$

are not only greater than zero, but less than ∞ .

In AdS, $\Phi_\omega = z^{\frac{D}{2}} J_\nu(qz)$, which, as $z \rightarrow \infty$, goes as

$$z^{\frac{D}{2}} J_\nu(qz) \sim z^{\frac{D-1}{2}} \cos\left(z + \frac{\pi\nu}{2} - \frac{\pi}{4}\right) \left(1 + O\left(\frac{1}{z}\right)\right) \tag{2}$$

with $q = \sqrt{\omega^2 - k^2} > 0$ so that

$$\int dz \phi^2 = \infty \tag{3}$$

which is the reason there is no mass gap in AdS.

2 What is a Black Hole?

A rough definition of a black hole is that it is a “region of space from which there is no escape (at least classically).” We can, to zeroth order, argue the existence of black holes even without GR. If a heavy object creates a gravitational potential well, the velocity that a test particle would need to escape at distance r is v_{esc} , where

$$\frac{v_{esc}^2}{r} \sim \frac{GM}{r^2}. \quad (4)$$

Under some reasonable assumptions about physics, no particle can go faster than c , and so if the test particle is within $r_s = \frac{2GM}{c^2}$ (the Schwarzschild radius (for the '2' we need GR)) of the mass, it will not be able to escape.

Thus, a BH forms if the massive object has a radius less than r_s . The boundary of the region of no escape is called the event horizon. If what we're interested in is just some sort of classical evolution outside of the horizon, we never have to ask what sits behind it.

An important question which might annoy us: when we say “escape”, escape to where? This, in general, depends on the asymptotics of the space. For example, one might be tempted to say that a massive particle in AdS can never escape, since the geodesics just oscillate around the center of AdS and never reach the boundary. However, we should probably be more careful in our definition of a BH so that this case is excluded. Most of the theorems we're about to talk about were proven for the asymptotically flat case (i.e. not AdS), so we may have to think about how to apply them.

The horizon is a global concept. To know where the horizon is, you actually need to know the whole future evolution of the spacetime. The definition of the horizon in the asymptotically flat case is “the boundary of the past of future null infinity”. It makes sense if you stare at it long enough. Really.

3 The Schwarzschild BH

The metric of the Schwarzschild BH is (in these particular coordinates),

$$ds^2 = -f dt^2 + \frac{dr^2}{f} + r^2 d\Omega_2^2 \quad (5)$$

with $f \equiv 1 - \frac{2GM}{r}$ (the “emblackening factor”). As $r \rightarrow \infty$, $f \rightarrow 1$ and the space becomes asymptotically flat (Minkowskian). In this spacetime, there exists a timelike killing vector $\xi = \frac{\partial}{\partial t}$ which generates time translations. At the horizon $r = r_s$, however, this Killing vector becomes null as

$$|\xi| = \sqrt{\xi^\mu \xi^\nu g_{\mu\nu}} = \sqrt{f(r_s)} = 0 \quad (6)$$

Behind the horizon, there's a real singularity at $r = 0$, but the classical evolution outside the horizon is unaffected. Which is nice. We can use GR even though we don't know what happens behind the horizon. This fact is so convenient that the relativists have elevated it to a grand conjecture.....

4 The Cosmic Censorship Conjecture (due to optimistic relativists everywhere)

The cosmic censorship conjecture states, very roughly, that naked singularities don't happen. That is to say, all singularities are hidden behind a horizon. More specifically, it says that they don't happen for some set of "nice" initial data on some spacelike hypersurface, where "nice" has some technical meaning. There exists some numerical and analytical evidence for this. We usually just talk about this conjecture and forget about it, but we should pause to observe how nice this is. It means we don't need to ask hard questions about the singularity, quantum gravity, etc., because we can always apply classical GR outside of a horizon.

But, like all good things, this conjecture has a flipside—the singularity theorems of Hawking and Penrose. These state that if at some point in the evolution a horizon forms (or more specifically a "trapped surface"—see Wald if you are into both beauty and pain), the the future evolution will always involve a singularity behind the horizon. This rules out the suspicion one might have that the singularity is some artifact of the highly idealized Schwarzschild solution, and that some realistic, asymmetric gravitational collapse might produce a spacetime with no singularity.

5 BH Uniqueness

In some sense, black holes are actually extremely simple. The uniqueness theorems (or "no hair" theorems) state that there are only a few numbers which parametrize stationary BH's (those with a timelike killing vector): its mass M , its angular momentum J , and its long range gauge charges (i.e. electric or magnetic charges), Q_i . Black holes are the simplest macroscopic objects there are—in this respect they are very much like elementary particles.

Also, excitations above stationary solutions evolve back to the stationary solutions by 'ringdown' via quasinormal modes (which have $Im\omega > 0$). The decay is due to energy falling into the horizon. In asymptotically flat space, the black holes can also radiate to infinity via gravitational waves.

You should think of the black hole as some equilibrium which is parametrized by a small number of state variables. If you perturb the equilibrium, it returns.

6 Hawking's Area Theorem

Assuming cosmic censorship, and the null energy condition (which states that $T_{ab}k^ak^b \geq 0 \forall$ null k), Hawking's area theorem states that during the evolution of any black hole, the area of the horizon can't decrease. The idea of the proof is in two parts.

If the area ever decreases ($\frac{dA}{dt} < 0$), then the horizon develops "crossing points" at some finite later time. Consider a bundle of null geodesics which span the horizon—that is to say, every point on the horizon lies on some null geodesic that stays on the horizon. Consider the congruence of such geodesics. These curves have some tangent factor $k^a = \frac{dx^a}{d\lambda}$ where λ is an affine parameter. The fact that this is a geodesic with affine parameter implies that $k^a \nabla_a k^b = 0$. It's useful to introduce a quantity known as the convergence of the congruence, which, roughly speaking, measures how the geodesics are spreading out, and hence of how the area is changing

$$\rho \equiv -\frac{d}{d\lambda} \ln(dA) \quad (7)$$

If $\rho \leq 0$ then the area of the horizon is increasing. We will next use some GR magic—the Raychaudhuri focusing equation:

$$\frac{d\rho}{dt} = \frac{1}{2}\rho^2 + \sigma^2 + R_{ab}k^ak^b \quad (8)$$

σ is something that's called the "shear of the congruence." Its precise definition need not concern us—all we need to know here is that $\sigma^2 \geq 0$. The third term in the equation is, via the Einstein equations, proportional to $T_{ab}k^ak^b$, and hence by the null energy condition, also greater than or equal to zero. Therefore

$$\frac{d\rho}{dt} \geq \frac{1}{2}\rho^2 \quad (9)$$

Integrating, this gives

$$\rho(\lambda) \geq \frac{2}{\frac{2}{\rho_0} - \lambda} \quad (10)$$

with $\rho(0) = \rho_0$. From this we can see that if the initial ρ_0 is positive (that is if $\frac{dA}{d\lambda} < 0$), then $\rho \rightarrow \infty$ for some finite $\lambda_* \leq \frac{2}{\rho_0}$. A place where the convergence blows up is known as a crossing point or a caustic—it is where two geodesics cross and is a singularity in the congruence. However, this does not necessarily mean that it is a singularity in the spacetime itself.

The second part of the proof of the area theorem states something like "this can't happen." One can use some global properties to show that the horizon generators would have to end before λ_* , and this would imply a singularity. However, we can't quite use cosmic censorship now, because it's not clear that some drunken, barely clothed singularity cannot exist. There's some small deformation of this argument that works, though.

NB: if the definition of the convergence (7) makes you slightly uncomfortable, as it did me, there's a completely kosher, if much less illuminating definition of it on p. 217 of Wald (where it is labeled θ and called the "expansion.") A nice derivation of Raychaduhuri's equation follows, along with a similar derivation of the first part of this proof.

7 Using Black Holes to Make Engines

There are two ways one might imagine using black holes to extract energy and, say, run your toaster.

7.1 Box Lowering

One can use a BH to extract the rest energy of a particle of mass m . Along a geodesic, a particle has a conserved energy given by $E = m\dot{x}^\mu(\tau)\xi_\mu$, where ξ is the timelike killing vector. If we hold the particle at a fixed r, θ, ϕ , then the velocity vector takes the form $\dot{x}^\mu = \frac{\xi^\mu}{|\xi|}$ (the particle is moving in the time direction only). Therefore the energy is $E = |\xi|m$, which goes to m as $r \rightarrow \infty$, and goes to zero as we go to the horizon of the black hole (as we have already discussed, ξ becomes lightlike at the horizon). So if you carefully, quasistatically lower a box to the horizon, you can extract the rest energy, m , in the process.

7.2 Penrose Process

Rotating (Kerr) black holes have an "ergosphere" that lies outside the horizon for which the timelike killing vector becomes spacelike. The associated energy, therefore, is now proportional to the product of a timelike vector and a spacelike vector and may be negative. The basic idea of the Penrose process is that you send into the ergosphere some mass of energy E_0 attached to a bomb. In the ergosphere, the bomb explodes and the mass fragments into two pieces, one of which falls into the horizon and has $E_2 < 0$. By energy conservation, the other piece, which can now leave the ergosphere has $E_1 > E_0$, more energy than it went in with. In this way you can extract the rotational energy of the BH. One can show that this process is most efficient when the area of the horizon doesn't change (something like the adiabatic limit). If the change in area is not zero, then the process is irreversible. This analogy with thermodynamics is becoming hard to ignore...

8 The First Law of Black Hole Thermodynamics [Due to Smarr]

Under classical evolution, the following is true

$$dM = \Omega dJ + \Phi dQ + \frac{\kappa}{8\pi G} dA \tag{11}$$

with M, J, Q, A the state variables mass, angular momentum, charge and area, and Ω the angular velocity, Φ the electrostatic potential and κ something known as the surface gravity, about which we will have more to say shortly. The final term in this equation is analogous to the $T dS$ term in the first law of thermodynamics, although it is unclear which constants to assign to T and which to dS ; we'll work this out in future lectures. Nevertheless the surface gravity is something like the temperature and the area is something like the entropy.

Here we sketch the proof for uncharged, non-rotating black holes with $\Omega = \Phi = 0$. Let's consider some quasistatic process in which we throw some small bit of mass into the black hole. The flux of the conserved energy current is $T_{ab}\xi^b$, where T is the stress energy tensor. Thus the flux through the horizon is

$$\Delta M = \int_{horizon} T_{ab}\xi^a k^b d\lambda dA \quad (12)$$

k is a normal vector to the horizon (it is also tangent, since the horizon is null). The (unproven-in-lecture, though plausible¹) claim here is that $\xi^a = \kappa\lambda k^a$, and so

$$\begin{aligned} \Delta M &= \frac{\kappa}{8\pi G} \int R_{ab} k^a k^b \lambda d\lambda dA \\ &= \frac{\kappa}{8\pi G} \int \frac{d\rho}{d\lambda} \lambda d\lambda dA + O(\Delta M^2) \\ &= \frac{\kappa}{8\pi G} \int (-\rho) d\lambda dA \\ &= \frac{\kappa}{8\pi G} \int \left(\frac{d}{d\lambda} \ln(dA) \right) d\lambda dA \\ &= \frac{\kappa}{8\pi G} \Delta A \end{aligned} \quad (13)$$

In the first line we have used Einstein's equations, in the second we've used an identity that comes from playing with the focusing equation (*i.e.* the focusing equation expanded to first order in the

¹Most of the proof is as follows (thanks go to Julian Sonner for help with this): the statement that the k^a are tangent vectors to affinely parametrized geodesics means that $k^b \nabla^a k^a = 0$ (the tangents are parallelly-transported and don't change length). On the other hand, a definition of surface gravity (for a non-rotating black hole; for the proof of equivalence to the one we discuss below, see Wald section 12.5 where he proves the zeroth law) is $\xi^a \nabla_a \xi^b = -\kappa \xi^b$. Putting these together gives

$$\xi^a = A e^{-\kappa t} k^a,$$

for some constant A . Next we use the definition of tangent vector and the chain rule:

$$k^a = \frac{dx^a}{d\lambda} = \frac{dt}{d\lambda} \frac{dx^a}{dt} = \frac{dt}{d\lambda} \xi^a \longrightarrow \frac{d\lambda}{dt} = B e^{-\kappa t}$$

for some constant B . Integrating again gives

$$\lambda - \lambda_0 = -\frac{B}{\kappa} e^{-\kappa t}$$

and we set the initial value λ_0 to zero for convenience. This gives:

$$\xi^a = \left(-\frac{A}{B} \right) \lambda \kappa k^a.$$

I actually don't know how to show that the constant is $\frac{A}{B} = -1$.

small change in the mass). In the third, we've integrated by parts, and then used the definition of convergence in the fourth. For the final line, we've used the usual shadiness that comes with manipulating differentials in thermodynamics (*i.e.* no shadiness at all really).

To tie up a loose end, surface gravity can be defined in a few different ways. One can take the near horizon limit of a black hole; the metric will have the form

$$ds^2 = d\rho^2 - \kappa^2 \rho^2 dt^2 \tag{14}$$

for some κ which we call the surface gravity. Alternatively, it is the thing that satisfies the equation

$$\xi^a \nabla_a \xi^b = -\kappa \xi^b \tag{15}$$

For a stationary uncharged, unrotating black hole, $\kappa = \frac{1}{4M}$.

Next time, we'll complete the laws of thermodynamics, and then we'll show how they're wrong (classically) and fix them with quantum mechanics.