BLACK HOLES

IAN LIM
LAST UPDATED JUNE 4, 2019

These notes were taken for the *Black Holes* course taught by Jorge Santos at the University of Cambridge as part of the Mathematical Tripos Part III in Lent Term 2018. I live-TEXed them using Overleaf, and as such there may be typos; please send questions, comments, complaints, and corrections to itel2@cam.ac.uk.
Many thanks to Arun Debray for the LATEX template for these lecture notes: as of the time of writing, you can find him at https://web.ma.utexas.edu/users/a.debray/.

Contents

Lecture 1.

Friday, January 18, 2019

"The integral curves of the timelike Killing vector don't intersect, or else you could go back in time and kill your own grandmother... which would make you a WEIRDO." –Jorge Santos

Note. Some very important administrative details for this course! Lectures will be Monday, Wednesday, *Thursday*, and Friday, with M/W/F lectures from 12:00-13:00 and Thursday lectures from 13:00-14:00. There will be no classes from 4th February to 15th February, due to Prof. Santos anticipating a baby.

Some useful readings include

- Harvey Reall's notes on black holes and general relativity
- Wald's "General Relativity"
- Witten's review, "Light Rays, Singularities and All That"

To begin with, some conventions. Naturally, we set $c = G = 1$. We use the $- + ++$ sign convention for the Minkowski metric. We shall use the abstract tensor notation where tensor expressions with Greek indices μ, ν, σ are only valid in a particular coordinate basis, while Latin indices a, b, c are valid in any basis, e.g. the Riemann scalar is defined to be $R = g^{ab}R_{ab}$, whereas the Christoffel connection takes the form $\Gamma^\mu_{\nu\rho} = \frac{g^{\mu\epsilon}}{2}(g_{\epsilon\nu,\rho} + g_{\epsilon\rho,\nu} - g_{\nu\rho,\epsilon})$. We also define $R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z$.

Stars Black holes are one possible endpoint of a star's life cycle. Let's start by assuming spherical stars. Now, stars radiate energy and burn out. However, even very cold stars can avoid total gravitational collapse because of *degeneracy pressure*. If you make a star out of fermions (e.g. electrons) then the Pauli exclusion principle says they can't be in the same state (or indeed get too close), and it might be that the degeneracy pressure is enough to balance the gravitational forces. When this happens, we call the star a *white dwarf*. It turns out this can only happen for stars up to $1.4M_\odot$ (solar masses). If a star is instead made of neutrons (naturally we call these *neutron stars*) then the pressure of the neutrons can prevent gravitational collapse in a mass range from $1.4M_\odot < 3M_\odot$. *Beyond* $3M_\odot$*, stars are doomed to collapse into black holes.* We'll spend some time understanding this limit.

Spherical symmetry A normal sphere is invariant under rotations in 3-space, $SO(3)$. The line element on the 2-sphere of unit radius is

$$d\Omega_2^2 = d\theta^2 + \sin^2\theta d\phi^2.$$

It is also invariant under reflections sending $\theta \to \pi - \theta$ (the full group $O(3)$), and perhaps some other symmetries.

Definition 1.1. A spacetime (M, g) is *spherically symmetric* if it possesses the same group of isometries as the round two-sphere $d\Omega_2^2$. That is, it has an $SO(3)$ symmetry where the orbits are S^2s (two-spheres). Important remark– there are spacetimes such as Taub-NUT spacetime which enjoy $SO(3)$ symmetry but are *not* spherically symmetric.

In a spherically symmetric spacetime, we shall define a "radius" $r : M \to \mathbb{R}^+$ defined by

$$r(p) = \sqrt{\frac{A(p)}{4\pi}},$$

where $A(p)$ is the area of the S^2 orbit from a point p. This only makes good sense to define under spherical symmetry, but the idea is that we invert the old relationship $A = 4\pi r^2$ to define a radius given an area.

Definition 1.2. A spacetime (M, g) is *stationary* if it admits a Killing vector field K^a which is everywhere timelike. That is,

$$K^a K^b g_{ab} < 0.$$

Using the assumptions of time independence and spherical symmetry, we'll show some constraints on the resulting spacetime. What could the metric look like in such a spacetime? Let us pick a hypersurface Σ which is nowhere tangent to the Killing vector K. We assign coordinates t, x^i where x^i is defined on the hypersurface, and t then describes a distance along the integral curves of K^a through each point on Σ. That is, we follow the curves such that $\frac{dx^a}{dt} = K^a$.

But in this coordinate system, K^a now takes the wonderfully simple form

$$K^a = \left(\frac{\partial}{\partial t}\right)^a$$

Since K^a is a Killing vector, the Lie derivative of the metric with respect to K vanishes, $\mathcal{L}_K g = 0$.[1]

[1]See Harvey Reall's notes for the definition of a Lie derivative– it's just a derivative, "covariant-ized.") In this case, $K^c\partial_c g_{ab} + K^c{}_{,a} g_{cb} + K^c{}_{,b} g_{ac} = 0$. For the Lie derivative of a metric, we get a particularly nice equation in coordinate notation: $(\mathcal{L}_K g)_{ab} = \nabla_a K_b + \nabla_b K_a$. Setting this equal to zero yields the coordinate form of Killing's equation.

With these assumptions, our metric takes the form

$$ds^2 = g_{tt}(x^k)dt^2 + 2g_{ti}(x^k)dtdx^i + g_{ij}(x^k)dx^i dx^j,$$

where $K^a K^b g_{ab} = g_{tt}(x^k) < 0$.

We can also consider *static spacetimes*. Take Σ to be a hypersurface defined by $f(x) = 0$ for some function $f : M \to \mathbb{R}, df \neq 0$. Then df is orthogonal to Σ. The proof is as follows: take t^a to be tangent to Σ. Thus

$$(df)(t) = t(f) = t^\mu \partial_\mu f = 0$$

on Σ since f is constant on Σ. A useful example might be to compute this for the two-sphere.

Now take a general 1-form normal to Σ. This 1-form can be written as

$$n = gdf + fn'. \tag{1.3}$$

That is, on the surface Σ the one-form n is precisely normal to Σ, but if we go off Σ a little bit then we can smoothly extend n off by a bit. We require that g is a smooth function and that n' is smooth but otherwise arbitrary.

Then the differential of n is

$$dn = dg \wedge df + df \wedge n' + f \wedge dn' \implies dn|_\Sigma = (dg - n') \wedge df.$$

We find that $n|_\Sigma = g\,df \implies (n \wedge dn)|_\Sigma = gdf \wedge (dg - n') \wedge df = 0$. So if n is orthogonal to a hypersurface Σ then $(n \wedge dn)_\Sigma = 0$.

The converse is also true (a theorem due to Frobenius)– if n is a non-zero 1-form such that $n \wedge dn = 0$ everywhere, then there exists f, g such that $n = g\,df$ so n is hypersurface-orthogonal.

Definition 1.4. A spacetime is *static* if it admits a hypersurface-orthogonal timelike Killing vector field. In particular, static $\implies$ stationary.

In practice, for a static spacetime, we know that K^a is hypersurface orthogonal, so when defining coordinates we shall choose Σ to be orthogonal to K^a. Equivalently, this means that we can choose a hypersurface Σ to be the surface $t = 0$, which implies that $K_\mu \propto (1, 0, 0, 0)$.[2] Of course, this means that $K_i = 0$. But recall that

$$K^a = \left(\frac{\partial}{\partial t}\right)^a \implies K^a g_{ai} = g_{ti} = K_i = 0.$$

So our generic metric simplifies considerably– the cross-terms g_{ti} go away and spherical symmetry will further constrain the spatial g_{ij} terms.

Lecture 2.

Monday, January 21, 2019

"Not only are we going to make our cow spherical, we're going to shoot it down so it doesn't move."
–Jorge Santos

So far, we discussed two key concepts. We discussed the condition of a spacetime being static, i.e. stationary and enjoying invariance under $t \leftrightarrow -t$, which forced the line element to take the form

$$ds^2 = g_{tt}(x^k)dt^2 + g_{ij}(x^k)dx^i dx^j. \tag{2.1}$$

We also required spherical symmetry, i.e. the $SO(3)$ orbits of points p in the manifold are S^2s.

Let us see what we get when we combine static and spherically symmetric solutions. We know from staticity that there is a timelike Killing vector $K = \left(\frac{\partial}{\partial t}\right)^a$. Suppose we take a hypersurface Σ_t which is normal to our timelike Killing vector K. Then take any point $p \in \Sigma_t$. By spherical symmetry, its $SO(3)$ orbit is a sphere S^2. Assign angular coordinates θ, ϕ on the S^2 orbit. Using spacelike geodesics normal to the sphere S^2, we can then extend θ, ϕ to the entire hypersurface, a process which is shown in Fig. 1.

Thus our metric on Σ_t takes the form

$$ds^2_{\Sigma_t} = e^{2\psi(r)}dr^2 + r^2 d\Omega_2^2, \tag{2.2}$$

[2] Note the index is down! These are the components of K with the lowered index, i.e. considered as a one-form $K_\mu = (dt)_\mu$.

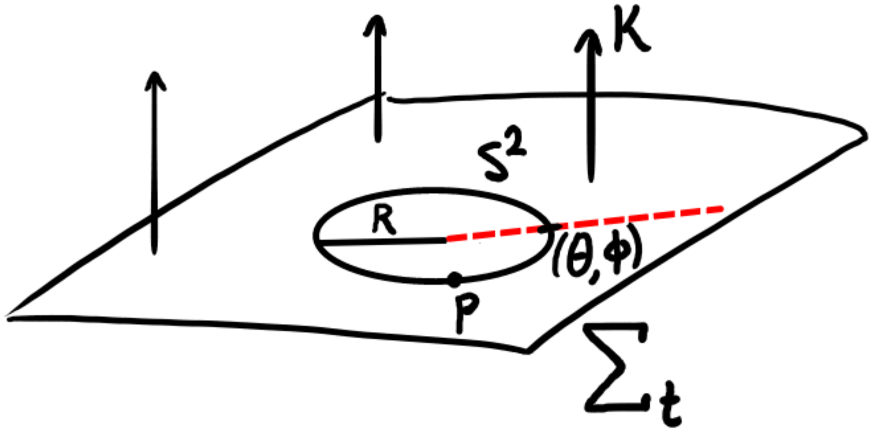

FIGURE 1. An illustration of our coordinates for static, spherically symmetric solutions. We can always choose a hypersurface Σ_t which is orthogonal to the timelike Killing vector K. On Σ_t, choose a point p and trace out its S^2 orbit (drawn here as a circle, S^1) under the action of the $SO(3)$ symmetry. On the S^2 orbit, we can define angular coordinates (θ, ϕ), and we can then extend these to the rest of Σ_t by defining θ, ϕ to be constant on spacelike geodesics normal to the S^2 orbit (red dashed line). The radial coordinate r is given by the area formula $r = \sqrt{A(p)/4\pi}$. This defines coordinates on Σ_t, which we can extend to the entire manifold by following the integral curves of K.

where the coefficient of dr^2 must only depend on r by spherical symmetry, and r is given by our old area relation, $r : \mathcal{M} \to \mathbb{R}^+$ with $r(p) = \sqrt{\frac{A(p)}{4\pi}}$. Now using the property our spacetime is static, we can write down the full spacetime metric,

$$ds^2 = -e^{2\Phi(r)}dt^2 + ds^2_{\Sigma_t}. \tag{2.3}$$

So far we have two degrees of freedom, $(\psi(r), \Phi(r))$. Let's now put some physics in and consider a fluid in our spherically symmetric spacetime. For fluids, recall that the stress-energy tensor takes the form

$$T_{ab} = (\rho + p)U_a U_b + p g_{ab} \tag{2.4}$$

where U_a is a four-velocity, ρ is an energy density and p is a pressure. By spherical symmetry, ρ and p can only be functions of the radial coordinate r, so $\rho = \rho(r)$ and $p = p(r)$. The four-velocity is always timelike, so

$$U^a U_a = U^a U^b g_{ab} = -1 \implies U^a = e^{-\Phi}\left(\frac{\partial}{\partial t}\right)^a \tag{2.5}$$

so that $p, \rho > 0$. This is an energy condition.

We want to describe spherical stars (with finite spatial extent), so outside the star both the pressure and energy density must vanish,

$$p = \rho = 0 \text{ for } r > R \tag{2.6}$$

with R the radius of the star. Now, we know that the defining property of stress-energy is that it is conserved– $\nabla^a T_{ab} = 0$. But the Einstein equation says that

$$R_{ab} - \frac{R}{2} g_{ab} = T_{ab}, \tag{2.7}$$

and by the contracted Bianchi identity we know that the divergence of the LHS always vanishes, so it suffices to look at the Einstein equation since it automatically implies the conservation equation for fluids. This is not generally true for other energy content since there may be other equations of motion that apply.

Let's look at a specific example, the tt component of the Einstein equations.

$$G_{tt} = \frac{e^{2(\Phi-\psi)}}{r^2}[e^{2\psi} + 2r\psi' - 1], \tag{2.8}$$

where the prime indicates a $\frac{\partial}{\partial R}$.

Let us also define a function $m(r)$, given by

$$e^{2\psi} \equiv \left[1 - \frac{2m(r))}{r}\right]^{-1}. \tag{2.9}$$

From the various components we learn that

$$tt : m' = 4\pi r^2 \rho(r) \tag{2.10}$$

$$nn : \Phi' = \frac{m + 4\pi r^3 p}{r(r - 2m)} \tag{2.11}$$

$$\theta\theta : p' = -(p + \rho)\frac{m + 4\pi n^3 p}{r[r - 2m(r)]}. \tag{2.12}$$

We call these the Tollman-Oppenheimer-Volkoff equations (TOV for short). We have three equations but four unknowns: m, Φ, p and ρ. We need one more bit of information– namely, an equation of state relating the pressure and energy density. Normally, p depends on ρ and also T the temperature. But for our purposes, we will assume cold stars so that p is only a function of the energy density ρ.

What can we figure out before imposing any sort of conditions on $p(\rho)$? Well, outside the star, $r > R$, we have $p = \rho = 0$. But we see immediately that in the region $\rho = 0$, the tt equation tells us that $m' = 0 \implies m = M$ for some constant M. This in turn implies that

$$\psi(r) = -\frac{1}{2}\log\left(1 - \frac{2M}{r}\right) = -\Phi(r). \tag{2.13}$$

We can now write down the line element in the exterior region,

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \frac{dr^2}{1 - \frac{2M}{r}} + r^2 d\Omega_2^2. \tag{2.14}$$

This is the *Schwarzschild line element.* We identify the parameter M with the mass of the system.

There's a bit of physics to extract from this– for *stars*, we need $R > 2M$ to keep the signs correct in the metric. For the sun, we have $2M_\odot = 3\,\text{km}$ and $R \simeq 7 \times 10^5\,\text{km}$, so this is a (very loose) bound which is easily satisfied.

Inside the star, life is not so easy. The mass now depends on the radius, and it has a solution

$$m(r) = 4\pi \int_0^r \rho(\tilde{r})\tilde{r}^2 d\tilde{r} + m_*, \tag{2.15}$$

with m_* some integration constant. Fortunately, we note that by physical concerns, $m(r) \to 0$ as $r \to 0$ in order to preserve regularity (the metric should look flat), which tells us that this integration constant is zero, $m_* = 0$.

At the surface of the star ($r = R$), the metric is continuous. This tells us that

$$M = 4\pi \int_0^R \rho(r) r^2 dr, \tag{2.16}$$

so the mass M is related to an integration of the energy density. It is not however the total energy, which is given by

$$E = \int_V \rho r^2 \sin\theta e^{\psi} > M.$$

The total energy differs by a factor which corresponds to the gravitational binding energy.

Restoring units to our $R < 2M$ bound on the star radius, we write

$$\frac{GM}{c^2 R} < \frac{1}{2}. \tag{2.17}$$

This isn't hard to satisfy but it's a start, considering we haven't assumed anything about the equations of state.

Let's add some assumptions. For reasonable matter,

$$\frac{dp}{d\rho} > 0, \quad \frac{dp}{dr} \leq 0 \implies \frac{d\rho}{dr} \leq 0. \tag{2.18}$$

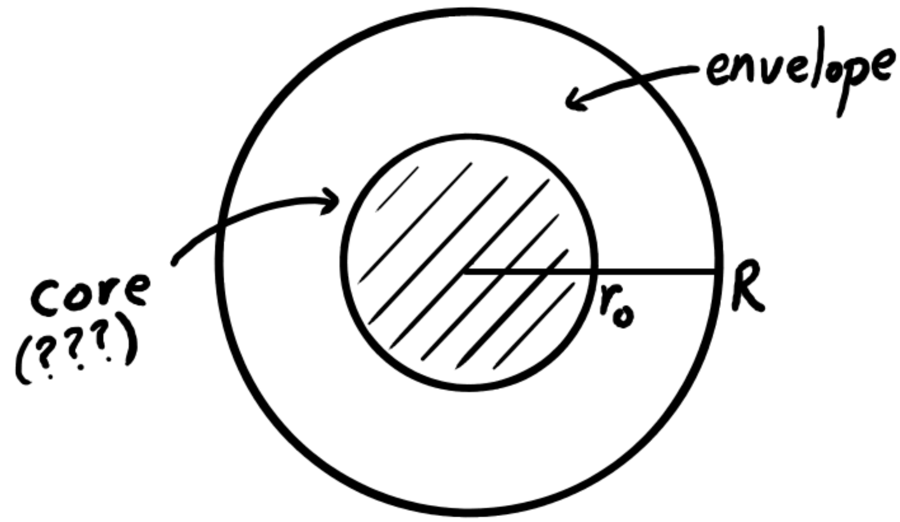

FIGURE 2. A schematic drawing of the interior + envelope model for a star. The interior region extends from $0 < r \leq r_0$ and the exterior region (envelope) goes from $r_0 < r < R$, to the surface of the star.

This first condition says that more stuff (density) means more pressure, and the second says that pressure decreases as we go towards the surface of the star. The $\theta\theta$ component then tells us that

$$\frac{m(r)}{r} < \frac{2}{9}\left[1 - 6\pi r^2 p + (1 + 6\pi r^2 p)^{1/2}\right], \tag{2.19}$$

which we will prove on Example Sheet 1. Knowing that the pressure vanishes at the surface of the star, $r = R$, we arrive at the Buchdahl bound,

$$R > \frac{9}{4}M. \tag{2.20}$$

This already improves on our naïve bound.

Now using the TOV equations, we could just consider the m' and p' equations. Recall that p is a function of ρ, so we can consider these as two first-order equations for p and m. Normally, each of these conditions would require a boundary condition. But recall we have one integration constant (our m_* from earlier) fixed to be zero, so really we just need to specify one boundary condition, $\rho(r = 0)$.

By the form of p', we see that the pressure decreases as we go towards the surface, so we just integrate outwards until p vanishes and we hit the surface of the star at some value R. This tells us that $M(\rho(0))$ and $R(\rho(0))$, so all the physical parameters of the star are fixed by just one number– the energy density at the center of the star, $\rho(0)$.

We could now introduce an equation of state, in principle. But let's try to be a bit more clever and deduce something independent of the equation of state of whatever this star is made of. This star could be super dense in its core, and maybe we don't know anything about physics in the interior, up to some radius r_0. But outside the core there's some envelope region $r_0 < r < R$ where we do know what's happening– see Fig. 2 for an illustration.

What could happen in the interior? If ρ takes on some value $\rho(r_0) = \rho_0$ on the surface of the core, then by integrating we can put the bound

$$m_0 \geq \frac{4\pi}{3} r_0^3 \rho_0 \tag{2.21}$$

on m_0 the mass contained in the core. That is, in the best case $\rho(r)$ is constant in the core region– the star certainly cannot be less dense in its core. But we have another inequality on $m(r)$, the Buchdahl limit 2.19, which we can see is a decreasing function of p. So we evaluate this condition at $r = r_0, m(r_0) = m_0$, noting that the most general bound we can put on m_0 in terms of r_0 occurs when $p = 0$. We find that

$$\frac{m_0}{r_0} < \frac{4}{9}. \tag{2.22}$$

These two inequalities in the space of core masses m_0 and core radii r_0 plus a value for the core density ρ_0 tell us that there is a limit on the total core mass– taking $\rho_0 = 5 \times 10^{14}\,\text{g}/\text{cm}^3$, the density of nuclear matter, we find that $m_0 < 5M_\odot$. Strictly, these are only limits on the core mass, but it turns out that the envelope region is generally insignificant, so

$$M \approx m_0 < 5M_\odot. \tag{2.23}$$

Lecture 3.

Wednesday, January 23, 2019

"If anyone starts calculating geodesics from here [the geodesic equation], I will give you a zero! You should pay for your own stupidity." –Jorge Santos

Today is an exciting class! For the first time, we will look at the Schwarzschild line element. We will not actually define what a black hole is for another ten lectures, but we'll plow ahead and learn some thing about them anyway.

In Schwarzschild coordinates $t, r, \theta, \phi)$, the line element of a Schwarzschild black hole takes the form

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\Omega_2^2. \tag{3.1}$$

For stars, we said this solution was only valid for $r > 2M$, but as you may already know, the $r = 2M$ singularity is just a coordinate singularity and can be defined away by a good choice of coordinates (as we'll do today). We take the parameter M to be a physical mass right now, and thus $M > 0$ (we will treat the case $M < 0$ later). The radius $r = 2M$ is known as the *Schwarzschild radius.*

Last time, we assumed our solutions were both spherically symmetric and stationary. What if we drop the assumption that the solution is stationary? The answer is the following theorem by Birkhoff.

Theorem 3.2 (Birkhoff's theorem)**.** *Any spherically symmetric solution of the vacuum Einstein equations is isometric to Schwarzschild.*

Proof. What could the spherically symmetric line element look like? The angular bit must look like $r^2 d\Omega_2^2$. The rest of it is unconstrained– a priori, we could have dt^2, dr^2, and $drdt$ cross terms. Thus

$$ds^2 = -f(t,r)[dt + \chi dr]^2 + \frac{dr^2}{g} + r^2 d\Omega_2^2. \tag{3.3}$$

Here, χ, f, and g can all depend on t, r. Let us begin by rescaling time t to set the factor $\chi = 0$. If you like, $dt + \chi dr = dt'$. We can still send $t \to p(t)$ a function of t. Now our line element reduces to

$$ds^2 = -f(t,r)dt^2 + \frac{dr^2}{g(t,r)} + r^2 d\Omega_2^2. \tag{3.4}$$

Let us now require that this metric solves the vacuum Einstein equations. Computing the Ricci tensor for the generic metric 3.4 is a bit of a pain (it is done in Carroll, for instance) but the upshot is this. From the tr component of the vacuum Einstein equations, we find that

$$0 = R_{tr} - g_{tr}\frac{R}{2} \implies \frac{\partial}{\partial t} g(t,r) = 0 \implies g(t,r) = g(r), \tag{3.5}$$

i.e. $g(r)$ does not depend on time. Looking at the tt component we get

$$1 - g(r) - rg'(r) = 0 \implies g(r) = 1 - \frac{2M}{r}, \tag{3.6}$$

where M is an integration parameter. From the rr component we get instead

$$1 - \frac{1}{g} + r\frac{f'}{f} = 0 \implies f(t,r) = C(t)\left[1 - \frac{2M}{r}\right]. \tag{3.7}$$

This is almost good– all we need to do is reparametrize t and we can set this function $C(t) = 1$. We find that with f, g defined in this way, our general spherically symmetric metric has been put into the form

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\Omega_2^2. \qquad \boxtimes$$

Gravitational redshift Suppose we have two observers Alice and Bob. They sit at some constant coordinates (r_A, θ, ϕ) and (r_B, θ, ϕ) respectively. Now Alice sends some signals (light pulses) to Bob, separated by a coordinate time Δt. Because $\frac{\partial}{\partial t}$ is a Killing vector, our scenario has time translation symmetry– each pulse will be the same as the last, just translated in time, so Alice and Bob agree on the difference in coordinate time Δt.

Let us notice that Alice measures the photons as being separated by a proper time

$$\Delta\tau_A = \sqrt{1 - \frac{2M}{r_A}}\Delta t. \tag{3.8}$$

An equivalent expression is true for Bob at r_B. Eliminating Δt, we find that

$$\frac{\Delta\tau_B}{\Delta\tau_A} = \sqrt{\frac{1 - \frac{2M}{r_B}}{1 - \frac{2M}{r_A}}} > 1 \tag{3.9}$$

for $r_B > r_A$. This interval is a proxy for a perceived wavelength, $\lambda_B > \lambda_A$. If Bob is far away, $R_B \gg 2M$, then the measured redshift z is given by

$$1 + z \equiv \frac{\lambda_B}{\lambda_A} = \frac{1}{\sqrt{1 - \frac{2M}{r_A}}}. \tag{3.10}$$

For an observer outside a physical star, $R = 9M/4$, $z = 2$ is the maximum redshift Bob can measure as Alice gets close to the surface of the star. But $R = 2M$ is a surface of *infinite* redshift, so it seems to represent something very strange. This is the *event horizon*. We'll explore the consequences of this calculation more as we go on.

Geodesics of Schwarzschild We'll introduce some nice coordinates to let us cross the event horizon. Let $x^\mu(\lambda)$ be an affinely parametrized geodesic with tangent vector

$$U^\mu \equiv \frac{dx^\mu}{d\tau}.$$

Since we have two Killing vectors $K = \frac{\partial}{\partial t}$ and $m = \frac{\partial}{\partial \phi}$, we get two conserved charges,

$$E = -K \cdot U = \left(1 - \frac{2M}{r}\right)\frac{dt}{d\tau} \tag{3.11}$$

$$h = m \cdot U = r^2 \sin^2\theta \frac{d\phi}{d\tau}. \tag{3.12}$$

Note that $U^a \nabla_a U_b = 0$ defines an affinely parametrized geodesic, and Killing vectors obey $\nabla_a K_b + \nabla_b K_a = 0$. These two facts are enough to prove the existence of a conserved charge:

$$U^c \nabla_c (U^a K_a) = (U^c \nabla_c U^a) K_a + U^c U^a \nabla_c K_a. \tag{3.13}$$

But this first term vanishes by the definition of an affinely parametrized geodesic, and the second can be symmetrized since U^c and U^a commute, so the second term vanishes by Killing's equation.[3] Thus $\nabla_U (U^a K_a) = 0$, so $U^a K_a$ is conserved along geodesics.

For timelike particles, if τ is the proper time, then E has the interpretation of energy per unit mass, and h is the angular momentum per unit mass. For null geodesics, the quantity

$$b = \left|\frac{h}{E}\right|$$

[3] Explicitly, $U^a U^c \nabla_c K_a = \frac{1}{2}(U^a U^c \nabla_c K_a + U^c U^a \nabla_c K_a) = \frac{1}{2} U^a U^c (\nabla_c K_a + \nabla_a K_c) = 0$, where we've just relabeled the dummy indices on the second term.

is the physical impact factor.

Note that there is a third conserved charge along geodesics– it is the value $\sigma = -U^a U_a = +1, 0$, or -1 depending on if the geodesic is timelike, null, or spacelike.

We can write an action

$$S = \int d\tau \, \dot{x}^a \dot{x}^b g_{ab}. \tag{3.14}$$

which naturally gives us a Lagrangian that we can apply our conserved charges to and calculate Euler-Lagrange equations for. Here, dots indicate derivatives with respect to the proper time τ. In Schwarzschild, the action takes the form

$$S = \int d\tau \left[g_{tt}\dot{t}^2 + g_{rr}\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2\theta \dot{\phi}^2 \right]. \tag{3.15}$$

Note also that

$$\frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} = 0 \implies r^2 \frac{d}{d\tau}(r^2\dot{\theta}) - \frac{\cos\theta}{\sin^3\theta} h^2 = 0. \tag{3.16}$$

WLOG we can always set $\theta(0) = \pi/2$ and $\dot{\theta}(0) = 0$, which means that $\ddot{\theta} = 0 \implies \theta(\tau) = \pi/2$ for all τ, so it suffices to consider orbits in the equatorial plane.

Substituting in our conserved quantities, we have only one equation which we need to consider: $\dot{r}$. For a geodesic such that

$$g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu = -\sigma, \tag{3.17}$$

we have

$$\dot{r}^2/2 + V(r) = E^2/2, \text{ with } V(r) = \frac{1}{2}\left(\sigma + \frac{h^2}{r^2}\right)\left(1 - \frac{2M}{r}\right). \tag{3.18}$$

What are the null radial free-fall paths of Schwarzschild? For radial trajectories, $h = 0$, and for null geodesics $\sigma = 0$, so the radial equation becomes

$$\dot{r} = \pm 1. \tag{3.19}$$

We can also rescale the affine parameter to normalize $E = 1$ so that

$$\dot{t} = \frac{1}{1 - \frac{2M}{r}}. \tag{3.20}$$

For the upper sign, $\dot{r} = +1$, we have $\dot{t}/\dot{r} > 0, r > 2M$, which represents outgoing trajectories. For the lower sign, $\dot{t}/\dot{r} < 0, r > 2M$ (ingoing). In any case, when $\dot{r} = -1$, we see that we can reach $r = 2M$ without a problem.[4]

In Schwarzschild, we can then write

$$\frac{dt}{dr} = \pm\left(1 - \frac{2M}{r}\right)^{-1}. \tag{3.21}$$

We define Regge-Wheeler coordinates (AKA tortoise coordinates) as follows:

$$dr_* = \frac{dr}{1 - \frac{2M}{r}} \implies r_* = r + 2M \log\left|\frac{r}{2M} - 1\right|. \tag{3.22}$$

We will show that in classical gravity, we can indeed cross the horizon in finite proper time.

Lecture 4.

Thursday, January 24, 2019

"This [tidal force divergence] is related to something that happens in AdS/CFT. ... don't tell anyone I mentioned that." –Jorge Santos

[4]At least from a pure gravity perspective. Once we introduce quantum mechanical effects, we have to worry about firewalls and such, and all bets are off. More on this later.

Quick announcement– office hours from this course will be held at 14:00 on Tuesdays. If you plan to attend the office hours, however, do send an email in advance.

Last time, we started looking at the geodesics of Schwarzschild. We found two particularly simple ones: the ingoing and outgoing radial null geodesics, with equation

$$\frac{dt}{dr} = \pm\left(1 - \frac{2M}{r}\right)^{-1}, \quad r > 2M. \tag{4.1}$$

We also defined the tortoise (Regge-Wheeler) coordinate r_* such that

$$dr_* = \frac{dr}{1 - \frac{2M}{r}}. \tag{4.2}$$

Thus our null geodesic equation becomes

$$\frac{dt}{dr_*} = \pm 1. \tag{4.3}$$

In tortoise coordinates, null geodesics therefore obey

$$t = \pm r_* + \text{constant}. \tag{4.4}$$

This seems kind of trivial. But now we introduce a new coordinate, the *ingoing Eddington-Finkelstein coordinate*, defining

$$v \equiv t + r_*. \tag{4.5}$$

This is clearly constant on ingoing geodesics, just by looking at the previous equation. That is, $dv = dt + dr_* = 0$. Good.[5] Now we eliminate the original coordinate time t from the line element using $dt = dv - \frac{dr}{1-\frac{2M}{r}}$. We get the new line element

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dv^2 + 2dvdr + r^2 d\Omega_2^2. \tag{4.6}$$

Let's write this in matrix notation. Nothing fancy.

$$g_{\mu\nu} = \begin{pmatrix} -\left(1 - \frac{2M}{r}\right) & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2 \sin^2\theta \end{pmatrix}. \tag{4.7}$$

We haven't done anything too extreme, just made a change of coordinates. But we see something very nice– none of the metric components are singular at $r = 2M$. In fact, the determinant of the metric is still perfectly nice at $r = 2M$– by an explicit computation, $\det g = -r^4 \sin^2\theta$. This only vanishes at $\theta = 0$, which is the regular coordinate badness[6] of spherical coordinates near the poles, and at $r = 0$, which may be a real problem (a priori, we don't know yet).

So we have found some coordinates which appear to extend r not just from $r > 2M$ but to all $r > 0$. Our metric is real and analytic (i.e. we've shown the determinant is nonsingular) so it is a nice analytic continuation of the old (bad) Schwarzschild coordinates. This is related to the problem of *extendibility*– are there other metrics which cover more of the spacetime manifold which are compatible with the solution that we've found?

However, something really bad does happen as $r \to 0$. The Kretchmann scalar $R^{abcd}R_{abcd} = \frac{48M^2}{r^6} \to \infty$ as $r \to 0$, and scalars by definition are invariant under coordinate transformations. So we cannot get rid of this by a simple redefinition of coordinates.

Moreover, let us observe that our Killing vector field becomes

$$K = \frac{\partial}{\partial t} = \frac{\partial x^\mu}{\partial t}\frac{\partial}{\partial x^\mu} = \frac{\partial}{\partial v} \implies K^2 = g_{vv} = -\left(1 - \frac{2M}{r}\right). \tag{4.8}$$

So our metric appears to be no longer static or stationary in these coordinates.

[5] We could have done the same for outgoing geodesics taking $u \equiv t - r_*$, and indeed we will do so later.

[6] AKA degeneracy. Basically, what is the value of ϕ at the north pole? It's not well-defined, hence our coordinates do not define an invertible map from coordinates to manifold points.

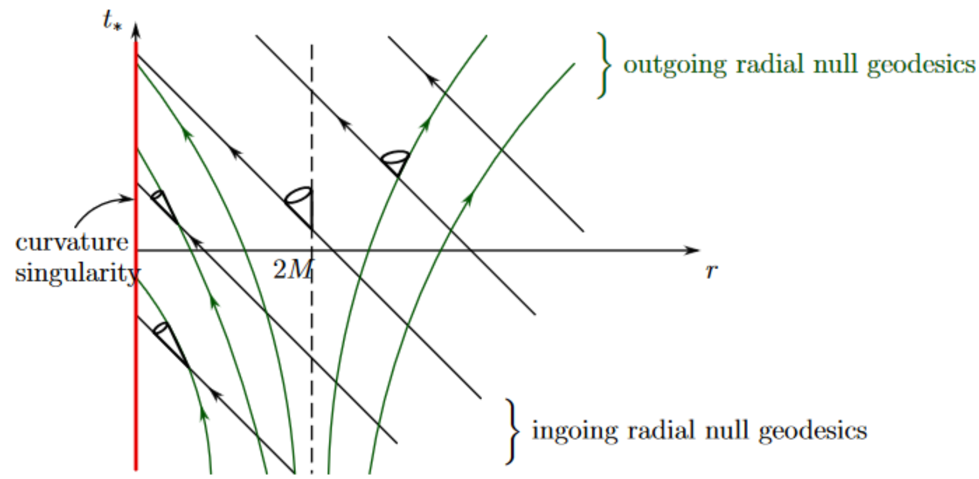

FIGURE 3. The Finkelstein diagram for the Schwarzschild spacetime. The sides of the light cones are given by lines of constant u and v. Notice the light cones "tip over" as we cross the event horizon at $r = 2M$, so that even "outgoing" radial null trajectories must proceed towards the $r = 0$ singularity.
Image credit to Prof. Reall's Black Holes notes, §2.5.

We now introduce the Finkelstein diagram, shown in Figure 3. Recall that for $r = 2M$, on outgoing geodesics we have $t - r_* = \text{constant} \implies v = 2r + 4M \log\left|\frac{r}{2M} - 1\right| + \text{constant}$. Let us draw a plot in the $t_* \equiv v - r, r$ plane. What we see is that the ingoing geodesics follow 45° paths, while the outgoing geodesics follow some curved paths in these coordinates.

In these coordinates, the light cones "tip over" at $r = 2M$– even the "outgoing" null geodesics are forced to proceed towards $r = 0$ for $r < 2M$. Now, this does not yet prove that this is a black hole. We've only examined radial geodesic trajectories, and it's not clear that adding even a little bit of angular momentum can't save us from a spaghettified death by black hole.[7]

Suppose that we[8] sit on the surface of a collapsing star. What we can show is that the time to singularity is $\Delta t = \mathbb{T}M$, where $M = M_\odot \implies \Delta t = 10 \times 10^{-5}$ s. It's a weird fact that we can indeed cross the horizon and hit the singularity in finite time, and yet because the event horizon is a surface of infinite redshift, a far-away observer will never see us actually cross the event horizon.

Black hole region How do we define a black hole? Before we can do that, we'll need some preliminaries.

Definition 4.9. A vector is *causal* if it is null or timelike and nonzero. A curve is causal if its tangent vector is everywhere causal.

Definition 4.10. A spacetime is *time-orientable* if it admits a time orientation, i.e. a causal vector field T^a. Another causal vector field x^a is then *future-directed* if it lies in the same light cone as T^a (i.e. $x^a T_a \leq 0$) and is past-directed otherwise.

Note that the old coordinate time t in Schwarzschild turned out to be a bad choice inside the event horizon. This is because $\frac{\partial}{\partial t}$ becomes spacelike inside the horizon, related to the change of sign of g_{tt} for $r < 2M$. So let us instead take $\pm\frac{\partial}{\partial r}$, and note that in our EF coordinates, $g_{rr} = 0$. Therefore $\frac{\partial}{\partial r}\frac{\partial}{\partial r} = 0$, meaning that we've found a vector field which is null everywhere.

In fact, recall that our timelike Killing vector K gave us a good sense of a time direction outside. We see that

$$K \cdot \left(-\frac{\partial}{\partial r}\right) = -g_{vr} = -1,$$

7 Strictly, death by "black hole-like object" at this point, in the same sense that the Higgs boson was first reported as a "Higgs-like particle." That is to say, out of an abundance of caution.

8 "Imagine that you, not me, are situated at the surface of this star. I don't want to go there." – Jorge Santos

where $K \equiv \frac{\partial}{\partial v}$. This tells us that $-\frac{\partial}{\partial r}$ and K lie in the same light cone, so we've found a null (hence causal) vector field which is good everywhere and defines a good time orientation for $r > 0$.

Proposition 4.11. *Let $x^\mu(\lambda)$ be any future-directed causal curve, i.e. one whose tangent vector is everywhere future-directed and causal. Assume $r(\lambda_0) \leq 2M$ in the Schwarzschild spacetime. Then $r(\lambda) \leq 2M$ for any $\lambda \geq \lambda_0$.*

Let us define the tangent vector $V^\mu = \frac{dx^\mu}{d\lambda}$. Since V^a is future-directed, we have

$$0 \geq \left(-\frac{\partial}{\partial r}\right) V = -g_{r\mu} V^\mu = -V^v = -\frac{dv}{d\lambda} \implies \frac{dv}{d\lambda} \geq 0. \tag{4.12}$$

Now

$$V^2 \equiv V^a V_a = -\left(1 - \frac{2M}{r}\right)\left(\frac{dv}{d\lambda}\right)^2 + 2\left(\frac{dv}{d\lambda}\right)\left(\frac{dr}{d\lambda}\right) + r^2 \left(\frac{d\Omega}{d\lambda}\right)^2, \tag{4.13}$$

where $\left(\frac{d\Omega}{d\lambda}\right)^2 \equiv \left(\frac{d\theta}{d\lambda}\right)^2 + \sin^2\theta \left(\frac{d\phi}{d\lambda}\right)^2$. Then

$$-2\left(\frac{dv}{d\lambda}\right)\left(\frac{dr}{d\lambda}\right) = -V^2 + \left(\frac{2M}{r} - 1\right)\left(\frac{dv}{d\lambda}\right)^2 + r^2\left(\frac{d\Omega}{d\lambda}\right)^2. \tag{4.14}$$

But if V is causal and $r \leq 2M$, then the right side of 4.14 is non-negative, so it follows that

$$\frac{dv}{d\lambda}\frac{dr}{d\lambda} \leq 0. \tag{4.15}$$

Let us assume that $\frac{dr}{d\lambda} > 0$ (our curve at any point is directed towards larger r). Then $\frac{dv}{d\lambda} = 0$, which means that by 4.14, $V^2 = 0$ and $\frac{d\Omega}{d\lambda} = 0$. But then the only nonvanishing component of V is $V^r = \frac{dr}{d\lambda} > 0 \implies V$ is a positive multiple of $\frac{\partial}{\partial r}$, and hence is past-directed.We have reached a contradiction.

Therefore $\frac{dr}{d\lambda} \leq 0$ if $r \leq 2M$. If $r < 2M$, the equality must be strict. If $\frac{dr}{d\lambda} = 0$ then by 4.14, $\frac{d\Omega}{d\lambda} = \frac{dv}{d\lambda} = 0 \implies V^\mu = 0$. Hence if $r(\lambda_0) < 2M$, then $r(\lambda)$ is monotonically decreasing for all $\lambda \geq \lambda_0$.

Lecture 5.

Friday, January 25, 2019

"[The white hole region is] like a norovirus. Nothing can stay in. Everything has to come out." –Jorge Santos

Last time, we made the proposition that the Schwarzschild spacetime really does feature a black hole region, i.e. a region where any future-directed causal curves with $r(\lambda) \leq 2M$ are forced to have $r(\lambda) \leq 2M \forall \lambda \geq \lambda_0$. We proved last time that if you are inside the horizon, $r(\lambda_0) < 2M$, the $r(\lambda)$ is monotonically decreasing for $\lambda \geq \lambda_0$.

Formally, we must consider the case $r(\lambda_0) = 2M$. Note that if $\frac{dr}{d\lambda} < 0|_{\lambda=\lambda_0}$, then we're done. In the next ϵ time later, we'll be inside the horizon and our proof holds. If $\frac{dr}{d\lambda} = 0$, then we sit at $r = 2M$ forever and we're also done.

So there's only one case we have to consider, $dr/d\lambda > 0$ at $\lambda = \lambda_0$.

We've seen that

$$-2\left(\frac{dv}{d\lambda}\right)\left(\frac{dr}{d\lambda}\right) = -V^2 + \left(\frac{2M}{r} - 1\right)\left(\frac{dv}{d\lambda}\right)^2 + r^2\left(\frac{d\Omega}{d\lambda}\right)^2.$$

(4.14 from last time). At $\lambda = \lambda_0$, this vanishes, so $\frac{d\Omega}{d\lambda} = V^2 = 0$. This means that $\frac{dv}{d\lambda} \neq 0$, or else $V^\mu = 0$ (which is a contradiction). Then

$$\frac{dv}{d\lambda}|_{\lambda=\lambda_0} > 0,$$

since we proved last time that $\frac{dv}{d\lambda} \geq 0$.

Hence at least near $\lambda = \lambda_0$, we can use v instead of λ as a parameter along the curve with $r = 2M$ at $v = v_0 = v(\lambda_0)$. Dividing 4.14 by $\left(\frac{dv}{d\lambda}\right)^2$ gives

$$-2\frac{dr}{dv} \geq \frac{2M}{r} - 1 \implies 2\frac{dr}{dv} \leq 1 - \frac{2M}{r}. \tag{5.1}$$

Hence for v_2, v_1 greater than v_0 with $v_2 > v_1$, we have

$$2\int_{r(v_1)}^{r(v_2)} \frac{dr}{1-\frac{2M}{r}} \leq v_2 - v_1. \tag{5.2}$$

This completes the $r = 2M$ case. ☒

Technically, this doesn't prove that we have a black hole because it is a local statement. To establish a global event horizon will take more work.

Detecting black holes Note the following facts.

(a) There is no upper bound on the mass of a black hole.
(b) Black holes are very small. For instance, a black hole with the mass of the Earth would have a radius of 0.9 cm.

We observe black holes by noticing their gravitational pull on stars, for instance, and considering their incredible compactness. Fun fact– there are massive and supermassive (billions of solar mass) black holes in the universe which we have detected astrophysically, and we have no idea where they came from. It seems like there isn't enough time in the age of the universe for them to have formed.

Orbits around a black hole Consider a timelike geodesic around a black hole. The turning points of the potential are given by points where $\dot{r} = 0$, i.e. in the radial equation,

$$V(r) = \frac{E^2}{2} = \frac{1}{2}\left(\sigma + \frac{h^2}{r^2}\right)\left(1 - \frac{2M}{r}\right), \tag{5.3}$$

and where

$$V'(r) = 0 \implies r_\pm = \frac{h^2 \pm \sqrt{h^4 - 12h^2\sigma M^2}}{2M\sigma}, \quad \sigma = 1. \tag{5.4}$$

If $h^2 < 12M^2$, then we are in free-fall– there are no turning points. However, if $h^2 > 12M^2$, then there is a local minimum of the potential at r_+, and a local maximum at r_-. One can show that

$$3M < r_- < 6M < r_+, \tag{5.5}$$

where $r = 6M$ is known as the ISCO (innermost stable circular orbit). There is no Newtonian analogue to this– $r = 6M$ lies well within the star.

The energy of the orbit is then

$$E_\pm = \frac{r_\pm - 2M}{r_\pm^{1/2}(r_\pm - 3M)^{1/2}}, \tag{5.6}$$

and for $r_+ \gg M$, we find that

$$E_+ \approx 1 - \frac{M}{2r_+} \to m - \frac{Mm}{(2r_+,)} \tag{5.7}$$

tells us that the energy at this orbit is the relativistic mass-energy $E = m$ minus a correction for the gravitational energy.

Let us approximate the accretion disk of the black hole as non-interacting so that particles basically travel along geodesics. Orbits will radiate off energy, decreasing towards $r \to 0$. When a particle hits $r = 6M$, it will suddenly turn towards the singularity, releasing a burst of energy (cf. brehmsstrahlung). It is these bursts of energy that we can detect in astrophysical systems.

White holes We looked at ingoing null geodesics in EF coordinates. What about outgoing (radial) null geodesics? We have the outgoing EF coordinate

$$u \equiv t - r_*, \tag{5.8}$$

which is constant along radial outgoing null geodesics. Here, $\frac{dt}{dr_*} = 1$. Then the metric takes the form

$$ds^2 = -\left(1 - \frac{2M}{r}\right)du^2 - 2du dr + r^2 d\Omega_2^2. \tag{5.9}$$

Importantly, this is *not* the same region that we found before! We have a new $r < 2M$ region, but outgoing null geodesics are forced to leave this region. However, we just spent all that time proving that geodesics inside the event horizon were trapped inside, so this cannot possibly be the same region as in the ingoing coordinates. For constant u (outgoing null radial geodesics), the metric tells us that

$$\frac{dr}{d\tau} = 1. \tag{5.10}$$

So these can propagate from the curvature singularity at $r = 0$ through the surface $r = 2M$. Weak cosmic censorship says that naked singularities are forbidden in nature, and to define a white hole we must have an initial singularity, so this seems to be unphysical. Another argument for this involves time reversal– it has been proved in some generality that black holes are stable to perturbations, so after perturbation they will settle down to another black hole state. If we take the time reversal of this statement, this suggests that white holes will be highly unstable to perturbation, so they are probably not physical.

The Kruskal extension Can we get both the black and white hole regions in a single set of coordinates? Yes! We start in the exterior region $r > 2M$, and define *Kruskal-Szekeres* coordinates (U, V, θ, ϕ), defining

$$U = -e^{-u/4M}, \quad V = e^{v/4M} \tag{5.11}$$

in terms of the EF coordinates from before. So for the exterior region $r > 2M$, $U < 0$ and $V > 0$.

Now we can directly compute the product

$$UV = -e^{r/2M}\left(\frac{r}{2M} - 1\right) = -e^{r/2M}\left(\frac{r}{2M}\right)\left(1 - \frac{2M}{r}\right). \tag{5.12}$$

Let us observe that the RHS of this equation is a monotonic function of r, and hence determines $r(U, V)$ uniquely. We also have

$$\frac{V}{U} = -e^{t/2M}, \tag{5.13}$$

which implicitly determines $t(U, V)$.[9]

Computing the differentials dU, dV in terms of du, dv, we find that

$$dU = -U\frac{du}{4M}, \quad dV = V\frac{dv}{4M}. \tag{5.14}$$

Therefore

$$\begin{aligned} dUdV &= \frac{1}{16M^2}(-UV)dudv \\ &= \frac{1}{16M^2}e^{r/2M}\frac{r}{2M}\left(1 - \frac{2M}{r}\right)dudv. \end{aligned}$$

To get this in terms of the original t, r coordinates, notice that since $u = t - r_*, v = t + r_*$, we have $du = dt - dr_*, dv = dt + dr_*$ and therefore

$$dudv = dt^2 - dr_*^2 = dt^2 - \frac{dr^2}{\left(1 - \frac{2M}{r}\right)^2}. \tag{5.15}$$

It follows that $\left(1 - \frac{2M}{r}\right)dudv = \left(1 - \frac{2M}{r}\right)dt^2 - \frac{dr^2}{\left(1 - \frac{2M}{r}\right)}$, so our metric takes the form

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \frac{dr^2}{\left(1 - \frac{2M}{r}\right)} + r^2 d\Omega^2 \tag{5.16}$$

$$= -\left(1 - \frac{2M}{r}\right)dudv + r^2 d\Omega^2 \tag{5.17}$$

$$= \frac{32M^3}{r(U,V)}e^{-r(U,V)/2M}dUdV + r^2(U,V)d\Omega^2, \tag{5.18}$$

which is exactly the metric in terms of the Kruskal-Szekeres coordinates (with r given implicitly in terms of U and V as above).[10]

[9] Explicitly, $t = 2M\log(-V/U)$. This is sensible in the exterior region since $U < 0, V > 0$ and so $V/U < 0$.

[10] I've cleaned up this derivation a bit from how we did it in class. Here is the original derivation, for reference.

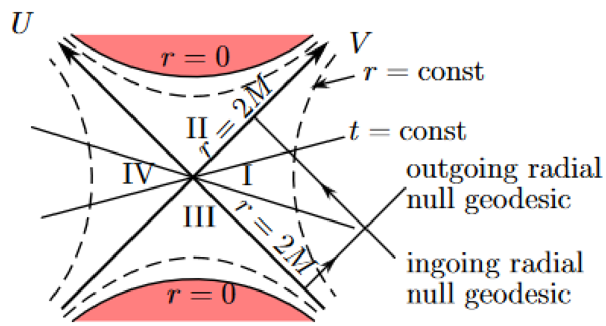

FIGURE 4. The Kruskal diagram for the Schwarzschild spacetime. There are four regions of interest: I, the original exterior; II, the black hole; III, the white hole; and IV, the mirror exterior region. (I'm not sure if there's a standard name for this.)

Lecture 6.

Monday, January 28, 2019

"What about the bifurcation sphere?" "It will be the throat of an Einstein-Rosen bridge. Those of you who watched Thor... this will be familiar." –a student and Jorge Santos

Last time, we tried to extend the Schwarzschild spacetime. Our first attempt took us inside the black hole with the Eddington-Finkelstein coordinates, and we then defined the Kruskal extension by a clever rescaling of the EF coordinates. Defining Kruskal coordinates U and V, we observe that

$$UV = -e^{\frac{+r}{2M}}\left(\frac{r}{2M}-1\right), \quad \frac{V}{U} = -e^{t/2M}. \tag{6.1}$$

Note that $U < 0$ and $V > 0$ for $r > 2M$ in Schwarzschild coordinates. But here's another surprise– if we switch the signs of both U *and* V, we can still sensibly define $r(U,V)$ for $U > 0, V < 0$ via 6.1.

This is an entirely new region of spacetime, which is isometric to our black hole exterior region with one notable caveat– time runs backwards. Notice that at the event horizon, $r = 2M \implies UV = 0$, so either $U = 0$ or $V = 0$. These lines intersect at $U = V = 0$, and note that at $r = 0$, we have $UV = +1$.

If we're interested in AdS/CFT, we should take all four regions seriously. However, if we're interested in astrophysical black holes, then most of the diagram is really the interior of the star as it proceeds to collapse along a geodesic.

Computing the differentials, we find that

$$dU = \frac{1}{4M}e^{-u/4M}du, \quad dV = \frac{1}{4M}e^{v/4M}dv.$$

Therefore

$$\begin{aligned} dUdV &= \frac{1}{16M^2}\exp\left(\frac{v-u}{4M}\right)dudv \\ &= \frac{1}{16M^2}\exp\left(\frac{r_*}{2M}\right)(dt^2 - dr_*^2) \\ &= \frac{1}{16M^2}\exp\left(\frac{r_*}{2M}\right)\left[dt^2 - \frac{dr^2}{(1-\frac{2M}{r})^2}\right]. \end{aligned}$$

This is almost what we want– now we just multiply by the appropriate factors to find that our new line element is

$$ds^2 = -32M^3\frac{\exp\left(-\frac{r(U,V)}{2M}\right)}{r(U,V)}dUdV + r^2(U,V)d\Omega_2^2.$$

Now recall that we have a Killing vector[11]

$$K = \frac{\partial}{\partial t} = \frac{1}{4M}\left(V\frac{\partial}{\partial V} - U\frac{\partial}{\partial U}\right). \tag{6.2}$$

What happens at the point $U = V = 0$?

The Einstein-Rosen bridge Let us make the following coordinate transformation. Define the coordinate ρ by

$$r = \rho + M + \frac{M^2}{4\rho} \tag{6.3}$$

so that as $\rho \to +\infty, r \to +\infty$. Naturally this is quadratic(ish) in ρ so we'll get two solutions for each value of r. We choose $\rho > M/2$ in region I, the exterior region, and $0 < \rho < M/2$ in region IV. Then our metric becomes[12]

$$ds^2 = -\left[\frac{1-\frac{M}{2\rho}}{1+\frac{M}{2\rho}}\right]^2 dt^2 + \left(1+\frac{M}{2\rho}\right)^4 \underbrace{(d\rho^2 + \rho^2 d\Omega_2^2)}_{\mathbb{R}^3} \tag{6.4}$$

So we've put the metric in a form where the spatial part looks like $\mathbb{R}^3$ up to a scaling factor. It's an exercise to check that the transformation $\rho \to M^2/4\rho$ leaves this metric unchanged.

Now what does a surface of constant t look like?

$$ds^2_{\Sigma_t} = \left(1+\frac{M}{2\rho}\right)^4 \underbrace{(d\rho^2 + \rho^2 d\Omega_2^2)}_{\mathbb{R}^3}, \tag{6.5}$$

As $\rho \to +\infty$, we get one asymptotically flat region, and as $\rho \to 0$ we get another asymptotically flat region. Connecting them we get a wormhole where there is an S^2 at $\rho = M/2$ with some minimum (nonzero) radius.[13]

Now, note that from our Kruskal diagram, this wormhole is non-traversable. In order to get to region IV, we must enter the black hole region. And once we're in the black hole region (region II), we are bound to hit the singularity. Tough luck.

Extendability and singularities

Definition 6.6. A spacetime $(\mathcal{M}, g)$ is *extendible* if it is isometric to a proper subset of another spacetime $(\mathcal{M}', g)$. The latter is called an *extension* of $(\mathcal{M}, g)$.

Thus the Kruskal spacetime is an extension of the Schwarzschild spacetime, and moreover it is the maximal analytic extension of Schwarzschild.

Let's talk a bit about singularities now. There are different sorts of singularities we might be interested in.

- Coordinate singularities: the metric (or determinant) is not smooth in some coordinate chart. Nothing physically bad has (necessarily) happened, but we just chose bad coordinates.
- Curvature singularities: a curvature scalar becomes singular. These are physically significant, since we cannot define them away by a coordinate choice.

[11]By the chain rule, $\frac{\partial}{\partial t} = \frac{\partial V}{\partial t}\frac{\partial}{\partial V} + \frac{\partial U}{\partial t}\frac{\partial}{\partial U}$. Now

$$\frac{\partial V}{\partial t} = \frac{\partial}{\partial t} e^{v/4M} = \frac{\partial v}{\partial t}\frac{1}{4M} V = \frac{1}{4M} V$$

and

$$\frac{\partial U}{\partial t} = \frac{\partial}{\partial t}(-e^{-u/4M}) = \left(-\frac{\partial u}{\partial t}\right)\frac{1}{4M} U = -\frac{1}{4M} U,$$

so adding it all up we have $\frac{\partial}{\partial t} = \frac{1}{4M}\left(V\frac{\partial}{\partial V} - U\frac{\partial}{\partial U}\right)$.

[12]This is a bit quick– see the end of this lecture's notes for the details.

[13]We know it's asymptotically flat as $\rho \to 0$ by the $\rho \to M^2/4\rho$ symmetry, and the value of ρ that's unchanged by $\rho \to M^2/4\rho$ is $\rho = M/2$. If we were being pretentious, we might say that there's a duality between the $\rho \to 0$ and $\rho \to \infty$ regions, and $\rho = M/2$ is the self-dual point. Fixing $\rho = M/2$, we get $ds^2 = 4M^2 d\Omega_2^2$, i.e. a round 2-sphere metric with radius $2M$.

- Conical singularities: consider the line element
$$g = dr^2 + \lambda^2 r^2 d\phi^2, \tag{6.7}$$
with $\lambda > 0$. We can always make this look like $dr^2 + r^2 d\tilde{\phi}^2$, which looks like $\mathbb{R}^2$. But suppose $\lambda \neq 1$. Then the period of our new $\tilde{\phi} = \lambda\phi$ coordinate is not 2π. If we take a "circle" of radius ϵ, we see that $\frac{\textit{circumference}}{\text{radius}} = \frac{2\pi\lambda\epsilon}{\epsilon} = 2\pi\lambda$, which does not go to 2π as $\epsilon \to 0$. What results is a conical singularity at the origin.
- There are certain metrics such that the components of the Riemann tensor are singular in every coordinate system, which means that the tidal forces become infinite as you approach the singularity. We will see an example of this on the examples sheet.

Definition 6.8. A *curve* is a smooth map from some interval of the real line to our manifold, $\gamma : (a, b) \to \mathcal{M}$, with $a, b \in \mathbb{R}$.

Definition 6.9. A point $p \in \mathcal{M}$ is a *future endpoint* of a future-directed causal curve if for any neighborhood $\mathcal{O}$ of p, there exists t_0 such that $\gamma(t) \in \mathcal{O}$ for all $t > t_0$. That is, our curve gets trapped arbitrarily close to p at late times.

Definition 6.10. We say that γ is future-inextendible if it has no future endpoints. Equivalent notions of past endpoints and past-inextendible can be defined by replacing future with past, and we say a curve is inextendible if it is future-inextendible and past-inextendible.

Example 6.11. Let $\gamma : (-\infty, 0) \to \mathcal{M}$, with $\gamma(t) = (t, 0, 0, 0)$. Clearly, this has a future endpoint at $(0, 0, 0, 0)$. If we remove that point from the manifold, however, there is no future endpoint– no other point on our manifold will satisfy the future endpoint condition since we can always take γ at some time arbitrarily close to zero. So the notion of extendibility depends crucially on our choice of manifold.

Definition 6.12. A *geodesic* is complete if an affine parameter for the geodesic exists to $\pm\infty$. A spacetime is *geodesically complete* if all inextendible causal geodesics are complete. A spacetime is *singular* if if it is inextendible and geodesically incomplete.

Example 6.13. The Kruskal spacetime is inextendible, and it has plenty of geodesics which hit the $r = 0$ singularity and are therefore incomplete. Therefore by our definition, the Kruskal spacetime is singular.[14]

Non-lectured aside: algebra for Einstein-Rosen Actually showing that the coordinate transformation gives the metric 6.4 is left as an exercise in Harvey Reall's notes, so I work out the details here.

First,
$$1 - \frac{2M}{r} = 1 - \frac{2M}{\rho + M + \frac{M^2}{4\rho}} = \frac{\rho - M + \frac{M^2}{4\rho}}{\rho + M + \frac{M^2}{4\rho}} = \frac{1 - \frac{M}{\rho} + \frac{M^2}{4\rho^2}}{1 + \frac{M}{\rho} + \frac{M^2}{4\rho^2}} = \left[\frac{1 - \frac{M}{2\rho}}{1 + \frac{M}{2\rho}}\right]^2.$$

That takes care of the dt^2 coefficient. Next, notice that
$$r = \rho\left(1 + \frac{M}{2\rho}\right)^2,$$

[14] In fact, incomplete geodesics are a sufficient but not necessary condition for a spacetime to be singular. We should also in principle be concerned about accelerating observers falling off the edge of spacetime into a singularity– cf. Hawking and Ellis, and also work by Geroch. According to Hawking and Ellis, the more appropriate notion is not geodesic completeness but a more general idea of "b-completeness," completeness of general inextendible causal curves in the spacetime.

so

$$\begin{aligned} dr &= d\rho\left(1+\frac{M}{2\rho}\right)^2 + \rho\left(2\left(1+\frac{M}{2\rho}\right)\left(-\frac{M}{2\rho^2}\right)d\rho\right) \\ &= d\rho\left(1+\frac{M}{2\rho}\right)^2 + \left(1+\frac{M}{2\rho}\right)\left(-\frac{M}{\rho}\right)d\rho \\ &= d\rho\left[1+\frac{M}{\rho}+\frac{M^2}{4\rho^2}-\frac{M}{\rho}-\frac{M^2}{2\rho^2}\right] \\ &= d\rho\left(1-\frac{M}{2\rho}\right)\left(1+\frac{M}{2\rho}\right). \end{aligned}$$

It follows that

$$\left(1-\frac{2M}{r}\right)^{-1}dr^2 = \left[\frac{1+\frac{M}{2\rho}}{1-\frac{M}{2\rho}}\right]^2\left(1-\frac{M}{2\rho}\right)^2\left(1+\frac{M}{2\rho}\right)^2 d\rho^2 = \left(1+\frac{M}{2\rho}\right)^4 d\rho^2$$

and

$$r^2 d\Omega^2 = \left(1+\frac{M}{2\rho}\right)^4 \rho^2 d\Omega^2.$$

Lecture 7.

Wednesday, January 30, 2019

"I could draw here AdS, but I know most of you here are allergic to that." –Jorge Santos

Today, we will introduce the initial value problem in general relativity. In classical physics, it's natural to say that we set initial conditions and (using some differential equations) evolve them to the future. But in GR, the question is more complicated because we are trying to evolve things to "the future," yet the future is part of what we are solving for! It will take some care to define precisely what we mean.

Definition 7.1. Let $(\mathcal{M}, g)$ be a time-orientable spacetime. A *partial Cauchy surface* Σ is a hypersurface for which no two points $p, p' \in \Sigma$ are connected by a causal curve in $\mathcal{M}$.

Definition 7.2. The *future domain of dependence* of Σ, denoted $D^+(\Sigma)$, is the set $p \in \mathcal{M}$ such that every past-inextendible causal curve through p intersects Σ. The past domain of dependence $D^-(\Sigma)$ is defined equivalently for future-inextendible causal curves, and the domain of dependence is then $D(\Sigma) = D^+(\Sigma) \cup D^-(\Sigma)$.

Example 7.3. Consider Minkowski space, with a partial Cauchy surface $\Sigma : t = 0, x > 0$. Then the domain of dependence $D(\Sigma)$ is the 45-degrees region with $x \geq |t|$– see Fig. 5 for an illustration.

Definition 7.4. As far as we (as physicists) are concerned, a *hyperbolic partial differential equation* (second order) is defined to be a differential equation of the form

$$g^{ef}\nabla_e\nabla_f T^{ab\ldots}{}_{cd\ldots} = \tilde{G}^{ab\ldots}{}_{cd}, \tag{7.5}$$

where the tensor we've put on the RHS depends only on T and first derivatives of T in some smooth way. That is, the second derivative is the highest derivative, and it appears only linearly. These are sometimes called quasilinear equations.

Definition 7.6. A *Cauchy surface* for a spacetime $(\mathcal{M}, g)$ is a partial Cauchy surface Σ with $D(\Sigma) = \mathcal{M}$ the entire manifold. Spacetimes $(\mathcal{M}, g)$ which admit Cauchy surfaces are called *globally hyperbolic*.

Note that if $D(\Sigma) \neq \mathcal{M}$, then the solution of hyperbolic equations will *not* be uniquely specified on $\mathcal{M} \setminus D(\Sigma)$ by data on Σ.

Example 7.7. A very simple example of a spacetime that is not globally hyperbolic is $\mathbb{R}^{1,1} \setminus 0$. There are points where a past-directed inextendible causal curve can drop off the manifold (i.e. hit the point we have removed).

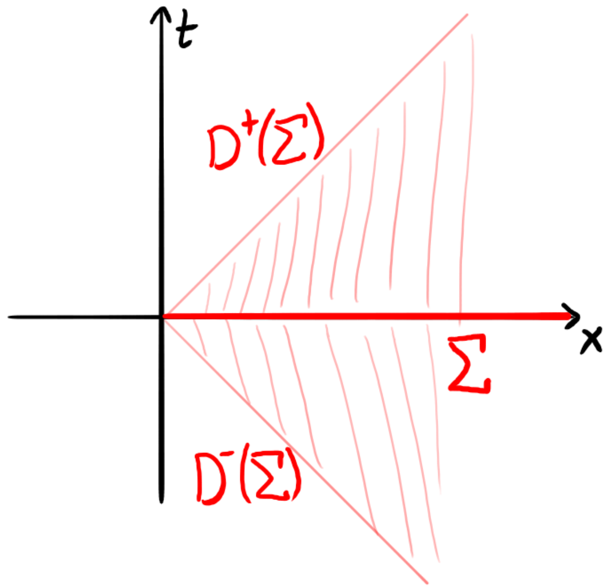

FIGURE 5. In Minkowski space, we select the partial Cauchy surface $\Sigma : t = 0, x > 0$ (in red). Given initial conditions on Σ, we can predict physics in the shaded region $D^+(\Sigma)$ and retrodict in the region $D^-(\Sigma)$. However, Σ is not a Cauchy surface because there are points in spacetime which lie outside the domain of dependence $D(\Sigma)$. Note also that the boundary of $D^+(\Sigma)$ (i.e. the Cauchy horizon) coincides with the boundary of the causal future of points outside Σ.

Theorem 7.8 (Wald). *Let $(\mathcal{M}, g)$ be a globally hyperbolic spacetime. Then*

(i) *There exists a global-time function $t : \mathcal{M} \to \mathbb{R}$ such that $-(dt)^a$ (the normal to surfaces of constant t) is future-directed and timelike.*
(ii) *Surfaces of constant t, Σ_t, are Cauchy surfaces and all have the same topology.*
(iii) *The topology of $\mathcal{M}$ is $\mathbb{R} \times \Sigma$.*

Remark: spacetimes which have singularities can still be globally hyperbolic. Consider the Kruskal spacetime– we can set initial conditions on the Cauchy surface $U + V =$ constant.

Extrinsic curvature Let Σ denote a spacelike or timelike hypersurface with unit normal n_a (i.e. such that $n^a n_a = \pm 1$). We can define a projector onto this hypersurface $h^a{}_b$, which has some nice properties.

Lemma 7.9. *For any $p \in \Sigma$, let $h^a{}_b = \delta^a{}_b \mp n^a n_b$ so that $h^a{}_b n^b = 0$ (where upper/lower signs apply for spacelike/timelike respectively). Then*

(a) $h^a{}_c h^c{}_b = h^a{}_b$
(b) *Any vector x^a at p can be written as $x^a_\parallel + x^a_\perp$, where $x^a_\parallel = h^a{}_b x^b$ and $x^a_\perp = \pm n_b x^b n^a$.*
(c) *If X^a, Y^a are tangent to Σ, then $h^{ab} X_a Y_b = g^{ab} X_a Y_b$. h is sometimes called the first fundamental form.*

That is, let N^a be normal to Σ at a point p. If we parallel transport N^a on Σ along X^a (i.e., obeying the equation $X^b \nabla_b N^a = 0$), will the parallel-transported N^a still be normal? Take Y tangent to Σ, so that $Y^a N_a = 0$ at the point p. Then

$$\nabla_X (Y^a N_a) = X^b \nabla_b (Y^a N_a) = N_a X^b \nabla_b Y^a. \tag{7.10}$$

Definition 7.11. Up to now, n_a has been defined only on Σ. First, let us extend it to a neighborhood of Σ in an arbitrary way. The *extrinsic curvature* (second fundamental form) is defined at $p \in \Sigma$ by

$$K(X, Y) = -n_a (\nabla_{X_\parallel} Y_\parallel)^a. \tag{7.12}$$

Here, X and Y need not be tangent to Σ, but we are interested in their projections onto Σ. K therefore represents the extent to which normal vectors on Σ parallel-transported by vectors tangent to Σ fail to be normal after parallel transport. In addition, K does not look manifestly symmetric in X and Y written in this form, but in fact it is.

Lemma 7.13. *K_{ab} is independent of how n_a is extended off Σ, and in particular*

$$K_{ab} = h_a{}^c h_b{}^d \nabla_c n_d. \tag{7.14}$$

Proof. The RHS of $K(X,Y)$ is

$$-n_d X^c_\parallel \nabla_c Y^d_\parallel = X^c_\parallel Y^d \nabla_c n_d \tag{7.15}$$

by the Leibniz rule, since $n_d Y^d_\parallel = 0$. So

$$K(X,Y) = X^c_\parallel Y^d_\parallel \nabla_c n_d = X^a Y^b h_a{}^c h_b{}^d \nabla_c n_d. \tag{7.16}$$

Therefore

$$K_{ab} = h_a{}^c h_b{}^d \nabla_c n_d. \tag{7.17}$$

To demonstrate that K_{ab} is independent of how n_a is extended, consider a different extension n'_a, and let $m_a \equiv n'_a - n_a$. Note that on Σ, $m_a = 0$. Then on Σ,

$$\begin{aligned} X^a Y^b (K'_{ab} - K_{ab}) &= X^c_\parallel Y^d_\parallel \nabla_c m_d \\ &= \nabla_{x_\parallel}(Y^d_\parallel m_d) = 0 \end{aligned}$$

since $Y^d_\parallel m_d = 0$ on Σ. ☒

We can use

$$n^b \nabla_c n_b = \frac{1}{2}\nabla_c (n_b n^b) = 0 \tag{7.18}$$

and conclude that

$$K_{ab} = h_a{}^c \nabla_c n_b. \tag{7.19}$$

At home, we should show that the extrinsic curvature is given by a Lie derivative,

$$K_{ab} = \frac{1}{2}\mathcal{L}_n(h_{ab}) \tag{7.20}$$

where h here is the first fundamental form.

Non-lectured aside: extrinsic curvature is Lie derivative We want to show that

$$K_{ab} = \frac{1}{2}\mathcal{L}_n(h_{ab}). \tag{7.21}$$

Let us compute the left side of this expression first:

$$K_{ab} = h_a{}^c \nabla_c n_b = (\delta^c_a + n_a n^c)\delta_c n_d = n_a n_b + n_a n^c \nabla_c n_b. \tag{7.22}$$

Note that by a lemma we'll prove in the next lecture, K_{ab} is symmetric, so

$$K_{ab} = K_{(ab)} = \frac{1}{2}(\nabla_a n_b + \nabla_b n_a + n_a n^c \nabla_c n_b + n_b n^c \nabla_c n_a). \tag{7.23}$$

Now the right side of the equality is

$$(\mathcal{L}_n h)_{\mu\nu} = n^\rho \partial_\rho h_{\mu\nu} + h_{\mu\rho}\partial_\nu n^\rho + h_{\rho\nu}\partial_\mu n^\rho \tag{7.24}$$

and promoting to covariant derivatives, we have

$$\begin{aligned} (\mathcal{L}_n h)_{ab} &= n^c \nabla_c h_{ab} + h_{ac}\nabla_b n^c + h_{cb}\nabla_a n^c \\ &= n^c \nabla_c (g_{ab} + n_a n_b) + (g_{ac} + n_a n_c)\nabla_b n^c + (g_{cb} + n_c n_b)\nabla_a n^c \\ &= n_a n^c \nabla_c n_b + n_b n^c \nabla_c n_a + \nabla_b n_a + \nabla_a n_b, \end{aligned} \tag{7.25}$$

where we have used metric compatibility and the property that $n_c \nabla_b n^c = 0$. By comparison to Eqn. 7.23, we see that

$$K_{ab} = \frac{1}{2}(\mathcal{L}_n h)_{ab}.$$

☒

Lecture 8.

Thursday, January 31, 2019

"I know that you all really hate my handwriting. And our equations have hundreds of indices. So today I will receive threatening emails by the end of the day." –Jorge Santos

Last time, we introduced the extrinsic curvature,

$$K_{ab} = h_a{}^c h_b{}^d \nabla_c n_d, \tag{8.1}$$

where $n^a n_a = \pm 1$, with n^a the unit normal to the hypersurface Σ. We also defined the projection (now with indices down)

$$h_{ab} = g_{ab} \mp n_a n_b. \tag{8.2}$$

Lemma 8.3. *$K_{ab} = K_{ba}$, so K is a symmetric 2-tensor.*

Proof. Let $f : \mathcal{M} \to \mathbb{R}$ be constant on Σ with $df \neq 0$ on Σ. Let X^a be tangent to Σ. Thus

$$X(f) = X^a \nabla_a f = 0. \tag{8.4}$$

This implies that $(df)^a$ is normal to Σ. Thus on Σ we can write

$$n_a = \alpha (df)_a, \tag{8.5}$$

where α is chosen such that $n_a n^a = \pm 1$. It follows that

$$\nabla_c n_d = \alpha \nabla_c \nabla_d f + (\nabla_c \alpha)(df)_d \tag{8.6}$$

$$= \alpha \nabla_c \nabla_d f + \frac{(\nabla_c \alpha)}{\alpha} n_d. \tag{8.7}$$

But this tells us that

$$K_{ab} = h_a{}^c h_b{}^d \nabla_c n_d \tag{8.8}$$

$$= h_a{}^c h_b{}^d (\alpha \nabla_c \nabla_d f + \frac{\nabla_c \alpha}{\alpha} n_d) \tag{8.9}$$

$$= \alpha h_a{}^c h_b{}^d \nabla_c \nabla_d f \tag{8.10}$$

since $h_b{}^d n_d = 0$, and for all torsion-free spacetimes, covariant derivatives commute on scalars, so K_{ab} is symmetric. ☒

As it turns out, the property that

$$K_{ab} = \frac{1}{2} (\mathcal{L}_n h)_{ab} \tag{8.11}$$

means that the extrinsic curvature has the right form for an initial condition in the initial value problem of general relativity.m

The Gauss-Coducci equation A tensor at a point $p \in \Sigma$ is invariant under projection $h_a{}^b$ if

$$T^{a_1 \ldots a_n}{}_{b_1 \ldots b_s} = h_{c_1}{}^{a_1} \ldots h_{c_n}{}^{a_n} h_{b_1}{}^{d_1} \ldots h_{b_s}{}^{d_s} T^{c_1 \ldots c_n}{}_{d_1} \ldots d_s. \tag{8.12}$$

Proposition 8.13. *A covariant derivative D on Σ can be identified by projection of the covariant derivative on $\mathcal{M}$. Thus*

$$D_a T^{b_1 \ldots b_n}{}_{c_1 \ldots c_s} = h_a{}^d h_{e_1}{}^{b_1} \ldots h_{e_n}{}^{b_n} h_{c_1}{}^{f_1} \ldots h_{c_s}{}^{f_s} \nabla_d T^{e_1 \ldots e_n}{}_{f_1 \ldots f_s}. \tag{8.14}$$

Lemma 8.15. *The covariant derivative D is precisely the Levi-Civita (metric) connection associated to the metric h_{ab} on the submanifold Σ:*

$$D_a h_{bc} = 0, \tag{8.16}$$

and D is torsion-free.

Proof. The covariant derivative D acting on the metric h_{ab} is the projection

$$D_a h_{bc} = h_e{}^a h_f{}^b h_g{}^c \nabla_a h_{bc}. \tag{8.17}$$

Recalling that $h_{ab} = g_{ab} \mp n_a n_b$, let's expand the ordinary covariant derivative:

$$\nabla_a h_{bc} = \mp n_c \nabla_a n_b \mp n_b \nabla_a n_c, \tag{8.18}$$

where the $\nabla_a g_{bc}$ term is zero by metric compatibility. But recall that $h_a{}^c n_c = 0$ (i.e. the projection of a normal vector onto the submanifold is zero), so when we compute the projection of $\nabla_a h_{bc}$ onto Σ, we find that

$$D_a h_{bc} = h_e{}^a h_f{}^b h_g{}^c (\mp n_c \nabla_a n_b \mp n_b \nabla_a n_c) = 0, \tag{8.19}$$

which tells us that D is metric compatible with respect to h_{ab}.

To prove it is torsion-free, let $f : \Sigma \to \mathbb{R}$ be a scalar function on Σ, and extend it off Σ to a function $f : \mathcal{M} \to \mathbb{R}$ so that $\nabla_a f$ is well-defined. Then consider $D_a D_b f$. This is

$$\begin{aligned} D_a D_b f &= D_a h_b{}^e \nabla_e f \\ &= h_a{}^c h_b{}^d \nabla_c (h_d{}^e \nabla_e f) \\ &= h_a{}^c h_b{}^e \nabla_c \nabla_e f + h_a{}^c h_b{}^d (\nabla_c h_d{}^e)(\nabla_e f). \end{aligned}$$

This first term is already manifestly symmetric in a and b since the original connection was torsion free (i.e. covariant derivatives $\nabla_c \nabla_e$ commute on scalars), so let us rewrite the second term (neglecting the $\nabla_e f$ part) as follows:

$$\begin{aligned} h_a{}^c h_b{}^d \nabla_c h_d{}^e &= g^{ef} h_a{}^c h_b{}^d \nabla_c h_{df} \\ &= \mp g^{ef} h_a{}^c h_b{}^d (n_f \nabla_c n_d + n_d \nabla_c n_f) \\ &= \mp n^e K_{ab}, \end{aligned}$$

where we have grouped together $h_a{}^c h_b{}^d \nabla_c n_d = K_{ab}$ and gotten rid of the $n_d \nabla_c n_f$ term by projection with $h_b{}^d$. But K_{ab} is symmetric, so it also vanishes under antisymmetrization. We conclude that

$$D_{[a} D_{b]} f = 0. \tag{8.20}$$

⊠

We'd like to relate the extrinsic curvature to our old-fashioned Riemann curvature. How do we do this?

Proposition 8.21. *Denote the Riemann tensor associated with D_a on Σ by $R'^a{}_{bcd}$. This is given by Gauss's equation:*

$$R'^a{}_{bcd} = h^a{}_e h_b{}^f h_c{}^g h_d{}^h R^e{}_{fgh} \pm 2 K_{[c}{}^a K_{d]b}. \tag{8.22}$$

Proof. Let x^a be tangent to Σ. Then the Ricci identity for D is

$$R'^a{}_{bcd} X^b = 2 D_{[c} D_{d]} X^a. \tag{8.23}$$

Let us compute the RHS of this equation.

$$\begin{aligned} D_c D_d X^a &= h_c{}^e h_d{}^f h_g{}^a \nabla_e (D_f X^g) &(8.24)\\ &= h_c{}^e h^f{}_d h^a{}_g \nabla_e (h_f{}^h h_i{}^g \nabla_h X^i) &(8.25)\\ &= h_c{}^e h^f{}_d h^a{}_g \nabla_e \nabla_h X^i + h_c{}^e h^f{}_d h^a{}_g (\nabla_e h^h{}_f) \nabla_h X^i + h_c{}^e h^f{}_d h^a{}_g (\nabla_e h^g{}_i) \nabla_h X^i. &(8.26) \end{aligned}$$

This first term is already looking good. Note that we have to project the indices everywhere, so in the first line we project c, d, a indices, and in the second line we project f, g indices, and in the final line we expand out the covariant derivative.[15]

We have seen that

$$h^c{}_d h^d{}_b \nabla_c h^e{}_d = \mp n^e K_{ab}. \tag{8.27}$$

[15] As a rule of thumb, all free indices (i.e. that are not contracted over) must be projected in *each* covariant derivative. In my (limited) experience, I have found it most useful to begin with the innermost covariant derivative and work my way out. So for instance we know that $D_d X^a$ will come with two projectors, one with d as the free index and another with a as the free index. Our final expression (say, $\overline{D_d X^a}$) still has two free indices d, a but it also has two projectors, so that when we take the covariant derivative $D_c(\overline{D_d X^a})$, there are three free indices to project. Hence we get three projectors outside all the derivatives and two inside a first derivative, as in Eqn. 8.25.

We can use this identity in the last two terms of 8.26 to get

$$D_c D_d X^a = h^e{}_c h_d{}^h h^a{}_i \nabla_e \nabla_h X^i \mp K_{cd} h^a{}_i n^h \nabla_h X^i \mp K_c{}^a n_a h^h{}_d \nabla_h X^i. \tag{8.28}$$

This second term already drops out since K_{cd} is symmetric and we're antisymmetrizing over c and d in the final expression. The last term can be recast as

$$K_c{}^a h^h{}_d \nabla_h (n_i X^i) \pm K^a{}_c X^i h^h{}_d \nabla_h n_i = \pm K_c{}^a K_{bd} X^b. \tag{8.29}$$

Antisymmetrizing, we find that

$$\begin{aligned} R'^a{}_{bcd} X^b &= 2h_{[c}{}^e h_{d]}{}^f h^a{}_g \nabla_e \nabla_f X^g \pm 2K_{[c}{}^a K_{d]b} X^b \\ &= h_c{}^e h_d{}^f h_g{}^a h^a{}_b R^g{}_{hef} x^b \pm 2K_{[c}{}^a K_{a]b} X^b. \end{aligned}$$

Since X^b is arbitrary, this holds as a tensor identity. ⊠

Lemma 8.30. *The Ricci scalar of* Σ *is*

$$R' = R \mp 2R_{ab} n^a n^b \pm K^2 \mp K^{ab} K_{ab}, \tag{8.31}$$

where $K = g^{ab} K_{ab}$, *the trace of* K_{ab}.

Proposition 8.32 (Codacci's equation)**.** $D_a K_{bc} - D_b K_{ac} = h_a{}^d h_b{}^e h_c{}^f n^g R_{defg}$.

The proof of Codacci's equation is homework but simple compared to Gauss's equation.

Lemma 8.33. $D_a K^a{}_b - D_b K = h_b{}^c R_{cd} n^d$. *This is sometimes referred to as the Codacci equation, though we will call it the contracted Codacci equation (if we refer to it by name at all).*

The constraint equations We now have several definitions and identities. What are they good for? Assume the hypersurface Σ is spacelike, with a timelike normal n^a. The Einstein equation is just

$$R_{ab} - \frac{R}{2} g_{ab} \equiv G_{ab} = 8\pi T_{ab}.$$

We're going to contract it with $n^a n^b$. Using the same notation as before, the Einstein equation takes the form

$$R' - K^{ab} K_{ab} + K^2 = 16\pi\rho. \tag{8.34}$$

A priori, we might have thought that we had two free choices for our initial conditions on Σ– we could have chosen a metric on Σ and also an extrinsic curvature (like a first derivative of the metric moving off of Σ). But in fact the choice is not entirely free.

The equation 8.34 is known as the *Hamiltonian constraint*. Contracting with n^a and projecting with h gives us instead

$$D_b K^b{}_a - D_a K = 8\pi h_a{}^b T_{bc} n^c, \tag{8.35}$$

which is sometimes called the momentum constraint.[16]

Theorem 8.36 (Choquet-Bruhat and Geroch (1969)**.** *Let* (Σ, h, K) *be initial data satisfying the vacuum Hamiltonian and momentum constraints* ($T_{ab} = 0$)*. Then there exists a unique (up to diffeomorphism) spacetime* $(\mathcal{M}, g_{ab})$ *called the maximal Cauchy development of* (Σ, h, K) *such that*

(i) $(\mathcal{M}, g)$ *satisfies the Einstein equation,*
(ii) $(\mathcal{M}, g)$ *is globally hyperbolic with Cauchy surface* Σ,
(iii) *the induced metric and the extrinsic curvature of* Σ *are* h *and* K*, respectively,*
(iv) *and any other spacetime satsfying (i), (ii), and (iii) is isometric to a subset of* $(\mathcal{M}, g)$.

Analogous conditions exist for matter that obeys reasonable energy conditions. However, *it is possible that the maximal Cauchy development is extendible.* In that case, the region of the manifold in the complement of the maximal Cauchy development is not unique. Physics is not predictable outside the maximal Cauchy development, and the boundary of this region is called a *Cauchy horizon.* Note that in order to avoid trivial Cauchy horizons from poor choices of Σ, we require the initial data to be inextendible.

[16] It's diffeomorphism invariance that gives us this constraint, analogous to the constraint on trying to solve QED in Coulomb gauge.

A final note. Look at the Schwarzschild solution with negative M, such that the line element is

$$ds^2 = -\left(1+\frac{2|M|}{r}\right)dt^2 + \frac{dr^2}{1+\frac{2|M|}{r}} + r^2 d\Omega_2^2. \tag{8.37}$$

But if we take any $t = 0$ surface, it is impossible for us to avoid the $r = 0$ singularity. Our initial data should not be singular! Look at outgoing geodesics,

$$\frac{dt}{dr} = \frac{1}{1+\frac{2|M|}{r}} \implies t \simeq t_0 + \frac{r^2}{4|M|}. \tag{8.38}$$

This tells us that a Cauchy horizon has appeared, and this one we cannot avoid by picking a different Cauchy surface.

Non-lectured aside: Codacci's equation Let us show that Codacci's equation holds by explicit computation.

$$\begin{aligned} D_a K_{bc} &= h_a{}^d h_b{}^e h_c{}^f \nabla_d K_{ef} \\ &= h_a{}^d h_b{}^e h_c{}^f \nabla_d (h_e{}^g h_f{}^h \nabla_g n_h) \\ &= h_a{}^d h_b{}^e h_c{}^f \left[(\nabla_d h_e{}^g)\nabla_g n_f + h_e{}^g \nabla_d \nabla_g n_f\right] \\ &= \mp n^g h_c{}^f K_{ab} \nabla_g n_f + h_a{}^d h_b{}^e h_c{}^f \nabla_d \nabla_e n_f, \end{aligned}$$

where in going from the third to fourth lines, we have used our previous computation from the proof that D is torsion-free and rewritten the second term using the fact that $h_b{}^e h_e{}^g = h_b^e \delta_e^g$. We see that this first term is symmetric in ab since K_{ab} is symmetric, while exchanging $a \leftrightarrow b$ in the second term is equivalent to exchanging d and e. It follows that

$$D_a K_{bc} - D_b K_{ac} = h_a{}^d h_b{}^e h_c{}^f (\nabla_d \nabla_e n_f - \nabla_e \nabla_d n_f) \tag{8.39}$$

$$= h_a{}^d h_b{}^e h_c{}^f n^g R_{defg} \tag{8.40}$$

by the Ricci identity. ⊠

Lecture 9.

Friday, February 1, 2019

"Look at this word, the most naughty word in mathematics. 'Generically.'"–Jorge Santos

Last time, we formulated the initial value problem in general relativity. We said that some initial data is bad– we require that it is at a minimum inextendible and geodesically complete. However, something else bad can happen. See the figure: if our initial data is prescribed on an asymptotically null surface, then it could be our domain of dependence is cut off by a light cone. Therefore, we requre that the initial data is also *asymptotically flat*. We haven't yet defined what this means, though.

Definition 9.1. An initial data set (Σ, h, K) has an asymptotically flat end if

(i) Σ is diffeomorphic to $\mathbb{R}^3 \setminus B$, where B is a closed ball centered on the origin in $\mathbb{R}^3$,

(ii) If we pull back the $\mathbb{R}^3$ coordinates to define coordinates x^i on Σ, then $h_{ij} = \delta_{ij} + O(1/r)$ (the Euclidean metric plus $1/r$ terms) and $K_{ij} = O(1/r^2)$ as $r \to +\infty$, where $r = \sqrt{x^i x_i}$.

(iii) Derivatives of the latter expression also hold, e.g.

$$\partial_k h_{ij} = O(1/r^2).$$

Definition 9.2. An initial data set is asymptotically flat with N ends if it is the union of a compact set with N asymptotically flat ends (e.g. the Kruskal spacetime has two asymptotically flat ends).

Having defined our wish list for an initial data set, it would be extremely disturbing if we made this data set as nice as we like, and yet ended up with an extendible spacetime as our maximal Cauchy development. That such spacetimes do not (generically) occur is the content of the strong cosmic censorship conjecture.

Theorem 9.3. *Let (Σ, h, K) be a geodesically complete, asymptotically flat, initial data set for the vacuum Einstein equations. Then generically the maximal Cauchy development if the initial data is inextendible.*

It has been proven (with much work) that strong cosmic censorship is true in asymptotically flat space (i.e. Minkowski space is stable to small perturbations). It has nearly been proved (by Dafermos et al) for the Kerr black hole. And violations are known for Reissner-Nordström solution in de Sitter space.

Singularity theorems Now, in Newtonian gravity the formation of singularities is generally avoidable. If we drop a bit of matter in a Newtonian potential, it will fall straight to $r = 0$. But if we add a bit of angular momentum, the angular momentum will prevent the matter from falling straight in, and so our chunk of stuff cannot actually hit the $r = 0$ singularity. This is *not* true in general relativity.

It is the focus of an amazing set of theorems originally due to Penrose, and later generalized in conjunction with Hawking,[17] that the formation of singularities is generic in general relativity given a set of very reasonable conditions.[18]

Definition 9.4. A *null hypersurface* is a hypersurface whose normal is everywhere null. For example, take the (inverse) metric

$$g^{\mu\nu} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 1-\frac{2M}{r} & 0 & 0 \\ 0 & 0 & 1/r^2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{9.5}$$

The 1-form $n = dr$ is normal to the surfaces of constant r. Thus note that

$$n^2 = g^{\mu\nu} n_\mu n_\nu = g^{rr} = \left(1 - \frac{2M}{r}\right). \tag{9.6}$$

So the surface $r = 2M$ is a null hypersurface.

Now let n_a be normal to a null hypersurface $\mathcal{N}$. Then any nonzero vector X^a tangent to the hypersurface obeys

$$X^a n_a = 0. \tag{9.7}$$

But in particular we could take $X^a = n^a$, which implies that either X^a is spacelike or X^a is parallel to n^a. In particular, note that n^a (i.e. the vector constructed from the 1-form n_a by raising an index) is *tangent* to the hypersurface. Hence on $\mathcal{N}$, the integral curves of n^a lie within $\mathcal{N}$.

Proposition 9.8. *The integral curves of n^a are null geodesics. These are called the generators of $\mathcal{N}$.*

Proof. Let the null hypersurface $\mathcal{N}$ be given by an equation $f =$ constant for some function f with $df \neq 0$. df is everywhere normal to $\mathcal{N}$, so we must have $n = h\,df$. Define $N \equiv df$. Since $\mathcal{N}$ is null, we must have

$$N_a N^a = 0 \text{ on } \mathcal{N}. \tag{9.9}$$

Hence the function $(N^a N_a)$ is constant on $\mathcal{N}$. This implies that the gradient of this function is normal to $\mathcal{N}$. Thus

$$\nabla_a (N_b N^b)|_{\mathcal{N}} = 2\alpha N_a \tag{9.10}$$

for some proportionality constant α. Now, we also have that

$$\nabla_a N_b = \nabla_a \nabla_b f = \nabla_b \nabla_a f = \nabla_b N_a,$$

where we have used the assumption that our spacetime is torsion-free, so that covariant derivatives commute on scalars. But then by the Leibniz rule, we conclude that

$$N^b \nabla_b N_a|_{\mathcal{N}} = \alpha N_a, \tag{9.11}$$

which tells us that N^a is the tangent vector of a non-affinely parametrized geodesic. We conclude that the integral curves of N are null geodesics. ⊠

[17] It's kind of great to learn about black holes from a guy who's on a first-name basis with Stephen Hawking.

[18] I gave a talk on this last term!

Geodesic deviation We first saw geodesic deviation last term in *General Relativity*– for "nearby" geodesics, we can define relative "velocities" and "accelerations" leading to tidal forces. We'll make this notion more precise now.

Definition 9.12. A 1-parameter family of geodesics is a map $\gamma : I \times I' \to \mathcal{M}$, where I, I' are both open intervals in $\mathbb{R}$, such that

(i) for fixed s, $\gamma(s, \lambda)$ is a geodesic with affine parameter λ,
(ii) the map $(s, \lambda) \mapsto \gamma(s, \lambda)$ is smooth and one-to-one with smooth inverse. Thus γ defines a surface in $\mathcal{M}$.

Let U^a be the tangent vector to the geodesics and S^a be the vector tangent to curves of constant λ (i.e. takes us between neighboring geodesics). In a chart x^μ, the geodesics are specified in coordinates $x^\mu(s, \lambda)$ where $S^\mu = \frac{\partial x^\mu}{\partial s}$. Hence

$$x^\mu(s + \delta s, \lambda) = x^\mu(s, \lambda) + \delta s S^\mu + O(\delta s^2),$$

where S^μ points towards the next geodesic at $s + \delta s$. For this reason, we call S^μ the *deviation vector*.

On the surface (i.e. the image of γ), we can use s and λ coordinates. This gives a coordinate chart in which

$$S^\mu = \left(\frac{\partial}{\partial s}\right)^\mu, \quad U^\mu = \left(\frac{\partial}{\partial \lambda}\right)^\mu \tag{9.13}$$

on the surface. But in this coordinate system, these are just partial derivatives, so they automatically commute, and the commutator is covariant, so S^a and U^a commute in general, independent of basis. Thus

$$0 = [S, U] \iff U^b \nabla_b S^a = S^b \nabla_b U^a. \tag{9.14}$$

This implies that

$$U^c \nabla_c (U^b \nabla_b S^a) = R^a{}_{bcd} U^b U^c S^d. \tag{9.15}$$

Solutions of 9.15 are called *Jacobi fields*.

Geodesic congruence We'll introduce a final definition for today.

Definition 9.16. Let $U \subset \mathcal{M}$ be open. A *geodesic congruence* on U is a family of geodesics such that exactly one geodesic passes through each $p \in U$.

A geodesic congruence provides us with a notion of a set of geodesics covering a subset of the manifold without overlap.

Lecture 10.

Monday, February 18, 2019

"We'll see that the shear in AdS tells us amazing things. Well, we won't see the AdS. But we'll see the amazing thing." –Jorge Santos

It's been 15 days since the last lecture. Where did we leave off? We introduced null surfaces, and suggested that null surfaces will be important to our understanding of singularity theorems.

Last time, we talked about geodesic congruences (families of geodesics which cover some submanifold of the spacetime). Consider a congruence where all geodesics are of the same type,

$$U^a U_a = \pm 1$$

in affine parametrization (for spacelike or timelike geodesics respsectively) or $U^a U_a = 0$ for null congruences. Now consider a 1-parameter family of geodesics which form such a congruence. In terms of the geodesic tangent vector U and the deviation vector S, we have

$$[S, U] = 0 \iff U^b \nabla_b S^a = S^b \nabla_b U^a = S^b B^a{}_b \tag{10.1}$$

where we have defined

$$B^a{}_b \equiv \nabla_b U^a. \tag{10.2}$$

That is, B measures the failure of S^a to be parallel propagated along U^b. This tensor has some nice properties:

$$B^a{}_b U^b = 0, \quad U_a B^a{}_b = \frac{1}{2}\nabla_b(U^2) = 0. \tag{10.3}$$

The first follows from the geodesic equation and the latter from $U^2 =$ constant.[19] Consider the following expression:

$$U \cdot \nabla(U \cdot S) = (U \cdot \nabla U^a)S_a + U^a(U \cdot \nabla S_a), \tag{10.4}$$

where $(\cdot)$ indicates contracted indices. But this first term is zero by the geodesic equation, and the second is zero by the property $U^a B_{ab} = 0$ from above. Thus $U \cdot S$ is constant along the integral curves of U.

The existence of some quantity along geodesics suggests to us there might be some gauge freedom we get to fix. In particular, we can make a very nice choice to fix the affine parameter along geodesics:

$$\lambda' = \lambda - a(s) \tag{10.5}$$

where a is some function of our choice. In particular this allows us to shift S:[20]

$$S'^a \equiv S^a + \frac{da}{ds}U^a \implies U \cdot S' = U \cdot S + \frac{da}{ds}U^2. \tag{10.6}$$

What's nice about this? $U^2 = \pm 1$ for timelike or spacelike coordinates, which means that WLOG we can set $U \cdot S' = 0$. But for the null case, $U^2 = 0$ means that the reparametrization cannot be fixed by a choice of a– we just get $U \cdot S' = U \cdot S$.

Gauge fixing, the easy way For null geodesic congruences, we have $U^2 = 0$ and thus

$$U \cdot S' = U \cdot S,$$

so a doesn't help us pick a nice reparametrization. Instead, we pick a spacelike hypersurface Σ which intersects each geodesic once– we can do this since we're looking at a congruence. Now let N^a be a vector field defined on Σ obeying $N^2 = 0, N \cdot U = -1$ on Σ.

Now extend N^a off Σ by parallel transport along the geodesic

$$U \cdot \nabla N^a = 0. \tag{10.7}$$

There's a nice discussion of equivalence classes and the full freedom we have to fix the gauge, but we'll leave it to Wald. For our purposes, notice that we have the three properties

$$N^2 = 0, \quad U \cdot N = -1, \quad U \cdot \nabla N^a = 0. \tag{10.8}$$

Take a deviation vector S^a and decompose it as follows:

$$S^a = \alpha U^a + \beta N^a + \hat{S}^a, \tag{10.9}$$

where

$$U \cdot \hat{S} = N \cdot \hat{S} = 0. \tag{10.10}$$

This is a bit like a Gram-Schmidt process– we can subtract off the bits parallel to the geodesic U^a and also the bit parallel to N^a. Note that $U \cdot S = -\beta$, where β is constant, as we found at the start of the calculation. It's also important to observe that any vector which is orthogonal to two null vectors is either spacelike or the zero vector.[21]

[19] Just one line of algebra. $U_a B^a{}_b = U_a \nabla_b U^a = \frac{1}{2}\nabla_b(U^2) = 0$.

[20] Recall that $S \equiv \frac{dx^\mu}{ds}$, and $U^\mu \equiv \frac{dx^\mu}{d\lambda}$. Thus for a geodesic parametrized by λ', s, we may write

$$\begin{aligned} x^\mu(\lambda', s) &= x^\mu(\lambda - a(s), s) \\ &= x^\mu(\lambda, s) + \frac{dx^\mu}{d\lambda}a(s) \\ &= x^\mu(\lambda, s) + U^\mu a(s), \end{aligned}$$

and taking a derivative with respect to s now yields

$$\frac{dx^\mu(\lambda', s)}{ds} = \frac{dx^\mu(\lambda, s)}{ds} + \frac{da}{ds}U^\mu.$$

[21] To see this, try the calculation in Minkowski space. I haven't worked out the details myself yet.

So we can write a deviation vector S^a as the sum of a part

$$\alpha U^a + \hat{S}^a$$

which is orthogonal to U^a and a part

$$\beta N^a$$

that is parallel transported along each geodesic. We are interested in a congruence containing the generators of a null hypersurface $\mathcal{N}$. In this case, if we pick a 1-parameter family of geodesics contained in $\mathcal{N}$, then the deviation vector S^a will be tangent to $\mathcal{N}$ and hence $U \cdot S = 0$. Since U^a is normal to $\mathcal{N}$, we have $\beta = 0$.

We can write

$$\hat{S}^a = P^a{}_b S^b \tag{10.11}$$

under the projection

$$P^a{}_b = \delta^a{}_b + N^a U_b + U^a N_b, \tag{10.12}$$

where it's a quick exercise to check that $P^a{}_b P^b{}_e = P^a{}_e$. This projection projects the tangent space at p onto the spacelike $T_\perp$, the space perpendicular to both null vectors N^a and U^a. One can also check that

$$U \cdot \nabla P^a{}_b = 0, \tag{10.13}$$

so P is parallel-propagated trivially along geodesics U^a.

Proposition 10.14. *A deviation vector for which $U \cdot S = 0$ satisfies*

$$U \cdot \nabla \hat{S}^a = \hat{B}^a{}_b \hat{S}^b, \tag{10.15}$$

where $\hat{B}^a{}_b$ is a projected version of $B^a{}_b$ into the perpendicular space,

$$\hat{B}^a{}_b = P^a{}_c B^c{}_d P_b{}^d. \tag{10.16}$$

Proof.

$$\begin{aligned} U \cdot \nabla \hat{S}^a &= U \cdot \nabla (P^a{}_c S^c) \\ &= P^a c U \cdot \nabla S^c \\ &= P^a{}_c B^c{}_d S^d \\ &= P^a{}_c B^c_d P^d{}_e S^e, \end{aligned}$$

where we used $U \cdot S = 0$ and $B^c{}_d U^d = 0$. Finally, we can use the fact that P is a projector so that $P^2 = P$. So we replace $P^d{}_e = P^d{}_f P^f{}_e$ and this gives us the projectors to write everything with hats as

$$U \cdot \nabla \hat{S}^a = P^a{}_c B^c{}_d P^d{}_f P^f{}_e S^e = \hat{B}^a{}_f \hat{S}^f.$$ ⊠

Expansion, rotation, and shear We've proved some nice properties about this $B^a{}_b$ matrix. But what exactly is it? As it turns out, $\hat{B}^a{}_b$ can be regarded as a matrix that acts on the 2D space $T_\perp$. It's very natural for us to decompose such a matrix into its symmetric (trace-free) part, its antisymmetric part, and its trace part.

Definition 10.17. Let us define

$$\theta \equiv \hat{B}^a{}_a, \quad \hat{\sigma}_{ab} \equiv \hat{B}_{(ab)} - \frac{1}{2} P_{ab} \theta, \quad \hat{\omega}_{ab} = \hat{B}_{[ab]}. \tag{10.18}$$

Note that the factor of $1/2$ works for $3+1$ spacetime dimensions because the trace of P_{ab} is 2 here. This then implies that

$$\hat{B}^a{}_b = \frac{1}{2}\theta P^a{}_b + \hat{\sigma}^a{}_b + \hat{\omega}^a{}_b. \tag{10.19}$$

With this definition, notice that

$$\theta \equiv g^{ab} \hat{B}_{ab} = g^{ab} B_{ab} = \nabla_a U^a, \tag{10.20}$$

which does not depend on the choice of N. In fact, this is true for any scalar like the eigenvalues of the rotation matrix.

Proposition 10.21. *If the congruence contains the generators of a null hypersurface $\mathcal{N}$, then $\hat{\omega}_{ab} = 0$ on $\mathcal{N}$. Conversely, if $\hat{\omega}_{ab} = 0$ everywhere on $\mathcal{N}$, then U^a is everywhere hypersurface orthogonal.*[22]

[22] This is not quite true in some extensions of general relativity, but it is exact in pure GR.

Proof. The definition of $\hat{B}$ and the fact that $U \cdot U = B \cdot U = 0$ implies htat

$$\hat{B}^b{}_c = B^b{}_c + U^b N_d B^d{}_c + U_c B^b{}_d N^d + U^b U_c N_d B^d{}_e N^e. \tag{10.22}$$

Using this we have

$$U_{[a}\hat{\omega}_{bc]} = U_{[a}\nabla_c U_{b]} = -\frac{1}{6}(U \wedge dU)_{abc}, \tag{10.23}$$

where we have just rewritten the expression in form notation. If U^a is normal to $\mathcal{N}$, then $U \wedge dU = 0$ on $\mathcal{N}$, and hence on $\mathcal{N}$,

$$0 = U_{[a}\hat{\omega}_{bc]} = \frac{1}{3}(U_a\hat{\omega}_{bc} + U_b\hat{\omega}_{ca} + U_c\hat{\omega}_{ab}), \tag{10.24}$$

where we've just expanded out the antisymmetrization and used the fact that $\hat{\omega}_{bc}$ is already antisymmetric to absorb factors of 2.

Contracting with N^a gives

$$\hat{\omega}_{bc} = 0 \tag{10.25}$$

on $\mathcal{N}$, where $U \cdot N = -1$ and $\hat{\omega} \cdot N = 0$. Conversely if $\hat{\omega} = 0$ everywhere, then 10.23 implies that U is hypersurface orthogonal by the Frobenius theorem. ⊠

Lecture 11.

Wednesday, February 20, 2019

"Roger (Penrose) noticed something very interesting about the signs of these guys. ...well, I have no idea if that was actually the case, I'm just making this up as I go along." –Jorge Santos

We showed last time that we can characterize the behavior of null congruences by three objects– the expansion, the shear, and the rotation. Having showed that the rotation was zero under certain conditions, we now consider the consequences of these objects.

Gaussian null coordinates For a null hypersurface $\mathcal{N}$, take a 2D spacelike surface S with coordinates y_i on $\mathcal{N}(U)$. Take another null vector field V on $\mathcal{N}$ satisfying

$$V \cdot \frac{\partial}{\partial y^i} = 0 \text{ and } V \cdot U = 1, V^2 = 0. \tag{11.1}$$

Now we assign coordinates (r, λ, y^i) to the point at affine parameter distance r along the null geodesic which starts at the point on $\mathcal{N}$ with coordinate (λ, y^i) and has tangent vector V^a. Thus we follow U along $\mathcal{N}$ $(r = 0)$ for distance λ and then go off $\mathcal{N}$ by following V for affine parameter r. See Fig. 6 for an illustration.

This defines a coordinate chart in a neighborhood of $\mathcal{N}$ such that the hypersurface $\mathcal{N}$ is at $r = 0$, with $U = \frac{\partial}{\partial \lambda}$ on $\mathcal{N}$ and $\frac{\partial}{\partial r}$ tangent to affinely parametrized null geodesics.

The latter condition implies that $g_{rr} = 0$ everywhere (that is, $g_{\mu\nu}(\frac{\partial}{\partial r})^\mu(\frac{\partial}{\partial r})^\nu = g_{rr} = 0$). It also follows from the geodesic equation for $\frac{\partial}{\partial r}$ that

$$g_{r\mu,r} = 0, \tag{11.2}$$

where μ is now a free index that runs over r, λ, y^i.[23] But at $r = 0$, we have $g_{r\lambda} = U \cdot V = 1$, and $g_{ri} = V \cdot \frac{\partial}{\partial y^i} = 0$. So these are the first indications these are good coordinates. We also know that $g_{\lambda\lambda} = 0$ at $r = 0$, as U^a is null, and $g_{\lambda i} = 0$ at $r = 0$, as $\frac{\partial}{\partial y^i}$ is tangent to $\mathcal{N}$ and hence orthogonal to U^a. Thus the metric components take the form

$$g_{\lambda\lambda} = rF, \quad g_{\lambda i} = rh_i, \tag{11.3}$$

[23] The proof is left as an exercise in Harvey Reall's notes. It's two lines: let $U^\mu = (1, 0, 0, 0)$ in (r, λ, y^1, y^2) coordinates. Then the geodesic equation says that

$$\begin{aligned} 0 &= \frac{dU^\mu}{d\tau} + \Gamma^\mu_{\alpha\beta}U^\alpha U^\beta \\ &= \Gamma^\mu_{rr} = g^{\mu\sigma}(g_{r\sigma,r} + g_{\sigma r,r} - g_{rr,\sigma}) = 2g^{\mu\sigma}g_{r\sigma,r} \end{aligned}$$

since $g_{rr} = 0$ and g is symmetric. We conclude that $g_{r\sigma,r} = 0$ for any index σ.

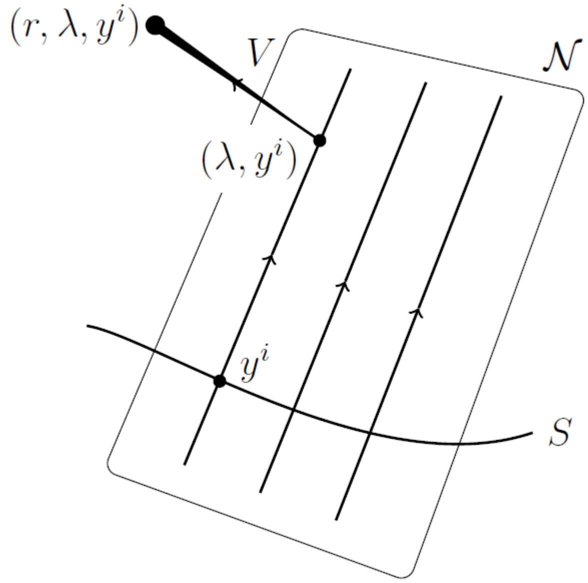

FIGURE 6. An illustration of Gaussian null coordinates. For a null hypersurface $\mathcal{N}$, take a 2D spacelike surface $S \subset \mathcal{N}$ with coordinates y_i. We construct the integral curves of U (the generators of $\mathcal{N}$) and from S, follow these curves for affine parameter distance λ. Finally, we take a null vector field V and follow its integral curves off of $\mathcal{N}$ for a distance r. Image credit to Prof. Reall's Black Holes notes, §4.6.

and we find the metric in these coordinates takes the form

$$ds^2 = 2drd\lambda + rFd\lambda^2 + 2rh_i d\lambda dy^i + h_{ij}dy^i dy^j. \tag{11.4}$$

Note that at $r = 0$, F must vanish. To see this, we use the fact that

$$\lambda \mapsto (0, \lambda, y^i) \tag{11.5}$$

are affinely parametrized null geodesics. For this, the only non-vanishing component of the geodesic equation is the r component,

$$\partial_r(rF) = 0. \tag{11.6}$$

Integrating this at $r = 0$ tells us that $F = 0$ at $r = 0$, so we can write $F = r\hat{F}$ for $\hat{F}$ some smooth function. WLOG we can therefore write

$$ds^2 = 2drd\lambda + r^2\hat{F}d\lambda^2 + 2rh_i d\lambda dy^i + h_{ij}dy^i dy^j. \tag{11.7}$$

What does this metric look like on our null hypersurface? On $\mathcal{N}$, the metric is simply

$$g|_{\mathcal{N}} = 2drd\lambda + h_{ij}dy^i dy^j. \tag{11.8}$$

Thus $U^\mu = (0, 1, 0, 0)$ on $\mathcal{N}$, which implies $U_\mu = (1, 0, 0, 0)$ (where we've lowered the index using $g|_{\mathcal{N}}$). Recall that

$$U \cdot B = B \cdot U = 0 \implies B^r{}_\mu = B^\mu{}_\lambda = 0. \tag{11.9}$$

We also know that $\theta = B^\mu{}_\mu$, but we've just shown that two of the trace components vanish, $B^r{}_r = B^\lambda{}_\lambda = 0$, so it must be the y^i components which contribute to the expansion. On $\mathcal{N}$,

$$\theta = B^i{}_i = \nabla_i U^i = \partial_i U^i + \Gamma^i_{i\mu} U^\mu \tag{11.10}$$

$$= \Gamma^i_{i\lambda} = \frac{1}{2} h^{ij} (g_{ji,\lambda} + g_{j\lambda,i} - g_{i\lambda,j}) \tag{11.11}$$

$$= \frac{1}{2} h^{ij} h_{ji,\lambda} \tag{11.12}$$

$$= \frac{\partial_\lambda \sqrt{h}}{\sqrt{h}} \tag{11.13}$$

where $h = \det h_{ij}$. To make the simplifications in the first and second lines, we notice that $\partial_i U^i = 0$ and $g_{i\lambda} = 0$ on $\mathcal{N}$ (remark: λ is a component, not a free index!). Hence we have

$$\frac{d\sqrt{h}}{d\lambda} = \theta \sqrt{h}. \tag{11.14}$$

Borrowing a notation from Witten, if we denote $V \equiv \sqrt{h}$, then $\theta = \dot{V}/V$. Since V represents the volume of a little area on S, θ represents the normalized expansion of that volume as we flow along the null geodesics in the congruence.

Trapped surfaces Consider a 2D spacelike surface S, i.e. a 2D submanifold for which all tangent vectors are spacelike. For any $p \in S$, there will be *precisely* two future-directed null vectors $U^a{}_1, U^a{}_2$ orthogonal to S (up to a freedom of rescaling).[24] This generalizes the notion of ingoing and outgoing trajectories.

Example 11.15. Let S be a 2-sphere with $U = U_0, V = V_0$ in the Kruskal spacetime. (That is, points in the Kruskal diagram represent 2-spheres.) By spherical symmetry, the generators of $\mathcal{N}_i$ will be radial null geodesics. Hence, $\mathcal{N}_i$ must be surfaces of constant U or constant V, with the generators tangent to dU and dV. Raising the index, we have

$$U^a{}_1 = re^{r/2M} \left(\frac{\partial}{\partial V} \right)^a, \quad U^a{}_2 = re^{r/2M} \left(\frac{\partial}{\partial U} \right)^a. \tag{11.16}$$

We have fixed signs such that both are future-directed. Now

$$\theta_1 = \nabla_a U^a{}_1 = \frac{1}{\sqrt{g}} \partial_a (\sqrt{-g} U^a{}_1) = -\frac{8M^2}{r} U, \tag{11.17}$$

$$\theta_2 = -\frac{8M^2}{r} V. \tag{11.18}$$

But note that U and V have different signs in the different quadrants of the Kruskal diagram. Something very interesting happens because of this.

For S in region I, we have

$$U < 0, V > 0 \implies \theta_1 > 0, \theta_2 < 0. \tag{11.19}$$

Region IV is similar. But region II (the black hole interior) is different. Here,

$$U > 0, V > 0 \implies \theta_1 < 0, \theta_2 < 0, \tag{11.20}$$

so now both the expansion coefficients are negative, and everything is shrinking as we flow along congruences.

Definition 11.21. A compact orientable 2D spacelike surface is *trapped* if both families of null geodesics orthogonal to S have negative expansion everywhere on S.

As it turns out, the existence of trapped surfaces is closely related with the presence of singularities– they are a condition in the famous Penrose singularity theorems.

---

[24]"If you want to see if you really understand the class, go home and think about why the word precisely is here."

The Raychaudhuri equation

Proposition 11.22. *With the expansion defined as before,*

$$\frac{d\theta}{d\lambda} = -\frac{1}{2}\theta^2 + \hat{\sigma}^{ab}\hat{\sigma}_{ab} + \hat{\omega}^{ab}\hat{\omega}_{ab} - R_{ab}U^aU^b. \tag{11.23}$$

Proof. We start out with

$$\begin{aligned}
\frac{d\theta}{d\lambda} &= U \cdot \nabla (B^a{}_b P^b{}_a) \\
&= P_a{}^b U \cdot \nabla^a{}_b \\
&= P_a{}^b U^c (\nabla_c \nabla_b U^a + R^a{}_{dcb} U^d) \\
&= P_a{}^b [\nabla_b (U^c \nabla_c U^a) - (\nabla_b U^c)(\nabla_c U^a)] + P_a{}^b R^a{}_{dcb} U^c U^d \\
&= -B^c{}_b P^b{}_a B^a{}_c - R_{cd} U^c U^d \\
&= -\hat{B}^c{}_a \hat{B}^a{}_c - R_{ab} U^a U^b.
\end{aligned}$$

where we've used the Ricci identity to commute the derivatives and moved a U inside a derivative with Leibniz in order to get something that vanishes by the geodesic equation.

Instead of just B^2, we've got $\hat{B}^2$, which is even nicer, and from here we simply insert the explicit expression for $\hat{B}$ in terms of $\theta, \hat{\sigma}$, and $\hat{\omega}$. ⊠

But notice that none of this proof depends on the Einstein equations! It can be formulated as a purely geometric statement. Where Einstein comes in is in recognizing that the Ricci tensor has appeared in the Raychaudhuri equation, and in particular the Einstein equations connect the Ricci tensor to the stress tensor. $R_{ab}U^aU^b$ is some contracted quantity, and what this starts to suggest to us is conditions on the matter that lives in spacetime, i.e. energy conditions.

Lecture 12.

Thursday, February 21, 2019

"You might not be familiar [with these topological concepts]. But you need to get familiar because we're going to squeeze physics out of this like there's no tomorrow." –Jorge Santos

We continue our discussion of focusing of geodesics. Last time, we wrote down Raychaudhuri's equation,

$$\frac{d\theta}{d\lambda} = -\frac{1}{2}\theta^2 - \hat{\sigma}^{ab}\hat{\sigma}_{ab} + \hat{\omega}^{ab}\hat{\omega}_{ab} - R_{ab}U^aU^b,$$

where $U^aU_a = 0$ (i.e. we consider null geodesics). Today, we will set up the proof that singularities are generic in general relativity.

The first three terms on the RHS of the Raychaudhuri equation have definite signs (e.g. we showed that the third is zero along the integral curves of the generators of a null hypersurface). These were purely geometric statements. But to say anything useful about this last term with the Ricci tensor in it, we'll need some physics. Let u^a be the 4-velocity of an observer. We now define a current

$$J^a = -T^a{}_b u^b. \tag{12.1}$$

Definition 12.2. The *Dominant Energy Condition (DEC)* states that $-T^a{}_bV^b$ is a future-directed causal vector (or zero) for all future-directed timelike vectors V^a.

Definition 12.3. The *Weak Energy Condition (WEC)* states that $T_{ab}V^aV^b \geq 0$ for any causal vector V^a.

Definition 12.4. The *Null Energy Condition (NEC)* states that $T_{ab}V^aV^b \geq 0$ for any null vector V^a.

There is also a strong energy condition, but note that it *does not* imply the weak energy condition.

Definition 12.5. The *Strong Energy Condition (SEC)* states that

$$(T_{ab} - \frac{1}{2}g_{ab}T^c{}_c)V^aV^b \geq 0 \tag{12.6}$$

for all causal vectors V^a.

Thus the DEC $\implies$ WEC $\implies$ NEC and SEC $\implies$ NEC, but there are no other implications. Fortunately, we will only need the weakest of these– the null energy condition.

Conjugate points

Lemma 12.7. *In a spacetime satisfying Einstein's equation with matter satisfying the null energy condition,*

$$\frac{d\theta}{d\lambda} \leq -\frac{\theta^2}{2}. \tag{12.8}$$

Consider the Raychaudhuri equation. The NEC fixes the sign of the the R_{ab} term so that $-R_{ab}U^aU^b < 0$, and the $\hat{\sigma}$ comes from projecting onto a null hypersurface, so $\hat{\sigma}$ has only spacelike indices (hence $\hat{\sigma}^{ab}\hat{\sigma}_{ab}$) is manifestly non-negative). The $\hat{\omega}$ contribution is zero, as argued previously. Thus all the terms on the RHS apart from $-\frac{1}{2}\theta^2$ are non-positive, so the lemma follows. ☒

Corollary 12.9. *If $\theta = \theta_0 < 0$ at a point p on a generator γ of a null hypersurface, then $\theta \to -\infty$ along γ within finite parameter distance $2/|\theta_0|$ provided that γ extends this far.*

Proof. Let $\lambda = 0$ at p, and notice that our lemma can be written as

$$\frac{d(\theta^{-1})}{d\lambda} \geq \frac{1}{2} \implies \theta^{-1} - \theta_0^{-1} \geq \frac{\lambda}{2} \implies \theta \leq \frac{\theta_0}{1+\lambda\frac{\theta_0}{2}}. \tag{12.10}$$

But now notice that this denominator becomes singular as $\lambda \to 2/|\theta_0|$, and since the numerator is positive, we see that $\theta \to -\infty$ as $\lambda \to 2/|\theta_0|$. ☒

Definition 12.11. Points p, q on a geodesic γ are *conjugate* if there exists a Jacobi field along γ that vanishes at p and q but is not identically zero.

These give us a sense of focal points, so they are natural to think about in the context of singularity theorems.

Theorem 12.12. *Consider a null geodesic congruence which includes all of the null geodesics through p.*[25] *If $\theta \to -\infty$ at a point q on a null geodesic γ through p, then q is conjugate to p along γ.*

Theorem 12.13. *Let γ be a causal curve with $p, q \in \gamma$. Then there* does not *exist a smooth 1-parameter family of causal curves γ_s connecting p, q with $\gamma_0 = \gamma$ and γ_s timelike for $s > 0$ if and only if γ is a null geodesic wth no point conjugate to p along γ between p and q.*

Example 12.14. Take the Einstein static universe with metric

$$ds^2 = -dt^2 + d\Omega_2^2.$$

Let us observe that null geodesics emitted from the south pole at time $t = 0$ all reconverge at the north pole at time $t = \pi$. However, if we now look at a point q where we would have to travel farther than R along a great circle, then by deforming the great circle into a shorter path, one can travel from p to q with a velocity less than the speed of light. Thus there exists a timelike curve from p to q.[26]

Consider now a 2D spacelike surface S. We have seen that (in $3+1$ dimensions) we can introduce two future-directed null vector fields (ingoing/outgoing trajectories, if you like) U_1^a and U_2^a on S that are normal to S. Consider the null geodesics which have one of these vectors as their tangent on S. A point is *conjugate* to a surface if there is a Jacobi field which is tangent on the surface and vanishes at that point. The analogue of Thm. 12.12 is then as follows.

Theorem 12.15. *A point P is conjugate to a surface S if $\theta \to -\infty$ at p along one of the geodesics parametrized by U_1^a, U_2^a.*

[25] This is not quite a congruence as we've defined it before. It is necessarily singular at p since p lies on more than one geodesic– by definition, it lies on all of them! However, we won't worry too much about the details. Just take the set of all null geodesics through p and work from there.

[26] If we like, we can say that timelike geodesics therefore only have the property of extremizing proper time until they hit their first focal (conjugate) point. As Witten writes, "A geodesic that is continued past a focal point and thus has gone more than half way around the sphere is no longer length minimizing, as it can be 'slipped off' the sphere."

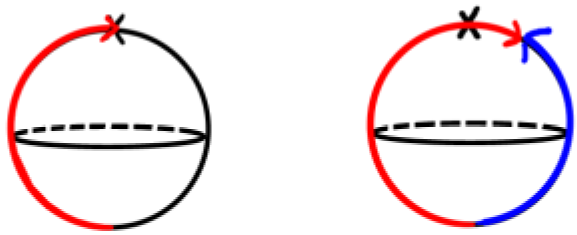

FIGURE 7. Left: a 2-sphere with a geodesic that travels from the south pole to the north pole (marked by an X). Right: a 2-sphere with a geodesic (red) that continues through the north pole. As this geodesic has passed through a focal point, it no longer minimizes the distance between its endpoints, and can be deformed into the geodesic in blue which is of shorter length.

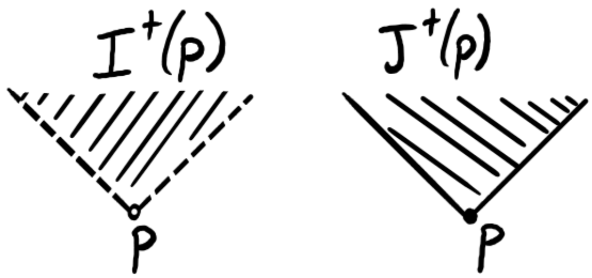

FIGURE 8. The chronological future I^+ and the causal future J^+ in Minkowski space. Notice that the chronological future $I^+(p)$ is strictly the interior of the light cone of the point p and does not include p itself, whereas the causal future includes the light cone and p itself.

Definition 12.16. Let $(\mathcal{M}, g)$ be a time-orientable spacetime $U \subset \mathcal{M}$. The *chronological future/past* of U, denoted $I^{\pm}(U)$, is the set of points of $\mathcal{M}$ which can be reached by a future-/past-directed timelike curve starting on U (not including U).

Definition 12.17. The *causal future/past* of U, denoted $J^{\pm}(U)$, is the union of U with the set of points of $\mathcal{M}$ that can be reached by a future-/past-directed *causal* curve starting on U.

Theorem 12.18. *Let S be a two-dimensional orientable spacelike submanifold of a globally hyperbolic spacetime $(\mathcal{M}, g)$. Then every point $p \in \dot{J}^+(S)$ (the topological boundary of $J^+(S)$) lies on a future-directed null geodesic starting from S which is orthogonal to S and has no point conjugate to S between S and p. Furthermore, $\dot{J}^+(U)$ is a submanifold of $\mathcal{M}$ and is achronal (no two points on $\dot{J}^+(U)$ are connected by a timelike curve), where $U \subset \mathcal{M}$.*

Lecture 13.

Friday, February 22, 2019

"Is this clear? Now go home and prove it." –Jorge Santos, on the Penrose singularity theorem

We stated an important theorem last time, and we'll use it to prove something great.

Theorem 13.1 (Penrose singularity theorem (1965))**.** *Let $(\mathcal{M}, g)$ be globally hyperbolic with a non-compact Cauchy surface Σ. Assume the Einstein equations and the null energy condition hold, and that $\mathcal{M}$ contains a trapped surface T. Let $\theta_0 < 0$ be the maximum value of θ on T for both sets of null geodesics orthogonal to T. Then at least one of these geodesics is future-directed inextendible and has a finite length no greater than $2/|\theta_0|$.*

Proof. We will prove this by contradiction. Assume that all inextendible null geodesics orthogonal to T have affine length greater than $2/|\theta_0|$. By the Raychaudhuri equation, we saw that along any of these

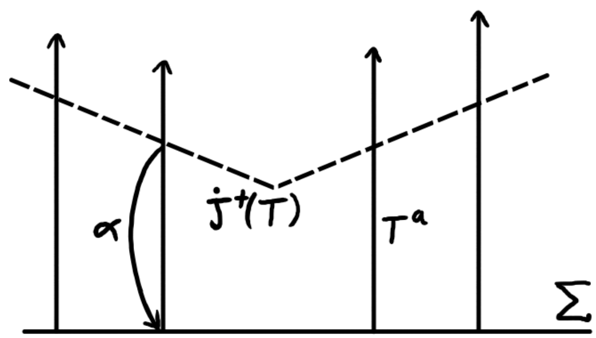

FIGURE 9. An illustration of the homeomorphism between $\dot{J}^+(T)$ and Σ. By following the timelike curves which are the integral curves of T^a (the time-orientation of our spacetime) backwards from $\dot{J}^+(T)$ to some subset of Σ, we get a continuous one-to-one map $\alpha : \dot{J}^+(T) \to \Sigma$.

geodesics, we will have $\theta \to -\infty$. Hence a point conjugate to T within affine parameter distance no greater than $2/|\theta_0|$.

It's time to use our theorem from last time. Let $p \in \dot{J}(T)$, $p \notin T$. From Thm. 12.18, we know that p lies on a future-directed null geodesic γ starting from T which is orthogonal to T and has no point conjugate to T between T and p. It follows that p cannot lie beyond the point on γ conjugate to T on γ.

Therefore $\dot{J}^+(T)$, the topological boundary of the causal future of T, is a subset of a compact set consisting of the set of points along the null geodesics orthogonal to T with affine parameter less than or equal to $2/|\theta_0|$. But $\dot{J}^+(T)$ is a boundary, hence closed, which implies that $\dot{J}^+(T)$ is in fact compact.[27]

Now recall that $J^+(T)$ is a (sub)manifold, which implies that it cannot have a boundary.[28] In addition, we said that Σ is a non-compact Cauchy surface. Therefore pick a timelike vector field T^a (we are free to do this since our manifolds are time-orientable). By global hyperbolicity, integral curves of T^a intersect Σ exactly once. Notice that they will intersect $\dot{J}^+(T)$ at most once because $\dot{J}^+(T)$ is achronal (otherwise, two points would be connected by a timelike curve). This defines a continuous one-to-one map $\alpha : \dot{J}^+(T) \to \Sigma$, as illustrated in Fig. 9. In particular this is a homeomorphism between $\dot{J}^+(T)$ and $\alpha(\dot{J}^+(T)) \subset \Sigma$. Since the former is a closed set, so is the latter.

Now $\dot{J}^+(T)$ is a 3D submanifold, and hence for any $p \in \dot{J}^+(T)$, we can find a neighborhood V of p with $V \subset \dot{J}^+(T)$. Then $\alpha(V)$ gives a neighborhood of $\alpha(p)$ in the image $\alpha(\dot{J}^+(T))$. Hence the latter set is open in Σ.

Since $\alpha(\dot{J}^+(T))$ is both open and closed in Σ, it must be the entire set,

$$\alpha(\dot{J}^+(T)) = \Sigma.$$

But we showed that $\alpha(\dot{J}^+(T))$ was compact, and by assumption Σ was noncompact. Therefore we have reached a contradiction. ⊠

It seems like we haven't quite gotten the result we were promised. Penrose requires a trapped surface, and maybe it's the case that trapped surfaces never form. However, due to an incredible result by Christodoulou, it turns out that trapped surfaces do form when you scatter enough gravitons, at least in Minkowski backgrounds. Moreover, there's the notion of Cauchy stability, i.e. given initial conditions that lead to a trapped surface, we can perturb them, solve the Einstein equations, and see how the new solution depends on the perturbation. It turns out that generically, perturbed spacetimes which had trapped surfaces before perturbation do still feature trapped surfaces, and we can do analytic calculations in spherical symmetry which do have trapped surfaces in them. Numerically, trapped surfaces turn up all the time. Therefore life is not so bad. Trapped surfaces are basically generic and so singularities are generic. QED.

[27] That is, a closed subset of a compact set is compact.

[28] Some people make a distinction between manifold with boundary and manifold without boundary. We won't bother here, and assume that a manifold just looks like $\mathbb{R}^n$ everywhere.

Asymptotic flatness We have discussed asymptotic flatness in the context of initial conditions. What does it mean for a spacetime to be asymptotically flat? To understand this, we'll need the notion of *conformal compactification*.

Given a spacetime $(\mathcal{M}, g)$, we can define a new metric

$$\bar{g} = \Omega^2 g. \tag{13.2}$$

In particular, we can choose Ω such that the resulting spacetime $(\mathcal{M}, \bar{g})$ is extendible into $(\bar{\mathcal{M}}, \bar{g})$. That is, M is a proper subset of $\bar{\mathcal{M}}$. Note that since Ω^2 is non-negative (generally, positive), the causal structure of the spacetime will be respected by this transformation. Spacelike separations in g will also be spacelike in $\bar{g}$, and so on.

Let's start with Minkowski space. Let $(\mathcal{M}, g)$ be Minkowski in spherical coordinates,

$$g = -dt^2 + dr^2 + r^2 d\omega^2, \tag{13.3}$$

where we denote the round 2-sphere metric by ω to avoid confusion with the conformal factor Ω. Define advanced and retarded (light cone) coordinates

$$u = t - r, \quad v = t + r. \tag{13.4}$$

Under this change of coordinates, we get the metric

$$g = -dudv + \frac{1}{4}(u - v)^2 d\omega^2. \tag{13.5}$$

Now we'll make a somewhat stranger change of coordinates,

$$u = \tan p, \quad v = \tan q \tag{13.6}$$

where in the (p, q) coordinates, $-\pi/2 < p \leq q < \pi/2$. Substituting back into the metric, we get something that looks kind of weird. It's

$$g = \frac{1}{(2\cos p \cos q)^2}\left[-4dpdq + \sin^2(q - p)d\omega^2\right]. \tag{13.7}$$

Notice that this denominator diverges when $q \to \pi/2$, so this suggests to us a very natural choice of conformal factor Ω. It's just

$$\Omega = 2\cos q \cos p, \tag{13.8}$$

which gives us a new metric

$$\bar{g} = -4dpdq + \sin^2(q - p)d\omega^2 = \Omega^2 g. \tag{13.9}$$

Finally, make the coordinate transformation

$$T = q + p \in (-\pi, \pi), \quad \chi = q - p \in [0, \pi). \tag{13.10}$$

What we get is the metric

$$\bar{g} = -dT^2 + \underbrace{d\chi^2 + \sin^2\chi d\omega^2}_{S^3}, \tag{13.11}$$

which looks very much like the Einstein static universe! Except there's a little caveat– the coordinate ranges mean that we don't get the whole of the cylinder. In fact, the range of χ also depends on T. An illustration of this appears in Fig. 10. Moreover, by projecting down into the T-χ plane (so that each point represents a 2-sphere), we get the Penrose diagram for Minkowski space, seen in Fig. 11. Penrose diagrams are very good for radial motion, but they come at the slight cost of hiding all angular motion behind our projection. In addition, note that under this projection, the points $(T = T_0, \chi = +\chi_0)$ and $(T = T_0, \chi = -\chi_0)$ are collapsed to the same point. I emphasize that there is nothing special about $\chi = 0$ in Minkowski space– it is utterly non-singular. This is just a quirk of our projection, taking advantage of the spherical symmetry of our spacetime.

It's also important to notice that we've introduced an unphysical, extendible metric $\bar{g}$. Minkowski space itself is inextendible, but after a conformal transformation we got a subset of the Einstein static universe (where this subset is very clearly extendible). Moreover, conformal transformations preserve the causal structure of spacetime, so we've found a very nice way to illustrate the causal structure of a spacetime. In the next lecture, we'll look at the solutions to wave equations living in curved spacetime.

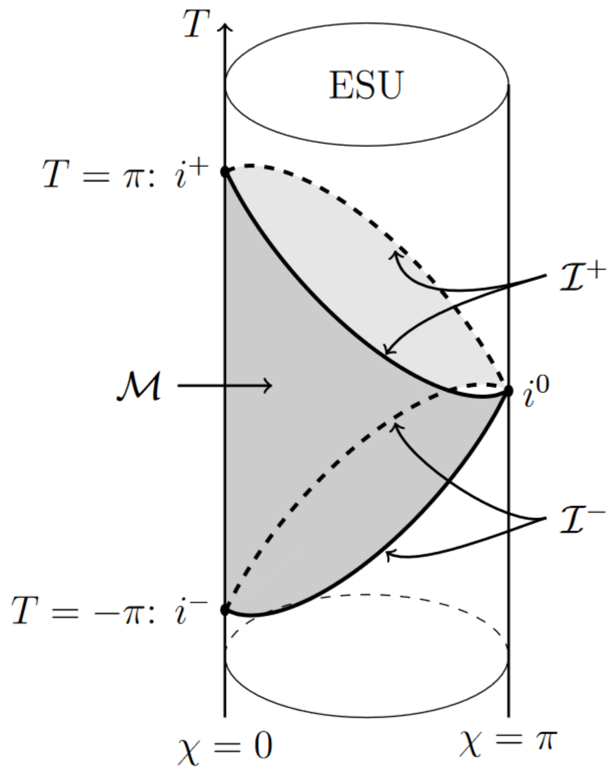

FIGURE 10. The patch of the Einstein static universe conformally equivalent to Minkowski space. The ESU has the topology of $\mathbb{R} \times S^3$, so we suppress two of the angular coordinates to draw this as a cylinder instead, $\mathbb{R} \times S^1$, where the T direction runs vertically and one of the angular coordinates (say, χ) runs around the circumference of the cylinder. Also indicated are future/past timelike infinity $i^{\pm}$, future/past null infinity $\mathcal{I}^{\pm}$, and spacelike infinity i^0.
Image credit to Prof. Reall's Black Holes notes, §5.1.

Lecture 14.

Monday, February 25, 2019

"I want to end this [lecture] with the definition of a black hole. Because you guys deserve it." –Jorge Santos

Last time, we saw that under a conformal compactification, we could bring infinity into a finite distance and draw the Penrose diagram to study its causal stucture.

We could study perturbations to the metric itself (i.e. the graviton), but we'll do something more straightforward– a massless scalar field. Take the massless Klein-Gordon equation:

$$\nabla^a \nabla_a \psi = 0. \tag{14.1}$$

Suppose we study spherically symmetric solutions, $\psi(t, r)$. The most general solution is

$$\psi = \frac{1}{r}[f(t-r) + g(t+r)]. \tag{14.2}$$

It's easy to check this– if $\psi = \frac{\Pi}{r}$ for some other field, then Π satisfies a regular wave equation which has left- and right-moving modes. Now, regularity at $r = 0$ demands that

$$\psi(t, r) = \frac{1}{r}[f(u) - f(v)] = \frac{1}{r}[F(p) - F(q)], \tag{14.3}$$

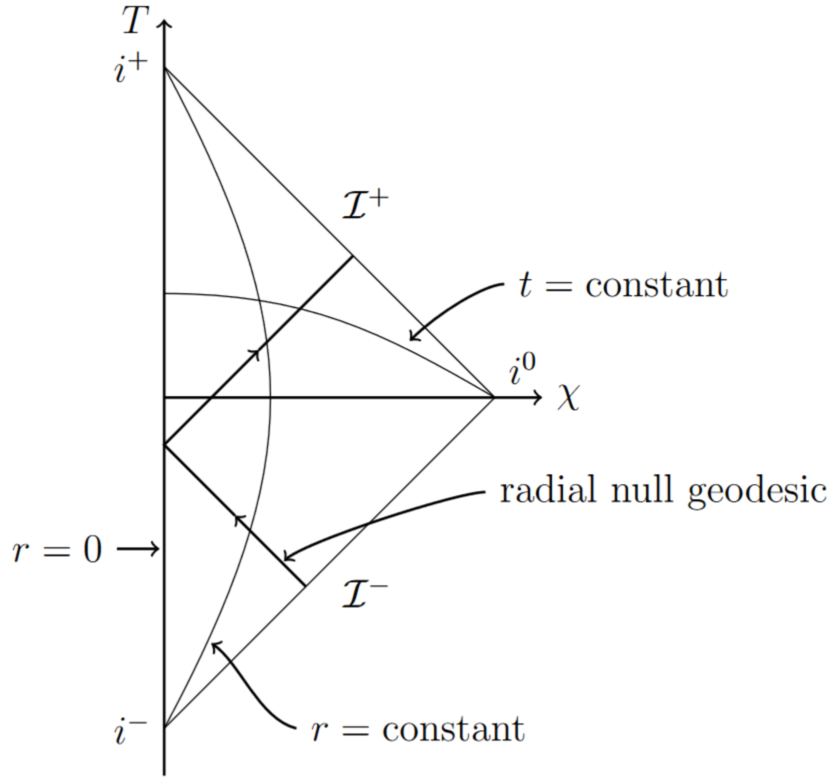

FIGURE 11. The Penrose diagram for Minkowski space. Note that the radial null geodesic drawn is not "reflecting" off anything; there is nothing special about $r = 0$ in Minkowski space. The geodesic simply passes through $r = 0$ and then proceeds back out to larger values of r as it heads towards future null infinity, $\mathcal{I}^+$.
Image credit to Prof. Reall's Black Holes notes, §5.1.

where $F(x) := f(\tan x)$. Let $F_0(q)$ denote the limiting value of $r\psi$ on $\mathcal{I}^-$, at $p = -\pi/2$. Thus

$$F(-\pi/2) - F(q) = F_0(q). \tag{14.4}$$

But this is valid for any q, so it is valid for any p. We conclude that

$$\psi(t,r) = \frac{1}{r}[F_0(p) - F_0(q)]. \tag{14.5}$$

Thus it suffices to specify initial data on $\mathcal{I}^-$, past null infinity.

Penrose diagram of Kruskal spacetime It's a bit fiddly to explicitly construct the Penrose diagram for the Kruskal spacetime, so instead we'll use our knowledge of the asymptotics to draw the diagram. Take the Kruskal diagram, and notice that regions I and IV are asymptotically flat, so must look like Minkowski. Null rays already travel at 45 degrees, so this is nice. Finally, we can straighten out the singularity by some choice of Ω to get the Penrose diagram for the Schwarzschild black hole in asymptotically flat space. It's pretty straightforward to see from the lines of constant r that i^+ is a singular point since it intersects $r = 0$, which is singular. But it's also true (though much harder to prove) that spacelike null infinity i^0 is also singular.

We can also draw the Penrose diagram for spherical collapse. The surface of the star follows a timelike geodesic, collapsing to $r = 0$. When it hits $r = 0$, we get a singularity, and there's also some horizon cutting off the black hole region.

Now, we'd like to discuss asymptotic flatness, but timelike and spacelike infinity are both singular. What's left? Null infinity, $\mathcal{I}^\pm$.

Asymptotic flatness

Definition 14.6. A *manifold-with-boundary* is defined in the same way as a manifold, except charts now map $\phi : \mathcal{M} \to \mathbb{R}^n/2 = \{(x^1, \dots, x^n); x^1 \leq 0\}$. The boundary $\partial\mathcal{M}$ of $\mathcal{M}$ is defined to be the set of points which have $x^1 = 0$ in some chart.

Definition 14.7. A time-orientable spacetime $(\mathcal{M}, g)$ is asymptotically flat at null infinity if there exists a spacetime $(\bar{\mathcal{M}}, \bar{g})$ such that

(a) $\exists$ a positive function Ω defined on $\mathcal{M}$ such that $(\bar{\mathcal{M}}, \bar{g})$ is an extension of $(\mathcal{M}, \Omega^2 g)$.
(b) Within $\mathcal{M}$, $\mathcal{M}$ can be extracted to obtain a manifold-with-boundary $\mathcal{M} \cup \partial\mathcal{M}$.
(c) Ω can be extended to a function on $\bar{\mathcal{M}}$ such that $\Omega = 0$ and $d\Omega \neq 0$ on $\partial\mathcal{M}$.
(d) The boundary $\partial\mathcal{M}$ is a disjoint union of two components of $\mathcal{I}^+, \mathcal{I}^-$, each diffeomorphic to $\mathbb{R} \times S^2$.
(e) No past- (future-)directed causal curve starting in $\mathcal{M}$ intersects $I^+(\mathcal{I}^-)$.
(f) $\mathcal{I}^\pm$ are 'complete.'

To recap, conditions a-c are conditions on how to choose Ω, while the other conditions describe the spacetime, and we haven't actually explained what 'complete' means yet.

Now, we've seen that

$$R_{ab} = \bar{R}_{ab} + 2\Omega^{-1}\bar{\nabla}_a\bar{\nabla}_b\Omega + \bar{g}_{ab}\bar{g}^{cd}\left(\Omega^{-1}\bar{\nabla}_c\bar{\nabla}_d\Omega - 3\frac{\nabla_c\Omega\nabla_d\Omega}{\Omega^2}\right). \tag{14.8}$$

In pure gravity, we take vacuum solutions of the Einstein equations and hence $R_{ab} = 0$. Note barred quantities are taken with respect to $(\mathcal{M}, \Omega^2 g)$. Multiplying by a single factor of Ω, we get

$$\Omega\bar{R}_{ab} + 2\bar{\nabla}_a\bar{\nabla}_b\Omega + \bar{g}_{ab}\bar{g}^{cd}\left(\bar{\nabla}_c\bar{\nabla}_d\Omega - 3\frac{\nabla_c\Omega\nabla_d\Omega}{\Omega}\right) = 0. \tag{14.9}$$

Regularity at $\Omega = 0$ implies that

$$\bar{g}^{cd}\bar{\nabla}_c\Omega\bar{\nabla}_d\Omega \propto \Omega. \tag{14.10}$$

So $d\Omega$ is null when $\Omega = 0$ ($\mathcal{I}^+$).

We deduce that $\mathcal{I}^+$ is a null hypersurface. There's still a residual freedom in our choice of Ω, though– we can still choose $\Omega' = \omega^2\Omega$. And this means we can be more clever about our choice of Ω. Choose ω^2 such that

$$\Omega^{-1}\bar{g}^{cd}\bar{\nabla}_c\Omega\bar{\nabla}_d\Omega = 0 \text{ on } \mathcal{I}^+. \tag{14.11}$$

That is, $\bar{g}^{cd}\bar{\nabla}_c\Omega\bar{\nabla}_d\Omega$ goes as Ω^2 so that $\Omega^{-1}\bar{g}^{cd}\bar{\nabla}_c\Omega\bar{\nabla}_d\Omega \to 0$ as $\Omega \to 0$.

Using this, we find that

$$\bar{\nabla}_a\bar{\nabla}_b\Omega = 0 \text{ on } \mathcal{I}^+, \tag{14.12}$$

and hence

$$\bar{\nabla}_a n^b = 0 \text{ on } \mathcal{I}^+, \tag{14.13}$$

which tells us that

$$n^a\bar{\nabla}_a n^b = 0 \text{ on } \mathcal{I}^+. \tag{14.14}$$

So the geodesics of n^a are null and they foliate $\mathcal{I}^+$. We introduce coordinates near $\mathcal{I}^+$ as follows: if we have some coordinates on $\mathcal{I}^+$, we can follow n^a on $\mathcal{I}^+$ a affine parameter distance u to get (u, θ, ϕ) coordinates, and we follow Ω off of $\mathcal{I}^+$ to get to points at $(\Omega, u, \theta, \phi)$. But in fact this is nothing more than Gaussian null coordinates, which means we can immediately write down the metric:

$$\bar{g}_{\Omega=0} 2du d\Omega + d\theta^2 + \sin^2\theta d\phi^2. \tag{14.15}$$

For small $\Omega \neq 0$, the metric components will differ from the above by $O(\Omega)$ terms. (This is hard to prove and not particularly illuminating, so we won't do it.)

If we now introduce $r = 1/\Omega$, then

$$g = -2du rd + r^2(d\theta^2 + \sin^2\theta d\phi^2) + \dots \tag{14.16}$$

where $\dots$ indicates corrections to higher order in Ω. The completeness condition on $\mathcal{I}^+$ says that $u \in (-\infty, +\infty)$.

Definition of a black hole

Definition 14.17. A *black hole interior* is the region

$$\mathcal{B} = \mathcal{M} \setminus [\mathcal{M} \cap J^-(\mathcal{I}^+)]. \tag{14.18}$$

The black hole *future event horizon* is

$$\mathcal{H}^+ = \mathcal{M} \cap \dot{J}^-(\mathcal{I}^+). \tag{14.19}$$

That is, take the causal past of future null infinity. Follow the light rays back everywhere you can get on your manifold, and take the complement. What remains is the black hole.

Lecture 15.

Wednesday, February 27, 2019

> *"If you want to work in mathematical GR, this is the problem to tackle. Your supervisor will never give you this problem. But this is the one." –Jorge Santos*

Last time, we finally defined a black hole. Recall that the (future) horizon is

$$\mathcal{H}^+ = \mathcal{M} \cap \dot{J}^-(\mathcal{I}^+).$$

The horizons $\mathcal{H}^+$ are null hypersurfaces, and the generators of $\mathcal{H}^+$ cannot have future endpoints. For instance, we saw this in the spherical collapse of a star.

Today we'll show in full generality that black holes cannot bifurcate. To prove this, we'll need some definitions.

Definition 15.1. An asymptotically flat spacetime $(\mathcal{M}, g)$ is *strongly asymptotically predictable* if $\exists$ an open region $\bar{V} \subset \bar{\mathcal{M}}$ such that $\overline{\mathcal{M} \cap J(I^+)} \subset \bar{V}$ and $(\bar{V}, g)$ is globally hyperbolic.

That is, if we start with a partial Cauchy surface and some initial data on the horizon, we can predict everything.

Theorem 15.2. *Let $(\mathcal{M}, g)$ be strongly asymptotically predictable, and let Σ_1, Σ_2 be Cauchy surfaces for $\bar{V}$ with $\Sigma_2 \subset I^+(\Sigma_1)$. Let B be a connected component of $\mathcal{B} \cap \Sigma_1$. Then $J^+(B) \cap \Sigma_2$ is contained within a connected component of $\mathcal{B} \cap \Sigma_2$.*

Proof. The proof is by contradiction. Global hyperbolicity implies that a causal curve from Σ_1 will intersect Σ_2. Note that $J^+(B) \subset B$, so $J^+(B) \cap \Sigma_2 \subset \mathcal{B} \cap \Sigma_2$. Thus the causal future of the black hole region on Σ_1 is contained within the black hole region on Σ_2.

Now assume that $J^+(B) \cap \Sigma_2$ is not constrained within a single connected component of $\mathcal{B} \cap \sigma_2$. Then we can find disjoint open sets $O, O' \subset \Sigma_2$ such that $J^+(B) \cap \Sigma_2 \subset O \cup O'$ with $J^+(B) \cap O \neq \emptyset$ and $J^+(B) \cap O' \neq \emptyset$. Then $B \cap I^-(O)$ and $B \cap I^-(O')$ are non-empty and $B \subset I^-(O) \cup I^-(O')$. Now $p \in B$ cannot lie in both $I^-(O)$ and $I^-(O')$, for then we could divide future-directed timelike geodesics into two sets according to whether they intersect O or O', and hence divide future-directed timelike vectors at p into two disjoint open sets, contradicting connectedness of the future light cone at p. ⊠

Weak cosmic censorship Weak cosmic censorship is the biggest outstanding problem in mathematical general relativity. Consider the Penrose diagram for the Schwarzschild black hole with $M < 0$. This spacetime has a naked singularity, a singularity that is visible from timelike infinity.. But maybe such spacetimes are unrealistic– can we form a naked singularity through gravitational collapse?

Birkhoff's theorem tells us we can't have a spherically symmetric collapse in pure gravity because the most general spherically symmetric solution is static. So suppose we add in a scalar field. The region where the naked singularity is visible is $\overline{D^+(\Sigma)} \setminus I(D^+(\Sigma))$. In fact, we really shouldn't talk about the singularity itself. Beyond the Cauchy horizon, solutions become highly non-unique, so we should remove that region from the spacetime. Instead, it's more appropriate to talk about $\mathcal{I}^+$– if $\mathcal{I}^+$ is incomplete, then we have a naked singularity.

Conjecture. Weak cosmic censorship (Horowitz and Geroch). *Let (Σ, h, K) be a geodesically complete, asymptotically flat initial data set. Let the matter fields obey hyperbolic equations and satisfy the dominant energy condition. Then generically the maximal development of this initial data is an asymptotically flat spacetime.*

Apparent horizons

Theorem 15.3. *Let T be a trapped surface in a strongly asymptotically predictable spacetime obeying the null energy condition. Then $T \subset \mathcal{B}$ (the trapped surface is within the black hole region).*

Proof. Assume there exists $p \in T$ such that $p \notin \mathcal{B}$. Thus p is in the causal past of $\mathcal{I}^+$, i.e. $p \notin J^-(\mathcal{I}^+)$. Then there exists a causal curve from p to $\mathcal{I}^+$.

One can then use strong asymptotic predictability to show that this implies that $\dot{J}^+(T)$ must intersect $\mathcal{I}^+$, i.e. there exists a $q \in \mathcal{I}^+$ with $q \in \dot{J}^+(T)$. But we have seen (from two lectures ago) that q lies on a null geodesic γ from $r \in T$ that is orthogonal to T and has no point conjugate to r along γ. Since T is trapped, the expansion of the null geodesics orthogonal to T is negative at r, and hence $\theta \to -\infty$ within finite affine parameter along γ. Thus $\exists$ a point s conjugate to r along γ, which is a contradiction. ☒

Definition 15.4. Let Σ_t be a Cauchy surface in a globally hyperbolic spacetime $(\mathcal{M}, g)$. The trapped region τ_t of Σ_t is the set of points $p \in \Sigma_t$ for which there exists a trapped surface S with $p \in S \subset \Sigma_t$. The apparent horizon $\mathcal{A}_t$ is the boundary of τ_t.

Note that the apparent horizon depends on the foliation of the spacetime, which depends on the choice of Cauchy surface Σ_t. Therefore we must be very careful about trusting numerical results on (weak) cosmic censorship, since one can only determine the apparent horizon and not the event horizon itself.

Lecture 16.

Thursday, February 28, 2019

"As soon as someone shows you this Penrose diagram, you should tell them, 'Hey, you've never been to Cambridge. People there will yell at you for doing this.'" –Jorge Santos

Today we will discuss the Reissner-Nordström solution, the metric describing an electrically (or magnetically) charged black hole.[29]

The Reissner-Nordström solution Astrophysically speaking, we do not expect to observe black holes with large electrical charges in nature. So why should we study them?

(a) RN black holes provide a close analogue of the Kerr (rotating) solutions, since they have inner (Cauchy) horizons and very similar Penrose diagrams (i.e. causal structures).
(b) RN black holes are also ubiquitous in string theory. In fact, string theory provides a correct counting of the microstates of the Reissner-Nordström solution (related to the entropy of the black hole).

The Reissner-Nordström solution solves the equations of motion which come from varying the Einstein-Maxwell action

$$S = \frac{1}{16\pi G} \int d^4x \sqrt{-g} \Big(R - F_{ab} F^{ab} \Big) \tag{16.1}$$

which describes a massless vector field (photon) coupled to gravity. Here, $F = dA$, where A is a 1-form. The corresponding equations of motion are

$$R_{ab} - \frac{R}{2} g_{ab} = 2 \left(F_{ac} F_b{}^c - \frac{1}{4} g_{ab} F^{cd} F_{cd} \right), \tag{16.2}$$

$$\nabla^b F_{ab} = 0. \tag{16.3}$$

That is, the electromagnetic field has a corresponding stress-energy tensor, and F_{ab} the electromagnetic field strength tensor obeys a conservation equation.

Theorem 16.4. *The unique spherically symmetric solution to the Einstein-Maxwell equations with a non-constant radius function r is the Reissner-Nordström solution*

$$ds^2 = -\left(1 - \frac{2M}{r} + \frac{e^2}{r^2} \right) dt^2 + \frac{dr^2}{1 - \frac{2M}{r} + \frac{e^2}{r^2}} + r^2 d\Omega_2^2 \tag{16.5}$$

[29] n.b. this lecture was transcribed from a handwritten set of notes I took after my laptop ran out of battery, so it may feature somewhat less commentary than usual. Whether or not this is an improvement is up to the reader to decide.

and

$$A = -\frac{Q}{r}dt - P\cos\theta d\phi, \quad e \equiv Q^2 + P^2, \tag{16.6}$$

where Q and P are electric and magnetic charges, respectively.

Note that this solution is static, i.e. $\left(\frac{\partial}{\partial t}\right)^a = K^a$ is a Killing field for this metric. It is also asymptotically flat, since it reduces to the Minkowski metric in the limit $r \to \infty$.

To discuss the properties of this solution, let us define

$$\Delta(r) \equiv r^2 - 2Mr + e^2 = (r - r_-)(r - r_+), \tag{16.7}$$

which is just a factorization of the g_{tt} component of the metric. The zeroes of $\Delta(r)$ are then

$$r_\pm = M \pm \sqrt{M^2 - e^2}, \tag{16.8}$$

which reduces to $r = 2M$ (the Schwarzschild event horizon) in the $e \to 0$ limit.

Notice that if $e > M$, $r_\pm$ are both complex, so there are no real horizons. However, in Reissner-Nordström there is always a curvature singularity at $r = 0$, which one can determine e.g. by calculating the Kretchmann scalar $R^{abcd}R_{abcd}$ at $r = 0$. Therefore by weak cosmic censorship, we can neglect this case since it features a naked singularity. In fact, we can't even dynamically form such a solution since near extremality, throwing in more charge causes the field to disperse.

Another interesting case is $M = e$, where the two horizons coincide (i.e. the extremal Reissner-Nordström black hole). We'll defer discussion of this case to the end of lecture, and focus our efforts on the $M > e$ case, where there are two real, positive roots of Δ, i.e. $0 < r_- < r_+$. To understand the structure of this metric, we introduce Eddington-Finkelstein coordinates for our metric

$$ds^2 = -\frac{\Delta}{r^2}dt^2 + \frac{r^2 dr^2}{\Delta} + r^2 d\Omega_2^2. \tag{16.9}$$

That is, introduce a tortoise coordinate r_* defined by

$$dr_* = \frac{r^2}{\Delta}dr \implies r_* = r + \frac{1}{2\kappa_+}\log\left|\frac{r - r_+}{r_+}\right| + \frac{1}{2\kappa_-}\log\left|\frac{r - r_-}{r_-}\right|, \tag{16.10}$$

where $\kappa_\pm = \frac{r_\pm - r_\mp}{2r_\pm^2}$ are constants which we will later identify as the surface gravities of the horizons.

Now introduce the ingoing and outgoing coordinates

$$u = t - r_*, \quad v = t + r_*. \tag{16.11}$$

In (v, r, θ, ϕ) coordinates, the metric takes the form

$$ds^2 = -\frac{\Delta}{r^2}dv^2 + 2dvdr + r^2 d\Omega_2^2. \tag{16.12}$$

In the original coordinates, this solution was non-singular up to r_+, but we can clearly extend this now to $r \in (0, r_+]$ (e.g. by noticing the metric determinant is non-singular at r_+, as with Schwarzschild).

A surface of constant r now defines a surface with normal $n_a = (dr)_a$. But notice that $n^2 = g^{rr} = \frac{\Delta}{r^2}$, so the surfaces with $r = r_\pm$ are then null hypersurfaces. The proof that these represent horizons is then the same as for the Schwarzschild solution. Thus $r = r_+$ is the future event horizon of the Reissner-Nordström spacetime.

The maximal extension of Reissner-Nordström In Schwarzschild, a bit of careful work with coordinate ranges and redefinitions led us to the Kruskal extension, the maximally extended Schwarzschild solution. Can we do the same for Reissner-Nordström?

The answer is yes, but the extended RN spacetime has a few surprises in store. Let us introduce Kruskal-like coordinates

$$U^\pm = -e^{-\kappa_\pm u}, V^\pm = \pm e^{\kappa_\pm v}. \tag{16.13}$$

Let us start in the exterior region, $r > r_+$. In $U^+, V^+, \theta, \phi)$ coordinates we find that the metric takes the form

$$ds^2 = -\frac{r_+ r_-}{\kappa_+^2 r^2}e^{-2\kappa_+ r}\left(\frac{r - r_-}{r_-}\right)dU^+ dV^+ + r^2 d\Omega_2^2, \tag{16.14}$$

where r should be considered as a function of U and V, given implicitly by

$$-U^+V^+ = e^{2\kappa_+ r}\left(\frac{r-r_+}{r_+}\right)\left(\frac{r_-}{r-r_-}\right)^{\kappa_+/|\kappa_-|}. \tag{16.15}$$

The RHS of our expression for U^+V^+ is a monotonically increasing function of r^* for $r > r_-$. Note that in ingoing coordinates, we could hit the $r = 0$ singularity, but in these Kruskal coordinates, we first have to stop at $r = r_-$. We can draw the Penrose diagram for this spacetime.

Something weird has happened– If we continue to analytically extend this solution, we find that if we cross the Cauchy horizon, there is yet another $r = r_-$ horizon to our future, whereupon crossing this new horizon drops us in another asymptotically flat region, and we can continue this process indefinitely. However, there is also a timelike singularity at $r = 0$ which we can see, but need not cross. This singularity still removes regions from the domain of dependence, creating a Cauchy horizon at $r = r_-$. Therefore strong cosmic censorship suggests that we shouldn't trust this diagram beyond the Cauchy horizon, since we generically expect a singularity to form at the Cauchy horizon, cutting off the new region.

Let us now revisit the extremal case. For $M = e$, we have $r_+ = r_-$, and the metric is then

$$ds^2 = -\left(1-\frac{M}{r}\right)^2 dt^2 + \frac{dr^2}{\left(1-\frac{M}{r}\right)^2} + r^2 d\Omega_2^2. \tag{16.16}$$

The Penrose diagram for this is a little different. We still get a single horizon, but there is an asymptotic region replacing $r = 0$ in the asymptotically flat diamonds of the Penrose diagram.

To make sense of this, we can study the near-horizon region by expanding $r = M(1+\lambda)$ to leading order in λ for $\lambda \ll 1$. In that case, the metric takes the form

$$ds^2 = \underbrace{-\lambda^2 dt^2 + M^2\frac{d\lambda^2}{\lambda^2}}_{\text{AdS}_2} + \underbrace{M^2 d\Omega^2}_{S^2}, \tag{16.17}$$

which has the form of $\text{AdS}_2 \times S^2$, i.e. a region with an AdS "throat" going to $r = 0$. This is known as the Robinson-Bertotti metric.

Majumdar-Papapetrou solutions Consider the extremal RN solution under the coordinate transformation $\rho = r - M$. If we take $P = 0$ (no magnetic charge) we get the metric

$$ds^2 = -H^2 dt^2 + H^2(d\rho^2 + \rho^2 d\Omega_2^2), \tag{16.18}$$

where

$$H \equiv 1 + \frac{M}{\rho}, \quad A = H^{-1}dt. \tag{16.19}$$

However, suppose now we write down a metric

$$ds^2 = -H(\mathbf{x})^2 dt^2 + H(\mathbf{x})^2 d\mathbf{x}^2, \quad A = H(\mathbf{x})^{-1}dt. \tag{16.20}$$

Plugging into the Einstein equations, we get a condition $\nabla_x^2 H = 0$, whose solutions are harmonic functions of $\mathbf{x}$. Therefore take

$$H(\mathbf{x}) = 1 + \sum_{i=1}^{N} \frac{M_i}{|\mathbf{x} - \mathbf{x}^{(i)}|}, \tag{16.21}$$

which is a solution by linearity. Now the $\mathbf{x}^{(i)}$s specify locations of black holes in our spacetime, and we can put them wherever we want.[30] A priori, there's no spherical symmetry in these solutions. The gravitational attraction is precisely balanced by the charge of these black holes– in a sense, they are supersymmetric (albeit highly unstable).

[30] "With the solutions I'm gonna tell you, you can write your name. Whatever sh*t you want, you can build it out of black holes." –Jorge Santos

Lecture 17.

Friday, March 1, 2019

"In fact, you become a dude if you are here because you are forced to surf the black hole. You HAVE to surf the black hole. So physicists are COOL." –Jorge Santos

Rotating black holes (Kerr) Last time, we described charged black holes. Highly charged black holes are probably not astrophysically relevant, but black holes with large angular momentum are. In the realm of stationary solutions, one might hope to make a classification of stationary solutions to the Einstein equations which contain black holes. To do this, we'll have to discuss *uniqueness theorems.*

Definition 17.1. A spacetime which is asymptotically flat at null infinity is *stationary* if it admits a Killing vector field K^a that is timelike in a neighborhood of $\mathcal{I}^+$.

This is a little different from our old definition, where we required a timelike Killing vector field everywhere (which we might now call strictly stationary). Here, we only care about what's happening at null infinity.

It is convenient to normalize this Killing vector field to

$$K^2 = -1 \text{ at } \mathcal{I}^+. \tag{17.2}$$

We need another definition to discuss rotating black holes.

Definition 17.3. A spacetime asymptotically flat at null infinity is *stationary* and *axisymmetric* if

(i) it is stationary,
(ii) it admits a Killing vector field m^a that is spacelike near $\mathcal{I}^+$,
(iii) m^a generates a 1-parameter group of isometries isomorphic to $U(1)$ (i.e. 2D rotations)
(iv) $[K, m] = 0$.

This tells us precisely that we have a generator of rotations m which commutes with the generator of time translations K.

Theorem 17.4 (Israel 1967, Bunting and Masood 1987)**.** *If $(\mathcal{M}, g)$ is a static, asymptotically flat vacuum black hole spacetime that is suitably regular on and outside the event horizon, then $(\mathcal{M}, g)$ is isometric to the Schwarzschild solution.*

However, let us assume that our spacetime is not static, just stationary. There is an analogue of this theorem for Einstein-Maxwell.

Theorem 17.5 (Hawking 1973, Wald 1991)**.** *If $(\mathcal{M}, g)$ is a stationary, non-static, asymptoticallly flat analytic solution of the Einstein-Maxwell equation that is suitably regular that is suitably regular on and outside the event horizon, then $(\mathcal{M}, g)$ is stationary and axisymmetric.*

This is sometimes called a "rigidity" theorem, and the analyticity condition is often considered too strong (since analyticity says that the entire spacetime can be determined by its behavior around any single point).

Theorem 17.6 (Carter 1971, Robinson 1975)**.** *If $(\mathcal{M}, g)$ is a stationary, axisymmetric asym. flat vacuum spacetime regular on and outside a connected event horizon, then $(\mathcal{M}, g)$ is a member of the 2-parameter Kerr (1963) family of solutions, with the two parameters mass M and angular momentum J.*

In fact, black holes have a remarkable property– no matter what sort of matter we use to build the black hole, all that remains after collapse are two parameters in pure gravity (or four in Einstein-Maxwell). These are the *no-hair theorems*, and they only strictly hold in four dimensions.

If we turn on charge, we get the Kerr-Newman solution, written in so-called Boyer-Lindquist coordinates. There are four parameters of this solution, a (angular momentum), M (mass), P, and Q (magnetic and electric charges). The Kerr-Newman solution enjoys two Killing vectors,

$$K^a = \left(\frac{\partial}{\partial t}\right)^a, \quad m^a = \left(\frac{\partial}{\partial \phi}\right)^a, \tag{17.7}$$

and there is also a nice symmetry where the line element is invariant under the symmetry

$$(t,\phi) \to (-t,-\phi). \tag{17.8}$$

The parameter $a = J/M$ is simply a normalized version of the old angular momentum.

The *Kerr solution* is the Kerr-Newman solution with no charge, $Q = P = 0$. Define $\Delta \equiv (r-r_+)(r-r_-), r_\pm = M \pm \sqrt{M^2-a^2}$. Notice the analogy to Reissner-Nordström– if $M^2 < a^2$ we have a naked singularity (ruled out by WCC conjecture). If $M^2 > a^2$, we have a black hole. If $M^2 = a^2$, we have the extremal Kerr solution.

From the form of Kerr-Newman, it looks like the metric becomes singular when $\Delta = 0$ (at $r = r_\pm$) or when $\Sigma = 0$ (at $r = 0, \theta = \pi/2$). The first two are coordinate singularities, whereas the $r = 0$ singularity is a proper curvature singularity.

We introduce now the equivalent of EF coordinates, which here are called Kerr coordinates:

$$dv = dt + \frac{r^2+a^2}{\Delta}dr, \quad d\chi = d\phi + \frac{a}{\Delta}dr. \tag{17.9}$$

In thse coordinates, the Killing vectors become $K = \frac{\partial}{\partial v}$ and $m = \frac{\partial}{\partial \chi}$, where χ is periodic with $\chi \sim \chi + 2\pi$. In these coordinates, we can extend our solution to $0 < r < r_+$.

Proposition 17.10. *The surface $r = r_+$ is a null hypersurface with normal $\xi^a = K^a + \Omega_H m^a$, where $\Omega_H = \frac{a}{r_+^2+a^2}$.*

Proof. First compute ξ_μ in Kerr coordinates. We will find that

$$\xi_\mu dx^\mu|_{r=r_+} \propto dr, \quad \xi^\mu \xi_\mu|_{r=r_+} = 0.$$

⊠

In BL coordinates, $\xi = \frac{\partial}{\partial t} + \Omega_H \frac{\partial}{\partial \phi}$ and hence

$$\xi^\mu \partial_\mu(\phi - \Omega_H t) = 0 \implies \phi = \text{const} + \Omega_H t. \tag{17.11}$$

Thus the generators of the horizon travel with an angular velocity given by Ω_H.

Maximal analytic extension Much of the structure of the Kerr solution will be similar to the RN solution– we'll still get an event horizon and a Cauchy horizon. However, the Kerr solution is not spherically symmetric, so we must restrict ourselves to a submanifold– we'll choose $(\theta = 0, \theta = \pi)$. Fortunately, this is totally geodesic (i.e. geodesics don't escape the submanifold).

But (as we'll show on the examples sheet) the singularity of Kerr is not a point but a ring. One can go "through" the center of the ring and end up in another asymptotically flat end, and near the singularity there are closed timelike curves. Again, strong cosmic censorship saves us– it tells us that we really shouldn't trust the diagram beyond the Cauchy horizon, and this result is strongly suggested (though not quite proven) by work by Dafermos.

The ergosurface and Penrose process Let us look at the time translation Killing field K:

$$K^2 = g_{tt} = -\left[1 - \frac{2Mr}{r^2 + a^2\cos^2\theta}\right]. \tag{17.12}$$

For what values of r is K^2 timelike? Taking $K^2 < 0$, we find that

$$r > M + \sqrt{M^2 - a^2\cos^2\theta} \geq r_+. \tag{17.13}$$

There is a boundary known as the *ergosphere* bounding a region known as the *ergoregion*, some of which lies outside the black hole. Notice that the sign of g_{tt} flips sign within the ergoregion, even though this region lies outside the event horizon.[31]

Suppose now we have a stationary observer on a timelike trajectory that follows the integral curves of K. Such an observer is dragged along with the co-rotating frame, and we'll moreover show that we can use this phenomenon to extract energy from the black hole.

[31] This is because the spacetime is stationary but not static. In a stationary spacetime like Schwarzschild, the Killing horizon coincides with the event horizon. So the sign flip of g_{tt} is associated with the fact that the event horizon is a point of no return, since the r component also changes sign. In Kerr, the r component changes sign at the event horizon but not at the ergosurface, so one can still escape the ergoregion in the radial direction.

Lecture 18.

Monday, March 4, 2019

"What is the area of the black hole?" "Zero?" "No, it's the other guy. Infinity." –Jorge Santos and a student

Penrose process Last time, we mentioned that rotating black holes drag along "stationary" observers in a region outside the black hole region. That is, the originally timelike killing vector $K = \frac{\partial}{\partial t}$ is such that within the ergoregion, $K^2 > 0$, so timelike observers are forced to "surf" the black hole.

More concretely, consider a particle with 4-momentum

$$P^a = \mu U^a \tag{18.1}$$

where μ is the rest mass and U^a is the four-velocity. The energy along such a geodesic is conserved,

$$E = -K \cdot P. \tag{18.2}$$

Suppose this particle carries a bomb and breaks apart so that

$$P^a = P_1^a + P_2^a \implies E = E_1 + E_2, E_i = -K \cdot P_i. \tag{18.3}$$

Since K is spacelike in the ergoregion, $E_1 < 0$. But we see that

$$E_2 = E + |E_1| > E, \tag{18.4}$$

so the particle with energy E_2 leaves with more energy that went in. This is known as the *Penrose process*. How much energy can we steal from the black hole? A particle that crosses $\mathcal{H}^+$ must have $-P \cdot \xi \geq 0$,[32] which tells us that

$$E - \Omega_H L \geq 0 \tag{18.5}$$

with $L = m \cdot P$ the angular momentum of the particle that fell in. After the system settles, it must be that

$$\delta M = E_1, \quad \delta J = L. \tag{18.6}$$

Comparing to our inequality, we find that

$$\delta J \leq \frac{\delta M}{\Omega_H} = \frac{2M(M^2 + \sqrt{M^4 - J^2})}{J} \delta M, \tag{18.7}$$

or with a bit of manipulation,

$$\delta M_{\text{irr}} \geq 0, \quad M_{\text{irr}} \equiv \left[\frac{1}{2}(M^2 + \sqrt{M^4 - J^2})\right]^{1/2}, \tag{18.8}$$

where we have defined the irreducible mass M_{irr}. Solving for M^2, we find that

$$M^2 = M_{\text{irr}}^2 + \frac{J^2}{4M_{\text{irr}}^2} \geq M_{\text{irr}}^2 \tag{18.9}$$

[32] This just says that the horizon generator ξ and the particle momentum P lie in the same light cone.

Thus the irreducible mass places a bound on how much mass and angular momentum we can extract from the black hole. The area of the intersection of a partial Cauchy surface $t =$ constant with $\mathcal{H}^+$ is[33]

$$A = 16\pi M_{\text{irr}}^2 \implies \delta A \geq 0. \tag{18.10}$$

Mass, charge, and angular momentum For our next trick, we'll show that the quantities we have been calling M, Q, and J correspond to the ideas of mass, charge, and angular momentum in a sensible way. Let us begin with electromagnetism.

$$\nabla^a F_{ab} = -4\pi J_b, \quad \nabla_{[a} F_{bc]} = 0. \tag{18.11}$$

This is just Maxwell's equations and the Bianchi identity on the field strength tensor. The *Hodge dual* (Hodge star) of a p-form ω_p is a $d - p$ form defined by

$$(*\omega_p)^{a_1 \ldots a_{d-p}} \equiv \frac{1}{p!} \frac{\epsilon^{a_1 \ldots a_{d-p} b_1 \ldots b_p}}{\sqrt{-g}} \omega_{b_1 \ldots b_p}, \tag{18.12}$$

where ϵ is the totally antisymmetric rank-p tensor (fixing some orientation for the spacetime). Using differential forms, we can write Maxwell's equations compactly as

$$d * F = -4\pi * j, \quad dF = 0. \tag{18.13}$$

The second of these tells us by the Poincaré lemma that $dF = 0 \implies F = dA$ for some one-form A, at least locally. Consider a spacelike surface Σ, such that

$$Q \equiv -\int_\Sigma *j \tag{18.14}$$

defines a charge. The key thing about a charge is that it should be conserved, i.e. if we take another spacelike surface Σ', the charge is unchanged. By the Maxwell equation,

$$Q = -\int_\Sigma *j = \frac{1}{4\pi} \int_\Sigma d * F = \frac{1}{4\pi} \int_{\partial\Sigma} *F, \tag{18.15}$$

where we have applied Stokes's theorem. This tells us that the value of the charge indeed depends only on the value of $*F$ on the boundary $\partial\Sigma$.

Definition 18.16. Let (Σ, h, K) be an asymptotically flat end. Then the electric and magnetic charges associated with this end are

$$Q = \frac{1}{4\pi} \lim_{r\to+\infty} \int_{S_r^2} *F, \quad P = \frac{1}{4\pi} \lim_{r\to+\infty} \int_{S_r^2} F, \tag{18.17}$$

where S_r^2 is an S^2 of radius r.

[33] To see this, notice that the induced metric on the horizon $r = r_+, dt = 0, dr = 0$ is

$$\gamma_{ij} dx^i dx^j = (r_+^2 + a^2 \cos^2\theta) d\theta^2 + \left[\frac{(r_+^2 + a^2)^2 \sin^2\theta}{r_+^2 + a^2\cos^2\theta} \right] d\phi^2.$$

The horizon area is

$$A = \int \sqrt{|\gamma_{ij}|} d\theta d\phi$$

and the determinant of the induced metric is just $(r_+^2 + a^2)^2 \sin^2\theta$ so

$$A = \int (r_+^2 + a^2) \sin\theta d\theta d\phi = 4\pi(r_+^2 + a^2).$$

Substituting in the definition of r_+, we find that

$$A = 4\pi(2(M^2 + \sqrt{M^4 - J^2}) = 16\pi M_{\text{irr}}^2.$$

This quick derivation comes from Carroll, pg. 270.

Komar integrals If a spacetime $(\mathcal{M}, g)$ is stationary, then there exists a conserved energy momentum current

$$J_a = -T_{ab}K^b \implies \nabla_a J^a = 0. \tag{18.18}$$

One can show this since T_{ab} is conserved and K^b is Killing. In the language of differential forms,

$$d * J = 0. \tag{18.19}$$

Note that T_{ab} need not be the stress-energy tensor on the RHS of the Einstein equations. We will simply treat it as a symmetric 2-tensor that is conserved, i.e. obeys $\nabla_a T^{ab} = 0$.

Hence we can define the total energy of matter on a spacelike hypersurface Σ as

$$E[\Sigma] = -\int_\Sigma *J. \tag{18.20}$$

Then

$$E[\Sigma'] - E[\Sigma] = -\int_{\partial R} *J = -\int_R d * J = 0 \tag{18.21}$$

by Stokes's theorem. In electromagnetism, we said that $*J = dX$ for some X. But GR is harder– the current here will not generally admit a form $*J = dX$. However, we can get close. Look at

$$\begin{aligned}(*d * dK)_a = -\nabla^b (dK)_{ab} &= -\nabla^b \nabla_a K_b + \nabla^b \nabla_b K_a \\ &= 2\nabla^b \nabla_b K_a\end{aligned}$$

where we have used the fact that K is Killing to rewrite the final line.

Lemma 18.22. *A Killing vector field obeys*

$$\nabla_a \nabla_b K^c = R^c{}_{bad} K^d. \tag{18.23}$$

Hence we have $(*d * dK)_a = -2R_{ab}K^b \equiv 8\pi J'_a$, where $J'_a = -2(T_{ab} - \frac{1}{2} T g_{ab})K^b$ by the Einstein equations. It follows that

$$d * dK = 8\pi * J'. \tag{18.24}$$

We know $*J'$ is exact and conserved, so

$$-\int_\Sigma *J' = -\frac{1}{8\pi}\int_\Sigma d * dK = -\frac{1}{8\pi}\int_{\partial\Sigma} *dK. \tag{18.25}$$

Definition 18.26. Let (Σ, h, K) be an asymptotically flat end in a stationary spacetime. The *Komar mass* (or energy) is

$$M_{\text{Komar}} = -\frac{1}{8\pi} \lim_{r\to+\infty} \int_{S^2_r} *dK \tag{18.27}$$

Definition 18.28. Let (Σ, h, K) be an asymptotically flat end in an axisymmetric spacetime. The *Komar angular momentum* is

$$J_{\text{Komar}} = \frac{1}{16\pi} \lim_{r\to+\infty} \int_{S^2_r} *dm, \quad m = \frac{\partial}{\partial\phi}. \tag{18.29}$$

Lecture 19.

Wednesday, March 6, 2019

"This is where we really start to see Stephen's influence on general relativity. From now on there is not a single class where I will not mention the name Stephen Hawking." –Jorge Santos

Last time, we talked about the Komar mass (energy), which let us define a mass for black hole systems that admit a Killing vector.

Hamiltonian formulation of GR For today, we'll take a new choice of units, $16\pi G = 1$, in order to avoid some awkward 16πs floating around. We begin with the action

$$ds^2 = -N^2 dt^2 + h_{ij}(dx^i + N^i dt)(dx^j + N^j dt), \tag{19.1}$$

where N is called the *lapse*, j_{ij} is some general 3×3 metric, and N^i is the *shift*. Note this is not Kaluza-Kelin because N, h_{ij}, and N^i all generically depend on time, e.g. $\partial_t N \neq 0$.

Now let us rewrite our action:

$$S = \int dt d^3x \mathcal{L} = \int dt d^3x \sqrt{h} N \left[{}^{(3)}R + K_{ij}K^{ij} - K^2 \right] \tag{19.2}$$

where ${}^{(3)}R$ is the ricci scalar of H and

$$K_{ij} = \frac{1}{2N}\left(\dot{h}_{ij} - D_i N - J - D_j N_i\right), \quad \cdot \equiv \partial_t \tag{19.3}$$

and

$$K = h^{ij}K_{ij}. \tag{19.4}$$

Notice there are no time derivatives of N or N^i, so they are not dynamical variables (by the equations of motion)– we are free to specify them, which corresponds to a coordinate choice. We vary the action with respect to N to get a Hamiltonian constraint, and with respect to N^i to get the momentum constraint. Note we must vary the action first, and then choose N.

To write a Hamiltonian, we must also compute a conjugate momentum of our system:

$$\pi^{ij} \equiv \frac{\delta S}{\delta \dot{h}_{ij}} = \sqrt{h}(K^{ij} - Kh^{ij}. \tag{19.5}$$

From here we compute

$$\begin{aligned} H &= \int d^3x (\pi^{ij}\dot{h}_{ij} - \mathcal{L}) \\ &= \int d^3x \sqrt{h}\left(N\mathcal{H} + \mathcal{H}_i N^i\right), \end{aligned}$$

where

$$\mathcal{H} = -{}^{(3)}R + h^{-1}\pi^{ij}\pi_{ij} - \frac{1}{2}h^{-1}\pi^2 \tag{19.6}$$

$$\mathcal{H}_i = -2h_{ik}D_j(h^{-1/2}\pi^{jk}) \tag{19.7}$$

where π here is $\pi \equiv h^{ij}\pi_{ij}$. Thus N, N^i act like Lagrange multipliers, and they set their coefficients to zero:

$$\frac{\delta H}{\delta N} = 0, \frac{\delta H}{\delta N^i} = 0 \implies \mathcal{H} = 0, \mathcal{H}_i = 0. \tag{19.8}$$

But this might seem a little weird– didn't we just say that the Hamiltonian was the sum of a $\mathcal{H}$ term and a $\mathcal{H}_i$ term? This seems to imply that the energy is zero. Here's the solution. The Hamilton equations are

$$\dot{h}_{ij} = \frac{\delta H}{\delta \pi^{ij}}, \dot{\pi}^{ij} = -\frac{\delta H}{\delta h_{ij}}, \tag{19.9}$$

and to compute these functional derivatives we actually need to vary for instance ${}^{(3)}R$ with respect to π^{ij}, which requires that we perform some integrations by parts. If the system were closed, there would be no boundary term and the energy would indeed be zero, but in our asymptotically flat system, a time slice is not compact and so we are not guaranteed that the boundary term is zero. As it is, this Hamiltonian doesn't give well-defined variations and so we must integrate by parts in order for our Hamiltonian to give us a physically meaningful energy.

Recall that for asymptotically flat initial data,

$$h_{ij} = \delta_{ij} + O(1/r) \implies \delta h_{ij} = O(1/r), \tag{19.10}$$

$$\delta\pi^{ij} = O(1/r^2). \tag{19.11}$$

Hence

$$N = 1 + O(1/r), N^i \to 0 \text{ as } r \to +\infty. \tag{19.12}$$

Consider the region of a constant t surface contained within a sphere of constant r with boundary S^2_r. When we vary π^{ij}, we get

$$\int_{S^2_r} dA(-2N^i h_{ik} n_j h^{-1/2} \delta\pi^{jk}) \tag{19.13}$$

where dA is the area element of S^2_r and n^j is the outward unit normal of S^2_r. Now, $dA = O(r^2)$, but $\delta\pi^{ij} = O(1/r^2)$ so this boundary term goes to zero.

Note also that the Ricci scalar is made of second derivatives of h, which means that varying with respect to h_{ij} gives two boundary terms. THus

$$\delta^{(3)}R = -R^{ij}\delta h_{ij} + D^i D^j h_{ij} - D^k D_k(h^{ij}\delta h_{ij}) \tag{19.14}$$

where D is a covariant derivative of h. Then one of the boundary terms is

$$S_1 = -\int_{S^n_r} dA\, N\Big[n^i D^j \delta h_{ij} - n^k D_k(h^{ij}\delta h_{ij})\Big]. \tag{19.15}$$

There's a second boundary term S_2 which goes to zero as $r \to +\infty$, but this term we have computed S_1 does not! Instead,

$$\lim_{r\to+\infty} S_1 = -\delta E_{ADM} \tag{19.16}$$

where

$$E_{ADM} = \lim_{r\to+\infty} \int_{S^2_r} dA n_i(\partial_j h_{ij} - \partial_i h_{jj}). \tag{19.17}$$

In general E_{ADM} will not be zero. In fact, what we ought to have started with was

$$H' = H + E_{ADM}, \tag{19.18}$$

which now poses a well-defined variational problem. Note that the indices down doesn't matter– we can just raise e.g. a pair of the i, j indices by a Kronecker delta since the integral is taken in Euclidean space.

Black hole mechanics There are three laws of black hole mechanics (*not* thermodynamics!).

Definition 19.19. A null hypersurface $\mathcal{N}$ is a *Killing horizon* if there exists a Killing vector field ξ^a defined in a neighborhood of $\mathcal{N}$ such that ξ^a is normal to $\mathcal{N}$.

Notice that nowhere in the definition of a black hole did we specify such a Killing horizon had to exist, and yet we have found one in every black hole spacetime we have studied so far.

Theorem 19.20 (Hawking 1972). *In a stationary, analytic, asymptotically flat vacuum black hole spacetime, $\mathcal{H}^+$ is a Killing horizon.*

These are quite strong conditions– explicit counterexamples exist as soon as one relaxes the asymptotic flatness condition, for example.

Notice that if $\mathcal{N}$ is a Killing horizon with respect to a Killing vector field ξ^a, then it is also a Killing horizon with respect to $c\xi$ for c a constant. We can fix the normalization by specifying that (e.g. in Kerr) if

$$\xi = K + \Omega_H m, \tag{19.21}$$

then $K^2 = -1$ as $r \to +\infty$, where K is the time translation Killing field and m is the axisymmetry Killing field. This tells us horizons rotate rigidly (Ω_H is just a constant), or else ξ would not be Killing.

Since $\xi^a\xi_a = 0$ on $\mathcal{N}$, it follows that the gradient of $\xi^a\xi_a$ is normal to $\mathcal{N}$, i.e. it is proportional to ξ_a.

$$\nabla_a(\xi^b\xi_b)|_{\mathcal{N}} = -2\kappa_0\xi_a \tag{19.22}$$

A priori, κ_0 is a function. We call it the *surface gravity*. Since ξ is Killing we can write

$$\nabla(\xi^b\xi_b) = 2\xi^b\nabla_a\xi_b = -2\xi^b\nabla_b\xi_a, \tag{19.23}$$

and thus

$$\xi^b\nabla_b\xi^a|_{\mathcal{N}} = \kappa_0\xi^a. \tag{19.24}$$

The surface gravity therefore tells us how much the integral curves of ξ^a fail to be affinely parametrized at the horizon.

Example 19.25. Consider the surface gravity of the Reissner-Nordström solution. We have the metric

$$ds^2 = -\frac{\Delta}{r^2}dv^2 + 2dvdr + r^2 d\Omega^2 \tag{19.26}$$

with $\Delta(r) = (r - r_+)(r - r_-), r_\pm = M \pm \sqrt{M^2 - Q^2}$. The stationary killing field is $K = \frac{\partial}{\partial v}$. At $r = r_\pm, \Delta = 0$, so

$$(K)_a = (dr)_a \tag{19.27}$$

at $r = r\pm$ is null. Then

$$d(K^a K_a) = d(-\frac{\Delta}{r^2}) = -\left(-\frac{\Delta'}{r^2} + \frac{2\Delta}{r^3}\right)dr, \tag{19.28}$$

and so evaluating at $r = r_\pm$, we have

$$d(K^b K_b)|_{r=r_\pm} = -\frac{(r_\pm - r_\mp)}{r_\pm^2} K|_{r=r_\pm} \implies \kappa_{0\pm} = \frac{r_\pm - r_\mp}{2r_\pm^2}. \tag{19.29}$$

What we've learned is that κ is actually not a function but just a number. One might think this comes about because of spherical symmetry, but it's actually true for other solutions like Kerr-Newman as well. Next time, we will show that the surface gravity of a black hole is always a constant and does not depend on where on the horizon you look.

Lecture 20.

Thursday, March 7, 2019

"Wald proved a version of the first law [of black hole mechanics] which is the most badass, hard version to prove." –Jorge Santos

In our last class, we talked about Killing horizons. We have seen that if $\mathcal{H}^+$ is a Killing horizon with Killing vector field ξ^a, then

$$\xi^a \nabla_a \xi^b|_{\mathcal{H}^+} = \kappa_0 \xi^b. \tag{20.1}$$

For the RN and Kerr-Newman solutions, the surface gravity κ_0 (which a priori could have been a function) is in fact a constant.

Hawking thought that perhaps we could prove that in general, the surface gravity was a constant and not dependent on where on the horizon we looked. This turns out to be true, and it leads us to the zeroth law of black hole mechanics.

Zeroth law

Proposition 20.2. *Consider a null geodesic congruence that contains the generators of a Killing horizon $\mathcal{N}$. Then $\theta = \hat{\sigma} = \hat{\omega} = 0$.*

Proof. We know that $\hat{\omega} = 0$ since our congruence is hypersurface orthogonal. Let ξ^a be a Killing field normal to $\mathcal{N}$. On $\mathcal{N}$, we can write $\xi^a = hU^a$, where U^a is affinely parametrized and h is a function on $\mathcal{N}$. Let the horizon $\mathcal{N}$ be specified by an equation $f = 0$. Then we can write $U^a = h^{-1}\xi^a + fV^a$, where V^a is a smooth vector field. This lets us extend U^a off of $\mathcal{N}$ a bit.

We can calculate the optical matrix

$$B_{ab} = \nabla_b U_a = (\nabla_a h^{-1})\xi_a + h^{-1}\nabla_b \xi_a + (\nabla_b f)V_a + f\nabla_b V_a. \tag{20.3}$$

Evaluating on $\mathcal{N}$ and using the fact that ξ is Killing,

$$B_{ab}|_{\mathcal{N}} = B_{(ab)}|_{\mathcal{N}} = \xi_{(a}\nabla_{b)}h^{-1} + V_{(a}\nabla_{b)}f|_{\mathcal{N}}, \tag{20.4}$$

where we need only consider the symmetric part since the antisymmetric part vanishes. But both $\xi_a, \nabla_a f$ are parallel to U_a on $\mathcal{N}$. Hence when we project onto $T_\perp$, we find that

$$\hat{B}_{(ab)}|_{\mathcal{N}} = P_a{}^c B_{(cd)} P_b{}^d = 0, \tag{20.5}$$

so indeed $\hat{\sigma} = 0$. ⊠

Theorem 20.6 (Zeroth law of black hole mechanics). *κ_0 is constant on the future event horizon of a stationary black hole spacetime obeying the dominant energy condition.*

Proof. Note that Hawking's theorem implies that $\mathcal{H}^+$ is Killing with respect to some Killing vector field ξ^a. We have just seen that $\theta = 0$ along the generators of $\mathcal{H}^+$. Hence $\frac{d\theta}{d\lambda} = 0$ (where λ parametrizes the integral curves of the generators). We also have $\hat{\sigma} = \hat{\omega} = 0$. Now the Raychaudhuri equation says that

$$0 = R_{ab}\xi^a\xi^b|_{\mathcal{H}^+} = 8\pi T_{ab}\xi^a\xi^b|_{\mathcal{H}^+} \tag{20.7}$$

where we've applied the Einstein equation and dropped the trace term since we're on the horizon.

This implies that

$$J \cdot \xi|_{\mathcal{H}^+} = 0 \tag{20.8}$$

where $J_a \equiv -T_{ab}\xi^b$. Now ξ^a is a future-directed causal vector, so by the dominant energy condition, J_a is also future-directed. Hence the above equation implies that J^a is parallel to ξ^a on $\mathcal{H}^+$.

Since this is the case, we can look at the expression

$$0 = \xi_{[a}J_{b]}|_{\mathcal{H}^+} = -\frac{1}{8\pi}\xi_{[a}R_{b]c}\xi^c|_{\mathcal{H}^+}. \tag{20.9}$$

On Examples Sheet 4, problem 1, we will prove that the previous expression implies the following:

$$0 = \frac{1}{8\pi}\xi_{[a}\nabla_{b]}\kappa_0|_{\mathcal{H}^+}. \tag{20.10}$$

But this means that $\nabla_a\kappa_0$ is proportional to ξ_a, i.e.

$$t \cdot \nabla\kappa_0 = 0 \tag{20.11}$$

for any tangent vector to $\mathcal{H}^+$. Hence κ_0 is constant on the horizon. ⊠

In fact, one can relax the energy condition from the DEC a bit, but it requires other details about the spacetime (asymptotic flatness and others).

First law of black hole mechanics Consider the Kerr spacetime. The horizon area A is the area of the intersection of $\mathcal{H}^+$ with a partial Cauchy surface of $t =$ constant. It is also the area of the bifurcating Killing horizon. A will be a function of J and M, and we can consider the quantity $\frac{\kappa_0}{8\pi}$. This turns out to be related to J, M, and A by

$$\frac{\kappa_0}{8\pi}\delta A = \delta M - \Omega_H \delta J. \tag{20.12}$$

This leads us to an interesting question. Start with a particular Kerr solution. Take a slice of the Kerr solution, perturb the initial conditions, and solve the perturbed initial value problem. Can we show that the area of the black hole increases in this way? This is hard to do and in fact wasn't completed until the 1990s.However, there is a more physical argument we can make about this relationship, made much earlier by Hartle and others.

Suppose we have some matter with a stress-energy tensor of $O(\epsilon)$. That is, it is small compared to the stress-energy of the black hole. Let

$$J^a = -T^a{}_bK^b, \quad L^a = T^a{}_bm^b \tag{20.13}$$

where K is the stationary Killing vector field and m is the axisymmetric Killing vector field. We would like these currents to still be conserved after we throw some stuff in. In Kerr, ∇_aJ^a is exactly zero. In the new metric, we have instead

$$\nabla_aJ^a = O(\epsilon^2), \nabla_aL^a = O(\epsilon^2) \tag{20.14}$$

That is, the metric is perturbed from Kerr by $g = \bar{g}_K + \epsilon\hat{g}$ where $\bar{g}_K$ is exactly Kerr. Hence we get an $O(\epsilon)$ correction in the covariant derivative from the perturbation to the metric, and we get another factor of ϵ from the perturbation of the stress-energy tensor. Hence to order epsilon this really is a current.

One may compute (on examples sheet 3) that

$$\delta M = -\int_{\mathcal{N}} *J, \quad \delta J = -\int_{\mathcal{N}} *L. \tag{20.15}$$

Note that the J in the first expression is the current corresponding to the Killing vector K, while the second J is the angular momentum of the black hole.

Now, one may choose Gaussian null coordinates on the horizon. In these coordinates, $\mathcal{H}^+$ is the surface $r = 0$, and the metric on $\mathcal{H}^+$ takes the form

$$ds^2_{\mathcal{H}^+} = 2drd\lambda + h_{ij}(\lambda, y)dy^idy^j. \tag{20.16}$$

Using $\sqrt{-g} = \sqrt{h}$, we can write

$$\eta = \sqrt{h} d\lambda \wedge dr \wedge dy^1 \wedge dy^2 \tag{20.17}$$

and fix an orientation (i.e. choose whether or not to put a minus sign on our definition of η as a volume element).

The orientation of $\mathcal{N}$ used in defining $\delta M, \delta J$ is the one used in Stokes's theorem. Viewing $\mathcal{N}$ as the boundary of the region with $r > 0$ (the black hole exterior, if you like), the volume element on $\mathcal{N}$ is $d\lambda \wedge dy^1 \wedge dy^2$. Hence on $\mathcal{N}$,

$$(*J)_{\lambda 12} = \sqrt{h} J^r = \sqrt{h} J_\lambda = \sqrt{h} U \cdot J \tag{20.18}$$

where $U = \frac{\partial}{\partial \lambda}$ is tangent to the generators of $\mathcal{N}$ and we've lowered the index on J using the metric since $g_{r\lambda} = 1$.

We can now evaluate everything on Kerr, since the deviation is $O(\epsilon^2)$. Hence $\mathcal{N}$ is Killing with a Killing vector $\xi = K + \Omega m$ on $\mathcal{N}$, and we also have $\xi = fU$ for a function f. Using $\xi^a \nabla_a \xi^b|_{\mathcal{H}^+} = \kappa_0 \xi^b$, we find that

$$\xi \cdot \nabla(\log|f|) = \kappa_0 \implies \frac{\partial f}{\partial \lambda} = \kappa_0 \implies f = \kappa_0 \lambda + f_0(y) \tag{20.19}$$

(where we can integrate knowing $\kappa_0 = 0$).

So far, we have

$$\xi^a = [\kappa_0 \lambda + f_0(y)] U^a, \tag{20.20}$$

where $\lambda = 0$ is the bifurcating surface. Hence we find that the integration constant vanishes, $f_0(y) = 0$, and so

$$\xi^a = \kappa_0 \lambda U^a \text{ on } \mathcal{N}. \tag{20.21}$$

From the definition of δM, we have (substituting)

$$\begin{aligned}
\delta M &= \int_{\mathcal{N}} d\lambda d^2 y \sqrt{h} T_{ab} U^a (\xi^b - \Omega_H m^b) \\
&= \int_{\mathcal{N}} d\lambda dy^2 \lambda \sqrt{h} T_{ab} U^a U^b \kappa_0 \lambda - \Omega_H \int_{\mathcal{M}} d\lambda d^2 y \sqrt{h} U \cdot L,
\end{aligned}$$

where this second term is precisely $-\delta J$. However, we have not yet used the Einstein equation– we can turn T_{ab} into a curvature condition,

$$8\pi T_{ab} U^a U^b = R_{ab} U^a U^b \tag{20.22}$$

using the fact that since U is null, the $\frac{1}{2} g_{ab} R$ term is zero. Then our expression can be written in terms of R_{ab}, so that we can use Raychaudhuri to evaluate this explicitly. Thus

$$\begin{aligned}
\delta M - \Omega_H \delta J &= \frac{\kappa_0}{8\pi} \int_{\mathcal{N}} d\lambda d^2 y \sqrt{h} \lambda R_{ab} U^a U^b \\
\implies \delta M - \Omega_H \delta J &= -\frac{\kappa_0}{8\pi} \int d^2 y \int_0^{+\infty} \sqrt{h} \lambda \frac{d\theta}{d\lambda} d\lambda \\
&= -\frac{\kappa_0}{8\pi} \int d^2 y \left\{ [\sqrt{h} \lambda \theta]_0^{+\infty} - \int_0^{+\infty} (\sqrt{h} + \underbrace{\lambda \frac{d\sqrt{h}}{d\lambda}}_{O(\epsilon^2)}) \theta d\lambda \right\},
\end{aligned}$$

where we have applied Raychaudhuri and integrated by parts.

Now recall that by the definition of the expansion θ,

$$\frac{d\sqrt{h}}{d\lambda} = \theta \sqrt{h} = O(\epsilon). \tag{20.23}$$

By assumption, the final state is stationary, so that in the $\lambda \to +\infty$ limit, $\sqrt{h}$ is finite and thus $\delta(\sqrt{h})$ is also finite. Hence

$$\int_0^{+\infty} \sqrt{h} \theta d\lambda = \int_0^{+\infty} \frac{d\sqrt{h}}{d\lambda} d\lambda = \delta(\sqrt{h}). \tag{20.24}$$

Finiteness of the RHS tells us that $\theta = o(1/\lambda)$, i.e. decays at least as fast as $1/\lambda$, and hence the boundary term vanishes. The only remaining term is

$$\int d^2y\, \delta(\sqrt{h}) = \delta \int d^2y\, \sqrt{h} = \delta A, \tag{20.25}$$

so we learn that

$$\delta M - \Omega_H \delta J = \frac{\kappa_0}{8\pi} \delta A. \quad \boxtimes \tag{20.26}$$

Lecture 21.

Friday, March 8, 2019

"So far in the course everything has been done in mathematical rigor, at least to my standards. But this is the work of the devil. It smells like sulphur. Nevertheless we are going to roll with it." –Jorge Santos

Second law of black hole mechanics

Theorem 21.1 (Hawking 1972)**.** *Let $(\mathcal{M}, g)$ be a strongly asymptotically predictable spacetime satisfying the Einstein equation with the null energy condition. Let $U \subset \mathcal{M}$ be globally hyperbolic region for which $\bar{J}^-(\mathcal{I}^+) \subset U$. Let Σ_1, Σ_2 be spacelike Cauchy surfaces for U with $\Sigma_2 \subset J^+(\Sigma_1)$ and let $H_1 = \mathcal{H}^+ \cap \Sigma_i$. Then area$(H_2) \geq$ area(H_1).*

That is, the surface area of the black hole event horizon is monotonically increasing.

Proof. We will make the additional assumption that inextendible generators of $\mathcal{H}^+$ are future complete (i.e $\mathcal{H}^+$ is non-singular). Recall that the generators for a null congruence of geodesics, and their behavior is in part governed by the expansion θ. So it might be a good start to show that $\theta \geq 0$ on $\mathcal{H}^+$.

The proof is by contradiction. Suppose $\theta < 0$ at $p \in \mathcal{H}^+$, and let γ be the (inextendible) generator of $\mathcal{H}^+$ through p and let q be slightly to the future of p along γ. But then we know that there exists a point r conjugate to p on γ using our assumption and the Raychaudhuri equation.

Recall our theorem about conjugate points, however. This tells us that we can deform γ to obtain a timelike curve from p to r, but this violates achronality of $\mathcal{H}^+$.

Let $p \in H_1$. The generators of $\mathcal{H}^+$ through p cannot leave $\mathcal{H}^+$, so they must intersect H_2 as Σ_2 is a Cauchy surface. This therefore defines a map $\phi : H_1 \to H_2$ (just follow the generators of $\mathcal{H}^+$). Now area$(H_2) \geq$ area$(\phi(H_1)) \geq$ area(H_1), where the first inequality follows because $\phi(H_1) \subset H_2$ and the second inequality follows from $\theta > 0$. $\boxtimes$

Suppose we have two black holes at some time in the past with masses M_1, M_2 and they merge to form a new black hole with mass M_3. The theorem we have just proved told us that

$$A_3 \geq A_1 + A_2 \implies M_3^2 \geq M_1^2 + M_2^2. \tag{21.2}$$

We define the efficiency of the merger process to be

$$\frac{M_1 + M_2 - M_3}{M_1 + M_2} \leq 1 - \frac{1}{\sqrt{2}}. \tag{21.3}$$

By looking at a BH merger (e.g. by LIGO), we could have checked the area theorem. We have very sensitive measurements of M_1 and M_2, but we currently have no way to measure M_3. In fact, the M_3 that was reported comes from numerical simulations (i.e. input M_1 and M_2 and simulate the merger)– with a more sensitive detector, we might have seen the quasinormal ringdown (decaying oscillations) of the post-merger black hole. Unfortunately, we weren't able to at the time, and hence Stephen Hawking didn't win a Nobel Prize.

Let us now recap the three laws of black hole mechanics.

- 0th law: κ_0 is constant.
- $dM = \frac{\kappa_0}{8\pi} dA + \Omega dJ$
- $\Delta A \geq 0$.

Notice the first law looks suspiciously like the first law of thermodynamics,

$$dE = TdS + \mu dJ \tag{21.4}$$

under the identification

$$T = \lambda\kappa_0, \quad S = \frac{A}{8\pi\lambda} \tag{21.5}$$

where λ is some undetermined number. The second law checks out here as well– $\delta A \geq 0$ tells us entropy is increasing. It's also reasonable to think that black holes are thermodynamic objects– there's entropy in the stuff we throw in, so there should still be entropy in the black hole system by the second law of thermodynamics. This argument is due to Bekenstein.

But we just proved that black holes were black– classically, nothing escapes the black hole. But thermodynamic objects must radiate. What's the solution?

It is the following formula:

$$T = \frac{\hbar\kappa_0}{2\pi}, \tag{21.6}$$

taking $\lambda = \hbar/2\pi$. We see now that this is something quantum. In the classical limit as $\hbar \to 0$, the entropy becomes infinite.

Quantization of the free scalar field Let $(\mathcal{M}, g)$ be a globally hyperbolic spacetime with metric

$$ds^2 = -N^2 dt^2 + h_{ij}(dx^i + N^i dt)(dx^j + n^j dt). \tag{21.7}$$

Let Σ be a Cauchy surface with normal $n_a = -N(dt)_a$. The metric on Σ_t is simply h_{ij}, with $\sqrt{-g} = N\sqrt{h}$. Now we can write down the action for a free scalar field,

$$S = \int_{\mathcal{M}} dt d^3x \sqrt{-g}\left[-\frac{1}{2}\nabla^a\Phi\nabla_a\Phi - \frac{1}{2}m^2\Phi^2\right]. \tag{21.8}$$

We compute the conjugate momentum to the field; it is

$$\pi(x) = \frac{\delta S}{\delta(\partial_t\Phi)} = \sqrt{h}n^\mu\nabla_\mu\Phi. \tag{21.9}$$

Next, to quantize we promote Φ, π to operators (in $\hbar = 1$ units) such that they satisfy equal-time commutation relations

$$[\Phi(t,x), \pi(t,x')] = i\delta^{(3)}(x - x'). \tag{21.10}$$

We now introduce the Hilbert space where the field lives. Let S be the space of complex solutions to the Klein-Gordon equation ($\Box\Phi = m^2\Phi$, where $\Box = \nabla_a\nabla^a$). Global hyperbolicity implies that a point in S is specified uniquely by the initial conditions $\Phi, \partial_t\Phi$ on Σ_0. Note that the Hilbert space comes equipped with an inner product:

$$(\alpha, \beta) = -\int_{\Sigma_0} d^3x\sqrt{h}n_a j^a(\alpha,\beta), \quad j(\alpha,\beta) \equiv -i(\bar{\alpha}d\beta - \beta d\bar{\alpha}). \tag{21.11}$$

In fact, one now finds that this current j we have defined is actually conserved, i.e. from Klein-Gordon it follows that

$$\nabla^a J_a = 0. \tag{21.12}$$

This implies our norm is actually independent of t– while it may be convenient to calculate on Σ_0, (α, β) does not depend on t.

Note that

$$(\alpha, \beta) = \overline{(\beta, \alpha)}, \tag{21.13}$$

which implies that $(\cdot, \cdot)$ is a Hermitian form. It is non-degenerate, i.e. if $(\alpha, \beta) = 0$ for all $\beta \in S$, then $\alpha = 0$. However,

$$(\alpha, \beta) = -(\bar{\beta}, \bar{\alpha}), \text{ so } (\alpha, \alpha) = -(\bar{\alpha}, \bar{\alpha}), \tag{21.14}$$

which tells us our inner product is not positive-definite. In Minkowski space, the inner product $(,)$ is positive-definite on the subpsace S_p of S consisting of positive frequency solutions. A basis for S_p is the set of plane waves,

$$\psi_{\mathbf{p}}(x) = \frac{1}{(2\pi)^{3/2}(2p^0)^{1/2}}e^{ip\cdot x}, \quad p^0 = \sqrt{\mathbf{p}^2 + m^2}. \tag{21.15}$$

These modes (i.e. solutions of the K-G equation) are positive frequency in the sense that if we take $K = \frac{\partial}{\partial t}$ to be our stationary Killing field, then these modes have negative imaginary eigenvalues with respect to the Lie derivative $\mathcal{L}_K$. That is,

$$\mathcal{L}_K \psi_{\mathbf{p}} = -ip^0 \psi_{\mathbf{p}} \tag{21.16}$$

Now observe that the complex conjugate of $\psi_{\mathbf{p}}$ is a negative frequency plane wave. By linearity, we can decompose

$$S = S_p + \bar{S}_p. \tag{21.17}$$

However, note that in curved spacetime, we do not have a definition of "positive frequency" except if the spacetime is stationary (in which case the Killing field provides us a "time" direction). Instead, we simply choose a subspace S_p for which $(\cdot,\cdot)$ is positive-definite and S can be decomposed $S = S_p \oplus \bar{S}_p$. In general there are many ways to do this, and moreover physics may look very different depending on how we decompose the space.

In the QFT, we define the creation and annihilation operators to be modes $f \in S_p$ of a real scalar field ($\Phi^\dagger = \Phi$) by

$$a(f) = (f, \Phi), \quad a^\dagger(f) = -(\bar{f}, \Phi). \tag{21.18}$$

These imply that

$$[a(f), a(g)^\dagger] = (f, g) \tag{21.19}$$

which in Minkowski space reduces to

$$[a_{\mathbf{p}}, a^\dagger_{\mathbf{q}}] = \delta^{(3)}(\mathbf{p} - \mathbf{q}). \tag{21.20}$$

The last ingredient we need is a vacuum state. Once we have a vacuum state, we can say that the Hilbert space is just the Fock space we get by acting on the vacuum repeatedly with the one-particle operators. We therefore define the (normalized) vacuum state $|0\rangle$ by the condition

$$a(f)|0\rangle = 0 \,\forall f \in S_p, \quad \langle 0|0\rangle = 1. \tag{21.21}$$

However, suppose we chose a different positive frequency space, S'_p. The for any $f' \in S'_p$, we can decompose this function in the original basis, $f' = f + \bar{g}$ for $f \in S_p, \bar{g} \in \bar{S}_p$. Now the annihilation operator of our f' is

$$a(f') = (f, \Phi) + (\bar{g}, \Phi) = a(f) - a(g)^\dagger. \tag{21.22}$$

And hence what we would have called a positive frequency solution relative to S_p now gives us an $a(g)^\dagger$, which creates a particle relative to the original vacuum. We see that something as innocuous-sounding as a decomposition into positive and negative frequency solutions means that we won't agree even on the number of particles in the system.

Lecture 22.

Monday, March 11, 2019

> *"Here is a form where you can complain about me. You can complain about me any day, but today is the day I will hear your complaints." –Jorge Santos*

Last time, we said that a free real massive scalar field in a curved spacetime obeys a Klein-Gordon equation. In particular, we defined a space of solutions S, and we found that because the norm of states under the inner product

$$(\alpha, \beta) = -\int_\Sigma d^3x \sqrt{h} n_a j^a(\alpha, \beta) \tag{22.1}$$

is not positive-definite, we do not have a unique way to decompose S into positive and negative frequency states. In fact, we showed that

$$a(f') = a(f) - a(g)^\dagger, \tag{22.2}$$

so observers moving relative to one another will disagree on the vacuum.

However, there are no such problems in Minkowski space. In a stationary spacetime, one can use time translation symmetry to identify a preferred choice of S_p (the space of positive frequency states). Let K^a be the (future-directed) time translation Killing Vector field of Minkowski space. It can be shown that the Lie derivative $\mathcal{L}_K$ commutes with the Klein-Gordon operator ($\Box - m^2$) and therefore maps the space of solutions onto itself, $S \to S$.

It can be shown (on the examples sheets) that $\mathcal{L}_K$ is anti-Hermitian and hence has purely imaginary eigenvalues.

Definition 22.3. We say that an eigenfunction has positive frequency if its eigenvalue under $\mathcal{L}_K$ is negative imaginary,

$$\mathcal{L}_K u = -i\omega u, \omega > 0. \tag{22.4}$$

Because $\mathcal{L}_K$ is anti-Hermitian, this gives us a preferred decomposition of $S = S_p \oplus \bar{S}_p$ into positive and negative frequency solutions.

Bogoliubov transformations Let $\{\psi_i\}$ be an orthonormal[34] basis for S_p. That is,

$$(\psi_i, \psi_j) = \delta_{ij}, \quad (\psi_i, \bar{\psi}_j) = 0. \tag{22.5}$$

We can define creation and annihilation operators $a_i, a_i^\dagger$ such that

$$\Phi = \sum_i (a_i \psi_i + a_i^\dagger \bar{\psi}_i), \tag{22.6}$$

which is like a Fourier decomposition of a general $\Phi \in S$, with $a_i \equiv (\Phi, \psi_i)$. In such a basis,

$$[a_i, a_j^\dagger] = \delta_{ij}, \quad [a_i, a_j] = 0. \tag{22.7}$$

Let S'_p be a different choice for the positive frequency subspace, with its own (orthonormal) basis $\{\psi'_i\}$. Since we are still decomposing the same space of solutions, there had better be a linear transformation relating this decomposition and the other one:

$$\psi'_i = \sum_j (A_{ij}\psi_j + B_{ij}\bar{\psi}_j), \tag{22.8}$$

with $\bar{\psi}'_i$ similar. Note that the ψ'_is must still satisfy orthogonality relations,

$$(\psi'_i, \psi'_j) = \delta_{ij}, \quad (\psi'_i, \bar{\psi}'_j) = 0, \tag{22.9}$$

and so we get some conditions on A_{ij} and B_{ij}, known as the Bogoliubov coefficients. One finds that

$$AA^\dagger - BB^\dagger = 1, \quad AB^T - BA^T = 0. \tag{22.10}$$

Why are these transformations interesting? Consider some spacetime which we divide into three regions by time: $(\mathcal{M}_-, g_-), (\mathcal{M}_0, g_0), (\mathcal{M}_+, g_+)$. Here, we have a preferred time direction, and we want a globally hyperbolic spacetime $(\mathcal{M}, g)$ where

$$\mathcal{M} = \mathcal{M}_- \cup \mathcal{M}_0 \cup \mathcal{M}_+. \tag{22.11}$$

In particular, the first and last regions are stationary, while the middle one is dynamic. We imagine e.g. a black hole in the past that eats some matter and then settles down to a steady state. In the spacetime $(\mathcal{M}_\pm, g)$, stationarity implies that there is a preferred choice of positive frequency modes $S_p^\pm$ in both $\mathcal{M}_+, \mathcal{M}_-$.

Hence we have two choices of positive frequency subspaces for $(\mathcal{M}, g)$ given by S_p^+, S_p^-. Notice that these work for the *entire* manifold $\mathcal{M}$ because of global hyperbolicity. Let $\{U_i^\pm\}$ denote an orthonormal basis for $S_p^\pm$ and let $a_i^\pm$ be the annihilation operators

$$u_i^+ = \sum_j (A_{ij} u_j^- + B_{ij} \bar{u}_j^-), \tag{22.12}$$

with

$$a_i^+ = \sum_j (\bar{A}_{ij} a_j^- - \bar{B}_{ij} a_j^{-\dagger}). \tag{22.13}$$

That is, we can relate the subspace decompositions $S_p^\pm$ by a Bogoliubov transformation.

Now denote the respective vacua by $|0\pm\rangle$, such that

$$a_i^\pm |0\pm\rangle = 0 \quad \forall i. \tag{22.14}$$

[34] We're writing the indices like they're discrete, but really they could be some sort of continuum.

In the past, start with the vacuum $|0-\rangle$. The particle number for the ith late time mode is

$$N_i^+ = a_i^{+\dagger} a_i^+. \tag{22.15}$$

Thus the particle number seen in the $+$ region is

$$\langle 0-|N_i^\dagger|0-\rangle = \sum_{j,k} \langle 0-|a_k^-(-B_{ij})(-\bar{B}_{ij})a_j^{-\dagger}|0-\rangle = (BB^\dagger)_{ii}. \tag{22.16}$$

Hence the expected total number of particles is

$$\sum_i (BB^\dagger)_{ii} = \text{Tr}(B^\dagger B), \tag{22.17}$$

so if $B \neq 0$, we will have particle production. If $B = 0$, then $S_p^+ = S_p^-$ and the decompositions agree. But this is generally not true.

To really understand the consequences of this, we want to know these Bogoliubov coefficients for the Schwarzschild black hole. That is, we will solve the wave equation around the black hole.

Wave equation in Schwarzschild spacetime Notice that the Schwarzschild spacetime has spherical symmetry. Thus we might like to decompose a massless Klein-Gordon field Φ into spherical harmonics,

$$\Phi = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \frac{1}{r} \phi_{lm}(t,r) Y_{lm}(\theta,\phi). \tag{22.18}$$

The wave equation $\Box\Phi = 0$ then reduces to

$$\left[\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial r_*^2} + V_l(r_*)\right]\phi_{lm}(t,r_*) = 0 \tag{22.19}$$

where

$$V_l(r_*) = \left[1 - \frac{2M}{r(r_*)}\right]\left[\frac{l(l+1)}{r(r_*)^2} + \frac{2M}{r(r_*)^3}\right]. \tag{22.20}$$

Notice that when $r_* \to -\infty, r \to 2M$ (the near-horizon limit). Moreover, $V_l(r_*)$ vanishes both as $r_* \to +\infty$and $r_* \to -\infty$.

However, note that if we prepare a wavepacket at the maximum of the potential, it will generically split into two wavepackets under time evolution. One packet will fall in towards the horizon, and the other will go out towards $\mathcal{I}^+$.

At late time $t \to +\infty$ we therefore expect the solution to consist of a superposition of two wavepackets propagating to the "left" ($r_* \to -\infty$) and to the "right" ($r_* \to +\infty$) and hence in the limit $t \to \pm\infty$ we expect a decomposition of the form

$$\phi_{lm}(t,r_*) = f_\pm(t - r_*) + g_\pm(t + r_*) = f_\pm(u) + g_\pm(v), \tag{22.21}$$

ingoing and outgoing wavepackets. $f_+(u)$ is an outgoing wavepacket propagating to $\mathcal{I}^+$, while $g_+(v)$ describes an ingoing wavepacket propagating to $\mathcal{H}^+$. Hence the solution is uniquely determined by specifying its behavior on $\mathcal{I}^+ \cup \mathcal{H}^+$.

We will define an "out" mode to be a solution which vanishes on the horizon $\mathcal{H}^+$ and a "down" mode to be a solution which vanishes on $\mathcal{I}^+$. Similarly, at early times, we can evaluate the integral defining the Klein-Gordon inner product to find the "in" mode which vanishes on $\mathcal{H}^-$ and the "up" mode which vanishes on $\mathcal{I}^-$.

Lecture 23.

Wednesday, March 13, 2019

"This all starts with very obvious statements. But they build up, and complicate towards the end."
–Jorge Santos

Our previous class was entirely classical. We studied waves on curved spacetime, defining an important set of modes– the out and down modes, defined in the future (on $\mathcal{I}^+, \mathcal{H}^+$) and the in and up modes, defined in the past (on $\mathcal{I}^-$ and $\mathcal{H}^-$). Note also that these modes have nothing to do with quasinormal modes, which you will come across if you continue to study general relativity.[35]

Now, the Krukal spacetime is stationary, since it admits a Killing vector $K \equiv \frac{\partial}{\partial t}$. Hence

$$\Phi_{lm}(t, r_*) = e^{-i\omega t} R_{\omega l m}(r) \tag{23.1}$$

are eigenfunctions of $\mathcal{L}_K$, and for $\omega > 0$ they have positive norm.

Input this into the Klein-Gordon equation $\Box\Phi = 0$, and we get a time-independent Schrödinger equation:

$$\left[-\frac{d^2}{dr_*^2} + V_l(r_*) \right] R_{\omega l m} = \omega^2 R_{\omega l m}. \tag{23.2}$$

We impose the boundary condition that these modes vanish either on the future event horizon $\mathcal{H}^+$ or at future null infinity, $\mathcal{I}^+$ to get e.g. the out and down modes.

We emphasize that these are different from quasinormal modes, which obey infalling and outgoing boundary conditions and describe the ringing of a black hole after perturbation.

Hawking radiation Let us now venture into the quantum consequences of the calculation we have done. A priori, it's not surprising that we have particle production when the spacetime is time-dependent. What is surprising is that particle production is not a transient behavior but persists indefinitely into the future (up to the evaporation of the BH).

We will now introduce a basis for the spacetime describing a collapsing star, analogous to the modes of Kruskal. In the past, there are no "up" modes since there is no past event horizon. But we do have "in" modes!

Notice the geometry near $\mathcal{I}^-$ is static, so there is a natural notion of positive frequency there. Let $\{f_i\}$ be a basis of the "in" modes. After the black hole forms, we must deal with the out and down modes.

At late times, we can define "out" and "down" modes as before. The geometry near $\mathcal{I}^+$ is also static, so we can define a notion of positive frequency there. Call a basis for such modes $\{p_i\}$. The geometry is *not* static everywhere on $\mathcal{H}^+$, however, so there is no natural notion of positive frequency modes. We can nevertheless pick there a basis $\{q_i\}$.

First we impose orthonormality:

$$(p_i, p_j) = (q_i, q_j) = \delta_{ij}, \quad (f_i, f_j) = \delta_{ij}. \tag{23.3}$$

Moreover,

$$(p_i, q_j) = 0 \tag{23.4}$$

since they have mutually exclusive supports.

Let a_i, b_i be annihilation operators for the "in" and "out" modes respectively,

$$a_i = (f_i, \Phi), b_i(p_i, \Phi). \tag{23.5}$$

Using a Bogoliubov transformation, we can expand

$$p_i = \sum_j (A_{ij} f_j + B_{ij} \bar{f}_j) \tag{23.6}$$

so that

$$b_i = (p_i, \Phi) = \sum_j \left(\bar{A}_{ij} a_j - \bar{B}_{ij} a_j^\dagger \right). \tag{23.7}$$

In the past, we are in the vacuum state $a_i|0\rangle = 0$. But we showed that the number of expected particles in the future is related to the Bogoliubov coefficients. The expected number of particels present in the ith "out" mode is then

$$\langle 0 | b_i^\dagger b_i | 0 \rangle = (BB^\dagger)_{ii}. \tag{23.8}$$

[35]If you want to learn about quasinormal modes, come to my talk on them on March 14, 2019! (Pi Day!) In MR2, 1:45 PM. For most readers this will be impossible without closed timelike curves or the like. But I will be there at that point in spacetime.

We will choose our "out" basis p_i so that at $\mathcal{I}^+$ they are wavepackets localized around some particular retarded time u_i, and containing only positive frequencies localized around some value ω_i. What you should have in mind is a cosine modulated by a Gaussian, $e^{-(u-u_i)^2}\cos(\omega_i u)$.

Consider first Kruskal. Imagine propagating a wavepacket back in time from $\mathcal{I}^+$. Part of the wavepacket could be 'reflected" to give a wavepacket on $\mathcal{I}^-$ (which we call $p_i^{(1)}$), and part of it will be "transmitted" and reach $\mathcal{H}^-$ (as $p_i^{(2)}$). More generally we could study a wavepacket from $\mathcal{I}^+ \cup \mathcal{H}^+$.

We want to look at the case of a wavepacket that is localized around a late retarded time. We can also decompose the solutions that end up on $\mathcal{I}^-$ as the sum of modes:

$$p_i = p_i^{(1)} + p_i^{(2)}, \tag{23.9}$$

where $p_i^{(1)}$ is simply "reflected" off the Schwarzschild geometry, while $p_i^{(2)}$ passes through the time-dependent region before reaching $\mathcal{I}^-$. It must be that B (the Bogoliubov coefficients) is related to $p_i^{(2)}$, the part of p_i that sees the time dependence.

Let us define transmission and reflection coefficients

$$T_i = \sqrt{(p_i^{(2)}, p_i^{(2)}}, \quad R_i = \sqrt{(p_i^{(1)}, p_i^{(1)}} \tag{23.10}$$

such that

$$R_i^2 + T_i^2 = 1. \tag{23.11}$$

From what we have seen,

$$A_{ij} = A_{ij}^{(1)} + A_{ij}^{(2)}, \quad B_{ij} = B_{ij}^{(2)}. \tag{23.12}$$

We want to compute B_{ij}, which means we have to find the behavior of $p_i^{(2)}$ on $\mathcal{I}^-$.

Notice that with the form of p we have specified on $\mathcal{I}^+$, as we approach the horizon H^+, $u \to +\infty$, and our wavepacket oscillates infinitely. Because there are infinitely many oscillations, we can apply the WKB approximation (i.e. we study a quickly varying solution on a slowly varying background potential). Hence

$$\Phi(x) = A(x)e^{i\lambda S(x)}, \lambda \gg 1. \tag{23.13}$$

Applying the Klein-Gordon equation gives

$$\Box\Phi = 0 \implies \nabla^a S \nabla_a S = 0. \tag{23.14}$$

Surfaces of constant phase S are null hypersurfaces. The generators of these hypersurfaces are null geodesics.

In our next class, we will show that since surfaces of constant S have null geodesics, we can solve for these null geodesics and derive a great consequence.

Lecture 24.

Thursday, March 14, 2019 (The End)

"Look! This has G, ħ, and c. That's all the fundamental constants. This is the one statement about quantum gravity we know is true." –Jorge Santos

Today, we conclude the calculation of Hawking radiation. Using WKB, we took our modes to have the form $\Phi = A(x)e^{i\lambda S(x)}, \lambda \gg 1$, and applying $\Box\Phi = 0$ we have $\nabla^a S \nabla_a S = 0$ to leading order in λ. Surfaces of constant phase S are null hypersurfaces.

Now, we can introduce a null vector N which is parallel propagated along the generators U

$$U^b \nabla_b N^a = 0, \quad N \cdot U = -1. \tag{24.1}$$

We can import some of our results from Kruskal and take

$$U^a = \left(\frac{\partial}{\partial V}\right)^a, \quad N = c\frac{\partial}{\partial U} \tag{24.2}$$

with some constant c. Hence any deviation vector for this congruence can be decomposed into the sum of a part orthogonal U^a and a term βN^a which is parallel transported along the geodesics.

Choose $\beta = -\epsilon$ where $\epsilon > 0$ and small. A nearby geodesic γ' then has tangent

$$U = -C\epsilon \implies u = -\frac{1}{\kappa_0}\log(-U) = -\frac{1}{\kappa_0}(C\epsilon). \tag{24.3}$$

Let $F(u)$ denote the phase of our wavepacket p_i on $\mathcal{I}^+$. Then the phase everywhere on γ' must be

$$S = F(-\frac{1}{\kappa_0}\log(C\epsilon)), \tag{24.4}$$

so we know what the value of the phase is in the region $v < 0$. We want to translate this into a statement in terms of v.

In fact, we've already seen the technology to do this calculation and find ϵ in terms of v. This is because on $\mathcal{I}^-$, we have another null hypersurface and we can therefore apply the same reasoning as we did near $\mathcal{I}^+$. Recall that the metric on $\mathcal{I}^-$ is

$$ds^2 = -dudv + \frac{1}{4}(u-v)^2 d\Omega_2^2, \tag{24.5}$$

so we again introduce N as before,

$$N = D^{-1}\frac{\partial}{\partial v} \tag{24.6}$$

at $\mathcal{I}^-$. We conclude that

$$v = -D^{-1}\epsilon \implies \epsilon = -Dv. \tag{24.7}$$

Hence the phase on $\mathcal{I}^-$ in terms of v is

$$S = F\left[-\frac{1}{\kappa_0}\log(-CDv)\right]. \tag{24.8}$$

We have

$$p_i^{(2)} \approx \begin{cases} 0, & v > 0 \\ A_i(v)\exp\left[iF\left(-\frac{1}{\kappa_0}\log(-D(v))\right)\right], & v < 0. \end{cases} \tag{24.9}$$

Notice that most of the wavepacket is very close to $v = 0$, where the argument of the log becomes very large.

We'll now cheat "very badly" by converting everything to plane waves. This is not really valid– some of the integrals we will write down diverge, breaking some of our original assumptions. For plane waves,

$$F(u) = -\omega_i u. \tag{24.10}$$

This leads to p_i that are neither normalizable nor localized at late times. This is more of an illustrative calculation– the integrals for e.g. the modulated cosine can be computed, but they are hard.

$$p_\omega \simeq \begin{cases} 0, & v > 0 \\ A_\omega(v)\exp\left[i\frac{\omega}{\kappa_0}\log(-CDv)\right], & v < 0. \end{cases} \tag{24.11}$$

Similarly, we will use a basis of "in" modes f_σ such that v_σ has frequency $\sigma > 0$. THus

$$f_\sigma = \frac{1}{(2\pi N_\sigma)}e^{-i\sigma v}. \tag{24.12}$$

These are just related by a Fourier transform:

$$\tilde{p}_\omega^{(2)}(\sigma) = A\omega\int_{-\infty}^0 dv e^{i\sigma v}\exp\left[i\frac{\omega}{\kappa_0}\log(-CDv)\right]. \tag{24.13}$$

with inverse transform

$$\begin{aligned} p_\omega^{(2)}(v) &= \int_{-\infty}^{\infty}\frac{d\sigma}{2\pi}\tilde{p}_\omega^{(2)}(v)e^{-i\sigma v} \\ &= \int_0^{+\infty} d\sigma N_\sigma \tilde{p}_\omega^{(2)} f_\sigma(v) + \int_0^{+\infty} d\sigma \bar{N}_\sigma \tilde{p}_\omega^{(2)}(v)\bar{f}_\sigma(v). \end{aligned}$$

But we see that A is just related to the positive frequency ($f(v)$) components and B is related to the negative frequency components, hence

$$A^{(2)}\omega_\sigma = N_\sigma \tilde{p}^{(2)}_\omega(\sigma), \quad B\omega_\sigma = \bar{N}_\sigma \tilde{p}^{(2)}_\omega(-\sigma), \quad \omega, v > 0. \tag{24.14}$$

We now have to choose where the branch cut lies for the log in the complex plane. Define

$$\log z = \log |z| + i\text{ang} z, \quad \text{ang} \in (-\pi/2, 3\pi/2). \tag{24.15}$$

Here's the first integral that diverges. We're integrating over $-\infty$ to 0 in v, and we close the contour by adding a contour over $0 \to \infty$ and close the contour in the lower half-plane. The log is otherwise analytic, so we assume that I_1 the integral $-\infty \to 0$ is just $-I_2$ the integral from $0 \to \infty$. Hence

$$\begin{aligned} \tilde{p}^{(2)}_\omega(-\sigma) &= -A\omega \int_0^{+\infty} dve^{-\sigma v} \exp\left[i\frac{\omega}{\kappa_0}\log(-CDv)\right] \\ &= -A\omega \int_0^{+\infty} dve^{-i\sigma v} \exp\left\{i\frac{\omega}{\kappa_0}[\log(CDv)]\right\} e^{2\omega\pi/\kappa_0} = -e^{-\omega\pi/\kappa_0} p^{(2)}_\omega(v). \end{aligned}$$

We find that the Bs and As are related:

$$|B\omega\sigma| = e^{-\omega\pi/\kappa_0}|A^{(2)}_{\omega\sigma}|. \tag{24.16}$$

If we were braver, we might have repeated this calculation with the actual wavepackets and found that

$$|B_{ij}| = e^{-\omega_i\pi/\kappa_0}|A^{(2)}_{ij}|. \tag{24.17}$$

Then

$$T_i^2 = (p_i^{(2)}, p_i^{(2)}) = \sum_j (|A^{(2)}_{ij}|^2 - |B_{ij}|^2) \implies (BB^\dagger)_{ii} = \frac{T_i^2}{e^{2\omega_i\pi/\kappa_0} - 1}. \tag{24.18}$$

This gives (almost) blackbody radiation at the *Hawking temperature*

$$T_H = \frac{\kappa_0}{2\pi}. \tag{24.19}$$

For a solar mass black hole, we find that

$$T_H = 6 \times 10^{-8} \frac{M.}{M} \text{ Kelvin}. \tag{24.20}$$

This can be generalized to the fermionic case, the massive case, and the gravitational case, amongst others. This is remarkable.

Black hole thermodynamics Now that we have a concept of temperature, we can define an entropy! Putting constants back in,

$$S_{BH} = \frac{c^3 A}{4G\hbar}. \tag{24.21}$$

This is an incredible result. Any theory of quantum gravity must reproduce the entropy formula.

Notice that for a solar mass black hole, $S_{BH} \sim 10^{77}$, whereas $S_\odot \sim 10^{58}$. Why isn't everything in the universe just black holes? It's because the initial state of the universe was very special, as postulated by Roger Penrose.

Moreover, if we calculate the energy change from Hawking radiation, we find that

$$\frac{dE}{dt} \approx -\alpha A T^4. \tag{24.22}$$

This leads us to a generalized second law– because of evaporation, it's not the black hole entropy alone which is nondecreasing but the entropy of the BH-universe system.