Guide to and brief synopsis of 8.962 lectures, Spring 2020

Especially now that we have transitioned to an online-only "asynchronous" mode of instruction, it is important everyone have a good idea what is contained in each set of posted lecture notes and in the recorded lecture videos. The hand-written notes and recorded videos are, of necessity, heavy on calculational detail. The goal of this guide is to provide a brief narrative synopsis of each lecture as an accompaniment to the posted material.

Lecture 1: Introduction; the geometric viewpoint on physics. Review of Lorentz transformations and Lorentz-invariant intervals, which leads to the definition of the displacement "4-vector" $\Delta\vec{x}$ between two events. Definition of 4-vector as a set that transforms between inertial reference frames in same way as the components of $\Delta\vec{x}$. Introduction of component notation: $\vec{A} \doteq (A^0, A^1, A^2, A^3)$, denoted by A^α. Definition of basis vectors; definition of the inner product between two four vectors; using the inner product to define the metric tensor $\eta_{\alpha\beta}$.

Lecture 2: The notion of "*coordinate*" bases: basis objects such that the displacement $\Delta\vec{x}$ is built only from basis vectors and coordinate differentials, i.e. such that $\Delta\vec{x} = \Delta x^\alpha \vec{e}_\alpha$. Several important 4-vectors for physics: 4-velocity, 4-momentum, 4-acceleration, and their properties. Definition of (M, N) tensors: a linear mapping between N 4-vectors, M 1-forms and the frame-independent real numbers. Using the metric and its inverse to raise and lower tensor indices. Considerations on derivatives of tensor fields – simple at this point, since we are working in special relativity and only considering rectilinear coordinates. Number flux 4-vector; its use in defining a conservation law (both differential and integral forms).

Lecture 3: More on tensors, derivatives, 1-forms. Contraction of tensor indices; the dual nature of vectors and the associated 1-form found by lowering the vector index.

Lecture 4: Volumes and volume elements, covariant construction using the Levi-Civita tensor[1]. How to go between differential and integral formulations of conservation laws. Electrodynamics in geometric language (4-current, Faraday field tensor). Introduction of the stress-energy tensor, with the *perfect fluid* stress-energy tensor presented as a particularly important example.

Lecture 5: More on the stress-energy tensor: symmetry, physical meaning of its components in a given representation. Differential formulation of conservation of energy and conservation of momentum. Prelude to curvature: special relativity and tensor analyses in curvilinear coordinates. The distinction between coordinate basis and other basis is particularly important here; for example, we see that $\vec{e}_\phi$ must have the dimensions of length in order that $\Delta\phi\,\vec{e}_\phi$ make a dimensionally sensible contribution to $\Delta\vec{x}$. The Christoffel symbol: the quantity (not a tensor!) which relates derivatives of basis objects to basis objects. Introduction to the covariant derivative: how to make a derivative whose components transform like tensor components. By appealing to the principle of equivalence ("I can find a local representation in which spacetime looks flat"), we require the metric to have zero covariant derivative. This leads to a simple rule for building the Christoffel symbol from derivatives of the metric.

Lectures 6 and 7: Introduction to the principle of equivalence, in particular the use of *freely falling frames* as our generalization of the inertial frames that play an important role in special relativity. The physical meaning of this arises from the fact that gravity couples to the same mass m that determines an object's inertia in $\mathbf{F} = m\mathbf{a}$. Hence, in a freely falling frame, all objects experience the same $\mathbf{a}$; it is effects *relative to* $\mathbf{a}$ that are of fundamental physical interest. Several variants of the equivalence principle (EP) exist: The *weak* EP tells us that one cannot distinguish free fall under gravity from uniform acceleration, at least over "sufficiently small" regions (meaning small enough that tides can be neglected); the *Einstein* EP tells us that over a sufficiently small region, the laws of physics in freely falling frames are identical to those in special relativity. (The *strong* EP also exists, but won't be discussed much in 8.962; it essentially tells us that gravitational energy falls in gravitational fields just like any other kind of energy. This is most important in analyzing the motion of bodies that are very strongly gravitationally bound, like neutron stars and black holes, for which a substantial fraction or even the majority of its mass/energy content is gravitational.)

Lecture 7: Additional material in Lecture 7 demonstrates that a general coordinate transformation has enough functional freedom to make the spacetime metric look flat at a particular point, up to quadratic corrections. In other words, there exist coordinate transformations such that $g_{\mu\nu} \to \eta_{\mu\nu} + (\partial^2 g)(\delta x)^2$, where $\partial^2 g$ schematically indicates two derivatives of the metric. Worth noting: When we clear out the metric, there

[1] A tensor in special relativity formulated in inertial coordinates, *not* a tensor in more general spacetimes.

are 6 leftover degrees of freedom. These correspond to the freedom to set 3 boosts and 3 rotations. When we clear out the first derivative of the metric, the number of degrees of freedom in the transformation is exactly the right number to satisfy the constraints imposed by the coordinate transformation. At second order, we cannot satisfy all of the constraints needed to flatten the spacetime: we would need an additional 20 degrees of freedom in general to flatten spacetime at this order. We will later see that, in 4 dimensional spacetime, the tensor which describes spacetime curvature has 20 independent components, exactly corresponding to the 20 constraints which cannot be eliminated by a coordinate transformation at this order.

Lecture 7 also discusses the need for a law of *transport* to connect two points in order to define a useful notion of derivative for tensor fields on a curved manifold. Fundamentally, this arises because basis objects "live" in the tangent space to points on a manifold. When we compare fields at two different points, we need to account for the fact that the bases are different at these points. Transport method 1 introduces a set of *connection coefficients* which account for how a vector field A^α is transported over a displacement δx^β. If we require that the metric have zero derivative under this transport law, then the connection coefficients are in fact the Christoffel symbols we found earlier, and the derivative that emerges from this transport analysis is the covariant derivative. This mechanism is then known as "parallel transport," since it amounts to holding the components of a vector constant in the freely falling or locally Lorentz reference frame. Generalization to more complicated tensors than 4-vectors is straightforward, and amounts to including a connection coefficient / Christoffel symbol for each tensor index.

Lecture 8: This lecture discusses a second transport method, *Lie transport*, which leads to the Lie derivative. This is a form of derivative that shows up a lot in discussions of fluid flow. For us, its most important application is when the Lie derivative of the metric along a vector $\vec{\xi}$ is zero, which shows that the direction $\vec{\xi}$ is associated with a symmetry of the spacetime. Such a vector field is called a *Killing vector*; the Lie derivative of the metric along $\vec{\xi}$ can be turned into a relation called Killing's equation, $\nabla_{(\alpha}\xi_{\beta)} = 0$. Killing vectors play an important role later in the course identifying quantities that are constant for bodies moving in a spacetime.

Lecture 8 also discusses the notion of *tensor densities*: quantities with transformation laws similar to a tensor, but with a slight modification: they involve a power of the determinant of the metric. For our purposes, the two most important tensor densities are the determinant of the metric itself, and the Levi-Civita symbol. By combining these quantities, we can make a properly tensorial quantity which measures volumes: the tensor $\epsilon_{\alpha\beta\gamma\delta} = \sqrt{|g|}\tilde{\epsilon}_{\alpha\beta\gamma\delta}$ does exactly this (with $\tilde{\epsilon}$ marking the Levi-Civita symbol we used to compute volumes in special relativity); the absolute value insures that we don't take the square root of a negative quantity. We also derive some useful identities ("party tricks") based on the determinant of the metric which come in handy for some important calculations.

Lecture 9: Here we begin discussing the kinematics of bodies that move through spacetime. We consider the motion of a highly idealized *test mass*: a body with no charge, no spatial extent, no spin − just a pure point mass. The only "force" which such a body can experience is gravity, which means that in a freely falling frame, its motion is purely inertial: in the freely falling frame, it moves in a "straight line," so $x^\alpha(\tau) = x^\alpha(0) + u^\alpha\tau$, where $x^\alpha(\tau)$ denotes the sequence of events through which it moves as a function of its own proper time τ, and u^α is its 4-velocity.

This representation of its motion only holds in the freely falling frame. A frame-independent formulation is to say that the body parallel transports the tangent to its worldline along its worldline, or $u^\alpha\nabla_\alpha u^\beta = 0$. Such motion is called a *geodesic* of the spacetime. In the timelike case, one can show that a geodesic extremizes (in this case, maximizes) the proper time that accumulates between all timelike trajectories between two events. Replacing the 4-velocity u^α with 4-momentum p^α, it is simple to reformulate the geodesic equation for null or lightlike trajectories.

If the spacetime is independent of a particular coordinate x^a (where a is some particular choice of index), then it can be shown that p_a is a constant along the geodesic. For example, if the metric is time independent, then p_t is constant — a fact we can and will exploit in a later lecture. This constancy is related to the fact that such a metric has a Killing vector $\vec{\xi}$; one can show that $p^\alpha\xi_\alpha$ is conserved along a geodesic worldline.

Lecture 10: This lecture begins by examining geodesics in a particular spacetime, $ds^2 = -(1+2\Phi)dt^2 + (1-2\Phi)(dx^2+dy^2+dz^2)$, with $\Phi = \Phi(x,y,z) \ll 1$. Considering the slow motion limit shows that such geodesics yield the Newtonian equation of motion, with Φ the gravitational potential. (We derive and justify this spacetime in a later lecture.)

We next derive a tensor which describes spacetime *curvature* by consider the parallel transport of a vector around a closed figure. Take the vector to be V^μ, and transport it around a parallelogram with sides δx^α,

δy^β. When it comes back to its stating point, the vector will have changed by $\delta V^\mu = R^\mu{}_{\nu\alpha\beta} V^\nu \delta x^\alpha \delta y^\beta$, where $R^\mu{}_{\nu\alpha\beta}$ is the *Riemann curvature* tensor. Riemann has certain important symmetries catalogued in this lecture; carefully counting them up shows that it has 20 independent components, exactly accounting for the 20 constraints that, at second order, cannot be "transformed away" by going into a freely falling frame of reference.

It's worth noting that this is equivalent to saying that the commutator of covariant derivatives is non zero, with the action producing the Riemann curvature: $[\nabla_\mu, \nabla_\nu] p^\alpha = R^\alpha{}_{\beta\mu\nu} p^\beta$.

Lecture 11: More curvature: by taking traces, we define the *Ricci curvature*: $R_{\mu\nu} \equiv R^\alpha{}_{\mu\alpha\nu} = g^{\alpha\beta} R_{\alpha\mu\beta\nu}$. The Ricci scalar is in turn found by tracing the Ricci tensor: $R \equiv R^\mu{}_\mu = g^{\mu\nu} R_{\mu\nu}$.

Spacetime curvature describes tides. We see this manifested most clearly by considering two nearby geodesics, along which events at the same affine parameter λ are separated by a vector X^α. (I've changed notation slightly from what is in the lecture notes to avoid confusion with the symbol usually used to denote the Killing vector.) Take the tangent vector to these geodesics to be u^α (they are close enough that they tangent vectors are the same to first order in separation). Propagating along these geodesics, we see that the separation vector evolves according to $u^\beta \nabla_\beta \left(u^\alpha \nabla_\alpha X^\mu \right) = R^\mu{}_{\gamma\delta\nu} u^\gamma u^\delta X^\nu$.

Finally, by applying the covariant derivative commutator rules in a somewhat complicated way, one can show that the Riemann tensor obeys the *Bianchi identity*: $\nabla_\alpha R_{\beta\gamma\mu\nu} + \nabla_\beta R_{\gamma\alpha\mu\nu} + \nabla_\gamma R_{\alpha\beta\mu\nu} = 0$.

Lecture 12: By tracing over pairs of indices, we show that the Bianchi identity can be written $\nabla^\mu G_{\mu\nu} = 0$, where $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R$ is the *Einstein curvature* tensor. Note that the trace of Einstein is the same as the trace of Ricci, up to a minus sign: $G^\mu{}_\mu = R - \frac{1}{2} g^\mu{}_\mu R = -R$ (using the fact that $g^\mu{}_\mu = 4$).

With the Einstein tensor in hand, we at last have the tool we need to make a field equation for spacetime. Our governing principle is that the stress-energy tensor is the most natural source for spacetime. We need a left-hand side of the equation; it must be a two-index, symmetric, divergence-free curvature tensor. This tells us that our field equation is of the form $G_{\mu\nu} = \alpha T_{\mu\nu}$. By demanding that this equation yield Newtonian gravity in the appropriate limit, we find $\alpha = 8\pi G$ (or $8\pi G/c^4$ if we work in units where $c \neq 1$), yielding at last the Einstein equations of general relativity: $G_{\mu\nu} = 8\pi G\, T_{\mu\nu}$.

Notice that the left-hand side of this equation just needs to be any divergence-free tensor with the dimensions of curvature. The metric (modulo an appropriate constant) is actually such a tensor, since the metric has zero covariant derivative. We can thus add $\Lambda g_{\mu\nu}$ to the left-hand side, where Λ is known as the *cosmological constant*. This term can be interpreted as a uniform stress-energy filling all of spacetime; it is in fact a perfect fluid with $\rho = -P = \Lambda/8\pi G$.

Lecture 13: This lecture is somewhat more advanced material, presenting a second route to the Einstein field equations: via a *variational principle*. We first review how one can extremize a Lagrangian density $\hat{\mathcal{L}}$ that depends on some field and thereby derive Euler-Langrange equations which turn into equations governing that field. For gravity, we take that field to be the spacetime metric. For our Lagrangian density, we require that it be some scalar derived from the curvature. The simplest such choice is to put $\hat{\mathcal{L}} = R$, the Ricci scalar; an almost straightforward calculation shows that this choice yields the Einstein field equations.

Several comments are in order here. First, on the calculation being "almost" straightforward: one step, showing that a particular term which arises in the calculation can be neglected, is quite subtle. At the level of 8.962, it suffices to say that the term can in fact be eliminated. Interested students are referred to Appendix E of the textbook by Wald for detailed discussion. One could also proceed by treating both the metric and its derivative (in the form of the connection) as quantities which are separately varied (much as one separately varies a particle's position and velocity when particle kinematics is computed in Lagrangian mechanics). Doing so is called performing the *Palatini* variation; the problematic term discussed above does not appear in this case, and one finds that the connection and the metric are related via the Christoffel symbol formula. Your lecturer was hoping to update his lecture notes to present the Palatini variation this year, but thanks to the COVID-19 emergency was unable to do so.

Second, it is worth thinking about the significance of the fact that equating the Lagrangian density to the simplest possible curvature scalar, the Ricci scalar R, yields the Einstein field equation. This demonstrates that general relativity is, in a very quantifiable sense, the *simplest* possible relativistic theory of gravity. One can certainly imagine more complicated Lagrangians (e.g., include terms that go as $1/R$, or R^2, or that include gravitational coupling to different fields); indeed, if one expects that if gravity can be unified with the other fundamental forces, it is is likely that there will be corrections to this "leading" description. The

Lagrangian formulation gives us a systematic way of building relativistic theories of gravity that go beyond Einstein's general relativity.

Lecture 14: The remainder of this course is dedicated to solving the Einstein field equations (the "EFEs") to construct spacetimes, and to examine the properties of these spacetimes. As a system of differential equations, the EFEs are tremendously complicated: they are ten coupled, nonlinear partial differential equations, with complicated boundary conditions. We will discuss in detail two ways of solving them. The first method is to assume that spacetime is "close to" the flat spacetime of special relativity, and linearize around this "flat background." The second method is to assume a symmetry, and use that symmetry to reduce the complexity of these equations. (A third method, which is discussed briefly in some "extra" lectures toward the end of the course, is to simply attack the full coupled nonlinear complications of these equations head on. With effort, one can then re-write the equations in a form amenable to numerical analysis.)

We begin by linearizing around flat spacetime: we put $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, and assume the components $h_{\mu\nu}$ are small enough that any term of order h^2 can be neglected. In this limit, an important coordinate transformation is the *infinitesimal* transformation: We change coordinates according to $x^\alpha \to x^\alpha + \xi^\alpha$, and take the generator of the transformation to have the property that $\partial_\alpha \xi^\beta$ is likewise small. Applying this coordinate transformation to $g_{\mu\nu}$, we find that it is equivalent to shifting the perturbation according to the rule $h_{\mu\nu} \to h_{\mu\nu} - \partial_\mu \xi_\nu - \partial_\nu \xi_\mu$. A straightforward exercise shows that this shift leaves curvature tensors unchanged. As such, this operation is essentially identical to a gauge transformation in electrodynamics which changes potentials but leaves fields unchanged; it is in fact often called a gauge transformation in linearized gravity.

Constructing the Einstein tensor, rewriting it in terms of the "trace-reversed" metric perturbation $\bar{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu} h$ (where $h = \eta^{\mu\nu} h_{\mu\nu}$; notice that $\bar{h}^\mu{}_\mu = -h$, hence the name trace reversed), and choosing our gauge appropriately, we find the EFEs reduce to the simple wave equation $\Box \bar{h}_{\mu\nu} = -16\pi G\, T_{\mu\nu}$. Solving this in the time-independent limit (for which the wave operator $\Box$ goes over to the Poisson operator ∇^2) yields the spacetime $ds^2 = -(1+2\Phi)dt^2 + (1-2\Phi)(dx^2 + dy^2 + dz^2)$ which describes the Newtonian limit of general relativity.

Lecture 15: Here we examine the linearized EFE $\Box \bar{h}_{\mu\nu} = -16\pi G\, T_{\mu\nu}$ for a non-static source. Any equation of this form can be solved using a radiative Green's function (discussed in electrodynamics texts like Jackson, for example). This gives us a solution in which all ten components of $\bar{h}_{\mu\nu}$ appear to be radiative.

Looks can be deceiving: this appearance is a consequence of the gauge we used to write the linearized EFE in this simple-to-solve fashion. This lecture then presents a synopsis of a somewhat advanced topic: how to characterize the *gauge-invariant* degrees of freedom encoded in the metric components $h_{\mu\nu}$. The result of this exercise is that we find there are six such degrees of freedom. Four are governed by Poisson-type equations (∇^2"field" = "source"), and as such describe non-radiative contributions to the spacetime (much as a Coulomb electric field or an Ampere magnetic field describes the non-radiative bits of an electrodynamic system). The other two degrees are radiative (governing equation of the form $\Box$"field" = "source"), and are encoded in the spatial, transverse, and traceless components of the metric perturbation $h_{\mu\nu}$. Of the ten components $h_{\mu\nu}$, only six represent gauge-invariant contributions to spacetime physics; the remaining four components are purely gauge.

It's worth reiterating that much of this lecture is on the advanced side; we do not expect students to follow every detail. We will use the result characterizing radiation in the next lecture.

Lecture 16: In this lecture, we study gravitational radiation. We first write down a metric perturbation $h_{\mu\nu}$ whose components describe fields that propagate in the z direction at the speed of light (as we expect for radiation), and we require the metric to be transverse (components parallel to the propagation direction — the z and t components — set to zero) and traceless (sum of the diagonal elements is zero). These requirements tell us that the only non-zero components are $h_{xx} = -h_{yy} \equiv h_+$, and $h_{xy} = h_{yx} \equiv h_\times$.

The geodesic equation quickly reveals that test bodies which are initially at rest in this spacetime apparently remain at rest. Bear in mind, though, that the geodesic equation tells us about motion with respect to a given coordinate system. A more meaningful calculation is to compute the proper separation of two test bodies in this spacetime, or to compute the geodesic deviation of these bodies. This reveals that the separation of the bodies stretches and squeezes, depending on the relative values of h_+ and $h_\times$ and the geometry of the bodies' separation.

We next revisit the linearized EFE in order to understand how to compute radiation given a particular solution. We see that radiation depends at leading order on the second time derivative of a source's mass quadrupole moment. This should be reminiscent of electrodynamics, in which the leading radiative potential depends on the first time derivative of a source's charge dipole moment. The solution also involves a set of projection tensors (first analyzed on problem set 1, problem 3), which enforce the condition that radiation be transverse and traceless.

Lecture 17: This lecture is considerably more advanced than most of the material we discuss in this course, and as such is presented somewhat schematically. The goal of this analysis is two fold: first, to understand how to describe gravitational waves propagating on a background more general than the flat background we used in Lecture 16; and second, to characterize the energy and momentum carried by gravitational radiation as it propagates across the universe.

On the first point, it is important to keep in mind that on a general background, defining what part of spacetime is "radiation" and what part is "not radiation" is quite ambiguous. Both represent spacetime curvature; both can be spatially and temporally varying. For a separation into "radiation" and "background" to make sense, there must be a notion of separation of scales: the radiation varies on short times and lengthscales; the background must vary on long time and lengthscales. Doing so, we define the radiative content of the spacetime metric, the associated curvature tensors, and generalize the notion of gauge transformation to this general background.

These foundations are needed for us to study the energy and momentum carried by gravitational waves. A key to this must be *nonlocality*: the energy in gravitational waves cannot be localized to a single point. One can always go to a freely falling frame at that point, representing spacetime there in nearly inertial coordinates. At that single point, there is no wave! As we saw in the previous lecture, we need to examine separated points in order for the wave's effects to show up. We will need to average the wave's effects over a region that is several wavelengths in size. At linear order, the wave will vanish when we so average it; this means that we must go to second order in perturbation theory.

What follows in these notes schematically presents how to organize the Einstein field equations. We sketch how the second order term acts, after averaging over the correct length and time scales, as an effective stress-energy tensor back-reacting on the background spacetime. After organizing terms, the leading contribution to the power carried by gravitational waves takes the form of three time derivatives of the quadrupole moment, squared. This again is reminiscent of electrodynamics, in which the leading power carried from a source takes the form of two time derivatives of the dipole momentum squared.

Lecture 18: Cosmology and cosmological spacetimes. This is the first problem for which we solve the EFE by imposing a symmetry. In this part of the class we consider *maximally symmetric spacetimes* (or spaces). Such a space is a manifold whose geometry has the largest number of possible Killing vectors given the space's dimension: $n(n+1)/2$ Killing vectors in n dimensions. Flat spacetime is one example — the 10 Killing vectors ($n = 4$) correspond to 3 boosts, 3 rotations, and 4 translations. Euclidean space is another — the 6 Killing vectors ($n = 3$) correspond to 3 rotations and 3 translations. A maximally symmetric space must have a Riemann tensor $R_{\alpha\mu\beta\nu} = R(g_{\alpha\beta}g_{\mu\nu} - g_{\alpha\nu}g_{\beta\mu})/[n(n-1)]$ (where R is the Ricci scalar).

Our universe is spatially homogeneous and isotropic on the largest scales[2], but is not temporally isotropic: the past looks different from the present (and presumably the future). We choose a spacetime which reflects this, writing $ds^2 = -dt^2 + R^2(t)\gamma_{ij}dx^i dx^j$. The spatial coordinates x^i are dimensionless; all dimensions of length are absorbed into the scale factor[3] $R(t)$. Notice we have set $g_{tt} = -1$ and $g_{ti} = 0$. This means we have chosen "comoving coordinates": the separation of two observers at rest in these coordinates will change depending on how the scale factor $R(t)$ behaves. It is worth noting that Earth is *not* comoving, thanks to the Solar System's orbit around the Milky Way and the Milky Way's motion within the local group of galaxies.

We require γ_{ij} to describe a maximally symmetric 3-space. Maximally symmetric implies that it be spherically symmetric, so we can write $\gamma_{ij}dx^i dx^j = f(\bar{r})d\bar{r}^2 + \bar{r}^2 d\Omega^2$, where $\bar{r}$ is a dimensionless radial coordinate and $d\Omega^2$ is the usual spherical angular interval. Computing the Ricci tensor two different ways (one: simply "turn the crank" using the mathematics of curvature tensors given a metric; two: enforce maximal symmetry) we can solve for $f(\bar{r})$, yielding

$$ds^2 = -dt^2 + R^2(t)\left[\frac{d\bar{r}^2}{1 - k\bar{r}^2} + \bar{r}^2 d\Omega^2\right] .$$

[2] Tens of megaparsecs and larger.

[3] Caution: overloaded nomenclature abounds in this subject. We will soon shift to a different definition of scale factor, but for now be aware that this $R(t)$ is *not* the Ricci curvature scalar.

As shown in the notes, k takes on three possible values: $-1, 0, 1$. By changing coordinates using $d\chi = d\bar{r}/\sqrt{1-k\bar{r}^2}$, one can re-write this as

$$ds^2 = -dt^2 + R^2(t)\left[d\chi^2 + S_k^2(\chi)d\Omega^2\right] .$$

When $k = +1$, $S_k(\chi) = \sin\chi$. Each spatial slice has the geometry of a 3-sphere; note that such a space has a maximum radius. This is called a "closed" universe. When $k = +1$, $S_k(\chi) = \chi$. Each spatial slice is Euclidean; this is called a "flat[4]" universe. When $k = -1$, $S_k(\chi) = \sinh\chi$. Each spatial slice is a hyperboloid; this is called an "open" universe.

A common notation is pick a particular value of the scale factor $R(t)$, say $R_0 = R(\text{now})$, then define $a(t) = R(t)/R_0$, $r = R_0\bar{r}$, $\kappa = k/R_0^2$. The line element becomes

$$ds^2 = -dt^2 + a^2(t)\left[\frac{dr^2}{1-\kappa r^2} + r^2 d\Omega^2\right] .$$

Note that $a(\text{now}) \equiv a_0 = 1$. This form in fact is probably the most common way of denoting these spacetimes, which are called *Robertson-Walker* metrics. It's worth noting that the parameter R_0, which provides the overall scale to all dimensionful quantities, cannot be measured.

To proceed further, we need to think about the right-hand side of the EFE. We will take our source to be a perfect fluid, which satisfies the requirement of spatial homogeneity and isotropy, requiring that all fluid elements be at rest with respect to the comoving coordinates. The requirement that $\nabla_\mu T^\mu{}_0 = 0$ leads to an important formulation of local conservation of energy: we find it tells us that $\partial_t(\rho R^3) = -P\,\partial_t(R^3)$. This is simply the first law of thermodynamics, $dU = -P\,dV$, written in a funny way.

The Einstein field equations yield two equations which govern the scale factor of the universe:

$$\begin{aligned}\left(\frac{\dot{a}}{a}\right)^2 &= \frac{8\pi G}{3}\rho - \frac{\kappa}{a^2} , \\ \left(\frac{\ddot{a}}{a}\right) &= -\frac{4\pi G}{3}(\rho + 3P) .\end{aligned}$$

These are known as the *Friedmann equations*; a Robertson-Walker metric with $a(t)$ determined according to these equations is known as a Friedmann-Robertson-Walker (FRW) spacetime.

At this point it is useful to introduce some notation. $H \equiv \dot{a}/a$ is the Hubble expansion parameter; its value now H_0 is called the Hubble constant. We define[5] $\Omega = \rho/\rho_{\rm crit}$, where $\rho_{\rm crit} \equiv 3H^2/8\pi G$ is the "critical density." The first Friedmann equation can then be written

$$\Omega - 1 = \frac{\kappa}{H^2 a^2} .$$

Notice that the value of Ω determines whether $k = -1$, 0, or $+1$: If $\Omega > 1$ ($\rho > \rho_{\rm crit}$), then we must have $k = +1$; if $\Omega = 1$ ($\rho = \rho_{\rm crit}$), $k = 0$; if $\Omega < 1$ ($\rho < \rho_{\rm crit}$), then $k = -1$. The overall density of the universe compared to the critical value determines whether our universe is open, flat, or closed.

Further progress requires an equation of state that relates P and ρ. Cosmology generally uses $P = w\rho$, with w a constant. One can imagine a universe with multiple kinds of "stuff," ie contributions with different values of w. For intuition, imagine a single value of w dominating the universe. Our first law formulation can then be integrated up to find

$$\frac{\rho}{\rho_0} = \left(\frac{a}{a_0}\right)^{-3(1+w)} .$$

Consider now different kinds of material that can be sources of stress-energy. The three most commonly considered are

1. Pressureless matter (or simply "matter"), $w = 0$. This actually gives a good description of a universe filled with galaxies and galaxy clusters that very weakly interact with each other (except via gravity). Notice that $\rho_m \propto a^{-3}$ as a changes. This says that if one considers a chunk of the universe, the number of matter "particles" in it is fixed while the volume evolves as a^3.

[4]Caution: Although each spatial slice is flat, space*time* is curved! When a cosmologist calls the universe "flat," the meaning is bit different from how we've used the term to date.

[5]Ω is used for both this normalized density, and for a solid angle element. Context generally makes the meaning clear.

2. Radiation: radiation pressure has an equation of state $P_r = \frac{1}{3}\rho_r$, so $w = 1/3$ in this case, and we find $\rho_r \propto a^{-4}$. Here, as a changes, not only does the volume change with a^3, but the energy of each "particle" of radiation scales with $1/a$. In the next lecture, we rigorously show that radiation *redshifts* with the scale factor a exactly as this scaling suggests.

3. Cosmological constant: As discussed when we derived the Einstein field equation, a cosmological constant is equivalent to perfect fluid with $P_\Lambda = -\rho_\Lambda$, meaning $w = -1$. This leads to $\rho_\Lambda =$ constant! A cosmological constant is akin to a vacuum energy density (with negative pressure) that is independent of the universe's overall scale.

Lecture 19: Cosmology continued. Cosmology as a science is all about understanding the large scale structure of the universe and its constituents, which essentially boils down to understanding what are the sources of stress-energy that determine the metric of spacetime on the largest scales, the value of k in the Robertson-Walker metric (ie, whether our universe is open, flat, or closed), and how the scale factor $a(t)$ behaves. The "forward" problem is conceptually simple: given a stress energy tensor that contains a mixture of different kinds of matter at some initial time, simply integrate the Friedmann equations, with the constraint of $\partial_t(\rho a^3) = -P\,\partial_t(a^3)$ that arises from local energy conservation. This can even be done analytically in some simple limits: if $k = 0$, and the stress energy tensor is all matter or all radiation, then $a(t)$ grows as a power law in time; if the stress-energy tensor is cosmological constant, then $a(t)$ grows exponentially with time. (Calculation of this is actually given at the end of the notes for Lecture 18.)

Of more interest is the inverse problem: given what we can measure, can we determine what our universe is made of? Essentially, we would like to measure the scale factor a at various t. This means we need an observationally useful surrogate for both the scale factor a, and for the time t — when we look at, for example, a distant object (like a quasar or a galaxy), we need to know the scale factor at which that object emitted its radiation, and how long ago the radiation was emitted.

Consider the scale factor first. A useful tool for us is that an FRW spacetime has a *Killing tensor*. Recall a Killing vector $\vec{\xi}$ is related to a symmetry of the spacetime, and satisfies the equation $\nabla_{(\alpha}\xi_{\beta)} = 0$. A Killing tensor generalizes this to more indices: a rank n Killing tensor has n indices and satisfies $\nabla_{(\alpha}K_{\beta\gamma\delta\ldots)} = 0$. With this, it's a straightforward exercise to show that, if u^α is the tangent to a geodesic trajectory, then $\mathcal{K} = K_{\alpha\beta\gamma\delta\ldots}u^\alpha u^\beta u^\gamma u^\delta \ldots$ is constant along that trajectory.

The Killing tensor we wish to use is $K_{\mu\nu} = a^2(t)(g_{\mu\nu} + u_\mu u_\nu)$, where $\vec{u}$ is the 4-velocity of a comoving observer in the spacetime. We contract this with the 4-momentum of a null geodesic, and conclude that $\mathcal{K} = K_{\mu\nu}p^\mu p^\nu = a^2[g_{\mu\nu}p^\mu p^\nu + (u_\mu p^\mu)(u_\nu p^\nu)] = a^2E^2$ is constant along that null geodesic, where E is the energy measured by a comoving observer. Hence, $a(t)E$ is a constant, meaning that the null geodesic's energy as measured by comoving observers varies as $1/a$.

Note that this justifies part of the intuitive explanation for why the energy density of radiation evolves with a^{-4} as the scale factor a evolves. More importantly, this shows how we can directly measure the scale factor a: It is encoded as the *redshift* of radiation that we can measure. In particular, if we imagine measuring radiation that is emitted when the scale factor is a and with a particular energy spectrum, every feature in that spectrum will be redshifted by a factor of a when we measure it (using the convention that the scale factor now is 1). Astronomers typically measure redshift z, defined by a shift in the wavelength of a source: $z = (\lambda_{\rm obs} - \lambda_{\rm emitted})/\lambda_{\rm emitted}$. As shown in the notes, the scale factor at emission is $a_{\rm emitted} = 1/(1+z)$.

We also need to know the time t at which light is emitted. This is hard to measure, but it should be possible at least in principle to determine the distance d over which the light traveled to reach us. By using null geodesics, we should be able to convert the distance that the light travels to the time light is emitted.

In Euclidean geometry, there are at least three ways that one can determine how far away a distant object is. First, you can compare an object's intrinsic luminosity (assuming it is known somehow) to the radiated flux we measure from it: $F = L/(4\pi D_L^2)$. The distance we determine this way is called the *luminosity distance*. Second, we can compare the physical size of the object (assuming it is known somehow) to the angular size: $\Delta\Theta = \Delta L/D_A$. Distance determined this way is the *angular diameter distance*. Finally, we can compare a known transverse speed to the apparent angular speed of an object on the sky: $\dot{\Theta} = v_\perp/D_M$. This is called the *proper motion distance*. These three notions of distance give the same result in Euclidean geometry, but differ in an FRW spacetime. After going through the detailed calculation, find that $D_L = (1+z)D_M = (1+z)^2 D_A$.

Cosmology then becomes a task of precisely measuring z and various distance measures for a large number of objects across a range of redshift, and comparing with models to understanding the nature of the universe. We are fortunate that Nature in fact provides some objects with known standard (or standardizable) luminosities and sizes ("standard candles" and "standard rulers"), making it possible to measure both luminosity distances and angular diameter distances. Our present best fit suggests a universe that is spatially flat ($k = 0$), with a Hubble constant $H_0 \simeq 70\,\mathrm{km/(sec\text{-}Mpc)}$ (although there are some odd discrepancies showing up in this parameter in recent measurements), with about 30% of the energy density in the form of matter (and only about 15% of that as matter that fits the standard model), and the remaining 70% of the energy density as some form of "dark energy" that behaves like a cosmological constant. Optional readings describing up-to-date cosmological models, and presenting the detailed challenge of how to build such models, have been posted to the 8.962 website.

Two lingering mysteries are why it is that the universe is so flat, and why it is so homogeneous. If one considers the fundamental parameter in the FRW metric to be κ, then $\kappa = 0$ is a single point in a broad range of possible values. One might think that perhaps $|\kappa|$ is simply very small. It is not hard to show, however, that in a matter- or radiation-dominated universe (as it appears our universe was over much of its early history), then $|\kappa|$ tends to grow as a power of the scale factor, pushing away from zero to become either more positive or more negative. Homogeneity (particularly at the earliest times) indicates that all of the observable sky must have been in thermal equilibrium at the earliest moments in the universe's history. However, it is not terribly difficult to show that, if the universe were matter or radiation dominated, then large patches of the sky would have be "out of causal contact" (ie, unable to exchange information) early in its history. This cannot be sensibly reconciled with the idea that the early universe was in thermal equilibrium.

Cosmic inflation offers a solution to both of these problems. If, at the earliest moments in its history, the universe were in a "false vacuum" state, then spacetime would be filled with an energy density that acts like a cosmological constant. As is explored on pset #9, by having the universe expand exponentially, you can drive its spatial curvature so close to zero that it is unobservable, and you can inflate a patch of early universe that is easily in causal contact to a large enough scale to explain the universe we observe today.

Lecture 20[6]*:* The spacetime of a spherically symmetric *compact* body. In this lecture, we continue to simplify the EFE by imposing symmetry, but now we consider a source that is "compact": the body occupies a finite volume of space, and spacetime asymptotically approaches the flat spacetime of special relativity far away. In practice, this means that we imagine the source has some non-zero stress-energy tensor for $r \le R_S$ (where R_S is the body's surface), and has $T_{\mu\nu} = 0$ for $r > R_S$.

Begin by imagining that spacetime is static (nothing varies with time). The most general such spacetime has a line element with the form $ds^2 = -e^{2\Phi(r)}dt^2 + e^{2\Lambda(r)}dr^2 + R(r)^2 d\Omega^2$. Notice that all functions depend only the radial coordinate r, and that the angular section is proportional to the 2-sphere metric $d\Omega^2 = d\theta^2 + \sin^2\theta\, d\phi^2$.

At this point, we choose $R(r) \equiv r$. This means that r is an *areal* radius: events at radius r all lie on the surface of a sphere centered on the origin and with proper area $4\pi r^2$. Note that this is *not* a unique choice. In fact, the weak-field solution yielding the Newtonian limit that we studied earlier was written using *isotropic coordinates*, a choice which emphasizes the fundamental isotropy of the three spatial directions.

With the metric selected, it is a straightforward exercise to construct the curvature tensor components and to assemble the Ricci and Einstein tensors. We begin by examining what the EFE tells us in the exterior $r > R_S$, where $T_{\mu\nu} = 0$. As shown in the notes, we find

$$\begin{aligned} \partial_r \Phi = -\partial_r \Lambda \qquad &\rightarrow \qquad \Phi = -\Lambda + k\ , \\ \partial_r \left[r e^{2\Phi}\right] = 1 \qquad &\rightarrow \qquad \Phi = \frac{1}{2}\ln\left(1 + \frac{A}{r}\right)\ , \end{aligned}$$

where k and A are constants of integration. The constant k ends up being just a coordinate rescaling, so we can set $k = 0$ without loss of generality. To fix A, note that the spacetime now takes the form

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

and consider non-relativistic ($dt/d\tau \simeq 1$, $dr/d\tau \ll 1$) radial ($d\theta/d\tau = d\phi/d\tau = 0$) infall in the weak field ($r \gg |A|$) of this spacetime. The geodesic equation in this situation becomes $d^2r/dt^2 = A/2r^2$. The

[6] Note that the handwritten label on this PDF file is "Lecture 21." Lecture 20 was an optional guest lecture the year I wrote these notes, coinciding with travel to the April APS meeting. All the lecture notes from Lecture 20 onward are shifted by 1 in their handwritten label.

expectation from Newtonian free fall is $d^2r/dt^2 = -GM/r^2$, where M is the mass of the body. This leads us to identify $A = -2GM$. The exterior spacetime becomes

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

an extremely important and famous result known as the *Schwarzschild metric*, derived by Karl Schwarzschild under rather arduous circumstances shortly after Einstein announced the field equations of general relativity (I have a minute or so digression on this in the video recording of this lecture). One reason that this result is so important is that this spacetime describes the exterior vacuum region of *any* spherically symmetric body — even ones that are time-varying, as long as the variations preserve the symmetry ("Birkhoff's theorem").

For the interior, we use our old friend the perfect fluid, writing the stress-energy tensor in the form $T^\mu{}_\nu = \mathrm{diag}[-\rho(r), P(r), P(r), P(r)]$. The equation $G^t{}_t = 8\pi G T^t{}_t$ takes a simple form provided we define

$$e^{-2\Lambda(r)} \equiv 1 - \frac{2G\,m(r)}{r} .$$

The interior mass function $m(r)$ is determined by the fluid's density:

$$\frac{dm}{dr} = 4\pi r^2 \rho(r) \qquad \text{or} \qquad m(r) = 4\pi \int_0^r \rho(r')(r')^2 dr' .$$

[Note that $m(0) = 0$. This may seem obvious, but in fact there are solutions in which $m(0) \neq 0$ which we will discuss soon.] Putting $m(R_S) = M$, this solution for $m(r)$ allows the interior to very nicely connect with the exterior at the surface radius $r = R_S$.

The equation $G^r{}_r = 8\pi G T^r{}_r$ yields an equation for the metric function $\Phi(r)$:

$$\frac{d\Phi}{dr} = \frac{G[m(r) + 4\pi r^3 P(r)]}{r[r - 2Gm(r)]} .$$

On an old homework exercise, you showed that for a perfect fluid $(\rho + P)u^\beta \nabla_\beta u_\alpha = -\partial_\alpha P - u_\alpha u^\beta \partial_\beta P$. Combining this with the fact that for a static fluid in the spacetime each fluid element has a 4-velocity $u^\alpha \doteq (e^{-\Phi}, 0, 0, 0)$, we find

$$\frac{dP}{dr} = -\frac{G[\rho(r) + P(r)][m(r) + 4\pi r^3 P(r)]}{r[r - 2Gm(r)]} .$$

The equations for $m(r)$, dP/dr, and $d\Phi/dr$ are known as the *Tolman-Oppenheimer-Volkov* (TOV) equations of stellar structure. To solve them, we we need an *equation of state* which relates pressure and density, and an "initial" condition describing the matter at $r = 0$. The solution then gives us a model for a spherical "star" in fully relativistic gravity.

Lecture 21: In this lecture, we continue our discussion of the spacetimes describing spherically symmetric compact bodies. To get some insight into solutions of the TOV equations, consider a highly idealized, unphysical limit: a star with $\rho = \text{constant}$. Such a star would have an infinite speed of sound, which violates the law that no signal or information can travel faster than light. Despite this rather unphysical behavior, these stars yield important insights into general relativistic bodies.

As described in the previous lecture, once the equation of state is selected, relativistic stars are described by a 1-parameter family of solutions: pick the central pressure or central density, and the rest of the star's properties are determined. For our constant density stars, pressure is the useful parameter. The TOV equations yield the following solution relating the star's radius R_S and the central pressure P_c:

$$R_S^2 = \frac{3}{8\pi G\rho}\left[1 - \frac{(\rho + P_c)^2}{(\rho + 3P_c)^2}\right] \qquad \text{or} \qquad P_c = \frac{\rho\left[1 - (1 - 2GM/R_S)^{1/2}\right]}{3\sqrt{1 - 2GM/R_S} - 1} .$$

(Note that $M = \frac{4}{3}\pi\rho R_S^3$.) The second form tells us the central pressure we must have in order for the star to have a particular radius R_S. What's particularly interesting here is that P_c *diverges* if the star is too compact:

$$\begin{aligned} P_c \text{ finite} \quad \text{requires} \quad 3\sqrt{1 - 2GM/R_S} - 1 &> 0 \\ \rightarrow \frac{GM}{R_S} &< \frac{4}{9} . \end{aligned}$$

This means that for uniform density stars, *we* **cannot** *have physically reasonable pressure profiles if* $R_S < 9GM/4$. Although the proof goes beyond 8.962, this limit turns out to be quite general, and is expressed by a result known as Buchdahl's theorem: *No stable spherical fluid configuration can exist for a body with a surface radius* $R_S < 9GM/4$. If such a star existed, it would not be stable, and would collapse to "something else" (to be discussed soon).

More realistic bodies are described by an equation of state[7] $P = P(\rho)$. An approximate form that is often used at least for test purposes is a power law, $P = K\rho_0^\Gamma$, where K and Γ are constants. This form is known as a *polytrope*. Note that the density which appears here is the *rest mass* density; it does not take into account the work that is done in compressing a fluid to density ρ. As described in supplemental notes[8], the relationship between rest density and density ρ for a polytrope is given by

$$\rho = \rho_0 + \frac{P}{\Gamma - 1} = \rho_0 + \frac{K\rho_0^\Gamma}{\Gamma - 1} \, .$$

With this in hand, it is a straightforward computational exercise to numerically integrate the TOV equations to build relativistic stellar models.

What if we had a spacetime that was given by the Schwarzschild form

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

for *all* r, not just the vacuum exterior of some body? This is, after all, an exact solution for a vacuum $T_{\mu\nu} = 0$; but it is a vacuum that somehow has a mass M.

This situation is somewhat reminiscent of the Coulomb point charge in electrodynamics — a field for which the charge density is zero everywhere, but the total charge is q. In electrodynamics, we cured this apparent contradiction by invoking a Dirac delta function density distribution; in general relativity, it is harder thanks to the nonlinear nature of the field equations.

Considering the spacetime metric itself, we see two radii that appear to be problematic: $r = 0$ and $r = 2GM$. Since metric components can be deceiving, let's assemble a scalar invariant from curvature tensors:

$$I \equiv R_{\alpha\beta\gamma\delta} R^{\alpha\beta\gamma\delta} = \frac{48G^2M^2}{r^6} \, .$$

The square root of this quantity (called the Kretschmann scalar) is, roughly, a measure of the tidal forces felt by a freely falling body in the spacetime. Notice that I diverges as $r \to 0$; that point is, indeed, singular. But there's nothing particularly "special" about $r = 2GM$; an observer falling into this spacetime would pass $r = 2GM$ without any particular notification that anything interesting happened there.

As the lecture notes describe, this metric has a *coordinate singularity* at this radius. Insight into the nature of this singularity can be found by examining the motion of a test body dropped from some starting radius $r = r_0$, but parameterizing this motion with both proper time (time measured according to the body's own clock) and coordinate time. The result is that

$$\begin{aligned}
\tau &= \frac{4GM}{3}\left[\left(\frac{r_0}{2GM}\right)^{3/2} - \left(\frac{r}{2GM}\right)^{3/2}\right] , \\
t &= 2GM\left\{\ln\left[\frac{(r/2GM)^{1/2}+1}{(r/2GM)^{1/2}-1}\right] - 2\sqrt{\frac{r}{2GM}}\left(1 + \frac{r}{6GM}\right) - (\text{same expression but with } r \to r_0)\right\} .
\end{aligned}$$

Notice that the infalling body reaches $r = 0$ in finite proper time τ. **However**, as $r \to 2GM$, the coordinate time diverges: $t \to \infty$ as $r \to 2GM$.

How is that possible??? The answer, as we'll discuss in the next lecture, is tied up with what the coordinate time t means, and how clocks behave in a gravitational field.

[7]Strictly speaking, this is a "cold" equation of state, in which the density decouples from entropy because the fluid is "cold." The scale defining "hot" and "cold" for relativistic fluids is the Fermi temperature. For the objects for which general relativity is important, the Fermi temperature tends to be trillions or tens of trillions of Kelvin. Even the hottest newly born neutron stars are orders of magnitude colder than this, so the "cold" equation of state approximation is quite adequate.

[8]The lecture notes in this section need revision — please instead consult the document entitled "Polytropes and the first law" under "Materials" on the website.

Lecture 22: To address the apparent paradox that emerged at the end of the previous lecture, it is worth remembering what the coordinate t means. The Schwarzschild spacetime is asymptotically flat, meaning that it asymptotes to the metric of special relativity when $r \gg 2GM$. In that limit, t is simply the time measured on the clock of a stationary observer, and so the coordinate t is best thought of as time as measured on the clocks of observers who are very far away from the mass M. Any pathologies in this coordinate are thus tied up in how one connects t at one spatial location to t at another. Recalling that in special relativity we synchronized clocks using the "Einstein synchronization procedure," which relied on the fact that the propagation of light has invariant properties we can exploit, we see that if we want to understand why t behaves as it does, we need to understand what happens to light as it propagates in this spacetime.

Let us imagine that as the body falls in it emits a radially directed radio pulse, and let's examine the energy of this pulse as measured by a sequence of static observers. These observers have 4-velocity

$$u^\alpha \doteq \left[\left(1 - \frac{2GM}{r}\right)^{-1/2}, 0, 0, 0\right] .$$

(Notice that this becomes ill-behaved for $r \leq 2GM$, an aspect of its behavior we will return to shortly.) The energy measured by an observer at radius r is given by $E(r) = -p_\alpha u^\alpha = -p_t u^t(r)$, where p_α are the components of the 4-velocity of the radio pulse, and $u^\alpha(r)$ is the 4-velocity of the static observer at radius r. Let us use this to compare the energy of the pulse at emission, $r = r_0$, to its value when it is very far away:

$$\frac{E_{\rm obs}}{E_{\rm emit}} = \frac{-p_t u^t(r \to \infty)}{-p_t u^t(r_0)} = \sqrt{1 - \frac{2GM}{r_0}} .$$

(We used the fact that p_t is a constant along the light-ray, which follows from $\partial_t g_{\mu\nu} = 0$.) This tells us that the pulse of light *redshifts away* as its point of emission r_0 approaches $2GM$. From a similar analysis, it is not hard to show that if the radio pulse is repeated emitted after an interval ΔT at radius r_0, then the spacing between pulses far away is given by

$$\Delta T_\infty = \Delta T \left(1 - \frac{2GM}{r_0}\right)^{-1/2} .$$

The propagation of light is significantly affected by the strong gravity of this spacetime as $r \to 2GM$. Both of these effects are, at heart, nothing more than gravitational redshift, which has been experimentally tested to very high precision in weak fields, but now carried into an extremely strong gravity regime. Because t, by design, encodes the behavior of light propagation, this coordinates inherits this significant impact — in particular, become totally singular and pathological as $r \to 2GM$. (Further discussion and considerations on light propagation can be found in the 8.962 notes entitled "Behavior of light as it propagates out of Schwarzschild," written to elaborate on some questions asked in the Spring 2019 semester.)

This tells us why when parameterized by t, the infalling body never crosses the event horizon: the distant observers can never see this happen. Information describing events near and approaching $r = 2GM$ take a divergingly long time to reach these observers, and the packets of information which communicate the details of infall are redshifted away as this radius is approached. Rather than seeing the infalling body cross $r = 2GM$, a distant observer sees that body slowly approach this radius and fade from view as the photons which carry information about the body redshift away.

Different coordinate systems can be written down which ameliorate these difficulties and help to clarify what is going on. Leaving the details to the lecture notes, the critical conclusion is that for $r \leq 2GM$, there are *no* timelike or null trajectories which allow an observer to reach or even communicate with larger radii. Events at $r \leq 2GM$ are "out of causal contact" with the rest of the spacetime. The radius $r = 2GM \equiv r_H$ is called an **event horizon**: no events in this part of the spacetime can communicate with any events outside of it. Spacetimes with mass and event horizons are called **black holes**.

The Schwarzschild black hole is the simplest member of a broader family. Black holes can spin ("Kerr" black holes), and they can be charged ("Reissner-Nordstrom"). A charged black hole is spherically symmetric, but is endowed with a Coulomb electric field; the horizon radius becomes $r_H = G(M + \sqrt{M^2 - Q^2})$. A spinning black hole is *not* spherically symmetric, though the horizon is at constant coordinate radius $r_H = GM + \sqrt{G^2M^2 - a^2}$, where $a = \mathbf{S}/M$ is a parameter describing its spin angular momentum. A spinning, charged solution also exists (the "Kerr-Newman" solution).

This finite set of solutions turns out to be enough to describe all black holes in our universe thanks to a set of remarkable results. First, much work over the past several decades has shown that the *only* stationary spacetimes (at least in 3 space plus 1 time dimension) which have event horizons are the Kerr-Newman solutions. Second, if an event horizon forms and it is not of Kerr-Newman form, then the spacetime is not stationary: it is dynamic, and the radiation associated with these dynamics backreacts on the spacetime in such a way as to drive it very rapidly to the Kerr-Newman solution. **The Kerr-Newman spacetime is the generic outcome of black hole formation.**

In fact, charge is certain to be astrophysically irrelevant, since any macroscopic charged object in an astrophysical environment will be rapidly neutralized by nearby plasma. Because of this, we expect that the Kerr solution gives an essentially *exact* solution describing black holes in the universe. It is remarkable that such a simple mathematical object should describe so many massive objects in our universe. As physicists, we of course regard this as a hypothesis that must be tested. A substantial body of work in gravitational wave astronomy (including a good chunk of your lecturer's career) has as its goal probing the strong-field nature of black hole spacetimes and testing the hypothesis that Kerr accurately describes these objects.

Lecture 23: Testing the nature of black hole spacetimes largely boils down to modeling the motion of light and bodies in their vicinity. This could be done by computing all the connection coefficients and studying the geodesic equation. However, black holes' highly symmetric nature means that there are other tools which can be used. In this lecture, we develop the details for motion in the Schwarzschild spacetime. Similar results can be found for motion in the Kerr spacetime; because it has less symmetry (it is axially but not spherically symmetric), the results for Kerr are somewhat more complicated.

The spherical symmetry of Schwarzschild means that we can always rotate our coordinate system so that any point-particle orbit lies in the equatorial plane, $\theta = \pi/2$. It is simple to show that an orbit which begins at $\theta = \pi/2$ with $d\theta/d\tau = 0$ will remain in this plane forever. In essence, the symmetry means that gravity cannot exert a torque to change the orientation of the orbital plane. For all black hole solutions, $\partial_t g_{\mu\nu} = 0$ and $\partial_\phi g_{\mu\nu} = 0$. This guarantees that there exist Killing vectors associated with the time and axial directions, and that the geodesics of these spacetimes have both a conserved energy E and a conserved angular momentum L_z.

The 4-momentum of a body moving on a timelike trajectory in Schwarzschild can in general be written

$$p^\mu \doteq m\left(\frac{dt}{d\tau}, \frac{dr}{d\tau}, 0, \frac{d\phi}{d\tau}\right) .$$

Because the metric is time-independent, we know that $p_t = g_{tt}p^t \equiv -E$ is a constant of the motion. This tells us that

$$E = m\left(1 - \frac{2GM}{r}\right)\frac{dt}{d\tau} .$$

Notice that this relates $dt/d\tau$ to a constant, E, and a simple function of r. Likewise, because the metric is ϕ-independent, we know that $p_\phi = g_{\phi\phi}p^\phi \equiv L_z$ is a constant of the motion:

$$L_z = m\,r^2\sin^2\theta\frac{d\phi}{d\tau} = m\,r^2\frac{d\phi}{d\tau}$$

(using $\theta = \pi/2$ in the final simplification). Finally, use the fact that $g_{\mu\nu}p^\mu p^\nu = -m^2$: expanding this equation, using the connection between E and $dt/d\tau$ and between L_z and $d\phi/d\tau$, we find

$$\begin{aligned}\left(\frac{dr}{d\tau}\right)^2 &= \hat{E}^2 - \left(1 - \frac{2GM}{r}\right)(\left(1 + \frac{\hat{L}_z^2}{r^2}\right) \\ &= \hat{E}^2 - V_{\rm eff}(\hat{L}_z, r) .\end{aligned}$$

We have defined the energy and angular momentum per unit rest mass, $\hat{E} \equiv E/m$ and $\hat{L}_z \equiv L_z/m$, as well as the effective potential $V_{\rm eff}$. Studying trajectories of bodies near Schwarzschild black holes boils down to a simple recipe: First, pick the energy $\hat{E}$ and angular momentum $\hat{L}_z$. Second, pick an initial position (r, ϕ). Finally, integrate up the equations for $dr/d\tau$, $d\phi/d\tau$, and $dt/d\tau$.

As described in the notes, all of the key features of the behavior of black hole orbits is bound up in the effective potential. For example, it's not hard to show using that if $\hat{E} \geq 1$, then the motion is unbound: the orbiting body comes in from far away, turns around at the radius where $\hat{E} = \sqrt{V_{\rm eff}}$, then returns to very far away. An orbit with $\hat{E} < 1$ generalizes an eccentric orbit, turning around at the two radii which solve $\hat{E} = \sqrt{V_{\rm eff}}$. By choosing the conditions just right, one can put the orbit at a particular radius such that $dr/d\tau = 0$ for all time — in other words, a *circular* orbit. As shown in the notes, one must have

$$\hat{E} = \frac{1 - 2GM/r}{\sqrt{1 - 3GM/r}}\ , \quad \hat{L}_z = \pm\sqrt{\frac{GM}{1 - 3GM/r}}\ .$$

For such orbits, it is not difficult to show that the angular frequency of the orbit as seen by distant observers,

$$\frac{d\phi}{dt} = \frac{d\phi/d\tau}{dt/d\tau} = \pm\sqrt{\frac{GM}{r^3}}\ .$$

This, amusingly, is exactly the same as the formula for orbital frequency yielded by Kepler's law. It should be emphasized that this is not deep; different results emerge if one uses, for example, isotropic radial coordinates rather than the areal Schwarzschild coordinate. It is quite convenient, however, and certainly easy to remember. Stable circular orbits cease to exist for $r \le 6GM$, a starkly non-Newtonian behavior.

We repeat this exercise for photon orbits, for which $m = 0$ and we define the 4-momentum as $p^\mu = dx^\mu/d\lambda$. Energy and angular momentum are still conserved, telling us that

$$E = \frac{dt}{d\lambda}\left(1 - \frac{2GM}{r}\right)^{-1}\ , \quad L_z = r^2\frac{d\phi}{d\lambda}\ .$$

We also require $g_{\mu\nu}p^\mu p^\nu = 0$, yielding

$$\left(\frac{dr}{d\lambda}\right)^2 = E^2 - \frac{L^2}{r^2}\left(1 - \frac{2GM}{r}\right)\ .$$

It is a bit disturbing that this appears to predict that the trajectory depends on the energy — as long as we are in the geometric optics limit, the trajectory should be independent of E. To account for this, divide both sides by L_z^2, redefine the affine parameter via $\lambda' = L_z\lambda$ (then drop the prime), and define $b \equiv L_z/E$. The equation becomes

$$\begin{aligned}\left(\frac{dr}{d\lambda}\right)^2 &= \frac{1}{b^2} - \frac{1}{r^2}\left(1 - \frac{2GM}{r}\right)\\ &= \frac{1}{b^2} - V_{\rm phot}(r)\ .\end{aligned}$$

The photon potential has a maximum at $r = 3GM$, and the height at that maximum is $V_{\rm phot} = 1/(3\sqrt{3}GM)^2$. The parameter b is an *impact parameter*: as discussed in the notes and in lecture, it parameterizes the offset of an ingoing trajectory from the center of the black hole. If $b > 3\sqrt{3}GM$ a photon directed inward toward the black hole will scatter around it, propagating back out to infinity. If $b < 3\sqrt{3}GM$, it will fall inside, never returning to large radius. If $b = 3\sqrt{3}GM$, the photon will circulate forever at $r = 3GM$, a special radius known as the "light ring." Generalizations of this notion exist for Kerr black holes, though the details are necessarily somewhat more complicated.

This in fact has astrophysical importance. If a black hole is surrounded by hot, luminous matter, the light from that matter will tend to be trapped at the light ring. The conditions need to be tuned so carefully to stay at this radius that almost every photon so trapped will eventually fall out of the light ring, either being eventually captured by the black hole, or else escaping and eventually reaching distant observers. Thanks to the symmetry of the light ring, one expects that a distant observer will actually see a ring of light, whose radius is exactly the impact parameter $b = 3\sqrt{3}GM$. This is what was observed using high-resolution radio interferometry by the Event Horizon Telescope.

Lecture 24 (not video recorded): The previous 23 lectures cover the most important topics in a one semester academic presentation of general relativity. There are certain topics which, if the course ran longer, we would explore in greater depth, but the contents of those 23 lectures (plus the associated problem sets) should put you on a strong footing for learning about any topic in classical general relativity.

In a normal semester, we would spend the final two lectures covering some advanced material. Both of these lectures describe how to analyze realistic compact sources, particularly highly dynamical, strong-field binaries. These lectures serve two purposes. First, they would present you with material that is related to ongoing research in modern classical general relativity; this background is particularly valuable to understanding the astrophysics of gravitational-wave sources, a focus of your lecturer's research. Second, these lectures would show you a snapshot of techniques that are used to solve the Einstein field equations in situations beyond the circumstances we studied in class (linearizing about a flat background; exploiting a symmetry). In the pandemic term of Spring 2020, I was not able to record videos corresponding to these two lectures, but I have posted the lecture material. These notes are the (relatively) less technical synopsis of this material. Note that no assignments rely on these two lectures; this is truly "bonus" material.

The first of these lectures describe what are essentially modifications of the "perturbation to flat spacetime" idea, describing how one can iterate from small deviations from flat spacetime to not-so-small small deviations, as well as describing perturbations about exact strong-field spacetimes. The exact strong-field spacetimes we examine are black hole solutions; perturbations around cosmological spacetimes have a similar character. The second lecture describes how to reformulate the EFEs in a way that allows one to build a numerical spacetime by direct numerical integration of the field equations.

The first method we discuss is called *post-Newtonian* ("pN") theory, since it shows how one can iterate from the Newtonian limit of general relativity to a solution that progressively becomes closer and closer to the solution of the field equations. PN theory begins by defining a quantity that looks like a metric perturbation:

$$h^{\alpha\beta} = \sqrt{-g}g^{\alpha\beta} - \eta^{\alpha\beta} .$$

Note, however, that we do *not* assume the components of $h^{\alpha\beta}$ are small in any sense. We impose one condition on this tensor: we require $\partial_\alpha h^{\alpha\beta} = 0$. This is called the "deDonder gauge," and the coordinate system we use which respects this is called "harmonic coordinates." (You examined the linearized gravity limit of this on problem set 7.)

When one defines the field $h^{\alpha\beta}$ in this way and imposes deDonder gauge, a seeming miracle happens: the **exact** Einstein field equations become

$$\Box h^{\alpha\beta} = 16\pi G\tau^{\alpha\beta} ,$$

where $\Box \equiv \eta^{\mu\nu}\partial_\mu\partial_\nu$ is the wave operator in **flat** spacetime, and where

$$\tau^{\alpha\beta} = (-g)T^{\alpha\beta} + \frac{\Lambda^{\alpha\beta}}{16\pi G} .$$

The tensor $\Lambda^{\alpha\beta}$ encodes all the nonlinearities of general relativity, and can be written schematically

$$\Lambda^{\alpha\beta} = N^{\alpha\beta}[h,h] + M^{\alpha\beta}[h,h,h] + L^{\alpha\beta}[h,h,h,h] + \ldots$$

(i.e., the N term involves two fields of order $h^{\alpha\beta}$ coupling to one another; the M term involves three such fields; etc.). The exact detailed form of this term is given in the posted PDF file "Lecture 24, slide batch 2." The formal solution can be immediately written down using the radiative Green's function, which solves any sourced wave equation:

$$h^{\alpha\beta}(\mathbf{x},t) = -4G\int \frac{\tau^{\alpha\beta}(\mathbf{x}';t'-|\mathbf{x}-\mathbf{x}'|)}{|\mathbf{x}-\mathbf{x}'|}d^3x' .$$

Like many exact solutions, this may seem to be useless at first sight: it's actually an integro-differential equation for $h^{\alpha\beta}$, in which you need to know the solution in order to build the source which you must know to compute the solution. As detailed in the posted PDF slides, the secret to solving it is to realize that the RHS of this equation introduces a factor of G, which plays the role of a "small parameter" in defining an *iterative* solution. By writing

$$h^{\alpha\beta} = \sum_{n=1}^{\infty} G^n h_n^{\alpha\beta} ,$$

And inserting into this equation, one finds that that $h_1^{\alpha\beta}$ is simply the linear theory solution we worked out earlier (ie, it encodes the Newtonian limit), and that $h_n^{\alpha\beta}$ only depends on knowledge of $h_m^{\alpha\beta}$, where $m < n$.

This means that one can iteratively correct the spacetime, and iteratively build solutions describing the motion of (for example) two bodies in orbit about one another. For example, by taking things to order G^2 and developing the geodesic equation in this spacetime, one finds the acceleration a^i of body 1 in a binary system is given by

$$\begin{aligned} a_1^i &= -\frac{Gm_2 r_{12}^i}{r_{12}^3} \\ &+\frac{1}{c^2}\left\{\left[\frac{5G^2 m_1 m_2}{r_{12}^3}+\frac{4G^2 m_2^2}{r_{12}^3}+\frac{Gm_2}{r_{12}^3}\left(\frac{3}{2}\left(\frac{\mathbf{r}_{12}\cdot\mathbf{v}_2}{r_{12}}\right)^2 - v_1^2 + 4\left(\mathbf{v}_1\cdot\mathbf{v}_2\right) - 2v_2^2\right)\right] r_{12}^i + \right. \\ &\left. +\frac{Gm_2}{r_{12}^3}\left[4(\mathbf{r}_{12}\cdot\mathbf{v}_1) - 3(\mathbf{r}_{12}\cdot\mathbf{v}_2)\right]\left(v_1^i - v_2^i\right)\right\} . \end{aligned}$$

The first line is of course just the "normal" Newtonian acceleration of body 1 due to the gravitational attraction of body 2. The second line shows all the terms that are of order G^2 (or Gv^2, which by the virial theorem is of the same order) which correct Newton's gravity in this gauge. Here, r_{12}^i is the i component of the separation vector between the two bodies in deDonder gauge, and $\mathbf{v}_j$ gives the velocity $d\mathbf{r}_j/dt$ of body j in this coordinate system. One can continue, and in fact this expansion has been done so far to order G^8. The results for order G^3 and G^4 take a paragraph to write out; those beyond that become increasingly voluminous, filling multiple journal pages at order G^5 and higher.

It is then "simply" a matter of very careful[9] analysis to compute things like the gravitational waveforms such a binary generates, and the backreaction of gravitational waves on the binary's dynamics. Generically, this approach works well (meaning that the expansion has good convergence properties) as long as the separation of the members of the binary isn't too small — in other words, we require the typical separation r to be at many times larger than the gravitational lengthscale GM (or equivalently, that the typical orbit speed v be small compared to c). As we'll revisit briefly in the next lecture synopsis, post-Newtonian analyses have played an extremely important role in understanding binary dynamics in general relativity, and were in fact the only effective tool for modeling such systems (at least when the binary's members are of comparable mass) prior to some breakthroughs in computational relativity that occurred circa 2005.

Another approach is to do linear perturbation theory, but to linearize around an exact solution which describes a strong-field object in general relativity. We put $g_{\alpha\beta} = \hat{g}_{\alpha\beta} + \epsilon h_{\alpha\beta}$, where $\hat{g}_{\alpha\beta}$ is some exact solution, expand the EFEs to linear order in ϵ and then set $\epsilon = 1$ to develop equations describing the perturbation $h_{\alpha\beta}$. This technique is particularly well-developed when the exact solution describes a black hole, and is the foundation for black hole perturbation theory.

The posted notes and accompanying slides sketch how this can be done when the black hole "background" is taken to be a Schwarzschild black hole. By requiring that background plus perturbation be a valid solution of the vacuum EFE, one finds that the perturbation to Ricci must satisfy $\delta R_{\alpha\beta} = 0$. Further, one can organize the functional form of the perturbation by its properties with respect to rotations (the background is spherically symmetric), and by their parity properties. Focusing for now on odd-parity perturbations, one finds that *all* critical metric perturbations can be derived from a "master function" Q that is governed by the equation

$$\frac{\partial^2 Q}{\partial t*^2} - \frac{\partial^2 Q}{\partial r_*^2} + \left(1 - \frac{2GM}{r}\right)\left[\frac{\ell(\ell+1)}{r^2} - \frac{6GM}{r^3}\right] = 0 .$$

The integer ℓ is a spherical harmonic index, and reflects the fact that a spherical harmonic decomposition has been introduced. The coordinate

$$r_* = r + 2GM \ln\left[\frac{r}{2GM} - 1\right]$$

is called the *tortoise coordinate*. Notice that as $r \to \infty$, $r_* \to \infty$; but, as $r \to 2GM$, $r_* \to -\infty$. The tortoise coordinate is not very different from the Schwarzschild radial coordinate at large radius, but it puts the black hole event horizon infinitely far away (in coordinates! — not in proper distance, of course). The equation governing Q is called the Regge-Wheeler equation; a similar analysis done for even parity modes yields a similar (but somewhat messier) equation known as the Zerilli equation.

[9]The term "straightforward but tedious" is designed for such an analysis, which makes one praise computer algebra systems with great vigor.

In the limits $r \to \infty$ and $r \to 2GM$, the equation for Q has the asymptotic form

$$\frac{\partial^2 Q}{\partial t*^2} - \frac{\partial^2 Q}{\partial r_*^2} = 0 ,$$

which has the solutions $Q = \exp[-i\omega(t \pm r_*)]$. We expect the solution of the form $Q \propto \exp[-i\omega(t - r_*)] \equiv Q_{\rm out}$ to describe Q as $r \to \infty$, since this corresponds to purely outgoing radiation far from the black hole; likewise, we expect $Q \propto \exp[-i\omega(t + r_*)] \equiv Q_{\rm in}$ as $r \to 2GM$, since this corresponds to purely ingoing radiation coming into the black hole.

Of course, one can solve the equation for Q at all r. By enforcing as a boundary condition $Q \to Q_{\rm out}$ as $r \to \infty$ and $Q \to Q_{\rm in}$ as $r \to 2GM$, one finds that only certain values of ω "work." Such omegas are in general complex, and thus describe damped oscillations. These solutions are called *quasi-normal modes* of the black hole spacetime, and represent an oscillation of the black hole's geometry.

As described here, this exercise does not work very well for Kerr black holes because the equations one gets for Kerr metric perturbations are very difficult to work with. However, in something of a miracle, it turns out that one can make substantial progress by working with curvature perturbations.

You worked out some of the core details of this on a problem set. The idea is to take a derivative of the Bianchi identity to obtain a wave equation for the Riemann curvature with the schematic form $\Box R_{\alpha\beta\mu\nu} =$ terms involving Riemann squared (focusing on the vacuum case, so that the stress-energy and Ricci tensors vanish). By writing $R_{\alpha\beta\mu\nu} = \hat{R}_{\alpha\beta\mu\nu} + \delta R_{\alpha\beta\mu\nu}$ (where $\hat{R}_{\alpha\beta\mu\nu}$ is the Riemann curvature of the background Kerr black hole, choosing a set of basis vectors adapted to radiative and non-radiative degrees of freedom, and then projecting Riemann onto appropriate combinations of these basis vectors, one can pick out components of curvature which describe radiation, and ones which describe non-radiative aspects of the spacetime. The components which describe radiation in this spacetime turn out to be governed by a fairly simple differential equation. This analysis, for example, yields a generalization of the quasi-normal modes described above, which shows that the modes have a frequency and a damping time which depend upon and thus encode a black hole's mass and spin.

One does not need to focus on purely vacuum solutions. By imagining that the wave equation has a source, one can build solutions which describe how a black hole's spacetime is modified by orbiting matter. A particularly fruitful direction has been to take the source to be a mass μ that is much less massive than the black hole, $\mu \ll M$, and to place it is in orbit about the black hole. One can then construct solutions to the perturbed geometry which describe the binary. This method complements post-Newtonian theory, as it does not restrict the class of orbits that can be considered, as long as the mass ratio is large enough ($\mu/M \ll 1$) that the perturbative expansion is valid. This also turns out to provide a very accurate representation of binary sources with large mass ratios, a limit that is in fact of great astrophysical interest.

Lecture 25 (not video recorded): What do we do when our spacetime has no particular symmetry, and no obvious "small parameter" by which one can expand around some exact background solution? In other fields of analysis, the answer is simple: numerically solve the equations which govern the system. For systems of partial differential equations, this typically involves some scheme for discretizing the time and space representation of all quantities under study. For example, one replaces derivatives by finite difference approximants:

$$\partial_x f(t, x, y, z) \to \frac{f(x_{i+1}, y_j, z_k, t_n) - f(x_{i-1}, y_j, z_k, t_n)}{x_{i+1} - x_{i-1}} .$$

The differential equations relating fields at different points in space and different moments in time become algebraic relations which can be solved on a computer. Provided that one can come up with an accurate scheme for solving these equations, and that one has the computational resources needed, it is then straightforward in principle (though often *very* challenging in practice) to model your system. Certain purists[10] sometimes sneer at such computational approaches. Such purists are not just wrong, but deluded. A tremendous number of important problems that are described by nonlinear systems of equations must be modeled numerically, including (for example) fluid dynamics and plasma problems.

Complicating the problem in relativity is the principle of covariance: we have tremendous freedom to choose how to divide "spacetime" into "space" and "time." How do we "integrate forward in time" when the notion

[10]Including past 8.962 students, one of whom told me that no "real" physicists use numerical methods.

of "time" is not unique? On a past problem set, you showed that this is possible in principle: the contacted Bianchi identity[11], $\nabla_a G^{ab} = 0$ becomes

$$\nabla_0 G^{0b} = -\nabla_i G^{ib} \ .$$

The terms on the right-hand side have at most two time derivatives in them. Since the operator ∇_0 explicitly contains a time derivative, we know that G^{0b} has at most one time derivative. This shows us that the Einstein field equations split into two "flavors":

$$\begin{aligned} G^{0b} &= 8\pi T^{0b} \qquad \text{"Constraint" equations ;} \\ G^{ij} &= 8\pi T^{ij} \qquad \text{"Evolution" equations .} \end{aligned}$$

The constraint equations have at most one derivative of the metric in them; we can think of them as tell us how the geometry and its first derivative behave as a function of "space" at a single moment of "time." These equations play a mathematical role similar to the divergence Maxwell equations. The evolution equations have two time derivatives, and tell us how the geometry evolves from "moment to moment." These equations are akin to Maxwell's curl equations.

To proceed, we need to break the beautiful covariance of spacetime and pick what is time and what is space. However, we wish to use the powerful mathematical machinery of tensor analysis, so we wish to do by formulating tensor equations that are adapted to the time and space coordinates that we choose. Breaking spacetime into space and time is described on pages 2 and 3 of lecture 25, as well as slides 2, 3, and 4 of the accompanying slides (labeled "Lecture 26, slide batch 2"). The basic idea is to imagine two nearby "slices" of spacetime. Each corresponds to a particular moment of the time coordinate you have choose; the first is at t, the second at $t + dt$. Imagine that event A is at spatial coordinate x^i in the t slice, and that event B is at spatial coordinate $x^i + dx^i$ in the $t + dt$ slice. At each event, one can define[12] a vector n^a which points normal to the event's timeslice.

Suppose you are sitting at event A and move along the normal from slice t to slice $t + dt$. You do not necessarily find yourself at coordinate x^i in the new slice, but you may be *shifted* by an amount $-\beta^i\,dt$ from that coordinate. In addition, the amount of proper time you experience in moving along this shift defined to be $d\tau = \alpha\,dt$. The quantities α and β^i are called the *lapse* and the *shift*, respectively. The lapse gives us freedom to "run time" at different speeds in different parts of our manifold (so that clocks at the edge of our grid, presumably far from any source, run fast compared to those which may be close to the source and experiencing gravitational redshift). The shift allows us slide our spatial coordinates around in each slice in a way that is most convenient to our analysis. For example, if an object is undergoing gravitational collapse, it may be useful to have our coordinates densely packed near the collapsing matter, but leave things more sparse far away (similarly to how the tortoise coordinate r_* used in black hole perturbation theory becomes infinitely dense compared to r near an event horizon). The lapse and shift can be freely chosen, and can be regarded as a generalized notion of "gauge choice" for strong-field relativity. With them chosen, the spacetime interval between events A and B is

$$ds^2 = -\alpha^2 dt + \gamma_{ij}(dx^i + \beta^i dt)(dx^j + \beta^j dt) \ ,$$

where γ_{ij} is the 3-dimensional metric in a given constant time slice.

Pages 4, 5, and 6 of the posted notes and slides 4 and 5 of the accompanying slide deck define these quantities more precisely. Key concepts are to regard the time coordinate t as level surfaces of some scalar field that fills all of spacetime, and to define n^a as the normal to these level surfaces. The quantity $\gamma_{ab} = g_{ab} + n_a n_b$ is a projection tensor; the field $[T^a{}_b]_{\text{in slice}} = \gamma^a{}_c \gamma^d{}_b T^c{}_d$ is defined only in the time slice to which n^a is the normal. This tells us that γ_{ab} is the metric in that slice (and that with the coordinate properly chosen, only the spatial components γ_{ij} are non zero). We denote by D_a the covariant derivative associated with this spatial metric.

We next need to develop all the curvature tensors in this language. First is the "in-slice" Riemann tensor. This is actually quite easy: it's just the normal Riemann tensor, but built up using the metric γ_{ab}. We call this tensor, $R^a{}_{bcd}$ the *intrinsic* curvature associated with a given time slice.

---

[11] In this lecture, we follow much of the numerical relativity literature and use what is sometimes called the "Fortran convention." All indices are written as lower-case latin letters; spacetime indices are designated with the set $(a, b, \ldots, h)$; space indices are designated with (i, j, k, l, m). If you understand why this is called the Fortran convention, you are either way older than the typical 8.962 student, or you have been cursed to work with archaic code and have my condolences.

[12] The index on this vector is incorrectly written i in the lecture notes. With a little thought, you can see it must point in the timelike direction, so we should use the spacetime index label rather than a spatial index label.

The second contribution is more subtle. We have tremendous freedom to select each time slice. If we picture spacetime as a 4-dimensional "slab," we can make each 3-dimensional time slice flat or wiggly depending upon how we choose to chop up that original slab. The *extrinsic* curvature is the curvature in each slice that arises from how we decided to "cut" each of these slices. Some intuition comes from thinking about the surface of a cylinder. A cylinder's surface is intrinsically flat: Two geodesic trajectories on it which start out parallel will stay parallel forever. However, it's also clearly round in an intuitive sense. This roundness comes from how we embed this 2-dimensional surface into 3-dimensional space.

The extrinsic curvature is quantified by examining how the normal vectors expand or diverge as we move from timeslice to timeslice. We define the extrinsic curvature tensor as

$$K_{ab} = -\gamma^c{}_a \gamma^d{}_b \nabla_c n_d \ .$$

By piecing together various definitions which are given in the notes, one can show that this is equivalent to a Lie derivative of γ_{ab} along the normal direction:

$$K_{ab} = -\frac{1}{2}\mathcal{L}_{\vec{n}}\gamma_{ab} \ .$$

Since $\vec{n}$ points from slice to slice, one can intuitively regard K_{ab} as a kind of first time derivative of the spatial geometry. The fields γ_{ab} and K_{ab} completely describe the geometry of space at any given moment.

Once one knows the extrinsic and intrinsic curvature, it becomes possible to build the 4-dimensional curvature tensor ${}^{(4)}R^a{}_{bcd}$ from them. The results, which I leave in the notes, are known as *Gauss's equation* (which tells us about ${}^{(4)}R^a{}_{bcd}$ with all four indices projected into a spatial slice), the *Codazzi equation* (which tells us about ${}^{(4)}R^a{}_{bcd}$ with three indices projected spacelike), and *Ricci's equation* (${}^{(4)}R^a{}_{bcd}$ with two indices projected spacelike). Thanks to Riemann's symmetries, this completely characterizes this curvature tensor.

This gives us everything we need to build the Einstein field equations, ${}^{(4)}G_{ab} = 8\pi G\,{}^{(4)}T_{ab}$. We break it up into 3 pieces, depending on how many components are parallel or perpendicular to the normal n^a:

$$n^a n^b\,{}^{(4)}G_{ab} = 8\pi G\,{}^{(4)}T_{ab} n^a n^b \equiv 8\pi G\rho$$

becomes

$$R + K^2 - K_{ab}K^{ab} = 16\pi G\,\rho \ .$$

Here, $K = \gamma^{ab}K_{ab}$ is the trace of the extrinsic curvature, and ρ is the energy density in the spacetime as measured by a "normal" observer (i.e., an observer who moves along the normal to the slice). This equation is known as the *Hamiltonian constraint*, and is the fully tensorial version of the relation $G^{00} = 8\pi G T^{00}$. Next we look at $\gamma^b{}_a n^c$ acting on the Einstein field equation, yielding

$$D_b K^b{}_a - D_a K = 8\pi G\, j_a$$

where $j_a = -\gamma^b{}_a n^c\,{}^{(4)}T_{bc}$ is the momentum density measured by a normal observer. This is called the *momentum constraint*.

The final Einstein equations come from making two spatial projections. Before doing so, we define the "time direction":

$$t^a = \alpha n^a + \beta^a \ ,$$

An observer who moves along n^a is an "Eulerian observer," who remains at rest in a time slice; an observer who moves along t^a is a "coordinate observer," who slices along the grid maintaining a constant (spatial) coordinate position x^i (even if that coordinate's position is changing from slice to slice). The result turns out to be

$$\begin{aligned}\mathcal{L}_{\vec{t}}K_{ab} &= -D_a D_b\alpha + \alpha\left(R_{ab} - 2K_{ac}K^c{}_b + KK_{ab}\right)\\ &\quad -8\pi G\alpha\left(S_{ab} - \frac{1}{2}\gamma_{ab}(S-\rho)\right) + \mathcal{L}_{\vec{\beta}}K_{ab} \ .\end{aligned}$$

(where $S_{ab} \equiv \gamma^c{}_a\gamma^d{}_b{}^{(4)}T_{cd}$, and $S = \gamma^{ab}S_{ab}$). This is the *evolution equation*. Notice that it is effectively a time derivative of the extrinsic curvature. Since the extrinsic curvature is a kind of time-derivative of the spatial geometry, this equation provides, in fully tensorial form, a relationship for the second derivative of the spatial geometry.

Formally, this solves the problem. These systems of equations can be used to prove theorems on the existence of solutions to Einstein's equations. For example, given an initial geometry in our spacetime manifold, these equations tell us how to build the geometry at later times.

Some lingering concerns remain. For example, the equations predict the existence of singularities — points in the manifold where the geometry becomes ill-behaved, and beyond which the equations cannot be integrated. However, in all known generic cases, these singularities have been found to be hidden behind an event horizon. As such, singular parts of the manifold are removed from causal contact with the rest of the manifold, and are thereby rendered harmless.

It is not known whether this is a generic feature of general relativity. The hypothesis that singularities which form from the collapse of non-singular initial conditions are always hidden by event horizons is known as the "Cosmic Censorship Conjecture." One counterexample, which requires extreme fine tuning of initial conditions, is known[13]. This fine-tuned case leads to the formation of a strange singularity, a structure of zero mass but infinite tidal stresses exactly at the structure. No counterexample is known which forms from a distribution which is not extraordinarily fine tuned. Whether such an example will be found has been the subject of moderately famous bets by moderately famous scientists being moderately silly[14].

Practically, there is still a tremendous amount to be done. The exercise described above is what one needs to do to *set up* the problem of numerically computing a spacetime. In practice, implementing this proved to be very challenging. For decades, only highly constrained problems (typically of reduced dimension, with symmetries imposed) could be solved. Whenever one tried to evolve a generic case, numerical instabilities, seeded by discretization and round-off error and magnified by the nonlinear nature of the equations, grew in an unbounded fashion. An example discussed in the posted slides shows that even the "easy" cases didn't really work. Imagine starting with initial data describing a static black hole doing absolutely nothing. Formally, we know that nothing should happen: it should just sit there. In 1995, one found that the equations describing the "evolution" of this system became artificially dynamical due to numerical error. The analysis code crashed after a typical time interval of $t \sim 50GM$. Highly asymmetric and dynamical problems were even worse.

This all changed, rather dramatically, in late 2004 and 2005. In that year, various groups found ways of representing the data (essentially, "good" choices of lapse and shift) which kept numerical instabilities under control. It must be noted that the decades of frustrating challenges produced many good ideas which helped tremendously once good gauge conditions were known. Your lecturer remembers in the span of one year that leading senior people in the field were beginning to discuss whether one should give up on the field, to suddenly beginning to mass produce astrophysically important simulations of interesting strong-gravity binary dynamics. The field has now reached the point where for many problems it is practically a matter of engineering. Computing the dynamics of binary black holes, for example (two black holes which orbit one another, generating strong gravitational waves which backreact on the system and drive them to merge into a single object), can be done so well that computational models play a large role informing the analysis of data from gravitational-wave detectors like LIGO and Virgo.

[13]See arXiv:9612015 for discussion

[14]http://www.theory.caltech.edu/people/preskill/bets.html

MIT OpenCourseWare
https://ocw.mit.edu

8.962 General Relativity
Spring 2020

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