ZHANPEI FANG

A BRIEF SKETCH OF RELATIVITY

PREPARED FOR THE SUMMER SCIENCE PROGRAM, JULY 2020

0.1 Introduction

These lecture notes provide a glancing survey of the basic mathematical framework of special and general relativity. I will try to keep things relatively light on the mathematical side for conceptual clarity, and avoid too much historical detail, though the historical development of the theory is of definite intellectual interest.

The sources I consulted can be found listed at the end of these lecture notes. I have basically cribbed all of Carroll's notes in "A No-Nonsense Introduction to General Relativity", also referencing his textbook *Spacetime and Geometry*, and I will be using his notational conventions, such as letting the speed of light $c = 1$, and in particular his sign convention for the metric. So this really cannot be considered my own work. I have done my best to make the material palatable and readily digestible to somebody with minimal prior exposure to things such as tensor calculus; I am not sure if I've succeeded in this regard. I would additionally encourage you to read Wikipedia's non-technical introduction to the subject, which can be found at `https://en.wikipedia.org/wiki/Introduction_to_general_relativity`.

1
Special relativity

1.1 Space and time

All of the strangeness of special relativity arises from the invariance of the speed of light c in all reference frames. This leads to our considering the notion of a unified **spacetime**, for which we want to be able to talk about space and time coordinates in a mathematical sense.

The **spacetime interval** is given by $\Delta s^2 = -\Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2$, and is what we use when we talk about the "distance" between two "things"; it is a "distance" measure which also includes information on how events are separated in time. Instead of using t's and x, y, z's however, which can get kind of clunky later on, we can instead write all of the space and time quantities (the *coordinates* of spacetime) in a single vector[1]. For Cartesian coordinates:

$$x^\mu \begin{cases} x^0 \equiv ct = t \\ x^1 = x \\ x^2 = y \\ x^3 = z \end{cases} \tag{1.1}$$

[1] Yes, in this lecture you're going to see why index notation is *extremely* useful. Note that the superscripts here are *indices* and not exponents.

The collection of all four coordinates is denoted as x^μ; since this vector has four components we call it a **4-vector**. So our usual space components x^i are given by:[2]

$$x^i \begin{cases} x^1 = x \\ x^2 = y \\ x^3 = z \end{cases} \tag{1.2}$$

[2] Convention has it that we use Latin indices as free/dummy indices when we're just talking about the space components $i \in \{1, 2, 3\}$, and Greek letters when we're talking about both space and time components, for example $\mu \in \{0, 1, 2, 3\}$.

Special relativity is set within the **Minkowski spacetime**, which combines 3-dimensional space and time into a 4-dimensional manifold[3] and contains all information about the geometry of the manifold; we say that the x^μ are coordinates on this manifold. The **Minkowski met-**

[3] Informally, a manifold is a "possibly curved space which in small enough (infinitesimal) regions looks like flat space."

ric is given as a matrix:

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.3}$$

and so, moving away from the discrete form Δs^2, the spacetime interval can be succinctly written using the metric as[4]

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu = -dt^2 + dx^2 + dy^2 + dz^2 \tag{1.4}$$

[4] An equation of this form is also, unfortunately, often called "the metric", particularly in cosmology; for example the Friedmann–Lemaître–Robertson–Walker metric (or FLRW for short).

In this I have introduced our implementation of the dot product:

$$A \cdot B \equiv \eta_{\mu\nu} A^\mu B^\nu = -A^0 B^0 + A^1 B^1 + A^2 B^2 + A^3 B^3 \tag{1.5}$$

The Minkowski metric has the Lorentzian **signature** $(-+++)$. This is why the interval elapsed as a particle fixed in space moves forward in time is actually negative: $ds^2 = -dt^2 < 0$. The **proper time** is the time measured by a clock following a timelike worldline in the metric, $d\tau^2 \equiv -ds^2$, and the discrete proper time interval between any two events is given by $(\Delta\tau)^2 = -(\Delta s)^2 = -\eta_{\mu\nu}\Delta x^\mu \Delta x^\nu$. The proper time elapsed along a trajectory through spacetime will be the actual time measured by an observer on that trajectory.

$$\tau = \int \sqrt{-ds^2} = \int \sqrt{-\eta_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda \tag{1.6}$$

Finally, the 3D Cartesian space we typically consider in classical mechanics (and during SSP), we discard the time component and the interval is given simply by $ds^2 = (\mathrm{d}x)^2 + (\mathrm{d}y)^2 + (\mathrm{d}z)^2 = \mathrm{d}r^2 + r^2\mathrm{d}\theta^2 + r^2 \sin^2\theta \mathrm{d}\phi^2$, where the second equality is for a spherical coordinate system like the ones we've been considering throughout the program.

A little more on 4-vectors

Vectors in spacetime are always fixed at a particular **event** in our spacetime. A four-vector V^μ is known as a **timelike** vector if it has negative norm, **null** vector if the norm is zero, and **spacelike** if the norm is positive.

$$\eta_{\mu\nu} V^\mu V^\nu \begin{cases} < 0 & V^\mu \text{ timelike} \\ = 0 & V^\mu \text{ lightlike or null} \\ > 0 & V^\mu \text{ spacelike} \end{cases} \tag{1.7}$$

Similarly, trajectories with $ds^2 < 0$ are timelike, $ds^2 = 0$ null, and $ds^2 > 0$ spacelike. The set of null trajectories in/out of an event are called its **light cone** (Fig. 1.1).

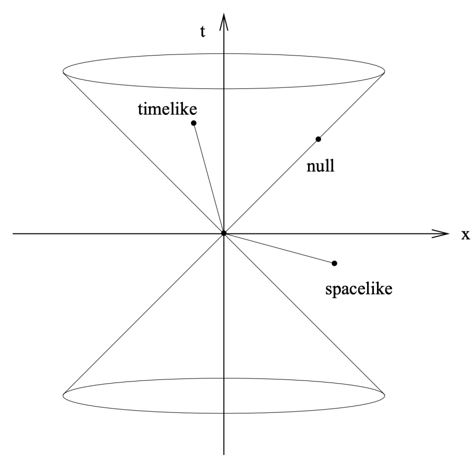

Figure 1.1: A lightcone as it appears on a spacetime diagram: a single space coordinate x on the horizontal, time coordinate t on the vertical axis. [c/o Carroll]

A path through spacetime is specified by giving the four spacetime coordinates as a function of some parameter λ, so $x^\mu(\lambda)$. A path is timelike/null/spacelike if its tangent vector $dx^\mu/d\lambda$ is timelike/null/spacelike. The tangent vector $dx^\mu/d\tau = U^\mu$ is the **4-velocity**, and the **4-momentum** is $p^\mu = mU^\mu$, with m the mass of the particle.

1.2 *Lorentz transformations*

In an introductory course in special relativity,[5] you might see the Lorentz transformation written as, for a boost in the $+x$-direction:

$$\begin{aligned} x' &= \gamma(x - vt) \\ y' &= y \\ z' &= z \\ t' &= \gamma(t - vx) \end{aligned}$$

where γ is the Lorentz factor[6] $\gamma = 1/\sqrt{1-v^2}$. The Lorentz factor is the factor by which time, length and relativistic mass change for an object in motion.

However, the above transformation can be more succinctly written as:

$$x^{\mu'} = \Lambda^{\mu'}{}_\nu x^\nu \tag{1.8}$$

where the matrix

$$\Lambda^{\mu'}{}_\nu = \begin{pmatrix} \cosh\phi & -\sinh\phi & 0 & 0 \\ -\sinh\phi & \cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.9}$$

for the $+x$-boost. Another example of a transformation matrix is rotation in the xy plane:

$$\Lambda^{\mu'}{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.10}$$

In general, Lorentz transformations are *defined* as those matrices Λ which satisfy the equation $\eta = \Lambda^T \eta \Lambda$, or, writing all the indices out, $\eta_{\rho\sigma} = \Lambda^{\mu'}{}_\rho \eta_{\mu'\nu'} \Lambda^{\nu'}{}_\sigma = \Lambda^{\mu'}{}_\rho \Lambda^{\nu'}{}_\sigma \eta_{\mu'\nu'}$. When we say that a quantity or physical law is **invariant** under Lorentz transformation, we mean that said thing does not change when Λ is applied to it.

And that's probably all you need to know about special relativity for right now.[7]

[5] I don't know if they still do this, but when I was a freshman in the Physics 60 series at Stanford, we *began* with special relativity and then did your typical classical mechanics with free-body diagrams, blocks on inclines, rolling without slipping, etc. Talk about a trial by fire!

[6] There's a hidden speed-of-light c in here—you'll typically first see the Lorentz factor first written as $\gamma = 1/\sqrt{1-v^2/c^2} = 1/\sqrt{1-\beta^2}$. But again, I'm suppressing all c's.

[7] Apologies for not going deeper into the typical thought experiments—clocks on a spaceship for time dilation, pole-barn paradox for length contraction, the twin paradox and so on. Feel free to look those up on your own time! They are essential for getting an intuition for the 'weirdness' of special relativity.

2

Some formalism

Some more proper mathematical formalisms before we get any further into the mathematical weeds of GR, which is on some level an exercise in applied differential geometry.

2.1 *Vectors*

Suppose that for each tangent space we set up a basis of four vectors $\hat{e}_{(\mu)}$, $\mu \in \{0, 1, 2, 3\}$. Then any abstract vector A can be written as a linear combination of basis vectors:

$$A = A^{\mu}\hat{e}_{(\mu)} \tag{2.1}$$

The coefficients A^{μ} are the components of the vector A, and often we refer to them interchangeably (ignoring the basis and saying things like "the vector A^{μ}".)

A standard example of a vector in spacetime is the **tangent vector** to a curve; as mentioned earlier, a **parameterized curve** through spacetime is specified by the coordinates as a function of the parameter, for example $x^{\mu}(\lambda)$. The tangent vector $V(\lambda)$ has components

$$V^{\mu} = \frac{dx^{\mu}}{d\lambda} \tag{2.2}$$

In this case, the entire vector is formally understood as $V = V^{\mu}\hat{e}_{(\mu)}$. Under a Lorentz transformation the coordinates x^{μ} change according to Eq. (1.8), while the parameterization λ doesn't change; so the components of the tangent vector change as $V^{\mu} \to V^{\mu'} = \Lambda^{\mu'}{}_{\nu}V^{\nu}$. However, this is just for V's components in some coordinate system; the vector V itself is invariant under Lorentz transformation.

Basis vectors transform as follows. The old basis $\hat{e}_{(\mu)}$ transforms into the new one $\hat{e}_{(\nu')}$ by multiplying by the Lorentz transformation:

$$\hat{e}_{(\nu')} = \Lambda^{\mu}{}_{\nu'}\hat{e}_{(\mu)} \tag{2.3}$$

Dual vectors (one-forms)

Once we have set up a tangent vector space T_p, there is an associated space known as the **dual vector space**, usually denoted by an asterisk as T_p^*. Every dual vector (also known as a one-form) can be written in terms of its components $\omega = \omega_\mu \hat{\theta}^{(\mu)}$ and transforms like $\omega_{\mu'} = \Lambda^\nu{}_{\mu'} \omega_\nu$. Like with vectors, we typically write ω_μ to represent the entire dual vector.

Now is also probably a good time to define contravariant and covariant vectors, and explain the upper/lower index notation I've used extremely liberally thus far. Elements of T_p ("vectors") are referred to as **contravariant** and are the ones with upper indices, and elements of T_p^* ("dual vectors") are referred to as **covariant** and have their indices down below as subscripts.

In spacetime the simplest example of a dual vector is the **gradient** of a scalar function, the set of partial derivatives with respect to the spacetime coordinates.

$$d\phi = \frac{\partial \phi}{\partial x^\mu} \hat{\theta}^{(\mu)} \tag{2.4}$$

The components of dual vectors transform like

$$\frac{\partial \phi}{\partial x^{\mu'}} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial \phi}{\partial x^\mu} = \Lambda^\mu{}_{\mu'} \frac{\partial \phi}{\partial x^\mu} \tag{2.5}$$

Because the gradient is a dual vector, we can write the partial derivative in this shorthand, using commas:

$$\frac{\partial \phi}{\partial x^\mu} = \partial_\mu \phi = \phi_{,\mu} \tag{2.6}$$

The gradient of a tangent vector to a curve is just the ordinary derivative of the function along the curve, as one might intuit.

$$\partial_\mu \phi \frac{\partial x^\mu}{\partial \lambda} = \frac{d\phi}{d\lambda} \tag{2.7}$$

2.2 *Tensors*

As we make our transition from flat to curved spacetime, we can no longer use our trusty old Cartesian coordinate system; we are motivated to make our equations *coordinate-invariant*. Tensors help us write equations in a coordinate-invariant form; they can be understood as like a vector, just with more indices. More abstractly, **tensors** can be *defined* as those objects which transform under a change of coordinates $x^\mu \to x^{\mu'}$ as tensors. For example, this is a tensor:

$$S^{\mu'}{}_{\nu'\rho'} = \frac{\partial x^{\mu'}}{\partial x^\mu} \frac{\partial x^\nu}{\partial x^{\nu'}} \frac{\partial x^\rho}{\partial x^{\rho'}} S^\mu{}_{\nu\rho} \tag{2.8}$$

The unprimed indices in the RHS are the dummy indices we're summing over, so that we can convert everything to the primed coordinate system.

GR has no preferred coordinate system: if an equation relating two tensors holds in one coordinate system, it holds in *all* coordinate systems. As with vectors, the upper indices are contravariant, and the lower indices are covariant.

Again, the index notation is really quite powerful. For instance, all of classical electromagnetism can be summarized in just two lines:

$$\partial_\mu F^{\nu\mu} = J^\nu \tag{2.9}$$

$$\partial_{[\mu} F_{\nu\lambda]} = \partial_m u F_{\mu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = 0 \tag{2.10}$$

where $F_{\mu\nu}$ is the electromagnetic field strength tensor:

$$F_{\mu\nu} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix} = -F_{\mu\nu} \tag{2.11}$$

and $J^\nu = (\rho, \mathbf{J})$ is the four-current. This formulation of E&M turns out to hold in curved space so long as we replace the partial derivatives ∂_μ by covariant derivatives ∇_μ (defined later).

Manipulating tensors

An (n, m) tensor (that is, one with n upper and m lower indices) can be **contracted** to form a $(n-1, m-1)$ tensor by summing over one upper and lower index:

$$S^\mu = T^{\mu\lambda}{}_\lambda \tag{2.12}$$

The contraction of a 2-index tensor is called the **trace**. A tensor is **symmetric** in two indices if we can interchange them without changing the tensor, $S_{\dots\alpha\beta\dots} = S_{\dots\beta\alpha\dots}$, and **antisymmetric** if it changes sign, $S_{\dots\alpha\beta\dots} = -S_{\dots\beta\alpha\dots}$. For an arbitrary tensor (without any particular symmetry properties), we can pick out the antisymmetric and symmetric pieces by taking appropriate linear combinations.[1]

[1] Included for completeness; *please* don't worry about this too much.

$$T_{(\mu_1\mu_2\dots\mu_n)} = \frac{1}{n!}(T_{\mu_1\mu_2\dots\mu_n} + \text{sum over permutations of } \mu_1\dots\mu_n)$$

$$T_{[\mu_1\mu_2\dots\mu_n]} = \frac{1}{n!}(T_{\mu_1\mu_2\dots\mu_n} + \text{alternating sum over permutations } \mu_1\dots\mu_n)$$

The metric tensor

The most important tensor in GR is the **metric tensor** $g_{\mu\nu}$, which generalizes the Minkowski metric $\eta_{\mu\nu}$ we met before. As in Minkowski space, we can use the metric to raise or lower indices:

$$A_\mu \equiv g_{\mu\nu} A^\mu, A^\mu = g^{\mu\nu} A_\nu \tag{2.13}$$

and take dot products:

$$A \cdot B \equiv g_{\mu\nu} A^\mu B^\nu = A_\nu B^\nu \tag{2.14}$$

We can define the inverse metric $g^{\mu\nu}$ (already spoiler-ed a few lines above) as the matrix inverse of the metric tensor:

$$g^{\mu\nu} g_{\nu\rho} = \delta^\mu_\rho \tag{2.15}$$

where δ^μ_ρ is the Kronecker delta in this spacetime (!!).

2.3 *Curvature*

Covariant derivatives

The partial derivative (Eq. 2.6) is, unfortunately, not a tensor. Transforming the partial derivative of a scalar returns a reasonable result (a $(0, 1)$ tensor):

$$\partial_\mu \phi \to \partial_{\mu'} \phi = \frac{\partial x^\mu}{\partial x^{\mu'}} \partial_\mu \phi \tag{2.16}$$

but if we try the same move on a vector, using $V^\mu \to \frac{\partial x^{\mu'}}{\partial x^\mu} V^\mu$:

$$\begin{aligned} \partial_\mu V^\nu \to \partial_{\mu'} V^{\nu'} &= \left(\frac{\partial x^\mu}{\partial x^{\mu'}} \partial_\mu \right) \left(\frac{\partial x^{\nu'}}{\partial x^\nu} V^\nu \right) \\ &= \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^{\nu'}}{\partial x^\nu} (\partial_\mu V^\nu) + \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial^2 x^{\nu'}}{\partial x^\nu \partial x^\mu} V^\mu \end{aligned}$$

which has an extra second term we don't want. Instead we need to define a **covariant derivative** to be a partial derivative plus a correction which is linear in the original tensor:

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda \tag{2.17}$$

where the new symbol I've introduced $\Gamma^\nu_{\mu\lambda}$ stands for a collection of numbers called **connection coefficients**, which transform not as tensors but as

$$\Gamma^{\nu'}_{\mu'\lambda'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\lambda}{\partial x^{\lambda'}} \frac{\partial x^{\nu'}}{\partial x^\nu} \Gamma^\nu_{\mu\lambda} - \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\lambda}{\partial x^{\lambda'}} \frac{\partial^2 x^{\nu'}}{\partial x^\mu \partial x^\lambda} \tag{2.18}$$

If we set the transform of $\Gamma^\nu_{\mu\lambda}$ as such, $\nabla_\mu V^\nu$ is guaranteed to transform like a tensor. Covariant derivatives of tensors with lowered indices are defined in a similar manner; for a 1-form,

$$\nabla_\mu \omega_\nu = \partial_\mu \omega_\nu - \Gamma^\lambda_{\mu\nu} \omega_\lambda \tag{2.19}$$

Let's return to these connection coefficients I introduced. The connection coefficients of the Levi-Civita connection (or pseudo-Riemannian connection) expressed in a coordinate basis are also called **Christoffel symbols**; feel free to use the terms interchangeably for now.[2] The Christoffel symbol is symmetric in its lower indices:

$$\Gamma^{\sigma}_{\mu\nu} = \frac{1}{2} g^{\sigma\rho} (\partial_\mu g_{\nu\rho} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu}) \tag{2.20}$$

[2] The third name for these objects is the *Levi-Civita* connection coefficients. :)

Using these, we have that the covariant derivative of the metric and its inverse are always zero, an incredibly useful behavior known as **metric compatibility**:

$$\nabla_\sigma g_{\mu\nu} = 0, \nabla_\sigma g^{\mu\nu} = 0 \tag{2.21}$$

If there exists a metric-compatible and torsion-free ($\Gamma^{\lambda}_{\mu\nu} = \Gamma^{\lambda}_{(\mu\nu)}$) connection, it must be of the form of a Christoffel symbol.[3]

[3] Luckily, there are ways to automate the calculating of Christoffel symbols and other quantities; see Mathematica notebook.

The Riemann curvature tensor

I briefly gave a definition of a manifold in the footnote on page 3, but let's talk about it more. One example of a manifold could be the Earth, which is so large that we (puny) observers may reasonably assume that it's flat.[4] Some other examples of manifolds are n-dimensional flat space $\mathbb{R}^n$, or the n-dimensional sphere S^n. $\mathbb{R}^1$ is the real line, $\mathbb{R}^2$ is a plane, etc.; S^1 is a circle, S^2 a sphere, etc.[5]

[4] [Insert flat-Earther joke of your choice.]

[5] For example, if we use θ and ϕ as coordinates for a sphere S^2 with radius $r = 1$, the metric is $ds^2 = d\theta^2 + \sin^2\theta d\phi^2$.

Manifolds can be (and often are) curved, which means we need to whip out some of the formalisms we've already laid out in the preceding pages for a non-Cartesian basis. As I said on page 3, all "information" about the curvature of a manifold is contained in the metric. How do we get this information back? Turns out, the information about curvature is contained in a four-component tensor which we call the **Riemann curvature tensor**:

$$R^{\rho}{}_{\sigma\mu\nu} \equiv \partial_\mu \Gamma^{\rho}_{\nu\sigma} - \partial_\nu \Gamma^{\rho}_{\mu\sigma} + \Gamma^{\rho}_{\mu\lambda}\Gamma^{\lambda}_{\nu\sigma} - \Gamma^{\rho}_{\nu\lambda}\Gamma^{\lambda}_{\mu\sigma} \tag{2.22}$$

This tensor is built up from non-tensors (partial derivatives and Christoffel symbols) but miraculously transforms as a tensor. Importantly, the Riemann tensor works well as a measure of curvature because *all of its components vanish iff the space is flat*; "flat" meaning that $\exists$ a global coordinate system in which the metric components are everywhere constant. We can give names to two particularly useful contractions of the Riemann tensor: the **Ricci tensor** is given by $R_{\mu\nu} = R^{\lambda}{}_{\mu\lambda\nu}$, and the **Ricci scalar** by $R = R^{\mu}{}_{\mu} = g^{\mu\nu} R_{\mu\nu}$.

While there are many possible components of the Riemann tensor due to its many indices, it turns out that its symmetry properties simplify calculations a fair bit. In particular:

$$\begin{aligned} R_{\rho\sigma\mu\nu} &= -R_{\sigma\rho\mu\nu} = -R_{\rho\sigma\nu\mu} \\ R_{\rho\sigma\mu\nu} &= R_{\mu\nu\rho\sigma} \\ R_{\rho[\sigma\mu\nu]} &= R_{\rho\sigma\mu\nu} + R_{\rho\mu\nu\sigma} + R_{\rho\nu\sigma\mu} = 0 \\ R_{[\rho\sigma\mu\nu]} &= 0 \end{aligned}$$

This means that the Ricci tensor is symmetric, $R_{\mu\nu} = R_{\nu\mu}$. Additionally the Riemann tensor behaves according to a differential identity called the **Bianchi identity**:

$$\nabla_{[\lambda} R_{\rho\sigma]\mu\nu} = 0 \tag{2.23}$$

If we define a new tensor, called the **Einstein tensor** ('Einstein' denoting its importance for reasons which will soon become clear!),

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \tag{2.24}$$

the Bianchi identity implies that the divergence of this tensor vanishes.

$$\nabla^{\mu} R_{\rho\mu} = \frac{1}{2} \nabla_{\rho} R \Longleftrightarrow \nabla^{\mu} G_{\mu\nu} = 0 \tag{2.25}$$

Parallel transport and geodesics

To be brief and informal, a geodesic is "the shortest path between two points."[6] More formally, the geodesic extremizes the length functional $\int ds$. For a path $x^{\mu}(\lambda)$, the infinitesimal distance along the curve is given by

$$ds = \sqrt{\left| g_{\mu\nu} \frac{dx^{\mu}}{d\lambda} \frac{dx^{\nu}}{d\lambda} \right|} d\lambda \tag{2.26}$$

which, if integrated, gives the full length of the curve $L = \int ds$. Finding extrema[7] of L requires calculus of variations, which I am going to skip for very rational reasons, but the upshot is that $x^{\mu}(\lambda)$ is a geodesic iff it satisfies the **geodesic equation**:

$$\frac{d^2 x^{\mu}}{d\lambda^2} + \Gamma^{\mu}_{\rho\sigma} \frac{dx^{\rho}}{d\lambda} \frac{dx^{\sigma}}{d\lambda} = 0 \tag{2.27}$$

This turns out to only hold iff λ is a **affine parameter**, i.e. if it can be related to the proper time by $\lambda = a\tau + b$.[8] Geodesics are extremely important because, in GR, a freely moving or falling body (a test body) always moves along a geodesic in a manner governed by the geodesic equation (2.27). If the body is massless, $ds^2 = 0$ and the geodesic is null, and if the body is massive then $ds^2 < 0$ and the geodesic is timelike. Upshot: in GR, particles move along geodesics unless acted on by an external force, and the geodesic equation is "our" version of Newton's second law.

[6] For example, Earth's "great circles" are the *geodesics* of the Earth.

[7] I say "extrema" and not "minima" because massive test particles move on geodesics of *maximum* proper time. In the twin paradox, a pair of twins take two paths through flat spacetime, one along a geodesic (sitting on their couch at home) and the other traveling into space and back. The couch-bound twin is older upon their reunion, because geodesics maximize proper time.

[8] We nearly always use τ *as* the affine parameter, when parameterizing timelike geodesics.

3
General relativity

Finally, the promised land! The core of the theory of general relativity can be given in two statements.

(1) Spacetime is a curved pseudo-Riemannian manifold with a metric of Lorentzian signature $(-+++)$.

(2) The relationship between matter and the curvature of spacetime is wholly contained in the **Einstein equation**:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{3.1}$$

The left-hand side of Eq. 3.1 is the Einstein tensor $G_{\mu\nu}$, defined in Eq. (2.24).[1] $T_{\mu\nu}$ is a symmetric tensor called the **stress-energy tensor**[2], and contains all information about the energy and momentum of matter fields, which are the source of gravity. So Eq. 3.1 basically says:

LHS $\leftrightarrow$ curvature of spacetime;

RHS $\leftrightarrow$ energy and momentum contained in the spacetime.

[1] Be careful to not confuse it with G which is the familiar Newtonian gravitational constant (and not the trace of $G_{\mu\nu}$).

[2] Also known as the energy-momentum tensor.

Eq. (2.24) functions as our "equation of motion" for the metric. One elegant and succinct formulation is "Spacetime tells matter how to move, matter tells spacetime how to curve."[3]

[3] Quote attributed to J.A. Wheeler.

What is the stress-energy tensor, exactly? The components of $T^{\mu\nu}$ are given by the flux of the μth component of momentum p^μ across a surface of constant x^ν (in the νth direction). But what does this mean? Consider a **perfect fluid** which is isotropic[4] in its rest frame. In any frame, the stress-energy tensor of this perfect fluid can be specified entirely in terms of its rest-frame energy density ρ and rest-frame pressure p, which is isotropic and equal in all directions:

[4] *Isotropy* = uniform in all orientations; no viscosity or heat flow.

$$T_{\mu\nu} = (\rho + p)U_\mu U_\nu + p g_{\mu\nu} \tag{3.2}$$

Normalizing by $g^{\mu\nu}U_\mu U_\nu = -1$ and raising the second index, this becomes:

$$T_\mu{}^\nu = \begin{pmatrix} -\rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix} \tag{3.3}$$

We claim that $T^{\mu\nu}$ contains all information about energy and momentum, so it must also contain the conservation laws. This is done by asserting that the covariant divergence of $T^{\mu\nu}$ vanishes:

$$\nabla_\mu T^{\mu\nu} = 0 \tag{3.4}$$

Since the Bianchi identity (Eq 2.23) guarantees that the divergence of the Einstein tensor vanishes, Eq. (3.1) guarantees energy-momentum conservation.

Let us remind ourselves that Einstein's formulation of GR should be a natural extension of Newtonian gravity, and should approximate Newton's laws in some limiting case. For the Newtonian limit, we assume that (1) the particles are moving slowly, (2) the gravitational field is weak, and (3) that it is static. Doing the requisite math, and using Poisson's equation for the Newtonian potential $\nabla^2\Phi = 4\pi G\rho$, we will find that $\Phi = -\frac{GM}{r}$, recovering the Newtonian gravitational law.

3.1 *Schwarzschild solution*

Unfortunately the TA lecture has probably gone on far too long by this point, so I won't be able to cover the Schwarzschild solution and black holes in any depth, so consider this supplementary reading.

How do we even go about solving an equation like (3.1)? The **Schwarzschild metric** is one idealized solution to the Einstein field equations, which describes the gravitational field outside of a spherically symmetric mass, assuming that the electric charge of the mass, angular momentum of the mass, and the cosmological constant are all zero.

$$ds^2 = -\left(1 - \frac{2GM}{r}\right) dt^2 + \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{3.5}$$

where M is the mass of the body with radius r. Perhaps it would be interest for you to see the nonzero Christoffels written out for this metric:

$$\begin{array}{lll} \Gamma^r_{tt} = \frac{GM}{r^3}(r - 2GM) & \Gamma^r_{rr} = \frac{-GM}{r(r-2GM)} & \Gamma^t_{tr} = \frac{GM}{r(r-2GM)} \\ \Gamma^\theta_{r\theta} = \frac{1}{r} & \Gamma^r_{\theta\theta} = -(r - 2GM) & \Gamma^\phi_{r\phi} = \frac{1}{r} \\ \Gamma^r_{\phi\phi} = -(r - 2GM)\sin^2\theta & \Gamma^\theta_{\phi\phi} = -\sin\theta\cos\theta & \Gamma^\phi_{\theta\phi} = \frac{\cos\theta}{\sin\theta} \end{array} \tag{3.6}$$

As it turns out, the Schwarzschild metric is the *unique* solution to Einstein's equation in a vacuum with a spherically symmetric matter distribution.[5] As such, the Schwarzschild solution is capable of describing things as un-exotic as our Solar System (notably the precession of Mercury's orbit, a problem of definite historical interest), as well as things like black holes.

[5] This is known as Birkhoff's theorem.

It turns out that the metric components blow up (go to infinity) at $r = 0$ and $r = 2GM$ (can you see why?), two points which we classify as **singularities**. **True** singularities point to an actual breakdown of the geometry, while **coordinate** singularities, which are just a side effect of choosing a bad coordinate system. In this case, $r = 0$ is a true singularity, while $r = 2GM$ is a coordinate singularity which also happens to specify the event horizon of a black hole.[6]

[6] In order to show this, we can make a coordinate transformation to what is known as the **Kruskal coordinates** and observe that nothing blows up at $r = 2GM$ in that basis.

What happens at $r = 2GM$? To an external observer, a clock falling into a black hole will appear to move more and more slowly as it approaches r, never actually crossing the surface; but for the observer who is actually falling into the black hole, time will progress at its usual rate since their watch will be inertial in their frame. They will readily pass through $r = 2GM$ and fall straight into $r = 0$. This is in fact inevitable, as r becomes a timelike coordinate for $r > 2GM$; falling to the center of the black hole is identical and equivalent to moving forwards in time. Assuming that you haven't yet been torn to unrecognizable strips of flesh by tidal forces, you would travel along a geodesic which maximizes the proper time τ.

There are other metric solutions to slightly more complex black holes, which build from the Schwarzschild black hole—such as the Reissner-Nordstrom metric for charged black holes, or the Kerr metric for rotating black holes.

3.2 *Cosmology*

In cosmology we make the simplifying assumption that the Universe is homogeneous and isotropic,[7] from which we can claim a "rest frame" of the Universe and write down the **Robertson-Walker** metric in comoving coordinates:

[7] Whether it is either, both, or neither of these things is a matter of considerable active research today.

$$ds^2 = -dt^2 + a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right] \tag{3.7}$$

where $a(t)$ is the scale factor, a measure of the relative expansion of the Universe, and

$$k = \begin{cases} -1 & \text{for an open universe;} \\ 0 & \text{for a flat universe;} \\ +1 & \text{for a closed universe.} \end{cases}$$

This is the only possible homogeneous and isotropic metric, and we just need to solve for $a(t)$ using Einstein's equation, in which case Einstein's equation becomes two differential equations known as the **Friedmann equations**:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2} \tag{3.8}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3\rho) \tag{3.9}$$

These equations determine the evolution of RW metrics, in what we call the Friedmann-Robertson-Walker cosmology. From solving the Friedmann equations we can begin to do things such as examine the behavior of different universes (matter-, radiation-, or vacuum-dominated), characterize their evolution and their final fates.

4
Bibliography

S. M. Carroll. A No-Nonsense Introduction to General Relativity. URL `https://www.astro.caltech.edu/~george/ay21/readings/grtiny.pdf`.

S. M. Carroll. Lecture Notes on General Relativity. *arXiv preprint gr-qc/9712019*, 1997.

S. M. Carroll. *Spacetime and Geometry: An Introduction to General Relativity*. Cambridge University Press, 2019. DOI: 10.1017/9781108770385.