GENERAL RELATIVITY

IAN LIM
MICHAELMAS 2018

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

Contents

Lecture 1.

Friday, October 5, 2018

Unlike in previous years, this course is intended to be a stand-alone course on general relativity, building up the mathematical formalism needed to construct the full theory and explore some examples of interesting spacetime metrics. It is linked to the Black Holes course taught in Lent term, which I will also be writing notes for.

Some recommended course materials and readings include the following:

- Sean Carroll, Spacetime and Geometry[1]
- Misner, Thorne, and Wheeler, Gravitation
- Wald, General Relativity
- Zee, Einstein Gravity in a Nutshell
- Hawking and Ellis, "The Large Scale Structure of Spacetime"

In Minkowski[2] spacetime (flat space) we specify points in spacetime by spatial coordinates in $\mathbb{R}^3$, i.e. the Cartesian coordinates (x, y, z), plus a time coordinate t. The line element (spacetime separation) is given by the metric

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2.$$

ds is the proper distance between x and $x + dx$, y and $y + dy$, z and $z + dz$, and t and $t + dt$. (As is typical in relativity, we work in units where $c = 1$. Note that the metric sign convention here is flipped from my QFT notes, which uses the "mostly minus" convention– this is arbitrary and so long as one is consistent it makes no difference.) Using the Einstein summation convention, the metric is usually written more compactly as

$$ds^2 = \eta_{\alpha\beta} x^\alpha x^\beta,$$

[1] I should point out that Sean Carroll's textbook is based off a set of GR notes which are available for free online. They are really excellent and pedagogically structured, and I have cross-referenced them frequently when revising these notes after lecture. Here is a link to the PDF on the arXiv: https://arxiv.org/pdf/gr-qc/9712019.pdf

[2] I've heard some USAmericans pronounce this "min-cow-ski." In German, it is "min-koff-ski."

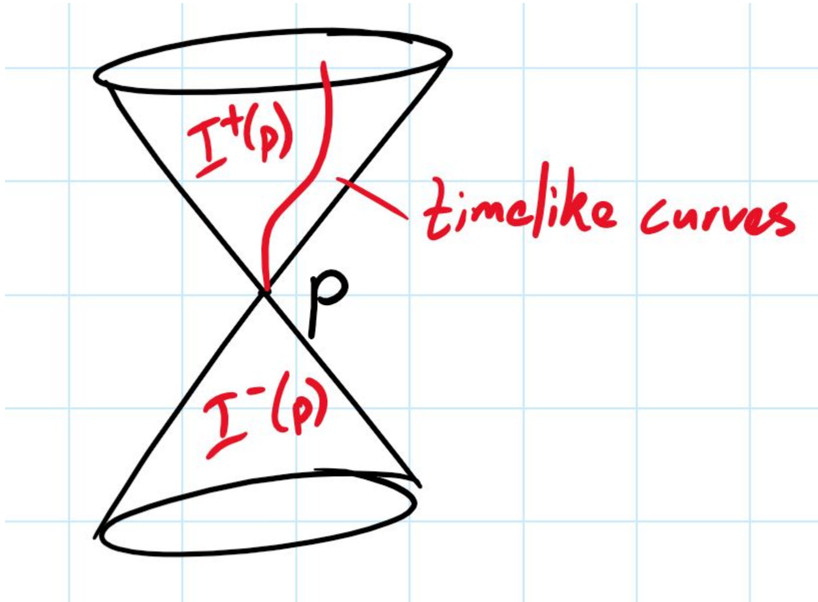

FIGURE 1. An illustration of the light cones from a point p, plus the chronological future I^+ and chronological past I^-. Also depicted in red is a timelike curve (e.g. a possible particle trajectory in spacetime).

with $\eta_{\alpha\beta}$ the Minkowski space metric.

Let's recall from special relativity that we call separations with $ds^2 > 0$ "spacelike," with $ds^2 < 0$ "timelike," and $ds^2 = 0$ null (or occasionally lightlike).

Definition 1.1. The *chronological future* of a point p is the set of all points that can be reached from p along future directed timelike lines, and we call this $I^+(p)$. It is the interior of the future-directed light cone. Conversely we have the chronological past of p, $I^-(p)$, which is the interior of the past-directed light cone. We also have the *causal future* of p, which is the set of all points that can be reached from p along future-directed timelike *or* null lines, and we call this $J^+(p)$. Similarly we have the causal past, $J^-(p)$. Thus J is the closure of I and is the interior *plus* the light cone itself.

Let $x^a(\tau)$ be a curve in spacetime.[3] Then the tangent vector to the curve is $u^a = \frac{dx^a}{d\tau}$. For timelike curves, $u^a u^b \eta_{ab} = -1 \iff \tau$ is the proper time along the curve. [4] We also know that $\int_p^q d\tau = \Delta\tau$, which just says that the integral of $d\tau$ along a curve from p to q yields the proper time interval, what a clock actually measures.

We also remark that Minkowski space has some very nice symmetries. Since x, y, and z do not appear explicitly in the metric, our spacetime is invariant under translations. It is also invariant under rotations in $\mathbb{R}^3$. It would be nice to extend rotations to include the time coordinate t as well– this is exactly what a Lorentz transformation does.

Lorentz transformations in general involve time– they are defined by the matrices Λ which satisfy

$$\Lambda^T \eta \Lambda = \eta,$$

i.e. they preserve the inner product η in Minkowski space, forming the group $O(3,1)$. Lorentz transformations consist of rotations in $\mathbb{R}^3$ and boosts. This is equivalent to the defining property of rotation matrices

[3]Evidently we are not using the convention that Greek indices range from 0 to 3 and Latin indices range from 1 to 3. I have copied the lecturer's convention here, but may change to more traditional notation if it becomes relevant.

[4]The property that $U^\alpha U^\beta \eta_{\alpha\beta} = -1$ is easy to prove. See the Special Relativity catch-up sheet found here for some nice exercises in SR: this is exercise 3. Assuming the result of exercise 2 which states that the four-velocity of a massive particle is $U^\mu = \gamma(1, v^i)$, we then have $U \cdot U = \gamma^2(-1 + v^2) = \frac{v^2-1}{1-v^2} = -1$. Since this is a fully contracted expression (no indices floating around), it is true in all frames.

R that $R^T \delta R = \delta$, meaning that rotation matrices preserve the standard Euclidean inner product in $\mathbb{R}^3$ and form the group $O(3)$.[5] Written explicitly, the Lorentz boost in the x-direction to a frame moving with velocity v is

$$\begin{aligned} t \to t' &= \frac{t - vx}{\sqrt{1-v^2}} \\ x \to x' &= \frac{x - vt}{\sqrt{1-v^2}} \\ y \to y' &= y \\ z \to z' &= z \end{aligned}$$

We may also write it in matrix notation,

$$\Lambda^a{}_b = \begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

where γ is defined in the usual way by $\gamma \equiv \frac{1}{\sqrt{1-v^2}}$.

Rather than constructing the (in general complicated) Lorentz boost in an arbitrary direction, it is often more convenient to rotate one's frame of reference in $\mathbb{R}^3$ so the boost is in the new x-direction, perform the Lorentz boost, and then transform back:

$$R^T \Lambda R = \Lambda_R,$$

where Λ_R is a new Lorentz transformation.[6]

Definition 1.2. The Lorentz transformations taken together form the *Lorentz group*. It satisfies the group axioms of identity, unique inverses (since $\det \Lambda \neq 0$), associativity (from associativity of matrix multiplication), and closure (see footnote for proof).[7]

Λ can include reflections in time or space. To avoid such complications, we sometimes refer to the *proper orthochronous Lorentz group*, i.e. to exclude space and time reversals, but often we are more careless and simply call it the Lorentz group.

Definition 1.3. The *Poincaré group* is then the semidirect product of Lorentz transformations and translations. This is the group of symmetries of Minkowski space.

We have translations defined as

$$x^a \to x^{a\prime} = x^a + \Delta x^a$$

and also Lorentz transformations, with the property

$$(\Lambda^T)_a{}^c \eta_{cd} \Lambda^d{}_b = \eta_{ab}.$$

Definition 1.4. We also have *contravariant vectors* (indices up) written u^a and their corresponding *covariant* vectors (indices down)

$$u_a \equiv \eta_{ab} u^b,$$

where we have used the metric to lower an index. These are sometimes equivalently called simply vectors and covectors. We can also raise indices using the inverse metric η^{ab} (defined by $\eta^{ab}\eta_{bc} = \delta^a_c$). Thus

$$u^b = \eta^{ba} u_a.$$

[5]Strictly, $O(3)$ also includes reflections– for matrices which preserve both orientation and the inner product, we must also require that $\det R = +1$, defining the group $SO(3)$. We'll see a similar caveat with the Lorentz group in just a second.

[6]It's easy to check that Λ_R really is a Lorentz transformation– just observe that rotations alone are a subset of Lorentz transformations, since they preserve the inner product on $\mathbb{R}^3$ and do not affect the time coordinate. In the language of group theory, rotations form an $SO(3)$ subgroup of the full Lorentz group $O(3,1)$– see Definition 1.2. Therefore any combination of rotations and Lorentz boosts will form another valid Lorentz transformation by the group closure property.

[7]More precisely, we know that the determinant is nonzero since $-1 = \det \eta = \det(\Lambda^T \eta \Lambda) = \det(\Lambda^T)\det(\eta)\det(\Lambda) = (-1)\det(\Lambda)^2 \implies \det(\Lambda) = \pm 1 \neq 0$. To prove closure, suppose Λ_1, Λ_2 are Lorentz transformations. The product $\Lambda_1\Lambda_2$ then satisfies $(\Lambda_1\Lambda_2)^T \eta (\Lambda_1\Lambda_2) = \Lambda_2^T \Lambda_1^T \eta \Lambda_1 \Lambda_2 = \Lambda_2^T \eta \Lambda_2 = \eta$, so $\Lambda_1\Lambda_2$ is also a Lorentz transformation.

We define the Lorentz transformation of a contravariant vector as $u^a \to u^{a\prime} = \Lambda^a{}_b u^b$. For instance, x^a is an example of a contravariant vector.

Definition 1.5. A *scalar* is an object which is invariant under a Lorentz transformation. We saw that a covariant vector transforms with right multiplication by the Lorentz transformation, whereas a contravariant vector transforms by left multiplication.

More generally, a *tensor of type* (r,s) transforms with r copies of the Lorentz transformation on the r up indices and s copies of the Lorentz transformation on the s down indices,

$$T^{\mu_1\mu_2\ldots\mu_r}{}_{\nu_1\nu_2\ldots\nu_s} \to T^{\alpha_1\alpha_2\ldots\alpha_r}{}_{\beta_1\beta_2\ldots\beta_s} = \Lambda^{\alpha_1}_{\mu_1}\ldots\Lambda^{\alpha_r}_{\mu_r} T^{\mu_1\mu_2\ldots\mu_r}{}_{\nu_1\nu_2\ldots\nu_s} \Lambda^{\nu_1}_{\beta_1}\ldots\Lambda^{\nu_s}_{\beta_s} \tag{1.6}$$

By this definition, a scalar may be thought of as a type $(0,0)$ tensor, a contravariant vector a type $(1,0)$ tensor, and a covariant vector a type $(0,1)$ tensor.

Lecture 2.

Monday, October 8, 2018

Today, we'll start by remarking that Maxwell's equations can be written compactly in 4-vector format. Recall from a good course on electrodynamics that we define the electromagnetic field strength tensor $F^{\mu\nu}$ as

$$F^{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{pmatrix}.$$

$F^{\mu\nu}$ is a totally antisymmetric rank two tensor. Defining the four-current $j^\mu \equiv (\rho, \mathbf{j})$ with $\mathbf{j}$ the ordinary current density and ρ the charge density, we see that

$$\partial_a F_{bc} + \partial_b F_{ca} + \partial_c F_{ab} = 0$$

and

$$\partial_a F^{ab} = -j^b.$$

But there's something strange about this– these equations in their current form hold for cartesian coordinates only. Of course, the laws of physics (i.e. as expressed through observable results in experiments) cannot depend on the coordinate system used.

Example 2.1. The Minkowski metric takes the Cartesian form

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2$$

but if we pass to spherical coordinates, the metric now takes the form

$$ds^2 = -dt^2 + dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2 = g_{ab} dx^a dx^b.$$

General relativity is thus motivated by a desire to understand how the laws of physics are invariant not just under Lorentz transformations but general coordinate transformations. It is also motivated by the *weak equivalence principle*, which states that inertial mass and gravitational mass are the same thing– the m in $F = ma$ and the m in $F = -\frac{GMm}{r^2}$ are the same mass! This is closely related to the *Einstein equivalence principle*, which states that in a freely falling frame, the laws of physics are those of special relativity. One cannot distinguish between being in freefall under a gravitational field and simply being at rest in no gravitational field.

We consider spacetime to be a 4-dimensional system $(3+1$ dimensions, if you like) and in particular it has a manifold structure. We may make an explicit choice of x^a some coordinates that label points in M, but it would be nice to define vectors in a way that is independent of the coordinates. This will lead us to revisit vectors and covectors.

Consider a curve $\lambda(\tau) : \mathbb{R} \to M$ a parametrized curve sitting in M. Now take $f = f(x^a)$ a differentiable function of the coordinates, and define an operator that maps f into df/dt: by applying the chain rule, we have

$$df/dt = \frac{\partial x^a}{\partial t}\left(\frac{\partial}{\partial x^a} f\right).$$

Thus a vector is a differential operator that acts on f: explicitly, it is $\frac{\partial x^a}{\partial t}\frac{\partial}{\partial x^a}$, where the $\frac{\partial x^a}{\partial t}$ are the components of the vector.

A general vector may therefore be written in its components in some basis x^a as

$$V = V^a \frac{\partial}{\partial x^a}.$$

Thinking back to our curve $\lambda(\tau)$, we may expand our coordinates locally as $x^a(\tau) = x^a(\tau_0) + V^a(\tau - \tau_0) + O((\tau - \tau_0)^2)$, where V is the tangent vector to some curve through the point τ_0. (Okay, we're being a bit careless with notation here– the instructor has written $\lambda(t)$, but sometimes t is a coordinate on the manifold.) Therefore we may also interpret (tangent) vectors as describing how our manifold curves locally about a point.

Vectors (somewhat obviously) form a vector space. If W, Y are vectors, α, β real numbers, then $\alpha W + \beta Y$ is another vector with components

$$(\alpha W^a + \beta Y^a)\frac{\partial}{\partial x^a}.$$

As (multi)linear differential operators, vectors obey the Leibniz rule

$$V^a \frac{\partial}{\partial x^a}(fg) = V^a \frac{\partial f}{\partial x^a} g + f V^a \frac{\partial g}{\partial x^a}.$$

So they form a vector space (check the vector space axioms again).

The space of tangent vectors at a point p is called $T_p(M)$. Recall that we defined our tangent vectors with respect to its components in some basis x^a. But if we now change to $\tilde{x}^b = \tilde{x}^b(x^a)$, then by the chain rule our basis vectors $\frac{\partial}{\partial x^a}$ transform as

$$\frac{\partial}{\partial x^a} = \frac{\partial \tilde{x}^b}{\partial x^a}\frac{\partial}{\partial \tilde{x}^b}.$$

But V as an operator is invariant– it does not depend on our choice of coordinates, so only its decomposition into basis vectors can change. This means that if we rewrite V in a different set of coordinates, we find that

$$V = V^a \frac{\partial}{\partial x^a} = \tilde{V}^a \frac{\partial}{\partial \tilde{x}^a} = V^a \frac{\partial \tilde{x}^b}{\partial x^a}\frac{\partial}{\partial \tilde{x}^b},$$

so by comparison the components of V transform as

$$V^a \to \tilde{V}^{a'} = \frac{\partial \tilde{x}^{a'}}{\partial x^a} V^a.$$

In other words, tangent vectors transform as contravariant vectors, which is a generalization of the formula in special relativity where we had

$$\frac{\partial \tilde{x}^{a'}}{\partial x^a} = \Lambda^{a'}_a$$

with $\Lambda^{a'}_a$ the Lorentz transform.

Definition 2.2. We may also define *one-forms*, which are covariant vectors at some point p. Thus the inner product $\langle \omega, V\rangle$ is a real number, with ω a 1-form and V a vector. The inner product is bilinear: if $V = \alpha Y + \beta W$, then

$$\langle \omega, \alpha Y + \beta W\rangle = \alpha\langle \omega, Y\rangle + \beta\langle \omega, W\rangle$$

and similarly for the first argument, if $\omega = \alpha\eta + \beta\xi$

$$\langle \alpha\eta + \beta\xi, V\rangle = \alpha\langle \eta, V\rangle + \beta\langle \xi, V\rangle.$$

Let us write V in a basis, $V = V^a E_a$ with E_a some set of basis vectors. Then $\omega = \omega_a E^a$ has components in some basis of one forms E^a. We have that $\langle E^a, E_b\rangle = \delta^a_b$, where E^a forms a basis of 1-forms which is dual to the ordinary basis vectors. (Remark: the components V^a of a vector transform like coordinate functions, while the components of a one-form ω_a transform like basis vectors E_a.) We can then compute the inner product of a generic one-form and a vector,

$$\begin{aligned}\langle \omega, V\rangle &= \langle \omega_a E^a, V^b E_b\rangle \\ &= \omega_a V^b \delta^a_b \\ &= \omega_a V^a.\end{aligned}$$

Lecture 3.

Wednesday, October 10, 2018

A quick admin note There is no lecture Monday 15 October. In addition, office hours will be Tuesdays at 4 PM in B1.26. Moving on.

Let us recall that we have a multiplication law on one-forms and vectors,

$$\langle \omega, X \rangle = \omega_a X^a$$

for ω any one-form, X any vector. That is, we can write this product in terms of the components of ω and X.

Definition 3.1. With this in mind, we define the *differential* of a function $f : M \to \mathbb{R}$ to be the one-form df, such that

$$\langle df, X \rangle = Xf$$

(that is, X as a differential operator acting on f).

Example 3.2. Non-lectured example: consider the function $f = x + y$ in $\mathbb{R}^3$ and let $X = \frac{\partial}{\partial y}$. (We have chosen a coordinate basis to make the computation clearer.) Then $df = dx + dy$ (a one-form) and now

$$\langle df, X \rangle = Xf = \frac{\partial}{\partial y}(x + y) = 1.$$

Recall we have a basis of 1-forms E^a and a basis of vectors E_b with $\langle E^a, E_b \rangle = \delta^a_b$. In a coordinate basis, the basis vectors take the form

$$E_a = \frac{\partial}{\partial x^a} \text{ and } E^b = dx^b.$$

Thus

$$\langle dx^a, \frac{\partial}{\partial x^b} \rangle = \delta^a_b.$$

Definition 3.3. A one-form is *exact* if it can be written as df for some scalar f. For instance, dt and dr are exact because they are the differentials of t and r, but $rd\theta$ is not exact. However, the one-form rdr is exact, since it can be written $d(r^2/2)$.

In Minkowski space with Cartesian coordinates, the natural basis of one-forms dt, dx, dy, dz forms a coordinate basis since each of these is exact, and the basis of vectors dual to this is $\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$.

However, in spherical coordinates the Minkowski metric looks different. It takes the form

$$ds^2 = -dt^2 + dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2.$$

The basis of one-forms here,

$$dt, dr, rd\theta, r\sin\theta d\phi$$

is not a coordinate basis because these are not all of the form df. The set of basis vectors dual to the one-forms in spherical coordinates is also kind of bad. They take the form

$$\frac{\partial}{\partial t}, \frac{\partial}{\partial r}, \frac{1}{r}\frac{\partial}{\partial \theta}, \frac{1}{r\sin\theta}\frac{\partial}{\partial \phi},$$

and these are not a coordinate basis because they are not of the form $\frac{\partial}{\partial x^a}$ (equivalently, they are not dual to exact one-forms).

However, we remark that our defining equation for the product of a one-form and vector produces an ordinary scalar, which must be invariant under coordinate transformations:

$$\langle \omega, X \rangle = \omega_a X^a \text{ in any basis.}$$

This determines how the components of a one-form ω_a change under coordinate transformations. In a coordinate basis, we know that the components of a vector transform like coordinate functions:

$$X^a \to \tilde{X}^{a'} = \frac{\partial \tilde{x}^{a'}}{\partial x^a} X^a.$$

Therefore in a coordinate basis, the components of a one-form must transform in the inverse way,

$$\omega_a \to \tilde{\omega}_{a'} = \frac{\partial x^a}{\partial \tilde{x}^{a'}} \omega_a.$$

Note where the primed indices lie and which coordinates are the new coordinates $\tilde{x}$ versus the old coordinates x. The factor here $\frac{\partial x^a}{\partial \tilde{x}^{a'}}$ is analogous to how the Lorentz transformation acts (as the Lorentz transformation is a particular coordinate transformation satisfying certain constraints).

Suppose that $\langle df, X\rangle = 0$ for some df. If one is working in n dimensions, this gives one constraint equation on the n components of X. Thus, there are $(n-1)$ different linearly independent choices of X which solve this equation and therefore span an $n-1$-dimensional space. We have put one constraint specified by f on our space of all possible X such that df is the normal to the surface $f =$ constant.

Example 3.4. Again, a non-lectured concrete example. Let us again work in $\mathbb{R}^3$ and set $f = x$. Then a general X can be written as $X^a \frac{\partial}{\partial x^a}$ and the condition that $\langle df, X\rangle = 0$ can be computed explicitly as

$$\langle df, X\rangle = \left(X^1 \frac{\partial}{\partial x} + X^2 \frac{\partial}{\partial y} + X^3 \frac{\partial}{\partial z} \right)(x) = X^1(1) = 0.$$

Therefore our surface is defined by $X^1 = 0$ but we may choose X^2 and X^3 freely ($n - 1 = 2$ free choices). Indeed, we see that $df = dx$ is normal to the surface $f = x =$ constant.

A tensor of type (r, s) in a basis of 1-forms E^a and vectors E_a takes the form

$$T = T^{a_1 \ldots a_r}{}_{b_1 \ldots b_s} E_{a_1} \otimes E_{a_2} \otimes \ldots \otimes E_{a_r} \otimes E^{b_a} \otimes \ldots \otimes E^{b_s},$$

where $\otimes$ is the tensor product (not just a direct product!).[8]

T is coordinate invariant, so in a coordinate basis the components of T transform as

$$\tilde{T}^{a'_1 \ldots a'_r}_{b'_1 \ldots b'_s} = \frac{\partial \tilde{x}^{a'_1}}{\partial x^{a_1}} \cdots \frac{\partial \tilde{x}^{a'_r}}{\partial x^{a_r}} \frac{\partial x^{b_1}}{\tilde{x}^{b'_1}} \cdots \frac{\partial x^{b_s}}{\tilde{x}^{b'_s}} T^{a_1 \ldots a_n}{}_{b_1 \ldots b_s}.$$

In a non-coordinate basis, these $\frac{\partial \tilde{x}^{a'}}{\partial x^a}$ are replaced by some general functions $\Phi^{a'}_a$ where $\tilde{x}^{a'} = \Phi^{a'}_a x^a$.

We can perform the symmetrization operation, denoted by putting indices to be symmetrized in parentheses:

$$X_{(a_1 \ldots a_r)} \equiv \frac{1}{r!} \left[\text{sum of all permutations of } a_1 \ldots a_r\right].$$

For example, $X_{(ab)} = \frac{1}{2}\left[X_{ab} + X_{ba}\right]$. Similarly we have the antisymmetrization operation, denoted by putting indices to be antisymmetrized in square brackets:

$$X_{[a_1 \ldots a_r]} = \frac{1}{r!} \left[\text{sum over all even permutations } - \text{ sum of all odd permutations}\right].$$

For example, $X_{[ab]} = \frac{1}{2}[X_{ab} - X_{ba}]$. Having defined symmetrization and antisymmetrization, we now consider a special class of tensor– the totally antisymmetric $(0, p)$ tensor.

Definition 3.5. A *differential p-form* is a tensor of type $(0, p)$ which is antisymmetric on all indices, i.e. $A_{a_1 \ldots a_p} = A_{[a_1 \ldots a_p]}$. Some familiar p-forms include the 2-form $F_{\mu\nu}$ from electromagnetism and the Levi-Civita symbol ϵ_{ijk}.

We can describe A in terms of basis vectors E^a using a construction called the wedge product.

Definition 3.6. The *wedge product* is a special kind of antisymmetrizing multiplication of a p-form and a q-form. For a p-form $A = A_{a_1 \ldots a_p}$ and a q-form $B = B_{b_1 \ldots b_q}$, the wedge product $A \wedge B$ is given by

$$(A \wedge B)_{a_1 \ldots a_p b_1 \ldots b_q} \equiv A_{[a_1 \ldots a_p} B_{b_1 \ldots b_q]}.$$

[8] Tensor products are more complicated than direct products because their addition structure is multilinear, i.e. linear in each argument individually but not all simultaneously. Where it might make sense to add $(2,1) + (1,2) = (3,3)$ in $\mathbb{R} \times \mathbb{R}$, the equivalent tensor product in $\mathbb{R} \otimes \mathbb{R}$ would have $2 \otimes 1 + 1 \otimes 2 = 2 \otimes 1 + 2 \otimes 1 = (2+2) \otimes 1 = 4 \otimes 1$. So this is quite a different beast. More info on tensor products and tensors as mathematical constructions can be found at `https://jeremykun.com/2014/01/17/how-to-conquer-tensorphobia/`.

For instance $A \wedge B = (-1)^{pq} B \wedge A$ (this is easy to prove– we simply switch the q indices of B past the p indices of A and pick up the appropriate pq sign flips along the way).

As an invariant object, the p-form A can be written as

$$A = A_{a_1 \dots a_p} E^{a_1} \wedge \dots \wedge E^{a_p},$$

where $A_{a_1 \dots a_p}$ are now the components of the p-form A.

Definition 3.7. We also define the exterior derivative, a generalization of the usual derivative ∂_μ:

$$(dA)_{ba_1 \dots a_p} \equiv \frac{\partial}{\partial x^{[b}} A_{a_1 \dots a_p]} = \partial_{[b} A_{a_1 \dots a_p]}$$

defines a $p+1$-form, as it is by definition antisymmetric in its $p+1$ indices. The exterior derivative of a product follows a variation of the Leibniz rule:

$$d(A \wedge B) = dA \wedge B + (-1)^p A \wedge dB.$$

Note that $ddA = 0$, so d is nilpotent (it kills all exact differentials).[9]

The gradient is a simple example of an exterior derivative of a 0-form (AKA a scalar):

$$(d\phi)_\mu = \partial_\mu \phi.$$

From prior experiences with special (or general) relativity, we might have an intuition that the metric has something to do with gravitation. The line element ds (defined by $ds^2 = g_{ab} dx^a dx^b$) is invariant and is therefore a (symmetric) tensor. In a freely falling frame, the metric of Minkowski space is

$$\tilde{\eta}_{a'b'} = \frac{\partial x^a}{\partial \tilde{x}^{a'}} \frac{\partial x^b}{\partial \tilde{x}^{b'}} g_{ab}.$$

Do such $\frac{\partial x^a}{\partial \tilde{x}^{a'}}$ always exist? The answer turns out to be yes– g_{ab} is not degenerate, so one may diagonalize it and then rescale the eigenvalues. Sylevester's theorem states that if g has r positive eigenvalues, s negative eigenvalues, then diagonalizing preserves this.

Therefore given g_{ab} that is non-degenerate, the inverse metric g^{ab} can be define with $g^{ab} g_{bc} = \delta^a_c$ the Kronecker delta. One may use the metric to raise and lower indices: $V_b = g_{bc} V^c$ and $V^a = g^{ab} V_b$.

"There are more unknowns than there are knowns." A brief summary of this course.

Lecture 4.

Friday, October 12, 2018

Previously, we defined the exterior derivative, which took a p-form to a $p+1$-form. Now we will define the covariant derivative, an operation which in general takes a tensor of type (r, s) to a tensor of type $(r, s+1)$.

Suppose we start with a scalar field $\phi(x)$. The ordinary derivative is just

$$\partial_a \phi = \frac{\partial \phi}{\partial x^a}.$$

Let us change coordinates to $\tilde{x}^{a'}$ some function of the original coordinates. Then this derivative transforms as

$$\partial_{a'} \phi = \frac{\partial x^a}{\partial \tilde{x}^{a'}} \frac{\partial}{\partial x^a} \phi = \frac{\partial x^a}{\partial \tilde{x}^{a'}} \partial_a \phi.$$

We might ask whether the derivative of a vector transforms in the same way. But instead, we get something a little different.

[9]Suppose we compute ddA: then we will have two derivatives in our expression $\partial_{[\mu} \partial_\nu A_{a_1 \dots a_p]}$. But derivatives commute, so to every $\partial_\alpha \partial_\beta$ term in the antisymmetrization sum there will be a corresponding $-\partial_\beta \partial_\alpha$ term. These terms cancel no matter what A is, so $ddA = 0$ identically.

$$\begin{aligned}\partial_{b'}\tilde{V}^{a'} &= \frac{\partial x^b}{\partial \tilde{x}^{b'}}\frac{\partial}{\partial x^b}\left(\frac{\partial \tilde{x}^{a'}}{\partial x^a}V^a\right)\\ &= \frac{\partial x^b}{\partial \tilde{x}^{b'}}\frac{\partial V^a}{\partial x^b}\frac{\partial \tilde{x}^{a'}}{\partial x^a}+\frac{\partial x^b}{\partial \tilde{x}^{b'}}\frac{\partial^2 \tilde{x}^{a'}}{\partial x^a \partial x^b}V^a.\end{aligned}$$

This first part is tensorial, but the second part is not (it has a term which is a second derivative of the coordinates). In order to get a tensor from the derivative, we need to add a correction term. Let us define

$$\nabla_b V^a \equiv \partial_b V^a + \Gamma^a_{bc}V^c$$

where Γ^a_{bc} is called a *connection*. We can figure out how Γ transforms under coordinate transformations:

$$\tilde{\Gamma}^{a'}_{b'c'} = \frac{\partial \tilde{x}^{a'}}{\partial x^a}\frac{\partial x^a}{\partial \tilde{x}^{b'}}\frac{\partial x^c}{\partial \tilde{x}^c}\Gamma^a_{bc} - \frac{\partial^2 \tilde{x}^{a'}}{\partial x^b \partial x^c}\frac{\partial x^b}{\partial \tilde{x}^{b'}}\frac{\partial x^c}{\partial \tilde{x}^{c'}}.$$

So Γ does *not* transform as a tensor, but that's actually what we want– this correction term allows us to get a proper tensor when we take the covariant derivative of a vector.

We'd like ∇ to be linear and obey the Leibniz rule: for two tensors T, S and two real numbers $\alpha, \beta \in \mathbb{R}$, we should have

$$\nabla(\alpha T + \beta S) = \alpha \nabla T + \beta \nabla S$$

and also

$$\nabla(T \otimes S) = \nabla T \otimes S + T \otimes \nabla S.$$

For a vector V and a one-form W, let $S = V^a W_a$. Then

$$\begin{aligned}\nabla_a S &= \partial_a S\\ &= (\partial_a V^b)W_b + V^b(\partial_a W_b)\\ &= (\nabla_a V^b)W_b - \Gamma^b_{ac}V^c W_b + V^b(\partial_a W_b)\\ &= (\nabla_a V^b)W_b + V^b \nabla_a W_b.\end{aligned}$$

Therefore for the Leibniz rule to hold on the product of a vector and a one-form, it must be that

$$\nabla_b W_a \equiv \partial_b W_a - \Gamma^c_{ba}W_c.$$

Note the sign flip from the vector definition! More generally, we can use Leibniz to deduce what the covariant derivative operator is on a general tensor of type (r, s).

$$\nabla_c T^{a_1\ldots a_r}_{b_1\ldots b_s} = \partial_c T^{a_1\ldots a_r}_{b_1\ldots b_s} + \Gamma^{a_1}_{cd}T^{da_2\ldots a_r} + \Gamma^{a_2}_{cd}T^{a_1 d a_3\ldots a_r}_{b_1\ldots} + \ldots + \Gamma^{a_r}_{cd}T^{a_1 a_2\ldots d}_{\ldots} - \Gamma^d_{cb_1}T^{\ldots}_{db_2\ldots b_s} - \Gamma^d_{cb_2}T^{\ldots}_{b_1 d\ldots b_s} - \ldots - \Gamma^d_{cb_s}T^{\ldots}_{b_1 b_2\ldots d}.$$

So every upstairs indices we swap out gets a $+\Gamma$ and every downstairs index we swap gets a $-\Gamma$. Let's return to our expression for the transformation of Γ,

$$\tilde{\Gamma}^{a'}_{b'c'} = \frac{\partial \tilde{x}^{a'}}{\partial x^a}\frac{\partial x^a}{\partial \tilde{x}^{b'}}\frac{\partial x^c}{\partial \tilde{x}^c}\Gamma^a_{bc} - \frac{\partial^2 \tilde{x}^{a'}}{\partial x^b \partial x^c}\frac{\partial x^b}{\partial \tilde{x}^{b'}}\frac{\partial x^c}{\partial \tilde{x}^{c'}}.$$

Note that the second part is symmetric under the interchange of b', c'. Therefore take just the part antisymmetric in b', c':

$$\Gamma^{a'}_{b'c'} - \Gamma^{a'}_{c'b'} = \frac{\partial \tilde{x}^{a'}}{\partial x^a}\frac{\partial x^b}{\partial \tilde{x}^{b'}}\frac{\partial x^c}{\partial \tilde{x}^{c'}}(\Gamma^a_{bc} - \Gamma^a_{cb}).$$

Definition 4.1. The antisymmetric part of Γ transforms like a tensor, and so we define the torsion tensor

$$T^a_{bc} \equiv \Gamma^a_{bc} - \Gamma^a_{cb} = 2\Gamma^a_{[bc]}.$$

Some definitions define this up to a factor of 2 or with different signs.

Consider an arbitrary scalar S.

$$(\nabla_a \nabla_b - \nabla_b \nabla_a)S = \nabla_a \partial_b S - \nabla_b \partial_a S.$$

If these were just partial derivatives, this would be zero. But working it out explicitly, we see that

$$(\nabla_a \nabla_b - \nabla_b \nabla_a)S = \partial_a \partial_b S - \Gamma^c_{ab}\partial_c S - \partial_b \partial_a S + \Gamma^c_{ba}\partial_c S = T^c_{ba}\partial_c S = T^c_{ba}\nabla_c S.$$

Therefore the torsion measures how much covariant derivatives fail to commute on scalars. In general relativity, the torsion is usually taken to be zero. However, a treatment of fermions naturally requires non-zero torsion, and in local supersymmetry or "superspace formulations of anything," non-zero torsion is essential.

We haven't yet actually found what the connection is in terms of things we actually care about.

Definition 4.2. Let us define the metric connection as the Γ such that

$$\nabla_c g_{ab} = 0.$$

This will allow us to find a formula for Γ in terms of the metric g.

We'll work it out explicitly.

$$\begin{aligned}
\nabla_a g_{bc} &= \partial_a g_{bc} - \Gamma^d_{ab} g_{dc} - \Gamma^d_{ac} g_{bd} = 0, \\
\nabla_b g_{ca} &= \partial_b g_{ca} - \Gamma^d_{bc} g_{da} - \Gamma^d_{ba} g_{cd} = 0, \\
\nabla_c g_{ab} &= \partial_c g_{ab} - \Gamma^d_{ca} g_{bd} - \Gamma^d_{cb} g_{ad} = 0.
\end{aligned}$$

If we add the first two of these and subtract the third, we end up

$$\partial_a g_{bc} + \partial_b g_{ca} - \partial_c g_{ab} = 2\Gamma^d_{ab} g_{dc},$$

using the fact that $\Gamma^d_{bc} = \Gamma^d_{cb}$ since we have set the torsion to zero.

Now we simply multiply by g^{ce} to find that

$$\frac{1}{2} g^{ce}(-\partial_c g_{ab} \partial_a g_{bc} + \partial_b g_{ca}) = \Gamma^d_{ab} g_{dc} g^{ce} = \Gamma^d_{ab} \delta^e_d = \Gamma^e_{ab}.$$

This gives us explicitly the metric connection, which we sometimes call the Christoffel connection or Christoffel symbols.[10] Thus

$$\Gamma^a_{bc} = \frac{1}{2} g^{ad}(-\partial_d g_{bc} + \partial_b g_{cd} + \partial_c g_{bd}).$$

It is, as expected, symmetric under exchange $b \leftrightarrow c$ since the metric is symmetric, $g_{ab} = g_{ba}$.

So now on scalars,

$$(\nabla_a \nabla_b - \nabla_b \nabla_a) S = 0,$$

i.e. covariant derivatives commute on scalars. Moreover using the metric connection if we have $V_a = g_{ab} V^b$, then

$$\nabla_c(V_a) = \nabla_c(g_{ab} V^b) = (\nabla_c g_{ab}) V^b + g_{ab} \nabla_c V^b = g_{ab} \nabla_c V^b,$$

since $\nabla_c g_{ab} = 0$. Therefore with the metric connection, the metric commutes with the covariant derivative. This is also true of the inverse metric– prove this as an exercise.

Exercise 4.3. Prove[11] that the covariant derivative of the inverse metric is also zero,

$$\nabla_c g^{ab} = 0.$$

Lecture 5.

Wednesday, October 17, 2018

We defined the symmetric metric connection (i.e. the Christoffel connection) such that

$$\nabla_a g_{bc} = 0 \text{ with } \Gamma^c_{ab} = \Gamma^c_{ba}.$$

We also found that if ϕ is a scalar field, the covariant derivatives commute:

$$(\nabla_a \nabla_b - \nabla_b \nabla_a)\phi = \partial_a \partial_b \phi - \Gamma^c_{ab} \partial_c \phi - \partial_b \partial_a \phi + \Gamma^c_{ba} \partial_c \phi = 0$$

[10] They are a pain to compute by hand, hence why one professor of mine once referred to them as the "Christ-awful symbols."

[11] I think we can do this with Leibniz, actually. $\nabla_c(g_{ab} g^{ab}) = \nabla_c(g_{ab}) g^{ab} + g_{ab} \nabla_c g^{ab} = 0$ since the trace of the metric is just a scalar. If g_{ab} is not identically zero, it must be that $\nabla_c g^{ab}$ vanishes.

where we used the symmetry of the Christoffel symbols to cancel the second and fourth terms. We might then ask if this is true for the covariant derivatives on a 1-form as well.

$$\begin{aligned}(\nabla_a\nabla_b - \nabla_b\nabla_a)V_c &= \nabla_a(\partial_b V_c - \Gamma^e_{bc}V_e) - (a \leftrightarrow b)\\ &= \partial_a\partial_b V_c - \Gamma^e_{bc}\partial_a V_e - (\partial_a\Gamma^e_{bc})V_e - \Gamma^f_{ab}(\partial_f V_c - \Gamma^e_{fc}V_e) - \Gamma^f_{ac}(\partial_b V_f - \Gamma^e_{bc}V_e) + \Gamma^f_{bc}(\partial_a V_f - \Gamma^e_{af})\\ &\quad -\partial_b\partial_a V_c + \Gamma^e_{ac}\partial_b V_e + (\partial_b\Gamma^e_{ac})V_e + \Gamma^f_{ba}(\partial_f V_c - \Gamma^e_{fc}V_e)\end{aligned}$$

It's not quite zero (in general), but what we find after a bit of close inspection is that all the second derivative terms cancel, and all the terms with derivatives of V also cancel.[12] We're left with products of the Christoffel symbols and also derivatives thereof:

$$(\nabla_a\nabla_b - \nabla_b\nabla_a)V_c = (-\partial_a\Gamma^e_{bc} + \partial_b\Gamma^e_{ac} - \Gamma^f_{bc}\Gamma^e_{af} + \Gamma^f_{ac}\Gamma^e_{bf})V_e.$$

Since the expression on the LHS is a tensor, the RHS must also be a tensor. (We can check this explicitly using the transformation properties of Γ, though I don't recommend it.)

Definition 5.1. We therefore define the curvature tensor ${R_{abc}}^e$ by the following:

$$(\nabla_a\nabla_b - \nabla_b\nabla_a)V_c \equiv {R_{abc}}^e V_e,$$

where R^e_{abc} is given explicitly by

$$R^e_{abc} = -\partial_a\Gamma^e_{bc} + \partial_b\Gamma^e_{ac} - \Gamma^f_{bc}\Gamma^e_{fa} + \Gamma^f_{ac}\Gamma^e_{fb}.$$

Roughly speaking, the curvature tensor measures how much the covariant derivatives of tensors fail to commute.[13]

On arbitrary tensors $T^{ab\dots}_{cd\dots}$ one can write down a rather long expression for the commutator of the covariant derivatives:

$$(\nabla_e\nabla_f - \nabla_f\nabla_e)T^{ab\dots}_{cd\dots} = {R_{ef}}^a{}_p T^{pb\dots}_{cd\dots} + R^b_{ef\,p} T^{ap\dots}_{cd\dots} + R^p_{efc}T^{ab\dots}_{pd\dots} + R^p_{efd}T^{ab\dots}_{cp\dots} + \dots$$

similar to our formula for taking individual covariant derivatives. There is no further interesting content in these arbitrary tensors, however– all the interesting physics seems to already be captured in the curvature tensor.

The Riemann tensor (i.e. the curvature tensor with an index lowered) also has some nice symmetries which you may like to check.

$$\begin{aligned} R_{abcd} &= -R_{bacd}\\ R_{abcd} &= -R_{abdc}\\ R_{abcd} &= R_{cdab}\\ R_{abcd} + R_{acdb} + R_{adbc} &= 0\end{aligned}$$

These can be recovered from the explicit form of the curvature tensor with sufficient patience.

As a consequence of these identities, the Riemann tensor has many components (though they are somewhat constrained by symmetry). In d dimensions, it has $\frac{1}{12}d^2(d^2-1)$ components, so in 4 spacetime dimensions there are 20 independent components. In $d = 3$ there are only 6 and in $d = 2$, just 1.

In general there are many, many terms one needs to work out to actually compute the Riemann tensor. There are very nice computer programs like Mathematica which can automate the process, or if you have some time on your hands it is a decent exercise to write the code yourself.

Definition 5.2. We also define the *Ricci tensor*:

$$R_{bd} \equiv R_{abcd}g^{ac}$$

[12]Actually, we could do this for a general connection without assuming that the connection is symmetric in its lower two indices. If we do so, we pick up a term with the torsion tensor in it.

[13]This is only technically true for torsion-free connections. To quote Sean Carroll, "The Riemann tensor measures that part of the commutator of covariant derivatives that is proportional to the vector field, while the torsion tensor measures the part that is proportional to the covariant derivative of the vector field."

where we contract on the first and third indices of the Riemann tensor. The Ricci tensor is symmetric, $R_{ab} = R_{ba}$.

Definition 5.3. If we contract once more, we get the *Ricci scalar*,

$$R \equiv R^{ab} g_{ab}.$$

In two-dimensional calculations, R is the same as the Gaussian curvature up to a numerical factor. In addition, computing the Ricci scalar for Minkowski space reveals that it is zero– Minkowski space is flat.

We can now discuss geodesics, curves which extremize the proper distance between two endpoints p, q.

Definition 5.4. The proper distance along the line from p to q is given by

$$\int_p^q ds = \int_p^q \sqrt{\left| g_{ab} \frac{dx^a}{d\lambda} \frac{dx^b}{d\lambda} \right|} d\lambda$$

since $ds^2 = g_{ab} dx^a dx^b$. This is a functional of the path $x^a(\lambda)$ we take through the space, and when it is extremized[14] we call the resulting path a *geodesic*.

Geodesics generalize the concept of a straight line to curved space. For instance, a great circle is an example of a geodesic for the surface of the Earth.

Extremizing the integral of ds is hard because of the square root, so we usually just extremize $\int_p^q ds^2$ instead. That is, we extremize

$$I = \int_p^q g_{ab} \frac{dx^a}{d\lambda} \frac{dx^b}{d\lambda} d\lambda = \int_p^q L d\lambda.$$

By the Euler-Lagrange equations, we can compute

$$\frac{d}{d\lambda} \left(\frac{\partial L}{\partial \dot{x}^a} \right) - \frac{\partial L}{\partial x^a} = 0$$

where $\cdot = \partial / \partial \lambda$. Therefore

$$\frac{d}{d\lambda} (2 g_{ab} \dot{x}^b) - \partial_a (g_{bc} \dot{x}^b \dot{x}^c) = 0$$

Equivalently

$$\frac{d}{d\lambda} (g_{ab} \dot{x}^b) - \frac{1}{2} (\partial_a g_{bc}) \dot{x}^b \dot{x}^c = 0.$$

By the chain rule we rewrite the first term as

$$g_{ab} \ddot{x}^b + \left(\frac{\partial}{\partial x^c} g_{ab} \right) \frac{\partial x^c}{\partial \lambda} \frac{\partial x^b}{\partial \lambda} - \frac{1}{2} \partial_a g_{bc} \dot{x}^b \dot{x}^c = 0.$$

Regrouping,

$$\begin{aligned} 0 &= g_{ab} \ddot{x}^b + (\partial_c g_{ab} - \frac{1}{2} \partial_a g_{bc}) \dot{x}^b \dot{x}^c \\ &= g^{ae} g_{ab} \ddot{x}^b + \frac{1}{2} g^{ae} (\partial_c g_{ab} + \partial_b g_{ac} - \partial_a g_{bc}) \dot{x}^b \dot{x}^c. \end{aligned}$$

Recognizing that the first term becomes a delta function δ^e_b and the second term includes a Christoffel symbol, we rewrite this as

$$\frac{d^2 x^e}{d\lambda^2} + \Gamma^e_{bc} \frac{dx^b}{d\lambda} \frac{dx^c}{d\lambda} = 0,$$

which we often call the *geodesic equation.* The name of the game is to construct the tangent vector V^a to the curve, with $V^a = \frac{dx^a}{d\lambda}$. We can also write

$$V^b \nabla_b V^a = 0,$$

which is also sometimes called the geodesic equation.

[14] When this refers to a path length in just space it's minimized, but when we are computing proper time it is maximized.

Lecture 6.

Friday, October 19, 2018

Last time, we wrote down the geodesic equation. That is, we defined curves $x^a(x)$ such that

$$\frac{d^2x^a}{ds^2} + \Gamma^a_{bc}\frac{dx^b}{ds}\frac{dx^c}{ds} = 0.$$

In terms of the tangent vector to the curve $V^a = \frac{dx^a}{ds}$, we can equivalently write

$$V^a \nabla_a V^b = 0.$$

What happens if we reparametrize the curve? For instance, we could change to some new variables

$$\tilde{s} = \tilde{s}(s).$$

Then $d/ds = \frac{d\tilde{s}}{ds}\frac{d}{d\tilde{s}}$ and $\frac{d^2}{ds^2} = \frac{d^2\tilde{s}}{ds^2}\frac{d}{d\tilde{s}^2} + (\frac{d\tilde{s}}{ds})^2\frac{d^2}{d\tilde{s}^2}$.

By substituting, we get a new version of the geodesic equation.

$$\left(\frac{d\tilde{s}}{ds}\right)^2 \frac{d^2}{d\tilde{s}^2}x^a + \frac{d^2\tilde{s}}{ds^2}\frac{d}{d\tilde{s}}x^a + \left(\frac{d\tilde{s}}{ds}\right)^2 \Gamma^a_{bc}\frac{dx^b}{d\tilde{s}}\frac{dx^c}{d\tilde{s}} = 0.$$

A little rewriting reveals that in terms of $\tilde{s}$, we get

$$\frac{d^2x^a}{d\tilde{s}^2} + \Gamma^a_{bc}\frac{dx^b}{d\tilde{s}}\frac{dx^c}{d\tilde{s}} = -\frac{\frac{d^2\tilde{s}^2}{ds^2}}{\left(\frac{d\tilde{s}}{ds}\right)^2}\frac{dx^a}{d\tilde{s}}.$$

But $\tilde{s}$ is arbitrary, so in terms of our new tangent vector $\tilde{V}^a = dx^a/d\tilde{s}$, we have

$$\tilde{V}^b \nabla_b \tilde{V}^a = f(s)\tilde{V}^a,$$

a more general form of the geodesic equation.

Definition 6.1. If $f(s) = 0$, we say the geodesic is *affinely parametrized.* If a geodesic is affinely parametrized, then it remains so for $\tilde{s}$ such that $d^2\tilde{s}/ds^2 = 0$, e.g. for $\tilde{s} = as + b$, a, b constants.

If a geodesic is affine, then

$$V^a \nabla_a (V^b V_b) = 2V_b V^a \nabla_a V^b = 0$$

and so $V^b V_b$ is constant along the geodesic– if a curve is initially timelike(/spacelike/null) it will remain timelike(/spacelike/null) along the geodesic.

Recalling how the Christoffel symbol transforms under arbitrary changes of coordinates, we find that under $x^a \mapsto \tilde{x}^{a'} \equiv \tilde{x}^{a'}(x^b)$, we have

$$\Gamma^{a'}_{b'c'} = \frac{\partial \tilde{x}^{a'}}{\partial x^a}\frac{\partial x^b}{\partial \tilde{x}^{b'}}\frac{\partial x^c}{\partial \tilde{x}^{c'}}\Gamma^a_{bc} - \frac{\partial^2 \tilde{x}^{a'}}{\partial x^b \partial x^c}\frac{\partial x^b}{\partial \tilde{x}^{b'}}\frac{\partial x^c}{\partial \tilde{x}^{c'}}.$$

Is it possible to make the new Christoffel symbol vanish in the $\tilde{x}$ coordinates? This is equivalent to the condition that

$$\frac{\partial \tilde{x}^{a'}}{\partial x^a}\Gamma^a_{bc} = \frac{\partial^2 \tilde{x}^{a'}}{\partial x^b \partial x^c}.$$

Suppose we want it to vanish at x^a_0. Let us choose coordinates $\tilde{x}$ defined by

$$\tilde{x}^{a'} = (x^a - x^a_0) + \frac{1}{2}\Gamma^a_{bc}(x^b - x^b_0)(x^c - x^c_0).$$

Taking the first derivative with respect to the original coordinates yields

$$\frac{\partial \tilde{x}^a}{\partial x^e} = \delta^a_e + \Gamma^a_{bc}(x^b - x^b_0)\delta^c_e + \ldots$$

where the ... denotes derivatives of the Christoffel symbols, and similarly

$$\frac{\partial^2 \tilde{x}^a}{\partial x^e \partial x^f} = \Gamma^a_{ef} + \ldots$$

We therefore see that

$$\frac{\partial \tilde{x}^{a'}}{\partial x^a}\Gamma^a_{bc} = \Gamma^a_{bc}(\delta^{a'}_a + \Gamma^{a'}_{b'c'}(x^{b'} - x^{b'}_0)\delta^{c'}_a) = \Gamma^{a'}_{bc} = \frac{\partial^2 \tilde{x}^{a'}}{\partial x^b \partial x^c}$$

when $x^b = x^b_0$, so we have found coordinates where the Christoffel symbol vanishes at a point of our choosing.

Definition 6.2. If we choose coordinates so that Γ^a_{bc} vanishes at a point, then those coordinates are called *Gaussian normal coordinates.* In normal coordinates, the metric takes the form

$$g_{ab} = C_{ab} + O(x - x_0)^2,$$

where C_{ab} is some set of constants. These coordinates have forced the terms linear in $x - x_0$ to vanish, and by applying the Lorentz transform and rotations (and possibly a scale transformation) we can diagonalize C_{ab} so that at p,

$$C_{ab} = \eta_{ab}.$$

What we have learned is that spacetime can always be made to look like Minkowski spacetime at a given point (up to higher-order corrections). We also call such a choice of coordinates *inertial coordinates*.

In normal coordinates, the geodesic equation for an affinely parametrized curve is

$$\frac{d^2 x^a}{ds^2} = 0,$$

the equation of motion for a freely falling particle in Minkowski space. This confirms our intuition that a geodesic is really a generalization of a straight line.

Of course, it's also apparent that this equation is not covariant since it depends on our choice of coordinates x^a. In order to be coordinate-independent, we need to write this as a tensorial equation, which is just

$$V^a \nabla_a V^b = 0.$$

If we further choose $V^a V_b = -1$ (this is true for the trajectory of a massive particle) then the parametrization s is simply the proper time along the curve.

One can now consider families of curve parametrized by both time t along the curve and another parameter s. The tangent vectors to each geodesic are given by

$$T^a(s) \equiv \frac{dx^a(t,s)}{dt}$$

and if we derive with respect to s instead we get a tangent vector relating neighboring geodesics,

$$S^a(t) \equiv \frac{dx^a(t,s)}{ds},$$

sometimes called the deviation vector. How does S^a change as one moves along the geodesics? If we consider

$$V^b \equiv (\nabla_T S)^b = T^a \nabla_a S^b$$

this is like the "relative velocity of geodesics" (where I've used ∇_T to represent a directional covariant derivative)– it measures how much the deviation vector S^a points in the direction of the tangent vector T^a.

One can equivalently define the "relative acceleration of geodesics" as

$$A^b \equiv (\nabla_T V)^b = T^a \nabla_a V^b.$$

We'll revisit this quantity shortly.

Lemma 6.3. *Consider the following quantity:*

$$S^a \nabla_a T^b - T^a \nabla_a S^b.$$

Expanding out, we find that

$$S^a \nabla_a T^b - T^a \nabla_a S^b = S^a(\partial_a T^b + \Gamma^b_{ac} T^c) - T^a(\partial_a S^b + \Gamma^b_{ac} S^c).$$

The terms with Christoffel symbols cancel by the symmetry of Γ^b_{ac} (since a and c are dummy indices, we can swap them and relabel). But

$$S^a \partial_a T^b = S^a \frac{\partial T^b}{\partial x^a} = \frac{\partial^2 x^b}{\partial s \partial t} = T^a \partial_a S^b,$$

so our expression

$$S^a \nabla_a T^b - T^a \nabla_a S^b = 0.$$

This tells us that we can swap S^a and T^b through the covariant derivative so long as we keep the indices fixed.

Now what is the acceleration between neighboring geodesics,

$$A^a = \frac{d^2 S^a}{dt^2}?$$

Expanding out, we find that it is

$$\begin{aligned} \frac{d^2 S^a}{dt^2} &= T^c \nabla_c (T^b \nabla_b S^a) \\ &= T^c (\nabla_c S^B)(\nabla_b T^a) + T^c S^b \nabla_c \nabla_b T^a \\ &= T^c (\nabla_c S^b)(\nabla_b T^a) + T^c S^b (\nabla_b \nabla_c T^a + R_{cb}{}^a{}_d T^d). \end{aligned}$$

In the next step, we'll move the T^c inside the ∇_b derivative to get

$$T^c (\nabla_c S^b)(\nabla_b T^a) + S^b \nabla_b (T^c \nabla_c T^a) - S^b (\nabla_b T^c) \nabla_c T^a + T^c S^b R_{cb}{}^a{}_d T^d.$$

But the second term vanishes by the geodesic equation, and the third term can be written as $-S^c \nabla_c T^b \nabla_b T^a = -T^c \nabla_c S^b \nabla_b T^a$ (by the identity we proved earlier) so the third term cancels the first one.

What remains is a function of the Riemann tensor. Using the symmetries of the Riemann tensor, we can rewrite our acceleration equation as

$$\frac{d^2 S^a}{dt^2} = R^a{}_{bcd} T^b T^c S^d,$$

which we call the *equation of geodesic deviation.* If the Riemann tensor is zero, then the relative velocity between geodesics is constant. If the Riemann tensor $\neq 0$, then we get a "stretching force" between neighboring geodesics, sometimes called a tidal force.