

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- \TeX ed 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

1.	Friday, October 5, 2018	2
2.	Monday, October 8, 2018	4

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
- Misner, Thorne, and Wheeler, Gravitation
- Wald, General Relativity
- Zee, Einstein Gravity in a Nutshell
- Hawking and Ellis, “The Large Scale Structure of Spacetime”

In Minkowski¹ 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

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

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.² 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.³ 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

¹I’ve heard some USAmericans pronounce this “min-cow-ski.” In German, it is “min-koff-ski.”

²Evidently 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.

³The property that $U^a U^b \eta_{ab} = -1$ 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.

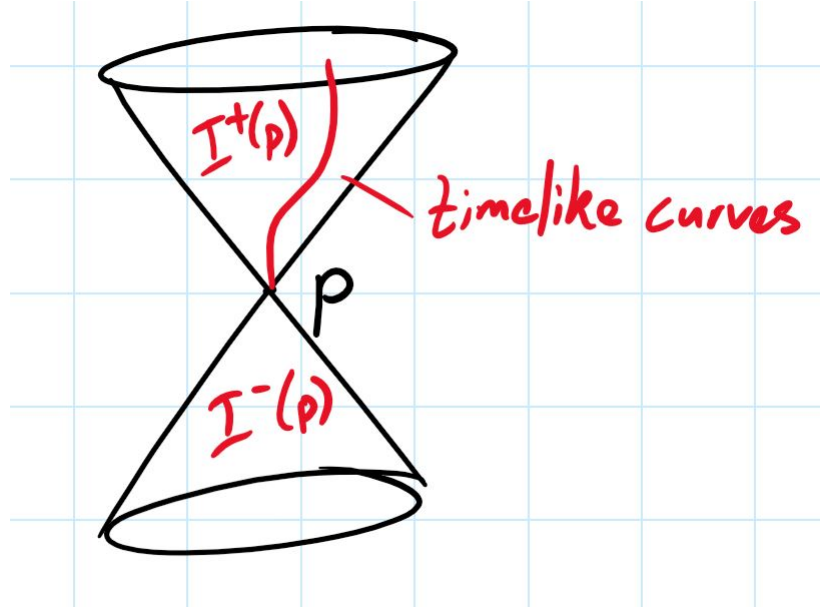


FIGURE 1. An illustration of the light cones from a point p , plus the chronological future I^+ and chronological past I^- . Also depicted in red is a timelike curve (e.g. a possible particle trajectory in spacetime).

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 $O(3)$.⁴ Written explicitly, the Lorentz boost in the x -direction to a frame moving with velocity v is

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

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.⁵

⁴Strictly, $O(3)$ 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.

⁵It'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.

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).⁶

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

$$x^a \rightarrow x^{a'} = x^a + \Delta x^a$$

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

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

We define the Lorentz transformation of a contravariant vector as $u^a \rightarrow u^{a'} = \Lambda^a{}_b u^b$. 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,

$$T^{\mu_1 \mu_2 \dots \mu_r}{}_{\nu_1 \nu_2 \dots \nu_s} \rightarrow T^{\alpha_1 \alpha_2 \dots \alpha_r}{}_{\beta_1 \beta_2 \dots \beta_s} = \Lambda^{\alpha_1}{}_{\mu_1} \dots \Lambda^{\alpha_r}{}_{\mu_r} T^{\mu_1 \mu_2 \dots \mu_r}{}_{\nu_1 \nu_2 \dots \nu_s} \Lambda^{\nu_1}{}_{\beta_1} \dots \Lambda^{\nu_s}{}_{\beta_s} \quad (1.6)$$

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

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

⁶More 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.

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 $\lambda(\tau) : \mathbb{R} \rightarrow M$ 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.)

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

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

so by comparison the components of V transform as In other words, tangent vectors transform as contravariant vectors, which is a generalization of the formula in special relativity where $\Lambda_a^{a'} = \frac{\partial x^{a'}}{\partial x^a}$

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. We remark that 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 $\omega = \alpha \eta + \beta \zeta$

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

Let us write $V = V^a E_a$ 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 $\langle E^a, E_b \rangle = \delta_b^a$, where E^a forms a basis of 1-forms which is dual to the ordinary basis vectors. As the components of a 1-form are real numbers (and the same is true of vectors) we may compute

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