GR lecture 1-1

Historical & conceptual intro, tensor indices in a Cartesian frame

I. HISTORICAL INTRO

Arik Einstein was cool, retreated from public life in the 80's to avoid celebrity status. Albert Einstein was cool too, has immense celebrity status. Is it justified? Um, yes.

- 1666: Year of Miracles. Newton does calculus, mechanics, gravity, geometric optics.
- 1905: 2nd Year of Miracles. Einstein does Special Relativity, photons, Brownian motion.
- 1916: Einstein finishes GR.

In 1905, Einstein is established as the greatest scientist of his generation. Special Relativity needs no introduction. Photons get a Nobel. The Brownian motion thing is very cool, and often overlooked. It leads to the first-ever estimate of the absolute size of atoms, or, in other words, the actual value of Avogadro's Number. All these activities are remarkable, but still within the normal scope of science. In particular, SR is sometimes depicted as an out-of-the-blue lone genius' breakthrough, but it absolutely wasn't. The 19th century has already produced a full-fledged special-relativistic theory – Maxwell's electromagnetism. It was constructed gradually, through a century's worth of interplay between theory and experiment. SR is simply the general framework that electromagnetism was suggesting. Several others were working it out at the same time – Lorentz, Poincare. They lacked some of Einstein's insight, but they could have gotten there eventually. So, the 1905 Einstein is merely the greatest scientist of his time. Boring.

1916 is a different story. Here, Einstein is not quite doing science as usually conceived. Here, there is no preceding century of experiments and theory, just a conceptual gap – between SR and Newtonian gravity. GR is ahead of its time in many respects. Eliezer Yudkowski has an essay about it on LessWrong titled "faster than science". Einstein did not produce GR by experimenting, not by model-building, not by trying to fit observational anomalies. He was working on a higher level of abstraction – he was meditating on the nature of physical law.

GR was ahead of its time in other ways too. These days, in field theory class, we teach a succession of theories in increasing order of subtlety – scalars (spin 0), fermions (spin 1/2), electromagnetism (spin-1 that interacts with lower spins), Yang-Mills (spin-1 that also interacts with itself), and finally GR (spin-2). But historically, Yang-Mills was invented only 40 years after GR. Even in the case of electromagnetism, its full theoretical structure – in particular, its gauge symmetry – was only understood in the light of GR.

The reason I rant about this stuff is that we in quantum gravity / string theory are all trying to be Einstein of the 1910's. We are trying to achieve the next revolution in theoretical physics without experimental input, by meditating on the nature of physical law (which in practice involves a lot of calculations set in imaginary worlds). God help us.

II. STATEMENT OF ATTITUDE, EXPECTATIONS, LEARNING GR IS A SOLITARY ENTERPRISE

III. CONCEPTUAL INTRO – CURVED COORDINATES, CURVED SPACE

In this course, we will deal with curved spacetime. This is a good time to inject some initial intuition about the concept. You know a lot about working with flat space. You also know that sometimes, curved coordinates can be useful – e.g. cylindrical and spherical. But they've always been just a trick to more efficiently solve some particular problem. The laws of Nature always look simpler when written in flat, Cartesian coordinates. However, the existence of such coordinates is actually a good definition for flat space. When such coordinates do not exist, we say that space is curved. So, space is curved when all coordinates are curved. Then we have no choice but to write our laws in curved coordinates!

All of this has a precise analogue in the discussion of inertial reference frames. We know that sometimes, a non-inertial frame can be useful – e.g. an accelerated or a rotating one. However, the laws are always simplest in inertial frames. Curved spacetime is when all frames are non-inertial. This is not just an analogy of words. Inertial motion at constant velocity describes a straight line in spacetime. Velocity is spacetime slope. Accelerated motion is when the line through spacetime curves. So, an inertial frame is defined by straight lines in spacetime, and a non-inertial one by curved lines! When there is no inertial frame, spacetime itself must be curved.

IV. TENSOR INDICES IN FLAT 3D SPACE

Before we start discussing inertial frames, let's make sure we're all synchronized on tensor notation in 3d space. For now, we're not worrying about the distinction between upper and lower indices – that will come soon. We use indices $(i, j, \dots)$ that can take the 3 values (x, y, z), or $(1, 2, 3)$, or however you like to label your 3 Cartesian axes. So, a vector $\mathbf{v} = (v_x, v_y, v_z)$ is denoted as v_i. The scalar product of two vectors is:

$$\mathbf{u} \cdot \mathbf{v} = u_x v_x + u_y v_y + u_z v_z = u_i v_i \ , \tag{1}$$

where a repeated index always means that we sum over all 3 values. We sometimes say that repeated indices are "contracted", or "traced over". Indices that aren't repeated are called "free". An equation must have the same free indices on both sides, and is meant to hold for any substitution of values for the indices. For example:

$$\mathbf{a} = \mathbf{b} - 2(\mathbf{b} \cdot \mathbf{c})\mathbf{c} \quad \Longleftrightarrow \quad a_i = b_i - 2b_j c_j b_i \ . \tag{2}$$

Exercise 1.

1. *Assume that c_i is a unit vector. What is the geometric meaning of eq. (2)?*

2. *What is the geometric meaning if we remove the factor of 2?*

All of matrix algebra can be captured in this language, if we denote a matrix as a quantity with two indices:

$$A_{ij} : \quad \begin{pmatrix} A_{xx} & A_{xy} & A_{xz} \\ A_{yx} & A_{yy} & A_{yz} \\ A_{zx} & A_{zy} & A_{zz} \end{pmatrix} \tag{3}$$

In particular, we have:

$$\begin{aligned} &A\mathbf{u} = \mathbf{v} \ \Leftrightarrow \ A_{ij} u_j = v_i \ ; \quad AB = C \ \Leftrightarrow \ A_{ik} B_{kj} = C_{ij} \ ; \\ &A = B^T \ \Leftrightarrow \ A_{ij} = B_{ji} \ ; \quad \operatorname{tr} A = A_{ii} \ . \end{aligned} \tag{4}$$

We leave determinants and inverse matrices for later, as those are a bit tricky.

The identity matrix is denoted as δ_{ij}, so that e.g. $u_i v_i = \delta_{ij} u_i v_j$. In addition to scalars, vectors and matrices, we can write quantities with arbitrary numbers of indices. These are called tensors. The largest interesting tensor that we'll encounter in this course will have 4 indices. Among all tensors, δ_{ij} has a special property: its components are unchanged under rotations or reflections of the Cartesian axes. This is just a statement of the invariance of the scalar product.

Spatial derivatives form a vector $\boldsymbol{\partial}$, which in index notation becomes:

$$\partial_i = \frac{\partial}{\partial x_i} \tag{5}$$

(we reserve the notation ∇ for the covariant derivative, which will appear later on). Thus, the gradient of a scalar is $\partial_i f$, and the divergence of a vector is $\partial_i f_i$.

One additional piece of notation that will be very useful is symmetrization and anti-symmetrization brackets, which pick out the symmetric or anti-symmetric component of a tensor under the permutation of some set of indices. For example:

$$A_{(ij)} = \frac{1}{2}(A_{ij} + A_{ji}) \ ; \quad A_{[ij]} = \frac{1}{2}(A_{ij} - A_{ji}) \ ; \tag{6}$$

$$\begin{aligned} A_{(ijk)} &= \frac{1}{6}(A_{ijk} + A_{ikj} + A_{jik} + A_{jki} + A_{kij} + A_{kji}) \ ; \\ A_{[ijk]} &= \frac{1}{6}(A_{ijk} - A_{ikj} - A_{jik} + A_{jki} + A_{kij} - A_{kji}) \ . \end{aligned} \tag{7}$$

Exercise 2. *Write an arbitrary matrix A_{ij} as a sum of three terms: a multiple of δ_{ij} ("spin-0"), an antisymmetric matrix ("spin-1") and a traceless symmetric matrix ("spin-2").*

V. THE LEVI-CIVITA TENSOR AND AXIAL STUFF

In addition to δ_{ij}, there is another "universal" tensor. It is also invariant under rotations, but flips sign under reflections. This is the "Levi-Civita tensor" ϵ_{ijk}, which we define as the totally antisymmetric tensor $\epsilon_{ijk} = \epsilon_{[ijk]}$ with $\epsilon_{xyz} = 1$.

Exercise 3. *Deduce all the components of ϵ_{ijk}.*

While δ_{ij} defines the scalar product, ϵ_{ijk} defines the vector product:

$$\mathbf{u} \times \mathbf{v} = \mathbf{w} \quad \Longleftrightarrow \quad \epsilon_{ijk} u_j v_k = w_i \ . \tag{8}$$

In particular, the curl of a vector field f_i is $\epsilon_{ijk}\partial_j f_k$. The Levi-Civita tensor is also closely related to the concept of volume: $\epsilon_{ijk} u_i v_j w_k$ is the volume of the parallelogram spanned by u_i, v_i and w_i, up to a sign that depends on their mutual orientation.

Exercise 4.

1. *Show that this is true when the 3 vectors are orthogonal to each other.*

2. *Now, keeping one of the orthogonal vectors fixed, "slant" the others via* $v_i \to v_i + \alpha u_i$ *and* $w_i \to w_i + \beta u_i + \gamma v_i$. *Show that the volume* $\epsilon_{ijk} u_i v_j w_k$ *is unchanged.*

A lot of elementary physics is written with vector products, forcing poor students to remember "right-hand rules", and creating the impression that Nature is not left-right symmetric. That impression is actually true, but only for the weak interactions. In mechanics and electromagnetism, all the ϵ_{ijk}'s and right-hand rules always cancel out, via the identity:

$$\epsilon_{ijk}\epsilon_{lmn} = \delta_{il}\delta_{jm}\delta_{kn} - \delta_{il}\delta_{jn}\delta_{km} - \delta_{im}\delta_{jl}\delta_{kn} + \delta_{im}\delta_{jn}\delta_{kl} + \delta_{in}\delta_{jl}\delta_{km} - \delta_{in}\delta_{jm}\delta_{kl} \ , \tag{9}$$

which means simply: "if ijk is some permutation of xyz and so is lmn, then ijk must be either lmn, or lnm, or mln, or…".

Exercise 5. *Similarly, write down* $\epsilon_{ijm}\epsilon_{klm}$, $\epsilon_{ikl}\epsilon_{jkl}$ *and* $\epsilon_{ijk}\epsilon_{ijk}$.

We can write most of physics without any ϵ_{ijk}'s at all, but that requires a small sacrifice. Recall that $\mathbf{u}\times\mathbf{v}$ is <u>the vector orthogonal to</u> the plane element spanned by u_i and v_i. Instead of taking that extra step, we can just talk about the plane element directly! This will not be a vector, but an antisymmetric matrix (also called a bivector) $2u_{[i}v_{j]} = u_i v_j - u_j v_i$. If we allow such objects in our equations, then the right-hand rule becomes unnecessary. In addition, this is good practice for life in 4d spacetime, where there is <u>no</u> unique vector orthogonal to a given plane element.

In elementary physics, there are three main "axial vectors", i.e. vectors that contain in them a right-hand rule: angular velocity ω_i, angular momentum L_i, and the magnetic field B_i. We can replace them once and for all by bivectors via:

$$\omega_{ij} = \epsilon_{ijk}\omega_k \ ; \quad \omega_i = \frac{1}{2}\epsilon_{ijk}\omega_{jk} \ . \tag{10}$$

Exercise 6. *Rewrite the following equations using index notation and the bivectors* ω_{ij}, L_{ij} *an* B_{ij}, *without any* ϵ_{ijk}*'s in the final result.*

1. *Rotational velocity:* $\mathbf{v} = \boldsymbol{\omega}\times\mathbf{r}$.

2. *Angular momentum from angular velocity and moment of inertia:* $L_i = I_{ij}\omega_j$.

3. *Magnetic force:* $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$.

4. *Maxwell's equations:*

$$\boldsymbol{\partial} \cdot \mathbf{E} = \rho/\epsilon_0 \ ; \quad \boldsymbol{\partial} \times \mathbf{E} = -\dot{\mathbf{B}} \ ; \quad \boldsymbol{\partial} \cdot \mathbf{B} = 0; \quad \boldsymbol{\partial} \times \mathbf{B} = \mu_0(\mathbf{j} + \epsilon_0 \dot{\mathbf{E}}) \ . \tag{11}$$

A. Determinants and inverse matrices

The determinant of a matrix A_{ij} can be nicely defined using the product (9) of two Levi-Civitas:

$$\det A = \frac{1}{6}\epsilon_{ijk}\epsilon_{lmn}A_{il}A_{jm}A_{kn} \ . \tag{12}$$

Exercise 7.

1. *Make sure you understand what the factor of* $1/6$ *is doing here. What would be the corresponding expression for* 2×2 *matrices? How about* 4×4*?*

2. *Use the decomposition* (9) *to express* $\det A$ *without using* ϵ_{ijk}*'s.*

The inverse of a matrix can be expressed via the formula:

$$(A^{-1})_{ij} = \frac{(\mathrm{adj}\,A)_{ij}}{\det A} \ , \tag{13}$$

where adj A_{ij} is the adjugate matrix:

$$(\mathrm{adj}\,A)_{ij} = \frac{1}{2}\epsilon_{jkl}\epsilon_{imn}A_{km}A_{ln} \ . \tag{14}$$

If you look closely, you'll see that these are the standard expressions from Linear Algebra. We can also use index notation to prove from scratch that (13) is indeed the inverse matrix:

Exercise 8 (TRICKY)**.** *Prove that* (13) *is the inverse of* A_{ij}*, i.e. that:*

$$(\mathrm{adj}\,A)_{ik}A_{kj} = (\det A)\delta_{ij} \ . \tag{15}$$

As an intermediate step, it helps to prove:

$$\epsilon_{lmn}A_{li}A_{mj}A_{nk} = \epsilon_{ijk}\det A \ . \tag{16}$$

To prove the latter, note that anything antisymmetric in ijk *must be proportional to* ϵ_{ijk}*, and you can use* (12) *to find the proportionality coefficient.*