GR lecture 3-1

Lorentz matrices, relativistic particle action, 4-current and stress-energy tensor, particle in EM field

I. MORE ON LORENTZ TRANSFORMATIONS

Let us bridge a small gap in our discussion of Lorentz transformations. In Lecture 2-2, we discussed rotations, in a context where upper and lower indices are the same. We then had the transformation law $x_i \to R_{ij}x_j$, or $\mathbf{x} \to R\mathbf{x}$ in matrix notation, with the orthogonality constraint $RR^T = 1$, i.e. $R^{-1} = R^T$. On the other hand, in Lecture 1-2, we discussed general basis transformations, without any constraints, which acted differently on upper vs. lower indices: $u_i \to M_i{}^j u_j$ ($u \to Mu$ in matrix notation) vs. $v^i \to (M^{-1})_j{}^i v^j$ ($v \to (M^{-1})^T v$ in matrix notation).

Lorentz transformations occupy a middle ground in this respect. On one hand, they are not arbitrary basis transformations: they are constrained to preserve the Minkowski metric $\eta_{\mu\nu}$. On the other hand, since $\eta_{\mu\nu}$ is not the identity matrix, there is a difference between upper and lower indices. Let us now sort out this slightly confusing situation.

To begin with, let's obey our convention for general basis transformations: vectors transform as $v^\mu \to (\Lambda^{-1})_\nu{}^\mu v^\nu$, covectors as $u_\mu \to \Lambda_\mu{}^\nu u_\nu$. In particular, the Minkowski metric transforms as:

$$\eta_{\mu\nu} \ \to \ \Lambda_\mu{}^\rho \Lambda_\nu{}^\sigma \eta_{\rho\sigma} \ . \tag{1}$$

The Lorentz transformations are those matrices $\Lambda_\mu{}^\nu$ that preserve $\eta_{\mu\nu}$, i.e. satisfy $\Lambda_\mu{}^\rho \Lambda_\nu{}^\sigma \eta_{\rho\sigma} = \eta_{\mu\nu}$. In matrix notation, this reads:

$$\Lambda\eta\Lambda^T = \eta \quad \Longleftrightarrow \quad \Lambda^{-1} = \eta\Lambda^T\eta^{-1} \ . \tag{2}$$

This is the Lorentzian generalization of the orthogonality condition we had for rotations. Converting the last equation back into index notation, we get:

$$(\Lambda^{-1})_\mu{}^\nu = \eta_{\mu\rho}\Lambda_\sigma{}^\rho \eta^{\sigma\nu} \equiv \Lambda^\nu{}_\mu \ , \tag{3}$$

where $\Lambda^\nu{}_\mu$ had its indices raised and lowered using the metric $\eta_{\mu\nu}$ and its inverse $\eta^{\mu\nu}$. Thus, the Lorentz transformation of vectors and covectors can be written in essentially the same way:

$$v^\mu \;\to \Lambda^\mu{}_\nu v^\nu \; ; \quad u_\mu \;\to\; \Lambda_\mu{}^\nu u_\nu \; . \tag{4}$$

In fact, this is another way of saying that the Lorentz transformations preserve the metric: they commute with the raising and lowering of indices.

Next, let's consider the infinitesimal version of a Lorentz transformation:

$$\Lambda_\mu{}^\nu = \delta_\mu^\nu + \epsilon M_\mu{}^\nu \; ; \quad (\Lambda^{-1})_\mu{}^\nu = \delta_\mu^\nu - \epsilon M_\mu{}^\nu \; , \tag{5}$$

where ϵ is a small parameter, and $M_\mu{}^\nu$ is a Lorentz generator. The condition (3) now becomes:

$$M^\nu{}_\mu = -M_\mu{}^\nu \quad \Longleftrightarrow \quad M_{\mu\nu} = -M_{\nu\mu} \; . \tag{6}$$

Thus, the Lorentz generators are antisymmetric matrices $M_{\mu\nu} = M_{[\mu\nu]}$, just like ordinary rotation generators! Note, however, that this is only true after we use the metric $\eta_{\mu\nu}$ to lower the second index of $M_\mu{}^\nu$. The antisymmetric generators $M_{\mu\nu}$ have the same meaning as in ordinary space: they specify the plane in which the infinitesimal rotation is taking place.

Exercise 1. *Consider a Lorentz boost:*

$$t \;\to\; \frac{t - vx}{\sqrt{1-v^2}} \; ; \quad x \;\to\; \frac{x - vt}{\sqrt{1-v^2}} \; , \tag{7}$$

in the limit of infinitesimal v (but otherwise relativistically, i.e. without assuming $|t| \gg |x|$). Parameterizing this boost as $x^\mu \to x^\mu + vM^\mu{}_\nu x^\nu$, write the components of the generator $M^\mu{}_\nu$. Write also the lowered-index components $M_{\mu\nu}$.

II. ACTION OF A FREE MASSIVE PARTICLE

The action of a free relativistic particle is simply $-m$ times the length of its path through spacetime (also called its worldline):

$$S = -m \int d\tau = -m \int \sqrt{-dx_\mu dx^\mu} \; . \tag{8}$$

Exercise 2. *Show that this action leads to uniform motion in a straight line. You can use your geometric intuition from ordinary space – there is no need for fancy calculations.*

Exercise 3. *Why is there a minus sign in front of the action? What difference between spacetime and ordinary space does it reflect?*

Exercise 4. *Again, use geometric intuition from ordinary space to argue that $p_\mu = mu_\mu$ is the canonical momentum derived from the action* (8) *by varying the trajectory's final point.*

Exercise 5. *In the limit of small velocities, show how the action* (8) *relates to the usual action $\frac{m}{2}\int v^2 dt$ of a non-relativistic particle.*

We can also derive uniform motion in a straight line from the action (8) by brute force, following the prescription of Lagrangian mechanics:

$$S = \int L(q, \dot{q})dt \quad \Longrightarrow \quad \frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = 0 \ . \tag{9}$$

However, in the relativistic setting, it would be unfortunate to treat the time coordinate t as special and separate from the spatial position $\mathbf{x}$. What we can do instead is introduce an arbitrary parameter λ that runs along the particle's worldline, and parameterize the trajectory as $x^\mu(\lambda)$. Then t is treated together with $\mathbf{x}$ as part of the "configuration variables q", while λ assumes the old role of t. It is possible to define λ as the proper time τ along the worldline, but enforcing that actually leads to unnecessary complications. The action (8) now becomes:

$$S = \int L(\dot{x}^\mu)d\lambda \ ; \quad L = -m\sqrt{-\dot{x}_\mu \dot{x}^\mu} \ , \tag{10}$$

where the dots now represent $d/d\lambda$ derivatives. The Euler-Lagrange equations then read:

$$0 = -\frac{d}{d\lambda}\frac{\partial L}{\partial \dot{x}^\mu} = -m\frac{d}{d\lambda}\left(\frac{\dot{x}_\mu}{\sqrt{-\dot{x}_\mu \dot{x}^\mu}}\right) = -m\frac{du_\mu}{d\lambda} \ . \tag{11}$$

Where we recognized $\dot{x}_\mu/\sqrt{-\dot{x}_\mu \dot{x}^\mu}$ as the unit tangent vector to the worldline, i.e. the 4-velocity $u_\mu = dx_\mu/d\tau$. Thus, the equations of motion demand that u_μ remains constant along the trajectory, as expected.

III. 4-CURRENTS AND CONSERVATION LAWS

In non-relativistic physics, we often talk about the density ρ of some scalar quantity, such as electric charge or the number of atoms, i.e. the quantity per unit volume. In non-static situations, we also talk about the current density $\mathbf{j}$, i.e. the quantity flowing per unit time through a unit area in each direction. Local conservation laws take the form:

$$\frac{\partial\rho}{\partial t} = -\boldsymbol{\partial}\cdot\mathbf{j} \; . \tag{12}$$

When integrated over a volume, the LHS becomes the time derivative of the charge in a region, and the RHS becomes the flux of current into the region.

In SR, the charge density ρ and current density $\mathbf{j}$ become unified into a 4-vector $j^\mu = (\rho, \mathbf{j})$, which we refer to as the 4-current.

Exercise 6. *Show that j^μ indeed transforms a 4-vector. Consider a uniform charge density at rest, $j^\mu = (\rho_0, \mathbf{0})$. Using the Lorentz transformation of the coordinates $x^\mu = (t, \mathbf{x})$, find the components of $j^\mu = (\rho, \mathbf{j})$ in a boosted frame, and compare with the expected transformation of a 4-vector's components.*

Like dx^μ, u^μ and p^μ, the 4-current j^μ tends to be associated with the motion of particles. It is useful to note a property that all these 4-vectors share. Their spatial components are related to the timelike one via:

$$\mathbf{dx} = \mathbf{v}dt \; ; \quad \mathbf{p} = E\mathbf{v} \; ; \quad \mathbf{j} = \rho\mathbf{v} \; . \tag{13}$$

Note that $\mathbf{p} = E\mathbf{v}$ is a relativistic generalization of the non-relativistic $\mathbf{p} = m\mathbf{v}$, since, at small velocities, we have $E = m + mv^2/2 + \cdots \approx m$.

Let's now return to the current conservation law (12). In spacetime notation, this becomes simply:

$$\partial_\mu j^\mu = 0 \; . \tag{14}$$

More generally, for charges that are not necessarily conserved, $\partial_\mu j^\mu$ is the amount of charge created per unit time per unit volume after taking into account the ingoing/outgoing flux $\boldsymbol{\partial}\cdot\mathbf{j}$.

Let us understand this in more detail. Consider integrating (14) over some spacetime 4-volume Ω. The result should be the amount of charge created inside this 4-volume. On the other hand, we know from Gauss' law that the integral of a divergence is a flux:

$$\int_\Omega \partial_\mu j^\mu d^4x = \int_{\partial\Omega} j^\mu dV_\mu \; . \tag{15}$$

Here, $\partial\Omega$ is the 3d boundary of the 4d region Ω, and dV_μ is a 3d volume element with direction. Note that dV_μ is naturally a covector, since its "direction" is that of a 3d surface, not a 1d line. The flux element $j^\mu dV_\mu$ is analogous to familiar 3d expressions such as $\mathbf{E}\cdot\mathbf{dS}$, where $\mathbf{E}$ is an electric field, and $\mathbf{dS}$ is a directed 2d area element.

Let us now get more specific, and consider a prism-like spacetime region Ω, composed of a spatial volume V that evolves in time for some interval $\Delta t = t_f - t_i$. The boundary $\partial\Omega$ then consists of an "initial snapshot" of $V(t_i)$ in the past, a "final snapshot" of $V(t_f)$ in the future, and a timelike "wall" consisting of the 2d boundary ∂V times the interval Δt. The flux (15) then decomposes as:

$$\int_{\partial\Omega} j^\mu dV_\mu = \int_{V(t_f)} j^t dV - \int_{V(t_i)} j^t dV + \int_{t_i}^{t_f} dt \int_{\partial V} \mathbf{j}\cdot\mathbf{dS} \tag{16}$$

Recalling that j^t is charge per unit volume and $\mathbf{j}$ is current per unit area, this becomes:

$$\int_{\partial\Omega} j^\mu dV_\mu = Q(t_f) - Q(t_i) + \int_{t_i}^{t_f} I(t)dt \ , \tag{17}$$

where Q is total charge, and I is total outgoing current. We see that the integral indeed describes the charge produced between t_i and t_f, having taken into account any ingoing/outgoing flow in the intervening time.

To get slightly philosophical, Special Relativity teaches us that "existence is a flow through time": the property of something like charge to exist in a given place – its local volume density – is actually the time component of its current! This is of a piece with the perspective switch from thinking about point particles in space to thinking about worldlines in spacetime.

IV. THE STRESS-ENERGY TENSOR

Note that the construction of 4-currents doesn't apply to every kind of non-relativistic density. For instance, mass density isn't part of any 4-vector. At best, we can think of it as an approximation for the energy density. But we can't make a 4-vector out of energy density, either: energy itself is not a scalar, but the time component of the 4-momentum p^μ. Thus, its density and current must be incorporated into a rank-2 tensor $T^{\mu\nu}$, which includes the density and current of both energy and spatial momentum. Thus, T^{tt} is energy density, T^{it} is energy current density, T^{ti} is momentum density, and T^{ij} is momentum current density.

This latter quantity is well-loved in the physics of solids, and is called the stress tensor: since the current of momentum $\dot{\mathbf{p}}$ is a force, T^{ij} measures force per unit area! As a result, $T^{\mu\nu}$ as a whole is called the "stress-energy tensor". In the same way that $\partial_\mu j^\mu = 0$ encodes the conservation of charge, $\partial_\mu T^{\mu\nu} = 0$ encodes the conservation of 4-momentum.

Let us get some initial intuition about $T^{\mu\nu}$. Consider a uniform distribution of n particles per unit volume, each with energy E, moving at the same velocity $\mathbf{v}$. The energy density is then $T^{tt} = nE$. By the logic of $\mathbf{j} = \rho\mathbf{v}$, the energy current density is then $T^{it} = T^{tt}v^i = nEv^i$. On the other hand, since $\mathbf{p} = E\mathbf{v}$, the momentum density is $T^{ti} = nEv^i$. Finally, employing $\mathbf{j} = \rho\mathbf{v}$ again, the momentum current density is $T^{ij} = T^{tj}v^i = nEv^iv^j$. We see that $T^{\mu\nu}$ is symmetric: we have $T^{ti} = T^{it}$ and $T^{ij} = T^{ji}$.

Let's repeat this construction in a more spacetime-covariant way. Consider a uniform distribution of particles, each with mass m and 4-velocity u^μ, whose density per unit volume in their rest frame is n_{rest}. As opposed to the density n in an arbitrary frame, which changes under boosts (see Exercise 6), n_{rest} is an invariant spacetime scalar. In the rest frame, $T^{\mu\nu}$ is clearly given by $T^{tt} = mn_{\text{rest}}$, with all other components vanishing. There is exactly one tensor that can be constructed out of the scalars m, n_{rest} and the 4-vector u^μ that satisfies this property: $T^{\mu\nu} = mn_{\text{rest}}u^\mu u^\nu$. By the same logic, if the particles have charge q, the associated charge 4-current is $j^\mu = qn_{\text{rest}}u^\mu$.

Exercise 7. *Consider a gas of n particles per unit volume. Each particle has the same mass m and velocity of magnitude v. The directions of the particles' velocities are uniformly distributed. Show that the stress-energy tensor has the form:*

$$T^{\mu\nu} = \begin{pmatrix} \varepsilon & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix} , \tag{18}$$

and find the values of ε and p. What is the physical meaning of p?

Exercise 8. *Now, consider a gas of n photons per unit volume. Each photon has energy E, and the directions of the photons' velocities are again uniformly distributed. Show that $T^{\mu\nu}$ is again of the form* (18)*, and find ε and p. What is the trace of the stress-energy tensor T^μ_μ? Can you relate the answer to some property of the 4-momentum of a single photon?*

Fields also have a stress-energy tensor, but that is a more delicate issue, as we will see in the next lecture.

V. ELECTROMAGNETISM

In electrostatics, we learn about the electric potential ϕ. Quite a bit later, we learn about the magnetic potential $\mathbf{A}$. In full electrodynamics, the electric and magnetic field are derived from these potentials via:

$$\mathbf{E} = -\boldsymbol{\partial}\phi - \frac{\partial \mathbf{A}}{\partial t} \ ; \quad \mathbf{B} = \boldsymbol{\partial} \times \mathbf{A} \ . \tag{19}$$

It turns out that, just like energy and momentum, the electric and magnetic potentials also combine into a 4-vector $A^\mu = (\phi, \mathbf{A})$ – the electromagnetic potential. The electric and magnetic field strengths (19) can now be rewritten as:

$$E_i = \partial_i A_t - \partial_t A_i \ ; \quad B_{ij} = 2\partial_{[i}A_{j]} = \partial_i A_j - \partial_j A_i \ , \tag{20}$$

where we replaced the axial vector B_i with a bivector B_{ij}, as in Lecture 1-1. We see that $\mathbf{E}$ and $\mathbf{B}$ are also components of a single spacetime object – an antisymmetric matrix $F_{\mu\nu}$:

$$F_{\mu\nu} = 2\partial_{[\mu}A_{\nu]} = \partial_\mu A_\nu - \partial_\nu A_\mu \ ; \quad E_i = F_{it} = -F_{ti} \ ; \quad \epsilon_{ijk}B_k = B_{ij} = F_{ij} \ . \tag{21}$$

In other words, an electric field is just like a magnetic field, but in a timelike plane! We refer to $F_{\mu\nu}$ as the electromagnetic field strength.

The 4-potential A_μ "wants" to have a lower index, due to the way in which it enters the action of a charged particle:

$$S = -m\int\sqrt{-dx_\mu dx^\mu} + q\int A_\mu dx^\mu \ . \tag{22}$$

Quite remarkably, the second term in (22) completely captures the interaction between a charged particle and an electromagnetic field: the covector A_μ simply defines an "extra bit of action" for a charged particle traveling along an interval dx^μ!

The derivation of the charge's equation of motion from (22) isn't difficult, but we will simply state the results here. A very flexible form of the equations that is also close to a Newtonian force law is given by:

$$dp_\mu = qF_{\mu\nu}dx^\nu \ , \tag{23}$$

where $p_\mu = m u_\mu$ is the particle's 4-momentum. Dividing by the proper time $d\tau$, we get the 4-acceleration in an EM field as:

$$\alpha_\mu = \frac{q}{m} F_{\mu\nu} u^\nu \ . \tag{24}$$

Exercise 9. *Derive from* (23) *the Lorentz force law* $\mathbf{F} = q(\mathbf{E}+\mathbf{v}\times\mathbf{B})$*, which is exact even at relativistic velocities, if we define the force* $\mathbf{F}$ *as the time derivative of the correct relativistic momentum* $\mathbf{p} = m\mathbf{v}/\sqrt{1-v^2}$*.*