

12 Relativity

Postulates

- (a) All inertial observers have the same equations of motion and the same physical laws. Relativity explains how to translate the measurements and events according to one inertial observer to another.
- (b) The speed of light is constant for all inertial frames

12.1 Elementary Relativity

Mechanics of indices, four-vectors, Lorentz transformations

- (a) We describe physics as a sequence of events labelled by their space time coordinates:

$$x^\mu = (x^0, x^1, x^2, x^3) = (ct, \mathbf{x}) \quad (12.1)$$

The space time coordinates of another inertial observer moving with velocity \mathbf{v} relative to the first measures the coordinates of an event to be

$$\underline{x}^\mu = (\underline{x}^0, \underline{x}^1, \underline{x}^2, \underline{x}^3) = (\underline{ct}, \underline{\mathbf{x}}) \quad (12.2)$$

- (b) The coordinates of an event according to the first observer x^μ determine the coordinates of an event according to another observer \underline{x}^μ through a linear change of coordinates known as a Lorentz transformation:

$$x^\mu \rightarrow \underline{x}^\mu = L^\mu_\nu(\mathbf{v})x^\nu \quad (12.3)$$

I usually think of x^μ as a column vector

$$\begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \quad (12.4)$$

so that without indices the transform

$$x \rightarrow \underline{x} = (L) x \quad (12.5)$$

Then to change frames from K to an observer \underline{K} moving to the right with speed v relative to K the transformation matrix is

$$L^\mu_\nu = \begin{pmatrix} \gamma_v & -\gamma\beta & & \\ -\gamma\beta & \gamma & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \quad (12.6)$$

with $\beta = v/c$ and $\gamma = 1/\sqrt{1 - \beta^2}$.

A short exercise done in class shows that a this boost contracts the $x^+ \equiv x^0 + x^1$ direction (*i.e.* $ct + x$) and expands the $x^- \equiv x^0 - x^1$ direction (*i.e.* $ct - x$). Thus, x^+ and x^- are eigenvectors of Lorentz

boosts in the x direction

$$\underline{x}^+ = \sqrt{\frac{1-\beta}{1+\beta}} x^+ \quad (12.7)$$

$$\underline{x}^- = \sqrt{\frac{1+\beta}{1-\beta}} x^- \quad (12.8)$$

(c) Instead of using v we sometimes use the rapidity y

$$\tanh y = \frac{v}{c} \quad \text{or} \quad y = \frac{1}{2} \ln \frac{1+\beta}{1-\beta} \quad (12.9)$$

and note that $y \simeq \beta$ for small β

With this parametrization we find that the Lorentz boost appears as a hyperbolic rotation matrix

$$L^\mu_\nu = \begin{pmatrix} \cosh y & -\sinh y & & \\ -\sinh y & \cosh y & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \quad (12.10)$$

Then

$$\underline{x}^+ = e^{-y} x^+ \quad \underline{x}^- = e^y x^- \quad (12.11)$$

(d) Since the speed of light is constant for all observers we demand that

$$-(ct)^2 + \mathbf{x}^2 = -(\underline{ct})^2 + \underline{\mathbf{x}}^2 \quad (12.12)$$

under Lorentz transformation. We also require that the set of Lorentz transformations satisfy the follow (group) requirements:

$$L(-\mathbf{v})L(\mathbf{v}) = \mathbb{I} \quad (12.13)$$

$$L(\mathbf{v}_2)L(\mathbf{v}_1) = L(\mathbf{v}_3) \quad (12.14)$$

here \mathbb{I} is the identity matrix. These properties seem reasonable to me, since if I transform to frame moving with velocity \mathbf{v} and then transform back to a frame moving with velocity $-\mathbf{v}$, I should get back the same result. Similarly two Lorentz transformations produce another Lorentz transformation.

(e) Since the combination

$$-(ct)^2 + \mathbf{x}^2 \quad (12.15)$$

is invariant under Lorentz transformation, we introduced an index notation to make such invariant forms manifest. We formalized the lowering of indices

$$x_\mu = g_{\mu\nu} x^\nu \quad x_\mu = (-ct, \mathbf{x}) \quad (12.16)$$

with a metric tensor:

$$g_{00} = -1 \quad g_{11} = g_{22} = g_{33} = 1 \quad (12.17)$$

In this way we define a dot product

$$x \cdot x = x^\mu x_\mu = -(ct)^2 + \mathbf{x}^2 \quad (12.18)$$

is manifestly invariant.

Similarly we raise indices

$$x^\mu = g^{\mu\nu} x_\nu \quad (12.19)$$

with

$$g^{\mu\nu} = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \quad (12.20)$$

Of course the process of lowering and index and then raising it again does nothing:

$$g^\mu_\nu = g^{\mu\sigma} g_{\sigma\nu} = \delta^\mu_\nu = \text{identity matrix} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \quad (12.21)$$

- (f) Generally the upper indices are “the normal thing”. We will try to leave the dimensions and name of the four vector, corresponding to that of the spatial components. Examples: $x^\mu = (ct, \mathbf{x})$, $A^\mu = (\Phi, \mathbf{A})$, $J^\mu = (c\rho, \mathbf{j})$, and $P^\mu = (E/c, \mathbf{p})$.
- (g) Four vectors are anything that transforms according to the lorentz transformation $A^\mu = (A^0, \mathbf{A})$ like coordinates

$$A^\mu = L^\mu_\nu A^\nu \quad (12.22)$$

Given two four vectors, A^μ and B^μ one can always construct a Lorentz invariant quantity.

$$A \cdot B = A_\mu B^\mu = A^\mu g_{\mu\nu} B^\nu = -A^0 B^0 + \mathbf{A} \cdot \mathbf{B} = -\underline{A}^0 \underline{B}^0 + \underline{\mathbf{A}} \cdot \underline{\mathbf{B}} = \underline{A}^\mu g_{\mu\nu} \underline{B}^\nu = \underline{A}_\mu \underline{B}^\mu = \underline{A} \cdot \underline{B} \quad (12.23)$$

- (h) From the invariance of the inner product we see that the lower (covariant) components of four vectors transform with the inverse transformation and as a row,

$$x_\mu \rightarrow \underline{x}_\nu = x_\mu (L^{-1})^\mu_\nu. \quad (12.24)$$

I usually think of x_μ (with a lower index) as a row

$$(x_0 \ x_1 \ x_2 \ x_3) \quad (12.25)$$

So the transformation rule in terms of matrices is

$$(\underline{x}_0 \ \underline{x}_1 \ \underline{x}_2 \ \underline{x}_3) = (x_0 \ x_1 \ x_2 \ x_3) \begin{pmatrix} L^{-1} \end{pmatrix} \quad (12.26)$$

In this way the inner product

$$\underline{A}_\mu \underline{B}^\mu = (A_0 \ A_1 \ A_2 \ A_3) \begin{pmatrix} L^{-1} \end{pmatrix} \begin{pmatrix} L \end{pmatrix} \begin{pmatrix} B^0 \\ B^1 \\ B^2 \\ B^3 \end{pmatrix} = A_\mu B^\mu \quad (12.27)$$

is invariant. If you wish to think of x_μ as a column, then it transforms under lorentz transformation with the inverse transpose matrix

$$\begin{pmatrix} \underline{x}_0 \\ \underline{x}_1 \\ \underline{x}_2 \\ \underline{x}_3 \end{pmatrix} = \begin{pmatrix} L^{-1\top} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad (12.28)$$

- (i) As is clear from Eq. (12.23), the metric tensor is an invariant tensor, *i.e.*

$$g^{\mu\nu} = L^\mu_\rho L^\nu_\sigma g^{\rho\sigma} \quad (12.29)$$

is the same tensor $\text{diag}(-1, 1, 1, 1)$ in all frames (so I don't need to put an underline $\underline{g}^{\mu\nu}$ on the LHS). From Eq. (12.29) it follows that the inverse (transpose) Lorentz transform can be found by raising and lowering the indices of the transform matrix, *i.e.*

$$L_\rho^\sigma \equiv g_{\rho\mu} L_\nu^\mu g^{\nu\sigma} = (L^{-1\top})_\rho^\sigma \quad (12.30)$$

where we have defined L_ρ^σ . Thus if one wishes to think of a lowered four vector A_μ as a column, one has

$$\underline{A}_\nu = L_\nu^\mu A_\mu \quad (12.31)$$

Thus, a short exercise (done) in class shows that if

$$\underline{T}^{\mu\nu} = L_\rho^\mu L_\sigma^\nu T^{\rho\sigma} \quad (12.32)$$

then there is a consistency check

$$\underline{T}_\nu^\mu = L_\sigma^\mu L_\nu^\rho T_\rho^\sigma = L_\sigma^\mu T_\rho^\sigma (L^{-1})_\nu^\rho \quad (12.33)$$

Doppler shift, four velocity, and proper time.

- (a) The frequency and wave number form a four vector $K^\mu = (\frac{\omega}{c}, \mathbf{k})$, with $|\mathbf{k}| = \omega/c$. This can be used to determine a relativistic doppler shift.
- (b) For a particle in motion with velocity \mathbf{v}_p and gamma factor γ_p , the space-time interval is

$$ds^2 \equiv dx_\mu dx^\mu = -(cdt)^2 + d\mathbf{x}^2 = -(cd\tau)^2. \quad (12.34)$$

ds^2 is associated with the clicks of the clock in the particles instantaneous rest frame, $ds^2 = -(cd\tau)^2$, so we have in any other frame

$$d\tau \equiv \sqrt{-ds^2}/c = dt \sqrt{1 - \left(\frac{dx}{dt}\right)^2 / c^2} \quad (12.35)$$

$$= \frac{dt}{\gamma_p} \quad (12.36)$$

- (c) The four velocity of a particle is the distance the particle travels per proper time

$$U^\mu \equiv \frac{dx^\mu}{d\tau} = (u^0, \mathbf{u}) = (\gamma_p c, \gamma_p \mathbf{v}_p) \quad (12.37)$$

so

$$\underline{U}^\mu = L_\nu^\mu U^\nu \quad (12.38)$$

Note $U_\mu U^\mu = -c^2$.

- (d) The transformation of the four velocity under Lorentz transformation should be compared to the transformation of velocities. For a particle moving with velocity \mathbf{v}_p in frame K , then in another frame \underline{K} moving to the right with speed v the particle moves with velocity

$$\mathbf{v}_p^\parallel = \frac{v_p^\parallel - v}{1 - v_p^\parallel v / c^2} \quad (12.39)$$

$$\mathbf{v}_p^\perp = \frac{v_p^\perp}{\gamma_p (1 - v_p^\parallel v / c^2)} \quad (12.40)$$

where v_p^\parallel and v_p^\perp are the components of \mathbf{v}_p parallel and perpendicular to v . These are easily derived from the transformation rules of U^μ and the fact that $\mathbf{v}_p = \mathbf{u}/u^0$.

Energy and Momentum Conservation

- (a) Finally the energy and momentum form a four vector

$$P^\mu = \left(\frac{E}{c}, \mathbf{p} \right) \quad (12.41)$$

The invariant product of P^μ with itself the rest energy

$$P^\mu P_\mu = -(mc)^2 \quad (12.42)$$

This can be inverted giving the energy in terms of the momentum, *i.e.* the dispersion curve

$$\frac{E(p)}{c} = \sqrt{p^2 + (mc)^2} \quad (12.43)$$

- (b) The relation between energy and momentum determines the velocity. At rest $E = mc^2$. Then a boost in the negative $-\mathbf{v}_p$ direction shows that a particle with velocity \mathbf{v}_p has energy and momentum

$$P^\mu = \left(\frac{E}{c}, \mathbf{p} \right) = mc(\gamma_p, \gamma_p \beta_p) = mU^\mu \quad (12.44)$$

i.e.

$$\mathbf{v}_p = c \frac{\mathbf{p}}{(E/c)} = \frac{\partial E(p)}{\partial \mathbf{p}} \quad (12.45)$$

Thus as usual the derivative of the dispersion curve is the velocity.

- (c) Energy and Momentum are conserved in collisions, e.g. for a reaction $1 + 2 \rightarrow 3 + 4$ we have

$$P_1^\mu + P_2^\mu = P_3^\mu + P_4^\mu \quad (12.46)$$

Usually when working with collisions it makes sense to suppress c or just make the association:

$$\begin{pmatrix} E \\ p \\ m \end{pmatrix} \quad \text{is short for} \quad \begin{pmatrix} E \\ cp \\ mc^2 \end{pmatrix} \quad (12.47)$$

A starting point for analyzing the kinematics of a process is to “square” both sides with the invariant dot product $P^2 \equiv P \cdot P$. For example if $P_1 + P_2 = P_3 + P_4$ then:

$$(P_1 + P_2)^2 = (P_3 + P_4)^2 \quad (12.48)$$

$$P_1^2 + P_2^2 + 2P_1 \cdot P_2 = P_3^2 + P_4^2 + 2P_3 \cdot P_4 \quad (12.49)$$

$$-m_1^2 - m_2^2 - 2E_1 E_2 + 2\mathbf{p}_1 \cdot \mathbf{p}_2 = -m_3^2 - m_4^2 - 2E_3 E_4 + 2\mathbf{p}_3 \cdot \mathbf{p}_4 \quad (12.50)$$

12.2 Covariant form of electrodynamics

(a) The players are:

i) The derivatives

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = \left(\frac{1}{c} \frac{\partial}{\partial t}, \nabla \right) \quad (12.51)$$

$$\partial^\mu \equiv \frac{\partial}{\partial x_\mu} = \left(-\frac{1}{c} \frac{\partial}{\partial t}, \nabla \right) \quad (12.52)$$

ii) The wave operator

$$\square = \partial_\mu \partial^\mu = \frac{-1}{c^2} \frac{\partial}{\partial t^2} + \nabla^2 \quad (12.53)$$

iii) The four velocity $U^\mu = (u^0, \mathbf{u}) = (\gamma_p, \gamma_p \mathbf{v}_p)$

iv) The current four vector

$$J^\mu = (c\rho, \mathbf{J}) \quad (12.54)$$

v) The vector potential

$$A^\mu = (\Phi, \mathbf{A}) \quad (12.55)$$

vi) The field strength is a tensor

$$F^{\alpha\beta} = \partial^\alpha A^\beta - \partial^\beta A^\alpha \quad (12.56)$$

which ultimately comes from the relations

$$\mathbf{E} = -\frac{1}{c} \partial_t \mathbf{A} - \nabla \Phi \quad (12.57)$$

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (12.58)$$

In indices we have

$$F^{0i} = E^i \quad E^i = F^{0i} \quad (12.59)$$

$$F^{ij} = \epsilon^{ijk} B_k \quad B_i = \frac{1}{2} \epsilon_{ijk} F^{jk} \quad (12.60)$$

In matrix form this anti-symmetric tensor reads

$$F^{\alpha\beta} = \begin{pmatrix} 0 & E^x & E^y & E^z \\ -E^x & 0 & B^z & -B^y \\ -E^y & -B^z & 0 & B^x \\ -E^z & B^y & -B^x & 0 \end{pmatrix} \quad (12.61)$$

Raising and lowering indices of $F^{\mu\nu}$ can change the sign of the zero components, but does not change the ij components, *e.g.*

$$E^i = F^{0i} = -F^{i0} = F^i{}_0 = -F_0{}^i = -F_{0i} = F^0{}_i = F^{0i} \quad (12.62)$$

vii) The dual field tensor implements the replacement

$$\mathbf{E} \rightarrow \mathbf{B} \quad \mathbf{B} \rightarrow -\mathbf{E} \quad (12.63)$$

As motivated by the maxwell equations in free space

$$\nabla \cdot \mathbf{E} = 0 \quad (12.64)$$

$$-\frac{1}{c} \partial_t \mathbf{E} + \nabla \times \mathbf{B} = 0 \quad (12.65)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (12.66)$$

$$-\frac{1}{c} \partial_t \mathbf{B} - \nabla \times \mathbf{E} = 0 \quad (12.67)$$

which are the same before and after this duality transformation. The dual field strength tensor is

$$\mathcal{F}^{\alpha\beta} = \begin{pmatrix} 0 & B^x & B^y & B^z \\ -B^x & 0 & -E^z & E^y \\ -B^y & E^z & 0 & -E^x \\ -B^z & -E^y & -E^x & 0 \end{pmatrix} \quad (12.68)$$

The dual field strength tensor

$$\mathcal{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}F_{\rho\sigma} \quad (12.69)$$

where the totally anti-symmetric tensor $\epsilon^{\mu\nu\rho\sigma}$ is

$$\epsilon^{\mu\nu\rho\sigma} = \begin{cases} +1 & \text{even perms } 0,1,2,3 \\ -1 & \text{odd perms } 0,1,2,3 \\ 0 & 0 \text{ otherwise} \end{cases} \quad (12.70)$$

viii) The stress tensor is

$$\Theta_{\text{em}}^{\mu\nu} = F^{\mu\lambda}F_{\lambda}^{\nu} + g^{\mu\nu} \left(-\frac{1}{4}F_{\alpha\beta}F^{\alpha\beta}\right) \quad (12.71)$$

Or in terms of matrices

$$\Theta_{\text{em}}^{\mu\nu} = \left(\begin{array}{c|c} u_{\text{em}} & \mathbf{S}_{\text{em}}/c \\ \hline \mathbf{S}_{\text{em}}/c & T^{ij} \end{array} \right) \quad (12.72)$$

Note that $\Theta^{0i} = \mathbf{S}_{\text{em}}^i/c = c\mathbf{g}_{\text{em}}^i$, and $T^{ij} = (-E^iE^j + \frac{1}{2}\delta^{ij}E^2) + (-B^iB^j + \frac{1}{2}\delta^{ij}B^2)$. You can remember the stress tensor $\Theta^{\mu\nu}$ by recalling that it is quadratic in F , symmetric under interchange of μ and ν , and traceless $\Theta_{\mu}^{\mu} = 0$. These properties fix the stress tensor up to a constant.

(b) The equations are

i) The continuity equation:

$$\partial_{\mu}J^{\mu} = 0 \quad (12.73) \qquad \partial_t\rho + \nabla \cdot \mathbf{J} = 0 \quad (12.74)$$

ii) The wave equation in the covariant gauge

$$-\square A^{\mu} = J^{\mu}/c \quad (12.75) \qquad -\square\Phi = \rho \quad (12.76)$$

$$-\square\mathbf{A} = \mathbf{J}/c \quad (12.77)$$

This is true in the covariant gauge

$$\partial_{\mu}A^{\mu} = 0 \quad (12.78) \qquad \frac{1}{c}\partial_t\Phi + \nabla \cdot \mathbf{A} = 0 \quad (12.79)$$

iii) The force law is:

$$\frac{dP^{\mu}}{d\tau} = eF_{\nu}^{\mu}\frac{U^{\nu}}{c} \quad (12.80) \qquad \frac{1}{c}\frac{dE}{dt} = e\mathbf{E} \cdot \frac{\mathbf{v}}{c} \quad (12.81)$$

$$\frac{d\mathbf{p}}{dt} = e\mathbf{E} + e\frac{\mathbf{v}}{c} \times \mathbf{B} \quad (12.82)$$

If these equations are multiplied by γ they equal the relativistic equations to the left.

iv) The sourced field equations are :

$$-\partial_\mu F^{\mu\nu} = \frac{J^\nu}{c} \quad (12.83) \quad \nabla \cdot \mathbf{E} = \rho \quad (12.84)$$

$$-\frac{1}{c} \partial_t \mathbf{E} + \nabla \times \mathbf{B} = \frac{\mathbf{J}}{c} \quad (12.85)$$

v) The dual field equations are :

$$-\partial_\mu \mathcal{F}^{\mu\nu} = 0 \quad (12.86) \quad \nabla \cdot \mathbf{B} = 0 \quad (12.87)$$

$$-\frac{1}{c} \partial_t \mathbf{B} - \nabla \times \mathbf{E} = 0 \quad (12.88)$$

as might have been inferred by the replacements $\mathbf{E} \rightarrow \mathbf{B}$ and $\mathbf{B} \rightarrow -\mathbf{E}$. The dual field equations can also be written in terms $F_{\mu\nu}$, and this is known as the Bianchi identity:

$$\partial_\rho F_{\mu\nu} + \partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} = 0, \quad (12.89)$$

where ρ, μ, ν are cyclic.

Or (for the mathematically inclined) the Bianchi identity reads

$$\partial_{[\mu_1} F_{\mu_2\mu_3]} = 0, \quad (12.90)$$

where the square brackets denote the fully antisymmetric combination of μ_1, μ_2, μ_3 , *i.e.* the order is like a determinant

$$\begin{aligned} \partial_{[\mu_1} F_{\mu_2\mu_3]} \equiv \frac{1}{3!} [& (\partial_{\mu_1} F_{\mu_2\mu_3} - \partial_{\mu_2} F_{\mu_1\mu_3} + \partial_{\mu_3} F_{\mu_1\mu_2}) \\ & + (-\partial_{\mu_1} F_{\mu_3\mu_2} + \partial_{\mu_2} F_{\mu_3\mu_1} - \partial_{\mu_3} F_{\mu_2\mu_1})] \end{aligned} \quad (12.91)$$

The second line is the same as the first since $F_{\mu\nu}$ is antisymmetric. Eq. (12.90) is the statement that $F_{\mu\nu}$ is an exact differential form.

vi) The dual field equations are equivalent to the statement that that $F_{\mu\nu}$ (or \mathbf{E}, \mathbf{B}) can be written in terms of the gauge potential A_μ (or Φ, \mathbf{A})

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (12.92) \quad \mathbf{B} = \nabla \times \mathbf{A} \quad (12.93)$$

$$\mathbf{E} = -\frac{1}{c} \partial_t \mathbf{A} - \nabla \Phi \quad (12.94)$$

The potentials are not unique as we can always make a gauge transform:

$$A_\mu \rightarrow A_\mu + \partial_\mu \Lambda \quad (12.95) \quad \mathbf{A} \rightarrow \mathbf{A} + \nabla \Lambda \quad (12.96)$$

$$\Phi \rightarrow \Phi + \frac{1}{c} \partial_t \Lambda \quad (12.97)$$

vii) The conservation of energy and momentum can be written in terms of the stress tensor:

$$-\partial_\mu \Theta_{\text{em}}^{\mu\nu} = F_\nu^\mu \frac{J^\nu}{c} \quad (12.98) \quad -\left(\frac{1}{c} \frac{\partial u_{\text{em}}}{\partial t} + \nabla \cdot (\mathbf{S}_{\text{em}}/c) \right) = \mathbf{E} \cdot \mathbf{J}/c \quad (12.99)$$

$$-\left(\frac{1}{c} \frac{\partial \mathbf{S}_{\text{em}}^j}{\partial t} + \partial_i T^{ij} \right) = \rho E^j + (\mathbf{J}/c \times \mathbf{B})^j \quad (12.100)$$

The energy and momentum transferred from the fields $F^{\mu\nu}$ to the particles is

$$\partial_\mu \Theta_{\text{mech}}^{\mu\nu} = F_\nu^\mu \frac{J^\nu}{c} \quad (12.101)$$

Or

$$\partial_\mu \Theta_{\text{mech}}^{\mu\nu} + \partial_\mu \Theta_{\text{em}}^{\mu\nu} = 0 \quad (12.102)$$