C7.5 Lecture 4: Special relativity

Joe Keir

Joseph.Keir@maths.ox.ac.uk

Notation and conventions

- The signature of the metric is $(-1, 1, 1, 1)$.
- We *always* use the Einstein summation convention: repeated indices (one appears raised and one lowered) are summed over, e.g.

$$A^a \mu_a := \sum_{i=0}^{3} A^i \mu_i.$$

- Greek letters will be used for *abstract indices*: these don't refer to any particular coordinate system but just tell us what kind of object we're dealing with: e.g. A^μ means "the vector A", ω_μ means "the covector ω (vectors, covectors etc. defined later).
- Latin letters from the start of the alphabet $(a, b, c \dots)$ refer to *concrete indices* – these some specific coordinate system, and run from 0 to 3.
- Latin letters from the middle of the alphabet $(i, j, k \dots)$ refer to *spatial* components within some chosen coordinate system, and run from 1 to 3.

The metric and causal structure

In inertial coordinates $x^a = (x^0, x^1, x^2, x^3)$, the Minkowski metric m_{ab} has components

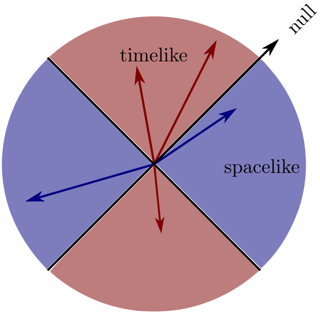
$$m := \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The inverse metric m^{-1} has the same components, so, for example, $m_{00} = (m^{-1})^{00} = -1$, and $m_{23} = (m^{-1})^{13} = 0$. The fact that these are inverses of one another is expressed in the relation

$$m_{ac}(m^{-1})^{cb} = \delta_a^b,$$

where $\delta_a^b = \mathrm{diag}(1, 1, 1, 1)$ are the components of the identity map.

Given a nonzero spacetime vector v, we say v is *spacelike* if $m(v,v) = m_{ab}v^a v^b > 0$, v is *timelike* if $m(v,v) < 0$ and v is *null* if $m(v,v) = 0$.
Likewise, the points p and q in spacetime are *spacelike separated* if the vector $(p-q)$ is spacelike, *timelike separated* if $(p-q)$ is timelike, and *null separated* if $(p-q)$ is null.
The set of all points q which are null separated from the point p is the *light cone* of the point p.



Timelike (red), spacelike (blue) and null (black) vectors, in a sketch where two spatial dimensions have been suppressed. The set of timelike vectors has two connected components, which we call *past directed* and *future directed* vectors. Despite its appearance in this sketch, the set of spacelike vectors has only a single connected component: one spatial dimension can be restored by revolving this diagram around the vertical axis.

Lorentz transformations

Transformations from one set of inertial coordinates (x^a) to another $(y^{a'})$, fixing the origin. Given by linear transformations

$$y^{a'} = \Lambda_a{}^{a'} x^a$$

where the form of the metric is preserved:

$$m_{ab} = \Lambda_a{}^{a'} \Lambda_b{}^{b'} m_{a'b'}.$$

There are boosts, which mix the time and space coordinates - e.g. a boost with relative velocity v in the x direction is

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

$$\gamma = \frac{1}{\sqrt{1-v^2}}$$

and rotations, which leave time invariant:

$$\Lambda_a{}^{a'} = \begin{pmatrix} 1 & 0 \\ 0 & R_i{}^j \end{pmatrix}$$

with $R_i{}^j \in SO(3)$.

Curves and tangent vectors

A *curve* is a map $\gamma : [0,1]$ (or $\mathbb{R}$ or $\mathbb{R}^+$) $\to \mathbb{M}^4$. In inertial coordinates, we can write $\gamma(\lambda) = x^a(\lambda)$.

The *tangent vector* to the curve γ at the point p with coordinate $x^a(\lambda_0)$ is

$$v^a\big|_p := \frac{\mathrm{d}}{\mathrm{d}\lambda} x^a(\lambda)\big|_{\lambda=\lambda_0}.$$

A tangent vector can be timelike, spacelike or null, and this characteristic is independent of the parametrisation of the curve.

If the tangent to a curve is *everywhere* timelike/spacelike/null, then the curve is said to be timelike/spacelike/null.

Derivatives along curves

We can use tangent vectors to measure the rate of change of a function along a curve: suppose $f : \mathbb{M}^4 \to \mathbb{R}$. Then, using the chain rule,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(f \circ \gamma(\lambda)\right) = \frac{\mathrm{d}\gamma^a}{\mathrm{d}\lambda}\frac{\partial f}{\partial x^a},$$

where $\gamma^a(\lambda) = x^a(\lambda)$ are the coordinates of the curve in the coordinates (x^a). Hence

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(f \circ \gamma(\lambda)\right) = v^a \frac{\partial f}{\partial x^a}.$$

Proper time

If γ is a timelike curve, then the *proper time* along γ, τ, is the parameter along the curve so that the tangent vector satisfies

$$m(v, v) = m_{ab} \left(\frac{\mathrm{d}}{\mathrm{d}\tau} x^a(\tau) \right) \left(\frac{\mathrm{d}}{\mathrm{d}\tau} x^b(\tau) \right) = -1.$$

Such a parameter can always be found along a timelike curve. It is unique up to the choice of origin, that is, the point along the curve at which $\tau = 0$. It has a physical meaning in special relativity, given by:

The clock postulate

An accurate clock moving along a timelike worldline measures the proper time along the worldline.

Proper length and affine parameters

If γ is a spacelike curve, then the *proper length s* of γ is the parameter defined so that

$$m_{ab}\frac{\mathrm{d}x^a}{\mathrm{d}s}\frac{\mathrm{d}x^b}{\mathrm{d}s} = 1.$$

No analogue to proper time or distance along a generic null curve. However, there are special null curves, which are *generated* by a null vector as follows: let $p \in \mathbb{M}^4$ be some fixed point, and let v be a null vector. Consider the curve

$$\gamma(\lambda) - p = \lambda v.$$

This is a null curve with tangent vector v. The parameter λ is called an *affine parameter*. We can find another affine parametrisation for such a curve by choosing a different point p along the curve, and choosing a different null vector which is proportional to v. Under such a reparametrisation, $\lambda \mapsto a\lambda + b$, for constants a and b.

For $p \in \mathbb{M}^4$, the *tangent space* at p is the vector space consisting of all the vectors from the point p:

$$T_p(\mathbb{M}^4) := \{(q - p) \in \mathbb{R}^4 \,\big|\, q \in \mathbb{M}^4\}$$

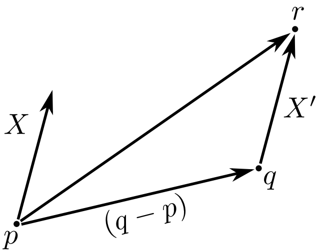
Equivalently, we can define $T_p(\mathbb{M}^4)$ as the set of all tangent vectors to curves through the point p.

There is a natural way to identify 'different' tangent spaces, i.e. the tangent spaces $T_p(\mathbb{M}^4)$ and $T_q(\mathbb{M}^4)$ with $q \neq p$.

Let $X \in T_p(\mathbb{M}^4)$; then there is some point $r \in \mathbb{M}^4$ such that $X + (q - p) = (r - p)$. Then we can define the vector $X' \in T_q(\mathbb{M}^4)$, which "corresponds" to $X \in T_p(\mathbb{M}^4)$, as

$$X' := (r - q).$$

Although this identification seems trivial, the lack of an analogue in curved spacetimes leads to a lot of complications in GR.



The vector transformation rule

Take inertial coordinates x^a, chosen so that p is at the origin $x^a = 0$. The point q has coordinates $x^a = q^a$. Under a Lorentz transformation (at p) $y^{a'} = \Lambda_a{}^{a'} x^a$, the new coordinates of the point q are

$$\begin{aligned} q^a &= x^a \\ &= (\Lambda^{-1})_{a'}{}^{a} y^{a'} \\ \Rightarrow y^{a'} &= \Lambda_a{}^{a'} q^a \end{aligned}$$

In other words, with respect to the $y^{a'}$ coordinates, the coordinates of q are

$$q'^{a'} = \Lambda_a{}^{a'} q^a.$$

Hence, the components of the vector $X = q - p$ transform as

$$X'^{a'} = \Lambda_a{}^{a'} X^a. \tag{1}$$

This is the *transformation law for the components of a vector*. Equation (1) is sometimes used to *define* vectors, i.e. a vector is any quantity which transforms under this rule.

Covectors and their transformation rule

The *cotangent space* at the point p, $T_p^*(\mathbb{M}^4)$, is the dual of the tangent space $T_p(\mathbb{M}^4)$, i.e. the space of linear maps from $T_p(\mathbb{M}^4)$ to $\mathbb{R}$. Elements of this space are *covectors* (or sometimes *1-forms*).

We can use the transformation law for vector components to find the corresponding transformation law for covectors: the components of a covector

$$\eta \in T_p^*(\mathbb{M}^4)$$

with respect to some inertial coordinates (chosen so that p is at the origin) are the four numbers η_a where, for any vector $X \in T_p(\mathbb{M}^4)$,

$$\eta(X) = \eta_a X^a,$$

where X^a are the components of the vector X in the inertial coordinates.

$\eta(X)$ is just a real number (a *scalar*) – it doesn't transform at all under Lorentz transformations! So, if $\eta'_{a'}$ are the components of η with respect to the inertial coordinates $y^{a'} = \Lambda_a{}^{a'} x^a$, we must have

$$\begin{aligned} \eta_a X^a &= \eta'_{a'} X'^{a'} \\ &= \eta'_{a'} \Lambda_a{}^{a'} X^a \end{aligned}$$

and so the *transformation law for covectors* is

$$\eta'_{a'} = (\Lambda^{-1})_{a'}{}^{a} \eta_a.$$

Sometimes say that vectors transform *contravariantly*, while covectors transform *covariantly*.

Tensors

A *tensor* at the point p is an element of $\left(T_p(\mathbb{M}^4)\right)^n \times \left(T_p^*(\mathbb{M}^4)\right)^m$ for some n, $m \geq 0$. Such a tensor is of *rank* (n, m), or has *valency* (n, m).

For example, given a vector X and a covector η we can form the tensor $X\eta$, with components (in *any* coordinate system)

$$(X\eta)^\mu{}_\nu := X^\mu \eta_\nu.$$

It is often useful to view a (n, m) tensor as a linear map from $\left(T_p(\mathbb{M}^4)\right)^m \times \left(T_p^*(\mathbb{M}^4)\right)^n \to \mathbb{R}$.

Tensor transformations

Tensors transform under Lorentz transformations in the obvious way: for a rank (n, m) tensor T, its components transform as

$$(T')^{a'_1 a'_2 \dots a'_n}{}_{b'_1 b'_2 \dots b'_m} = \Lambda_{a_1}{}^{a'_1} \dots \Lambda_{a_n}{}^{a'_n} (\Lambda^{-1})_{b'_1}{}^{b_1} \dots (\Lambda^{-1})_{b'_m}{}^{b_m} T^{a_1 \dots a_n}{}_{b_1 \dots b_m}.$$

A *contraction* of a tensor is formed by summing over a pair of indices, with one "up" index and one "down" index. Using the Einstein summation convention, this is written as a tensor with the same letter used in one of the "up" indexes and one of the "down" indexes, e.g. $T^b{}_{ba}$. These are also the components of a tensor (check the transformation rules).

Indices can be *lowered* and *raised* using the metric m and its inverse m^{-1}. The metric defines an isomorphism $T_p(\mathbb{M}) \to T_p^*(\mathbb{M})$ as follows: for a vector X, and an arbitrary vector Y

$$X \mapsto X^\flat$$
$$X^\flat(Y) = m(X, Y),$$

or, in terms of indices (in which case it is conventional to avoid the 'flat' sign)

$$X_\mu := m_{\mu\nu} X^\nu.$$

Similarly for a covector η and an arbitrary vector Y,

$$\eta \mapsto \eta^\sharp$$
$$m(\eta^\sharp, Y) = \eta(Y),$$

or in terms of indices

$$\eta^\mu := (m^{-1})^{\mu\nu} \eta_\nu.$$

Tensor fields

A *tensor field* is an assignment of a tensor to all points in spacetime.

Example 1: the metric m is a rank $(0,2)$ tensor field whose action on the vector fields X, Y is given by

$$m(X,Y) := m_{ab}X^aX^b$$

where, on the right hand side, we work in an inertial coordinate system, and $m_{ab} = \mathrm{diag}(-1,1,1,1)$. Note that, because of the special properties of the metric tensor, the metric takes this form in *all* inertial coordinate systems.

Example 2” the *identity* or *Kronecker delta* is a $(1,1)$ tensor field whose action on any vector field X and covector field η is

$$\delta(X,\eta) := \eta(X).$$

Easy to check that, at each point p, this is a linear map $T_p(\mathbb{M}^4) \times T_p^*(\mathbb{M}^4) \to \mathbb{R}$. Note that, in *any* coordinate system (not just inertial ones!) the components of δ are given by[1]

$$\delta_a^b = \mathrm{diag}(1,1,1,1).$$

[1] It is conventional *not* to stagger the indices on the Kronecker delta, unlike other tensor fields.

The exterior derivative

Example 3: Let $f : \mathbb{M}^4 \to \mathbb{R}$ be a smooth function on $\mathbb{M}^4$ (a "scalar field"), and consider the covector field $\mathrm{d}f$ whose components in some coordinates x^a are

$$(\mathrm{d}f)_a = \frac{\partial f}{\partial x^a} = \partial_a f.$$

By construction this is a covector field: its action on the vector X is

$$\mathrm{d}f(X) = (\partial_a f) X^a.$$

Working in another set of inertial coordinates $y^{a'} = \Lambda_a{}^{a'} x^a$, the components of $\mathrm{d}f$ are

$$\begin{aligned}
(\mathrm{d}f)'_{a'} &= (\Lambda^{-1})_{a'}{}^{a}(\mathrm{d}f)_a \\
&= (\Lambda^{-1})_{a'}{}^{a}\frac{\partial f}{\partial x^a} \\
&= (\Lambda^{-1})_{a'}{}^{a}\frac{\partial y^{b'}}{\partial x^a}\frac{\partial f}{\partial y^{b'}} \\
&= (\Lambda^{-1})_{a'}{}^{a}\Lambda_a{}^{b'}\frac{\partial f}{\partial y^{b'}} \\
&= \frac{\partial f}{\partial y^{a'}},
\end{aligned}$$

so in fact, in *any* inertial coordinates, the components of $\mathrm{d}f$ are given by an expression of the form $\frac{\partial f}{\partial x^a}$. So the definition of the covector $\mathrm{d}f$ doesn't require us to first pick out some preferred inertial coordinates.

Derivatives of tensor fields

Consider a general tensor field of rank (n, m), with components in the x^a coordinate system $T^{a_1 \dots a_n}{}_{b_1 \dots b_m}$. We can construct a rank $(n, m+1)$ tensor field with components in the x^a coordinate system

$$\partial_c T^{a_1 \dots a_n}{}_{b_1 \dots b_m}.$$

It is easy to check that this definition is actually independent of the inertial coordinates in which we work, i.e. this expression transforms as an $(n, m+1)$ tensor field.

Integral curves

If X is a vector field then, in inertial coordinates x^a, the *integral curve of the vector field X through the point with coordinates x_0^a* is the curve defined by the ODE

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}x^a(\lambda) = X^a\big|_{x^a(\lambda)}$$
$$x^a(0) = (x_0)^a.$$

Recall that $\frac{\mathrm{d}}{\mathrm{d}\lambda}x^a(\lambda)$ is the tangent vector to the curve $x^a(\lambda)$.

Standard ODE theory ensures that this equation has a unique solution, if the vector field X is smooth and has bounded components.

Worldlines

Suppose a particle moves along a curve with coordinates $x^a(\lambda)$ in some inertial frame. For a massive particle, we can parametrize this curve by the proper time τ instead of the parameter λ.

The *velocity* of the particle v as the tangent vector to its worldline, parametrized by proper time. In inertial coordinates:

$$v^a = \frac{\mathrm{d}x^a(\tau)}{\mathrm{d}\tau}.$$

By the definition of proper time, we have $m_{ab}v^a v^b = -1$. Hence

$$v = \gamma \begin{pmatrix} 1 \\ \boldsymbol{v} \end{pmatrix}$$
$$\gamma = \frac{1}{\sqrt{1 - |\boldsymbol{v}|^2}}.$$

The *four-momentum* of the particle is the covector $p_a := \mu m_{ab} v^b$, where μ is the rest mass of the particle. Thus

$$p = \begin{pmatrix} -\mu\gamma \\ \mu\gamma \boldsymbol{v} \end{pmatrix} = \begin{pmatrix} -\mu - \frac{1}{2}\mu|\boldsymbol{v}|^2 + \mathcal{O}(|\boldsymbol{v}|^4) \\ \mu\boldsymbol{v} + \mathcal{O}(|\boldsymbol{v}|^3) \end{pmatrix}.$$

If $|\boldsymbol{v}| \ll 1$, (velocities much slower than the speed of light), p_0 is (negative) the usual expression for the kinetic energy (plus a constant), while p_i is the momentum. Thus we write

$$p = \begin{pmatrix} -E \\ \boldsymbol{p} \end{pmatrix}.$$

On the other hand, we can choose inertial coordinates so that, at some time τ_0, the velocity vector has components $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. In these coordinates, at this instant of time

$$p = \begin{pmatrix} -\mu \\ 0 \end{pmatrix}.$$

The quantity $-(m^{-1})^{ab}p_a p_b$ is a scalar: its value does not depend on which inertial coordinates we use to evaluate it. So we have Einstein's famous formula

$$E^2 - |\boldsymbol{p}|^2 = \mu^2$$

or, restoring the speed of light using dimensional arguments (recall that we set $c = 1$),

$$E^2 = \mu^2 c^4 + |\boldsymbol{p}|^2 c^2.$$

The *acceleration* of a massive particle is the vector

$$a^a := \frac{\mathrm{d}^2}{\mathrm{d}\tau^2} x^a(\tau).$$

Exercise: show that the acceleration of a particle is orthogonal to its four-velocity in the Lorentzian sense, i.e. $m(v, a) = m_{ab} v^a a^b = 0$.

The energy-momentum tensor

For a continuous distribution of matter, the energy density, energy flux, momentum density and pressure are encoded in a *symmetric* rank $(2,0)$ tensor field $T^{\mu\nu}$.

If v^a is the tangent vector of the worldline of an observer moving through spacetime, then

- The vector $j^a := -T^{ab}v_b$ is the *four-momentum density*.
- The scalar $\rho := -v_a j^a = T^{ab}v_a v_b$ is the *energy density*. For normal matter $\rho \geq 0$ (the *weak energy condition*).

Moreover, if n and N are spacelike vectors defined along the worldline of an observer, normalised so that $m(n,n) = m(N,N) = 1$, and which are orthogonal to the velocity of the observer (i.e. $m(n,v) = m(N,v) = 0$), then

- The scalar $p = T^{ab}n_a n_b$ is the pressure measured by the observer in the n direction.
- The stress in the n direction across a surface orthogonal to N is $S = T^{ab}n_a N_b$.

For example, for a *perfect fluid* moving along the integral curves of a vector field u (normalised so that $m(u,u) = -1$) the energy-momentum tensor is

$$T^{ab} := (\rho + p)u^a u^b + p(m^{-1})^{ab}$$

where ρ and p are the fluid density and pressure in its rest frame. The *equation of state* specifies the scalar field p (the pressure of the fluid) in terms of the scalar field ρ (the density of the fluid), i.e. $p = p(\rho)$.
A second example is given by a *massless scalar field*, whose the energy-momentum tensor is

$$T^{ab} := (\partial^a \phi)(\partial^b \phi) - \frac{1}{2}(m^{-1})^{ab}(m^{-1})^{cd}(\partial_c \phi)(\partial_d \phi)$$

where here ϕ is the scalar field, and $\partial^a \phi = (m^{-1})^{ab}\partial_b \phi$.

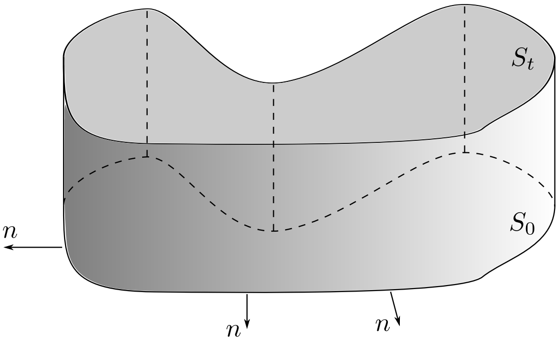
The conservation of energy and momentum is ensured by energy-momentum tensor being divergence free:

$$\partial_a T^{ab} = 0.$$

Consider some region S_0, with smooth boundary ∂S_0, in the "time slice" $t = 0$. Write S_t for the time translation of this surface. Let n be the outwards pointing unit normal to ∂_S. Then, by Stoke's theorem, we have

$$\int_{S_{t'}} T^{a0} \mathrm{d}x^1 \mathrm{d}x^2 \mathrm{d}x^3 = \int_{S_0} T^{a0} \mathrm{d}x^1 \mathrm{d}x^2 \mathrm{d}x^3 + \int_{t=0}^{t'} \int_{\partial S_t} T^{ab} n_b \, \mathrm{d}\Sigma \mathrm{d}t$$

where $\mathrm{d}\Sigma$ is the surface element of ∂S_t. Choosing $a = 0$ we find that the energy in the region S_t (or "in the region S at the time t) is equal to the initial energy in the region S_0, plus the integral of the flux of energy through the bounary ∂_S. Similarly, choosing $a = 1, 2, 3$ we obtain the same conclusion for the momentum in the region S.


S_t
S_0
n
n
n