1 Inertial frames

The relativity of simultaneity

We begin by considering the Newtonian case for how velocities transform. Consider a particle moving in a certain inertial frame described by the velocity components (V^x, V^y, V^z). Next, we consider the components for the velocity of that same particle, but as measured in a reference frame moving at a velocity v along the x-axis of the first frame. Then, the components of the velocity measured in this frame, $(V^{x'}, V^{y'}, V^{z'})$, will be related to the components in the first frame via

$$V^{x'} = \frac{dx'}{dt'} \tag{1}$$
$$= \frac{dx'}{dt} \tag{2}$$
$$= \frac{dx - vdt}{dt} \tag{3}$$
$$= \frac{dx}{dt} - v\frac{dt}{dt} \tag{4}$$
$$= V^x - v. \tag{5}$$

The other components trivially transform according to

$$V^{y'} = V^y \tag{6}$$
$$V^{z'} = V^z \tag{7}$$

since there is no relative motion between the two frames in these directions. Here, we implicitly assume that there is a universal notion of time between these two frames. That is, the label for time is the same between these two frames. We know that this cannot be the case though, if we accept Einstein's postulate that the speed of light is the same in all inertial reference frames. To see why, consider the case of Alice and Bob, standing at either end of a spaceship with a light equidistant between them. Suppose that inside the rocket ship, the light turns on and shines in both directions. Since they are equidistant from the light source, the light will reach them at the same time, and the light bulbs above their heads that are sensitive to these light rays will light up at the same time. Therefore, someone inside the spaceship sees their light bulbs turn on at the same time, and they would therefore report that the light reached Alice and Bob simultaneously.

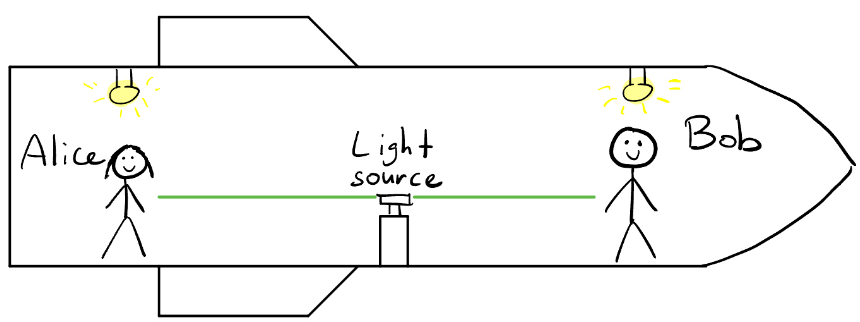

FIGURE 1: Illustration of the spaceship paradox from inside the spaceship.

But consider the perspective from someone standing outside the spaceship, who sees the spaceship going past them at velocity v, with Bob leading Alice. Suppose the light shoots out rays in both directions. Then, since light travels the same speed in both directions, the light will travel at the same speed towards Alice and Bob. But since Alice is moving towards where you saw the light was sent from, and Bob is moving away from where the light was sent from, you see Alice catching up to the light ray. In this way, for an observer outside the spaceship, they will see the light reach Alice first, and so his light bulb will light up before Bob's. This is a simple example of the concept of the relativity of simultaneity. That events that occur at the same time coordinate in one inertial frame do not necessarily occur at the same time coordinate in a different inertial frame.

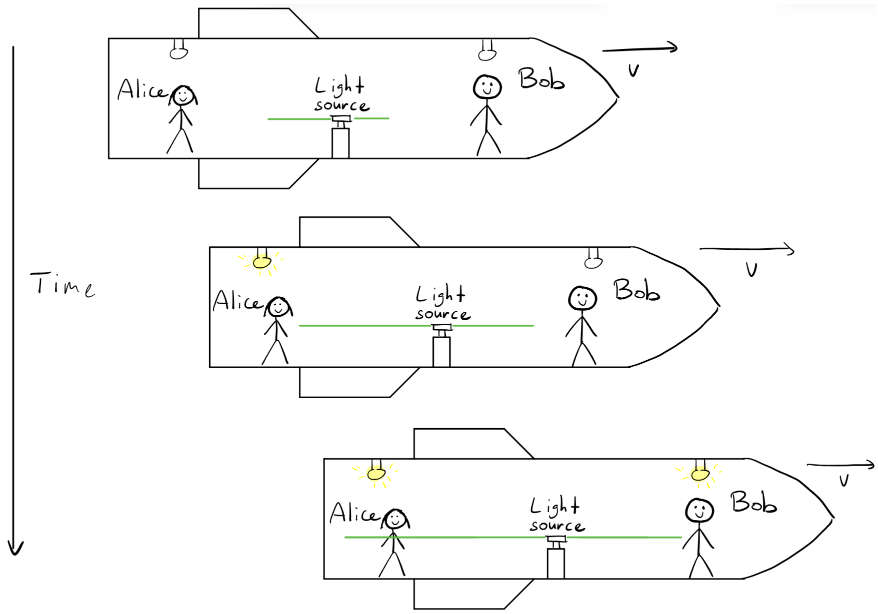

FIGURE 2: Illustration of the spaceship paradox from outside the spaceship.

This thought experiment suggests that instead of there being a universal notion of time that is shared between all inertial frames, instead there are four Cartesian coordinates, (t, x, y, z), describing events in spacetime in each inertial frame.

1.1 Spacetime diagrams

Since the time coordinate is different in each inertial frame, we can thus represent events in spacetime – coordinates in space and time (t, x, y, z) – on a spacetime diagram, with, for simplicity, ct on the vertical axis, and x on the horizontal axis. Particles that move in spacetime trace out a worldline, that describes how they move in space and time. The gradient of a particle's worldline on a spacetime diagram tells you something fundamental about how the particle moves since

$$\frac{d(ct)}{dx} = c\frac{dt}{dx} \tag{8}$$
$$= c\left(\frac{dx}{dt}\right)^{-1} \tag{9}$$
$$= \frac{c}{V^x} \tag{10}$$

and so a particle travelling at velocity $V^x = c$, the speed of light, will have a worldline with gradient 1, and so will move at a 45° angle to the axes. A particle not moving at all will have velocity $V^x = 0$ and so will have a vertical slope, as evident from the division by zero that occurs in Eq. (10).

2 The geometry of special relativity

2.1 The metric of flat spacetime

Special relativity is greatly simplified from the way it is taught in high school and first year university course when an emphasis is placed on the *geometry* that underlies it. In normal Euclidean space, we can describe the geometry of space completely by specifying the line element that connects two points in space. In flat, Euclidean space, the line element is given by Pythagoras's theorem

$$(\Delta s)^2 = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 \tag{11}$$

where Δs is the total distance between two points in Euclidean space. In special relativity, we require an analogous formula that tells us how to relate the distance, Δs, between two events in spacetime. We can derive this heuristically by considering two events in two different inertial frames that occur along the world lines of a light ray. In one of these frames, we will describe events by the coordinates (ct, x). Since these events are separated along the worldline of light, the events must be separated in such a way so that

$$V^x = \frac{\Delta x}{\Delta t} = c. \tag{12}$$

Following the example of the Euclidean interval, with the squares of distances, we can square both sides and rearrange to get

$$\frac{(\Delta x)^2}{(\Delta t)^2} = c^2 \tag{13}$$

$$(\Delta x)^2 = (c\Delta t)^2 \tag{14}$$

$$0 = -(c\Delta t)^2 + (\Delta x)^2. \tag{15}$$

Now, imagine these two points in a different inertial frame described by coordinates (ct', x'). Since the speed of light is an invariant between inertial frames, these two events which lie along the worldline of light should also lie along the worldline of light in this new frame so that

$$\frac{\Delta x'}{\Delta t'} = c \implies 0 = -(c\Delta t')^2 + (\Delta x')^2. \tag{16}$$

Hence, we see that between these two arbitrary frames,

$$-(c\Delta t)^2 + (\Delta x)^2 = -(c\Delta t')^2 + (\Delta x')^2. \tag{17}$$

Although this was derived from a specific case, this relationship actually holds in any inertial reference frame, and between any two events in spacetime. We thus define

$$\boxed{(\Delta s)^2 = -(c\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2} \tag{18}$$

or, for infinitesimal changes in the coordinates,

$$ds^2 = -(c\,dt)^2 + dx^2 + dy^2 + dz^2, \tag{19}$$

to be the line interval, or metric, of flat spacetime. This describes completely the geometry of special relativity. From this, every feature of special relativity can be derived.

One important thing to note here is that the sign of $(\Delta s)^2$ is not fixed, like it is in Euclidean geometry where $(\Delta s)^2 \geq 0$. In special relativity, we have the potential of all three cases

$$\begin{aligned} (\Delta s)^2 > 0 &\quad \text{spacelike separated,} \\ (\Delta s)^2 = 0 &\quad \text{null/lightlike separated,} \\ (\Delta s)^2 < 0 &\quad \text{timelike separated.} \end{aligned}$$

To understand what these different cases denote, we start with the case in which $(\Delta s)^2 > 0$ and restrict our attention to motion in only one spatial dimension

$$(\Delta s)^2 > 0 \tag{20}$$
$$-(c\Delta t)^2 + (\Delta x)^2 > 0 \tag{21}$$
$$(\Delta x)^2 > (c\Delta t)^2 \tag{22}$$
$$\left(\frac{\Delta x}{\Delta t}\right)^2 > c^2 \tag{23}$$
$$\left|\frac{\Delta x}{\Delta t}\right| > c. \tag{24}$$

Thus, if two events in spacetime are spacelike separated, then a particle that travelled between the two events must have been moving faster than light, and so these events cannot be made to fall in the same light cone i.e. a particle or light could not have travelled between the two events – and remember that this separation is the same in *every* inertial reference frame so there is no reference frame in which information, which travels at the speed of light or slower, could have been communicated between the two events. For the case in which $(\Delta s)^2 = 0$, we have

$$\left|\frac{\Delta x}{\Delta t}\right| = c \tag{25}$$

and so particles that move at the speed of light travel along lightlike worldlines. For the final case

$$\left|\frac{\Delta x}{\Delta t}\right| < c \tag{26}$$

which implies that a particle could travel between these two events at slower than light speed.

2.2 Lorentz transformations

The most important feature of this line element for PHYS2100 is a particular class of coordinate transformations called Lorentz transformations. In three-dimensional Euclidean space, you can choose to rotate your coordinate system by defining the x'- and y'-axes of your new coordinate system in terms of the coordinates of the old coordinate system. For example, we could simply define a new coordinate system to be a simple translation of the old coordinate system

$$x' = x + \alpha \tag{27}$$
$$y' = y + \beta \tag{28}$$
$$z' = z + \varepsilon \tag{29}$$

where α, β and ε are constants you are shifting your axes by. All that matters for this to define a new coordinate system however is that the Euclidean line element in Eq. (11) is unchanged. We see that it is indeed unaltered since

$$dx' = dx + d\alpha \tag{30}$$
$$dy' = dy + d\beta \tag{31}$$
$$dz' = dz + d\varepsilon \tag{32}$$

and since α, β and ε are constants, $d\alpha = d\beta = d\varepsilon = 0$ and hence

$$dx' = dx \tag{33}$$
$$dy' = dy \tag{34}$$
$$dz' = dz \tag{35}$$

which gives

$$(dx')^2 + (dy')^2 + (dz')^2 = dx^2 + dy^2 + dz^2. \tag{36}$$

Hence, under this coordinate translation, the way that distances are measured is unchanged. A more complicated coordinate transformation that preserves distances in Euclidean space is the rotation about the z-axis by an angle θ

$$x' = \cos(\theta)x + \sin(\theta)y \tag{37}$$
$$y' = -\sin(\theta)x + \cos(\theta)y \tag{38}$$
$$z' = z. \tag{39}$$

You can confirm by simple computation that this coordinate transformation satisfies

$$(dx')^2 + (dy')^2 + (dz')^2 = dx^2 + dy^2 + dz^2 \tag{40}$$

and so distances remain invariant under this coordinate transformation. But this coordinate transformation definitely feels different to the simple case of just translating the axes. In this case, you are defining your new x'-coordinate, for example, in terms of the x- and y-coordinates of your old coordinate system. But remember that coordinates are arbitrary labels you place on, in this case, Euclidean space. So long as you are describing the same physical location in this case, everything works out fine. How you choose to describe it is entirely up to you.

Entirely analogous to coordinate transformations in Euclidean space, we can perform coordinate transformations between *inertial frames* themselves in special relativity. Since the line element of special relativity maintains the $dx^2 + dy^2 + dz^2$, we can also perform rotations of x, y and z about each other in a way that preserves distances in special relativity. But we also have the extra time coordinate in special relativity, so is there anything interesting we can do to perform a sort of rotation in the (ct, x) plane? And if so, what would this correspond to physically? It turns out the coordinate transformation given by

$$ct' = \cosh(\theta)ct - \sinh(\theta)x \tag{41}$$
$$x' = -\sinh(\theta)ct + \cosh(\theta)x \tag{42}$$

in which space and *time* are being rotated into each other, preserves the line element of special relativity. So it works mathematically, but to find what it actually stands for, we have to define $v = c\tanh(\theta)$. Doing so, these coordinate transformations can be written in the more familiar form

$$\boxed{\begin{aligned} ct' &= \gamma\Big(ct - \frac{vx}{c}\Big) && (43)\\ x' &= \gamma(x - vt) && (44)\end{aligned}}$$

where γ is the Lorentz factor, given by

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{45}$$

Note that these can be also used to give us

$$c\Delta t' = \gamma\left(c\Delta t - \frac{v\Delta x}{c}\right) \tag{46}$$

$$\Delta x' = \gamma(\Delta x - v\Delta t) \tag{47}$$

or, for infinitesimal changes in the coordinates,

$$c\,dt' = \gamma\left(c\,dt - \frac{v\,dx}{c}\right) \tag{48}$$

$$dx' = \gamma(dx - vdt). \tag{49}$$

Thus, we see that a boost in velocity v between inertial reference frames corresponds to rotating space and time into each other in the context of the geometry of flat spacetime. These are called the Lorentz transformations and relate the coordinates of an event in spacetime in some frame with coordinates (ct, x) to the coordinates of the same event in the inertial frame with coordinates (ct', x'). We can see that these formulae make explicit the relativity of simultaneity if we try and calculate the time interval measured in frame (ct', x') between two events that are simultaneous in frame (ct, x). If they are simultaneous in this frame, then $\Delta t = 0$ and hence

$$c\Delta t' = \gamma\left(c\Delta t - \frac{v\Delta x}{c}\right)\Big|_{\Delta t=0} \tag{50}$$

$$= -\frac{\gamma v\Delta x}{c}. \tag{51}$$

Hence, despite the events being simultaneous in the unprimed frame, they are not simultaneous in the primed frame.

2.3 Relativistic velocity addition

Using the formulae for the Lorentz transformations, we can also determine how velocities transform in different inertial frames. Suppose we want to know how the velocity of a particle that's travelling with velocity V^x in frame S transforms into frame S' that's moving with velocity v along the x-axis of frame S. We wish to calculate dx'/dt' which, using the Lorentz formulae in Eqs. (48) and (49), can be written as

$$\frac{dx'}{dt'} = \frac{\gamma(dx - vdt)}{\gamma\left(dt - \frac{v\,dx}{c^2}\right)} \tag{52}$$

$$= \frac{\frac{dx}{dt} - v\frac{dt}{dt}}{\frac{dt}{dt} - \frac{v}{c^2}\frac{dx}{dt}} \tag{53}$$

$$\boxed{\frac{dx'}{dt'} = \frac{V^x - v}{1 - vV^x/c^2}.} \tag{54}$$

In the limit that $vV^x \ll c^2$, so when either the frames are moving slowly with respect to each other or the particle has a low velocity in frame S, the factor vV^x/c^2 is negligible and so we recover

$$V^{x'} = V^x - v \tag{55}$$

which is what you'd expect from Sir Isaac Newton.

3 Time dilation and length contraction

3.1 Proper time

The geometric distinction between timelike and spacelike curves is mirrored in the devices used to measure them. Along timelike curves, clocks are used to measure distances while along spacelike curves, rulers are used to measure distances. To measure the distance along a particle's timelike worldline, it is useful to introduce the concept of proper time

$$\boxed{d\tau^2 = -\frac{ds^2}{c^2}} \tag{56}$$

which we note has units of time. To understand what the proper time physically represents, we note that in our formula for the line element of special relativity, if we imagine the interval between two events that occur at the same place along a particle's worldline, we will have $dx^2 = dy^2 = dz^2 = 0$ in which case we find

$$ds^2 = -(c\,dt)^2 \tag{57}$$

$$dt^2 = -\frac{ds^2}{c^2} \tag{58}$$

$$= d\tau^2 \tag{59}$$

and so we see that the proper time along a timelike curve is the distance that is measured by a clock travelling along that worldline.

3.2 Time dilation

Given our expression for the proper time along a worldline, if we wish to actually calculate it along some arbitrary worldline between two events A and B, we simply integrate the differential form of τ along that worldline

$$\tau_{AB} = \int d\tau \tag{60}$$

$$= \int \left(-\frac{ds^2}{c^2}\right)^{1/2} \tag{61}$$

$$= \int \left\{-\frac{1}{c^2}\left[-(c\,dt)^2 + dx^2 + dy^2 + dz^2\right]\right\}^{1/2} \tag{62}$$

$$= \int \left[dt^2 - \frac{1}{c^2}\left(dx^2 + dy^2 + dz^2\right)\right]^{1/2} \tag{63}$$

$$= \int dt\left\{1 - \frac{1}{c^2}\left[\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2\right]\right\}^{1/2} \tag{64}$$

$$= \int dt\left(1 - \frac{V^2}{c^2}\right)^{1/2} \tag{65}$$

where we have used $V^2 = (V^x)^2 + (V^y)^2 + (V^z)^2$. For short enough intervals where the velocity is constant, we have the usual formula for time dilation

$$\boxed{d\tau = dt\left(1 - \frac{V^2}{c^2}\right)^{1/2}.} \tag{66}$$

Since the factor multiplying dt is always less than 1 (unless $V^2 = 0$), we find that the proper time is always smaller than the time interval measured by an observer in any other inertial frame.

3.3 Length contraction

Now we consider the phenomenon of length contraction. Consider a measurement of the length of a rod in the frame in which it is stationary. This measurement of the length of the rod corresponds to the interval between two simultaneous events occurring at each end of the rod. Let us denote this interval by L_*. To determine what an observer in a frame in which the rod is moving with velocity V will measure its length to be, we note that a measurement of the length of the rod in this frame corresponds to simultaneous events in this reference frame at either end of the rod. But the notion of simultaneity is relative and so the interval between events at the end of the rod will be different in this frame. The length L of the rod in this frame is given by

$$\boxed{L = L_* \sqrt{1 - \frac{V^2}{c^2}}.} \tag{67}$$

3.4 Natural units convention

Up until now, we've used SI units in which the speed of light takes on some large number that's not really nice to work with. If we instead use natural units, however, in which we set $c = 1$ and instead give velocities as a fraction of the speed of light. In this system of units then, the metric of flat spacetime becomes

$$\boxed{ds^2 = -dt^2 + dx^2 + dy^2 + dz^2} \tag{68}$$

and the Lorentz transformation formulae become

$$\boxed{\begin{aligned} t' &= \gamma(t - vx) \quad (69) \\ x' &= \gamma(x - vt) \quad (70) \end{aligned}}$$

which reveals the symmetry between the time and position coordinates.

4 Four-vectors and mechanics in special relativity

4.1 Four-vectors

Now, we can turn to how we represent worldlines of particles using vectors in special relativity. In non-relativistic classical physics we can represent the path of a particle through space by a vector with three components. In the context of special relativity, we instead represent particles by vectors with four components, the extra component being for the time component of the particle. Thus, we represent some arbitrary four-vector as a linear combination of the unit basis vectors, $\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2$ and $\mathbf{e}_3$ or, equivalently, $\mathbf{e}_t, \mathbf{e}_x, \mathbf{e}_y$ and $\mathbf{e}_z$

$$\mathbf{a} = a^t \mathbf{e}_t + a^x \mathbf{e}_x + a^y \mathbf{e}_y + a^z \mathbf{e}_z. \tag{71}$$

We call the list of numbers (a^0, a^1, a^2, a^3) or (a^t, a^x, a^y, a^z) are called the (contravariant) components of the four-vector. More compactly, we can write the arbitrary four-vector $\mathbf{a}$ as

$$\mathbf{a} = \sum_{\alpha=0}^{3} a^\alpha \mathbf{e}_\alpha \tag{72}$$

which can be made even more compact by introducing Einstein's summation convention where repeated indices are implicitly summed over

$$\mathbf{a} = a^\alpha \mathbf{e}_\alpha. \tag{73}$$

4.2 Scalar product and the metric tensor

We can take the inner product between two four-vectors using the formula

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} &= (a^\alpha \mathbf{e}_\alpha) \cdot (b^\beta \mathbf{e}_\beta) && (74) \\ &= (\mathbf{e}_\alpha \cdot \mathbf{e}_\beta) a^\alpha b^\beta. && (75) \end{aligned}$$

We denote the mathematical object $\mathbf{e}_\alpha \cdot \mathbf{e}_\beta$ by a special symbol which we call the Minkowski metric tensor of flat spacetime

$$\boxed{\eta_{\alpha\beta} = \mathbf{e}_\alpha \cdot \mathbf{e}_\beta.} \tag{76}$$

Therefore, the dot product between two four-vectors can be written as

$$\mathbf{a} \cdot \mathbf{b} = \eta_{\alpha\beta} a^\alpha b^\beta. \tag{77}$$

We can thus choose to represent $\eta_{\alpha\beta}$ in matrix form as

$$\boxed{\eta_{\alpha\beta} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.} \tag{78}$$

From this, we can express the line element of special relativity in the more compact form

$$(\Delta s)^2 = \Delta \mathbf{x} \cdot \Delta \mathbf{x}. \tag{79}$$

4.3 The metric tensor to raise and lower indices

In the previous section, I exclusively talked about 4-vectors as having superscript components. These are the so-called **contravariant components** of the positional 4-vector[1].There are also the **covariant components** of the positional 4-vector, denoted by x_μ, and related to the contravariant components by

$$x_\mu = \sum_{\nu=0}^{3} \eta_{\mu\nu} x^\nu$$

where $\eta_{\mu\nu}$ are the *covariant* components of the metric tensor. We see also that in this expression that we have a repeated index of ν both 'upstairs' and 'downstairs.' Indeed, whenever we wish to raise any index (in this case, raise one of the indices of a rank-two covariant tensor), we can apply a contravariant metric tensor in the following way

$$A^\mu_\nu = \eta^{\mu\rho} A_{\rho\nu}.$$

[1]By convention, spacetime indices are usually Greek letters, however, there are some exceptions. For example, ϕ, ψ, and η are used in classical and quantum field theories as representations of fields or wave functions. In general, we want to chose indices that avoid any confusion between fields and different sorts of indices, so don't reuse the same index in one calculation if they are not related.

where we note that we have omitted the summation sign since the repeated upstairs and downstairs index imply summation. Notice also in this summation that the index ν remains unchanged on both sides of this equation. We call this a **free index** since it is the same on both sides of the equation and isn't affected by our summation. We would call ρ a **summation index** since it is summed over and disappears on the left-hand side of the expression. Therefore, we could relabel our summation indices to be whatever we want, since they end up disappearing and are 'replaced' by the remaining (free) index on the metric tensor. We can also lower indices by applying a covariant metric tensor

$$A_{\mu\nu} = \eta_{\mu\alpha} A^{\alpha}_{\nu}.$$

Although this procedure works with higher rank tensors with more than one or two indices, when we are dealing with rank two tensors, with two indices like we have been doing here, we see that this resembles matrix multiplication. Indeed, we can think of this as being matrix multiplication when working with these rank one and rank two tensors. For example, the metric tensor can be written as

$$\eta_{\mu\nu} = \eta^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

and we can think of contravariant rank one vectors as being column vectors

$$x^{\mu} = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}.$$

We can think of covariant vectors as being row vectors[2]

$$x_{\mu} = \begin{pmatrix} x_0 & x_1 & x_2 & x_3 \end{pmatrix}.$$

The reason that we make this connection is that we wish to understand what it means to take a dot product between two 4-vectors a^{μ} and b^{ν}. We can't simply sum over the contravariant components $a^{\mu} b^{\mu}$ since summations only work with repeated upstairs and downstairs indices, so we need to lower the index of one of these using a metric tensor to give

$$\mathbf{a} \cdot \mathbf{b} = a_{\mu} b^{\mu} = \eta_{\mu\alpha} a^{\alpha} b^{\mu}$$

so if we wish to do everything with matrices, the left-hand side of this equation, to give a scalar, requires that the components a_{μ} are the entries of a row vector, with the last three entries having a negative sign from the metric tensor

$$a_{\mu} = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 \end{pmatrix} = \begin{pmatrix} a^0 & -a^1 & -a^2 & -a^3 \end{pmatrix}.$$

Extending this to the positional 4-vector, we have

$$\begin{aligned} x_{\mu} &= \begin{pmatrix} x_0 & x_1 & x_2 & x_3 \end{pmatrix} \\ &= \begin{pmatrix} x^0 & -x^1 & -x^2 & -x^3 \end{pmatrix} \\ &= \begin{pmatrix} ct & -x & -y & -z \end{pmatrix}. \end{aligned}$$

[2] One mnemonic to remember the difference is 'co-low-row', meaning, covariant vectors have lowered indices and are row vectors. It does not quite work the other way around for 'cont-high-col,' but hey, nobody's perfect.

This means that our dot product between x^μ and itself is equal to

$$x_\mu x^\mu = \begin{pmatrix} ct & -x & -y & -z \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = (ct)^2 - x^2 - y^2 - z^2.$$

4.4 Lorentz transformations of 4-vectors

Having seen that we can represent special collections of numbers, such as the coordinates of an object in spacetime x^μ, as 4-vectors, we might wonder how 4-vectors transform under Lorentz transformations. Remember that x^μ contains the time and spatial coordinates of some object in Minkowski space, and we know that under a velocity boost in the x-direction of velocity v, for example, the coordinates of such an object transform as

$$ct' = \gamma\Big(ct - \frac{vx}{c}\Big) \tag{80}$$

$$x' = \gamma(x - vt) \tag{81}$$

$$y' = y \tag{82}$$

$$z' = z. \tag{83}$$

We note that on the left-hand side of these four equations are the coordinates of the 4-position in the new reference frame after the boost, and they are related on the right-hand side to the coordinates of the particle in the unboosted reference frame linearly. This means that we can write the above four equations simply as a matrix equation,

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma v/c & 0 & 0 \\ -\gamma v/c & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}. \tag{84}$$

The matrix on the right-hand side of this equation (which is a function of the velocity of the boost in the x-direction) encodes the information about how the contravariant components of the 4-vector $\mathbf{x}$ should transform under a velocity boost in the x-direction. We can therefore refer to this matrix as the Lorentz transformation itself, and we usually denote it by $\mathbf{\Lambda}$, so that

$$\mathbf{x}' = \mathbf{\Lambda}\mathbf{x} \;\longleftrightarrow\; x^{\mu'} = \Lambda^{\mu'}_{\nu} x^\nu. \tag{85}$$

Strictly speaking, if we were to think of the Lorentz transformation $\mathbf{\Lambda}$ as being represented by a matrix with components $\Lambda^{\mu'}_{\nu}$ with respect to the ordered basis $\{\mathbf{e}_\alpha\}$ then we should really make it clear which of the indices on $\Lambda^{\mu'}_{\nu}$ represent the row and which represents the column, since in general Lorentz transformations will not be symmetric. However, for the purposes of Lorentz boosts in special relativity, and particularly in PHYS2100, you will never have to worry about this distinction. If you look to nonsymmetric Lorentz transformations, such as those representing a spatial rotation about a given axis, then you will need to specify which index represents the row and which the column. In Eq. (85) this would be done by shifting the lower (covariant) index on Λ to the right, so that we write the Lorentz transformation as $x^{\mu'} = \Lambda^{\mu'}{}_{\nu} x^\nu$. This implies that we read $\Lambda^{\mu'}{}_{\nu}$ as being the matrix entry corresponding to the Lorentz transformation $\mathbf{\Lambda}$ in the μ'th row and the νth column.

4.5 Kinematics in special relativity

When describing the worldline of a particle in special relativity, we will normally do so using the concept of four-vectors whose components give the four coordinates of the particle as a function of some parameter σ. Many parameters are possible but a natural one to choose is the proper time τ, in which case we have $x_\alpha(\tau)$. The distance τ could be measured by a clock carried along the worldline since it's just the proper time measured along that worldline. The four-velocity is the four-vector $\mathbf{u}$ whose components u_α are the derivatives of the position along the worldline with respect to the proper time

$$\boxed{u_\alpha = \frac{dx_\alpha}{d\tau}.} \tag{86}$$

The components become

$$u_0 = \frac{dt}{d\tau} = \left(1 - \vec{V}^2\right)^{-1/2} = \gamma. \tag{87}$$

Also, for the other components, we can use the chain rule

$$\frac{dx}{d\tau} = \frac{dx}{dt}\frac{dt}{d\tau} = \gamma V^x. \tag{88}$$

Therefore, we find

$$\boxed{u_\alpha = (\gamma, \gamma\vec{V})} \tag{89}$$

where $\vec{V}$ is the particle's three-velocity. We therefore find

$$\begin{aligned} \mathbf{u}\cdot\mathbf{u} &= -\gamma^2 + \gamma^2\vec{V}^2 && (90)\\ &= -\frac{1}{1-\vec{V}^2} + \frac{\vec{V}^2}{1-\vec{V}^2} && (91)\\ &= -1. && (92) \end{aligned}$$

This so-called normalisation of the four-velocity also follows directly from the definition of the scalar product

$$\begin{aligned} \mathbf{u}\cdot\mathbf{u} &= \eta_{\alpha\beta}\frac{dx_\alpha}{d\tau}\frac{dx_\beta}{d\tau} && (93)\\ &= \frac{ds^2}{d\tau^2} && (94)\\ &= -1 && (95) \end{aligned}$$

using the definition $d\tau^2 = -ds^2$ (with $c = 1$).

4.6 Special relativistic mechanics

Newton's laws of motion in special relativity take on an analogous form in special relativity to how they do in non-relativistic, Newtonian physics. Firstly, the equation of motion for a free particle experiencing no external forces is given by

$$\frac{d\mathbf{u}}{d\tau} = 0 \tag{96}$$

which is exactly what you'd expect from Newton's first law of motion. Similarly, when a particle is acted on by external forces, the equations of motion are

$$\boxed{\begin{aligned}\mathbf{f} &= m\mathbf{a} && (97)\\ &= m\frac{d\mathbf{u}}{d\tau} && (98)\\ &= m\gamma\frac{d\mathbf{u}}{dt} && (99)\end{aligned}}$$

where m is the mass of the particle and $\mathbf{f}$ is called the four-force. This is the correct law of motion for special relativistic mechanics. Although this represents four equations, you really only have three independent equations since the normalisation of the four-velocity implies

$$m\frac{d(\mathbf{u}\cdot\mathbf{u})}{d\tau} = 0 \tag{100}$$

which implies $\mathbf{u}\cdot\mathbf{a} = 0$ or

$$\mathbf{u}\cdot\mathbf{f} = 0 \tag{101}$$

so we really only have three independent equations to solve. These equations of motion imply the notion of four-momentum

$$\boxed{\mathbf{p} = m\mathbf{u}} \tag{102}$$

which we see has the normalisation condition

$$\boxed{\begin{aligned}\mathbf{p}\cdot\mathbf{p} &= m\mathbf{u}\cdot m\mathbf{u} && (103)\\ &= m^2(\mathbf{u}\cdot\mathbf{u}) && (104)\\ &= -m^2. && (105)\end{aligned}}$$

We thus have

$$p_t = \frac{m}{\sqrt{1-\vec{V}^2}} \quad \text{and} \quad \vec{p} = \frac{m\vec{V}}{\sqrt{1-\vec{V}^2}}. \tag{106}$$

For small speeds in which $V \ll 1$, we can Taylor expand these expressions to get

$$p_t = m + \frac{1}{2}m\vec{V}^2 + \ldots \quad \text{and} \quad \vec{p} = \vec{V} + \ldots \tag{107}$$

Thus, at small velocities, $\vec{p}$ reduces to the kinetic energy plus the rest mass. For this reason $\mathbf{p}$ is also called the energy-momentum four-vector, and its components in an inertial frame

$$\boxed{p^\alpha = (E, \vec{p}) = (m\gamma, m\gamma\vec{V}).} \tag{108}$$

Taking the dot product between $\mathbf{p}$ and itself gives

$$\begin{aligned}\mathbf{p}\cdot\mathbf{p} = -m^2 &= -E^2 + \vec{p}^2 && (109)\\ E^2 &= \vec{p}^2 + m^2 && (110)\\ \Aboxed{E &= \sqrt{\vec{p}^2 + m^2}.} && (111)\end{aligned}$$

The connection between the relativistic equations of motion can be made more explicit by defining the three-force

$$\vec{F} = \frac{d\vec{p}}{dt} \tag{112}$$

which is the same expression as Newton's laws but with the relativistic version of the three-momentum. The four-force can be written in terms of the three-force as

$$\boxed{f^{\alpha} = (\gamma \vec{F} \cdot \vec{V}, \gamma \vec{F})} \tag{113}$$

The time component of the four-force is thus given by

$$\frac{dE}{dt} = \vec{F} \cdot \vec{V} \tag{114}$$

which is the familiar formula from Newtonian mechanics.

4.7 Light rays

When parametrising the worldline of a light ray, we can no longer use the proper time since the spacetime interval between any two points on its worldline will be zero. To parametrise the worldline of a photon, we use the fact Einstein's energy relation for photons

$$E = \hbar\omega. \tag{115}$$

Since we have

$$\begin{aligned} \vec{p} &= \frac{m\vec{V}}{\sqrt{1-\vec{V}^2}} && (116)\\ &= p_0 \vec{V} && (117)\\ &= E\vec{V} && (118) \end{aligned}$$

this implies that $|\vec{p}| = E$ for a photon (with $V = 1$) so the three-momentum can be written as

$$\vec{p} = \hbar\vec{k} \tag{119}$$

where $\vec{k}$ points in the direction of propagation, has magnitude $|\vec{k}| = \omega$, and is called the wave three-vector. In any inertial frame the components of the four-momentum of a photon $\mathbf{p}$ can therefore be written as

$$\boxed{p^{\alpha} = (E, \vec{p}) = (\hbar\omega, \hbar\vec{k}) = \hbar k^{\alpha}.} \tag{120}$$

The four-vector $\mathbf{k}$ is called the wave four-vector. Evidently,

$$\mathbf{p} \cdot \mathbf{p} = \mathbf{k} \cdot \mathbf{k} = 0. \tag{121}$$

The first of these equations implies that $m = 0$ for photons. The equation of motion for a photon can be written in terms of $\mathbf{p}$ or $\mathbf{k}$ as

$$\frac{d\mathbf{p}}{d\lambda} = 0 \quad \text{or} \quad \frac{d\mathbf{k}}{d\lambda} = 0 \tag{122}$$

where λ is an affine parameter. This is the same equation of motion as for a free particle with mass acted on by no forces.