Unit 7: Electromagnetism in Special Relativity

In this unit we will see how to formulate the laws of electromagnetism in the language of special relativity. We will also see how to generalize Larmor's formula to get the corresponding result when the charge is moving at relativistic speeds. It is assumed that this is *not* your first exposure to the theory of special relativity.

Unit 7-1: Review of Special Relativity

The theory of special relativity is concerned with how one describes the laws of physics in differing *inertial frames of reference.* To describe what we mean by inertial frames of reference, consider one frame of reference that is "at rest." The set of inertial frames of reference are then those reference frames that move with some constant velocity $\mathbf{v}$ with respect to the frame at rest.

Of course it is a question, when we say "at rest" what exactly do we mean? – at rest with respect to what? Sitting in a chair in your room, you may consider yourself to be at rest. But of course your chair is in your room, which is on the earth, which is spinning on its axis, and rotating around the sun, which itself is rotating around in the galaxy, and the galaxy is rotating around in the universe. So really you are hardly at rest! One could say that "at rest" should mean: at rest with respect to the center of mass of the universe. And if the universe is self contained, then the center of mass of the universe is presumably not being accelerated. We will not worry about such subtleties, and for the sake of our discussion, you are at rest if you are at rest with respect to the earth and all accelerations you are interested in are large compared to the acceleration of the earth with respect to the center of mass of the universe! In this case, a train moving with constant $\mathbf{v}$ would be an inertial frame with respect to our rest frame. This will be good for thinking about experiments conducted on our everyday time and length scales, but of course not good if we are talking about astrological times and length scales.

The theory of special relativity is then based on two assumptions:

1) The speed of light in a vacuum is the same constant in all inertial frames of reference.

2) The laws of physics must look the same in all inertial frames of reference – there is no experiment that one can do that will determine if one is "at rest" or if one is moving with a constant velocity. There is no meaningful notion of the "absolute velocity" of any inertial frame. For example, if you are in a train car moving with constant velocity, or you are in the elevator moving between floors at constant velocity, you can bounce balls, and shine light, and do any other physical experiment, and all the results you find would come out exactly as if you were at rest – there is no way you can detect and measure the velocity of your inertial frame of reference. You will only know that you are moving, and not at rest, when your frame accelerates or decelerates; i.e. if the train speeds up or slows down, or if the elevator accelerates as it starts to move, or decelerates as it comes to a stop.

Consider the inertial frame of reference $\mathcal{K}$ in which the coordinates to measure events are (x, y, z, t).

Consider the inertial frame $\mathcal{K}'$ which travels with constant velocity $\mathbf{v} = v\hat{\mathbf{x}}$ with respect to $\mathcal{K}$, and which uses coordinates (x', y', z', t'). The coordinates are chosen so that the spatial origins of $\mathcal{K}$ and $\mathcal{K}'$ coincide at time $t = t' = 0$.

An event happens and an observer in $\mathcal{K}$ measures it to occur at (x, y, z, t). An observer in $\mathcal{K}'$ measures the same event to occur at coordinates (x', y', z', t'). What is the transformation that relates the coordinates measured in $\mathcal{K}'$ to the coordinates measured in $\mathcal{K}$?

Since $\mathcal{K}'$ moves with velocity $v\hat{\mathbf{x}}$ as seen by $\mathcal{K}$, we can conclude that for the coordinates transverse to the direction of motion,

$$y' = y \qquad \text{and} \qquad z' = z \tag{7.1.1}$$

To see this, let us imagine the following. The observer at rest in $\mathcal{K}$ draws a red line parallel to his x-axis at height $y = H$. The observer at rest in $\mathcal{K}'$ draws green line parallel to her x-axis at height $y' = H$. Now suppose it was the case that lengths appear to shrink in directions transverse to the direction of relative motion. Then the observer at rest in $\mathcal{K}$, who sees $\mathcal{K}'$ moving, would expect that the green line in $\mathcal{K}'$ should lie below his own red line in $\mathcal{K}$. However the observer at rest in $\mathcal{K}'$, who sees $\mathcal{K}$ moving, would expect that the red line in $\mathcal{K}$ should lie below her own green line in $\mathcal{K}'$ (because otherwise, they would be able to tell which of the two is the one who is moving, and which is the one at rest; that would violate (2)). Clearly both things cannot happen! The green line is either above or below the red line; it can't be one way in frame $\mathcal{K}$ and the other way in frame $\mathcal{K}'$. The only conclusion consistent with assumption (2) is that the green line and the red line are overlapping, i.e. the coordinates transverse to the direction of relative motion are the same, $y = y'$.

Now consider what happens when a pulse of light goes off at $t = t' = 0$, when the origins of $\mathcal{K}$ and $\mathcal{K}'$ coincide. By the above assumptions (1) and (2), the outgoing wavefronts must look spherical in both frames of reference (since otherwise one could distinguish which frame was at rest, and which was the frame moving with a constant velocity). The equation of the wavefront in each frame is therefore,

$$\text{in } \mathcal{K}\text{:} \quad r^2 - c^2 t = 0 \qquad \text{and} \qquad \text{in } \mathcal{K}'\text{:} \quad {r'}^2 - c^2 {t'}^2 = 0 \tag{7.1.2}$$

so we can write,

$$r^2 - c^2 t \;=\; {r'}^2 - c^2 {t'}^2 \qquad \Rightarrow \qquad x^2 + y^2 + z^2 - c^2 t^2 \;=\; {x'}^2 + {y'}^2 + {z'}^2 - c^2 {t'}^2 \tag{7.1.3}$$

and since $y = y'$ and $z = z'$, this becomes,

$$x^2 - c^2 t^2 \;=\; {x'}^2 - c^2 {t'}^2 \qquad \Rightarrow \qquad (x - ct)(x + ct) = (x' - ct')(x' + ct') \tag{7.1.4}$$

and so

$$\frac{(ct + x)}{(ct' + x')} \frac{(ct - x)}{(ct' - x')} \;=\; 1 \tag{7.1.5}$$

We expect that the transformation relating x' and t' to x and t must be *linear*. If it was not linear, than an object moving with constant speed in frame $\mathcal{K}$, $x(t) = ut$, would look like it was accelerated in frame $\mathcal{K}'$.

Since the transition must be linear, we have,

$$ct' + x' = (ct + x) f \qquad \text{for some constant } f. \tag{7.1.6}$$

Then to satisfy Eq. (7.1.5), we must have

$$ct' - x' = (ct - x) f^{-1} \tag{7.1.7}$$

We will write the constant f as $f = e^{-y}$, where y is called the *rapidity* (not to be confused with the coordinate y!).

Solving the pair of linear Eqs. (7.1.6) and (7.1.7) to write ct' and x' in terms of ct and x, we get,

$$ct' = \quad ct \left(\frac{e^y + e^{-y}}{2} \right) - x \left(\frac{e^y - e^{-y}}{2} \right) \tag{7.1.8}$$

$$x' = -ct \left(\frac{e^y - e^{-y}}{2} \right) + x \left(\frac{e^y + e^{-y}}{2} \right) \tag{7.1.9}$$

or

$$ct' = \quad ct \cosh y - x \sinh y \tag{7.1.10}$$

$$x' = -ct \sinh y + x \cosh y \tag{7.1.11}$$

The Lorentz Transformation

We now want to see what is the physical meaning of the rapidity y.

Since frame $\mathcal{K}'$ moves with speed v to the right, as seen by the frame $\mathcal{K}$, then the frame $\mathcal{K}$ moves with speed v to the left, as seen by the frame $\mathcal{K}'$. Therefore the origin of frame $\mathcal{K}$ (which is at $x = 0$) has a trajectory as seen in frame $\mathcal{K}'$,

$$x' = -vt' \qquad \Rightarrow \qquad \frac{x'}{t'} = -v \tag{7.1.12}$$

Appling the transformation of Eq. (7.1.11) with $x = 0$, we find that the origin of $\mathcal{K}$ as seen in frame $\mathcal{K}'$ satisfies,

$$\frac{x'}{ct'} = \frac{-ct\sinh y}{ct\sinh y} = -\tanh y \tag{7.1.13}$$

Comparing with Eq. (7.1.12) we conclude,

$$\frac{v}{c} = \tanh y \tag{7.1.14}$$

Using the identities for the hyperbolic trig functions, we have,

$$\cosh y = \frac{1}{\sqrt{1-\left(\frac{v}{c}\right)^2}} \equiv \gamma \qquad \text{and} \qquad \sinh y = \frac{v}{c}\gamma \tag{7.1.15}$$

We thus get the *Lorentz transformation* from the coordinates in $\mathcal{K}$ to the coordinates in $\mathcal{K}'$,

$$\begin{aligned} ct' &= \gamma ct - \gamma\left(\frac{v}{c}\right)x \\ x' &= -\gamma\left(\frac{v}{c}\right)ct + \gamma x \\ y' &= y \\ z' &= z \end{aligned} \tag{7.1.16}$$

The inverse transformation from $\mathcal{K}'$ to $\mathcal{K}$ is obtained by taking $v \leftrightarrow -v$ in the above, and interchanging primed with unprimed variables,

$$\begin{aligned} ct &= \gamma ct' + \gamma\left(\frac{v}{c}\right)x' \\ x &= \gamma\left(\frac{v}{c}\right)ct' + \gamma x' \\ y &= y' \\ z &= z' \end{aligned} \tag{7.1.17}$$

From the above Lorentz transformation we can derive the well known behaviors of special relativity.

<u>FitzGerald contraction</u>

Consider a stick at rest in frame $\mathcal{K}'$. The length of the stick in its rest frame $\mathcal{K}'$ is L_0. What length does the observer in $\mathcal{K}$ measure the stick to have, when the stick is oriented parallel to the relative motion of $\mathcal{K}'$ with respect to $\mathcal{K}$?

In frame $\mathcal{K}'$ the stick is at rest, and its left end is at $x'_1 = 0$ while its right end is at $x'_2 = L_0$. At $t = t' = 0$ the observer in $\mathcal{K}$, watching the stick move to his right with velocity $v\hat{\mathbf{x}}$, marks where the left and right ends of the stick are. He measures the left end to be at $x_1 = 0$ (since the origins of $\mathcal{K}$ and $\mathcal{K}'$ coordinates coincide at $t = t' = 0$). And he measures the right end of the stick to be at $x_2 = L$. We call "event 1" the measurement by $\mathcal{K}$ of the position of the left end of the stick, and "event 2" the measurement by $\mathcal{K}$ of the position of the right end of the stick. Since $\mathcal{K}$ measures both ends *simultaneously* the times of event 1 and event 2 in $\mathcal{K}$ are $t_1 = t_2 = 0$. From the Lorentz transformation we know,

$$x'_2 = -\gamma\left(\frac{v}{c}\right)ct_2 + \gamma x_2 = \gamma x_2 \qquad \text{since } t_2 = 0 \tag{7.1.18}$$

Since, by definition $x'_2 = L_0$ and $x_2 = L$, we therefore have the *FitzGerald contraction*,

$$L = \frac{L_0}{\gamma} \tag{7.1.19}$$

Since $\gamma > 1$, we have $L < L_0$. The length of the stick L, as seen by the observer in $\mathcal{K}$, appears to have decreased compared to its length L_0 in the frame $\mathcal{K}'$ in which it is at rest $\Rightarrow$ sticks oriented parallel to their direction of motion appear to shrink in length!

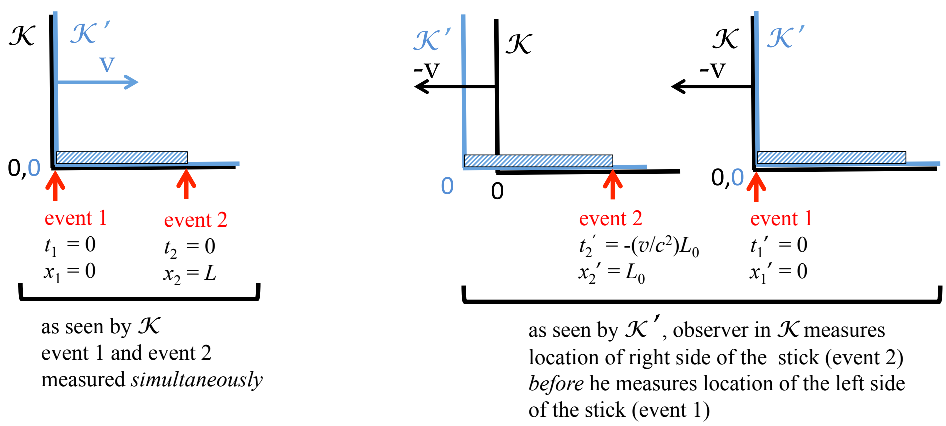
What does an observer in frame $\mathcal{K}'$ see as the observer in $\mathcal{K}$ makes his two measurements. In frame $\mathcal{K}'$, event 1 occurs at coordinates $t'_1 = 0$ and $x'_1 = 0$, since $t_1 = x_1 = 0$. But the coordinates of event 2 occur at,

$$x'_2 = -\gamma\left(\frac{v}{c}\right)ct_2 + \gamma x_2 = \gamma x_2 \qquad \text{and} \qquad ct'_2 = \gamma ct_2 - \gamma\left(\frac{v}{c}\right)x_2 = -\gamma\left(\frac{v}{c}\right)x_2 \tag{7.1.20}$$

since $t_2 = 0$. Using $x_2 = L$, the observer in $\mathcal{K}'$ thus sees the observer in $\mathcal{K}$ measure the position of the right end of the stick at position $x'_2 = \gamma L = L_0$, but the measurement takes place at the time $t'_2 = -\gamma\left(\frac{v}{c^2}\right)L = -\left(\frac{v}{c^2}\right)L_0 < t'_1 = 0$. So, although the observer in $\mathcal{K}$ believes he is making his measurements of the two ends of the stick *simultaneously*, the observer in $\mathcal{K}'$ sees $\mathcal{K}$ measure the position of the right end of the stick *before* he measures the position of the left end of the stick. Since the stick has moved in between $\mathcal{K}$'s two measurements, $\mathcal{K}'$ will see $\mathcal{K}$ measure the stick to be shorter than it really is! This is a second important result of special relativity:

Simultaneity of two events:

Two events which are simultaneous in one inertial frame of reference, are *not* simultaneous in another inertial frame of reference.



Note, if the stick if oriented perpendicular to its direction of motion, then at $t' = t = 0$, the ends of the stick in $\mathcal{K}'$ are at $y'_1 = 0$ and $y'_2 = L_0$. The Lorentz transformation then gives that the observer in $\mathcal{K}$ will measure the ends to be at $y_1 = 0$ and $y_2 = L_0$. Sticks oriented perpendicular to their direction of motion do not appear to change their length.

Time dilation

Consider a clock at rest in frame $\mathcal{K}'$. The time between ticks of the clock in its rest frame $\mathcal{K}'$ is Δt_0. What will be the time between ticks Δt as measured by an observer in frame $\mathcal{K}$?

Let the clock be at the origin of $\mathcal{K}'$. Let event 1 be the first tick of the clock which occurs at time $t'_1 = 0$ and position $x'_1 = 0$. Let event 2 be the second tick of the clock which occurs at time $t'_2 = \Delta t_0$ and position $x'_2 = 0$. Using the inverse Lorentz transformation

$$ct = \gamma ct' + \gamma\left(\frac{v}{c}\right)x' \tag{7.1.21}$$

we can compute the times of these two events as seen by the observer in frame $\mathcal{K}$.

Event 1 occurs at time $t_1 = 0$ since $t'_1 = x'_1 = 0$. Event 2 occurs at time $ct_2 = \gamma ct'_2$, since $x'_2 = 0$. Since $t'_2 = \Delta t_0$, and $t_2 - t_1 = \Delta t$, we conclude that,

$$\Delta t = \gamma \Delta t_0 \tag{7.1.22}$$

Since $\gamma > 1$, the time between ticks of the clock is greater in $\mathcal{K}$ than in the clock's rest frame $\mathcal{K}'$. The moving clock appears to have slowed down! This is known as *time dilation*.

4-Vector Notation

3-vectors: We will start by reviewing some of the properties of ordinary three-component spatial vectors.

For specifying the location of a point in *space*, we set up a right handed orthonormal coordinate system, and then specify the coordinates of the point with reference to that orthonormal basis, (x, y, z). To simplify notation, we call (x, y, z) a *vector* and denote it as $\vec{r}$ or $\mathbf{r} = (x, y, z)$. If we want to refer to a particular component of the vector, we write r_i, with $i = 1, 2$, or 3. We say that any triple of numbers (a_1, a_2, a_3), that transforms under a rotation of the coordinate basis the same way as the position vector (x, y, z), is also a vector.

If we want to take a dot product between two vectors we write $\mathbf{v} \cdot \mathbf{u}$, or using the summation convention we could write $v_i u_i$. The dot product is independent of which orthonormal coordinates we use to express the vector, i.e., if $\mathbf{v}$ and $\mathbf{u}$ are given by (v_1, v_2, v_3) and (u_1, u_2, u_3) with respect to one coordinate basis, and (v_1', v_2', v_3') and (u_1', u_2', u_3') are the coordinates of the same vectors with respect to a rotated coordinate basis, then $\mathbf{v} \cdot \mathbf{u} = u_i v_i = u_i' v_i'$.

If we have the coordinates (x, y, z) of the vector with respect to one coordinate basis, and we would like the coordinates (x', y', z') of the same vector with respect to a rotated coordinate basis, that transformation is given by a linear matrix operation, $(x', y', z') = \mathbb{M} \cdot (x, y, z)$, with $\mathbb{M}$ the rotation matrix. In terms of components, and using the summation convention, we could write this as $x_i' = M_{ij} x_j$. Since the dot product of two vectors must have the same value in any orthonormal coordinate basis, we must have

$$|\mathbf{r}|^2 = \mathbf{r} \cdot \mathbf{r} = x_i'^2 = x_i' x_i' = M_{ij} x_j M_{ik} x_k = x_j x_j = x_j^2 \tag{7.1.23}$$

For this to be true it must be that

$$M_{ij} M_{ik} = \delta_{ik} \qquad \text{with } \delta_{jk} \text{ the Kronecker delta.} \tag{7.1.24}$$

In matrix notation, with the transpose $\mathbb{M}^t$ of the matrix $\mathbb{M}$ defined by $[\mathbb{M}^t]_{ij} = M_{ji}$, we can write this condition as

$$\mathbb{M}^t \cdot \mathbb{M} = \mathbb{I}, \qquad \text{with } \mathbb{I} \text{ the identity matrix.} \tag{7.1.25}$$

This then shows that for rotation matrices, the transpose of the matrix must be equal to the inverse of the matrix, $\mathbb{M}^t = \mathbb{M}^{-1}$. Such a matrix is said to be *orthonormal.*

4-vectors: We now extend the same ideas to describe the coordinates of events in *space-time*.

In special relativity we refer to the coordinates of an event in space-time with a quadruple of numbers, (x, y, z, ict). In these coordinates, the time t is measured in units of ct to give it the same dimension of length as the other coordinates. We take the time component to be *imaginary* for reasons to be made clear soon. We call this quadruple of numbers the position *4-vector*, and denote it as x_μ, with $\mu = 1, 2, 3$ or 4. So $x_2 \equiv y$ and $x_4 \equiv ict$. Unlike 3-vectors, the custom (following Einstein) is not to represent the 4-vector some with some notation like $\mathbf{x} = (x, y, z, ict)$, but rather to represent it simply as x_μ. Depending on the context, x_μ could mean specifically the μ-th component of the 4-vector, or it could represent the entire 4 components of the 4-vector.

If x_μ are the coordinates of an event in one inertial frame of reference $\mathcal{K}$, then the coordinates of the same event as measured in another inertial frame of reference $\mathcal{K}'$ are given by $x_\mu' = a_{\mu\nu} x_\nu$, where $a_{\mu\nu}$ is the matrix of the linear Lorentz transformation from $\mathcal{K}$ to $\mathcal{K}'$. For the special case where $\mathcal{K}'$ moves with velocity $v\hat{\mathbf{x}}$ with respect to $\mathcal{K}$, the elements of the Lorentz transformation matrix are given by Eq. (7.1.16). In terms of our notation x_μ, with $x_4 = ict$, we can rewrite Eq. (7.1.16) as,

$$\begin{aligned} x_1' &= \gamma \left(x_1 + i \left(\frac{v}{c} \right) x_4 \right) \\ x_2' &= x_2 \\ x_3' &= x_3 \\ x_4' &= \gamma \left(x_4 - i \left(\frac{v}{c} \right) x_1 \right) \end{aligned} \tag{7.1.26}$$

If we denote as $\mathcal{L}$ the Lorentz transformation from $\mathcal{K}$ to $\mathcal{K}'$, then the above equations give the elements of the matrix $a(\mathcal{L})$ of this transformation as,

$$a(\mathcal{L}) = \begin{pmatrix} \gamma & 0 & 0 & i\left(\frac{v}{c}\right)\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -i\left(\frac{v}{c}\right)\gamma & 0 & 0 & \gamma \end{pmatrix} \tag{7.1.27}$$

The inverse Lorentz transformation $\mathcal{L}^{-1}$ from $\mathcal{K}'$ to $\mathcal{K}$ has a matrix that is obtained by taking $v \to -v$ in the above. We thus see that the matrix of the inverse transformation, $a_{\mu\nu}(\mathcal{L}^{-1}) = a_{\nu\mu}(\mathcal{L})$, or $a(\mathcal{L}^{-1}) = a^t(\mathcal{L})$ – the matrix of the inverse transformation $\mathcal{L}^{-1}$ is the transpose of the matrix of $\mathcal{L}$.

We define the square of a 4-vector by,

$$x_\mu^2 \equiv x_\mu x_\mu = x^2 + y^2 + z^2 + (ict)^2 = x^2 + y^2 + z^2 - c^2t^2 = r^2 - c^2t^2 \tag{7.1.28}$$

where $\mathbf{r} = (x, y, z)$ stands for the spatial components of x_μ, and $r = |\mathbf{r}|$. The equation $x_\mu x_\mu = 0$ then defines the coordinates of the outward traveling wavefront of a pulse of light that leaves the origin at $t = 0$. Since this must be the same in all inertial frames of reference, we conclude that we must have,

$$x'_\mu x'_\mu = x_\mu x_\mu \tag{7.1.29}$$

That is, the square of a 4-vector must have the same value in any inertial frame of reference. This is the reason why we defined the time component of the position 4-vector to have the imaginary factor of i. With the factor i, we can then represent the square of the 4-vector as simply the sum of the squares of its four components, as in Eq. (7.1.28).

Sometimes, particularly when one is discussing *general relativity* one writes the 4-vector as (x, y, z, ct), without the factor i in the time component. The square of the 4-vector is then written as $x_\mu g_{\mu\nu} x_\nu$, with

$$g_{\mu\nu} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{7.1.30}$$

called the *metric tensor*. But for special relativity we do not need such a fancy notation, and we will stick with $x_\mu = (x, y, z, ict)$.

Note, the condition $|\mathbf{r}|^2 = 0$ for a 3-vector $\mathbf{r}$ defines a unique point in three dimensional space. However the condition $x_\mu^2 = x_\mu x_\mu = 0$ defines a surface in four dimensional space-time. This surface is known as the *light cone* of the origin at $x_\mu = 0$. For $t > 0$, the light cone is all the points in space-time that will be reached by a pulse of light emitted at the origin at $t = 0$. For $t < 0$, the light cone is all the points in space-time that, if a pulse of light was emitted on this surface, the pulse would reach the origin at $t = 0$. The $t > 0$ part of the surface is called the *forward* light cone; the $t < 0$ part of the surface is called the backwards light cone.

Now the requirement of Eq. (7.1.29) imposes a condition on the Lorentz transformation matrix $a(\mathcal{L})$. We must have,

$$x'_\mu x'_\mu = a_{\mu\nu}(\mathcal{L}) x_\nu \; a_{\mu\lambda}(\mathcal{L}) x_\lambda = x_\nu x_\nu \tag{7.1.31}$$

This requires that the matrix $a(\mathcal{L})$ must satsify,

$$a_{\mu\nu}(\mathcal{L})a_{\mu\lambda}(\mathcal{L}) = \delta_{\nu\lambda} \qquad \Rightarrow \qquad a^t_{\nu\mu}(\mathcal{L})a_{\mu\nu}(\mathcal{L}) = \delta_{\nu\lambda} \qquad \Rightarrow \qquad a^t(\mathcal{L}) = a^{-1}(\mathcal{L}) \tag{7.1.32}$$

The transpose of the Lorentz transformation matrix is equal to the inverse of the matrix. The Lorentz transformation matrices $a(\mathcal{L})$ are therefore 4×4 orthonormal matrices operating on space-time coordinates.

Following Eq. (7.1.27) we noted that $a(\mathcal{L}^{-1}) = a^t(\mathcal{L})$. Since we now have $a^t(\mathcal{L}) = a^{-1}(\mathcal{L})$, we therefore conclude what may seem obvious, $a(\mathcal{L}^{-1}) = a^{-1}(\mathcal{L})$. The matrix of the inverse transformation $\mathcal{L}^{-1}$ is the inverse of the matrix of $\mathcal{L}$.

Now if $\mathcal{L}_1$ is the Lorentz transformation from $\mathcal{K}$ to $\mathcal{K}'$, and if $\mathcal{L}_2$ is the Lorentz transformation from $\mathcal{K}'$ to $\mathcal{K}''$, then the Lorentz transformation $\mathcal{L}_2\mathcal{L}_1$ from $\mathcal{K}$ to $\mathcal{K}''$ is given by the matrix,

$$a_{\mu\nu}(\mathcal{L}_2\mathcal{L}_1) = a_{\mu\lambda}(\mathcal{L}_2)\, a_{\lambda\nu}(\mathcal{L}_1) \tag{7.1.33}$$

If we apply the above with $\mathcal{L}_1 \equiv \mathcal{L}$ and $\mathcal{L}_2 \equiv \mathcal{L}^{-1}$, so that $\mathcal{L}_2\mathcal{L}_1 = \mathcal{I}$ the identity transformation, we then get,

$$a_{\mu\nu}(\mathcal{L}^{-1}\mathcal{L}) = \delta_{\mu\nu} = a_{\mu\lambda}(\mathcal{L}^{-1})\, a_{\lambda\nu}(\mathcal{L}) \qquad \Rightarrow \qquad a(\mathcal{L}^{-1}) = a^{-1}(\mathcal{L}) \tag{7.1.34}$$

So again we have that the matrix of the inverse transformation $\mathcal{L}^{-1}$ is the inverse of the matrix of $\mathcal{L}$.

Difference 4-Vector

If we have two position 4-vectors, $x_{\mu 1}$ and $x_{\mu 2}$, we can define the difference 4-vector,

$$\Delta x_\mu = x_{\mu 2} - x_{\mu 1} = (x_2 - x_1,\ y_2 - y_1,\ z_2 - z_1,\ ic[t_2 - t_1]) = (\Delta\mathbf{r},\ ic\Delta t) \tag{7.1.35}$$

Let us label these two positions so that $t_2 \geq t_1$.

The square of the difference 4-vector is,

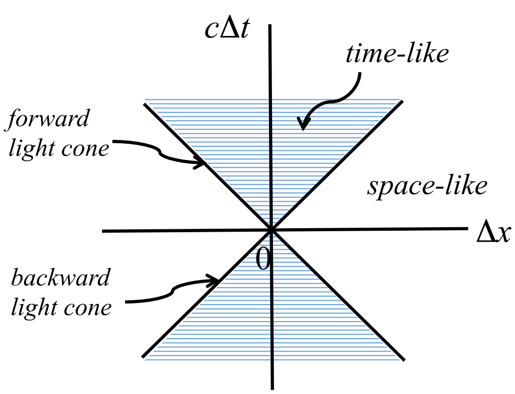
$$\Delta x_\mu^2 = |\Delta\mathbf{r}|^2 - c^2\Delta t^2 \tag{7.1.36}$$

When $\Delta x_\mu^2 = 0$, $x_{\mu 2}$ is on the *light cone* of $x_{\mu 1}$. A pulse of light leaving position $\mathbf{r}_1$ at time t_1 will arrive at position $\mathbf{r}_2$ exactly at time t_2.

When $\Delta x_\mu^2 < 0$, a pulse of light leaving position $\mathbf{r}_1$ at time t_1 will arrive at position $\mathbf{r}_2$ at a time t *earlier* than t_2. The positions $x_{\mu 1}$ and $x_{\mu 2}$ are said to be *time-like*. There exists an inertial frame of reference $\mathcal{K}'$ in which $x_{\mu 1}$ and $x_{\mu 2}$ occur at the same point in space, $\mathbf{r}'_1 = \mathbf{r}'_2$, but at different times, $t'_1 \neq t'_2$. To find this frame, just transform from the original frame $\mathcal{K}$ to a new $\mathcal{K}'$ that moves with velocity $\mathbf{v} = \Delta\mathbf{r}/\Delta t$ as seen by $\mathcal{K}$. For two time-like points $x_{\mu 1}$ and $x_{\mu 2}$, what happens at $x_{\mu 1}$ can effect what will happen at $x_{\mu 2}$, if the agent causing that effect travels slower than the speed of light.

When $\Delta x_\mu^2 > 0$, a pulse of light leaving position $\mathbf{r}_1$ at time t_1 will arrive at position $\mathbf{r}_2$ at a time t *later* than t_2. The positions $x_{\mu 1}$ and $x_{\mu 2}$ are said to be *space-like*. There exists an inertial frame of reference in which $x_{\mu 1}$ and $x_{\mu 2}$ occur at the same time, $t'_1 = t'_2$, but at different points in space, $\mathbf{r}'_1 \neq \mathbf{r}'_2$. What happens at $x_{\mu 1}$ can never effect what happens at $x_{\mu 2}$ because it would require an agent traveling faster than the speed of light to cause such an effect.

We sketch these different regions in the space-time diagram below, where we also include the case $\Delta t < 0$ (i.e. $t_2 < t_1$).



An event happening at the origin can effect what happens in the time-like region bounded by the forward light cone.

An event happening in the time-like region bounded by the backward light cone can effect what happens at the origin.

An event in the space-like region can never effect, nor be effected by, what happens at the origin.

Differential 4-Vector and the Proper Time

For two position 4-vectors infinitesmally close to each other, we can define the differential 4-vector.

If $x_\mu = (x, y, z, ict)$, and $x_\mu + dx_\mu = (x + dx, y + dy, z + dz, ic[t + dt])$, then the differential 4-vector is,

$$dx_\mu = (dx, dy, dz, icdt) \tag{7.1.37}$$

We can then write for the square of this differential 4-vector,

$$-(dx_\mu)^2 = c^2dt^2 - dr^2 \equiv c^2ds^2 \tag{7.1.38}$$

Since this is the square of a 4-vector, it is a Lorentz invariant scalar, i.e. it has the same value in all inertial frames of reference.

When x_μ and $x_\mu + dx_\mu$ are time-like, i.e. $dx_\mu^2 < 0$, we can write,

$$ds^2 = dt^2\left[1 - \frac{1}{c^2}\left(\frac{dx_1}{dt}\right)^2 - \frac{1}{c^2}\left(\frac{dx_2}{dt}\right)^2 - \frac{1}{c^2}\left(\frac{dx_2}{dt}\right)^2\right] = dt^2\left[1 - \frac{v^2}{c^2}\right] = \frac{dt^2}{\gamma^2} \qquad \text{with } v < c \tag{7.1.39}$$

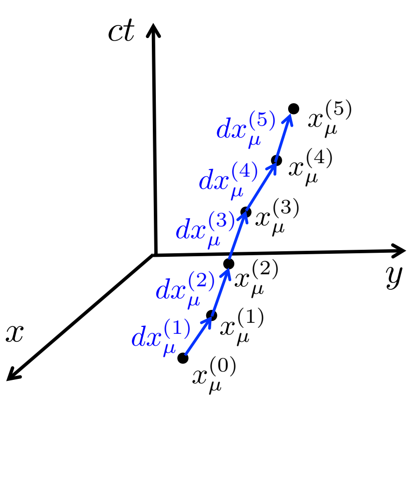
and so,

$$\boxed{ds = \frac{dt}{\gamma}} \qquad \text{is called the } \textit{proper time} \text{ interval between two infinitesmally close time-like events.} \tag{7.1.40}$$

The speed v that appears in the factor γ is the speed it takes to go from $\mathbf{r}$ to $\mathbf{r} + d\mathbf{r}$ in the time dt. If the points x_μ and $x_\mu + dx_\mu$ are two nearby points on the space-time trajectory of a moving particle, then v is just the speed of the particle.

Because ds is defined in terms of the square dx_μ^2, it is a Lorentz invariant scalar, having the same value in all inertial frames of reference.

If, in frame $\mathcal{K}$, the two events at x_μ and $x_\mu + dx_\mu$ occur at the exact same spatial position, i.e. $d\mathbf{r} = (dx, dy, dz) = 0$, then the proper time interval ds between the two events is just the same as the time interval dt between the two events in frame $\mathcal{K}$. Or to put it another way, the proper time interval ds between two events, is the time interval between those events as measured in the inertial frame of reference in which those two events occur at the same position in space. To put it yet another way, ds is the time between ticks on a clock that is at rest, and so $ds = dt/\gamma$ is essentially the same as the time dilation result of Eq. (7.1.22), $\Delta t_0 = \Delta t/\gamma$.



If x_μ represents the position of a particle moving on a space-time trajectory, then it is often convenient to parameterize that trajectory by the proper time $s = \int ds$. To see how to do this, imagine labeling discrete nearby points on the trajectory by an index i, so that the trajectory is given by the set of points $\{x_\mu^{(i)}\}$. Then define $dx_\mu^{(i)} = x_\mu^{(i)} - x_\mu^{(i-1)}$, and $ds_i = \sqrt{-\left[dx_\mu^{(i)}\right]^2/c^2}$. Note, because the particle speed must satisfy $v < c$, $-\left[dx_\mu^{(i)}\right]^2$ is always positive. Then, if we define the initial point $x_\mu^{(0)}$ to be at proper time s_0, and define $s_i = \sum_{j=1}^{i} ds_j + s_0$, we can then label the points on the trajectory $x_\mu^{(i)}$ as $x_\mu(s_i)$. Letting the spacing between the points go to zero, we have a continuous parameterization of the trajectory, $x_\mu(s)$. Viewed this way, the proper time s is the analog of the arc length along a curve in three dimensional space, only using the distance metric appropriate to space-time. Physically, s is the time that would be measured on a clock that is moving along with the particle.

Other 4-Vectors

Above we defined the spatial 4-vector x_μ. We now generalize the concept of 4-vector to include any quadruple of numbers (u_1, u_2, u_3, u_4) that transform under a Lorentz transformation the same way as x_μ, i.e. $u'_\mu = a_{\mu\nu} u_\nu$.

<u>4-velocity</u>:

Consider a particle that is moving on a trajectory $x_\mu(s)$ in space-time. We define the 4-velocity u_μ of the particle as the derivative of the 4-position with respect to the proper time interval,

$$u_\mu \equiv \frac{dx_\mu}{ds} \equiv \dot{x}_\mu \tag{7.1.41}$$

From here on, a dot over a symbol will denote a derivative with respect to the proper time interval ds.

Since x_μ is a 4-vector, and ds is a Lorentz invariant scaler, then it follows that u_μ must transform like x_μ, and so it is a 4-vector.

Using Eq. (7.1.40) for ds we have,

$$u_\mu = \frac{dx_\mu}{ds} = \gamma \frac{dx_\mu}{dt} \tag{7.1.42}$$

So the spatial components of u_μ are $\mathbf{u} = \gamma \mathbf{v}$, and the time component is $u_4 = \gamma d(ict)/dt = ic\gamma$.

Note, in the above expressions, the speed v that enters the factor $\gamma = \dfrac{1}{\sqrt{1 - v^2/c^2}}$ is just the speed of the particle.

The square of the 4-velocity is thus,

$$u_\mu u_\mu = \gamma^2 v^2 - c^2\gamma^2 = \gamma^2(v^2 - c^2) = \frac{v^2 - c^2}{1 - v^2/c^2} = -c^2 \tag{7.1.43}$$

<u>4-acceleration</u>:

We can similarly define the 4-acceleration of a particle as,

$$\alpha_\mu \equiv \frac{du_\mu}{ds} = \dot{u}_\mu = \ddot{x}_\mu = \gamma \frac{du_\mu}{dt} \tag{7.1.44}$$

We will write α_μ in terms of the ordinary acceleration $\mathbf{a} = d\mathbf{v}/dt$ later.

<u>4-gradient</u>:

We define the 4-gradient as,

$$\frac{\partial}{\partial x_\mu} \equiv \left(\boldsymbol{\nabla}, -\frac{i}{c}\frac{\partial}{\partial t} \right) \tag{7.1.45}$$

We now will prove that the 4-gradient is indeed a 4-vector.

Using the chain rule for differentiation, we can write,

$$\frac{\partial}{\partial x'_\mu} = \frac{\partial x_\lambda}{\partial x'_\mu} \frac{\partial}{\partial x_\lambda} \tag{7.1.46}$$

But since x_λ is obtained from x'_μ by the inverse Lorentz transformation $\mathcal{L}^{-1}$, we have $x_\lambda = a_{\lambda\mu}(\mathcal{L}^{-1}) x'_\mu$, and so

$$\frac{\partial x_\lambda}{\partial x'_\mu} = a_{\lambda\mu}(\mathcal{L}^{-1}) = a^{-1}_{\lambda\mu}(\mathcal{L}) = a_{\mu\lambda}(\mathcal{L}) \tag{7.1.47}$$

where in the second step we use the result that the matrix of the inverse transformation $\mathcal{L}^{-1}$ is the inverse of the matrix of $\mathcal{L}$, and in the third step we used the result that the inverse of the matrix of the transformation is equal to the transpose of the matrix. Using this in Eq. (7.1.46) we then get,

$$\frac{\partial}{\partial x'_\mu} = a_{\mu\lambda}(\mathcal{L})\,\frac{\partial}{\partial x_\lambda} \tag{7.1.48}$$

which is just the transformation law of a 4-vector. Hence the 4-gradient is a 4-vector.

<u>wave equation operator</u>:

We can now consider the square of the 4-gradient, which is a Lorentz invariant scalar differential operator,

$$\left(\frac{\partial}{\partial x_\mu}\right)^2 = \nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} \tag{7.1.49}$$

This is just the wave equation operator!

<u>inner products</u>:

For any 4-vector u_μ we showed that $u_\mu u_\mu$ is a Lorentz invariant scalar. Consider now a second 4-vector v_μ. The inner product of the two 4-vectors is defined to be,

$$u_\mu v_\mu \tag{7.1.50}$$

You should show that this is also a Lorentz invariant scalar, i.e. that $u'_\mu v'_\mu = u_\mu v_\mu$.