Thus far we have not dealt explicitly in this class with relativity. We have implicitly, because Maxwell's equations are in fact a relativistically correct theory, and in principle we could use them to tackle our next problem: extending our calculation of emission from accelerating charges to the case where the charges are moving relativistically. However, we will take a somewhat different route, which is to develop a fully (special) relativistic version of classical electromagnetism, and then use this to deduce the nature of emission from relativistic emitters directly. We will begin this class with a brief review of special relativity, then develop an explicitly-relativistic version of Maxwell's Laws.

I. Lorentz transformations

We begin with a review of special relativity and its application to light. Since this is a review, we will not attempt to demonstrate many results in detail. We will simply remind ourselves of results that we will need later.

A. Position and velocity transformations

The most basic concept in special relativity involves the relationships (hence the name) between different reference frames. Einstein postulated that the laws of physics must be the same in all inertial references frames, but that, in addition, the speed of light must be the same in all reference frames. This second postulate means that rules for transforming between frames cannot be the Galilean ones. In fact, one can show that this postulate requires (together with the assumptions that space is homogenous and isotropic, so the laws of physics do not depend on position or direction) that the correct transformations between reference frames are the Lorentz transformations. Consider two Cartesian reference frames, one described by the coordinates (t, x, y, z) and another by (t', x', y', z'). The primed frame is moving at constant velocity $v\hat{\mathbf{x}}$ as viewed from the unprimed frame. We refer to the velocity separating the two frames as the boost. The coordinates in the two frames are related by

$$t' = \gamma\left(t - \frac{v}{c^2}x\right) \tag{1}$$

$$x' = \gamma(x - vt) \tag{2}$$

$$y' = y \tag{3}$$

$$z' = z \tag{4}$$

where

$$\gamma \equiv \left(1 - \frac{v^2}{c^2}\right)^{-1/2}. \tag{5}$$

The inverse transformation from the primed to the unprimed coordinates is identical except that the $-$ signs in the x and t transformations become plus signs. In the limit $v/c \ll 1$, we have $\gamma \approx 1$, and the transformation rules reduce to the Galilean transformation with which are familiar, in which t and t' are identical (so all observers agree on time), and $x' = x - vt$, so that an object that appears stationary in the unprimed frame will appear to be moving toward negative x' in the primed coordinate system.

The Lorentz transformation can be generalised to an arbitrary direction just by taking the boost $\mathbf{v}$ to be a vector, so $\mathbf{v} = v\hat{\mathbf{n}}$ for some unit vector $\hat{\mathbf{n}}$. In this case the

Lorentz transformation from position $(\mathbf{x}, t)$ to $(\mathbf{x}', t')$ just becomes

$$
\begin{aligned}
t' &= \gamma \left(t - \frac{v \mathbf{x} \cdot \hat{\mathbf{n}}}{c^2} \right) && (6) \\
\mathbf{x}' &= \mathbf{x} + (\gamma - 1)\,(\mathbf{x} \cdot \hat{\mathbf{n}})\hat{\mathbf{n}} - \gamma t \mathbf{v}. && (7)
\end{aligned}
$$

The combination $\mathbf{x} \cdot \hat{\mathbf{n}}$ is just the component of the position vector $\mathbf{x}$ along the velocity direction $\hat{\mathbf{n}}$.

We can also use the Lorentz transformations to deduce the law for how velocities transform from one frame to another. Let $\mathbf{u}'$ be velocity of an object measured by an observer in the primed frame. What will an observer in the unprimed frame measure? For simplicity we will again align our coordinate system so that the velocity between the two frames (but not $\mathbf{u}'$) lies in the x direction.

The observer in the primed frame measures that the object moves a distance

$$
d\mathbf{x}' = \mathbf{u}'\, dt' \tag{8}
$$

over a time dt'. Transforming these distances to the unprimed frame using the Lorentz transformations, we have

$$
\begin{aligned}
dx &= \gamma(dx' + v\, dt') && (9) \\
dy &= dy' && (10) \\
dz &= dz'. && (11)
\end{aligned}
$$

Similarly, the time dt' measured in the primed frame corresponds to a time interval

$$
dt = \gamma \left(dt' + \frac{v}{c^2} dx' \right) \tag{12}
$$

in the unprimed frame. Thus the three components of velocity measured in the unprimed frame are

$$
\begin{aligned}
u_x = \frac{dx}{dt} &= \frac{\gamma(dx' + v\, dt')}{\gamma(dt' + v\, dx'/c^2)} = \frac{u'_x + v}{1 + v u'_x / c} && (13) \\
u_y = \frac{dy}{dt} &= \frac{u'_y}{\gamma(1 + v u'_x / c^2)} && (14) \\
u_z = \frac{dy}{dt} &= \frac{u'_z}{\gamma(1 + v u'_x / c^2)} && (15)
\end{aligned}
$$

The generalisation to an arbitrary orientation of $\mathbf{v}$ and $\mathbf{u}$ is obvious if we just map from u_x in the formulae we have written to $u_{\|} = u(\mathbf{u} \cdot \hat{\mathbf{n}})$, where $\hat{\mathbf{n}}$ is the unit vector parallel to $\mathbf{v}$. The combined magnitude of the y and z components together make up $u_{\perp} = |\mathbf{u} - u_{\|}\hat{\mathbf{n}}|$. This gives the transformation formulae

$$
\begin{aligned}
u_{\|} &= \frac{u'_{\|} + v}{1 + v u'_{\|} / c^2} && (16) \\
u_{\perp} &= \frac{u'_{\perp}}{\gamma(1 + v u'_{\|} / c^2)} && (17)
\end{aligned}
$$

B. Relativistic beaming

One immediate application of the transformation of velocities is to derive relativistic beaming, the phenomenon whereby the radiation produced by relativistic emitters is beamed into a narrow angle around their direction of motion. The key point to realise here is that nothing we did in our derivation of the velocity transformation required that the velocity refer to any material object. We simply considered two points in spacetime separated by position $d\mathbf{x}'$ and time dt' as measured by one particular observer, and then asked what separations $d\mathbf{x}$ and dt would be measured by a different observer. The velocity is just what we get by dividing one of these by the other. The separations in spacetime can be arbitrary, so we are free to apply them to the trajectory taken by photons just as much as by material particles.

Consider an isotropically-emitting radiation source as seen in the primed frame. What does this look like in the unprimed frame, which is separated from the primed frame by a boost $\mathbf{v}$ in direction $\hat{\mathbf{n}}$? To answer this, consider radiation emitted at an angle θ' relative to the boost direction, as measured in the primed frame. The perpendicular and parallel components of the velocity in this frame are

$$u'_\perp = c\sin\theta' \qquad u'_\parallel = c\cos\theta'. \tag{18}$$

If the radiation source is isotropic in the primed frame, the distribution of number of photons emitted versus angle in this frame follows

$$\frac{dN}{d\theta'} \propto \sin\theta' \qquad \Longrightarrow \qquad \frac{dN}{d\cos\theta'} \propto \text{const} \tag{19}$$

where θ runs from 0 (radiation emitted along the direction of the boost) to π (radiation emitted opposite the direction of the boost). The fact that the distribution is uniform in $\cos\theta'$ is simply because of the fact that there is more solid angle available near $\theta' = \pi/2$ (the equator) than near $\theta' = 0$ (the North Pole) or $\theta' = \pi$ (the South Pole).

The corresponding angle as measured in the unprimed frame obeys

$$\cos\theta = \frac{u_\parallel}{\sqrt{u_\parallel^2 + u_\perp^2}} = \frac{(u'_\parallel + v)/(1 + vu'_\parallel/c^2)}{c} = \frac{\cos\theta' + \beta}{1 + \beta\cos\theta'} \tag{20}$$

where $\beta = v/c$. In the second from last step we made use of the velocity transformation formulae, and in the last step we just used the definition of $u'_\parallel$. Note that nothing in the argument used to derive this relationship requires that we be talking about light; this relationship holds for any objects moving in straight lines, and thus is in fact a generic relationship between angles in one reference frame and angle in another. We will make use of this angle transformation rule extensively later on.

We can use this result to derive the distribution of angles in the unprimed frame, which the particle emitter is seen to be moving. In this frame, we have

$$\frac{dN}{d\cos\theta} \propto \frac{dN}{d\cos\theta'}\frac{d\cos\theta'}{d\cos\theta} \propto \frac{d\cos\theta'}{d\cos\theta}, \tag{21}$$

where the last step follows because we have assumed that $dN/d\cos\theta'$ is constant. Using our expression for the relationship between $\cos\theta$ and $\cos\theta'$, it is straightforward to take the derivative and re-arrange to get

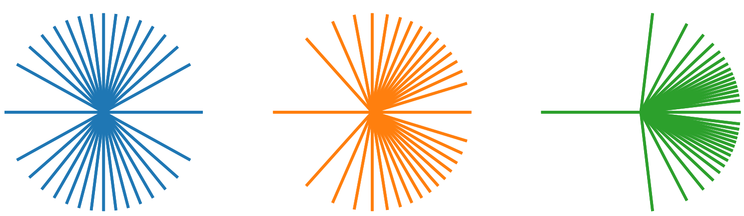
$$\frac{dN}{d\cos\theta} = \frac{1-\beta^2}{2(1-\beta\cos\theta)^2}, \tag{22}$$

where the normalisation factor has been set by required $\int_{-1}^{1} (dN/d\cos\theta)\, d\cos\theta = 1$. For reference, the corresponding fraction of photons beamed into an angular patch of size $d\Omega$ at an angle θ relative to the direction of motion is

$$\frac{dN}{d\Omega} = \frac{(1-\beta^2)\sin\theta}{4\pi(1-\beta\cos\theta)^2}, \tag{23}$$

where the normalisation has been set so that integration over all angles gives 1.

To get an intuitive sense for what these results mean, it is helpful to make some plots. In the plot below, the boost is to the right, and the lines are spaced so that equal numbers of photons are emitted between every pair of lines. Note that, even for $\beta = 0$, the lines are not uniformly distributed around a circle because of the effect that there is more real estate at the equator than at the poles. However, as β increases the radiation becomes increasingly concentrated toward the direction of motion. There are 32 wedges in the figure, and in the case $\beta = 0.9$, 30 of them lie entirely within 90° of the direction of motion, so more than 90% of the photons go in the forward or sideways direction relative to the emitter's direction of motion, and fewer than 10% go behind it.



C. Relativistic Doppler effect

In addition to changing the angular distribution of photons, boosting changes their frequencies. Suppose that an electromagnetic wave being emitted has angular frequency ω' in the emitter's rest frame, which is moving at velocity $\mathbf{v}$ relative to an observer in the unprimed frame. Measured in the emitter's frame, the time between two successive maxima of the electric field is $\Delta t' = 2\pi/\omega'$. In the observer's frame, the Lorentz transform tells us that this time interval corresponds to

$$\Delta t = \frac{2\pi\gamma}{\omega'}. \tag{24}$$

During this time the observer will see the emitter move by a distance $\mathbf{r} = \mathbf{v}\,\Delta t$. The projection of this distance on the direction from the emitter to the observer is

$$d = v\,\Delta t\,\cos\theta, \tag{25}$$

where θ is the angle between $\mathbf{v}$ and the observer. Thus if $\theta = 0$, so the emitter is coming directly at the observer, the emitter gets a distance $v\,\Delta t$ closer to the observer between two successive electric field maxima. If $\theta = \pi$, so the emitter is going directly away, the distance will increase by $v\,\Delta t$. Since the signal is moving at c, the difference in travel time induced by this distance is d/c.

Given this analysis, we can immediately write down the difference in arrival times between two successive pulses in the observer's frame: it is

$$\Delta t_{\mathrm{arr}} = \Delta t - \frac{d}{c} = \Delta t\,(1 - \beta\cos\theta) \tag{26}$$

The angular frequency measured by the observer is therefore

$$\omega = \frac{2\pi}{\Delta t_{\rm arr}} = \frac{\omega'}{\gamma(1-\beta\cos\theta)}. \tag{27}$$

Using the relationship between the angles θ and θ', this can equivalently be written as

$$\omega = \omega'\gamma(1+\beta\cos\theta'). \tag{28}$$

II. Tensor formulation of Lorentz transforms

Having reviewed Lorentz transformations and used them to derive some elementary results, we now proceed with a much more high-level and general discussion of Lorentz transformations, before using them to formulate a relativistic form of Maxwell's equations. The ultimate goal here is to write Maxwell's Laws in a covariant form (where we will define what that means shortly) so that we know how to transform between reference frames that are moving relativistically. This will give us the tools we need to compute emission from relativistic particles in the next lecture. To do that, however, we first need to develop some general mathematical tools.

A. Four-vectors

A crucial property of the Lorentz transformations is that the leave the quantity

$$s^2 = -c^2t^2 + x^2 + y^2 + z^2 \tag{29}$$

unchanged. One can verify this immediately simply by substitution:

$$\begin{aligned}
s'^2 &= -c^2t'^2 + x'^2 + y'^2 + z'^2 \\
&= -\gamma^2(ct - vx/c)^2 + \gamma^2(x-vt)^2 + y^2 + z^2 \\
&= \gamma^2\left(-c^2t^2 - v^2x^2/c^2 + v^2t^2 + x^2\right) + y^2 + z^2 \\
&= \gamma^2\left(1 - v^2/c^2\right)\left(-c^2t^2 + x^2\right) + y^2 + z^2 \\
&= -c^2t^2 + x^2 + y^2 + z^2 \\
&= s^2.
\end{aligned} \tag{30}$$

The quantity s is known as the separation between any two points in space-time, and quantities such as s that are unchanged by a Lorentz transformation are called Lorentz invariants. Just as the invariance of the length from the origin to a position (x, y, z) in three-dimensional space suggests that (x, y, z) can be put together into a single object that is governed by rules about how it transforms (e.g., by rotating our coordinate system), the invariance of the separation under Lorentz transformations suggests the same about vectors in relativity. We define the four-vector $\vec{x}$ describing any space-time position as

$$\vec{x} = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{31}$$

The significance of writing the indices in superscript will become apparent momentarily. Note that we write four-vectors with the vector symbol, $\vec{x}$, to distinguish them from ordinary three-vectors, which we wrote in boldface, $\mathbf{x}$. We can also denote $\vec{x}$ as x^μ, where Greek indices are by convention understood to run from 0 to 3. A space-time point defined in this way can also be called an "event."

In order to capture the − sign in the definition of the separation, we need to introduce another object, called the metric. In this case it is the Minkowski metric, which is

$$\eta_{\mu\nu} = \eta^{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{32}$$

The purpose of the metric is to specify how a given coordinate component contributes to length. The separation is found by multiplying $\mathbf{x}$ by itself and the metric. Specifically, we can write

$$s^2 = \sum_{\mu=0}^{3}\sum_{\nu=0}^{3} \eta_{\mu\nu} x^\mu x^\nu \equiv \eta_{\mu\nu} x^\mu x^\nu \tag{33}$$

where in the last step we have introduced the Einstein summation convention, whereby repeated Greek indices, with one above and one below, are always summed over from 0 to 3, even if we don't write out the summation symbol. One implication of this convention is that the choice of Greek letter doesn't matter, because we're going to sum over it anyway. Thus if we changed all the μ's to σ's or some other Greek letter in this equation, the meaning would not change in the slightest.

The combination $\eta_{\mu\nu} x^\nu$ can also be used to define a new vector

$$x_\mu = \eta_{\mu\nu} x^\nu = \begin{pmatrix} -ct \\ x \\ y \\ z \end{pmatrix}. \tag{34}$$

This now explains the distinction between indices above and below: a vector with an index below is just a vector with the index above that has been multiplied by the metric. We can therefore also write s^2 as

$$s^2 = x^\mu x_\mu. \tag{35}$$

An index in superscript is called a covariant index, and one in subscript is called a contravariant index. A four-vector has only one index, so we can also refer to four-vectors as being covariant or contravariant.

With these definitions, we can now re-introduce the Lorentz transformation in the form of a matrix multiplication. One can verify by direct substitution that boosting by a velocity v in the x direction transforms the four-vector $\vec{x}$ to $\vec{x}'$ as

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \tag{36}$$

where

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{37}$$

The generalisation of this to a boost in an arbitrary direction is (we will not prove this – see Jackson's textbook *Classical Electrodynamics* for an explicit calculation) is

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & -\gamma\beta_1 & -\gamma\beta_2 & -\gamma\beta_3 \\ -\gamma\beta_1 & 1+\frac{(\gamma-1)\beta_1^2}{\beta^2} & \frac{(\gamma-1)\beta_1\beta_2}{\beta^2} & \frac{(\gamma-1)\beta_1\beta_3}{\beta^2} \\ -\gamma\beta_2 & \frac{(\gamma-1)\beta_1\beta_2}{\beta^2} & 1+\frac{(\gamma-1)\beta_2^2}{\beta^2} & \frac{(\gamma-1)\beta_2\beta_3}{\beta^2} \\ -\gamma\beta_3 & \frac{(\gamma-1)\beta_1\beta_3}{\beta^2} & \frac{(\gamma-1)\beta_2\beta_3}{\beta^2} & 1+\frac{(\gamma-1)\beta_3^2}{\beta^2} \end{pmatrix} \tag{38}$$

where $\beta_{1,2,3}$ are the components of $\mathbf{v}/c$ in the x, y, and z directions.

The equivalent transformation for contravariant vectors is just

$$x'_\mu = \eta_{\mu\nu}x'^\nu = \eta_{\mu\nu}\Lambda^\nu{}_\sigma x^\sigma = \eta_{\mu\nu}\Lambda^\nu{}_\sigma \eta^{\sigma\tau}x_\tau \equiv \tilde{\Lambda}_\mu{}^\tau x_\tau, \tag{39}$$

where

$$\tilde{\Lambda}_\mu{}^\tau \equiv \eta_{\mu\nu}\Lambda^\nu{}_\sigma \eta^{\sigma\tau} \tag{40}$$

is the boost matrix for contravariant components. Since s^2 is invariant, we can immediately deduce an important constraint on the relationship between Λ and $\tilde{\Lambda}$:

$$s^2 = x^\mu x_\mu = x'^\mu x'_\mu = \Lambda^\mu{}_\sigma \tilde{\Lambda}_\mu{}^\tau x^\sigma x_\tau \tag{41}$$

If therefore immediately follows that the combination

$$\Lambda^\mu{}_\sigma \tilde{\Lambda}_\mu{}^\tau = \delta^\sigma{}_\tau, \tag{42}$$

where $\delta^\sigma{}_\tau$ is 1 if the indices are the same, and zero otherwise. This is known as the Kroenecker δ.

We can use this relation to write down inverse transformations as well. We start with the forward transformation

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \tag{43}$$

and multiply on both sides by $\tilde{\Lambda}_\mu{}^\sigma$, giving

$$\tilde{\Lambda}_\mu{}^\sigma x'^\mu = \tilde{\Lambda}_\mu{}^\sigma \Lambda^\mu{}_\nu x^\nu = \delta^\sigma_\nu x^\nu = x^\sigma \qquad \Longrightarrow \qquad x^\mu = \tilde{\Lambda}^\mu{}_\nu x'^\nu. \tag{44}$$

Thus the inverse transform is the same as the forward one, but with Λ replaced by $\tilde{\Lambda}$.

We define a general four-vector $\vec{A}$ as any object that transforms like $\vec{x}$ under boosts; that is, a four-vector is any quantity where the rule for transforming it from one frame to another is exactly the same as the rule for transforming space-time coordinates. Because any four-vector $\vec{A}$ follows the same transformation rules as $\vec{x}$, and because the transformation rules are what guarantee that $x^\mu x_\mu$ is invariant under boosts, it immediately follows that $A_\mu A^\mu$ is constant under Lorentz transformations for any four-vector. This is completely analogous to the way that, in ordinary three-dimensional space, a vector is not just any collection of three numbers; it is a collection of three numbers that has the property that, as you rotate it or translate it, its magnitude stays constant. A four-vector is just the generalisation of this idea to Minkowski space.

B. Four-velocity and four-momentum

There are many examples of four-vectors, but we will pause here to define two that we will need in our calculations later on. One arises from the question "How should we define velocities in relativity?" Clearly the spatial and time separations between any two points depends on the frame of the observer, and the velocity transformation rules we wrote down earlier are not the same transform rules we wrote for $\vec{x}$. Can we define something that is a four-vector, and does transform like $\vec{x}$, but that acts like the velocity we're used to in the non-relativistic limit?

To do so, first note that the separation in space and time between any two positions $\vec{x}_1$ and $\vec{x}_2$, $d\vec{x} = \vec{x}_2 - \vec{x}_1$, clearly is a four-vector, since the Lorentz transformation rules are linear. So if a particular object passes through two points in space-time, the separation between those points is a four-vector. To make this into something like a velocity, we need to consider two points that are infinitesimally-close to one another and divide by an infinitesimal time; the question is, which time, since observers disagree about time? The answer is to use the separation s between the two points to define the time. Since s invariant, we can see immediately that for any two points in space-time that are an infinitesimal time t and infinitesimal distances (dx, dy, dz) apart, the quantity

$$d\tau^2 = dt^2 - c^{-2}\left(dx^2 + dy^2 + dz^2\right) \tag{45}$$

must be invariant as well, since it is just $-s^2/c^2$. We refer to $d\tau$ as the proper time; it is the time between two events in an object's life that would be measured in a frame in which the object is at rest, so $dx = dy = dz = 0$. We therefore define the four-velocity of any object as

$$U^\mu = \frac{dx^\mu}{d\tau}, \tag{46}$$

This is automatically a four-vector because the numerator is a four-vector and the denominator is a Lorentz invariant. In the frame co-moving with the object, proper time $d\tau$ is equal to the time dt, and thus we immediately have

$$U^\mu = c\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{47}$$

but we can boost this to get the velocity in any frame we want. In particular, if we boost by a velocity u in the x direction, then the velocity becomes

$$U'^\mu = \Lambda^\mu{}_\nu U^\nu = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} c \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = c\begin{pmatrix} \gamma \\ -\beta\gamma \\ 0 \\ 0 \end{pmatrix}. \tag{48}$$

In the non-relativistic limit $\gamma \approx 1$, so the new x velocity is $-u$, exactly as we would have expected for Galilean rather than Einsteinian relativity.

Keeping the same logic, we can now immediately define two more four-vectors: the four-acceleration is

$$a^\mu = \frac{d^2 U^\mu}{d\tau^2}, \tag{49}$$

and the four-momentum is

$$P^\mu = m_0 U^\mu \tag{50}$$

where m_0 is the particle rest mass, which is also an invariant since it is defined in the particle's rest frame. The zeroth component is P^μ has a natural physical interpretation: in a frame where the particle velocity $v \ll c$,

$$P^0 = m_0\gamma c = m_0 c\left(1 - \frac{v^2}{c^2}\right)^{-1/2} = \frac{1}{c}\left(m_0 c^2 + \frac{1}{2}m_0 v^2 + \cdots\right), \tag{51}$$

where in the last step we took the Taylor expansion in the limit of small v/c. We can identify the quantity in parentheses as the energy: $m_0 c^2$ is the rest energy and $m_0 v^2/2$ is the non-relativistic kinetic energy. Thus $P^0 = E/c$, where E is the particle energy. This last point allows us to extend the definition of the four-vector momentum even to massless particles like photons:

$$P^{\mu} = \begin{pmatrix} E/c \\ \mathbf{p} \end{pmatrix}, \tag{52}$$

where $\mathbf{p}$ is the ordinary three-momentum.

C. Tensors and gradients

We can also extend the concept of four-vectors to objects that have more than one index. We call these generalised objects tensors, and the number of indices the rank of the tensor, so that four-vectors are tensors of rank 1, and Lorentz-invariant scalars like the separation s and the proper time $d\tau$ are tensors of rank 0. Conceptually, a tensor of rank 2 can be thought of as being like a matrix, with rows and columns. However, just as not every set of four numbers is a four-vector, not every 4×4 matrix is a rank 2 tensor. What defines a four-vector is the way it transforms under boosts, and this is true of tensors in general. A rank 2 tensor is a matrix that obeys the transformation rules

$$T'^{\mu\nu} = \Lambda^{\mu}{}_{\sigma} \Lambda^{\nu}{}_{\tau} T^{\sigma\tau}. \tag{53}$$

Here the μ and ν are covariant indices, but we can equally well define contravariant indices, which obey transformation rules analogous to those for contravariant vectors:

$$T'_{\mu\nu} = \tilde{\Lambda}_{\mu}{}^{\sigma} \tilde{\Lambda}_{\nu}{}^{\tau} T_{\sigma\tau}. \tag{54}$$

One can also have mixed tensors containing both covariant and contravariant indices:

$$T'^{\ \nu}_{\mu} = \tilde{\Lambda}_{\mu}{}^{\sigma} \Lambda^{\nu}{}_{\tau} T_{\sigma}{}^{\tau}. \tag{55}$$

One can also define tensors of arbitrary rank just by adding more Λ boost factors. Rank 0 tensors have zero Λ factors, which makes sense since they are quantities that are the same in any reference frame.

One particularly important reason to introduce the idea of tensors is that, as we will now show, the gradient of a four-vector field (i.e., a vector that is a function of position in space) is a rank 2 tensor field, and in general the gradient of a tensor field of rank n is another tensor field of rank $n+1$. To show this, we must show that the object we produce by taking derivatives of a four-vector field obeys the transformation rules for tensors.

To show this, consider a covariant four-vector $\vec{A}(\vec{x})$ which is a function of the position $\vec{x}$. Let us define

$$B(\mu, \nu) \equiv \frac{\partial}{\partial x^{\mu}} A^{\nu}, \tag{56}$$

where we write $B(\mu, \nu)$ rather than $B^{\mu\nu}$ or something like that because we have not yet proven that B is tensor, or determined what type of tensor it is – right now it is just a matrix with two indices μ and ν. Now let us consider what an observer in a different reference frame would measure for B. This is

$$
\begin{aligned}
B'(\mu,\nu) &= \frac{\partial}{\partial x'^{\mu}} A'^{\nu} && (57)\\
&= \Lambda^{\nu}{}_{\tau} \frac{\partial}{\partial x'^{\mu}} A^{\tau} && (58)\\
&= \Lambda^{\nu}{}_{\tau} \left(\frac{\partial x^{\sigma}}{\partial x'^{\mu}} \right) \frac{\partial}{\partial x^{\sigma}} A^{\tau} && (59)\\
&= \Lambda^{\nu}{}_{\tau} \left(\tilde{\Lambda}_{\alpha}{}^{\sigma} \frac{\partial x'^{\alpha}}{\partial x'^{\mu}} \right) \frac{\partial}{\partial x^{\sigma}} A^{\tau} && (60)\\
&= \tilde{\Lambda}_{\mu}{}^{\sigma} \Lambda^{\nu}{}_{\tau} B(\sigma,\tau) && (61)
\end{aligned}
$$

It is helpful to walk through this argument step by step, to see how it follows. The first line is just the definition of B in the primed frame. The next, Equation 58 is just the application of the transformation rule for four-vectors to $\vec{A}'$; we can take the resulting $\Lambda^{\nu}{}_{\tau}$ term out of the derivative because Λ depends only on the boost between two frames, not on position. The line after that, Equation 59, is just the chain rule for derivatives. In the following line, Equation 60, we again used the transformation rule for four-vectors, to turn x^{σ} into x'^{α}, its equivalent in the primed frame. Finally, to obtain Equation 61, we used the fact that the different components of $\vec{x}'$ are all orthogonal to one another, so $\partial x'^{\alpha}/\partial x'\mu$ is 1 if α and μ are the same index (e.g., if this term reads $\partial y'/\partial y'$) and is zero if they are different indices (e.g., if this term reads $\partial z'/\partial y'$).

However, now notice that the transformation rule in Equation 61 is exactly what we already wrote down in Equation 55 for how mixed rank 2 tensors transform. Thus we learn that B does in fact transform like a tensor – a mixed covariant-contravariant tensor to be precise. Thus we learn that the gradient of a four-vector is a tensor, and by the same argument any gradient of any tensor is just another tensor of one higher rank. We can write this in shorthand form as

$$
\partial_{\mu} A^{\nu} \equiv \frac{\partial}{\partial x^{\mu}} A^{\nu}. \tag{62}
$$

In general we use the notation that ∂_{μ} means differentiation with respect to x^{μ}, while ∂^{μ} means differentiation with respect to x_{μ}. The partial derivative is also sometimes denoted with indices with commas, as

$$
\partial_{\mu} A^{\nu} \equiv A^{\nu}{}_{,\mu}. \tag{63}
$$

However, we will use the ∂_{μ} notation in these notes.

It is somewhat instructive to write out the partial derivatives in the more conventional notation we are used to. Recalling the definitions of x^{μ} and x_{μ}, for any quantity q

$$
\begin{aligned}
\partial_{\mu} q &= \left(\frac{1}{c}\frac{\partial q}{\partial t}, \nabla q \right) && (64)\\
\partial^{\mu} q &= \left(-\frac{1}{c}\frac{\partial q}{\partial t}, \nabla q \right) && (65)\\
\partial_{\mu}\partial^{\mu} q &= -\frac{1}{c^2}\frac{\partial^2 q}{\partial t^2} + \nabla^2 q = \Box^2 q. && (66)
\end{aligned}
$$

These operators should look familiar. In particular, note that the d'Alembertian operator we defined earlier is in fact just a covariant double-derivative.

III. Maxwell's Laws in covariant form

We have now developed the necessary mathematical machinery to recast Maxwell's laws in covariant form, meaning in a form where the equations are expressed solely in terms of tensors that obey the transformation laws we have just outlined. If we believe that relativity is a correct theory for physics, meaning that we think that the fundamental physical laws that describe electromagnetism are the same in all reference frames that are related to one another by Lorentz transformations, then this must be possible. The property of being invariant under Lorentz transformations is called covariance, and if relativity is a correct theory then all physical laws should be covariant.

The advantage of writing Maxwell's laws in terms of tensors is that the covariance becomes automatic and obvious, following simply from the transformation properties of the objects we use to write the equations. Equations that involve no operations besides arithmetic and partial derivatives of tensors are automatically covariant. Equations in this form are called manifestly covariant. A side-benefit of having such a manifestly-covariant formulation of Maxwell's Laws is that it becomes obvious how currents, fields, etc., transform between reference frames; we will make extensive use of this in our discussion of relativistic emitters.

To formulate these laws, let us begin with charge and current, the "source terms" in Maxwell's Laws. The equation for conservation of charge, which we wrote down at the beginning of our discussion of Maxwell's Laws, is

$$\frac{\partial \rho_e}{\partial t} + \nabla \cdot \mathbf{j} = 0. \tag{67}$$

Examining the definition of the operator ∂_μ (Equation 64), it is clear that if we define a four-vector for electric charge and current as

$$j^\mu = \begin{pmatrix} \rho_e c \\ \mathbf{j} \end{pmatrix}, \tag{68}$$

where $\mathbf{j}$ is the ordinary current density in three spatial dimensions, then the equation for conservation of charge can be written as simply

$$\partial_\mu j^\mu = 0, \tag{69}$$

which is manifestly covariant as desired.

Just to confirm to ourselves that this makes sense, let us verify that it does by considering a simple example. Suppose we have, in one reference frame, a box at rest that is size L on each side, and that contains a total charge q, so the charge density is $\rho_e = q/L^3$. Now consider what this looks like in another reference frame that is boosted by speed v along the x direction. Let us compute the current and charge density in this frame in two ways – one by using the transformation rules for the charge-current four-vector, one supposing we just knew the Lorentz transforms for coordinates, and not for charge or current. Using the first approach, the boost we are discussing can be written as

$$j'^\mu = \begin{pmatrix} \rho'_e c \\ j'_x \\ j'_y \\ j'_z \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \rho_e c \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{70}$$

so we learn that, in the primed frame, $\rho'_e = \gamma \rho_e$ and $j'_x = -\beta\gamma\rho_e c = -\gamma v \rho_e$. Now suppose that we didn't know this transformation, and just asked that the charge and current are in the primed frame. Since the box is stationary in the unprimed frame, in the primed frame it is seen to be moving at speed v in the $-x$ direction. The side of the box that lies along the x axis is contracted from size L to size L/γ, while the sides in the y and z directions are unchanged. Thus the volume of the box seen in the primed frame is L^3/γ, and the charge density is, assuming that charge is the same in all frames, $\rho_e' = \gamma q/L^3 = \gamma\rho_e$, i.e., exactly what we got by transforming the four-vector. Similarly, the current density in the x direction is $j_x' = -\rho_e' v = -\gamma v \rho_e$. Thus the four-vector definition of charge and current transforms between frames exactly was we would naively expect if we just know how to Lorentz transform coordinates, and if we though that charge should be the same in every frame.

Next consider the potential form of Maxwell's Laws. We had already written these (not by accident) in a form incredibly suggestive of manifest covariance:

$$\Box^2 \begin{pmatrix} \phi \\ \mathbf{A} \end{pmatrix} = -4\pi \begin{pmatrix} \rho_e \\ \mathbf{j}_e/c \end{pmatrix}. \tag{71}$$

This suggests that we identify

$$A^\mu = \begin{pmatrix} \phi \\ \mathbf{A} \end{pmatrix} \tag{72}$$

as a four-vector, known as the four-potential, and that we write Maxwell's Laws in potential form as simply

$$\partial_\nu \partial^\nu A^\mu = -\frac{4\pi}{c} j^\mu. \tag{73}$$

The Lorentz gauge, which is defined by

$$\frac{1}{c}\frac{\partial \phi}{\partial t} + \nabla \cdot \mathbf{A} = 0, \tag{74}$$

becomes

$$\partial_\mu A^\mu = 0. \tag{75}$$

Thus we have successfully written the potential form of Maxwell's Laws in covariant form.

What about the field version of Maxwell's Laws? To determine this, we just need to recall how the fields are related to the potentials:

$$\begin{aligned} \mathbf{E} &= -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} - \nabla\phi & (76) \\ \mathbf{B} &= \nabla \times \mathbf{A}. & (77) \end{aligned}$$

To see how to represent these in terms of tensors, it is helpful to write out the x components explicitly:

$$\begin{aligned} E_x &= -\frac{1}{c}\frac{\partial A_x}{\partial t} - \frac{\partial \phi}{\partial x} = -\left(\partial^0 A^1 - \partial^1 A^0\right) & (78) \\ B_x &= \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} = -\left(\partial^2 A^3 - \partial^3 A^2\right). & (79) \end{aligned}$$

This suggests that $\mathbf{E}$ and $\mathbf{B}$ should come from a tensor of rank 2, since there are two indices involved in the right hand sides we wrote down, and that it should be anti-symmetric, meaning that transposing the indices (e.g., changing T^{01} to T^{10}) introduces a minus sign – in matrix language, elements above and below the diagonal should have opposite signs.

With a bit of experimentation one can quickly work out that the following tensor fits the bill:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix}. \tag{80}$$

The quantity $F_{\mu\nu}$ is called the field strength tensor. We can now write Maxwell's Laws in terms of the field strength tensor. With a very little bit of algebra, one can show that the two equations

$$\nabla \cdot \mathbf{B} = 0 \tag{81}$$

$$\nabla \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = 0 \tag{82}$$

are equivalent to

$$\partial_\sigma F_{\mu\nu} + \partial_\mu F_{\nu\sigma} + \partial_\nu F_{\sigma\mu} = 0. \tag{83}$$

This is the covariant version of Maxwell's Laws in field form.