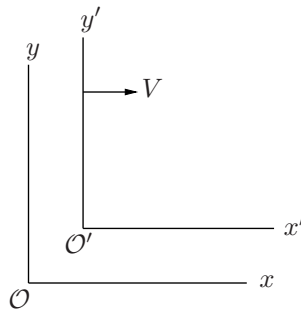


Supplementary Notes on Special Relativity and Maxwell's Equations

1 The Lorentz Transformation

This is a derivation of the Lorentz transformation of Special Relativity. The basic idea is to derive a relationship between the spacetime coordinates x, y, z, t as seen by observer \mathcal{O} and the coordinates x', y', z', t' seen by observer \mathcal{O}' moving at a velocity V with respect to \mathcal{O} along the positive x axis.



These observers are assumed to be *inertial*. In other words, they are moving at a constant velocity with respect to each other and in the absence of any external forces or accelerations (which is somewhat redundant). In particular, there is no rotational motion or gravitational field present.

Our derivation is based on two assumptions:

1. The Principle of Relativity: Physics is the same for all observers in all inertial coordinate systems.
2. The speed of light c in a vacuum is the same for all observers independently of their relative motion or the motion of the light source.

We first show that this transformation must be of the form

$$t' = at + bx \tag{1.1a}$$

$$x' = dt + ex \tag{1.1b}$$

$$y' = y$$

$$z' = z$$

where we assume that the origins coincide at $t = t' = 0$. The above figure shows the coordinate systems displaced simply for ease of visualization.

The first thing to note is that the y and z coordinates are the same for both observers. (This is only true in this case because the relative motion is along the x axis only. If the motion were in an arbitrary direction, then each spatial coordinate of \mathcal{O}' would depend on all of the spatial coordinates of \mathcal{O} . However, this is the case that is used in almost all situations, at least at an elementary level.) To see that this is necessary, suppose there is a yardstick at the origin of each coordinate system aligned along each of the y - and y' -axes, and suppose there is a paintbrush at the end of each yardstick pointed towards the other. If \mathcal{O}' 's yardstick along the y' -axis gets shorter as seen by \mathcal{O} , then when the origins pass each other \mathcal{O} 's yardstick will get paint on it. But by the Principle of Relativity, \mathcal{O}' should also see \mathcal{O} 's yardstick get shorter and hence \mathcal{O}' would get paint on his yardstick. Since this clearly can't happen, there can be no change in a direction perpendicular to the direction of motion.

The next thing to notice is that the transformation equations are *linear*. This is a result of space being homogeneous. To put this very loosely, "things here are the same as things there." For example, if there is a yardstick lying along the x axis between $x = 1$ and $x = 2$, then the length of this yardstick as seen by \mathcal{O}' should be the same as another yardstick lying between $x = 2$ and $x = 3$. But if there were a nonlinear dependence, say $\Delta x'$ goes like Δx^2 , then the first yardstick would have a length that goes like $2^2 - 1^2 = 3$ while the second would have a length that goes like $3^2 - 2^2 = 5$. Since this is also not the way the world works, equations (1.1) must be linear as shown. We now want to figure out what the coefficients a, b, d and e must be.

First, let \mathcal{O} look at the origin of \mathcal{O}' (i.e., $x' = 0$). Since \mathcal{O}' is moving at a speed V along the x -axis, the x coordinate corresponding to $x' = 0$ is $x = Vt$. Using this in (1.1b) yields $x' = 0 = dt + eVt$ or $d = -eV$. Similarly, \mathcal{O}' looks at \mathcal{O} (i.e., $x = 0$) and it has the coordinate $x' = -Vt'$ with respect to \mathcal{O}' (since \mathcal{O} moves in the negative x' direction as seen by \mathcal{O}'). Then from (1.1b) we have $-Vt' = dt$ and from (1.1a) we have $t' = at$, and hence $t' = -dt/V = at$ so that $-aV = d = -eV$ and thus also $a = e$ and $d = -aV$. Using these results in equations (1.1) now gives us

$$t' = a \left(t + \frac{b}{a} x \right) \quad (1.2a)$$

$$x' = a(x - Vt) \quad (1.2b)$$

$$y' = y$$

$$z' = z.$$

Now let a photon move along the x -axis (and hence also along the x' -axis) and pass both origins when they coincide at $t = t' = 0$. Then the x coordinate of the photon as seen by \mathcal{O} is $x = ct$, and the x' coordinate as seen by \mathcal{O}' is $x' = ct'$. Note that the value of c is the same for both observers. This is assumption (2). Using these in equations (1.2) yields

$$t' = a \left(t + \frac{b}{a} ct \right) = at \left(1 + \frac{bc}{a} \right)$$

$$x' = a(ct - Vt) = cat \left(1 - \frac{V}{c}\right)$$

so that

$$cat \left(1 - \frac{V}{c}\right) = x' = ct' = cat \left(1 + \frac{bc}{a}\right)$$

and therefore $-V/c = bc/a$ or

$$\frac{b}{a} = -\frac{V}{c^2}.$$

So now equations (1.2) become

$$t' = a \left(t - \frac{V}{c^2} x \right) \tag{1.3a}$$

$$x' = a(x - Vt) \tag{1.3b}$$

$$y' = y$$

$$z' = z.$$

We still need to determine a . To do this, we will again use the Principle of Relativity. Let \mathcal{O} look at a clock situated at \mathcal{O}' . Then $\Delta x = V\Delta t$ and from (1.3a), \mathcal{O} and \mathcal{O}' will measure time intervals related by

$$\Delta t' = a \left(\Delta t - \frac{V}{c^2} \Delta x \right) = a \left(1 - \frac{V^2}{c^2} \right) \Delta t.$$

Now let \mathcal{O}' look at a clock at \mathcal{O} (so $\Delta x = 0$). Then $\Delta x' = -V\Delta t'$ so (1.3b) yields

$$-V\Delta t' = \Delta x' = a(0 - V\Delta t) = -aV\Delta t$$

and hence

$$\Delta t = \frac{1}{a} \Delta t'.$$

By the Principle of Relativity, the relative factors in the time measurements must be the same in both cases. In other words, \mathcal{O} sees \mathcal{O}' 's time related to his by the factor $a(1 - V^2/c^2)$, and \mathcal{O}' sees \mathcal{O} 's time related to his by the factor $1/a$. This means that

$$\frac{1}{a} = a \left(1 - \frac{V^2}{c^2} \right)$$

or

$$a = \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}}$$

and therefore the final Lorentz transformation equations are

$$\begin{aligned}
 t' &= \frac{t - \frac{V}{c^2}x}{\sqrt{1 - \frac{V^2}{c^2}}} \\
 x' &= \frac{x - Vt}{\sqrt{1 - \frac{V^2}{c^2}}} \\
 y' &= y \\
 z' &= z.
 \end{aligned} \tag{1.4}$$

It is very common to define the dimensionless variables

$$\beta = \frac{V}{c} \quad \text{and} \quad \gamma = \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}} = \frac{1}{\sqrt{1 - \beta^2}}. \tag{1.5}$$

In terms of these variables, equations (1.4) become

$$\begin{aligned}
 t' &= \gamma \left(t - \frac{\beta}{c} x \right) \\
 x' &= \gamma (x - \beta ct) \\
 y' &= y \\
 z' &= z.
 \end{aligned} \tag{1.6}$$

Since c is a universal constant, it is essentially a conversion factor between units of time and units of length. Because of this, we may further change to units where $c = 1$ (so time is measured in units of length) and in this case the Lorentz transformation equations become

$$\begin{aligned}
 t' &= \gamma(t - \beta x) \\
 x' &= \gamma(x - \beta t) \\
 y' &= y \\
 z' &= z.
 \end{aligned} \tag{1.7a}$$

These equations give the coordinates as seen by \mathcal{O}' in terms of those of \mathcal{O} . If we want the coordinates as seen by \mathcal{O} in terms of those of \mathcal{O}' , then we let $\beta \rightarrow -\beta$ and we have

$$\begin{aligned}
 t &= \gamma(t' + \beta x') \\
 x &= \gamma(x' + \beta t') \\
 y &= y' \\
 z &= z'.
 \end{aligned} \tag{1.7b}$$

Note that $0 \leq \beta \leq 1$ so that $1 \leq \gamma < \infty$. We also see that

$$\gamma^2 = \frac{1}{1 - \beta^2}$$

so that

$$\gamma^2 - \gamma^2 \beta^2 = 1.$$

Then recalling the hyperbolic trigonometric identities

$$\cosh^2 \theta - \sinh^2 \theta = 1$$

and

$$1 - \tanh^2 \theta = \operatorname{sech}^2 \theta$$

we may define a parameter θ (sometimes called the *rapidity*) by

$$\beta = \tanh \theta$$

so that

$$\gamma = \cosh \theta$$

and

$$\gamma\beta = \sinh \theta.$$

In terms of θ , equations (1.7a) become

$$\begin{aligned} t' &= (\cosh \theta)t - (\sinh \theta)x \\ x' &= -(\sinh \theta)t + (\cosh \theta)x \\ y' &= y \\ z' &= z \end{aligned} \tag{1.8}$$

which looks very similar to a rotation in the xt -plane, except that now we have hyperbolic functions instead of the usual trigonometric ones. However, note that both \sinh terms have the same sign.

Next, let us consider motion as seen by both observers. In this case we write displacements in both space and time as

$$\begin{aligned} dt' &= \gamma(dt - \beta dx) \\ dx' &= \gamma(dx - \beta dt) \\ dy' &= dy \\ dz' &= dz. \end{aligned}$$

Then the velocity v'_x of a particle along the x' -axis as seen by \mathcal{O}' is

$$v'_x = \frac{dx'}{dt'} = \frac{dx - \beta dt}{dt - \beta dx} = \frac{dx/dt - \beta}{1 - \beta(dx/dt)} = \frac{v_x - \beta}{1 - \beta v_x}. \tag{1.9a}$$

Alternatively, we may write

$$v_x = \frac{dx}{dt} = \frac{dx' + \beta dt'}{dt' + \beta dx'} = \frac{dx'/dt' + \beta}{1 + \beta(dx'/dt')} = \frac{v'_x + \beta}{1 + \beta v'_x}. \tag{1.9b}$$

These last two equations are called the **relativistic velocity addition law**. It should be obvious that for motion along the x -axis we have $v'_y = v_y$ and $v'_z = v_z$.

Be sure to understand what these equations mean. They relate the velocity of an object as seen by two different observers whose relative velocity along their common x -axis is β . Note that even if $v'_x = 1$ (corresponding to $v'_x = c$), the velocity as seen by \mathcal{O} is still just $v_x = 1$. This is quite different from the classical Galilean addition of velocities, because nothing can go faster than light ($c = 1$).

One of the most important aspects of Lorentz transformations is that they leave the quantity $t^2 - x^2 - y^2 - z^2$ invariant. In other words, using equations (1.7a) you can easily show that

$$t'^2 - x'^2 - y'^2 - z'^2 = t^2 - x^2 - y^2 - z^2. \quad (1.10)$$

Note that setting this equal to zero, we get the equation of an outgoing sphere of light as seen by either observer. (Don't forget that if $c \neq 1$, then t becomes ct .) We refer to this as the **invariance of the interval** because it can be written as

$$(\Delta t')^2 - (\Delta x')^2 - (\Delta y')^2 - (\Delta z')^2 = (\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2.$$

If the primed frame is the *rest frame* of a particle, then we have $dx' = dy' = dz' = 0$ and dt' measures the time interval as seen by the particle, called the **proper time**. Because of this, we sometimes write

$$d\tau^2 := dt^2 - dx^2 - dy^2 - dz^2.$$

This is frequently also called the **proper distance** (or **proper length**) and written ds^2 . The difference between $d\tau$ and ds is if $c \neq 1$, we have

$$ds^2 = c^2 dt^2 - d\mathbf{x}^2 = c^2 dt^2 (1 - \mathbf{v}^2/c^2) = c^2 dt^2 / \gamma^2 := c^2 d\tau^2$$

so that $d\tau^2 = ds^2/c^2$. Since we are working with $c = 1$, we will write proper distance as

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2.$$

Now let's go to some modern notation. In units with $c = 1$, we first define our four spacetime components as the vector

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

(If $c \neq 1$ then $x^0 = ct$.) This vector is an element of a 4-dimensional vector space called **Minkowski space**. Then we have

$$ds^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2$$

or, defining the **Lorentz** (or **Minkowski**) **metric**

$$g_{\mu\nu} = \begin{bmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{bmatrix} \quad (1.11)$$

we write (using the summation convention)

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu. \quad (1.12)$$

Let me note that most particle physicists use this metric, which we can also write as simply $g_{\mu\nu} = \text{diag}(1, -1, -1, -1)$, but most relativists use the metric $g_{\mu\nu} = \text{diag}(-1, 1, 1, 1)$, and you need to be careful when reading equations. Many physicists also use the symbol $\eta_{\mu\nu}$ rather than $g_{\mu\nu}$ when dealing with the Lorentz metric (as opposed to more general metrics used in general relativity).

Vectors x^μ in Minkowski space are classified as **timelike** if $x_\mu x^\mu > 0$, **space-like** if $x_\mu x^\mu < 0$ and **null** (or **lightlike**) if $x_\mu x^\mu = 0$. Light rays are null, and hence we see that there are nonzero vectors with zero norm. Because of this, the Minkowski metric is not positive definite, and we say that Minkowski space is **semi-Riemannian**.

From linear algebra we know that the metric defines an inner product, and we can use this to raise or lower indices, for example, $x_\mu = g_{\mu\nu} x^\nu$. In the case of the Lorentz metric, we have the inverse metric with components $g^{\mu\nu} = g_{\mu\nu}$. Furthermore, there is no difference between x^0 and x_0 , but $x_i = -x^i$ for $i = 1, \dots, 3$. (Again, be careful because this is the opposite of what you get if you use the other metric.)

Using this notation, equation (1.10) is written

$$g_{\mu\nu} x'^\mu x'^\nu = g_{\mu\nu} x^\mu x^\nu$$

where the metric is the same in both frames. Lowering indices, we write this in its most compact form as

$$x'_\mu x'^\mu = x_\mu x^\mu$$

and we say that the length $x^2 = x_\mu x^\mu$ is an **invariant**. It is also important to understand that the scalar product of two vectors a^μ and b^μ is written in the equivalent forms

$$a \cdot b = g_{\mu\nu} a^\mu b^\nu = a_\nu b^\nu = a_0 b^0 + a_i b^i = a_0 b^0 - \sum_{i=1}^3 a^i b^i = a_0 b^0 - \mathbf{a} \cdot \mathbf{b}.$$

Note that the summation convention means that repeated Greek indices are to be summed from 0 to 3, and repeated Latin indices are to be summed from 1 to 3.

We now write our Lorentz transformation equations as

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu \quad (1.13)$$

where we have defined the **Lorentz transformation matrix**

$$\Lambda^\mu{}_\nu = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (1.14)$$

(You should be aware that some authors put the prime on the indices and write this in the form $x^{\mu'} = \Lambda^{\mu'}_{\nu} x^{\nu}$.) Using this, we write the invariant x^2 as $x'^{\mu} x'^{\mu} = \Lambda_{\mu\alpha} \Lambda^{\mu}_{\beta} x^{\alpha} x^{\beta}$. But this must equal $x^{\alpha} x_{\alpha}$, and hence we have

$$\Lambda_{\mu\alpha} \Lambda^{\mu}_{\beta} = (\Lambda^T)_{\alpha\mu} \Lambda^{\mu}_{\beta} = (\Lambda^T)_{\alpha}^{\nu} g_{\nu\mu} \Lambda^{\mu}_{\beta} = g_{\alpha\beta} \quad (1.15)$$

which can be written in matrix form as

$$\Lambda^T g \Lambda = g.$$

In fact, this can be taken as the *definition* of a Lorentz transformation Λ . Since $g_{\beta}^{\alpha} = \delta_{\beta}^{\alpha}$, we can write equation (1.15) as $(\Lambda^T)_{\alpha}^{\mu} \Lambda^{\mu}_{\beta} = \delta_{\beta}^{\alpha}$ which shows that Λ is an orthogonal transformation, i.e., $\Lambda^T = \Lambda^{-1}$. This is actually just what equations (1.8) say—a Lorentz transformation is a rotation in Minkowski space.

Since $(\Lambda^{-1})^{\mu}_{\nu} = (\Lambda^T)^{\mu}_{\nu} = \Lambda_{\nu}^{\mu}$, we see from equation (1.13) that

$$(\Lambda^{-1})^{\alpha}_{\mu} x'^{\mu} = (\Lambda^{-1})^{\alpha}_{\mu} \Lambda^{\mu}_{\nu} x^{\nu} = x^{\alpha}$$

or

$$x^{\alpha} = \Lambda_{\mu}^{\alpha} x'^{\mu}. \quad (1.16)$$

Equations (1.13) and (1.16) then give us the very useful results

$$\Lambda^{\mu}_{\nu} = \frac{\partial x'^{\mu}}{\partial x^{\nu}} \quad \text{and} \quad \Lambda_{\mu}^{\nu} = (\Lambda^{-1})^{\nu}_{\mu} = \frac{\partial x^{\nu}}{\partial x'^{\mu}}. \quad (1.17)$$

In order to define velocity in an invariant manner, we define the **4-velocity** in terms of the proper time by

$$u^{\mu} := \frac{dx^{\mu}}{d\tau}. \quad (1.18)$$

Note we can write $d\tau^2 = dt^2 - d\mathbf{x}^2 = dt^2(1 - \mathbf{v}^2)$. Here \mathbf{v} is the velocity of a particle as seen by \mathcal{O} . If we let \mathcal{O}' be the particle rest frame, then \mathbf{v} is just β and we have $d\tau^2 = dt^2(1 - \mathbf{v}^2) = dt^2/\gamma^2$ so that

$$\frac{dt}{d\tau} = \gamma.$$

Then

$$u^{\mu} = \frac{dx^{\mu}}{d\tau} = \frac{dt}{d\tau} \frac{dx^{\mu}}{dt} = \gamma \frac{dx^{\mu}}{dt}$$

and we can write the 4-velocity as the vector

$$u^{\mu} = \begin{bmatrix} \gamma \\ \gamma \mathbf{v} \end{bmatrix} \quad (1.19)$$

which has the magnitude

$$u_{\mu} u^{\mu} = \gamma^2 - \gamma^2 \mathbf{v}^2 = \gamma^2(1 - \mathbf{v}^2) = 1.$$

(Again, with $c \neq 1$ we have $x^0 = ct$ so that $u^\mu = \gamma[c, \mathbf{v}]$ and $u_\mu u^\mu = c^2$.)
 Since $\Lambda^\mu{}_\nu$ is a constant matrix, we have

$$u'^\mu = \frac{dx'^\mu}{d\tau} = \Lambda^\mu{}_\nu \frac{dx^\nu}{d\tau} = \Lambda^\mu{}_\nu u^\nu$$

so that u^μ transforms in exactly the same manner as x^μ . We call any vector that transforms in this way a **4-vector**, which justifies the term ‘4-velocity’ used above. Similarly, we define the **4-momentum** by

$$p^\mu := mu^\mu = m\gamma \begin{bmatrix} 1 \\ \mathbf{v} \end{bmatrix} \quad (1.20)$$

so that

$$p^2 = p^\mu p_\mu = m^2.$$

(If $c \neq 1$ then $p^2 = m^2 c^2$. Let me also emphasize that the mass m in all of our equations is the constant *rest mass*. We never talk about a “relativistic mass” γm that many older books use where they write our mass as m_0 and then define $m = \gamma m_0$.)

Expanding the square root we have

$$p^0 = m\gamma = \frac{m}{\sqrt{1-v^2}} = m \left(1 + \frac{1}{2}v^2 + \dots \right)$$

which is the sum of a rest energy term m ($= mc$) and a kinetic energy $mv^2/2$ ($= (1/2)mv^2/c$) plus higher order terms in v ($= v/c$). Because of this, we see that p^0 is the total energy $p^0 = m\gamma = E$ ($= E/c$) of the particle, so using $m\gamma\mathbf{v} = \mathbf{p}$ as the classical momentum, we have

$$p^\mu = \begin{bmatrix} E \\ \mathbf{p} \end{bmatrix}. \quad (1.21)$$

Therefore $m^2 = p^2 = E^2 - \mathbf{p}^2$ or

$$E^2 = \mathbf{p}^2 + m^2. \quad (1.22)$$

(If $c \neq 1$, then $p^\mu = m\gamma[c, \mathbf{v}] = [E/c, \mathbf{p}]$ and this becomes $E^2 = \mathbf{p}^2 c^2 + m^2 c^4$.)

Now, the gradient operator is *defined* as $\nabla = \partial/\partial\mathbf{x}$ so that

$$\nabla^i = \partial/\partial x^i := \partial_i.$$

Let us define $\partial_\mu = \partial/\partial x^\mu$. Then

$$\partial_\mu = \begin{bmatrix} \partial_0 \\ \nabla \end{bmatrix} \quad \text{and} \quad \partial^\mu = \begin{bmatrix} \partial^0 \\ -\nabla \end{bmatrix}. \quad (1.23)$$

Using equation (1.17) we have

$$\partial'_\mu = \frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial}{\partial x^\nu} = \Lambda_\mu{}^\nu \partial_\nu$$

or, equivalently,

$$\partial'^{\mu} = \Lambda^{\mu}_{\nu} \partial^{\nu}$$

which shows that indeed ∂_{μ} transforms as a 4-vector (which is implied by the notation). The operator

$$\partial_{\mu} \partial^{\mu} = (\partial_0)^2 + \partial_i \partial^i = (\partial_t)^2 - \nabla^2 = \frac{\partial^2}{\partial t^2} - \nabla^2$$

is called the **d'Alembertian**, and is frequently written as \square .

In quantum mechanics we have the momentum operator defined by $\mathbf{p} = -i\hbar\nabla$ and the energy operator defined by $E = i\hbar(\partial/\partial t)$. Then $p^i = -i\hbar\nabla^i = -i\hbar\partial_i = +i\hbar\partial^i$ and we can define the relativistic momentum operator

$$p^{\mu} = i\hbar\partial^{\mu}.$$

Using units with $\hbar = 1$, the expression $E^2 - \mathbf{p}^2 - m^2 = 0$ becomes $-\partial_t^2 + \nabla^2 - m^2 = 0$ or

$$(\partial_t^2 - \nabla^2 + m^2)\phi(x) = (\square + m^2)\phi(x) = 0$$

which is known as the **Klein-Gordon equation**.

Even though the two reference frames relative to which we are describing motion must be inertial, there is no reason we can't describe the motion of an accelerated object. As you might guess, we define the **4-acceleration** of an object by

$$a^{\mu} = \frac{du^{\mu}}{d\tau}.$$

Since $u^{\mu}u_{\mu} = 1$, it follows that the 4-acceleration is always orthogonal to the 4-velocity because

$$0 = \frac{d}{d\tau} u^{\mu}u_{\mu} = 2u_{\mu} \frac{du^{\mu}}{d\tau} = 2u_{\mu} a^{\mu}.$$

We also define the **4-force**

$$f^{\mu} = \frac{dp^{\mu}}{d\tau} = ma^{\mu}$$

so that

$$f^{\mu} = \frac{dp^{\mu}}{d\tau} = \gamma \frac{dp^{\mu}}{dt} = \gamma \begin{bmatrix} d(\gamma m)/dt \\ d\mathbf{p}/dt \end{bmatrix} = \begin{bmatrix} f^0 \\ \gamma \mathbf{f}_c \end{bmatrix} \quad (1.24)$$

where $\mathbf{f}_c = d\mathbf{p}/dt$ is the classical force on the particle. Since the 4-force obviously obeys $f^{\mu}u_{\mu} = 0$, we have

$$0 = f^{\mu}u_{\mu} = \gamma f^0 - \gamma^2 \mathbf{f}_c \cdot \mathbf{v}$$

and therefore

$$f^0 = \gamma \mathbf{f}_c \cdot \mathbf{v} \quad (1.25)$$

which, to within the factor of γ , is just the classical power (i.e., the rate at which work is done). (And if $c \neq 1$ we have $0 = f^{\mu}u_{\mu} = f^0\gamma c - \gamma^2 \mathbf{f}_c \cdot \mathbf{v}$ so that $f^0 = (\gamma/c)\mathbf{f}_c \cdot \mathbf{v}$.)

2 Maxwell's Equations

Experimentally, it is found that the charge to mass ratio $e/m\gamma$ of a particle moving at velocity β obeys the law

$$\frac{e}{m\gamma} = \frac{e}{m}(1 - \beta^2)^{1/2}.$$

(The two sides of this equation refer to different measurements, so it's not as trivial of a statement as it looks at first.) Therefore, e is a constant, and we have that charge is an invariant quantity. What we would now like to know is how charge *density* and electric current behave. Since charge density is charge/volume, we must find out how volumes transform.

Let frame 2 (the primed frame) be in motion with respect to frame 1 (the unprimed frame) along their mutual x -axis, and consider a small cube of side l_0 at rest in frame 2. In its rest frame, this cube has volume $d\tau_0$ (not to be confused with proper time). From (1.7a) we have

$$d\tau_1 = dx dy dz = \frac{1}{\gamma} dx' dy' dz' = \frac{1}{\gamma} d\tau_0 \quad (2.1)$$

where we used $dx' = \gamma(dx - \beta dt)$ together with $dt = 0$ for a measurement made in frame 1. Thus we have

$$d\tau_1 = d\tau_0(1 - \beta^2)^{1/2}.$$

Now suppose the volume is also moving with respect to frame 2, and let this motion be along the x_2 axis. Letting v_{2x} be the velocity of the box with respect to frame 2, and similarly for v_{1x} , we have

$$d\tau_2 = d\tau_0(1 - v_{2x}^2)^{1/2} \quad \text{and} \quad d\tau_1 = d\tau_0(1 - v_{1x}^2)^{1/2}.$$

But from (1.9b) we have

$$v_{1x} = \frac{v_{2x} + \beta}{1 + \beta v_{2x}}$$

and therefore

$$\begin{aligned} d\tau_1 &= d\tau_0 \left[1 - \left(\frac{v_{2x} + \beta}{1 + \beta v_{2x}} \right)^2 \right]^{1/2} \\ &= d\tau_0 \left[\frac{1 + 2\beta v_{2x} + \beta^2 v_{2x}^2 - v_{2x}^2 - 2\beta v_{2x} - \beta^2}{(1 + \beta v_{2x})^2} \right]^{1/2} \\ &= d\tau_0 \left[\frac{(1 - \beta^2)(1 - v_{2x}^2)}{(1 + \beta v_{2x})^2} \right]^{1/2} = d\tau_0 \frac{[(1 - \beta^2)(1 - v_{2x}^2)]^{1/2}}{1 + \beta v_{2x}} \\ &= \frac{(1 - \beta^2)^{1/2}}{1 + \beta v_{2x}} d\tau_2 \end{aligned}$$

where in going to the last line we used $d\tau_2 = d\tau_0(1 - v_{2x}^2)^{1/2}$. Rearranging, this is

$$d\tau_1 = \frac{d\tau_2}{\gamma(1 + \beta v_{2x})}. \quad (2.2a)$$

Reversing the frame point of view, we clearly also have

$$d\tau_2 = \frac{d\tau_1}{\gamma(1 - \beta v_{1x})}. \quad (2.2b)$$

Be sure to understand what these equations say. The velocities v_{1x} and v_{2x} are the observed velocities of a box with rest volume $d\tau_0$ moving along the common x -axis as seen in frames 1 and 2, which are moving with velocity β with respect to each other.

Now suppose that we have dn charges of Q coulombs each. As we stated above, Q is an invariant. Obviously, dn is also an invariant since it is just the number of charges. Then the charge densities as observed in frames 1 and 2 are

$$\rho_1 = \frac{Q dn}{d\tau_1} \quad \text{and} \quad \rho_2 = \frac{Q dn}{d\tau_2}$$

so that

$$\rho_1 d\tau_1 = \rho_2 d\tau_2$$

where $d\tau_1$ and $d\tau_2$ are the same volume containing the charge $Q dn$ as observed in the two different reference frames. Then, using equations 2.2, we have

$$\begin{aligned} \rho_1 &= \rho_2 \frac{d\tau_2}{d\tau_1} = \rho_2 \gamma(1 + \beta v_{2x}) \\ \rho_2 &= \rho_1 \frac{d\tau_1}{d\tau_2} = \rho_1 \gamma(1 - \beta v_{1x}). \end{aligned} \quad (2.3)$$

The **current density $\mathbf{J} := \rho\mathbf{v}$** is defined as the charge per area-time, so we can write $J_{1x} = \rho_1 v_{1x}$ and $J_{2x} = \rho_2 v_{2x}$. Using these definitions in equations (2.3) yields the **transformation of charge density**

$$\begin{aligned} \rho_1 &= \gamma(\rho_2 + \beta J_{2x}) \\ \rho_2 &= \gamma(\rho_1 - \beta J_{1x}). \end{aligned} \quad (2.4)$$

Now using equations (1.9) and (2.3) we also have

$$J_{1x} = \rho_1 v_{1x} = \rho_2 \gamma(1 + \beta v_{2x}) \frac{v_{2x} + \beta}{1 + \beta v_{2x}}$$

or

$$\begin{aligned} J_{1x} &= \gamma(J_{2x} + \beta \rho_2) \\ J_{2x} &= \gamma(J_{1x} - \beta \rho_1). \end{aligned} \quad (2.5a)$$

Similarly,

$$J_{1y} = \rho_1 v_{1y} = \rho_2 \gamma(1 + \beta v_{2x}) v_{1y}.$$

But

$$v_{1y} = \frac{dy_1}{dt_1} = \frac{dy_2}{\gamma(dt_2 + \beta dx_2)} = \frac{v_{2y}}{\gamma(1 + \beta v_{2x})}.$$

Hence

$$J_{1y} = \rho_2 v_{2y} = J_{2y} \quad \text{and similarly} \quad J_{1z} = J_{2z}. \quad (2.5b)$$

Comparing equations (2.4) and (2.5) with equations (1.7), we see that we have a **4-current density**

$$J^\mu = \begin{bmatrix} \rho \\ \mathbf{J} \end{bmatrix}. \quad (2.6)$$

(And again, if $c \neq 1$ this becomes $J^\mu = [\rho c, \mathbf{J}]$.)

Now that we have shown that J^μ does indeed define a 4-vector, let me show another way to arrive at this conclusion that is analogous to the definition of 4-momentum. From (2.1) we can write $d\tau = d\tau_0/\gamma$ where $d\tau_0$ is at rest in frame 2, and $\gamma = (1 - \beta^2)^{-1/2}$ where β is the velocity of frame 2 with respect to frame 1. Then the invariance of charge implies $\rho d\tau = \rho_0 d\tau_0$ or $\rho d\tau_0/\gamma = \rho_0 d\tau_0$, and hence

$$\rho = \gamma \rho_0.$$

This is analogous to the expression $m_{\text{relativistic}} = \gamma m_{\text{rest}}$ or simply $m = \gamma m_0$.

Now we also have $\mathbf{J} = \rho \mathbf{v} = \rho_0 \gamma \mathbf{v}$. Recalling that the 4-velocity is given by

$$u^\mu = \begin{bmatrix} \gamma \\ \gamma \mathbf{v} \end{bmatrix}$$

we see that letting \mathbf{v} be the velocity of the charge, i.e., $\mathbf{v} = \beta$, then (2.6) is the same as

$$J^\mu = \rho_0 u^\mu$$

which is analogous to $p^\mu = m u^\mu$. In other words, we have shown

$$J^\mu = \begin{bmatrix} \rho \\ \mathbf{J} \end{bmatrix} = \rho_0 \begin{bmatrix} \gamma \\ \gamma \mathbf{v} \end{bmatrix} = \rho_0 u^\mu. \quad (2.7)$$

Since we saw earlier that the derivative operator ∂_μ transforms as a 4-vector (technically, a co-vector), and we just showed that J^μ is a 4-vector, it follows that $\partial_\mu J^\mu$ is a Lorentz scalar. But

$$\partial_\mu J^\mu = \partial_0 \rho + \partial_i J^i = \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J}$$

and therefore the continuity equation may be written in the covariant form

$$\partial_\mu J^\mu = 0. \quad (2.8)$$

Now recall Maxwell's equations:

$$\nabla \cdot \mathbf{E} = 4\pi\rho \quad (2.9a)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (2.9b)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (2.9c)$$

$$\nabla \times \mathbf{B} = 4\pi\mathbf{J} + \frac{\partial \mathbf{E}}{\partial t} \quad (2.9d)$$

Using $\mathbf{B} = \nabla \times \mathbf{A}$, equation (2.9c) implies

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}$$

so that equations (2.9a) and (2.9d) then imply (using the identity $\nabla \times \nabla \times \mathbf{A} = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}$)

$$\nabla^2 \phi + \frac{\partial}{\partial t}(\nabla \cdot \mathbf{A}) = -4\pi\rho \quad (2.10a)$$

$$\nabla^2 \mathbf{A} - \frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla \left(\nabla \cdot \mathbf{A} + \frac{\partial \phi}{\partial t} \right) = -4\pi\mathbf{J}. \quad (2.10b)$$

The **gauge transformation** $\mathbf{A} \rightarrow \mathbf{A}' = \mathbf{A} + \nabla\Lambda$ leaves the physical field $\mathbf{B} = \nabla \times \mathbf{A}$ unchanged, so if $\mathbf{E} = -\nabla\phi - \partial\mathbf{A}/\partial t$ is also to remain unchanged, we must have $\phi \rightarrow \phi' = \phi - \partial\Lambda/\partial t$. This gives us the freedom to choose (ϕ, \mathbf{A}) such as to satisfy the **Lorentz gauge** (or **Lorentz condition**)

$$\nabla \cdot \mathbf{A} + \frac{\partial \phi}{\partial t} = 0.$$

In other words, we *demand* that the new potentials (ϕ', \mathbf{A}') satisfy

$$\begin{aligned} 0 &= \nabla \cdot \mathbf{A}' + \frac{\partial \phi'}{\partial t} \\ &= \nabla \cdot \mathbf{A} + \nabla^2 \Lambda + \frac{\partial \phi}{\partial t} - \frac{\partial^2 \Lambda}{\partial t^2}. \end{aligned}$$

Thus, if we can find a Λ that satisfies

$$\nabla^2 \Lambda - \frac{\partial^2 \Lambda}{\partial t^2} = -\left(\nabla \cdot \mathbf{A} + \frac{\partial \phi}{\partial t} \right)$$

the gauge transformed fields will satisfy the Lorentz condition. Fortunately, this is a straightforward problem to solve. All we need to do is find the Green's function for the wave equation, and then Λ will be the integral of the Green's function times the quantity on the right. (Very briefly, if you have a linear operator $L(x)$ acting on a function $f(x)$ such that $L(x)f(x) = g(x)$, and if you find a Green's function $G(x, x')$ defined by $L(x)G(x, x') = \delta(x - x')$, then the

solution to the problem is essentially $f(x) = \int G(x, x')g(x') dx'$. You can easily verify that acting on this with $L(x)$ will yield $L(x)f(x) = g(x)$.

In any case, choosing the Lorentz gauge, equations (2.10) become the well-known wave equations

$$\nabla^2 \phi - \frac{\partial^2 \phi}{\partial t^2} = -4\pi\rho \quad (2.11a)$$

$$\nabla^2 \mathbf{A} - \frac{\partial^2 \mathbf{A}}{\partial t^2} = -4\pi\mathbf{J}. \quad (2.11b)$$

Note that an equivalent way of writing these is

$$\square \phi = -4\pi J^0$$

$$\square \mathbf{A} = -4\pi\mathbf{J}.$$

So if we define the **4-potential** A^μ by

$$A^\mu = \begin{bmatrix} \phi \\ \mathbf{A} \end{bmatrix}$$

then the wave equations may be written in the concise form

$$\square A^\mu = -4\pi J^\mu \quad (2.12)$$

where the Lorentz condition becomes

$$\partial_\mu A^\mu = 0.$$

That A^μ is indeed a 4-vector follows because \square is a Lorentz invariant quantity and J^μ is a 4-vector. Thus A^μ must transform as a 4-vector so that (2.12) is covariant (i.e., so that both sides transform the same way). Also, A^μ is unchanged even if $c \neq 1$. This is because the right side of (2.11b) becomes $(-4\pi/c)\mathbf{J}$ and $J^0 = \rho c$.

Now we need a bit of terminology. Recall that a 4-vector was *defined* as a quantity v^μ that under a Lorentz transformation transformed as

$$v^\mu \rightarrow v'^\mu = \Lambda^\mu_\nu v^\nu.$$

As we shall see below, it is also possible to have quantities with more than one index such that under a Lorentz transformation, each index transforms like a 4-vector. For example, a quantity $F^{\mu\nu}$ (not necessarily the electromagnetic field tensor) with two indices that transforms like

$$F^{\mu\nu} \rightarrow F'^{\mu\nu} = \Lambda^\mu_\alpha \Lambda^\nu_\beta F^{\alpha\beta}$$

is called a (second rank) **tensor**. Higher rank tensors are defined in the obvious manner. Note also that all of the indices need not be superscripts (such indices

are called **contravariant**). We can equally have subscripts (called **covariant**) that transform like

$$F'_{\mu\nu} = \Lambda_{\mu}^{\alpha} \Lambda_{\nu}^{\beta} F_{\alpha\beta}.$$

And we can have a mixed tensor like $F^{\mu}{}_{\nu}$. Indices are raised and lowered by using the metric $g_{\mu\nu}$ and its inverse $g^{\mu\nu}$.

At last we are ready to write Maxwell's equations in covariant form. It is not hard to show that even though ∂_{μ} transforms as a 4-vector under a Lorentz transformation $\Lambda^{\mu}{}_{\nu}$, as does A_{μ} , the quantity $\partial_{\mu}A_{\nu}$ does not transform as a second-rank tensor. However, the antisymmetric quantity $F_{\mu\nu}$ defined by

$$F_{\mu\nu} := \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu} \quad (2.13a)$$

does indeed transform as a tensor. This is called the **electromagnetic field tensor**. Equivalently, we may consider the contravariant version

$$F^{\mu\nu} = \partial^{\mu}A^{\nu} - \partial^{\nu}A^{\mu}. \quad (2.13b)$$

I claim that equations (2.9a) and (2.9d) can be written in the form

$$\partial_{\mu}F^{\mu\nu} = J^{\nu}. \quad (2.14)$$

To see this, first recall that we are using the metric $g = \text{diag}(1, -1, -1, -1)$ so that $\partial/\partial t = \partial/\partial x^0 = \partial_0 = \partial^0$ and $\nabla^i := \partial/\partial x^i = \partial_i = -\partial^i$. Using $\mathbf{E} = -\nabla\varphi - \partial\mathbf{A}/\partial t$ we have

$$E^i = \partial^i A^0 - \partial^0 A^i = F^{i0}$$

and also $\mathbf{B} = \nabla \times \mathbf{A}$ so that

$$B^1 = \nabla^2 A^3 - \nabla^3 A^2 = -\partial^2 A^3 + \partial^3 A^2 = -F^{23}$$

plus cyclic permutations $1 \rightarrow 2 \rightarrow 3 \rightarrow 1$. Then the electromagnetic field tensor is given by

$$F^{\mu\nu} = \begin{bmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{bmatrix}. \quad (2.15)$$

(Be sure to note that this is the form of $F^{\mu\nu}$ for the metric $\text{diag}(1, -1, -1, -1)$. If you use the metric $\text{diag}(-1, 1, 1, 1)$ then all entries of $F^{\mu\nu}$ change sign. In addition, you frequently see the matrix $F^{\mu}{}_{\nu}$ which also has different signs.)

Now, for the $\nu = 0$ component of (2.14) we have

$$J^0 = \partial_{\mu}F^{\mu 0} = \partial_i F^{i0} = \partial_i E^i$$

which is Coulomb's law

$$\nabla \cdot \mathbf{E} = \rho.$$

Next consider the $\nu = 1$ component of (2.14). This is

$$\begin{aligned} J^1 &= \partial_\mu F^{\mu 1} = \partial_0 F^{01} + \partial_2 F^{21} + \partial_3 F^{31} \\ &= -\partial_0 E^1 + \partial_2 B^3 - \partial_3 B^2 \\ &= -\partial_t E^1 + (\nabla \times \mathbf{B})^1 \end{aligned}$$

and therefore we have

$$\nabla \times \mathbf{B} - \frac{\partial \mathbf{E}}{\partial t} = \mathbf{J}.$$

Finally, I leave it as an exercise for you to show that equations (2.9b) and (2.9c) can be written as (note the superscripts are cyclic permutations)

$$\partial^\mu F^{\nu\sigma} + \partial^\nu F^{\sigma\mu} + \partial^\sigma F^{\mu\nu} = 0$$

or simply

$$\partial^{[\mu} F^{\nu\sigma]} = 0. \quad (2.16)$$

Remark: There is another interesting way to arrive at the electromagnetic field tensor that we now describe, but you are free to skip over it. First we need to give a more careful definition of a tensor. To begin, given a vector space V , we can define the **dual space** V^* as the vector space of linear functionals on V . In other words, $\alpha \in V^*$ means that $\alpha : V \rightarrow \mathbb{R}$ is a linear map from V to \mathbb{R} . (We restrict consideration to real vector spaces.) Members of the dual space are frequently called **covectors**. If V has a basis $\{e_1, \dots, e_n\}$, we *define* the n linear functionals $\{\omega^1, \dots, \omega^n\}$ by

$$\omega^i(e_j) = \delta_j^i.$$

I will show that these n linear functionals form a basis for V^* , i.e., that they are linearly independent and span V^* .

To show this, let $\alpha \in V^*$ be arbitrary but fixed. Note that for any $v \in V$, using the linearity of α we have

$$\alpha(v) = \alpha(v^i e_i) = v^i \alpha(e_i) = a_i v^i$$

where we have defined the scalars a_i by $a_i = \alpha(e_i)$. On the other hand,

$$\omega^i(v) = \omega^i(v^j e_j) = v^j \omega^i(e_j) = v^j \delta_j^i = v^i$$

and hence we see that $\alpha(v) = a_i v^i = a_i \omega^i(v)$ so that $\alpha = a_i \omega^i$. This shows that the ω^i span V^* .

To show they are linearly independent, suppose we have $c_i \omega^i = 0$ for some set of scalars c_i . Then for any $j = 1, \dots, n$ we have $0 = c_i \omega^i(e_j) = c_i \delta_j^i = c_j$ which proves linear independence. Thus we have shown that any $\alpha \in V^*$ can be written in the form

$$\alpha = a_i \omega^i = \alpha(e_i) \omega^i.$$

As an example, consider the space $V = \mathbb{R}^2$ consisting of all column vectors of the form

$$v = \begin{bmatrix} v^1 \\ v^2 \end{bmatrix}.$$

Relative to the standard basis we have

$$v = v^1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + v^2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = v^1 e_1 + v^2 e_2.$$

If $\phi \in V^*$, then $\phi(v) = \sum \phi_i v^i$, and we may represent ϕ by the row vector $\phi = (\phi_1, \phi_2)$. In particular, if we write the dual basis as $\omega^i = (a_i, b_i)$, then we have

$$1 = \omega^1(e_1) = (a_1, b_1) \begin{bmatrix} 1 \\ 0 \end{bmatrix} = a_1$$

$$0 = \omega^1(e_2) = (a_1, b_1) \begin{bmatrix} 0 \\ 1 \end{bmatrix} = b_1$$

$$0 = \omega^2(e_1) = (a_2, b_2) \begin{bmatrix} 1 \\ 0 \end{bmatrix} = a_2$$

$$1 = \omega^2(e_2) = (a_2, b_2) \begin{bmatrix} 0 \\ 1 \end{bmatrix} = b_2$$

so that $\omega^1 = (1, 0)$ and $\omega^2 = (0, 1)$. Note, for example,

$$\omega^1(v) = (1, 0) \begin{bmatrix} v^1 \\ v^2 \end{bmatrix} = v^1$$

as it should.

As another very important example, let V be an inner product space. If $a, b \in V$, then the inner product of a and b is the number $\langle a, b \rangle$. Then given a fixed vector a , the quantity $\langle a, \cdot \rangle$ is a linear functional on V because it takes a vector $b \in V$ and gives back a number:

$$\langle a, \cdot \rangle : b \mapsto \langle a, b \rangle \in \mathbb{R}.$$

Since $\langle a, \cdot \rangle$ is in V^* , let us denote it by α , so that $\alpha(b) = \langle a, b \rangle$.

Given a basis $\{e_i\}$ for V , let us *define* the numbers g_{ij} by

$$g_{ij} := \langle e_i, e_j \rangle = \langle e_j, e_i \rangle = g_{ji}.$$

This is the proper definition of the **metric**. Then

$$\langle a, b \rangle = \langle a^i e_i, b^j e_j \rangle = a^i b^j \langle e_i, e_j \rangle = a^i b^j g_{ij} = b^j g_{ji} a^i.$$

But we also have

$$\alpha(b) = \alpha(b^j e_j) = b^j \alpha(e_j) = b^j a_j$$

and therefore

$$b^j a_j = \alpha(b) = \langle a, b \rangle = b^j g_{ji} a^i .$$

Hence we define

$$a_j = g_{ji} a^i .$$

This is called **lowering an index**.

Since the inner product is nondegenerate by definition, the metric g_{ij} must be nonsingular, and hence we can define its inverse which we denote by g^{ij} . Multiplying this last equation by g^{kj} we then have

$$g^{kj} a_j = g^{kj} g_{ji} a^i = \delta_i^k a^i = a^k$$

and thus we define **raising an index** by

$$a^k = g^{kj} a_j .$$

Now that we have an understanding of dual spaces, we are in a position to define tensors carefully. So, a **tensor** T is just a *multilinear* map

$$T : V^{*s} \times V^r = V^* \times \dots \times V^* \times V \times \dots \times V \rightarrow \mathbb{R} .$$

By multilinear, we mean that it is linear in each variable separately. This tensor is said to have covariant order r , and contravariant order s , or simply to be a tensor of **type** (s, r) . In other words, T takes as its argument s covectors and r vectors. Since it is multilinear, we see that

$$\begin{aligned} T(\alpha^{(1)}, \dots, \alpha^{(s)}, v_{(1)}, \dots, v_{(r)}) &= T(a_{i_1}^{(1)} \omega^{i_1}, \dots, a_{i_s}^{(s)} \omega^{i_s}, v_{(1)}^{j_1} e_{j_1}, \dots, v_{(r)}^{j_r} e_{j_r}) \\ &= a_{i_1}^{(1)} \dots a_{i_s}^{(s)} v_{(1)}^{j_1} \dots v_{(r)}^{j_r} T(\omega^{i_1}, \dots, \omega^{i_s}, e_{j_1}, \dots, e_{j_r}) \\ &= a_{i_1}^{(1)} \dots a_{i_s}^{(s)} v_{(1)}^{j_1} \dots v_{(r)}^{j_r} T^{i_1 \dots i_s}_{j_1 \dots j_r} \end{aligned}$$

where the last line *defines* the components of T . Thus we see that if we know the components of T , then we know the result of T acting on an arbitrary collection of vectors and covectors.

What happens to the components of T under a change of coordinates? From a practical standpoint, this is what really defines a tensor. A change of basis in V is of the form

$$e_i \rightarrow \bar{e}_i = e_j p^j_i$$

where (p^j_i) is called the **transition matrix**. Then any $x \in V$ can be written in terms of its components with respect to either e_i or \bar{e}_i , and we have

$$x = x^j e_j = \bar{x}^i \bar{e}_i = \bar{x}^i e_j p^j_i$$

and therefore we must have

$$x^j = p^j_i \bar{x}^i$$

or

$$\bar{x}^i = (p^{-1})^i_j x^j.$$

From these we easily see that

$$p^i_j = \frac{\partial x^i}{\partial \bar{x}^j} \quad \text{and} \quad (p^{-1})^i_j = \frac{\partial \bar{x}^i}{\partial x^j}.$$

When V undergoes a change of basis, what about V^* ? Let us write in general $\omega^i \rightarrow \bar{\omega}^i = b^i_j \omega^j$. Since we must also have $\bar{\omega}^i(\bar{e}_j) = \delta^i_j$, we see that

$$\delta^i_j = \bar{\omega}^i(\bar{e}_j) = \bar{\omega}^i(e_k p^k_j) = p^k_j \bar{\omega}^i(e_k) = p^k_j b^i_l \omega^l(e_k) = p^k_j b^i_l \delta^l_k = b^i_k p^k_j$$

so that $b^i_k = (p^{-1})^i_k$. In other words,

$$\omega^i \rightarrow \bar{\omega}^i = (p^{-1})^i_j \omega^j.$$

Finally we are in a position to derive the general transformation law of a tensor. We have

$$\begin{aligned} \bar{T}^{i_1 \dots i_s}_{j_1 \dots j_r} &= T(\bar{\omega}^{i_1}, \dots, \bar{\omega}^{i_s}, \bar{e}_{j_1}, \dots, \bar{e}_{j_r}) \\ &= T((p^{-1})^{i_1}_{k_1} \omega^{k_1}, \dots, (p^{-1})^{i_s}_{k_s} \omega^{k_s}, e_{l_1} p^{l_1}_{j_1}, \dots, e_{l_r} p^{l_r}_{j_r}) \\ &= (p^{-1})^{i_1}_{k_1} \dots (p^{-1})^{i_s}_{k_s} p^{l_1}_{j_1} \dots p^{l_r}_{j_r} T(\omega^{k_1}, \dots, \omega^{k_s}, e_{l_1}, \dots, e_{l_r}) \\ &= (p^{-1})^{i_1}_{k_1} \dots p^{l_1}_{j_1} \dots T^{k_1 \dots k_s}_{l_1 \dots l_r} \end{aligned}$$

or

$$\bar{T}^{i_1 \dots i_s}_{j_1 \dots j_r} = \frac{\partial \bar{x}^{i_1}}{\partial x^{k_1}} \dots \frac{\partial \bar{x}^{i_s}}{\partial x^{k_s}} \frac{\partial x^{l_1}}{\partial \bar{x}^{j_1}} \dots \frac{\partial x^{l_r}}{\partial \bar{x}^{j_r}} T^{k_1 \dots k_s}_{l_1 \dots l_r}.$$

This is the **classical transformation law** of a type (s, r) tensor.

In the particular case of a second rank tensor $F^{\mu\nu}$ under a Lorentz transformation, we have

$$\bar{x}^\mu = \Lambda^\mu_\nu x^\nu$$

so that

$$\frac{\partial \bar{x}^\mu}{\partial x^\nu} = \Lambda^\mu_\nu$$

and therefore

$$\bar{F}^{\mu\nu} = \frac{\partial \bar{x}^\mu}{\partial x^\alpha} \frac{\partial \bar{x}^\nu}{\partial x^\beta} F^{\alpha\beta} = \Lambda^\mu_\alpha \Lambda^\nu_\beta F^{\alpha\beta}.$$

Let us now return to the physics. We know that the Lorentz force law is

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) = \frac{d\mathbf{p}}{dt}$$

so consider (now τ is the proper time again)

$$\frac{d\mathbf{p}}{d\tau} = \frac{dt}{d\tau} \frac{d\mathbf{p}}{dt} = \gamma \frac{d\mathbf{p}}{dt}.$$

In terms of the 4-velocity

$$u^\mu = \begin{bmatrix} \gamma \\ \gamma \mathbf{v} \end{bmatrix} = \begin{bmatrix} u^0 \\ \mathbf{u} \end{bmatrix}$$

we can write

$$\begin{aligned} \frac{d\mathbf{p}}{d\tau} &= \gamma q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) = q(\gamma \mathbf{E} + \gamma \mathbf{v} \times \mathbf{B}) \\ &= q(u^0 \mathbf{E} + \mathbf{u} \times \mathbf{B}). \end{aligned} \quad (2.17)$$

Also note that if $W = p^0$ is the energy of the particle, then the change in energy is the rate at which work is done, so that

$$\frac{dW}{dt} = \mathbf{F} \cdot \frac{d\mathbf{x}}{dt} = \mathbf{F} \cdot \mathbf{v} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathbf{v} = q\mathbf{E} \cdot \mathbf{v}$$

and therefore

$$\frac{dp^0}{d\tau} = \frac{dW}{d\tau} = \gamma \frac{dW}{dt} = q\mathbf{E} \cdot \gamma \mathbf{v} = q\mathbf{E} \cdot \mathbf{u}. \quad (2.18)$$

Combining equations (2.17) and (2.18), we see that we can define a linear map $u^\mu \rightarrow dp^\mu/d\tau$ of a 4-vector to another 4-vector, and hence there exists a second rank mixed tensor $F^\mu{}_\nu$ such that

$$\frac{dp^\mu}{d\tau} = qF^\mu{}_\nu u^\nu. \quad (2.19)$$

Comparing (2.19) with (2.17) and (2.18) allows us to pick out the components of $F^\mu{}_\nu$:

$$F^\mu{}_\nu = \begin{bmatrix} 0 & E_x & E_y & E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{bmatrix}.$$

Finally, we note that we can also write $F^\mu{}_\nu$ in the alternate forms $F_{\mu\nu} = g_{\mu\alpha} F^\alpha{}_\nu$ and $F^{\mu\nu} = g^{\nu\alpha} F^\mu{}_\alpha$.

Now that we have the electromagnetic field tensor, it is straightforward to derive the transformation laws for the \mathbf{E} and \mathbf{B} fields. Starting from $F^{\mu\nu}$ which defines the fields \mathbf{E} and \mathbf{B} , we have $F'^{\mu\nu} = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta F^{\alpha\beta}$ which then gives the fields \mathbf{E}' and \mathbf{B}' in terms of \mathbf{E} and \mathbf{B} . In matrix notation, we can write this as

$$F' = \Lambda F \Lambda^T.$$

Using equations (1.14) and (2.15), it is easy to multiply out the matrices and

show that

$$\begin{aligned}
 F'^{\mu\nu} &= \begin{bmatrix} 0 & -E'^1 & -E'^2 & -E'^3 \\ E'^1 & 0 & -B'^3 & B'^2 \\ E'^2 & B'^3 & 0 & -B'^1 \\ E'^3 & -B'^2 & B'^1 & 0 \end{bmatrix} \\
 &= \begin{bmatrix} 0 & -E^1 & -\gamma(E^2 - \beta B^3) & -\gamma(E^3 + \beta B^2) \\ E^1 & 0 & -\gamma(B^3 - \beta E^2) & \gamma(B^2 + \beta E^3) \\ \gamma(E^2 - \beta B^3) & \gamma(B^3 - \beta E^2) & 0 & -B^1 \\ \gamma(E^3 + \beta B^2) & -\gamma(B^2 + \beta E^3) & B^1 & 0 \end{bmatrix}.
 \end{aligned}$$

This was for the special case of a boost along the x -axis, i.e., $\boldsymbol{\beta} = \beta \hat{\mathbf{x}}$. it is not hard to see that we can write down the field transformation laws for a boost $\boldsymbol{\beta}$ in an arbitrary direction (but with the coordinate axes still parallel) if we write this in terms of components parallel and perpendicular to the boost. This yields

$$\mathbf{E}'_{\perp} = \gamma(\mathbf{E}_{\perp} + \boldsymbol{\beta} \times \mathbf{B}) \qquad \mathbf{E}'_{\parallel} = \mathbf{E}_{\parallel} \qquad (2.20a)$$

$$\mathbf{B}'_{\perp} = \gamma(\mathbf{B}_{\perp} - \boldsymbol{\beta} \times \mathbf{E}) \qquad \mathbf{B}'_{\parallel} = \mathbf{B}_{\parallel} \qquad (2.20b)$$