The Relativistic Symmetry of the Electromagnetic Lagrangian

We have obtained the Lagrangian density for electromagnetism as

$$\mathcal{L} = -\rho\,\Phi + \vec{j}\cdot\vec{A} + \frac{1}{2}\,\varepsilon_0(\partial_i\Phi + \partial_t A_i)\,(\partial_i\Phi + \partial_t A_i) - \frac{1}{4\,\mu_0}\,(\partial_i A_j - \partial_j A_i)\,(\partial_i A_j - \partial_j A_i).$$

The very important relative sign of the electric and magnetic terms was resolved by making sure that, when matter is added, the force, defined as the time rate of change of the momentum, which in turn is defined as the partial derivative of the matter Lagrangian with respect to the matter's velocity is in agreement with experiment. Further, the overall sign of the above Lagrangian is adjusted so that the kinetic energy of matter is to be added to it when the dynamics of matter is taken into account. Experiment had to be invoked to get the signs right, and as we will see this point is crucial. Indeed, in some sense, the following is an exercise in keeping track of signs carefully!

■ Linear Transformations I

If you are sensitive to symmetries in a theory, then you will seek to invent transformations of the coordinates and/or fields under which the Lagrangian density is invariant (to within a divergence). In general this gives insight into the content of the theory. Although in this course I have not derived it , Nöther theorem in field theory shows how to generate a conservation law for each continuously variable parameter in a symmetry transformation of the Lagrangian density.

So, looking at the above with an eye to symmetries, we can notice that the two source terms

$$-\rho\,\Phi + \vec{j}\cdot\vec{A} = -\rho\,\Phi + j_x\,A_x + j_y\,A_y + j_z\,A_z$$

Immediately suggest one: any linear transformation of both the current and the potentials that leaves this form invariant. As we will see, such linear transformations are those of Einstein's special relativity. But first various formalities and definitions which make the discussion convenient are useful to introduce. This will insure that the final formulation will conform to standard practice.

■ Formalities, notations, and the metric tensor

Begin by making the notation more symmetric. Let

$$\begin{pmatrix} j^0 \\ j^1 \\ j^2 \\ j^3 \end{pmatrix} = \begin{pmatrix} c\rho \\ j_x \\ j_y \\ j_z \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} A^0 \\ A^1 \\ A^2 \\ A^3 \end{pmatrix} = \begin{pmatrix} \frac{\Phi}{c} \\ A_x \\ A_y \\ A_z \end{pmatrix}$$

so that $-\rho\,\Phi + \vec{j}\cdot\vec{A}$ becomes $-(j^0\,A^0 - j^1\,A^1 - j^2\,A^2 - j^3\,A^3)$. The use of superscripts may seem odd, but it is conventional and will be discussed shortly.

In matrix–vector notation we can produce this form in four different ways, "keeping things that go together, together":

$$\rho\,\Phi - j_x\,A_x - j_y\,A_y - j_z\,A_z$$

$$= (\rho \quad j_x \quad j_y \quad j_z)\begin{pmatrix}\Phi\\ -A_x\\ -A_y\\ -A_z\end{pmatrix}$$

$$= (\rho \quad -j_x \quad -j_y \quad -j_z)\begin{pmatrix}\Phi\\ A_x\\ A_y\\ A_z\end{pmatrix}$$

$$= (\rho \quad j_x \quad j_y \quad j_z)\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}\Phi\\ A_x\\ A_y\\ A_z\end{pmatrix}$$

$$= (\rho \quad -j_x \quad -j_y \quad -j_z)\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}\begin{pmatrix}\Phi\\ -A_x\\ -A_y\\ -A_z\end{pmatrix}.$$

Rather than adopting just one of these forms as the standard one, a notation has been invented that incorporates *all* of them. In particular, in addition to the definitions given above, we also introduce

$$\begin{pmatrix}j_0\\ j_1\\ j_2\\ j_3\end{pmatrix}=\begin{pmatrix}c\rho\\ -j_x\\ -j_y\\ -j_z\end{pmatrix} \quad \text{and} \quad \begin{pmatrix}A_0\\ A_1\\ A_2\\ A_3\end{pmatrix}=\begin{pmatrix}\frac{\Phi}{c}\\ -A_x\\ -A_y\\ -A_z\end{pmatrix}.$$

More generally, the systematic notation inventions that have been incorporated into the subject of relativistic linear transformations are as follows:

a) Let Greek letter indices have a range 0, 1, 2, 3.
b) Distinguish two kinds of tensorial indices: covariant indices which are written as subscripts like j_μ and contravariant indices which are written as superscripts like j^μ (more on this later).
c) Use the Einstein summation convention. Summing over a pair of indices is called a *contraction,* and in general, a contraction is legitimate if and only if one index is covariant and the other is contravariant. Thus by $j_\mu\,A^\mu$ we mean $j_0\,A^0 + j_1\,A^1 + j_2\,A^2 + j_3\,A^3 = \rho\,\Phi - j_x\,A_x - j_y\,A_y - j_z\,A_z$, and $j_\mu\,A_\mu$ **is a mistake**. This convention is extreme useful for avoiding errors.
d) Introduce the *metric tensor* $g^{\mu\nu} = g_{\mu\nu} = \begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix} = \begin{pmatrix}1 & \underset{\sim}{0}^T\\ \underset{\sim}{0} & -\underset{\approx}{1}\end{pmatrix}$. In the (1+3)×(1+3) matrix notation the upper left

"1" is just the number 1, the zero vector $\underset{\sim}{0}$ has three rows and one column, $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $\underset{\approx}{1}$ is a square 3×3 unit matrix, $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, and the zero vector, $\underset{\sim}{0}^T$, is the transpose of $\underset{\sim}{0}$ and so has one row and three columns; it is $(\,0 \;\; 0 \;\; 0\,)$. Note that both indices of the metric tensor may be covariant or contravariant and that its square is the unit matrix: $g_{\mu\nu}\, g^{\nu\lambda} = \delta_\mu^\lambda = \begin{pmatrix} 1 & \underset{\sim}{0}^T \\ \underset{\sim}{0} & \underset{\approx}{1} \end{pmatrix}$.
Finally, this form for the metric tensor (rather than its negative) is often called the *West Coast* convention; they tell me that another possibility is used in the never-never land of the Eastern seaboard.
e) The metric tensor acts as an *index lowering or raising operator*. In other words it may be used to convert a covariant index to a contravariant one. So we have $j_\mu = g_{\mu\nu}\, j^\nu$ and $j^\mu = g^{\mu\nu}\, j_\nu$. Similarly from $g_{\mu\nu}\, g^{\nu\lambda} = \delta_\mu^\lambda$ we get the mixed form:

$$g_\mu^\nu = \delta_\mu^\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

f) Finally, adopt the convention as given above for the contravariant components of j^μ and A^μ. In accordance with the definition given above of the covariant indices, notice that $j_\mu = g_{\mu\nu}\, j^\nu$ so that $(j_0,\, j_{1,},\, j_{2,},\, j_3) = (\rho,\, -j_x,\, -j_y,\, -j_z)$ and similarly for A_μ.

With all these notational devices (which were invented for more general situations that we have in special relativity) , we now see that the four ways of writing $\rho\,\Phi - j_x\, A_x - j_y\, A_y - j_z\, A_z$ are

$$\rho\,\Phi - j_x\, A_x - j_y\, A_y - j_z\, A_z = j^\mu\, A_\mu = j_\mu\, A^\mu = j^\mu\, g_{\mu\nu}\, A^\nu = j_\mu\, g^{\mu\nu}\, A_\nu.$$

Incidentally, this can also be written as $j^\mu\, g_\mu^\nu\, A_\nu = j^\mu\, \delta_\mu^\nu\, A_\nu = j_\nu\, g_\mu^\nu\, A^\mu = j_\nu\, \delta_\mu^\nu\, A^\mu$.

Finally we ought to verify the consistency of the raising–lowering definition given above and $g_{\mu\nu} = g^{\mu\nu}$. Using a matrix notation for convenience we have

$$g_{\mu'\mu}\, g^{\mu\nu}\, g_{\nu\nu'} = \begin{pmatrix} 1 & \underset{\sim}{0}^T \\ \underset{\sim}{0} & -\underset{\approx}{1} \end{pmatrix} \begin{pmatrix} 1 & \underset{\sim}{0}^T \\ \underset{\sim}{0} & -\underset{\approx}{1} \end{pmatrix} \begin{pmatrix} 1 & \underset{\sim}{0}^T \\ \underset{\sim}{0} & -\underset{\approx}{1} \end{pmatrix} = \begin{pmatrix} 1 & \underset{\sim}{0}^T \\ \underset{\sim}{0} & -\underset{\approx}{1} \end{pmatrix} \begin{pmatrix} 1 & \underset{\sim}{0}^T \\ \underset{\sim}{0} & \underset{\approx}{1} \end{pmatrix} = \begin{pmatrix} 1 & \underset{\sim}{0}^T \\ \underset{\sim}{0} & -\underset{\approx}{1} \end{pmatrix} = g_{\mu'\nu'}.$$

■ Units

The "c's" in the definition of the zeroth component of j^μ and A^μ were introduced to arrange that all of the related quantities carry the same units. We know that $c\rho$ has units of $\frac{m}{\text{sec}}\,\frac{\text{Coul}}{m^3} = \frac{\text{Amp}}{m^2}$, the units of current density. Similarly, Φ has units of Volts and A, units of $\frac{\text{Henry}}{m}\,\text{Amp} = \frac{\text{sec}}{m}\,\frac{\text{Henry}}{\text{sec}}\,\text{Amp}\left(\text{from}\,\vec{A} = \frac{\mu_0}{4\pi}\int \frac{\vec{j}(\vec{r}')}{|\vec{r}-\vec{r}'|}\, d^3 r'\right)$. But $\frac{\text{Henry}}{\text{sec}}$ is the "$\omega\, L$" of circuit theory or Ohm, and Ohm Amp is Volt. So we see that A has units of $\frac{\text{Volt}}{m/\text{sec}}$,
the same as $\frac{\Phi}{c}$. So, as defined, all of A^μ have the same units and similarly for j^μ. Of course this unit–consistency is not strictly necessary, but it is very convenient.

■ Linear Transformations II

Remember that our goal is to seek linear transformations which are symmetries and the above barrage of notation was introduced to help keep track of details in making linear transformations. So notice that if you make a linear transformation

$(j'(t,x,y,z))^{\mu} = L^{\mu\cdot}_{\cdot\nu'}(j(t,x,y,z))^{\nu'}$
and
$(A'(t,x,y,z))^{\nu} = L^{\nu\cdot}_{\cdot\mu'}(A(t,x,y,z))^{\mu'}$ (these are NOT relativistic transformations)

then $(j'(t,x,y,z))^{\mu} g_{\mu\nu}(A'(t,x,y,z))^{\nu} = (j(t,x,y,z))^{\nu'} L^{\mu\cdot}_{\cdot\nu'} g_{\mu\nu} L^{\nu\cdot}_{\cdot\mu'}(A(t,x,y,z))^{\mu'}$ and so if we demand that the transformation satisfy the requirement that

$$L^{\mu\cdot}_{\cdot\nu'} g_{\mu\nu} L^{\nu\cdot}_{\cdot\mu'} = g_{\nu'\mu'}$$

(require the linear transformation to preserve the metric tensor) then we get the beginnings of a symmetry; with this condition, we have $(j(t,x,y,z))^{\nu'} g_{\nu'\mu'}(A(t,x,y,z))^{\mu'} = (j'(t,x,y,z))^{\mu} g_{\mu\nu}(A'(t,x,y,z))^{\nu}$. Some important things to notice are:

a) in the linear transformation matrix the dots $L^{\mu\cdot}_{\cdot\nu}$ in index positions are place holders. They remove the ambiguity that would otherwise result when a matrix notation is used. Thus in a matrix notation we have, with the index μ preceding the ν,

$$L^{\mu\cdot}_{\cdot\nu} = \underset{\approx}{L} = \begin{pmatrix} L^{0\cdot}_{\cdot 0} & L^{0\cdot}_{\cdot 1} & L^{0\cdot}_{\cdot 2} & L^{0\cdot}_{\cdot 3} \\ L^{1\cdot}_{\cdot 0} & L^{1\cdot}_{\cdot 1} & L^{1\cdot}_{\cdot 2} & L^{1\cdot}_{\cdot 3} \\ L^{2\cdot}_{\cdot 0} & L^{2\cdot}_{\cdot 1} & L^{2\cdot}_{\cdot 2} & L^{2\cdot}_{\cdot 3} \\ L^{3\cdot}_{\cdot 0} & L^{3\cdot}_{\cdot 1} & L^{3\cdot}_{\cdot 2} & L^{3\cdot}_{\cdot 3} \end{pmatrix}$$

and $L^{\mu\cdot}_{\cdot\nu'} g_{\mu\nu} L^{\nu\cdot}_{\cdot\mu'} = g_{\nu'\mu'}$ can be written in convenient matrix notation as $\underset{\approx}{L}^T \underset{\approx}{g} \underset{\approx}{L} = \underset{\approx}{g}$.

Matrix notation is very useful in practice for making transformation calculations, but care must be taken to distinguish a matrix from its transpose. Incidentally there is no need for the place holders in a symmetric matrix such as δ^{μ}_{ν}.
b) the primes on the indices are used only to avoid index-name proliferation,
c) the primes on the physical quantities are to indicate that the values of the primed components have different values from the unprimed ones.
d) as I have defined this transformation here, the space-time arguments of the fields′ component functions are the same in both the primed and unprimed fields and it is for this reason that this in NOT a relativistic transformation.

If this procedure had turned out to leave the whole Lagrangian in exactly the same form as it had before the transformation, it would be called an "internal symmetry", i.e., one not involving space-time. An example of such a transformation is that of isotopic spin; it is an approximate symmetry in nature and, in its original form, it says that neutrons and protons – apart from E&M! – are just the same. In the case at hand, however, we have **not** found a symmetry since the last two terms of $\mathcal{L}$ involving derivatives do not map into the same functional form under this transformation. We will see how to remedy this shortly but first let's look at the linear transformation a bit more.

■ Formalities involving linear transformations, covariant/contravariant indices, and the metric tensor

Since we have introduced both covariant and contravariant indices, or covariant and contravariant *tensors* as people usually say, we ought to see how they transform. From $(j')^\mu = L^{\mu\cdot}_{\cdot\nu'}(j)^{\nu'}$ we can immediately write, using j'^μ to mean $(j')^\mu$ to simplify notation, and dropping the space-time arguments for conciseness,

$$g_{\lambda\mu}\, j'^\mu = g_{\lambda\mu}\, L^{\mu\cdot}_{\cdot\nu'}\, \delta^{\nu'}_\nu\, j^\nu = g_{\lambda\mu}\, L^{\mu\cdot}_{\cdot\nu'}\, g^{\nu'\lambda'}\, g_{\lambda'\nu}\, j^\nu = (g_{\lambda\mu}\, L^{\mu\cdot}_{\cdot\nu'}\, g^{\nu'\lambda'})\, g_{\lambda'\nu}\, j^\nu$$

or

$$j'_\lambda = (g_{\lambda\mu}\, L^{\mu\cdot}_{\cdot\nu'}\, g^{\nu'\lambda'})\, j_{\lambda'}.$$

This then is the transformation law for a covariant index. But we can re-express $g_{\lambda\mu}\, L^{\mu\cdot}_{\cdot\nu'}\, g^{\nu'\lambda'}$ nicely using the fact that the metric tensor is symmetric and that the linear transformation satisfies the strong condition $L^{\mu\cdot}_{\cdot\nu'}\, g_{\mu\nu}\, L^{\nu\cdot}_{\cdot\mu'} = g_{\nu'\mu'}$ as follows:

$$L^{\mu\cdot}_{\cdot\nu'}\, g_{\mu\nu}\, L^{\nu\cdot}_{\cdot\mu'} = g_{\nu'\mu'} \quad \Rightarrow \quad L^{\mu\cdot}_{\cdot\nu'}\, g_{\mu\nu}\, L^{\nu\cdot}_{\cdot\mu'}\, g^{\mu'\lambda'} = g_{\nu'\mu'}\, g^{\mu'\lambda'} = \delta^{\lambda'}_{\nu'} = L^{\mu\cdot}_{\cdot\nu'}(\, g_{\mu\nu}\, L^{\nu\cdot}_{\cdot\mu'}\, g^{\mu'\lambda'}).$$

In matrix notation, this last equation is $\underset{\approx}{L}^T \underset{\approx}{g}\, \underset{\approx}{L}\, \underset{\approx}{g} = \underset{\approx}{1}$ or $\underset{\approx}{g}\, \underset{\approx}{L}\, \underset{\approx}{g} = (\underset{\approx}{L}^T)^{-1} = (\underset{\approx}{L}^{-1})^T$, and so the transformation equation for a covariant vector is

$$j'_\lambda = (g_{\lambda\mu}\, L^{\mu\cdot}_{\cdot\nu'}\, g^{\nu'\lambda'})\, j_{\lambda'} = (L^{-1})^{\lambda'\cdot}_{\cdot\lambda}\, j_{\lambda'}.$$

In words, covariant vectors transform inversely as contravariant ones. In general, this is the fundamental difference between covariant and contravariant indices. I have shown it in the case that there is a metric that is invariant under the transformation, but the notion is more basic than that and has no need of the metric in spite of its use in this derivation.

The following remarks are not necessary for showing the relativistic invariance of E&M; they are included only for completeness.

The basic idea in linear algebra is that a vector $\vec{v}$ in any dimension N can be expressed as a weighted sum of a basis set of linearly independent vectors, $\vec{e}_n$ (note that the convention is that the unit vectors, with covariant indices, act as a row matrix and the weights, with contravariant indices, as a column matrix):

$$\vec{v} = \sum_{n=0}^{N} \vec{e}_n\, v^n = (\, \vec{e}_0 \quad \ldots \quad \vec{e}_N\,) \begin{pmatrix} v^0 \\ \ldots \\ v^N \end{pmatrix}.$$

If another linearly independent basis $\vec{e}'_n$ is defined by

$$\vec{e}'_n = \sum_{m=0}^{N} \vec{e}_m (\Lambda^{-1})^{m\cdot}_{\cdot n} \quad \text{so that} \quad \vec{e}_m = \sum_{\ell=0}^{N} \vec{e}'_\ell\, \Lambda^{\ell\cdot}_{\cdot m} \quad \text{where} \quad \sum_{m=0}^{N} (\Lambda)^{\ell\cdot}_{\cdot m}\, (\Lambda^{-1})^{m\cdot}_{\cdot n} = \delta^\ell_n,$$

then we have

$$\vec{v} = \sum_{m=0}^{N} \vec{e}_m\, v^m = \sum_{m=0}^{N} v^m \sum_{\ell=0}^{N} \vec{e}'_\ell\, \Lambda^{\ell\cdot}_{\cdot m} = \sum_{\ell=0}^{N} \left(\sum_{m=0}^{N} \Lambda^{\ell\cdot}_{\cdot m}\, v^m \right) \vec{e}'_\ell = \sum_{\ell=0}^{N} v'^\ell\, \vec{e}'_\ell \quad \text{where} \quad v'^\ell = \sum_{m=0}^{N} \Lambda^{\ell\cdot}_{\cdot m}\, v^m.$$

(Note that the components of a vector act as a column matrix). With this law of transformation for a vector's components we see that a vector $\vec{v}$ is a geometric quantity independent of the basis in which it is expressed. Any index that transforms the same way as the basis vectors (the "standard" of transformation) is said to "co-vary" and any that transforms as the components of a vector are said to "counter-vary", i.e., counter the effect of the "varying". Incidentally, it is a matter of choice on which of the two types of indices to use the matrix or its inverse. The convention I have chosen is the usual for relativity.

An inner product between vectors can be introduced by a symmetric, non-singular matrix $\underset{\approx}{G}$ by defining $\vec{e}_n \cdot \vec{e}_m = G_{nm}$. Obviously with this definition, the matrix $\underset{\approx}{G}$ must be symmetric, and the non-singularity condition ensures that the basis vectors are linearly independent (it is easy to show that if $\det(\underset{\approx}{G}) = 0$, then there is a zero vector with some of its components being non–zero). Thus $\underset{\approx}{G}$ can be diagonalized by an orthogonal matrix into $\text{Diag}(G_0, G_1, \ldots, G_N)$ with all of G_i being non–zero, and the orthogonal matrix can be used to create a new basis as a linear combination of the original basis vectors. Finally, by a change of scale of the i^{th} dimension with the factor $\frac{1}{|G_i|}$, we have in the new basis (I use the same letter for the final basis vectors to avoid name proliferation) $\vec{e}_n \cdot \vec{e}_m = g_{nm}$ where g_{nm} is diagonal, and its eigenvalues are either +1 or -1. The absolute value of the difference between the number of positive and negative eigenvalues is called the *signature* of the metric space of a given dimension, and clearly, a metric space is characterized by its signature and dimension. We see that relativistic space is four dimensional with a signature of two.

Back to the main job.

■ Writing the full Lagrangian more symmetrically

Before proceeding, we ought to join the last two terms of the Lagrangian more obviously – their numeric coefficients are not even the same as now written.

First pull out $\frac{1}{\mu_0}$ to get

$$\frac{1}{\mu_0}\left(\frac{1}{2}\,\frac{1}{c^2}\,(\partial_i \Phi + \partial_t A_i)\,(\partial_i \Phi + \partial_t A_i) - \frac{1}{4}\,(\partial_i A_j - \partial_j A_i)\,(\partial_i A_j - \partial_j A_i)\right).$$

Then work on the first term so its units are the same as the last and it has the same numeric factor.

$$\frac{1}{2}\,\frac{1}{c^2}\,(\partial_i \Phi + \partial_t A_i)\,(\partial_i \Phi + \partial_t A_i) = \frac{1}{2}\left(\frac{\partial(\frac{\Phi}{c})}{\partial x_i} + \frac{\partial A_i}{\partial(ct)}\right)\left(\frac{\partial(\frac{\Phi}{c})}{\partial x_i} + \frac{\partial A_i}{\partial(ct)}\right)$$

$$= \frac{1}{4}\left(\left(\frac{\partial(\frac{\Phi}{c})}{\partial x_i} + \frac{\partial A_i}{\partial(ct)}\right)\left(\frac{\partial(\frac{\Phi}{c})}{\partial x_i} + \frac{\partial A_i}{\partial(ct)}\right) + \left(\frac{\partial A_i}{\partial(ct)} + \frac{\partial(\frac{\Phi}{c})}{\partial x_i}\right)\left(\frac{\partial A_i}{\partial(ct)} + \frac{\partial(\frac{\Phi}{c})}{\partial x_i}\right)\right)$$

and so we can write the last two terms as

$$\frac{1}{4\,\mu_0}\left(\left(\frac{\partial(\frac{\Phi}{c})}{\partial x_i} + \frac{\partial A_i}{\partial(ct)}\right)\left(\frac{\partial(\frac{\Phi}{c})}{\partial x_i} + \frac{\partial A_i}{\partial(ct)}\right) + \left(\frac{\partial A_i}{\partial(ct)} + \frac{\partial(\frac{\Phi}{c})}{\partial x_i}\right)\left(\frac{\partial A_i}{\partial(ct)} + \frac{\partial(\frac{\Phi}{c})}{\partial x_i}\right) - (\partial_i A_j - \partial_j A_i)\,(\partial_i A_j - \partial_j A_i)\right).$$

Now we want to invent a transformation so that this combination is invariant. Under the transformations which left the source terms invariant, the set ($\frac{\Phi}{c}$, A_x, A_y, A_z) transformed like a vector; all we need to do is arrange for the *derivatives* in this combination to transform like suitable vectors.

The fundamental idea is to let the space-time arguments of the fields′ functions transform linearly also.

Since the zero index component of the transformation involved a non three-vector quantity (ρ and Φ) whereas the other three components were those of a three vector ($\vec{j}$ and $\vec{A}$), it is natural to make the same arrangement for the space–time quantities, i.e., (ct, x, y, z), but it is not obvious if we should use the covariant or the contravariant transformation for them. I'll leave that question open for now and make which ever choice will lead to a transformation which is a symmetry (if any!).

■ Assigning contravariant indices

Now explore the new transformation (note that I have introduced ct to replace time so that the four space-time coordinates have the same units):

$(j'(ct', x', y', z'))^\mu = L^{\mu\cdot}_{\cdot\nu}\,(j(ct, x, y, z))^\nu$
and
$(A'(ct', x', y', z'))^\mu = L^{\mu\cdot}_{\cdot\nu}(A(ct, x, y, z))^\vee$

and $\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$ is either $\begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}$ or $\begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}$ and similarly for the primed coordinates.

(Note that the space – time coordinates differ on the two sides of the equations in distinction to earlier,)

From $x'^\mu = L^{\mu\cdot}_{\cdot\nu}\, x^\nu$ and $x^\nu = (L^{-1})^{\nu\cdot}_{\cdot\mu}\, x'^\mu$ we get the fundamental relation

$$\frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu}\,\frac{\partial}{\partial x^\nu} = (L^{-1})^{\nu\cdot}_{\cdot\mu}\,\frac{\partial}{\partial x^\nu}$$

and we see that *the index on the derivative with respect to contravariant quantities acts covariantly*. A natural notation then is

$$\frac{\partial}{\partial x^\nu} = \partial_\nu$$

and similarly it is easy to see that the derivative with respect to a covariant quantity yields a contravariant index, so

$$\frac{\partial}{\partial x_\nu} = \partial^\nu.$$

With this notation we see that the quantities

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \quad \text{and} \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$$

where the fields are functions of x^λ (or x_λ) respectively transform as indicated by the indices. Notice that these two tensors are antisymmetric, i.e., $F_{\mu\nu} = -F_{\nu\mu}$ and similarly for $F^{\mu\nu}$. Under the transformation $x'^\mu = L^{\mu\cdot}_{\cdot\mu'}\, x^{\mu'}$, we have $F'^{\mu\nu}(x'^\lambda) = L^{\mu\cdot}_{\cdot\mu'}\, L^{\nu\cdot}_{\cdot\nu'}\, F^{\mu'\nu'}(x^\lambda)$. This then implies that the quantity

$$F^{\mu\nu}(x^\lambda)\, F_{\mu\nu}(x^\lambda)$$

is an invariant under the transformation, i.e.,

$$F^{\mu\nu}(x^\lambda)\, F_{\mu\nu}(x^\lambda) = F'^{\mu\nu}(x'^\lambda)\, F'_{\mu\nu}(x'^\lambda).$$

Now arrange this so that it is ordered for doing the calculation with matrices (and be very careful about covariant and contravariant indices, particularly in derivatives!). In the last step, convert all derivatives to be with respect to contravariant coordinates, and similarly, all potentials to be contravariant:

$$F^{\mu\nu}\, F_{\mu\nu} = -F^{\mu\nu}\, F_{\nu\mu} = -\mathrm{Tr}\left(\begin{pmatrix} 0 & \partial^0 A^1 - \partial^1 A^0 & \partial^0 A^2 - \partial^2 A^0 & \partial^0 A^3 - \partial^3 A^0 \\ \partial^1 A^0 - \partial^0 A^1 & 0 & \partial^1 A^2 - \partial^2 A^1 & \partial^1 A^3 - \partial^3 A^1 \\ \partial^2 A^0 - \partial^0 A^2 & \partial^2 A^1 - \partial^1 A^2 & 0 & \partial^2 A^3 - \partial^3 A^2 \\ \partial^3 A^0 - \partial^0 A^3 & \partial^3 A^1 - \partial^1 A^3 & \partial^3 A^2 - \partial^2 A^3 & 0 \end{pmatrix} \begin{pmatrix} 0 & \partial_0 A_1 - \partial_1 A_0 & \partial_0 A_2 - \partial_2 A_0 & \partial_0 A_3 - \partial_3 A_0 \\ \partial_1 A_0 - \partial_0 A_1 & 0 & \partial_1 A_2 - \partial_2 A_1 & \partial_1 A_3 - \partial_3 A_1 \\ \partial_2 A_0 - \partial_0 A_2 & \partial_2 A_1 - \partial_1 A_2 & 0 & \partial_2 A_3 - \partial_3 A_2 \\ \partial_3 A_0 - \partial_0 A_3 & \partial_3 A_1 - \partial_1 A_3 & \partial_3 A_2 - \partial_2 A_3 & 0 \end{pmatrix}\right)$$

$$\begin{aligned} = -(& \\ & + (\partial^0 A^1 - \partial^1 A^0)(\partial_1 A_0 - \partial_0 A_1) + (\partial^0 A^2 - \partial^2 A^0)(\partial_2 A_0 - \partial_0 A_2) + (\partial^0 A^3 - \partial^3 A^0)(\partial_3 A_0 - \partial_0 A_3) \\ & + (\partial^1 A^0 - \partial^0 A^1)(\partial_0 A_1 - \partial_1 A_0) + (\partial^1 A^2 - \partial^2 A^1)(\partial_2 A_1 - \partial_1 A_2) + (\partial^1 A^3 - \partial^3 A^1)(\partial_3 A_1 - \partial_1 A_3) \\ & + (\partial^2 A^0 - \partial^0 A^2)(\partial_0 A_2 - \partial_2 A_0) + (\partial^2 A^1 - \partial^1 A^2)(\partial_1 A_2 - \partial_2 A_1) + (\partial^2 A^3 - \partial^3 A^2)(\partial_3 A_2 - \partial_2 A_3) \\ & + (\partial^3 A^0 - \partial^0 A^3)(\partial_0 A_3 - \partial_3 A_0) + (\partial^3 A^1 - \partial^1 A^3)(\partial_1 A_3 - \partial_3 A_1) + (\partial^3 A^2 - \partial^2 A^3)(\partial_2 A_3 - \partial_3 A_2) \\ ) & \end{aligned}$$

$$\begin{aligned} = -\Bigg(& \\ & + \left(\frac{\partial A^1}{\partial x_0} - \frac{\partial A^0}{\partial x_1}\right)\left(\frac{\partial A_0}{\partial x^1} - \frac{\partial A_1}{\partial x^0}\right) + \left(\frac{\partial A^2}{\partial x_0} - \frac{\partial A^0}{\partial x_2}\right)\left(\frac{\partial A_0}{\partial x^2} - \frac{\partial A_2}{\partial x^0}\right) + \left(\frac{\partial A^3}{\partial x_0} - \frac{\partial A^0}{\partial x_3}\right)\left(\frac{\partial A_0}{\partial x^3} - \frac{\partial A_3}{\partial x^0}\right) \\ & + \left(\frac{\partial A^0}{\partial x_1} - \frac{\partial A^1}{\partial x_0}\right)\left(\frac{\partial A_1}{\partial x^0} - \frac{\partial A_0}{\partial x^1}\right) + \left(\frac{\partial A^2}{\partial x_1} - \frac{\partial A^1}{\partial x_2}\right)\left(\frac{\partial A_1}{\partial x^2} - \frac{\partial A_2}{\partial x^1}\right) + \left(\frac{\partial A^3}{\partial x_1} - \frac{\partial A^1}{\partial x_3}\right)\left(\frac{\partial A_1}{\partial x^3} - \frac{\partial A_3}{\partial x^1}\right) \\ & + \left(\frac{\partial A^0}{\partial x_2} - \frac{\partial A^2}{\partial x_0}\right)\left(\frac{\partial A_2}{\partial x^0} - \frac{\partial A_0}{\partial x^2}\right) + \left(\frac{\partial A^1}{\partial x_2} - \frac{\partial A^2}{\partial x_1}\right)\left(\frac{\partial A_2}{\partial x^1} - \frac{\partial A_1}{\partial x^2}\right) + \left(\frac{\partial A^3}{\partial x_2} - \frac{\partial A^2}{\partial x_3}\right)\left(\frac{\partial A_2}{\partial x^3} - \frac{\partial A_3}{\partial x^2}\right) \\ & + \left(\frac{\partial A^0}{\partial x_3} - \frac{\partial A^3}{\partial x_0}\right)\left(\frac{\partial A_3}{\partial x^0} - \frac{\partial A_0}{\partial x^3}\right) + \left(\frac{\partial A^1}{\partial x_3} - \frac{\partial A^3}{\partial x_1}\right)\left(\frac{\partial A_3}{\partial x^1} - \frac{\partial A_1}{\partial x^3}\right) + \left(\frac{\partial A^2}{\partial x_3} - \frac{\partial A^3}{\partial x_2}\right)\left(\frac{\partial A_3}{\partial x^2} - \frac{\partial A_2}{\partial x^3}\right) \end{aligned}$$

$$\Bigg)$$

$$= -\Bigg($$

$$+\left(\frac{\partial A^1}{\partial x^0}+\frac{\partial A^0}{\partial x^1}\right)\left(\frac{\partial A_0}{\partial x^1}-\frac{\partial A_1}{\partial x^0}\right)+\left(\frac{\partial A^2}{\partial x^0}+\frac{\partial A^0}{\partial x^2}\right)\left(\frac{\partial A_0}{\partial x^2}-\frac{\partial A_2}{\partial x^0}\right)+\left(\frac{\partial A^3}{\partial x^0}+\frac{\partial A^0}{\partial x^3}\right)\left(\frac{\partial A_0}{\partial x^3}-\frac{\partial A_3}{\partial x^0}\right)$$

$$+\left(-\frac{\partial A^0}{\partial x^1}-\frac{\partial A^1}{\partial x^0}\right)\left(\frac{\partial A_1}{\partial x^0}-\frac{\partial A_0}{\partial x^1}\right)-\left(\frac{\partial A^2}{\partial x^1}-\frac{\partial A^1}{\partial x^2}\right)\left(\frac{\partial A_1}{\partial x^2}-\frac{\partial A_2}{\partial x^1}\right)-\left(\frac{\partial A^3}{\partial x^1}-\frac{\partial A^1}{\partial x^3}\right)\left(\frac{\partial A_1}{\partial x^3}-\frac{\partial A_3}{\partial x^1}\right)$$

$$-\left(\frac{\partial A^0}{\partial x^2}+\frac{\partial A^2}{\partial x^0}\right)\left(\frac{\partial A_2}{\partial x^0}-\frac{\partial A_0}{\partial x^2}\right)-\left(\frac{\partial A^1}{\partial x^2}-\frac{\partial A^2}{\partial x^1}\right)\left(\frac{\partial A_2}{\partial x^1}-\frac{\partial A_1}{\partial x^2}\right)-\left(\frac{\partial A^3}{\partial x^2}-\frac{\partial A^2}{\partial x^3}\right)\left(\frac{\partial A_2}{\partial x^3}-\frac{\partial A_3}{\partial x^2}\right)$$

$$-\left(\frac{\partial A^0}{\partial x^3}+\frac{\partial A^3}{\partial x^0}\right)\left(\frac{\partial A_3}{\partial x^0}-\frac{\partial A_0}{\partial x^3}\right)-\left(\frac{\partial A^1}{\partial x^3}-\frac{\partial A^3}{\partial x^1}\right)\left(\frac{\partial A_3}{\partial x^1}-\frac{\partial A_1}{\partial x^3}\right)-\left(\frac{\partial A^2}{\partial x^3}-\frac{\partial A^3}{\partial x^2}\right)\left(\frac{\partial A_3}{\partial x^2}-\frac{\partial A_2}{\partial x^3}\right)$$

$$\Bigg)$$

$$= -\Bigg($$

$$+\left(\frac{\partial A^1}{\partial x^0}+\frac{\partial A^0}{\partial x^1}\right)\left(\frac{\partial A^0}{\partial x^1}+\frac{\partial A^1}{\partial x^0}\right)+\left(\frac{\partial A^2}{\partial x^0}+\frac{\partial A^0}{\partial x^2}\right)\left(\frac{\partial A^0}{\partial x^2}+\frac{\partial A^2}{\partial x^0}\right)+\left(\frac{\partial A^3}{\partial x^0}+\frac{\partial A^0}{\partial x^3}\right)\left(\frac{\partial A^0}{\partial x^3}+\frac{\partial A^3}{\partial x^0}\right)$$

$$+\left(\frac{\partial A^0}{\partial x^1}+\frac{\partial A^1}{\partial x^0}\right)\left(\frac{\partial A^1}{\partial x^0}+\frac{\partial A^0}{\partial x^1}\right)+\left(\frac{\partial A^2}{\partial x^1}-\frac{\partial A^1}{\partial x^2}\right)\left(\frac{\partial A^1}{\partial x^2}-\frac{\partial A^2}{\partial x^1}\right)+\left(\frac{\partial A^3}{\partial x^1}-\frac{\partial A^1}{\partial x^3}\right)\left(\frac{\partial A^1}{\partial x^3}-\frac{\partial A^3}{\partial x^1}\right)$$

$$+\left(\frac{\partial A^0}{\partial x^2}+\frac{\partial A^2}{\partial x^0}\right)\left(\frac{\partial A^2}{\partial x^0}+\frac{\partial A^0}{\partial x^2}\right)+\left(\frac{\partial A^1}{\partial x^2}-\frac{\partial A^2}{\partial x^1}\right)\left(\frac{\partial A^2}{\partial x^1}-\frac{\partial A^1}{\partial x^2}\right)+\left(\frac{\partial A^3}{\partial x^2}-\frac{\partial A^2}{\partial x^3}\right)\left(\frac{\partial A^2}{\partial x^3}-\frac{\partial A^3}{\partial x^2}\right)$$

$$+\left(\frac{\partial A^0}{\partial x^3}+\frac{\partial A^3}{\partial x^0}\right)\left(\frac{\partial A^3}{\partial x^0}+\frac{\partial A^0}{\partial x^3}\right)+\left(\frac{\partial A^1}{\partial x^3}-\frac{\partial A^3}{\partial x^1}\right)\left(\frac{\partial A^3}{\partial x^1}-\frac{\partial A^1}{\partial x^3}\right)+\left(\frac{\partial A^2}{\partial x^3}-\frac{\partial A^3}{\partial x^2}\right)\left(\frac{\partial A^3}{\partial x^2}-\frac{\partial A^2}{\partial x^3}\right)$$

$$\Bigg)$$

$$= -\Bigg($$

$$+\left(\frac{\partial A^0}{\partial x^1}+\frac{\partial A^1}{\partial x^0}\right)\left(\frac{\partial A^0}{\partial x^1}+\frac{\partial A^1}{\partial x^0}\right)+\left(\frac{\partial A^1}{\partial x^0}+\frac{\partial A^0}{\partial x^1}\right)\left(\frac{\partial A^1}{\partial x^0}+\frac{\partial A^0}{\partial x^1}\right)$$

$$+\left(\frac{\partial A^0}{\partial x^2}\,\frac{\partial A^2}{\partial x^0}\right)\left(\frac{\partial A^0}{\partial x^2}+\frac{\partial A^2}{\partial x^0}\right)+\left(\frac{\partial A^2}{\partial x^0}+\frac{\partial A^0}{\partial x^2}\right)\left(\frac{\partial A^2}{\partial x^0}+\frac{\partial A^0}{\partial x^2}\right)$$

$$+\left(\frac{\partial A^0}{\partial x^3}+\frac{\partial A^3}{\partial x^0}\right)\left(\frac{\partial A^0}{\partial x^3}+\frac{\partial A^3}{\partial x^0}\right)+\left(\frac{\partial A^3}{\partial x^0}+\frac{\partial A^0}{\partial x^3}\right)\left(\frac{\partial A^3}{\partial x^0}+\frac{\partial A^0}{\partial x^3}\right)$$

$$-\left(\frac{\partial A^1}{\partial x^2}-\frac{\partial A^2}{\partial x^1}\right)\left(\frac{\partial A^1}{\partial x^2}-\frac{\partial A^2}{\partial x^1}\right)-\left(\frac{\partial A^2}{\partial x^1}-\frac{\partial A^1}{\partial x^2}\right)\left(\frac{\partial A^2}{\partial x^1}-\frac{\partial A^1}{\partial x^2}\right)$$

$$-\left(\frac{\partial A^1}{\partial x^3}-\frac{\partial A^3}{\partial x^1}\right)\left(\frac{\partial A^1}{\partial x^3}-\frac{\partial A^3}{\partial x^1}\right)-\left(\frac{\partial A^3}{\partial x^1}-\frac{\partial A^1}{\partial x^3}\right)\left(\frac{\partial A^3}{\partial x^1}-\frac{\partial A^1}{\partial x^3}\right)$$

$$-\left(\frac{\partial A^2}{\partial x^3}-\frac{\partial A^3}{\partial x^2}\right)\left(\frac{\partial A^2}{\partial x^3}-\frac{\partial A^3}{\partial x^2}\right)-\left(\frac{\partial A^2}{\partial x^3}-\frac{\partial A^3}{\partial x^2}\right)\left(\frac{\partial A^2}{\partial x^3}-\frac{\partial A^3}{\partial x^2}\right)$$

$$\Bigg)$$

Compare this to the last two terms of the Lagrangian (with $\frac{1}{4\mu_0}$ divided out)

$$\left(\frac{\partial(\frac{\Phi}{c})}{\partial x_i}+\frac{\partial A_i}{\partial(ct)}\right)\left(\frac{\partial(\frac{\Phi}{c})}{\partial x_i}+\frac{\partial A_i}{\partial(ct)}\right)+\left(\frac{\partial A_i}{\partial(ct)}+\frac{\partial(\frac{\Phi}{c})}{\partial x_i}\right)\left(\frac{\partial A_i}{\partial(ct)}+\frac{\partial(\frac{\Phi}{c})}{\partial x_i}\right)-(\partial_i A_j-\partial_j A_i)(\partial_i A_j-\partial_j A_i)$$

$$=\left(\frac{\partial(\frac{\Phi}{c})}{\partial x}+\frac{\partial A_x}{\partial(ct)}\right)\left(\frac{\partial(\frac{\Phi}{c})}{\partial x}+\frac{\partial A_x}{\partial(ct)}\right)+\left(\frac{\partial A_x}{\partial(ct)}+\frac{\partial(\frac{\Phi}{c})}{\partial x}\right)\left(\frac{\partial A_x}{\partial(ct)}+\frac{\partial(\frac{\Phi}{c})}{\partial x}\right)$$
$$+\left(\frac{\partial(\frac{\Phi}{c})}{\partial y}+\frac{\partial A_y}{\partial(ct)}\right)\left(\frac{\partial(\frac{\Phi}{c})}{\partial y}+\frac{\partial A_y}{\partial(ct)}\right)+\left(\frac{\partial A_y}{\partial(ct)}+\frac{\partial(\frac{\Phi}{c})}{\partial y}\right)\left(\frac{\partial A_y}{\partial(ct)}+\frac{\partial(\frac{\Phi}{c})}{\partial y}\right)$$
$$+\left(\frac{\partial(\frac{\Phi}{c})}{\partial z}+\frac{\partial A_z}{\partial(ct)}\right)\left(\frac{\partial(\frac{\Phi}{c})}{\partial z}+\frac{\partial A_z}{\partial(ct)}\right)+\left(\frac{\partial A_z}{\partial(ct)}+\frac{\partial(\frac{\Phi}{c})}{\partial z}\right)\left(\frac{\partial A_z}{\partial(ct)}+\frac{\partial(\frac{\Phi}{c})}{\partial z}\right)$$

$$-\left(\frac{\partial A_x}{\partial y}-\frac{\partial A_y}{\partial x}\right)\left(\frac{\partial A_x}{\partial y}-\frac{\partial A_y}{\partial x}\right)-\left(\frac{\partial A_y}{\partial x}-\frac{\partial A_x}{\partial y}\right)\left(\frac{\partial A_y}{\partial x}-\frac{\partial A_x}{\partial y}\right)$$
$$-\left(\frac{\partial A_x}{\partial z}-\frac{\partial A_z}{\partial x}\right)\left(\frac{\partial A_x}{\partial z}-\frac{\partial A_z}{\partial x}\right)-\left(\frac{\partial A_z}{\partial x}-\frac{\partial A_x}{\partial z}\right)\left(\frac{\partial A_z}{\partial x}-\frac{\partial A_x}{\partial z}\right)$$
$$-\left(\frac{\partial A_y}{\partial z}-\frac{\partial A_z}{\partial y}\right)\left(\frac{\partial A_y}{\partial z}-\frac{\partial A_z}{\partial y}\right)-\left(\frac{\partial A_z}{\partial y}-\frac{\partial A_y}{\partial z}\right)\left(\frac{\partial A_z}{\partial y}-\frac{\partial A_y}{\partial z}\right)$$

Comparing the last lines, we see that they are the same if we make the correspondences

$$\begin{pmatrix}\frac{\Phi}{c}\\ A_x\\ A_y\\ A_z\end{pmatrix}=\begin{pmatrix}A^0\\ A^1\\ A^2\\ A^3\end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix}ct\\ x\\ y\\ z\end{pmatrix}=\begin{pmatrix}x^0\\ x^1\\ x^2\\ x^3\end{pmatrix}.$$

Finally then we can write the Lagrangian density for electromagnetism in a form that makes its relativistic symmetry manifest:

$$\mathcal{L}=-j^\mu A_\mu-\frac{1}{4\mu_0}F^{\mu\nu}F_{\mu\nu}$$

where

$$x^\mu=\begin{pmatrix}ct\\ x\\ y\\ z\end{pmatrix},\ A^\mu=\begin{pmatrix}\frac{\Phi}{c}\\ A_x\\ A_y\\ A_z\end{pmatrix},\ j^\mu=\begin{pmatrix}c\rho\\ j_x\\ j_y\\ j_z\end{pmatrix},\ F^{\mu\nu}=\partial^\mu A^\nu-\partial^\nu A^\mu,\ g_{\mu\nu}=\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix},$$

and the matrix in the relativistic transformation $x'^\mu = L^{\mu\cdot}_{\cdot\nu'}\, x^{\nu'}$ satisfies $L^{\mu\cdot}_{\cdot\nu'}\, g_{\mu\nu}\, L^{\nu\cdot}_{\cdot\mu'} = g_{\nu'\mu'}$.

To see that this encompasses the usual relativistic transformation $x' = \gamma(x - v\,t)$, $t' = \gamma(t - \frac{v\,x}{c^2})$ where $\gamma = \frac{1}{\sqrt{1-(v/c)^2}}$, note that it can be written using the (2+2)×(2+2) matrix $\underset{\approx}{L} = \begin{pmatrix} \begin{pmatrix} \gamma & -\gamma\,v \\ -\gamma\,v & \gamma \end{pmatrix} & \underset{\approx}{0} \\ \underset{\approx}{0} & \underset{\approx}{1} \end{pmatrix}$ as

$$\begin{pmatrix} \begin{pmatrix} ct' \\ x' \end{pmatrix} \\ \begin{pmatrix} y' \\ z' \end{pmatrix} \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} \gamma & -\gamma\,\frac{v}{c} \\ -\gamma\,\frac{v}{c} & \gamma \end{pmatrix} & \underset{\approx}{0} \\ \underset{\approx}{0} & \underset{\approx}{1} \end{pmatrix} \begin{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} \\ \begin{pmatrix} y \\ z \end{pmatrix} \end{pmatrix}$$

and the transformation matrix satisfies

$$\begin{pmatrix} \begin{pmatrix} \gamma & -\gamma\,\frac{v}{c} \\ -\gamma\,\frac{v}{c} & \gamma \end{pmatrix} & \underset{\approx}{0} \\ \underset{\approx}{0} & \underset{\approx}{1} \end{pmatrix}^T \begin{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} & \underset{\approx}{0} \\ \underset{\approx}{0} & -\underset{\approx}{1} \end{pmatrix} \begin{pmatrix} \begin{pmatrix} \gamma & -\gamma\,\frac{v}{c} \\ -\gamma\,\frac{v}{c} & \gamma \end{pmatrix} & \underset{\approx}{0} \\ \underset{\approx}{0} & \underset{\approx}{1} \end{pmatrix}$$

$$= \begin{pmatrix} \begin{pmatrix} \gamma & -\gamma\,\frac{v}{c} \\ -\gamma\,\frac{v}{c} & \gamma \end{pmatrix} & \underset{\approx}{0} \\ \underset{\approx}{0} & \underset{\approx}{1} \end{pmatrix} \begin{pmatrix} \begin{pmatrix} \gamma & -\gamma\,\frac{v}{c} \\ \gamma\,\frac{v}{c} & -\gamma \end{pmatrix} & \underset{\approx}{0} \\ \underset{\approx}{0} & -\underset{\approx}{1} \end{pmatrix}$$

$$= \begin{pmatrix} \begin{pmatrix} \gamma^2(1-(v/c)^2) & 0 \\ 0 & -\gamma^2(1-(v/c)^2) \end{pmatrix} & \underset{\approx}{0} \\ \underset{\approx}{0} & -\underset{\approx}{1} \end{pmatrix}$$

$$= \begin{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} & \underset{\approx}{0} \\ \underset{\approx}{0} & -\underset{\approx}{1} \end{pmatrix}.$$

This important special transformation is called a *boost* along the *x* axis. It is parametrized by the velocity v with $|v| < c$ that a fixed point in the primed frame is seen to move in the unprimed frame. An alternative and frequently useful parametrization is in terms of the *rapidity* $\mathcal{R}$ where

$$\cosh\mathcal{R} = \gamma,\ \ \sinh\mathcal{R} = \frac{v}{c}\,\gamma,\ \ \text{and}\ \tanh\mathcal{R} = \frac{v}{c},\ \ \text{with}\ -\infty < \mathcal{R} < \infty.$$

It is interesting to notice that this simple boost is represented by a symmetric matrix whereas the general Lorentz transformation is represented by a non-symmetric one. This suggests that there are attributes of the Lorentz transformation that are not evident from the simple boost case, and in fact there are! However, I will not explore them here.

The Lorentz transformation also clearly encompasses ordinary three–dimensional rotations which are represented by 3×3 orthogonal matrices, $\underset{\approx}{Q}$, satisfying $\underset{\approx}{Q}\,\underset{\approx}{Q}^T = \underset{\approx}{1}$. The calculation is very easy in a $(1+3)\times(1+3)$ matrix format

$$\begin{pmatrix} 1 & \underset{\sim}{0}^T \\ \underset{\sim}{0} & \underset{\approx}{Q} \end{pmatrix}^T \begin{pmatrix} 1 & \underset{\sim}{0}^T \\ \underset{\sim}{0} & -\underset{\approx}{1} \end{pmatrix} \begin{pmatrix} 1 & \underset{\sim}{0}^T \\ \underset{\sim}{0} & \underset{\approx}{Q} \end{pmatrix} = \begin{pmatrix} 1 & \underset{\sim}{0}^T \\ \underset{\sim}{0} & \underset{\approx}{Q}^T \end{pmatrix} \begin{pmatrix} 1 & \underset{\sim}{0}^T \\ \underset{\sim}{0} & -\underset{\approx}{Q} \end{pmatrix} = \begin{pmatrix} 1 & \underset{\sim}{0}^T \\ \underset{\sim}{0} & -\underset{\approx}{1} \end{pmatrix}$$

which shows that ordinary rotations are Lorentz transformations.

Of course, many other interesting things about Lorentz transformations can be demonstrated with this formalism, but I will only discuss the transformation of electromagnetic quantities under a boost. We already know about the charge-current density 4-vectors and the potential 4-vector. To learn about the electric and magnetic field transformations, note that

$$F^{\mu\nu} = \begin{pmatrix} 0 & \partial^0 A^1 - \partial^1 A^0 & \partial^0 A^2 - \partial^2 A^0 & \partial^0 A^3 - \partial^3 A^0 \\ -(\partial^0 A^1 - \partial^1 A^0) & 0 & \partial^1 A^2 - \partial^2 A^1 & \partial^1 A^3 - \partial^3 A^1 \\ -(\partial^0 A^2 - \partial^2 A^0) & -(\partial^1 A^2 - \partial^2 A^1) & 0 & \partial^2 A^3 - \partial^3 A^2 \\ -(\partial^0 A^3 - \partial^3 A^0) & -(\partial^1 A^3 - \partial^3 A^1) & -(\partial^2 A^3 - \partial^3 A^2) & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & \frac{1}{c}\left(\frac{\partial A_x}{\partial t} + \frac{\partial \Phi}{\partial x}\right) & \frac{1}{c}\left(\frac{\partial A_y}{\partial t} + \frac{\partial \Phi}{\partial y}\right) & \frac{1}{c}\left(\frac{\partial A_z}{\partial t} + \frac{\partial \Phi}{\partial z}\right) \\ -\frac{1}{c}\left(\frac{\partial A_x}{\partial t} + \frac{\partial \Phi}{\partial x}\right) & 0 & -\frac{\partial A_y}{\partial x} + \frac{\partial A_x}{\partial y} & -\frac{\partial A_z}{\partial x} + \frac{\partial A_x}{\partial z} \\ -\frac{1}{c}\left(\frac{\partial A_y}{\partial t} + \frac{\partial \Phi}{\partial y}\right) & -\left(-\frac{\partial A_y}{\partial x} + \frac{\partial A_x}{\partial y}\right) & 0 & -\frac{\partial A_z}{\partial y} + \frac{\partial A_y}{\partial z} \\ -\frac{1}{c}\left(\frac{\partial A_z}{\partial t} + \frac{\partial \Phi}{\partial z}\right) & -\left(-\frac{\partial A_z}{\partial x} + \frac{\partial A_x}{\partial z}\right) & -\left(-\frac{\partial A_z}{\partial y} + \frac{\partial A_y}{\partial z}\right) & 0 \end{pmatrix}.$$

So

$$F^{\mu\nu} = \begin{pmatrix} 0 & -\frac{1}{c} E_x & -\frac{1}{c} E_y & -\frac{1}{c} E_z \\ \frac{1}{c} E_x & 0 & -B_z & B_y \\ \frac{1}{c} E_y & B_z & 0 & -B_x \\ \frac{1}{c} E_z & -B_y & B_x & 0 \end{pmatrix}.$$

Similarly, the doubly covariant field matrix is

$$F_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & -\frac{1}{c} E_x & -\frac{1}{c} E_y & -\frac{1}{c} E_z \\ \frac{1}{c} E_x & 0 & -B_z & B_y \\ \frac{1}{c} E_y & B_z & 0 & -B_x \\ \frac{1}{c} E_z & -B_y & B_x & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{c} E_x & \frac{1}{c} E_y & \frac{1}{c} E_z \\ \frac{1}{c} E_x & 0 & B_z & -B_y \\ \frac{1}{c} E_y & -B_z & 0 & B_x \\ \frac{1}{c} E_z & B_y & -B_x & 0 \end{pmatrix}.$$

So

$$F_{\mu\nu} = \begin{pmatrix} 0 & \frac{1}{c} E_x & \frac{1}{c} E_y & \frac{1}{c} E_z \\ -\frac{1}{c} E_x & 0 & -B_z & B_y \\ -\frac{1}{c} E_y & B_z & 0 & -B_x \\ -\frac{1}{c} E_z & -B_y & B_x & 0 \end{pmatrix}$$

Thus the Lagrangian can be expressed as

$$\mathcal{L} = -j^\mu A_\mu + \frac{1}{4\mu_0} F^{\nu\mu} F_{\mu\nu}$$

$$= -j^\mu A_\mu + \frac{1}{4\mu_0} \mathrm{Tr}\left(\begin{pmatrix} 0 & -\frac{1}{c} E_x & -\frac{1}{c} E_y & -\frac{1}{c} E_z \\ \frac{1}{c} E_x & 0 & -B_z & B_y \\ \frac{1}{c} E_y & B_z & 0 & -B_x \\ \frac{1}{c} E_z & -B_y & B_x & 0 \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{c} E_x & \frac{1}{c} E_y & \frac{1}{c} E_z \\ -\frac{1}{c} E_x & 0 & -B_z & B_y \\ -\frac{1}{c} E_y & B_z & 0 & -B_x \\ -\frac{1}{c} E_z & -B_y & B_x & 0 \end{pmatrix} \right)$$

$$= -j^\mu A_\mu + \frac{2}{4\mu_0} \left(\frac{1}{c^2} E^2 - B^2 \right)$$

$$= -\rho \Phi + \vec{j} \cdot \vec{A} + \frac{1}{2} \varepsilon_0 E^2 - \frac{1}{2\mu_0} B^2.$$

This is a cross check on the calculation since we have known this form of the Lagrangian density for a long time. It also shows the interesting fact that the form $E^2 - c^2 B^2$ is a relativistic invariant. Thus if in any frame we have $|\vec{E}| = c|\vec{B}|$, this is true in all frames. Similarly, if $|\vec{E}| > c|\vec{B}|$ at a space-time point in one frame, then at the corresponding space-time point in any other frame, the relations is still true. We will find a second invariant which is bilinear in the electric and magnetic fields shortly. But first work out the transformation of the fields resulting from a boost along the x axis.We have in general

$$F'^{\mu\nu} = L^{\mu\cdot}_{\cdot\mu'} L^{\nu\cdot}_{\cdot\nu'} F^{\mu'\nu'} \quad \Rightarrow \quad \underset{\approx}{F}' = \underset{\approx}{L} \underset{\approx}{F} \underset{\approx}{L}^T$$

so for a simple boost

$$\begin{pmatrix} 0 & -\frac{1}{c} E'_x & -\frac{1}{c} E'_y & -\frac{1}{c} E'_z \\ \frac{1}{c} E'_x & 0 & -B'_z & B'_y \\ \frac{1}{c} E'_y & B'_z & 0 & -B'_x \\ \frac{1}{c} E'_z & -B'_y & B'_x & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \cosh\mathcal{R} & -\sinh\mathcal{R} & 0 & 0 \\ -\sinh\mathcal{R} & \cosh\mathcal{R} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & -\frac{1}{c} E_x & -\frac{1}{c} E_y & -\frac{1}{c} E_z \\ \frac{1}{c} E_x & 0 & -B_z & B_y \\ \frac{1}{c} E_y & B_z & 0 & -B_x \\ \frac{1}{c} E_z & -B_y & B_x & 0 \end{pmatrix} \begin{pmatrix} \cosh\mathcal{R} & -\sinh\mathcal{R} & 0 & 0 \\ -\sinh\mathcal{R} & \cosh\mathcal{R} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \cosh\mathcal{R} & -\sinh\mathcal{R} & 0 & 0 \\ -\sinh\mathcal{R} & \cosh\mathcal{R} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{c} E_x \sinh\mathcal{R} & -\frac{1}{c} E_x \cosh\mathcal{R} & -\frac{1}{c} E_y & -\frac{1}{c} E_z \\ \frac{1}{c} E_x \cosh\mathcal{R} & -\frac{1}{c} E_x \sinh\mathcal{R} & -B_z & B_y \\ \frac{1}{c} E_y \cosh\mathcal{R} - B_z \sinh\mathcal{R} & -\frac{1}{c} E_y \sinh\mathcal{R} + B_z \cosh\mathcal{R} & 0 & -B_x \\ \frac{1}{c} E_z \cosh\mathcal{R} + B_y \sinh\mathcal{R} & -\frac{1}{c} E_z \sinh\mathcal{R} - B_y \cosh\mathcal{R} & B_x & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -\frac{1}{c} E_x(\cosh^2\mathcal{R} - \sinh^2\mathcal{R}) & -(\frac{1}{c} E_y \cosh\mathcal{R} - B_z \sinh\mathcal{R}) & -\frac{1}{c} (E_z \cosh\mathcal{R} + B_y \sinh\mathcal{R}) \\ \frac{1}{c} E_x(\cosh^2\mathcal{R} - \sinh^2\mathcal{R}) & 0 & \frac{1}{c} E_y \sinh\mathcal{R} - B_z \cosh\mathcal{R} & \frac{1}{c} E_z \sinh\mathcal{R} + B_y \cosh\mathcal{R} \\ \frac{1}{c} E_y \cosh\mathcal{R} - B_z \sinh\mathcal{R} & -\frac{1}{c} E_y \sinh\mathcal{R} + B_z \cosh\mathcal{R} & 0 & -B_x \\ \frac{1}{c} E_z \cosh\mathcal{R} + B_y \sinh\mathcal{R} & -(\frac{1}{c} E_z \sinh\mathcal{R} + B_y \cosh\mathcal{R}) & B_x & 0 \end{pmatrix}$$

Thus, for a relativistic boost along the x axis, the various coordinates and electromagnetic quantities transform as follows:

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} ct\cosh\mathcal{R} - x'\sinh\mathcal{R} \\ -ct\sinh\mathcal{R} + x'\cosh\mathcal{R} \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma(ct - \frac{v}{c}x) \\ \gamma(-vt + x) \\ y \\ z \end{pmatrix},$$

$$\begin{pmatrix} E'_x(x'^\mu) \\ E'_y(x'^\mu) \\ E'_z(x'^\mu) \end{pmatrix} = \begin{pmatrix} E_x(x^\mu) \\ E_y(x^\mu)\cosh\mathcal{R} - cB_z(x^\mu)\sinh\mathcal{R} \\ E_z(x^\mu)\cosh\mathcal{R} + cB_y(x^\mu)\sinh\mathcal{R} \end{pmatrix} = \begin{pmatrix} E_x(x^\mu) \\ \gamma(E_y(x^\mu) - vB_z(x^\mu)) \\ \gamma(E_z(x^\mu) + vB_y(x^\mu)) \end{pmatrix},$$

$$\begin{pmatrix} B'_x(x'^\mu) \\ B'_y(x'^\mu) \\ B'_z(x'^\mu) \end{pmatrix} = \begin{pmatrix} B_x(x^\mu) \\ \frac{1}{c}E_z(x^\mu)\sinh\mathcal{R} + B_y(x^\mu)\cosh\mathcal{R} \\ -\frac{1}{c}E_y(x^\mu)\sinh\mathcal{R} + B_z(x^\mu)\cosh\mathcal{R} \end{pmatrix} = \begin{pmatrix} B_x(x^\mu) \\ \gamma(B_y(x^\mu) + \frac{v}{c^2}E_z(x^\mu)) \\ \gamma(B_z(x^\mu) - \frac{v}{c^2}E_y(x^\mu)) \end{pmatrix},$$

$$\begin{pmatrix} \frac{\Phi'(x'^\mu)}{c} \\ A'_x(x'^\mu) \\ A'_y(x'^\mu) \\ A'_z(x'^\mu) \end{pmatrix} = \begin{pmatrix} \frac{\Phi(x^\mu)}{c}\cosh\mathcal{R} - A_x(x^\mu)\sinh\mathcal{R} \\ -\frac{\Phi(x^\mu)}{c}\sinh\mathcal{R} + A_x(x^\mu)\cosh\mathcal{R} \\ A_y(x^\mu) \\ A_z(x^\mu) \end{pmatrix} = \begin{pmatrix} \gamma(\frac{\Phi(x^\mu)}{c} - \frac{v}{c}A_x(x^\mu)) \\ \gamma(-\frac{\Phi(x^\mu)}{c} + A_x(x^\mu)) \\ A_y(x^\mu) \\ A_z(x^\mu) \end{pmatrix},$$

and
$$\begin{pmatrix} c\rho'(x'^\mu) \\ j'_x(x'^\mu) \\ j'_y(x'^\mu) \\ j'_z(x'^\mu) \end{pmatrix} = \begin{pmatrix} c\rho(x^\mu)\cosh\mathcal{R} - j_x(x^\mu)\sinh\mathcal{R} \\ -c\rho(x^\mu)\sinh\mathcal{R} + j_x(x^\mu)\cosh\mathcal{R} \\ j_y(x^\mu) \\ j_z(x^\mu) \end{pmatrix} = \begin{pmatrix} \gamma(c\rho(x^\mu) - \frac{v}{c}j_x(x^\mu)) \\ \gamma(-c\rho(x^\mu) + j_x(x^\mu)) \\ j_y(x^\mu) \\ j_z(x^\mu) \end{pmatrix}.$$

It is easy to check from this that $\frac{1}{c^2}E^2 - B^2$ is invariant, i.e., $\frac{1}{c^2}E'^2 - B'^2 = \frac{1}{c^2}E^2 - B^2$. Another invariant that is easily seen from this is $\vec{E}\cdot\vec{B}$. This says that if the electric and magnetic fields are orthogonal in one frame, then they are so in all frames. Similarly, if either the electric or magnetic field vanishes in one frame, then either one of them vanishes in another frame or they are orthogonal in it.

■ The adjoint field tensor $\mathcal{F}^{\mu\nu}$

Another way to get the invariance of this dot product is to introduce the totally antisymmetric tensor $\epsilon^{\mu\nu\lambda\sigma} = \begin{pmatrix} 1 & \text{if } \mu\nu\lambda\sigma \text{ is an even permutation of } 0123 \\ -1 & \text{if } \mu\nu\lambda\sigma \text{ is an odd permutation of } 0123 \\ 0 & \text{otherwise} \end{pmatrix}$. The definition of the determinant of a 4×4 matrix can be given in terms of this tensor by the $\det(M)\,\epsilon^{\mu\nu\lambda\sigma} = M^{\mu\cdot}_{\cdot\mu'}\,M^{\nu\cdot}_{\cdot\nu'}\,M^{\lambda\cdot}_{\cdot\lambda'}\,M^{\sigma\cdot}_{\cdot\sigma'}\,\epsilon^{\mu'\nu'\lambda'\sigma'}$. We can first use this to verify that $\epsilon^{\mu\nu\lambda\sigma}$ takes on the same values (to within a sign) in all reference frames if we transform it like a tensor with the rank and variance shown. We have

$$\epsilon'^{\mu\nu\lambda\sigma} = L^{\mu\cdot}_{\cdot\mu'}\,L^{\nu\cdot}_{\cdot\nu'}\,L^{\lambda\cdot}_{\cdot\lambda'}\,L^{\sigma\cdot}_{\cdot\sigma'}\,\epsilon^{\mu'\nu'\lambda'\sigma'} = \left(\det \underset{\approx}{L}\right)\epsilon^{\mu\nu\lambda\sigma}.$$

But from the fundamental condition on the transformation matrix, namely $\underset{\approx}{L}^T\,\underset{\approx}{g}\,\underset{\approx}{L} = \underset{\approx}{g}$, we immediately obtain by taking determinates and remembering

a) the determinate of a product of matrices is the product of their determinates,
b) and the determinate of the transpose of a matrix is the same as that of the matrix itself,

$$\det\left(\underset{\approx}{L}^T \underset{\approx}{g} \underset{\approx}{L}\right) = \det(\underset{\approx}{g}) \quad \Rightarrow \quad \det(\underset{\approx}{L})^2 = 1 \quad \Rightarrow \quad \det(\underset{\approx}{L}) = \pm 1.$$

Thus we get $\epsilon'^{\mu\nu\lambda\sigma} = \pm\epsilon^{\mu'\nu'\lambda'\sigma'}$, with the plus sign if $\det(\underset{\approx}{L}) = 1$ (*proper* Lorentz transformation) and the minus sign in case $\det(\underset{\approx}{L}) = -1$. As in three dimensions a tensorial quantity whose transformation law involves the determinant of the transformation matrix, we call $\epsilon^{\mu\nu\lambda\sigma}$ 4th rank, pseudo–contravariant tensor.

We can also use the determinate to calculate the covariant form of this tensor:

$$\epsilon_{\mu\nu\lambda\sigma} = g_{\mu\mu'}\, g_{\nu\nu'}\, g_{\lambda\lambda'}\, g_{\sigma\sigma'}\, \epsilon^{\mu'\nu'\lambda'\sigma'} = \det(\underset{\approx}{g})\, \epsilon^{\mu\nu\lambda\sigma} = -\epsilon^{\mu\nu\lambda\sigma}.$$

Thus we have the rule $\epsilon_{\mu\nu\lambda\sigma} = \begin{pmatrix} -1 & \text{if } \mu\nu\lambda\sigma \text{ is an even permutation of } 0123 \\ 1 & \text{if } \mu\nu\lambda\sigma \text{ is an odd permutation of } 0123 \\ 0 & \text{otherwise} \end{pmatrix}$. Finally, we can use the totally antisymmetric 4th rank tensor to define the adjoint electromagnetic field tensor (note the $\frac{1}{2}$ because both ϵ and F are antisymmetric)

$$\mathcal{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\lambda\sigma} F_{\lambda\sigma} = \frac{1}{2}\begin{pmatrix} 0 & \epsilon^{01\lambda\sigma} F_{\lambda\sigma} & \epsilon^{02\lambda\sigma} F_{\lambda\sigma} & \epsilon^{03\lambda\sigma} F_{\lambda\sigma} \\ -\epsilon^{01\lambda\sigma} F_{\lambda\sigma} & 0 & \epsilon^{12\lambda\sigma} F_{\lambda\sigma} & \epsilon^{13\lambda\sigma} F_{\lambda\sigma} \\ -\epsilon^{02\lambda\sigma} F_{\lambda\sigma} & -\epsilon^{12\lambda\sigma} F_{\lambda\sigma} & 0 & \epsilon^{23\lambda\sigma} F_{\lambda\sigma} \\ -\epsilon^{03\lambda\sigma} F_{\lambda\sigma} & -\epsilon^{13\lambda\sigma} F_{\lambda\sigma} & -\epsilon^{23\lambda\sigma} F_{\lambda\sigma} & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & F_{23} & F_{31} & F_{12} \\ -F_{23} & 0 & F_{03} & F_{20} \\ -F_{31} & -F_{03} & 0 & F_{01} \\ -F_{12} & -F_{20} & -F_{10} & 0 \end{pmatrix}.$$

So

$$\mathcal{F}^{\mu\nu} = \begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & \frac{1}{c}E_z & -\frac{1}{c}E_y \\ B_y & -\frac{1}{c}E_z & 0 & \frac{1}{c}E_x \\ B_z & \frac{1}{c}E_y & -\frac{1}{c}E_x & 0 \end{pmatrix}$$

It then clear that nothing new arises by considering the invariant $\mathcal{F}^{\mu\nu}\mathcal{F}_{\mu\nu}$. However we do get something new from

$$\mathcal{F}^{\mu\nu} F_{\mu\nu} = \mathrm{Tr}\left(\begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & \frac{1}{c} E_z & -\frac{1}{c} E_y \\ B_y & -\frac{1}{c} E_z & 0 & \frac{1}{c} E_x \\ B_z & \frac{1}{c} E_y & -\frac{1}{c} E_x & 0 \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{c} E_x & \frac{1}{c} E_y & \frac{1}{c} E_z \\ -\frac{1}{c} E_x & 0 & -B_z & B_y \\ -\frac{1}{c} E_y & B_z & 0 & -B_x \\ -\frac{1}{c} E_z & -B_y & B_x & 0 \end{pmatrix}\right)$$

$$= \frac{4}{c} \vec{E} \cdot \vec{B}.$$

Consequently, as shown directly from the explicit boost looked at above, we see that in general $\vec{E} \cdot \vec{B}$ is a relativistic invariant.

■ Maxwell's equations in manifestly relativistic covariant form

It is easy to get Maxwell's field equations from the Lagrangian density:

$$\mathcal{L} = -j^\mu A_\mu - \frac{1}{4\mu_0} F^{\mu\nu} F_{\mu\nu} = -j^\mu A_\mu - \frac{1}{4\mu_0} (\partial^\mu A^\nu - \partial^\nu A^\mu) g_{\mu\mu'} g_{\nu\nu'} (\partial^{\mu'} A^{\nu'} - \partial^{\nu'} A^{\mu'}).$$

The field equations are $\partial^\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial^\mu A^\nu)}\right) - \frac{\partial \mathcal{L}}{\partial A^\nu} = 0$. They give

$$-\frac{1}{\mu_0} \partial^\mu (\partial_\mu A_\nu - \partial_\nu A_\mu) - (-j_\nu) = 0 \qquad \Rightarrow \quad \partial^\mu F_{\mu\nu} = \mu_0 j_\nu.$$

It is also easy to see from the antisymmetry of $\epsilon^{\mu\nu\lambda\sigma}$ and the symmetry of both $\partial_\mu \partial_\lambda$ and $\partial_\mu \partial_\sigma$ that

$$\partial_\mu \mathcal{F}^{\mu\nu} = \frac{1}{2} \partial_\mu \epsilon^{\mu\nu\lambda\sigma} F_{\lambda\sigma} = \frac{1}{2} \epsilon^{\mu\nu\lambda\sigma} (\partial_\mu \partial_\lambda A_\sigma - \partial_\mu \partial_\sigma A_\lambda) = 0$$

We therefore have Maxwell's equations in manifestly covariant form:

$$\partial^\mu F_{\mu\nu} = \mu_0 j_\nu$$
$$\partial^\mu \mathcal{F}_{\mu\nu} = 0.$$

To make a final check, convert these equations to normal vectorial notation.
First, Coulomb's law from the $\nu = 0$ equation:

$$\partial^\mu F_{\mu 0} = \mu_0 j_0 \quad \Rightarrow \quad \partial^1 F_{10} + \partial^2 F_{20} + \partial^3 F_{30} = \frac{\partial}{\partial x_1}\left(-\frac{1}{c} E_x\right) + \frac{\partial}{\partial x_2}\left(-\frac{1}{c} E_y\right) + \frac{\partial}{\partial x_3}\left(-\frac{1}{c} E_z\right) = \mu_0 c\rho$$

or

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = \mu_0 c^2 \rho = \frac{\rho}{\varepsilon_0} \quad \text{or}$$

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\varepsilon_0}$$

Next the Ampere-Maxwell law from the three equations with $\nu = 1, 2, 3$:

$$\begin{pmatrix} \partial^\mu F_{\mu 1} = \mu_0\, j_1 \\ \partial^\mu F_{\mu 2} = \mu_0\, j_2 \\ \partial^\mu F_{\mu 3} = \mu_0\, j_3 \end{pmatrix} \Rightarrow \begin{pmatrix} \partial^0 F_{01} + \partial^2 F_{21} + \partial^3 F_{31} = -\mu_0\, j_x \\ \partial^0 F_{02} + \partial^1 F_{12} + \partial^3 F_{32} = -\mu_0\, j_y \\ \partial^0 F_{03} + \partial^1 F_{13} + \partial^2 F_{23} = -\mu_0\, j_z \end{pmatrix} \Rightarrow \begin{pmatrix} \frac{\partial}{\partial x_0} F_{01} + \frac{\partial}{\partial x_2} F_{21} + \frac{\partial}{\partial x_3} F_{31} = -\mu_0\, j_x \\ \frac{\partial}{\partial x_0} F_{02} + \frac{\partial}{\partial x_1} F_{12} + \frac{\partial}{\partial x_3} F_{32} = -\mu_0\, j_y \\ \frac{\partial}{\partial x_0} F_{03} + \frac{\partial}{\partial x_1} F_{13} + \frac{\partial}{\partial x_2} F_{23} = -\mu_0\, j_z \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} \frac{\partial}{\partial x_0} (\frac{1}{c} E_x) + \frac{\partial}{\partial x_2} B_z - \frac{\partial}{\partial x_3} B_y = -\mu_0\, j_x \\ \frac{\partial}{\partial x_0} (\frac{1}{c} E_y) - \frac{\partial}{\partial x_1} B_z + \frac{\partial}{\partial x_3} B_x = -\mu_0\, j_y \\ \frac{\partial}{\partial x_0} (\frac{1}{c} E_z) + \frac{\partial}{\partial x_1} B_y - \frac{\partial}{\partial x_2} B_x = -\mu_0\, j_z \end{pmatrix} \Rightarrow \begin{pmatrix} \frac{\partial}{\partial (ct)} (\frac{1}{c} E_x) - \frac{\partial B_z}{\partial y} + \frac{\partial B_y}{\partial z} = -\mu_0\, j_x \\ \frac{\partial}{\partial (ct)} (\frac{1}{c} E_y) + \frac{\partial B_z}{\partial x} - \frac{\partial B_x}{\partial z} = -\mu_0\, j_y \\ \frac{\partial}{\partial (ct)} (\frac{1}{c} E_z) - \frac{\partial B_y}{\partial x} + \frac{\partial B_x}{\partial y} = -\mu_0\, j_z \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} (\vec{\nabla} \times \vec{B})_x = \mu_0\, j_x + \frac{1}{c^2} \frac{\partial E_x}{\partial t} \\ (\vec{\nabla} \times \vec{B})_y = \mu_0\, j_y + \frac{1}{c^2} \frac{\partial E_y}{\partial t} \\ (\vec{\nabla} \times \vec{B})_z = \mu_0\, j_z + \frac{1}{c^2} \frac{\partial E_z}{\partial t} \end{pmatrix} \Rightarrow$$

$$\vec{\nabla} \times \vec{B} = \mu_0\, \vec{j} + \mu_0\, \varepsilon_0 \frac{\partial \vec{E}}{\partial t}$$

To get the other usual Maxwell equations for Faraday's law and $\vec{\nabla} \cdot \vec{B} = 0$ we use the adjoint equation, $\partial_\mu \mathcal{F}^{\mu\nu} = 0$.

First the equation from the $\nu = 0$ case:

$$\partial_1 \mathcal{F}^{10} + \partial_2 \mathcal{F}^{20} + \partial_3 \mathcal{F}^{30} = 0 \Rightarrow \frac{\partial}{\partial x} \mathcal{F}^{10} + \frac{\partial}{\partial y} \mathcal{F}^{20} + \frac{\partial}{\partial y} \mathcal{F}^{30} = 0 \Rightarrow \frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial y} = 0$$

or the familiar

$$\vec{\nabla} \cdot \vec{B} = 0.$$

Finally, we get Faraday's law from the $\nu = 1, 2, 3$ cases.

$$\begin{pmatrix} \partial_0 \mathcal{F}^{01} + \partial_2 \mathcal{F}^{21} + \partial_3 \mathcal{F}^{31} = 0 \\ \partial_0 \mathcal{F}^{02} + \partial_1 \mathcal{F}^{12} + \partial_3 \mathcal{F}^{32} = 0 \\ \partial_0 \mathcal{F}^{03} + \partial_1 \mathcal{F}^{13} + \partial_2 \mathcal{F}^{23} = 0 \end{pmatrix} \Rightarrow$$

$$\begin{pmatrix} \frac{\partial}{\partial (ct)} \mathcal{F}^{01} + \frac{\partial}{\partial y} \mathcal{F}^{21} + \frac{\partial}{\partial z} \mathcal{F}^{31} = 0 \\ \frac{\partial}{\partial (ct)} \mathcal{F}^{02} + \frac{\partial}{\partial x} \mathcal{F}^{12} + \frac{\partial}{\partial z} \mathcal{F}^{32} = 0 \\ \frac{\partial}{\partial (ct)} \mathcal{F}^{03} + \frac{\partial}{\partial x} \mathcal{F}^{13} + \frac{\partial}{\partial y} \mathcal{F}^{23} = 0 \end{pmatrix} \Rightarrow \begin{pmatrix} -\frac{\partial B_x}{\partial (ct)} - \frac{\partial}{\partial y} (\frac{1}{c} E_z) + \frac{\partial}{\partial z} (\frac{1}{c} E_y) = 0 \\ -\frac{\partial B_y}{\partial (ct)} + \frac{\partial}{\partial x} (\frac{1}{c} E_z) - \frac{\partial}{\partial z} (\frac{1}{c} E_x) = 0 \\ -\frac{\partial B_z}{\partial (ct)} - \frac{\partial}{\partial x} (\frac{1}{c} E_y) + \frac{\partial}{\partial y} (\frac{1}{c} E_x) = 0 \end{pmatrix} \Rightarrow \begin{pmatrix} -\frac{\partial B_x}{\partial t} = \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} \\ -\frac{\partial B_y}{\partial t} = -\frac{\partial E_z}{\partial x} + \frac{\partial E_x}{\partial z} \\ -\frac{\partial B_z}{\partial t} = \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} \end{pmatrix} \Rightarrow$$

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$$

Finally then, we have seen that electromagnetism is covariant under relativistic transformations. Of course, it was Einstein who showed this in a wonderfully elegant manner that elevated this fundamental electromagnetic symmetry to a general symmetry that (so far as we know) *all* physical laws must exhibit.