Chapter 4

Electromagnetic Fields

Once it is recognized that the seventeenth century theory of mechanics and the nineteenth century theory of electromagnetism involve two different invariance groups, the Galilean group on the one hand and the Lorentz group on the other hand, it follows that to preserve the principle of relativity one must give up one theory or the other. After one elects to give up one of the two theories, one must explain why that theory appeared to work so well, and one should look for situations which put the old theory to the test.

Einstein retained the Maxwell theory of electromagnetism, and replaced Newtonian mechanics with a Lorentz invariant theory. The apparent successes of the Newtonian theory were attributed to the fact that one never considered bodies traveling at speeds approaching that of light. In today's world of high energy particle accelerators, the situation is entirely different. No particle physicist would ever contemplate describing his work-a-day world in any other terms but those of a Lorentz invariant theory. Thus, soon after the notions of quantum theory were first introduced, theorists were striving to construct Lorentz invariant quantum theories, and this effort led to the development of quantum field theories, most notably quantum electrodynamics.

Since we shall adopt the point of view that Maxwell's theory of electromagnetism is, except for quantization, an acceptable modern theory, we should demonstrate explicitly the Lorentz invariant nature of this particular theory. In Einstein's original paper this was done at least for simple Lorentz transformations in the x-direction. We shall attempt to do this for the entire Lorentz group, and in the process we shall display various modern ways of displaying the Maxwell field equations.

In describing the Maxwell theory we shall have to select one of many equivalent formulations, differing from one another only in how the basic fields are defined and how the units are chosen. Throughout these notes I shall employ the Gaussian units as defined, for example, in *Classical Electrodynamics* by J. D. Jackson.

In the Gaussian units, Maxwell's equations assume the following form in a vacuum, i.e., where there is no polarization or magnetization of the medium:

Homogeneous equations These equations embody the Faraday law and the non-existence of magnetic monopoles.

$$\begin{aligned}\vec{\nabla} \times \vec{E} &= -\frac{1}{c}\frac{\partial \vec{H}}{\partial t} \\ \vec{\nabla} \cdot \vec{H} &= 0\end{aligned}$$

Inhomogeneous equations These equations embody the Ampère circuital law (with Maxwell's *displacement current*) and the so-called Gauss theorem.

$$\begin{aligned}\vec{\nabla} \times \vec{H} &= \frac{4\pi}{c}\vec{\jmath} + \frac{1}{c}\frac{\partial \vec{E}}{\partial t} \\ \vec{\nabla} \cdot \vec{E} &= 4\pi\rho\end{aligned}$$

Moreover, the force per cubic centimeter exerted upon a small body with charge density ρ and current density $\vec{\jmath}$ is given by

$$\vec{F} = \rho\vec{E} + \frac{1}{c}\vec{\jmath} \times \vec{H}.$$

These equations have been expressed in a vectorial language, which clearly displays the $O(3)$ rotation invariance of the Maxwell equations. As they are written, the $O(3,1)$ Lorentz invariance of the equations is far from obvious. We wish to rectify this situation, without having the Lorentz invariance emerge mysteriously out of thin air. How can this objective be achieved?

The trick is to consider the variational principle from which the Maxwell equations may be derived, for then you only have to establish the invariance of a single quantity, the Lagrangian density.

Now, if the homogeneous Maxwell equations hold in a simply connected region of space, they permit the introduction of a scalar potential ϕ and a vector potential $\vec{A}$ such that

$$\begin{aligned}\vec{H} &= \vec{\nabla} \times \vec{A}, \\ \vec{E} &= -\vec{\nabla}\phi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t}.\end{aligned}$$

The remaining inhomogeneous Maxwell equations can be expressed as second order differential equations involving the scalar and vector potentials.

Suppose you want to derive these second order differential equations via a variational principle. The Lagrangian density $\mathcal{L}$ will involve ϕ and $\vec{A}$ and the first partial derivatives of these potentials with respect to the coordinates x, y, z and the time t. If one initially ignores the terms involving the charge and current densities, the equations are homogeneous second order equations in the potentials. This strongly suggests a Lagrangian density that is quadratic in the first derivatives of the potentials. On the other hand, the obvious $O(3)$ invariance of the Maxwell equations suggests that the Lagrangian density should be an $O(3)$ scalar quantity. Such considerations as these permit one to identify without great difficulty a Lagrangian density from which the equations (with $\rho = 0$ and $\vec{\jmath} = \vec{0}$) can be derived; namely,

$$\mathcal{L} = \frac{1}{8\pi}(E^2 - H^2).$$

The numerical factor was chosen so that the Maxwell stress tensor associated with this Lagrangian would have components of appropriate magnitudes, a point we shall not go into here, except to remark that in the Gaussian units the energy density of an electromagnetic field is equal to $\frac{1}{8\pi}(E^2 + H^2)$.

In carrying out the variational principle, it is important to remember that $\vec{E}$ and $\vec{H}$ are given in terms of first derivatives of the potentials ϕ and $\vec{A}$.

Let us consider the variation $\phi \rightarrow \phi + \delta\phi$, where $\delta\phi$ vanishes at temporal and spatial infinity, and where $\delta\phi$ is *small.* $\vec{H}$ does not depend upon ϕ or its derivatives, and $\vec{E}$ depends only upon the spatial derivatives of ϕ. For the variation being considered, the change in the *action*

$$I = \int\int\int\int \mathcal{L}\; dx\; dy\; dz\; dt$$

is simply

$$\delta I = -\frac{1}{4\pi}\int\int\int\int \vec{E}\cdot\vec{\nabla}(\delta\phi)\; dx\; dy\; dz\; dt.$$

Performing the usual integration by parts, and using the boundary conditions on $\delta\phi$, we obtain

$$\delta I = \frac{1}{4\pi}\int\int\int\int \vec{\nabla}\cdot\vec{E}\;\delta\phi\; dx\; dy\; dz\; dt.$$

Since the small variation $\delta\phi(x, y, z, t)$ is completely arbitrary (except for the boundary conditions at infinity), we conclude that $\delta I = 0$ implies

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

This is one of the Maxwell equations (for $\rho = 0$).

Next consider the variation $A_x \to A_x + \delta A_x$. In this case H_y and H_z depend on some of the spatial derivatives of A_x and E_x depends on the time derivative of A_x. Proceeding in a way similar to that which we used in the case of ϕ, one can show that $\delta I = 0$ now implies

$$\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} = \frac{1}{c}\,\frac{\partial E_x}{\partial t}\ .$$

This is easily seen to be one component of the Maxwell equation

$$\vec{\nabla} \times \vec{H} = \frac{1}{c}\,\frac{\partial \vec{E}}{\partial t}\ ,$$

valid when $\vec{\jmath} = \vec{0}$.

Having derived the Maxwell equations in the absence of charges and currents, we may now search for a generalization of the Lagrangian density

$$\mathcal{L} = \frac{1}{8\pi}(E^2 - H^2)$$

from which the complete Maxwell equations may be derived. A little study reveals that one possible Lagrangian density is

$$\mathcal{L} = \frac{1}{8\pi}(E^2 - H^2) + \frac{1}{c}\vec{\jmath} \cdot \vec{A} - \rho\phi\ .$$

If we can show that the action is a Lorentz invariant quantity, we shall have a proof that the Lorentz transformations are an invariance group of the Maxwell field equations. In the first place we notice that the four-dimensional volume element *dx dy dz dt* is Lorentz invariant, just as the three-dimensional volume element *dx dy dz* is rotationally invariant. It remains, therefore, to show that the Lagrangian density $\mathcal{L}$ is Lorentz invariant.

We shall show that the terms in the Lagrangian density which yielded the Maxwell equations for $\rho = 0$ and $\vec{\jmath} = \vec{0}$ and the terms which involve $\vec{\jmath}$ and ρ are separately Lorentz invariant. Consider the latter terms; namely, $\frac{1}{c}\vec{\jmath} \cdot \vec{A} - \rho\phi$. We already know how x, y, z, t behave under a Lorentz transformation. How can we use this knowledge to infer how the various electromagnetic fields and potentials transform?

Consider $\vec{\jmath}$ and ρ. How must they transform under a Lorentz transformation? An important clue is the so-called *continuity equation*, which results from conservation of charge, a significant feature of the Maxwell theory.

If you consider a surface S enclosing a volume V, and $\vec{n}$ is an outwardly directed unit vector, then conservation of charge implies that

$$\oint_S \vec{\jmath} \cdot \vec{n}\, dS = -\frac{d}{dt} \int_V \rho\, dV ,$$

since whatever charge flows out through the surface S must diminish the charge left inside the volume V. Using the divergence theorem, the left side can be rewritten as a volume integral. Since the equation must hold regardless of what closed surface S we select, it follows that

$$\vec{\nabla} \cdot \vec{\jmath} + \frac{\partial \rho}{\partial t} = 0 .$$

This is the continuity equation.

Now, we already know that the differential operator

$$\boldsymbol{\nabla} := \mathbf{e}_1 \frac{\partial}{\partial x} + \mathbf{e}_2 \frac{\partial}{\partial y} + \mathbf{e}_3 \frac{\partial}{\partial z} - \mathbf{e}_4 \frac{1}{c} \frac{\partial}{\partial t}$$

is invariant under a Lorentz transformation; i.e., the various unit vectors $\mathbf{e}_j$ transform, and the various partial derivatives transform, but this combination remains unchanged.

Therefore, if $\vec{\jmath}$ and ρ are joined into a single object

$$\mathbf{J} := j_x \mathbf{e}_1 + j_y \mathbf{e}_2 + j_z \mathbf{e}_3 + c\rho \mathbf{e}_4 ,$$

then the continuity equation will assume the form

$$\boldsymbol{\nabla} \cdot \mathbf{J} = 0 .$$

This equation will be manifestly Lorentz invariant, if we assume that $\mathbf{J}$ like $\boldsymbol{\nabla}$ is Lorentz invariant.

In this way we surmise that Lorentz invariance of the Maxwell theory necessitates that $\vec{\jmath}$ and ρ be combined into an invariant object $\mathbf{J}$, and this tells us how the various components j_x, j_y, j_z and ρ transform under Lorentz transformations of the tetrad $\mathbf{e}_j$.

Now that we know how charge density and current density transform, we can use that knowledge to deduce how the potentials $\vec{A}$ and ϕ must transform if the Lagrangian density is to remain invariant. You want to express the terms

$$\frac{1}{c} \vec{\jmath} \cdot \vec{A} - \rho\phi$$

as a four-dimensional scalar product, one factor of which is the object **J**. Clearly, $\vec{A}$ and ϕ must be united in the combination

$$\mathbf{A} := A_x\mathbf{e}_1 + A_y\mathbf{e}_2 + A_z\mathbf{e}_3 + \phi\mathbf{e}_4 \ ,$$

for then we can express the terms in the Lagrangian density in the manifestly invariant form $\frac{1}{c}\mathbf{J}\cdot\mathbf{A}$.

It remains to show that the remaining terms $\frac{1}{8\pi}(E^2 - H^2)$ in the Lagrangian density can be written in a manifestly Lorentz invariant form.

You may have noticed that so far we have been able to choose freely how we put together diverse three-dimensional objects to form four-dimensional objects. Essentially, we made things work out correctly! Now, however, the moment of truth is close at hand. The remaining terms in the Lagrangian density involve only $\vec{E}$ and $\vec{H}$, and these are expressible entirely in terms of $\vec{A}$ and ϕ and various partial derivatives, the transformation properties of all of which have already been determined. We have no further freedom to fool around! Either the additional terms will be Lorentz invariant, or they will not be!

We have seen that the quantities

$$\begin{aligned} \boldsymbol{\nabla} &:= \mathbf{e}_1\frac{\partial}{\partial x} + \mathbf{e}_2\frac{\partial}{\partial y} + \mathbf{e}_3\frac{\partial}{\partial z} - \mathbf{e}_4\frac{1}{c}\frac{\partial}{\partial t} \ , \\ \mathbf{A} &:= A_x\mathbf{e}_1 + A_y\mathbf{e}_2 + A_z\mathbf{e}_3 + \phi\mathbf{e}_4 \ , \end{aligned}$$

must be invariant objects under Lorentz transformations, if the Maxwell theory is to be in fact Lorentz invariant. Therefore, the *tensor product* of these objects must also be an invariant. The tensor product is just the dyadic product $\boldsymbol{\nabla}\mathbf{A}$. (You may have encountered rank two tensors or dyadics in classical mechanics when you studied the moment of inertia, or in electromagnetic theory when you studied quadrupole moments. Here we are just extending the notion from three to four dimensions.)

The skew symmetric part of the tensor product $\boldsymbol{\nabla}\mathbf{A}$ is what is of interest to us now. There are *six* independent components of this invariant object, not four as in the case of a four-vector.

One way of looking at such invariant six-component objects is as follows: We shall introduce a new type of product, denoted by the symbol $\wedge$. Like the three-dimensional vector product, it is an *antisymmetric* product. Let us introduce the symbol

$$\mathbf{F} := \boldsymbol{\nabla}\wedge\mathbf{A} \ .$$

Upon substituting $\boldsymbol{\nabla}$ and $\mathbf{A}$, and using the antisymmetry of the wedge product, we obtain

$$\begin{aligned}\mathbf{F} &= \left(\frac{\partial A_z}{\partial y}-\frac{\partial A_y}{\partial z}\right)\mathbf{e}_2\wedge\mathbf{e}_3+\left(\frac{\partial A_x}{\partial z}-\frac{\partial A_z}{\partial x}\right)\mathbf{e}_3\wedge\mathbf{e}_1\\ &\quad+\left(\frac{\partial A_y}{\partial x}-\frac{\partial A_x}{\partial y}\right)\mathbf{e}_1\wedge\mathbf{e}_2+\left(\frac{\partial \phi}{\partial x}+\frac{1}{c}\frac{\partial A_x}{\partial t}\right)\mathbf{e}_1\wedge\mathbf{e}_4\\ &\quad+\left(\frac{\partial \phi}{\partial y}+\frac{1}{c}\frac{\partial A_y}{\partial t}\right)\mathbf{e}_2\wedge\mathbf{e}_4+\left(\frac{\partial \phi}{\partial z}+\frac{1}{c}\frac{\partial A_z}{\partial t}\right)\mathbf{e}_3\wedge\mathbf{e}_4\ .\end{aligned}$$

You will immediately recognize that the six coefficients are, in fact, Cartesian components of the electric and magnetic fields; namely,

$$\mathbf{F}=H_x\,\mathbf{e}_2\wedge\mathbf{e}_3+H_y\,\mathbf{e}_3\wedge\mathbf{e}_1+H_z\,\mathbf{e}_1\wedge\mathbf{e}_2-E_x\,\mathbf{e}_1\wedge\mathbf{e}_4-E_y\,\mathbf{e}_2\wedge\mathbf{e}_4-E_z\,\mathbf{e}_3\wedge\mathbf{e}_4.$$

Before Einstein the electric field $\vec{E}$ and the magnetic field $\vec{H}$ were quite independent notions, but we now see that they are different aspects of a single invariant object $\mathbf{F}$. Under a Lorentz transformation the $\mathbf{e}_j$ change, and the electric and magnetic field components change, but the combination $\mathbf{F}$ remains invariant.

Can the remaining terms $\frac{1}{8\pi}(E^2-H^2)$ in the Lagrangian density be expressed in terms of $\mathbf{F}$? Since the former quantity is quadratic in $\vec{E}$ and $\vec{H}$, it is clear that a quadratic expression in $\mathbf{F}$ is required. How can we form a quadratic invariant from the *bivector* $\mathbf{F}$?

The concept of an inner product (scalar product) can easily be extended from vectors to bivectors (or even further). One simply defines

$$(\mathbf{A}\wedge\mathbf{B})\cdot(\mathbf{C}\wedge\mathbf{D}):=(\mathbf{A}\cdot\mathbf{C})(\mathbf{B}\cdot\mathbf{D})-(\mathbf{A}\cdot\mathbf{D})(\mathbf{B}\cdot\mathbf{C})\ .$$

Then it can be seen that the six basic bivectors,

$$\mathbf{e}_2\wedge\mathbf{e}_3,\mathbf{e}_3\wedge\mathbf{e}_1,\mathbf{e}_1\wedge\mathbf{e}_2,\mathbf{e}_1\wedge\mathbf{e}_4,\mathbf{e}_2\wedge\mathbf{e}_4,\mathbf{e}_3\wedge\mathbf{e}_4,$$

are mutually orthogonal, in the sense that their inner products vanish, while

$$(\mathbf{e}_2\wedge\mathbf{e}_3)\cdot(\mathbf{e}_2\wedge\mathbf{e}_3)=(\mathbf{e}_3\wedge\mathbf{e}_1)\cdot(\mathbf{e}_3\wedge\mathbf{e}_1)=(\mathbf{e}_1\wedge\mathbf{e}_2)\cdot(\mathbf{e}_1\wedge\mathbf{e}_2)=1$$

and

$$(\mathbf{e}_1\wedge\mathbf{e}_4)\cdot(\mathbf{e}_1\wedge\mathbf{e}_4)=(\mathbf{e}_2\wedge\mathbf{e}_4)\cdot(\mathbf{e}_2\wedge\mathbf{e}_4)=(\mathbf{e}_3\wedge\mathbf{e}_4)\cdot(\mathbf{e}_3\wedge\mathbf{e}_4)=-1\ .$$

In conclusion, it is possible to express the Lagrangian density in the manifestly Lorentz invariant form

$$\mathcal{L} = \frac{1}{8\pi}\,\mathbf{F}\cdot\mathbf{F} + \frac{1}{c}\,\mathbf{J}\cdot\mathbf{A}\ .$$

We have now accomplished our goal of showing that the Maxwell theory is in fact a Lorentz invariant theory, and in the course of demonstrating this we have identified certain geometrical objects in spacetime. These include the following:

$$\begin{aligned}
\mathbf{J} &= j_x\mathbf{e}_1 + j_y\mathbf{e}_2 + j_z\mathbf{e}_3 + c\rho\mathbf{e}_4\ ,\\
\mathbf{A} &= A_x\mathbf{e}_1 + A_y\mathbf{e}_2 + A_z\mathbf{e}_3 + \phi\mathbf{e}_4\ ,\\
\mathbf{F} &= H_x\mathbf{e}_2\wedge\mathbf{e}_3 + H_y\mathbf{e}_3\wedge\mathbf{e}_1 + H_z\mathbf{e}_1\wedge\mathbf{e}_2\\
&\quad -E_x\mathbf{e}_1\wedge\mathbf{e}_4 - E_y\mathbf{e}_2\wedge\mathbf{e}_4 - E_z\mathbf{e}_3\wedge\mathbf{e}_4\ .
\end{aligned}$$

Under a Lorentz transformation geometrical objects such as these remain unchanged, but their components are transformed. Thus, for example, under the simple Lorentz transformation

$$\begin{aligned}
\mathbf{e}'_1 &= \cosh\psi\mathbf{e}_1 + \sinh\psi\mathbf{e}_4\ ,\\
\mathbf{e}'_2 &= \mathbf{e}_2\ ,\\
\mathbf{e}'_3 &= \mathbf{e}_3\ ,\\
\mathbf{e}'_4 &= \sinh\psi\mathbf{e}_1 + \cosh\psi\mathbf{e}_4\ ,
\end{aligned}$$

we have

$$\begin{aligned}
\mathbf{e}'_2\wedge\mathbf{e}'_3 &= \mathbf{e}_2\wedge\mathbf{e}_3\ ,\\
\mathbf{e}'_3\wedge\mathbf{e}'_1 &= \cosh\psi\mathbf{e}_3\wedge\mathbf{e}_1 + \sinh\psi\mathbf{e}_3\wedge\mathbf{e}_4\ ,\\
\mathbf{e}'_1\wedge\mathbf{e}'_2 &= \cosh\psi\mathbf{e}_1\wedge\mathbf{e}_2 - \sinh\psi\mathbf{e}_2\wedge\mathbf{e}_4\ ,\\
\mathbf{e}'_1\wedge\mathbf{e}'_4 &= \mathbf{e}_1\wedge\mathbf{e}_4\ ,\\
\mathbf{e}'_2\wedge\mathbf{e}'_4 &= -\sinh\psi\mathbf{e}_1\wedge\mathbf{e}_2 + \cosh\psi\mathbf{e}_2\wedge\mathbf{e}_4\ ,\\
\mathbf{e}'_3\wedge\mathbf{e}'_4 &= \sinh\psi\mathbf{e}_3\wedge\mathbf{e}_1 + \cosh\psi\mathbf{e}_3\wedge\mathbf{e}_4\ ,
\end{aligned}$$

and hence (from the invariance of the bivector **F**)

$$\begin{aligned}
H_x &= H'_x, & E_x &= E'_x,\\
H_y &= H'_y\cosh\psi - E'_z\sinh\psi, & E_y &= E'_y\cosh\psi + H'_z\sinh\psi,\\
H_z &= H'_z\cosh\psi + E'_y\sinh\psi, & E_z &= E'_z\cosh\psi - H'_y\sinh\psi.
\end{aligned}$$

Solving for the primed fields and substituting for ψ in terms of the velocity v of the primed frame relative to the unprimed frame, we have

$$\begin{aligned}
H'_x &= H_x, & E'_x &= E_x,\\
H'_y &= \frac{H_y+vE_z/c}{\sqrt{1-v^2/c^2}}, & E'_y &= \frac{E_y-vH_z/c}{\sqrt{1-v^2/c^2}},\\
H'_z &= \frac{H_z-vE_y/c}{\sqrt{1-v^2/c^2}}, & E'_z &= \frac{E_z+vH_y/c}{\sqrt{1-v^2/c^2}}.
\end{aligned}$$

Notice that the components of $\vec{E}$ and $\vec{H}$ in the direction of relative motion are unaffected by the transformation, but the transverse components do change.

As a simple application, consider a conducting bar oriented along the y-axis, moving at velocity v in the x-direction. Suppose also that there is a magnetic field H in the z-direction. This is a simple sophomore physics problem, but let's solve it using a Lorentz transformation.

Recalling that the primed frame is moving with velocity v in the x-direction relative to the unprimed frame, let us select the unprimed frame to be the laboratory frame and the primed frame to be a rest with respect to the conducting bar. Assuming that $|v|$ is small compared to the speed of light c, our transformation formulas yield the following expression for the electric field in the bar's frame:

$$E'_x = 0, \quad E'_y = -vH/c, \quad E'_z = 0.$$

This is in agreement with the sophomore physics result (except that in sophomore physics it is likely that MKS units would have been used). There is an effective electric field along the negative y-direction in the bar. It is this electric field that is responsible for the force upon charges within the bar.

When you think about it, this is quite remarkable. In prerelativity physics, one thought of the terms in the force law $\vec{F} = \rho\vec{E} + \frac{1}{c}\vec{j} \times \vec{H}$ as being completely independent of one another. In Einstein's view these two terms are intimately coupled to one another. In this way, Einstein has given us a profound new insight into a classical theory.