§1.6 Tensors

A more typical example of a tensor is the **electromagnetic field strength tensor**. We all know that the electromagnetic fields are made up of the electric field vector E_i and the magnetic field vector B_i. (Remember that we use Latin indices for spacelike components 1, 2, 3.) Actually these are only "vectors" under rotations in space, not under the full Lorentz group. In fact they are components of a (0, 2) tensor $F_{\mu\nu}$, defined by

$$F_{\mu\nu} = \begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E_3 & B_2 & -B_1 & 0 \end{pmatrix} = -F_{\nu\mu}. \tag{1.69}$$

From this point of view it is easy to transform the electromagnetic fields in one reference frame to those in another, by application of (1.63). The unifying power of the tensor formalism is evident: rather than a collection of two vectors whose relationship and transformation properties are rather mysterious, we have a single tensor field to describe all of electromagnetism. (On the other hand, don't get carried away; sometimes it's more convenient to work in a single coordinate system using the electric and magnetic field vectors.)

§1.8

MAXWELL'S EQUATIONS

We have now accumulated enough tensor know-how to illustrate some of these concepts using actual physics. Specifically, we will examine **Maxwell's equations** of electrodynamics. In 19th-century notation, these are

$$\begin{aligned} \nabla \times \mathbf{B} - \partial_t \mathbf{E} &= \mathbf{J} \\ \nabla \cdot \mathbf{E} &= \rho \\ \nabla \times \mathbf{E} + \partial_t \mathbf{B} &= 0 \\ \nabla \cdot \mathbf{B} &= 0. \end{aligned} \tag{1.92}$$

Here, $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic field 3-vectors, $\mathbf{J}$ is the current, ρ is the charge density, and $\nabla\times$ and $\nabla\cdot$ are the conventional curl and divergence. These equations are invariant under Lorentz transformations, of course; that's how the whole business got started. But they don't look obviously invariant; our tensor notation can fix that. Let's begin by writing these equations in component notation,

$$\begin{aligned} \tilde{\epsilon}^{ijk}\partial_j B_k - \partial_0 E^i &= J^i \\ \partial_i E^i &= J^0 \\ \tilde{\epsilon}^{ijk}\partial_j E_k + \partial_0 B^i &= 0 \\ \partial_i B^i &= 0. \end{aligned} \tag{1.93}$$

In these expressions, spatial indices have been raised and lowered with abandon, without any attempt to keep straight where the metric appears, because δ_{ij} is the metric on flat 3-space, with δ^{ij} its inverse (they are equal as matrices). We can therefore raise and lower indices at will, since the components don't change. Meanwhile, the three-dimensional Levi–Civita symbol $\tilde{\epsilon}^{ijk}$ is defined just as the four-dimensional one, although with one fewer index (normalized so that $\tilde{\epsilon}^{123} = \tilde{\epsilon}_{123} = 1$). We have replaced the charge density by J^0; this is legitimate because the density and current together form the **current 4-vector**, $J^\mu = (\rho, J^x, J^y, J^z)$.

From (1.93), and the definition (1.69) of the field strength tensor $F_{\mu\nu}$, it is easy to get a completely tensorial 20th-century version of Maxwell's equations. Begin by noting that we can express the field strength with upper indices as

$$
\begin{aligned}
F^{0i} &= E^i \\
F^{ij} &= \tilde{\epsilon}^{ijk} B_k.
\end{aligned}
\tag{1.94}
$$

To check this, note for example that $F^{01} = \eta^{00}\eta^{11} F_{01}$ and $F^{12} = \tilde{\epsilon}^{123} B_3$. Then the first two equations in (1.93) become

$$
\begin{aligned}
\partial_j F^{ij} - \partial_0 F^{0i} &= J^i \\
\partial_i F^{0i} &= J^0.
\end{aligned}
\tag{1.95}
$$

Using the antisymmetry of $F^{\mu\nu}$, we see that these may be combined into the single tensor equation

$$
\partial_\mu F^{\nu\mu} = J^\nu. \tag{1.96}
$$

A similar line of reasoning, which is left as an exercise, reveals that the third and fourth equations in (1.93) can be written

$$
\partial_{[\mu} F_{\nu\lambda]} = 0. \tag{1.97}
$$

It's simple to verify that the antisymmetry of $F_{\mu\nu}$ implies that (1.97) can be equivalently expressed as

$$
\partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = 0. \tag{1.98}
$$

The four traditional Maxwell equations are thus replaced by two, vividly demonstrating the economy of tensor notation. More importantly, however, both sides of equations (1.96) and (1.97) manifestly transform as tensors; therefore, if they are true in one inertial frame, they must be true in any Lorentz-transformed frame. This is why tensors are so useful in relativity—we often want to express relationships without recourse to any reference frame, and the quantities on each side of an equation must transform in the same way under changes of coordinates. As a matter of jargon, we will sometimes refer to quantities written in terms of tensors as **covariant** (which has nothing to do with "covariant" as opposed to "contravariant"). Thus, we say that (1.96) and (1.97) together serve as the covariant form of Maxwell's equations, while (1.92) or (1.93) are noncovariant.

§1.10 Classical Field Theory

A slightly more elaborate example of a field theory is provided by electromagnetism. We mentioned that the relevant field is the **vector potential** A_μ; the timelike component A_0 can be identified with the electrostatic potential Φ, and the spacelike components with the traditional vector potential $\mathbf{A}$ (in terms of which the magnetic field is given by $\mathbf{B} = \nabla \times \mathbf{A}$). The field strength tensor, with components given by (1.69), is related to the vector potential by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{1.157}$$

From this definition we see that the field strength tensor has the important property of **gauge invariance**: when we perform a **gauge transformation** on the vector potential,

$$A_\mu \to A_\mu + \partial_\mu \lambda(x), \tag{1.158}$$

the field strength tensor is left unchanged:

$$F_{\mu\nu} \to F_{\mu\nu} + \partial_\mu \partial_\nu \lambda - \partial_\nu \partial_\mu \lambda = F_{\mu\nu}. \tag{1.159}$$

The last equality follows from the fact that partial derivatives commute, $\partial_\mu \partial_\nu = \partial_\nu \partial_\mu$. Gauge invariance is a symmetry that is fundamental to our understanding of electromagnetism, and all observable quantities must be gauge-invariant. Thus, while the dynamical field of the theory (with respect to which we vary the action to derive equations of motion) is A_μ, physical quantities will generally be expressed in terms of $F_{\mu\nu}$.

We already know that the dynamical equations of electromagnetism are Maxwell's equations, (1.96) and (1.97). Given the definition of the field stregth tensor in terms of the vector potential, (1.97) is actually automatic:

$$\partial_{[\mu} F_{\nu\sigma]} = \partial_{[\mu} \partial_\nu A_{\sigma]} - \partial_{[\mu} \partial_\sigma A_{\nu]} = 0, \tag{1.160}$$

again because partial derivatives commute. On the other hand, (1.96) is equivalent to Euler–Lagrange equations of the form

$$\frac{\partial \mathcal{L}}{\partial A_\nu} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \right) = 0, \tag{1.161}$$

if we presciently choose the Lagrangian to be

$$\mathcal{L} = -\tfrac{1}{4} F_{\mu\nu} F^{\mu\nu} + A_\mu J^\mu. \tag{1.162}$$

For this choice, the first term in the Euler–Lagrange equation is straightforward:

$$\frac{\partial \mathcal{L}}{\partial A_\nu} = \delta^\nu_\mu J^\mu = J^\nu. \tag{1.163}$$

The second term is tricker. First we write $F_{\mu\nu}F^{\mu\nu}$ as

$$F_{\mu\nu}F^{\mu\nu} = F_{\alpha\beta}F^{\alpha\beta} = \eta^{\alpha\rho}\eta^{\beta\sigma}F_{\alpha\beta}F_{\rho\sigma}. \tag{1.164}$$

We want to work with lower indices on $F_{\mu\nu}$, since we are differentiating with respect to $\partial_\mu A_\nu$, which has lower indices. Likewise we change the dummy indices on $F_{\mu\nu}F^{\mu\nu}$, since we want to have different indices on the thing being differentiated and the thing we are differentiating with respect to. Once you get familiar with this stuff it will become second nature and you won't need nearly so many steps. This lets us write

$$\frac{\partial(F_{\alpha\beta}F^{\alpha\beta})}{\partial(\partial_\mu A_\nu)} = \eta^{\alpha\rho}\eta^{\beta\sigma}\left[\left(\frac{\partial F_{\alpha\beta}}{\partial(\partial_\mu A_\nu)}\right)F_{\rho\sigma} + F_{\alpha\beta}\left(\frac{\partial F_{\rho\sigma}}{\partial(\partial_\mu A_\nu)}\right)\right]. \tag{1.165}$$

Then, since $F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha$, we have

$$\frac{\partial F_{\alpha\beta}}{\partial(\partial_\mu A_\nu)} = \delta^\mu_\alpha \delta^\nu_\beta - \delta^\mu_\beta \delta^\nu_\alpha. \tag{1.166}$$

Combining (1.166) with (1.165) yields

$$\begin{aligned}
\frac{\partial(F_{\alpha\beta}F^{\alpha\beta})}{\partial(\partial_\mu A_\nu)} &= \eta^{\alpha\rho}\eta^{\beta\sigma}\left[(\delta^\mu_\alpha \delta^\nu_\beta - \delta^\mu_\beta \delta^\nu_\alpha)F_{\rho\sigma} + (\delta^\mu_\rho \delta^\nu_\sigma - \delta^\mu_\sigma \delta^\nu_\rho)F_{\alpha\beta}\right] \\
&= (\eta^{\mu\rho}\eta^{\nu\sigma} - \eta^{\nu\rho}\eta^{\mu\sigma})F_{\rho\sigma} + (\eta^{\alpha\mu}\eta^{\beta\nu} - \eta^{\alpha\nu}\eta^{\beta\mu})F_{\alpha\beta} \\
&= F^{\mu\nu} - F^{\nu\mu} + F^{\mu\nu} - F^{\nu\mu} \\
&= 4F^{\mu\nu},
\end{aligned} \tag{1.167}$$

so

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} = -F^{\mu\nu}. \tag{1.168}$$

Then sticking (1.163) and (1.168) into (1.161) yields precisely (1.96):

$$\partial_\mu F^{\nu\mu} = J^\nu. \tag{1.169}$$

Note that we switched the order of the indices on $F^{\mu\nu}$ in order to save ourselves from an unpleasant minus sign.

You may wonder what the purpose of introducing a Lagrangian formulation is, if we were able to invent the equations of motion before we ever knew the Lagrangian (as Maxwell did for his equations). There are a number of reasons,

starting with the basic simplicity of positing a single scalar function of spacetime, the Lagrange density, rather than a number of (perhaps tensor-valued) equations of motion. Another reason is the ease with which symmetries are implemented; demanding that the action be invariant under a symmetry ensures that the dynamics respects the symmetry as well. Finally, as we will see in Chapter 4, the action leads via a direct procedure (involving varying with respect to the metric itself) to a unique energy-momentum tensor. Applying this procedure to (1.148) leads straight to the energy-momentum tensor for a scalar field theory,

$$T^{\mu\nu}_{\text{scalar}} = \eta^{\mu\lambda}\eta^{\nu\sigma}\partial_\lambda\phi\partial_\sigma\phi - \eta^{\mu\nu}\left[\tfrac{1}{2}\eta^{\lambda\sigma}\partial_\lambda\phi\partial_\sigma\phi + V(\phi)\right]. \tag{1.170}$$

Similarly, from (1.162) we can derive the energy-momentum tensor for electromagnetism,

$$T^{\mu\nu}_{\text{EM}} = F^{\mu\lambda}F^{\nu}{}_{\lambda} - \tfrac{1}{4}\eta^{\mu\nu}F^{\lambda\sigma}F_{\lambda\sigma}. \tag{1.171}$$

Using the appropriate equations of motion, you can show that these energy-momentum tensors are conserved, $\partial_\mu T^{\mu\nu} = 0$ (and will be asked to do so in the Exercises).

§1.11 Exercises

10. Using the tensor transformation law applied to $F_{\mu\nu}$, show how the electric and magnetic field 3-vectors **E** and **B** transform under

(a) a rotation about the y-axis,

(b) a boost along the z-axis.

11. Verify that (1.98) is indeed equivalent to (1.97), and that they are both equivalent to the last two equations in (1.93).

12. Consider the two field theories we explicitly discussed, Maxwell's electromagnetism (let $J^\mu = 0$) and the scalar field theory defined by (1.148).

(a) Express the components of the energy-momentum tensors of each theory in three-vector notation, using the divergence, gradient, curl, electric, and magnetic fields, and an overdot to denote time derivatives.

(b) Using the equations of motion, verify (in any notation you like) that the energy-momentum tensors are conserved.

13. Consider adding to the Lagrangian for electromagnetism an additional term of the form $\mathcal{L}' = \tilde{\epsilon}_{\mu\nu\rho\sigma}F^{\mu\nu}F^{\rho\sigma}$.

(a) Express $\mathcal{L}'$ in terms of **E** and **B**.

(b) Show that including $\mathcal{L}'$ does not affect Maxwell's equations. Can you think of a deep reason for this?

$$\mathcal{L} = -\tfrac{1}{2}\eta^{\mu\nu}(\partial_\mu\phi)(\partial_\nu\phi) - V(\phi). \tag{1.148}$$