7.8. Symmetries and Conservation Laws

In mechanics, we learned that a conserved quantity was an observable, a function of coordinates and momenta describing the phase space of our dynamical system, which remained constant as these evolved following the equations of motion. Conservation laws were important in that they could be used to simplify the process of solving simple problems (e.g. the central force). More importantly, they allowed us to constrain the possible solutions to problems without explicit, complete solutions (e.g. collision processes).

In the context of a field theory like electrodynamics, we suspect that the role of coordinates and momenta will be taken by the fields themselves and their time derivatives (as we shall see, there are some complications here, but in essence this is the case). Note that these are functions on space so that the system in question will have an infinite-dimensional configuration space. In the context of a field theory, the most interesting conservation laws will involve local observables, which can be associated to an arbitrary volume in space and depend only upon the fields in this volume. Our usual way to construct such observables is as the integral over the volume $V$ of some function of the fields and their derivatives (to finite order). A conservation law will follow if the equations of motion (Maxwell’s equations) relate the change in this integral to some surface integral over the boundary of $V$. If $V$ is taken to be the whole of space with fields vanishing at infinity, the integral will be conserved, but far more useful is the local interpretation which takes the boundary term to represent the flow of our conserved quantity into or out of $V$. This very abstract discussion can be made more concrete by recalling, for example, our law of conservation of electric charge

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0. \tag{7.1}$$

We interpret the integral of $\rho$ as the total charge in $V$, and (7.1) relates the rate at which this changes to the flow of charge across the boundaries of $V$. Of course, if $V$ is all of space and the current vanishes at infinity, (7.1) tells us that the total charge in the universe is conserved, but the local form is far more useful in constraining physical processes.

How do we find conservation laws? Again, we draw upon our experience with classical mechanics where we found that these are related to symmetries of the system, transformations of the dynamical degrees of freedom which, when applied to a solution of the equations of motion, produced another solution. We will now undertake a careful study of our equations of motion – Maxwell’s equations – to find their symmetry properties and the local conservation laws that follow. In fact, we have already produced one important
example of such a local conserved quantity. Under suitable conditions (we will consider these below) we found that the energy stored in the fields in a region $V$ could be expressed as the integral of the energy density

$$u(x, t) = \frac{1}{2} (E \cdot D + B \cdot H) . \quad (7.2)$$

Is this conserved? We formulated (7.2) by starting with the equations of motion. We showed that if we assume a linear medium, with $D = \epsilon E$ and $H = \mu^{-1} B$, Maxwell’s equations allowed us to compute

$$\frac{\partial u}{\partial t} + \nabla \cdot S = -J \cdot E . \quad (7.3)$$

This looks almost like a conservation law, except that the right-hand side is not zero. Before we interpret this, let’s review the conditions under which (7.3) holds.

Our discussion of dispersion and absorption last time showed that the assumptions we made in deriving (7.3) are quite strong. In essence we assumed a medium with no dispersion (instantaneous constitutive equations) and hence (remember Kramers-Kronig) no absorption. We should not be too surprised that energy conservation is far more complicated for macroscopic fields in the presence of a medium with nontrivial electromagnetic properties. Under these circumstances, we expect energy exchange between the fields and the medium. Unless we include the latter in our dynamics, these exchanges will “violate” our conservation law. To find conserved quantities we will need to include the state of the medium in our equations. The “violation” of a conservation law is thus seen to be an important insight in its own right. Rather than dispose with energy conservation, we consider violations of (7.3) as providing us insights on the exchange of energy with the medium, as heat, mechanical stress, internal excitations, etc. On a more formal level, we recall that conservation of energy is related to invariance under time translations. Our macroscopic equations include some assumptions on the state of the medium in which the fields propagate; if this state is allowed to change with time (e.g. if the medium is heated by absorbing energy from the fields) then the equations, considered as determining the fields given a prescribed state of the medium, are not invariant under this transformation. To regain our conservation law and our symmetry, we need to extend our notion of energy to include the state of the medium; in terms of the equations of motion we need to include the medium as part of the dynamical system.
Let us begin, for simplicity, by neglecting all of these complications and working with the microscopic equations, describing propagation in an inert medium (or in the vacuum). Our equations are now
\[
\nabla \cdot \mathbf{B} = 0, \\
\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0, \\
\n\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}, \\
\n\nabla \times \mathbf{B} - \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} = \mu_0 \mathbf{J}.
\]
(7.4)

Are these invariant under time translations? Not as we have been using them so far. We have been (and other than in this section, will for a while longer) considering the sources on the right-hand side of these equations as prescribed functions of space and time. As such, they clearly violate the symmetry. We thus should not expect energy conservation, and indeed our would-be conservation law (7.3) has a nonzero right-hand side. The discussion above suggests both the source of the problem and a course of action. The reason energy is not conserved is that we have neglected one of its forms. The equations remind us that what we have neglected is the exchange of energy between the fields and the sources. Indeed, the right-hand side of (7.3) is precisely the density of the rate at which work is being done by the fields on the charges comprising the current distribution \( \text{vec} \mathbf{J} \) (this is where we started in our previous discussions of this equation). The solution, of course, is to \textit{include} the sources as part of our dynamical system. We have mentioned before that ultimately we will be interested in considering (7.4) together with the Lorentz force law
\[
\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})
\]
(7.5)
as a coupled set of equations of motion for the fields and the charges with which they interact. This complicates the solutions, so we will not make the leap yet. But for considerations of symmetry, we note that if the sources are part of our dynamical system rather than prescribed quantities, (7.4) does exhibit an invariance under time translations. Correspondingly, if we include the mechanical energy \( E_{\text{mech}} \) of our charged particles (we do not specify its precise form, because we can thus allow for other interactions among the particles). If all of the non-electromagnetic energy of the particles is included in this, then the rate of change of the energy in a volume \( V \) is given by the integral over \( V \) of the right-hand side of (7.3), and moving this over to the left-hand side we have our conservation law
\[
\frac{\partial u_{\text{mech}}}{\partial t} + \frac{\partial u}{\partial t} + \nabla \cdot \mathbf{S} = 0.
\]
(7.6)
Note that we are here assuming that electromagnetic fields are the only way energy can be transported through space. Thus the change in the total energy in a volume $V$ is given by the Poynting vector as the flux of energy carried by the fields through the boundaries of $V$. This can be false, for example if particles can move, carrying energy with them. In this case there will be an extra current, the convection current of mechanical energy, whose divergence will appear together with that of $S$.

The density of electromagnetic energy in the fields is

$$u = \frac{\epsilon_0}{2} (E^2 + c^2B^2) . \quad (7.7)$$

The role of the current is played here by the Poynting vector

$$S = \mu_0^{-1} E \times B . \quad (7.8)$$

The integral of $S$ over a surface gives a power – the rate at which the fields are transporting energy through the surface. Note that even static fields, so long as both electric and magnetic fields exist at the same time, carry energy across surfaces.

We have been very detailed and formal in our discussion; the benefits ought to be that extensions come easily. Our equations (7.4) are also invariant under translations in space, in that they involve no choice of an origin. As with energy, this is true provided we treat the sources as part of the system. Before we start, let us note that in this case the conserved quantity is a vector (momentum) rather than a scalar like the energy. More precisely, there are three conserved quantities, but they are dependent on a choice of axes and transform under rotations as a vector. We will thus find a vector of conservation equations, involving a vector of density functions and a two-index tensor combining the three currents representing the fluxes of the three components of momentum. This tensor was named by Maxwell, somewhat following conventions in elasticity theory, and is known as the Maxwell stress tensor $T$.

The rate at which momentum is transferred from the fields to the charged particles in a volume $V$ is simply found using (7.5) and Newton’s second law of motion as

$$\frac{dP_{\text{mech}}}{dt} = \int_V d^3x \left( \rho E + J \times B \right) . \quad (7.9)$$

The total momentum in the particles is given by the integral of a momentum density

$$P_{\text{mech}} = \int_V d^3x p_{\text{mech}} \quad (7.10)$$
in terms of which (7.9) is
\[ \frac{dp_{\text{mech}}}{dt} = \rho E + J \times B . \] (7.11)

We want to write the right hand side of this as the sum of a time derivative and a divergence. Both should involve the fields and not the sources, so we start by using (7.4) to obtain
\[
\frac{dp_{\text{mech}}}{dt} = \epsilon_0 (\nabla \cdot E)E + \mu_0^{-1} \left( \nabla \times B - \frac{1}{c^2} \frac{\partial E}{\partial t} \right) \times B \\
= \epsilon_0 \left[ \frac{\partial}{\partial t} (B \times E) - \frac{\partial B}{\partial t} \times E + (\nabla \cdot E)E \right] + \mu_0^{-1} (\nabla \times B) \times B \\
= \epsilon_0 \left[ \frac{\partial}{\partial t} (B \times E) + (\nabla \times E) \times E + (\nabla \cdot E)E \right] + \mu_0^{-1} [(\nabla \times B) \times B + (\nabla \cdot B)B] .
\] (7.12)

Note that we have used each of the four Maxwell’s equations once. Now we need to do some tensor manipulations. Recall that we use the Einstein summation convention; indices repeating in an expression are to be summed over the values 1, 2, 3. We use the essential identity we have used before, which in this condensed notation is written as
\[ \epsilon_{ijk} \epsilon_{klm} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl} \] (7.13)

and the abbreviation \( \partial_i \equiv (\partial/\partial x_i) \) to find
\[
[(\nabla \times E) \times E + (\nabla \cdot E)E]_i = -\epsilon_{ijk} E_j \epsilon_{klm} \partial_l E_m + E_i \partial_j E_j \\
= -\left( \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl} \right) E_j \partial_l E_m + E_i \partial_j E_j \\
= -E_j \partial_i E_j + E_j \partial_j E_i + E_i \partial_j E_j \\
= \partial_j (E_i E_j) - \frac{1}{2} \partial_i (E_j E_j) \\
= \partial_j \left( E_i E_j - \frac{1}{2} E^2 \delta_{ij} \right) .
\] (7.14)

This exhibits the terms in \( E \) as the divergence of a tensor. The terms in \( B \) are identical in index structure, so that if we define the Maxwell stress tensor
\[ T_{ij} = \epsilon_0 \left[ E_i E_j + c^2 B_i B_j - \frac{1}{2}(E^2 + c^2 B^2) \delta_{ij} \right] \] (7.15)

we have in fact a conservation equation
\[
\frac{dp_i}{dt} + \frac{\partial}{\partial t} \epsilon_0 (E \times B)_i + \partial_j T_{ij} = 0 .
\] (7.16)
The density of momentum in the fields is thus given by

\[ g = \epsilon_0 \mathbf{E} \times \mathbf{B} \, . \quad (7.17) \]

The fact that this is exactly \( g = (1/c^2)\mathbf{S} \) is not coincidental, reflecting essentially the relation of energy to momentum for massless particles. The tensor \( T \) represents the flux of momentum. For a surface \( C \) the integral

\[ \int_C T_{ij} da_j \quad (7.18) \]

represents the rate at which momentum in the \( i \) direction is transported through the surface by the fields. More concretely, it is the force in the \( i \) direction applied by the fields across the surface. It is interesting to note that \( T \) is symmetric

\[ T_{ij} = T_{ji} \quad (7.19) \]

though it is not obvious that it had to be.

Our equations (7.4) are also invariant under rotations, and as expected this leads to a conservation law for angular momentum. In the HW you will show that

\[ \frac{\partial}{\partial t} (\mathbf{x} \times g)_i + \frac{\partial}{\partial x_j} M_{ij} = -\frac{\partial}{\partial t} (\mathbf{x} \times p_{\text{mech}})_i \, , \quad (7.20) \]

with the current given by

\[ M_{ij} = \epsilon_{jkl} T_{ik} x_l \, . \quad (7.21) \]

The rotational invariance underlies many of the techniques we have used, and deserves a bit more detailed investigation. Rotations are linear transformations on space, i.e. they act as

\[ \mathbf{x} \rightarrow \mathbf{x}' = R\mathbf{x} \, , \quad (7.22) \]

or in components

\[ x'_i = R_{ij} x_j \, . \quad (7.23) \]

Not all linear transformations are rotations, of course. Under rotations, the magnitude of \( \mathbf{x} \) is preserved, which means \( R \) is an orthogonal matrix

\[ x'_i x'_i = R_{ij} x_j R_{ik} x_k = x_j R_{ij} R_{ik} x_k = x_i x_i \, , \quad (7.24) \]
which holds for arbitrary \( \mathbf{x} \) iff

\[
R_{ij} R_{ik} = \delta_{ik} ,
\]

or in matrix form

\[
RR^T = 1 .
\]

Rotations preserve something else besides the magnitude of vectors. Mathematically, they preserve the orientation of space. This means a right-handed triple of vectors transforms into another right-handed triple. What does this mean? We can characterize a right-handed triple of vectors \( \mathbf{a}, \mathbf{b}, \mathbf{c} \) by saying that

\[
(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \epsilon_{ijk} a_i b_j c_k \geq 0 .
\]

After a transformation by \( R \), we find

\[
(\mathbf{a}' \times \mathbf{b}') \cdot \mathbf{c}' = \epsilon_{lmn} a'_l b'_m c'_n = \epsilon_{lmn} R_{li} a_i R_{mj} b_j R_{nk} c_k .
\]

Note that I have chosen my indices carefully. The coefficient of \( a_i b_j c_k \) here is

\[
C_{ijk} = \epsilon_{lmn} R_{li} R_{mj} R_{nk} .
\]

What can we say about this set of 27 numbers (labeled by \( ijk \))? First, it is clearly antisymmetric under exchange of any pair of indices \( C_{ijk} = -C_{jik} = -C_{ikj} \). This means that of the 27 numbers only six are nonzero, and these are all related by the symmetry. In fact, what we have shown is

\[
C_{ijk} = C \epsilon_{ijk}
\]

for some number \( C \). To find \( C \) we compute

\[
\epsilon_{ijk} C_{ijk} = C \epsilon_{ijk} \epsilon_{ijk} = 6C .
\]

On the other hand, we have

\[
\epsilon_{ijk} \epsilon_{lmn} R_{li} R_{mj} R_{nk} = 6|R| ,
\]

where \( |R| \) is the determinant of \( R \). This may not be obvious to you, but pick any one of the six possible sets of indices \( ijk \) that contribute to this and write out the contribution; you will find an expression for the determinant. Since \( R \) is orthogonal, we know that

\[
|RR^T| = |R|^2 = 1 ,
\]
so that $|R| = \pm 1$. What (7.32) tells us is that rotations are represented by orthogonal matrices with determinant $+1$.

Physical quantities are classified by their properties under rotations. The classification is important and useful, because the rotational invariance of the underlying theory means that physically valid equations must be valid in any set of coordinates; in other words, they have to relate properties with compatible transformation properties. It turns out that the most useful objects are collections of numbers that transform \textit{linearly} under rotations. Not too surprising. This means that setting such a thing to zero (or equivalently setting two of them equal) in one set of axes will make it zero in any set. A collection of numbers transforming linearly form a \textit{representation} of the group of rotations.

The simplest to describe are \textit{scalars}. These are properties like mass or charge that are invariant under rotations. More interesting are \textit{vector} quantities like velocity or momentum. A vector is represented, once we choose a basis of coordinates, by three numbers, with the property that under rotations they transform in the same way that position vectors transform. There is a point of confusion here, let’s try to clarify it at the risk of achieving the opposite. First, in the context of a field theory we discuss physical properties assigned to a specific point in space at a particular time. Let us take for example the scalar potential $\Phi(x, t)$ and electric field $E(x, t)$. What is meant by the statement that the potential is a scalar and the field a vector? There are two equivalent points of view that one can take regarding the action of symmetries in general. The \textit{active} view compares the system with another system obtained from it by a rotation $R$, i.e. an object located at $x$ in the original is now located at the position $x' = Rx$. The electric potential observed in the new system will be a new function $\Phi'(x, t)$. The scalar transformation law is

$$
\Phi'(x', t) = \Phi(x, t) = \Phi(R^{-1}x', t) .
$$

(7.34)

For the electric field, we will find that in the new system we have a field $E'(x, t)$ related to the original by

$$
E'(x', t) = RE(x, t) = RE(R^{-1}x', t) .
$$

(7.35)

Note that in neither case can one obtain a constraint on the function $\Phi$ or $E$ from this alone. An equivalent (but confusingly so) point of view is the \textit{passive} one in which we imagine the \textit{system} remains unchanged but the \textit{observer} is rotated, so that measurements are now made with respect to axes related to the original ones by a rotation. It is worth
writing down the transformation properties to conclude that a passive rotation $R^{-1}$ leads to the same new functions $\Phi'$ and $E'$ as an active rotation by $R$.

We have seen that more elaborate constructions exist. By taking two vectors and putting them next to each other we can make a two-index tensor, a collection of nine objects transforming into each other as

$$T_{ij} \rightarrow T_{ij}' = R_{ik}R_{jl}T_{kl}.$$  \hspace{1cm} (7.36)

It turns out that all of the representations we care about (in classical physics) can be formed in this way, with one improvement. The nine objects in $T$ transform into linear combinations under rotations, but some combinations do not mix with others. One note is that the symmetric part of $T$ does not mix with its antisymmetric part. That is, if $T_{ij} = \pm T_{ji}$ then this will be true of $T'$ as well. So the nine objects in $T$ can be broken up, in an invariant and hence physically meaningful way, into three (antisymmetric) and six (symmetric)

$$T_{ij}^S = \frac{1}{2} (T_{ij} + T_{ji})$$

$$T_{ij}^A = \frac{1}{2} (T_{ij} - T_{ji}).$$  \hspace{1cm} (7.37)

Each of these can be further manipulated. First, the symmetric piece includes six numbers but one linear combination is in fact left invariant by rotations. It’s the trace

$$T_{ii} \rightarrow T_{ii}' = R_{ik}R_{ii}T_{kl} = \delta_{kl}T_{kl} = T_{ii},$$  \hspace{1cm} (7.38)

because $R$ is orthogonal. So the six break up into a scalar and a traceless symmetric two-tensor (spin-2)

$$T_{ij}^{(2)} = T_{ij}^S - \frac{1}{3}T_{ii}\delta_{ij}.$$  \hspace{1cm} (7.39)

You will recognize this as the manipulation we conducted on the electric quadrupole moment to extract the trace leaving the five components of the pure (traceless) quadrupole.

What about the antisymmetric piece? Here we have a special trick that works only in three dimensions (more generally it works for an $n - 1$-tensor in $n$ dimensions). Construct

$$t_i = \epsilon_{ijk}T_{jk}.$$  \hspace{1cm} (7.40)

Note that this indeed involves only the antisymmetric part of $T$. Further, you showed in HW4 that it transforms under rotations like a vector. The three components of an antisymmetric two-tensor can be combined into a vector under rotations.
Similar ideas can be applied to tensors of higher rank, breaking them up by symmetry, removing traces, and trading any two antisymmetrized indices for one using an $\epsilon$. We can’t take the time to do much more at this point, though.

What other symmetry properties of (7.4) can we use to advantage? In constructing the equations we used Galilean invariance to predict Faraday’s law of induction. Yet, puzzlingly, the resulting equations are not invariant under Galilean transformations. That is, there is no way to assign transformation properties to $\mathbf{E}$ and $\mathbf{B}$ so that a solution of (7.4) in one frame transforms to a solution in another frame. Many attempts to reconcile this contradiction in the nineteenth century failed, and ultimately what had to change was the invariante itself – the equations are invariant under relativistic Lorentz transformations. More on this later, of course.

There are an additional set of discrete symmetries of (7.4) that may be used, as we did the continuous ones, to constrain physical equations. First among these is a symmetry under reflections. Consider the inversion $P$ that takes $\mathbf{x} \to -\mathbf{x}$. This is orthogonal but has negative determinant. It is easy, in fact, to show that any orthogonal transformation (length-preserving) can be written as the product of $P$ and a rotation. In general, transformation properties under reflection may be assigned separately from rotation properties. Since $P^2 = 1$ a nontrivial transformation under $P$ is constrained to be such that it also squares to the identity; for linear actions this means simply a change of sign. True scalar quantities are invariant under reflections as well as under rotations. Mass and electric charge are true scalars. Pseudoscalars are invariant under rotations but change sign under reflections. Similarly, true (polar) vectors transform under $P$ like the coordinate vector while pseudo (axial) vectors are invariant. An important example is the “vector product.”

If $\mathbf{a}$ and $\mathbf{b}$ are polar vectors (or both axial vectors), then under a reflection both change sign and the “vector product” is unchanged. Similar considerations apply to tensors of higher rank. We usually follow the notation that a “pure” tensor of rank $n$ is what you get by juxtaposing $n$ vectors, so that under $P$ it transforms by a factor $(-1)^n$. If it transforms by $(-1)^{n+1}$ it is a pseudotensor. Note that the vector we form from an antisymmetric two-tensor (7.40) is an axial vector if $T$ is a pure two-tensor. In fact, this is probably a healthier way to think of vector product quantities such as angular momentum and magnetic fields. Inspection of (7.4) shows that if $\rho$ is a true scalar, $\mathbf{E}$ and $\mathbf{J}$ polar vectors, and $\mathbf{B}$ an axial vector, the equations are invariant under $P$.

Finally, like all equations of classical physics, (7.4) is also invariant under time reversal $T: t \to -t$. This changes the sign of all time derivatives. Position vectors are invariant
under $T$ but velocities change sign. $\mathbf{E}$ and $\rho$ are even under time reversal, but $\mathbf{B}$ and $\mathbf{J}$ are odd. Transformation properties of other relevant quantities are listed in the text or easily found from these.

7.9. Energy and Momentum in Linear Media

As we mentioned when discussing conservation laws, in general the presence of a medium can violate invariance under both space and time translations, hence both energy and momentum conservation. In technical terms, in our derivation of energy conservation from (7.22) we showed in (7.24) that

$$-\mathbf{J} \cdot \mathbf{E} = \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} + \nabla \cdot \mathbf{S}. \tag{7.41}$$

In attempting to rewrite the right-hand side as the time derivative of an energy density, we now know that assuming an instantaneous linear relation between $\mathbf{E}$ and $\mathbf{D}$, for example, was too quick. In fact, in the presence of absorption into the medium, we do not expect the electromagnetic energy density to be conserved. Since Kramers-Kronig shows that a causal dispersive medium is also absorptive, the difficulty is expected.

It is still worthwhile to probe this question a little closer. Let us, then, consider a real causal medium and use our Fourier transform to study the propagation of the fields. Thus we write (note in maintaining consistent Fourier conventions we differ from the text here) for the electric quantities (a parallel treatment can be made for the magnetic ones)

$$\mathbf{E}(\mathbf{x}, t) = \frac{1}{2\pi} \int d\omega e^{-i\omega t} \tilde{\mathbf{E}}(\mathbf{x}, \omega)$$

$$\mathbf{D}(\mathbf{x}, t) = \frac{1}{2\pi} \int d\omega e^{-i\omega t} \tilde{\mathbf{D}}(\mathbf{x}, \omega), \tag{7.42}$$

with the Fourier transformed constitutive relation

$$\tilde{\mathbf{D}}(\mathbf{x}, \omega) = \tilde{\epsilon}(\omega) \tilde{\mathbf{E}}(\mathbf{x}, \omega). \tag{7.43}$$

For simplicity, we are here assuming a homogeneous and isotropic medium. We can then
compute in terms of the Fourier transforms

\[
E \cdot \frac{\partial D}{\partial t} = \frac{1}{4\pi^2} \int d\omega \int d\omega' e^{-i(\omega+\omega')t} \tilde{E}(\omega)(-i\omega')\tilde{D}(\omega')
= \frac{1}{4\pi^2} \int d\omega \int d\omega' e^{-i(\omega'-\omega)t} \tilde{E}^*(\omega)(-i\omega'\tilde{\epsilon}(\omega'))\tilde{E}(\omega')
= \frac{1}{8\pi^2} \int d\omega \int d\omega' e^{-i(\omega'-\omega)t} \left[ \tilde{E}^*(\omega)(-i\omega'\tilde{\epsilon}(\omega'))\tilde{E}(\omega') + \tilde{E}^*(-\omega')(i\omega\tilde{\epsilon}(-\omega))\tilde{E}(-\omega) \right]
= \frac{1}{8\pi^2} \int d\omega \int d\omega' e^{-i(\omega'-\omega)t} \left[ \tilde{E}^*(\omega)(-i\omega'\tilde{\epsilon}(\omega'))\tilde{E}(\omega') + \tilde{E}(\omega')(i\omega\tilde{\epsilon}^*(\omega))\tilde{E}^*(\omega) \right]
= \frac{1}{8\pi^2} \int d\omega \int d\omega' e^{-i(\omega'-\omega)t} \tilde{E}^*(\omega)[-i\omega'\tilde{\epsilon}(\omega') + i\omega\tilde{\epsilon}^*(\omega)]\tilde{E}(\omega')
\]

(7.44)

where I have used several times the condition that the Fourier transform of a real function \(f(t)\) satisfies \(\tilde{f}(-\omega^*) = \tilde{f}^*(\omega)\) and the fact that the integrals are over real \(\omega\).

If we are describing the propagation of an almost monochromatic pulse, i.e. if \(\tilde{E}(\omega)\) is narrowly centered about some frequency \(\omega_0\) with a width \(\Delta\) small compared to \(d\tilde{\epsilon}/d\omega\), then we may expand \(\tilde{\epsilon}\) about \(\omega'\) to find

\[
-i\omega'\tilde{\epsilon}(\omega') + i\omega\tilde{\epsilon}^*(\omega) \sim 2\omega'\text{Im}(\tilde{\epsilon}(\omega')) - i(\omega' - \omega)\frac{d(\omega\tilde{\epsilon}^*)}{d\omega}(\omega') + \cdots
\]

(7.45)

which upon inserting into (7.44) yields (recall that inside the Fourier transform the factor of \(-i(\omega' - \omega)\) acts as a time derivative)

\[
E \cdot \frac{\partial D}{\partial t} = \frac{1}{4\pi^2} \int d\omega \int d\omega' e^{-i(\omega'-\omega)t} \tilde{E}^*(\omega)\omega'\text{Im}(\tilde{\epsilon}(\omega'))\tilde{E}(\omega')
+ \frac{1}{8\pi^2} \frac{\partial}{\partial t} \int d\omega \int d\omega' e^{-i(\omega'-\omega)t} \tilde{E}^*(\omega)\frac{d(\omega'\tilde{\epsilon}^*(\omega'))}{d\omega'}\tilde{E}(\omega')
\]

(7.46)

The total time derivative is what we are after if we want to relate the power dissipated to free charge \(J \cdot E\) to an energy density. The presence of an additional term represents dissipation into the medium, and as expected it is proportional to \(\text{Im}(\tilde{\epsilon})\). An identical set of manipulations on the magnetic contribution to the right-hand side of (7.41) yields altogether

\[
-J \cdot E = \frac{1}{8\pi^2} \frac{\partial}{\partial t} \int d\omega \int d\omega' e^{-i(\omega'-\omega)t} \left[ \tilde{E}^*(\omega)\frac{d(\omega'\tilde{\epsilon}^*(\omega'))}{d\omega'}\tilde{E}(\omega') + \tilde{H}^*(\omega)\frac{d(\omega'\tilde{\mu}^*(\omega'))}{d\omega'}\tilde{H}(\omega') \right]
+ \frac{1}{4\pi^2} \int d\omega \int d\omega' e^{-i(\omega'-\omega)t} \left[ \tilde{E}^*(\omega)\omega'\text{Im}(\tilde{\epsilon}(\omega'))\tilde{E}(\omega') + \tilde{H}^*(\omega)\omega'\text{Im}(\tilde{\mu}(\omega'))\tilde{H}(\omega') \right]
\]

(7.47)
To clarify the meaning of this, we write our fields with their sharply-peaked Fourier transform, now adding $\Delta << \omega_0$ to our assumption that $\Delta (d\epsilon/d\omega) << 1$ as

\[ \begin{align*} E(t) &= E_0(t) \cos(\omega_0 t + \alpha) \\ H(t) &= H_0(t) \cos(\omega_0 t + \beta) , \end{align*} \tag{7.48} \]

where $E_0$ and $H_0$ are slowly varying, in the sense that their variation over a period of the central oscillation frequency $\omega_0$ is negligible. The Fourier transform of $E$ then has the form

\[ \tilde{E}(\omega) = \frac{1}{2} \int dt E_0(t) \left( e^{i\alpha} e^{i(\omega + \omega_0) t} + e^{-i\alpha} e^{i(\omega - \omega_0) t} \right) = \frac{1}{2} \left( e^{i\alpha} \tilde{E}_0(\omega + \omega_0) + e^{-i\alpha} \tilde{E}_0(\omega - \omega_0) \right) . \tag{7.49} \]

The fact that $E_0$ is slowly varying means the two terms are sharply peaked about $\omega = \pm \omega_0$. Inserting this into (7.47) we find inside the integral a sum of four terms, two of which are peaked around $\omega' = \omega = \pm \omega_0$ and two about $\omega' = -\omega = \mp \omega_0$. After integration, the latter two terms will be rapidly oscillatory $\sim e^{\mp 2i\omega_0 t}$ while the former two vary slowly, on the scale of the variation of $E_0$. We use the smallness of this to average (7.47) over a time spanning many periods of $\omega_0$ but over which $E_0$ is nearly constant. Then the oscillatory terms average to zero and we are left with integrals of the form

\[ \langle \int d\omega \int d\omega' e^{-i(\omega' - \omega) t} \tilde{E}^*(\omega) f(\omega') \tilde{E}(\omega') \rangle \]

\[ \sim \frac{1}{4} \int d\omega \int d\omega' e^{-i(\omega' - \omega) t} f(\omega') \left( \tilde{E}^*_0(\omega + \omega_0) \tilde{E}_0(\omega' + \omega_0) + \tilde{E}^*_0(\omega - \omega_0) \tilde{E}_0(\omega' - \omega_0) \right) \]

\[ \sim \frac{1}{4} \int d\omega \int d\omega' e^{-i(\omega' - \omega) t} \left( f(\omega' - \omega_0) + f(\omega' + \omega_0) \right) \tilde{E}_0(\omega') \tilde{E}_0(\omega) . \tag{7.50} \]

where in the last line we have shifted the integration variables appropriately. Finally, if we can assume that $f(\omega)$ is slowly varying about $\pm \omega_0$ in the sense that it may be assumed to be constant in the integrals above for which $\omega'$ is restricted to the range $\Delta$ of frequencies for which $\tilde{E}_0$ is nonzero, then we find

\[ \langle \int d\omega \int d\omega' e^{-i(\omega' - \omega) t} \tilde{E}^*(\omega) f(\omega') \tilde{E}(\omega') \rangle \]

\[ \sim \frac{1}{4} \left[ f(-\omega_0) + f(\omega_0) \right] \int d\omega e^{i\omega t} \tilde{E}^*_0(\omega) \int d\omega' e^{-i\omega' t} \tilde{E}_0(\omega') \]

\[ = \pi^2 \left[ f(-\omega_0) + f(\omega_0) \right] \langle E_0(t) \cdot E_0(t) \rangle \]

\[ = 2\pi^2 \left[ f(-\omega_0) + f(\omega_0) \right] \langle E(t) \cdot E(t) \rangle . \tag{7.51} \]
Inserting this into (7.47) together with an identical analysis for the magnetic term, we note that
\[
\frac{d(\omega'\tilde{\epsilon}^*)}{d\omega'}\bigg|_{\omega'=\omega_0} + \frac{d(\omega'\tilde{\epsilon}^*)}{d\omega'}\bigg|_{\omega'=-\omega_0} = 2\text{Re} \left[ \frac{d(\omega'\tilde{\epsilon})}{d\omega}(\omega_0) \right]
\]
so that we have in all
\[
\langle -\mathbf{J} \cdot \mathbf{E} \rangle = \omega_0 \text{Im} \tilde{\epsilon}(\omega_0) \langle \mathbf{E}(t) \cdot \mathbf{E}(t) \rangle \\
+ \omega_0 \text{Im} \tilde{\epsilon}(\mu_0) \langle \mathbf{H}(t) \cdot \mathbf{H}(t) \rangle \\
+ \frac{\partial}{\partial t} \frac{1}{2} \left\{ \text{Re} \left[ \frac{d(\omega'\tilde{\epsilon})}{d\omega}(\omega_0) \right] \langle \mathbf{E}(t) \cdot \mathbf{E}(t) \rangle \\
+ \text{Re} \left[ \frac{d(\omega'\tilde{\mu})}{d\omega}(\omega_0) \right] \langle \mathbf{H}(t) \cdot \mathbf{H}(t) \rangle \right\} .
\]
This gives a modified version of Poynting’s theorem of energy conservation for a linear dispersive medium. The last two lines give the modified form of the electromagnetic energy density, while the first two provide an additional dissipative term describing energy loss to the medium, usually in the form of heat.

Defining the momentum of the radiation field in a medium is more tricky, in that elastic properties of the medium can in principle enter into the calculation. We will not have time to get into this.

7.10. A Word on Magnetic Monopoles

It is an experimental fact that no magnetic monopoles have ever been detected, despite serious searches. Before discussing the motivation for these searches, we should pause to note what we mean. Maxwell’s equations (7.22) certainly exhibit an aesthetically displeasing asymmetry between the electric and magnetic fields. This in itself is motivation for considering the modified equations that would hold in a universe including a conserved magnetic charge, satisfying
\[
\frac{\partial \rho_m}{\partial t} + \nabla \cdot \mathbf{J}_m = 0 .
\]
The equations are now pleasantly symmetric
\[
\nabla \cdot \mathbf{D} = \rho \\
\nabla \cdot \mathbf{B} = \rho_m \\
\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J}_e \\
-\nabla \times \mathbf{E} = \frac{\partial \mathbf{B}}{\partial t} + \mathbf{J}_m .
\]
Not only is there a nice symmetry under exchange of electric and magnetic fields, there is now a more general duality symmetry, under the transformation

\[
\begin{align*}
    E' &= E \cos \xi + Z_0 H \sin \xi \\
    Z_0 H' &= -E \sin \xi + Z_0 H \cos \xi \\
    Z_0 \rho_e &= Z_0 \rho_e \cos \xi + \rho_m \sin \xi \\
    \rho_m &= -Z_0 \rho_e \sin \xi + \rho_m \cos \xi ,
\end{align*}
\] (7.56)

with the corresponding action on \( D, B \), and the currents, where \( Z_0 = \sqrt{\mu_0/\epsilon_0} \) adjusts the units, and \( \xi \) is an arbitrary (pseudoscalar) angle. This preserves not only the equations (7.55) but also the Poynting vector, energy density, and all the quadratic forms we have considered. This means that our determination that particles carry electric, but not magnetic, charge, is a matter of convention, a choice of \( \xi \). The invariant question, of course, is the relative charges of various particles. For two charges \((q_e, q_m)\) and \((q'_e, q'_m)\) the product

\[
L(q, q') = q_e q'_m - q_m q'_e
\] (7.57)

is invariant under duality and, as we shall see, defines a physical property of a system containing these two. The experimental fact is that for all known particles, this product vanishes. In other words, having decided that electrons carry only electric charge, fixing \( \xi \), we discover experimentally that all other known particles carry charges proportional to that of the electron. Now that we mention this, we discover more, in fact. The charges are not only proportional to the charge of the electron, but the charge ratios are integers. This property, referred to as charge quantization may be taken as simply another of nature’s mysteries. Dirac, however, brilliantly showed how it follows in quantum mechanics from the existence of monopoles. By inverting his argument, which we will briefly summarize below, we may take the observed fact that charge is quantized as evidence for the existence of monopoles. Hence the searches. Dirac’s argument is beautiful in that it involves only electromagnetism. We now understand this to be embedded in the more general electroweak interactions, and can make somewhat more direct arguments (first set down by Polyakov) that monopoles, satisfying the properties Dirac set out, in fact do exist. Their experimental absence then becomes the puzzle, and in fact is one of the original motivations behind the inflationary cosmological theories first developed by Guth.

Dirac’s construction of a monopole is ingenious. We have seen, in our discussions of magnetostatics, that the field near the tip of a long bar magnet (or solenoid) looks
approximately like a monopole. In fact, we saw in class that for a bar with uniform
magnetization \( \mathbf{M} = M\hat{z} \), cross-sectional area \( a^2 \) and length \( L \) the field at a
distance \( a << r << L \) from one end is approximately

\[
\mathbf{B}(\mathbf{x}) \sim \frac{g\hat{r}}{4\pi r^2}, \quad (7.58)
\]

with \( g = Ma^2 \). This is correct to within corrections of order \( a/r \) and of order \( r/L \).
Imagine, therefore, a long, infinitely thin, magnet (or solenoid) stretching out from the
origin down the negative \( z \) axis to infinity. In the region \( z \geq 0 \) this configuration generates the
field configuration of (7.58). Of course, the apparent fact that \( \nabla \cdot \mathbf{B} = g\delta^{(3)}(\mathbf{x}) \) is an artifact
of our limiting procedure. In fact, the field configuration, while singular along the entire
negative \( z \) axis (because we have an infinitely thin magnet with an infinite magnetization
so that \( Ma^2 \) is finite while \( a \to 0 \)), is divergenceless. In particular, we can find a vector
potential \( \mathbf{A}_+(\mathbf{x}) \) such that \( \mathbf{B} = \nabla \times \mathbf{A}_+ \). Explicitly, we can write this as

\[
\mathbf{A}_+(\mathbf{x}) = -\frac{g}{4\pi} \int_{-\infty}^{0} dz'\hat{z} \times \nabla \left( \frac{1}{|\mathbf{x} - \mathbf{x}'|} \right). \quad (7.59)
\]

Nevertheless, for \( z \geq 0 \) it looks for all the world like a monopole. In fact, it looks like
a monopole everywhere except along the negative \( z \) axis! This is not yet a complete
description of a the field of a monopole, though, because it is singular along the “Dirac
string” lying along the negative \( z \) axis. Can we find a configuration that satisfies (7.58)
everywhere? Not with solenoids and magnets, of course, but here is where the ingenuity
comes in.

Consider a different situation, this time with a long, thin magnet of magnetization
\( \mathbf{M} = -M\hat{z} \) and lying along the positive \( z \) axis. For this configuration, the field is given
by (7.58) everywhere away from the positive \( z \) axis. Here too we can write the field, away
from the singularity at \( \theta = 0 \), as \( \mathbf{B} = \nabla \times \mathbf{A}_- \), with

\[
\mathbf{A}_-(\mathbf{x}) = \frac{g}{4\pi} \int_{0}^{\infty} dz'\hat{z} \times \nabla \left( \frac{1}{|\mathbf{x} - \mathbf{x}'|} \right). \quad (7.60)
\]

The crucial point is that away from both axes \( (0 < \theta < \pi) \) the fields of both con-
figurations agree. So we can imagine a field that agrees with the first configuration for
\( \theta < \pi \) and with the second for \( \theta > 0 \), these are consistent. Our field now is given by (7.58)
everywhere in space and we take this to define the field of a monopole with strength \( g \).
As we expected, this \textit{has} a divergence, so it will not be possible to find a vector potential
A such that $\mathbf{B} = \nabla \times \mathbf{A}$. On the other hand, everywhere except on the $z$ axis, the same field (7.58) can be obtained from either of (7.59) or (7.60). What this means is that the two vector potentials differ by a gauge transformation

$$
A_+ = A_- + \nabla \Lambda_{+-},
$$

(7.61)

for some function $\Lambda_{+-}$ everywhere both potentials exist (away from the $z$ axis). If we work hard enough, we can even find the function $\Lambda_{+-}$ but we will not really need to do this. There is something we can compute more easily. Let us consider a sphere (shell!) of radius $R$ about the origin, and let $C$ be the circle in which it intersects the plane $z = 0$, with an orientation such that $\Gamma = \partial H_+ = -\partial H_-$ is the boundary of the upper hemisphere (with outward-pointing normal). Then we know that

$$
\frac{g}{2} = \int_{H_+} \mathbf{B} \cdot d\mathbf{a} = \oint_C A_+ \cdot d\mathbf{l},
$$

(7.62)

where the first equality used (7.58) and the second is Stokes’s theorem. Similarly, we have

$$
-\frac{g}{2} = -\int_{H_-} \mathbf{B} \cdot d\mathbf{a} = \oint_C A_- \cdot d\mathbf{l}.
$$

(7.63)

Subtracting these two, we see that

$$
g = \oint_C (A_+ - A_-) \cdot d\mathbf{l} = \oint_C \nabla \Lambda_{+-} \cdot d\mathbf{l}.
$$

(7.64)

This is puzzling, because the last integral is easy to do and yields, if we begin at a point $p \in C$ (and hence end at the same point) $\Lambda_{+-}(p) - \Lambda_{+-}(p)$. This means that we cannot interpret $\Lambda$ as a real function at all. Rather, it is a function from space to a circle of circumference $g$. When we say that the gauge group associated to electrodynamics is $U(1)$, the group of unitary $1 \times 1$ matrices, which is a circle, this is what we mean.

Now we can finish off Dirac’s argument by bringing in a little bit of quantum mechanics. For a particle of (electric) charge $e$, we obtain the Hamiltonian for interactions with the electromagnetic field by the minimal coupling substitution

$$
\hat{p} \to \hat{p} - e \mathbf{A}.
$$

(7.65)

This is always a problem, because as we have said often, $\mathbf{A}$ is not a physical field, so is undetermined by physics. What is determined by physics is $\mathbf{B}$. That means that physics
as written in terms of \( \mathbf{A} \) needs to be gauge invariant. In particular this must be true of the Hamiltonian, and the way we ensure this is by making the right-hand side of (7.65) invariant. Since \( \hat{p} \) acts on \( \psi \) as \( -i\hbar \nabla \), we see that (7.65) is under the simultaneous action

\[
\begin{align*}
\mathbf{A} &\rightarrow \mathbf{A} - \nabla \Lambda \\
\psi &\rightarrow e^{ie\Lambda/\hbar},
\end{align*}
\]

(7.66)

we have

\[
(-i\hbar \nabla - eA)\psi \rightarrow e^{ie\Lambda/\hbar} (-i\hbar \nabla - e\nabla \Lambda - eA + e\nabla \Lambda) \psi = e^{ie\Lambda/\hbar} (-i\hbar \nabla - eA)\psi .
\]

(7.67)

The overall phase will cancel when constructing real observables (more precisely, this will be true for all observables that are not charged). We now know the gauge transformation properties of \( \psi \). But as we have just seen, \( \Lambda \) need not be a single-valued function, because it should be viewed as a map to the circle; in particular for our monopole field we see that the gauge transformation relating \( \mathbf{A}_\pm \) will, when acting on \( \psi \), produce a single-valued wave function only if we have

\[
e^g/\hbar = 2\pi n ,
\]

(7.68)

for some integer \( n \). This is Dirac’s famous quantization condition. The existence of a monopole constrains the possible electric charges (as of course the existence of electrons constrains the possible magnetic charges of monopoles).