Chapter 8

Maxwell's Equations

8.1 The Maxwell Displacement Current

Maxwell's Equations (ME) consist of two inhomogeneous partial differential equations and two homogeneous partial differential equations. At this point you should be familiar at *least* with the "static" versions of these equations by name and function:

$$\begin{aligned}
\boldsymbol{\nabla} \cdot \boldsymbol{D} &= \rho \text{ Gauss's Law for Electrostatics} && (8.1)\\
\boldsymbol{\nabla} \times \boldsymbol{H} &= \boldsymbol{J} \text{ Ampere's Law} && (8.2)\\
\boldsymbol{\nabla} \cdot \boldsymbol{B} &= 0 \text{ Gauss's Law for Magnetostatics} && (8.3)\\
\boldsymbol{\nabla} \times \boldsymbol{E} + \frac{\partial \boldsymbol{B}}{\partial t} &= 0 \text{ Faraday's Law} && (8.4)
\end{aligned}$$

in SI units, where $\boldsymbol{D} = \epsilon \boldsymbol{E}$ and $\boldsymbol{H} = \boldsymbol{B}/\mu$.

The astute reader will immediately notice two things. One is that these equations are not *all*, strictly speaking, static – Faraday's law contains a time derivative, and Ampere's law involves moving charges in the form of a current. The second is that they are *almost* symmetric. There is a divergence equation and a curl equation for each kind of field. The inhomogenous equations (which are connected to *sources* in the form of electric charge) involve the electric displacement and magnetic field, where the homogeneous equations suggest that there is no magnetic charge and consequently no screening of the magnetic induction or electric field due to magnetic charge. One asymmetry is therefore the presence/existence of electric charge in contrast with the absence/nonexistence of magnetic charge.

The other asymmetry is that Faraday's law connects the curl of the $\boldsymbol{E}$ field to the time derivative of the $\boldsymbol{B}$ field, but its apparent partner, Ampere's Law, does *not* connect the curl of $\boldsymbol{H}$ to the time deriviative of vD as one might expect from symmetry alone.

If one examines Ampere's law in its integral form, however:

$$\oint_C \boldsymbol{B} \cdot d\boldsymbol{\ell} = \mu \int_{S/C} \boldsymbol{J} \cdot \hat{\boldsymbol{n}}\, dA \tag{8.5}$$

one quickly concludes that the current through the open surface S bounded by the closed curve C is *not invariant* as one chooses different surfaces. Let us analyze this and deduce an invariant form for the current (density), two ways.

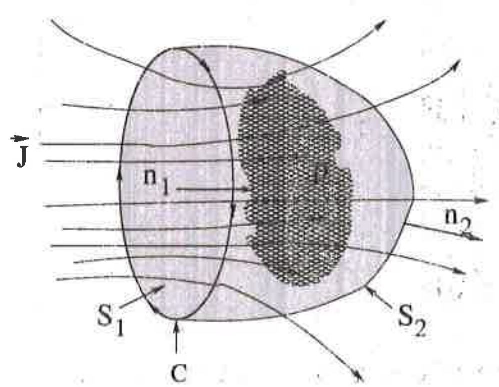

Figure 8.1: Current flowing through a closed curve C bounded by two surfaces, S_1 and S_2.

Consider a closed curve C that bounds *two distinct* open surfaces S_1 and S_2 that together form a *closed* surface $S = S_1 + S_2$. Now consider a current (density) "through" the curve C, moving from left to right. Suppose that *some* of this current accumulates inside the volume V bounded by S. The law of *charge conservation* states that the flux of the current density *out* of the closed surface S is equal to the rate that the total charge inside *decreases*. Expressed as an integral:

$$\oint_S \boldsymbol{J} \cdot \hat{\boldsymbol{n}}\, dA = -\frac{d}{dt} \int_{V/S} \rho\, dV \tag{8.6}$$

With this in mind, examine the figure above. If we rearrange the integrals on the left and right so that the normal $\hat{\boldsymbol{n}}_1$ points *in* to the volume (so we can compute the current through the surface S_1 moving from left to right) we can easily see that charge conservation tells us that the current in through S_1 minus the current out through S_2 must equal the rate at which the total charge inside this volume increases. If we express this as integrals:

$$\begin{aligned}\int_{S_1} \boldsymbol{J} \cdot \hat{\boldsymbol{n}}_1 \, dA - \int_{S_2} \boldsymbol{J} \cdot \hat{\boldsymbol{n}}_2 \, dA &= \frac{dQ}{dt} \\ &= \frac{d}{dt} \int_{V/S} \rho \, dV \end{aligned} \tag{8.7}$$

In this expression and figure, note well that $\hat{\boldsymbol{n}}_1$ and $\hat{\boldsymbol{n}}_2$ point through the loop in the *same* sense (e.g. left to right) and note that the volume integral is over the volume V bounded by the *closed* surface formed by S_1 and S_2 together.

Using Gauss's Law for the electric field, we can easily connect this volume integral of the charge to the flux of the electric field integrated over these two surfaces with *outward* directed normals:

$$\begin{aligned}\int_{V/S} \rho \, dV &= \epsilon \oint_S \boldsymbol{E} \cdot \hat{\boldsymbol{n}} \, dA \\ &= -\epsilon \int_{S_1} \boldsymbol{E} \cdot \hat{\boldsymbol{n}} \, dA + \epsilon \int_{S_2} \boldsymbol{E} \cdot \hat{\boldsymbol{n}} \, dA \end{aligned} \tag{8.8}$$

Combining these two expressions, we get:

$$\begin{aligned}\int_{S_1} \boldsymbol{J} \cdot \hat{\boldsymbol{n}}_1 \, dA \quad - \quad \int_{S_2} \boldsymbol{J} \cdot \hat{\boldsymbol{n}}_2 \, dA = \\ \frac{d}{dt}\left\{-\epsilon \int_{S_1} \boldsymbol{E} \cdot \hat{\boldsymbol{n}}_1 \, dA + \epsilon \int_{S_2} \boldsymbol{E} \cdot \hat{\boldsymbol{n}}_2 \, dA\right\} \end{aligned} \tag{8.9}$$

$$\begin{aligned}\int_{S_1} \boldsymbol{J} \cdot \hat{\boldsymbol{n}}_1 \, dA \quad + \quad \frac{d}{dt}\epsilon \int_{S_1} \boldsymbol{E} \cdot \hat{\boldsymbol{n}}_1 \, dA = \\ \int_{S_2} \boldsymbol{J} \cdot \hat{\boldsymbol{n}}_2 \, dA + \frac{d}{dt} \int_{S_2} \epsilon \boldsymbol{E} \cdot \hat{\boldsymbol{n}}_2 \, dA \end{aligned} \tag{8.10}$$

$$\int_{S_1} \left\{\boldsymbol{J} + \epsilon \frac{d\boldsymbol{E}}{dt}\right\} \cdot \hat{\boldsymbol{n}}_1 \, dA = \int_{S_2} \left\{\boldsymbol{J} + \epsilon \frac{d\boldsymbol{E}}{dt}\right\} \cdot \hat{\boldsymbol{n}}_2 \, dA \tag{8.11}$$

From this we see that the flux of the "current density" inside the brackets is *invariant* as we choose different surfaces bounded by the closed curve C.

In the original formulation of Ampere's Law we can clearly get a different answer on the right for the current "through" the closed curve depending on which surface we choose. This is clearly impossible. We therefore modify Ampere's Law to use the *invariant* current density:

$$\boldsymbol{J}_{\text{inv}} = \boldsymbol{J} + \epsilon \frac{d\boldsymbol{E}}{dt} \tag{8.12}$$

where the flux of the second term is called the *Maxwell displacement current* (MDC). Ampere's Law becomes:

$$\begin{aligned} \oint_C \boldsymbol{B} \cdot d\boldsymbol{\ell} &= \mu \int_{S/C} \boldsymbol{J}_{\text{inv}} \cdot \hat{\boldsymbol{n}}\, dA \\ &= \mu \int_{S/C} \left\{ \boldsymbol{J} + \epsilon \frac{d\boldsymbol{E}}{dt} \right\} \cdot \hat{\boldsymbol{n}}\, dA \end{aligned} \tag{8.13}$$

or

$$\oint_C \boldsymbol{H} \cdot d\boldsymbol{\ell} = \int_{S/C} \left\{ \boldsymbol{J} + \frac{d\boldsymbol{D}}{dt} \right\} \cdot \hat{\boldsymbol{n}}\, dA \tag{8.14}$$

in terms of the magnetic field $\boldsymbol{H}$ and electric displacement $\boldsymbol{D}$. The origin of the term "displacement current" is obviously clear in this formulation.

$(\bar{\nabla} \times \bar{H})_i = \epsilon_{ijk} \frac{\partial}{\partial x_j} H_k$

Using vector calculus on our old form of Ampere's Law allows us to arrive at this same conclusion much more simply. If we take the divergence of Ampere's Law we get:

$\bar{\nabla} \cdot (\bar{\nabla} \times \bar{H})$

$= \frac{\partial}{\partial x_i} \epsilon_{ijk} \frac{\partial}{\partial x_j} H_k$

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \boldsymbol{H}) = 0 = \boldsymbol{\nabla} \cdot \boldsymbol{J} \tag{8.15}$$

$= \underbrace{\epsilon_{ijk}}_{\text{ASYM}} \underbrace{\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j}}_{\text{SYM}} H_k$

$= 0$

If we apply the divergence theorem to the law of charge conservation expressed as a flux integral above, we get its differential form:

$$\boldsymbol{\nabla} \cdot \boldsymbol{J} - \frac{\partial \rho}{\partial t} = 0 \tag{8.16}$$

and conclude that in general we can *not* conclude that the divergence of $\boldsymbol{J}$ vanishes in general as this expression requires, as there is no guarantee that $\frac{\partial \rho}{\partial t}$ vanishes everywhere in space. It only vanishes for "steady state currents" on a background of uniform charge density, justifying our calling this form of Ampere's law a magnetostatic version.

If we substitute in $\rho = \boldsymbol{\nabla} \cdot \boldsymbol{D}$ (Gauss's Law) for ρ, we can see that it *is* true that:

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \boldsymbol{H}) = 0 = \boldsymbol{\nabla} \cdot \left\{ \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t} \right\} \tag{8.17}$$

as an identity. A sufficient (but *not necessary!*) condition for this to be true is:

$$\nabla \times \boldsymbol{H} = \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t} \tag{8.18}$$

or

$$\nabla \times \boldsymbol{H} - \frac{\partial \boldsymbol{D}}{\partial t} = \boldsymbol{J}. \tag{8.19}$$

This expression is identical to the magnetostatic form in the cases where $\boldsymbol{D}$ is constant in time but respects charge conservation when the associated (displacement) field is changing.

We can now write the *complete* set of Maxwell's equations, including the Maxwell displacement current discovered by requiring formal invariance of the current and using *charge conservation* to deduce its form. Keep the latter in mind; it should not be surprising to us later when the law of charge conservation pops *out* of Maxwell's equations when we investigate their formal properties we can see that we deliberately encoded it into Ampere's Law as the MDC.

Anyway, here they are. Learn them. They need to be second nature as we will spend a considerable amount of time using them repeatedly in many, many contexts as we investigate electromagnetic radiation.

$$\begin{aligned}
\nabla \cdot \boldsymbol{D} &= \rho \quad \text{(GLE)} && (8.20)\\
\nabla \times \boldsymbol{H} - \frac{\partial \boldsymbol{D}}{\partial t} &= \boldsymbol{J} \quad \text{(AL)} && (8.21)\\
\nabla \cdot \boldsymbol{B} &= 0 \quad \text{(GLM)} && (8.22)\\
\nabla \times \boldsymbol{E} + \frac{\partial \boldsymbol{B}}{\partial t} &= 0 \quad \text{(FL)} && (8.23)
\end{aligned}$$

(where I introduce and obvious and permanent abbreviations for each equation by name as used throughout the rest of this text).

Aren't they pretty! The no-monopoles asymmetry is still present, but we now have *two symmetric* dynamic equations coupling the electric and magnetic fields and are ready to start studying *electrodynamics* instead of *electrostatics*.

Note well that the two *inhomogeneous* equations use the in-media forms of the electric and magnetic field. These forms are already coarse-grain averaged over the microscopic distribution of point charges that make up bulk matter. In a truly microscopic description, where we consider only bare charges wandering around in free space, we should use the free space versions:

$$\nabla \cdot \boldsymbol{E} = \frac{1}{\epsilon_0}\rho \quad (8.24)$$

$$\nabla \times \boldsymbol{B} - \mu_0\epsilon_0 \frac{\partial \boldsymbol{E}}{\partial t} = \mu_0 \boldsymbol{J} \quad (8.25)$$

$$\nabla \cdot \boldsymbol{B} = 0 \quad (8.26)$$

$$\nabla \times \boldsymbol{E} + \frac{\partial \boldsymbol{B}}{\partial t} = 0 \quad (8.27)$$

It is time to make these equations jump through some hoops.

8.2 Potentials

We begin our discussion of potentials by considering the two *homogeneous* equations. For example, if we wish to associate a potential with $\boldsymbol{B}$ such that $\boldsymbol{B}$ is the result of differentiating the potential, we observe that we can satisfy GLM *by construction* if we suppose a *vector potential* $\boldsymbol{A}$ such that:

$$\boldsymbol{B} = \nabla \times \boldsymbol{A} \quad (8.28)$$

In that case:

$$\nabla \cdot \boldsymbol{B} = \nabla \cdot (\nabla \times \boldsymbol{A}) = 0 \quad (8.29)$$

as an identity.

Now consider FL. If we substitute in our expression for $\boldsymbol{B}$:

$$\nabla \times \boldsymbol{E} + \frac{\partial \nabla \times \boldsymbol{A}}{\partial t} = 0$$

$$\nabla \times (\boldsymbol{E} + \frac{\partial \boldsymbol{A}}{\partial t}) = 0 \quad (8.30)$$

We can see that if we define:

$$\boldsymbol{E} + \frac{\partial \boldsymbol{A}}{\partial t} = -\nabla\phi \quad (8.31)$$

then

$$\nabla \times (\boldsymbol{E} + \frac{\partial \boldsymbol{A}}{\partial t}) = -\nabla \times \nabla\phi = 0 \quad (8.32)$$

is also an identity. This leads to:

$$\boldsymbol{E} = -\boldsymbol{\nabla}\phi - \frac{\partial \boldsymbol{A}}{\partial t} \tag{8.33}$$

Our next chore is to transform the *inhomogeneous* MEs into equations of motion for these potentials – motion because MEs (and indeed the potentials themselves) are now potentially *dynamical* equations and not just static. We do this by substituting in the equation for $\boldsymbol{E}$ into GLE, and the equation for $\boldsymbol{B}$ into AL. We will work (for the moment) in *free space* and hence will use the *vacuum* values for the permittivity and permeability.

The first (GLE) yields:

$$\begin{aligned}
\boldsymbol{\nabla}\cdot(-\boldsymbol{\nabla}\phi - \frac{\partial \boldsymbol{A}}{\partial t}) &= \frac{\rho}{\epsilon_0} \\
\nabla^2\phi + \frac{\partial(\boldsymbol{\nabla}\cdot\boldsymbol{A})}{\partial t} &= -\frac{\rho}{\epsilon_0}
\end{aligned} \tag{8.34}$$

$(\nabla\times A)_i = \epsilon_{ijk}\frac{\partial}{\partial x_j}A_k$

$[\nabla\times(\nabla\times A)]_n$

$= \epsilon_{nli}\frac{\partial}{\partial x_l}\epsilon_{ijk}\frac{\partial}{\partial x_j}A_k$

$= \epsilon_{inl}\epsilon_{ijk}\frac{\partial}{\partial x_l}\frac{\partial}{\partial x_j}A_k$

$= (\delta_{nj}\delta_{lk} - \delta_{nk}\delta_{lj})\, \partial/\partial x_l\, \partial/\partial x_j\, A_k$

$= \frac{\partial}{\partial x_k}\frac{\partial}{\partial x_n}A_k - \partial/\partial x_l\, \partial/\partial x_l\, A_n$

$= [\vec\nabla(\vec\nabla\cdot\vec A)]_n - [\nabla^2 A]_n$

The second (AL) is a bit more work. We start by writing it in terms of $\boldsymbol{B}$ instead of $\boldsymbol{H}$ by multiplying out the μ_0:

$$\begin{aligned}
\boldsymbol{\nabla}\times\boldsymbol{B} &= \mu_0\boldsymbol{J} + \mu_0\epsilon_0\frac{\partial \boldsymbol{E}}{\partial t} \\
\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\boldsymbol{A}) &= \mu_0\boldsymbol{J} + \mu_0\epsilon_0\frac{\partial}{\partial t}(-\boldsymbol{\nabla}\phi - \frac{\partial \boldsymbol{A}}{\partial t}) \\
-\nabla^2\boldsymbol{A} + \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{A}) &= \mu_0\boldsymbol{J} - \frac{1}{c^2}\boldsymbol{\nabla}\frac{\partial\phi}{\partial t} - \frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2} \\
\nabla^2\boldsymbol{A} + -\frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2} &= -\mu_0\boldsymbol{J} + \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{A}) + \boldsymbol{\nabla}\frac{1}{c^2}\frac{\partial\phi}{\partial t} \\
\nabla^2\boldsymbol{A} + -\frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2} &= -\mu_0\boldsymbol{J} + \boldsymbol{\nabla}\left(\boldsymbol{\nabla}\cdot\boldsymbol{A} + \frac{1}{c^2}\frac{\partial\phi}{\partial t}\right)
\end{aligned} \tag{8.35}$$

8.2.1 Gauge Transformations

Now comes the tricky part. The following is *very important* to understand, because it is a common feature to nearly all differential formulations of any sort of potential-based field theory, quantum or classical.

We *know* from our extensive study of elementary physics that there must be some freedom in the choice of ϕ and $\boldsymbol{A}$. The fields are physical and can potential physical field

$\frac{d}{dx}\,\phi = $ [illegible]

can add "constant"

be "directly" measured, we know that they are *unique* and cannot change. However, they are both defined in terms of *derivatives* of the potentials, so there is an *infinite family* of possible potentials that will *all* lead to the *same fields*. The trivial example of this, familiar from kiddie physics, is that the electrostatic potential is only defined with an arbitrary additive constant. No *physics* can depend on the choice of this constant, but some choices make problems more easily solvable than others. If you like, experimental physics depends on potential *differences*, not the absolute magnitude of the potential.

So it is now in grown-up electrodynamics, but we have to learn a new term. This freedom to add a constant potential is called *gauge freedom* and the different potentials one can obtain that lead to the same physical field are generated by means of a *gauge transformation*. A gauge transformation can be broadly defined as any formal, systematic transformation of the potentials that leaves the fields invariant (although in quantum theory it can be perhaps a bit more subtle than that because of the additional degree of freedom represented by the quantum phase).

← Schroedinger Eqn comment

As was often the case in elementary physics were we freely moved around the origin of our coordinate system (a gauge transformation, we now recognize) or decided to evaluate our potential (differences) from the inner shell of a spherical capacitor (another choice of gauge) we will *choose a gauge* in electrodynamics to make the solution to a problem as easy as possible or to build a solution with some desired characteristics that can be enforced by a "gauge condition" – a constraint on the final potentials obtained that one can show is within the range of possibilities permitted by gauge transformations.

However, there's a price to pay. Gauge freedom in non-elementary physics is a *wee bit broader* than "just" adding a constant, because gradients, divergences and curls in multivariate calculus are not simple derivatives.

Consider $\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}$. $\boldsymbol{B}$ must be unique, but many $\boldsymbol{A}$'s exist that correspond to any given $\boldsymbol{B}$. Suppose we have one such $\boldsymbol{A}$. We can obviously make a new $\boldsymbol{A}'$ that *has the same curl* by adding the gradient of *any scalar function* Λ. That is:

$$\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A} = \boldsymbol{\nabla} \times (\boldsymbol{A} + \boldsymbol{\nabla}\Lambda) = \boldsymbol{\nabla} \times \boldsymbol{A}' \tag{8.36}$$

We see that:

$$\boldsymbol{A}' = \boldsymbol{A} + \boldsymbol{\nabla}\Lambda \tag{8.37}$$

is a *gauge transformation* of the vector potential that leaves the field invariant.

Note that it probably isn't true that Λ can be *any* scalar function – if this were a math class I'd add caveats about it being nonsingular, smoothly differentiable at least one time, and so on. Even if a physics class I might say a word or two about it, so I just did. The point being that before you propose a Λ that isn't, you at least need to think about this sort of thing. However, great physicists (like Dirac) have subtracted out irrelevant infinities from potentials in the past and gotten away with it (he invented "mass renormalization" – basically a gauge transformation – when trying to derive a radiation reaction theory), so don't be too closed minded about this either.

It is also worth noting that this only shows that this is *a* possible gauge transformation of $\boldsymbol{A}$, not that it is sufficiently general to encompass *all* possible gauge transformations of $\boldsymbol{A}$. There may well be tensor differential forms of higher rank that cannot be reduced to being a "gradient of a scalar function" that still preserve $\boldsymbol{B}$. However, we won't have the algebraic tools to think about this at least until we reformulate MEs in relativity theory and learn that $\boldsymbol{E}$ and $\boldsymbol{B}$ are *not, in fact, vectors!* They are components of a second rank tensor, where both ϕ and $\boldsymbol{A}$ combine to form a first rank tensor (vector) in four dimensions.

This is quite startling for students to learn, as it means that there are *many* quantities that they might have thought are vectors that are not, in fact, vectors. And it matters – the tensor character of a physical quantity is closely related to the way it transforms when we e.g. change the underlying coordinate system. Don't worry about this quite yet, but it is something for us to think deeply about later.

Of course, if we change $\boldsymbol{A}$ in arbitrary ways, $\boldsymbol{E}$ will change as well! Suppose we have an $\boldsymbol{A}$ and ϕ that leads to some particular $\boldsymbol{E}$ combination:

$$\boldsymbol{E} = -\boldsymbol{\nabla}\phi - \frac{\partial \boldsymbol{A}}{\partial t} \tag{8.38}$$

If we transform $\boldsymbol{A}$ to $\boldsymbol{A}'$ by means of a gauge transformation (so $\boldsymbol{B}$ is preserved), we (in general) will still get a different $\boldsymbol{E}'$:

$$\begin{aligned}
\boldsymbol{E}' &= -\boldsymbol{\nabla}\phi - \frac{\partial \boldsymbol{A}'}{\partial t} \\
&= -\boldsymbol{\nabla}\phi - \frac{\partial}{\partial t}(\boldsymbol{A} + \boldsymbol{\nabla}\Lambda) \\
&= \boldsymbol{E} - \frac{\partial \boldsymbol{\nabla}\Lambda}{\partial t} \neq \boldsymbol{E} \qquad (8.39)
\end{aligned}$$

as there is no reason to expect the gauge term to vanish. This is baaaaad. We want to get the *same* $\boldsymbol{E}$.

To accomplish this, *as* we shift $\boldsymbol{A}$ to $\boldsymbol{A}'$ we must *also* shift ϕ to ϕ'. If we substitute an unknown ϕ' into the expression for $\boldsymbol{E}'$ we get:

$$\begin{aligned}
\boldsymbol{E}' &= -\boldsymbol{\nabla}\phi' - \frac{\partial}{\partial t}(\boldsymbol{A} + \boldsymbol{\nabla}\Lambda) \\
\boldsymbol{E}' &= -\boldsymbol{\nabla}\phi' - \frac{\partial \boldsymbol{A}}{\partial t} - \boldsymbol{\nabla}\frac{\partial \Lambda}{\partial t} \qquad (8.40)
\end{aligned}$$

We see that in order to make $\boldsymbol{E}' = \boldsymbol{E}$ (so it doesn't vary with the gauge transformation) we have to *subtract* a compensating piece to ϕ to form ϕ':

$$\phi' = \phi - \frac{\partial \Lambda}{\partial t} \qquad (8.41)$$

so that:

$$\begin{aligned}
\boldsymbol{E}' &= -\boldsymbol{\nabla}\phi' - \frac{\partial \boldsymbol{A}'}{\partial t} = -\boldsymbol{\nabla}\phi + \boldsymbol{\nabla}\frac{\partial \Lambda}{\partial t} - \frac{\partial \boldsymbol{A}}{\partial t} - \boldsymbol{\nabla}\frac{\partial \Lambda}{\partial t} \\
&= -\boldsymbol{\nabla}\phi - \frac{\partial \boldsymbol{A}}{\partial t} = \boldsymbol{E} \qquad (8.42)
\end{aligned}$$

In summary, we see that a fairly general gauge transformation that preserves *both* $\boldsymbol{E}$ and $\boldsymbol{B}$ is the following pair of *simultaneous* transformations of ϕ and $\boldsymbol{A}$. Given an arbitrary (but well-behaved) scalar function Λ:

$$\begin{aligned}
\phi' &= \phi - \frac{\partial \Lambda}{\partial t} \qquad (8.43) \\
\boldsymbol{A}' &= \boldsymbol{A} + \boldsymbol{\nabla}\Lambda \qquad (8.44)
\end{aligned}$$

will leave the derived fields invariant.

As noted at the beginning, we'd like to be able to *use* this gauge freedom in the potentials to choose potentials that are easy to evaluate or that have some desired formal property. There are two choices for gauge that are very common in electrodynamics, and you should be familiar with both of them.

8.2.2 The Lorentz Gauge

The Lorentz gauge, for a variety of reasons, is in my opinion the "natural" gauge of electrodynamics. For one thing, it is elegant in four dimensional space-time, and we are gradually working towards the epiphany that we should have formulated all of physics in four dimensional space-time from the beginning, even if we're considering non-relativistic phenomena. Working in it, most problems are relatively tractible if not actually easy. We will therefore consider it first.

Above we derived from MEs and their definitions the two equations of motion for the potentials ϕ and $\boldsymbol{A}$:

$$\nabla^2\phi + \frac{\partial(\boldsymbol{\nabla}\cdot\boldsymbol{A})}{\partial t} = -\frac{\rho}{\epsilon_0} \tag{8.45}$$

$$\nabla^2\boldsymbol{A} + -\frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2} = -\mu_0\boldsymbol{J} + \boldsymbol{\nabla}\left(\boldsymbol{\nabla}\cdot\boldsymbol{A} + \frac{1}{c^2}\frac{\partial\phi}{\partial t}\right) \tag{8.46}$$

If we can guarantee that we can *always* find a gauge transformation from a given solution to these equations of motion, ϕ_0, $\boldsymbol{A}_0$, a new one such that new $\phi, \boldsymbol{A}$ such that the new ones satisfy the constraint (the Lorentz gauge condition):

$$\boldsymbol{\nabla}\cdot\boldsymbol{A} + \frac{1}{c^2}\frac{\partial\phi}{\partial t} = 0 \tag{8.47}$$

then the two equations of motion both became the *inhomogeneous wave equation* for potential waves that propagate at the speed of light into or out of the charge-current source inhomogeneities. This precisely corresponds to our intuition of what should be happening, is elegant, symmetric, and so on. Later we'll see how *beautifully* symmetric it really is.

We must, however, prove that such a gauge condition actually exists. We propose:

$$\phi = \phi_0 - \frac{\partial\Lambda}{\partial t} \tag{8.48}$$

$$\boldsymbol{A} = \boldsymbol{A}_0 + \boldsymbol{\nabla}\Lambda \tag{8.49}$$

and substitute it into the desired gauge condition:

$$\begin{aligned}\boldsymbol{\nabla}\cdot\boldsymbol{A} + \frac{1}{c^2}\frac{\partial\phi}{\partial t} &= \boldsymbol{\nabla}\cdot\boldsymbol{A}_0 + \nabla^2\Lambda + \frac{1}{c^2}\frac{\partial\phi_0}{\partial t} - \frac{1}{c^2}\frac{\partial^2\Lambda}{\partial t^2} \\ &= 0\end{aligned} \tag{8.50}$$

or

$$\nabla^2\Lambda - \frac{1}{c^2}\frac{\partial^2\Lambda}{\partial t^2} = \boldsymbol{\nabla}\cdot\boldsymbol{A}_0 + \frac{1}{c^2}\frac{\partial\phi_0}{\partial t} = f(\boldsymbol{x},t) \tag{8.51}$$

for some computable inhomogeneous sourcevfunction $f(\boldsymbol{x},t)$.

This equation is solvable for an enormous range of possible $f(\boldsymbol{x},t)$s (basically, all well-behaved functions will lead to solutions, with issues associated with their support or possible singularities) so it seems at the very least "likely" that such a gauge transformation always exists for reasonable/physical charge-current distributions.

Interestingly, the gauge function Λ that permits the Lorentz condition to be satisfied so that $\phi, \boldsymbol{A}$ satisfy wave equations is itself the solution to a wave equation! It is also interesting to note that there is *additional* gauge freedom within the Lorentz gauge. For example, if one's original solution $\phi_0, \boldsymbol{A}_0$ itself satisfied the Lorentz gauge condition, then a gauge transformation to $\phi, \boldsymbol{A}$ where Λ is *any free scalar wave*:

$$\phi = \phi_0 - \frac{\partial\Lambda}{\partial t} \tag{8.52}$$

$$\boldsymbol{A} = \boldsymbol{A}_0 + \boldsymbol{\nabla}\Lambda \tag{8.53}$$

$$\nabla^2\Lambda - \frac{1}{c^2}\frac{\partial^2\Lambda}{\partial t^2} = 0 \tag{8.54}$$

continues to satisfy the Lorentz gauge condition. Not only are we nearly guaranteed that solutions that satisfy the Lorentz gauge condition exist, we have discovered an *infinity* of them, connected by a *restricted gauge transformation.*

In the Lorentz gauge, then, everything is a wave. The scalar and vector potentials, the derived fields, and the scalar gauge fields all satisfy wave equations. The result is independent of coordinates, formulates beautifully in special relativity, and exhibits (as we will see) the causal propagation of the fields or potentials at the speed of light.

The other gauge we must learn is not so pretty. In fact, it is really pretty ugly! However, it is still *useful* and so we must learn it. At the very least, it has a few important things to teach us as we work out the fields in the gauge.

8.2.3 The Coulomb or Transverse Gauge

Let us return to the equations of motion:

$$\nabla^2 \phi + \frac{\partial(\boldsymbol{\nabla} \cdot \boldsymbol{A})}{\partial t} = -\frac{\rho}{\epsilon_0} \tag{8.55}$$

$$\nabla^2 \boldsymbol{A} + -\frac{1}{c^2}\frac{\partial^2 \boldsymbol{A}}{\partial t^2} = -\mu_0 \boldsymbol{J} + \boldsymbol{\nabla}\left(\boldsymbol{\nabla} \cdot \boldsymbol{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t}\right) \tag{8.56}$$

There is another way to make at least one of these two equations simplify. We can just insist that:

$$\boldsymbol{\nabla} \cdot \boldsymbol{A} = 0. \tag{8.57}$$

It isn't so obvious that we can always choose a gauge such that this is true. Since we know we can start with the Lorentz gauge, though, let's look for Λ such that it is. That is, suppose we've found ϕ, $\boldsymbol{A}$ such that:

$$\boldsymbol{\nabla} \cdot \boldsymbol{A} + \frac{1}{c^2}\frac{\partial \phi}{\partial t} = 0 \tag{8.58}$$

As before, we propose:

$$\phi' = \phi - \frac{\partial \Lambda}{\partial t} \tag{8.59}$$

$$\boldsymbol{A}' = \boldsymbol{A} + \boldsymbol{\nabla}\Lambda \tag{8.60}$$

such that

$$\boldsymbol{\nabla} \cdot \boldsymbol{A}' = \boldsymbol{\nabla} \cdot \boldsymbol{A} + \nabla^2 \Lambda = 0. \tag{8.61}$$

If we substitute in the Lorentz gauge condition:

$$\boldsymbol{\nabla} \cdot \boldsymbol{A} = -\frac{1}{c^2}\frac{\partial \phi}{\partial t} \tag{8.62}$$

we get:

$$\nabla^2 \Lambda = -\boldsymbol{\nabla} \cdot \boldsymbol{A} = \frac{1}{c^2}\frac{\partial \phi'}{\partial t} = g(\boldsymbol{x}, t) \tag{8.63}$$

As before, provided that a solution to the equations of motion in the Lorentz gauge *exists*, we can in principle solve this equation for a Λ that makes $\boldsymbol{\nabla} \cdot \boldsymbol{A} = 0$ true. It is therefore a legitimate gauge condition.

If we use the Coulomb gauge condition (which we are now justified in doing, as we know that the resulting potentials will lead to the same physical

8.4 Magnetic Monopoles

Let us think for a moment about what MEs might be changed into if magnetic monopoles were discovered. We would then expect all four equations to be inhomogeneous:

$$\begin{aligned}
\nabla\cdot\boldsymbol{D} &= \rho_e \quad \text{(GLE)} && (8.110)\\
\nabla\times\boldsymbol{H}-\frac{\partial\boldsymbol{D}}{\partial t} &= \boldsymbol{J}_e \quad \text{(AL)} && (8.111)\\
\nabla\cdot\boldsymbol{H} &= \rho_m \quad \text{(GLM)} && (8.112)\\
\nabla\times\boldsymbol{D}+\frac{\partial\boldsymbol{H}}{\partial t} &= -\boldsymbol{J}_m \quad \text{(FL)} && (8.113)
\end{aligned}$$

or, in a vacuum (with units of magnetic charge given as ampere-meters, as opposed to webers, where 1 weber $= \mu_0$ ampere-meter):

$$\begin{aligned}
\nabla\cdot\boldsymbol{E} &= \frac{1}{\epsilon_0}\rho_e \quad \text{(GLE)} && (8.114)\\
\nabla\times\boldsymbol{B}-\epsilon_0\mu_0\frac{\partial\boldsymbol{E}}{\partial t} &= \mu_0\boldsymbol{J}_e \quad \text{(AL)} && (8.115)\\
\nabla\cdot\boldsymbol{B} &= \mu_0\rho_m \quad \text{(GLM)} && (8.116)\\
\nabla\times\boldsymbol{E}+\frac{\partial\boldsymbol{B}}{\partial t} &= -\mu_0\boldsymbol{J}_m \quad \text{(FL)} && (8.117)
\end{aligned}$$

(where we note that if we discovered an elementary magnetic monopole of magnitude g similar to the elementary electric monopolar charge of e we would almost certainly need to introduce additional constants – or arrangements of the existing ones – to establish its quantized magnitude relative to those of electric charge in suitable units as is discussed shortly).

There are two observations we need to make. One is that nature could be rife with magnetic monopoles already. In fact, every single charged particle could have a *mix* of both electric and magnetic charge. As long as the ratio g/e is a *constant*, we would be unable to tell.

This can be shown by looking at the following *duality transformation* which "rotates" the magnetic field into the electric field as it rotates the magnetic charge into the electric charge:

$$\begin{aligned}
\boldsymbol{E} &= \boldsymbol{E}'\cos(\Theta)+Z_0\boldsymbol{H}'\sin(\Theta) && (8.118)\\
Z_0\boldsymbol{D} &= Z_0\boldsymbol{D}'\cos(\Theta)+\boldsymbol{B}'\sin(\Theta) && (8.119)
\end{aligned}$$

$$Z_0\boldsymbol{H} = -\boldsymbol{E}'\sin(\Theta) + Z_0\boldsymbol{H}'\cos(\Theta) \tag{8.120}$$

$$\boldsymbol{B} = -Z_0\boldsymbol{D}'\sin(\Theta) + \boldsymbol{B}'\cos(\Theta) \tag{8.121}$$

where $Z_0 = \sqrt{\frac{\mu_0}{\epsilon_0}}$ is the impedance of free space (and has units of ohms), a quantity that (as we shall see) appears frequently when manipulating MEs.

Note that when the angle $\Theta = 0$, we have the ordinary MEs we are used to. However, all of our measurements of *force* would remain unaltered if we rotated by $\Theta = \pi/2$ and $\boldsymbol{E} = Z_0\boldsymbol{H}'$ in the old system.

However, if we perform such a rotation, we must *also* rotate the charge distributions in exactly the same way:

$$Z_0\rho_e = Z_0\rho_e'\cos(\Theta) + \rho_m'\sin(\Theta) \tag{8.122}$$

$$\rho_m = -Z_0\rho_e'\cos(\Theta) + \rho_m'\sin(\Theta) \tag{8.123}$$

$$Z_0\boldsymbol{J}_e = -\boldsymbol{J}_e'\cos(\Theta) + \boldsymbol{J}_m'\sin(\Theta) \tag{8.124}$$

$$\boldsymbol{J}_m = -Z_0\boldsymbol{J}_e'\sin(\Theta) + \boldsymbol{J}_m'\cos(\Theta) \tag{8.125}$$

It is left as an exercise to show that the monopolar forms of MEs are left invariant – things come in just the right combinations on both sides of all equations to accomplish this. In a nutshell, what this means is that it is merely a matter of convention to call all the charge of a particle electric. By rotating through an arbitrary angle theta in the equations above, we can recover an equivalent version of electrodynamics where electrons and protons have only magnetic charge and the electric charge is zero everywhere, but *where all forces and electronic structure remains unchanged* as long as *all particles have the same g/e ratio.*

When we search for magnetic monopoles, then, we are really searching for particles where that ratio is *different* from the dominant one. We are looking for particles that have zero electric charge and only a magnetic charge in the current frame relative to $\Theta = 0$. Monopolar particles might be expected to be a bit odd for a variety of reasons – magnetic charge is a pseudoscalar quantity, odd under time reversal, where electric charge is a scalar quantity, even under time reversal, for example, field theorists would *really really like* for there to be at least *one* monopole in the universe. Nobel-hungry graduate students wouldn't mind if that monopole came wandering through their monopole trap, either.

However, so far (despite a few false positive results that have proven dubious or at any rate unrepeatable) there is a lack of actual experimental evidence for monopoles. Let's examine just a bit of why the idea of monopoles is exciting to theorists.

8.4.1 Dirac Monopoles

Consider a electric charge e at the origin and an monopolar charge g at an arbitrary point on the z axis. From the generalized form of MEs, we expect the electric field to be given by the well-known:

$$\boldsymbol{E} = \frac{e\hat{\boldsymbol{r}}}{4\pi\epsilon_0 r^2} \tag{8.126}$$

at an arbitrary point in space. Similarly, we expect the magnetic field of the monopolar charge g to be:

$$\boldsymbol{B} = \frac{g\hat{\boldsymbol{r}}'}{4\pi\mu_0 r'^2} \tag{8.127}$$

where $\boldsymbol{r} = \boldsymbol{z} + \boldsymbol{r}'$.

The *momentum density* of this pair of fields is given as noted above by:

$$\boldsymbol{g} = \frac{1}{c^2}(\boldsymbol{E} \times \boldsymbol{H}) \tag{8.128}$$

and if one draws pictures and uses one's right hand to determine directions, it is clear that the field momentum is directed *around* the $e - g$ axis in the right handed sense. In fact the momentum follows circular tracks around this axis in such a way that the field has a non-zero static *angular* momentum.

The system obviously has zero total momentum from symmetry. This means one can use any origin to compute the angular momentum. To do so, we compute the angular momentum density as:

$$\frac{1}{c^2}\boldsymbol{r} \times (\boldsymbol{E} \times \boldsymbol{H}) \tag{8.129}$$

and integrate it:

$$\begin{aligned}
\boldsymbol{L}_{\text{field}} &= \frac{1}{c^2}\int \boldsymbol{r} \times (\boldsymbol{E} \times \boldsymbol{H}) dV \\
&= \frac{\mu_0 e}{4\pi}\int \frac{1}{r}\hat{\boldsymbol{n}} \times (\hat{\boldsymbol{n}} \times \boldsymbol{H}) dV \\
&= -\frac{\mu_0 e}{4\pi}\int \frac{1}{r}\left\{\boldsymbol{H} - \hat{\boldsymbol{n}}(\hat{\boldsymbol{n}} \cdot \boldsymbol{H})\right\} dV
\end{aligned} \tag{8.130}$$

over all space. Using the vector identity:

$$(\boldsymbol{a}\cdot\boldsymbol{\nabla})\hat{\boldsymbol{n}}f(r) = \frac{f(r)}{r}\{\boldsymbol{a}-\hat{\boldsymbol{n}}(\hat{\boldsymbol{n}}\cdot\boldsymbol{a})\} + \hat{\boldsymbol{n}}(\hat{\boldsymbol{n}}\cdot\boldsymbol{a})\frac{\partial f}{\partial r} \tag{8.131}$$

this can be transformed into:

$$\boldsymbol{L}_{\text{field}} = -\frac{e}{4\pi}\int(\boldsymbol{B}\cdot\boldsymbol{\nabla})\hat{\boldsymbol{n}}dV \tag{8.132}$$

Integrating by parts:

$$\boldsymbol{L}_{\text{field}} = \frac{e}{4\pi}\int(\boldsymbol{\nabla}\cdot\boldsymbol{B})\hat{\boldsymbol{n}}dV - \frac{e}{4\pi}\int_S \hat{\boldsymbol{n}}(\boldsymbol{B}\cdot\hat{\boldsymbol{n}}')dA \tag{8.133}$$

The surface term vanishes from symmetry because $\hat{\boldsymbol{n}}$ is radially away from the origin and averages to zero on a large sphere. $\boldsymbol{\nabla}\cdot\boldsymbol{B} = g\delta(\boldsymbol{r}-\boldsymbol{z})$ Thus we finally obtain:

$$\boldsymbol{L}_{\text{field}} = \frac{eg}{4\pi}\hat{\boldsymbol{z}} \tag{8.134}$$

There are a variety of arguments that one can invent that leads to an important conclusion. The arguments differ in details and in small ways quantitatively, and some are more elegant than this one. But this one is adequate to make the point. If we require that this field angular momentum be *quantized* in *units of* $\hbar$:

$$\frac{eg}{4\pi}\hat{\boldsymbol{z}} = m_z\hbar \tag{8.135}$$

we can conclude that the *product* of eg must be quantized. This is an important conclusion! It is one of the few approaches in physics that can give us insight as to why charge is quantized.

This conclusion was originally arrived at by (who else?) Dirac. However, Dirac's argument was more subtle. He created a monopole as a *defect* by constructing a vector potential that led to a monopolar field everywhere in space but which was *singular* on a single line. The model for this vector potential was that of an infinitely long solenoid stretching in from infinity along the $-z$ axis. This solenoid was in fact a *string* – this was in a sense the first quantum string theory.

The differential vector potential of a differential magnetic dipole $d\boldsymbol{m} = g d\boldsymbol{\ell}$ is:

$$d\boldsymbol{A}(\boldsymbol{x}) = -\frac{\mu_0}{4\pi}d\boldsymbol{m}\times\boldsymbol{\nabla}\left(\frac{1}{|\boldsymbol{x}-\boldsymbol{x}'|}\right) \tag{8.136}$$

so

$$A(x) = -\frac{\mu_0 g}{4\pi} \int_L d\ell \times \nabla \left(\frac{1}{|x - x'|} \right) \tag{8.137}$$

This can actually be evaluated in coordinates for specific lines L, e.g. a line from $-\infty$ to the origin along the $-z$ axis (to put a "monopole") at the origin. If one takes the curl of this vector potential one does indeed get a field of:

$$B = \frac{\mu_0}{4\pi} \frac{\hat{r}}{r^2} \tag{8.138}$$

everywhere but *on* the line L, where the field is singular. If we subtract away this singular (but highly confined – the field is "inside" the solenoid where it carries flux in from $-\infty$) we are left with the true field of a monopole everywhere but on this line.

Dirac insisted that an electron near this monopole would have to not "see" the singular string, which imposed a condition on its wavefunction. This condition (which leads to the same general conclusion as the much simpler argument given above) is beyond the scope of this course, but it is an interesting one and is much closer to the real arguments used by field theorists wishing to accomplish the same thing with a gauge transoformation and I encourage you to read it in e.g. Jackson or elsewhere.