PY406 - Electromagnetic Fields and Waves II

Hongwan Liu

Table of Contents

1 Preamble, Recap and Index Notation

By now, you have already taken several courses in electromagnetism, and you're about to start on yet *another* course in electromagnetism. You might be wondering, why are you constantly repeating this subject as a physics major?

The most immediate answer is that the electromagnetic force governs pretty much everything that you encounter in your daily life, and so it is incredibly important to study it, not just for technological reasons, but for understanding many physical processes that you might be interested in studying as a physicist. From a theoretical perspective though, electromagnetism is your first example of a **field theory**: the fundamental objects are not discrete particles at fixed positions in time (as they are in mechanics), but rather fields that are defined at every point in space and time. Electromagnetism as taught in this class is an important example of what we call a **classical field theory**. These are models where the main actors are fields that exactly obey a set of equations of motion, which we call **Maxwell's equations** in E&M. To be honest, E&M isn't the simplest classical field theory to study, but it happens to be one of the most important ones in nature. In some sense, you can view this class as a stepping stone toward **quantum field theories**, where quantum mechanics and field theories are brought together; *it is no exaggeration to say that quantum field theories underpin all of modern physics, and is our main theoretical framework for understanding many systems of interest* in condensed matter physics, particle physics, cosmology and beyond. The particular combination of quantum mechanics and electromagnetism is known as **quantum electrodynamics** (QED), one of the most remarkable physical theories we have ever developed, but that will be left for a higher level class.

Instead, we will focus on understanding the classical theory first. In PY405, you've learned a lot about Maxwell's equations, the equations of motion of this theory, and the be all and end all of classical E&M, really. In some sense, there's nothing else to study! In PY406, however, we'll be focusing a lot on electromagnetic *waves*. One very interesting result that we'll see almost immediately is that the fields in this field theory not only interact with particles, but can also exist and propagate on their own as *waves*: it is in fact entirely self-consistent to study electromagnetism *even if charges didn't exist*. We'll study how these waves propagate (through different media, in different geometries, etc.), how they transport energy and momentum, and how they are produced.

Another goal of this course is to take a deeper theoretical dive into E&M as a classical field theory. We'll see how the EM fields can equivalently be described in terms of potentials, and a unique property of these potentials known as **gauge freedom**. We will also see that E&M is, inescapably, a **relativistic** field theory; nothing about E&M makes sense without relativity. That's a deep result that will require us to spend some time understanding special relativity, but is a deeply satisfying perspective once you see it.

So let's dive right in! I'll begin by sketching some of the main results from PY405 that you've seen a great detail, with the goal of reminding you not just of the results, but also to give you some intuition for them.

1.1 Integral Form of Gauss's Law

The electromagnetic force acts between objects that carry **electric charge** q. q can be positive or negative, and we measure the electric charge in units of **Coulombs** in the SI system, which we'll use for the first half of this course. For two point charges with charges q and q, with q at the origin, and Q at position

$\boldsymbol{r}$, then the force between these two particles is given by **Coulomb's law** as

$$\mathbf{F} = \frac{qQ}{4\pi\epsilon_0 r^2}\hat{\boldsymbol{r}}\,, \tag{1.1}$$

where ϵ_0 is an empirically measured constant,[1]

$$\epsilon_0 = 8.854\,187\,818\,8(14) \times 10^{-12}\,\mathrm{C}^2\,\mathrm{N}^{-1}\,\mathrm{m}^{-2}\,. \tag{1.2}$$

The electric force is therefore a central force, acting along the line connecting two charges, and follows an inverse square law.

Since the force exerted by a point charge q on a charge Q located at some arbitrary point in space is directly proportional to Q, we can think of every point charge as *sourcing an electric field across all of space*, defined as

$$\mathbf{E} \equiv \frac{q}{4\pi\epsilon_0 r^2}\hat{\boldsymbol{r}}\,, \tag{1.3}$$

with $\mathbf{F} = Q\mathbf{E}$.

Thinking in terms of electric fields is extremely helpful. Electric fields originate from positive charges which act as sources of these fields, and then must terminate on negative charges, which act as sinks. You would expect, intuitively, that in a small region of space with no charges, that somehow the "number of field lines" going in should cancel with the "number of field lines" going out, since there are no sources or sinks of the field in that region.[2]

We make this intuition more precise by defining the **electric flux**. Consider some small surface with area $\mathrm{d}S$, with a normal vector given by $\hat{\mathbf{n}}$. Then the electric flux through this small surface $\mathrm{d}\Phi$ is defined as

$$\mathrm{d}\Phi \equiv \mathrm{d}\mathbf{S}\cdot\mathbf{E}\,, \tag{1.4}$$

where $\mathbf{S} \equiv \hat{\mathbf{n}}\,\mathrm{d}S$. This rigorously defines what we had intuitively meant by "number of field lines" flowing through a surface (see Fig. 1). Let's then now consider a region of space enclosing no charges: we would then expect that the total electric flux through the closed surface around this region is

$$\oint_{\mathrm{vac}} \mathrm{d}\mathbf{S}\cdot\mathbf{E} = 0\,, \tag{1.5}$$

where the circle is a reminder that we're integrating over a closed surface, and vac indicates that it is a region with no charges.

What if there are charges in the enclosed region? Let's first consider a point charge q, and calculate the total electric flux through a spherical surface of radius r centered at the charge. The electric field of the point charge points radially outward, and always intersects the surface ∂B of the sphere B perpendicularly; we therefore have

$$\oint_{\partial B} \mathrm{d}\mathbf{S}\cdot\mathbf{E} = 4\pi r^2 \cdot \frac{q}{4\pi\epsilon_0 r^2} = \frac{q}{\epsilon_0}\,. \tag{1.6}$$

Now, let's place this point charge in some arbitrary enclosed region V, and draw a little spherical region B around it of radius r (see Fig. 2). In the region between the surface of the sphere ∂B and ∂V the boundary of V, which we denote $V - B$, there are no charges, and so we expect as above that

$$\oint_{\partial(V-B)} \mathrm{d}\mathbf{S}\cdot\mathbf{E} = 0\,. \tag{1.7}$$

[1] *Do not be mystified by* $4\pi\epsilon_0$: this is just a proportionality constant that ensures that when you plug in Coulombs for the charges, and meters for the distance, you'll get back the right answer for the force in Newtons. If we had chosen another system of units, which we will later on in the course, we would need to adopt a different proportionality constant.

[2] This is in all honesty extremely misleading, since things are not as inevitable as it sounds here, but it does give the right result in the end.

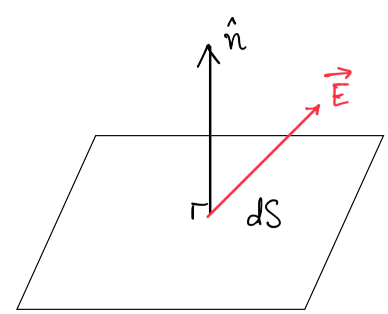

Figure 1: The definition of electric flux.

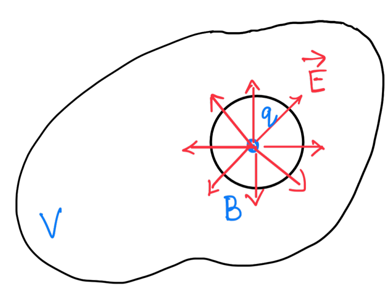

Figure 2: An arbitrary volume V with a point charge q surrounded by a spherical volume B.

But the total flux through $V - B$ can be divided into two parts: the flux penetrating $V - B$ from the point charge within B, and the flux exiting $V - B$ through the boundary of V, i.e.

$$\oint_{\partial(V-B)} \mathrm{d}\mathbf{S} \cdot \mathbf{E} = \oint_{\partial V} \mathrm{d}\mathbf{S} \cdot \mathbf{E} - \oint_{\partial B} \mathrm{d}\mathbf{S} \cdot \mathbf{E} \,. \tag{1.8}$$

Note the minus sign in front of the second term: the normal vector on the surface ∂B points inward *into* $V - B$, opposite to the direction you would define the normal vector on the surface with respect to the enclosed region $V - B$. We therefore conclude that

$$\oint_{\partial V} \mathrm{d}\mathbf{S} \cdot \mathbf{E} = \frac{q}{\epsilon_0} \,. \tag{1.9}$$

The intuitive generalization of this result gives us the following result, our first encounter with one of Maxwell's equations:

> The **integral form of Gauss's law** states that the total electric flux through any closed surface is proportional to the total charge Q_{enc} enclosed within that surface, i.e.
>
> $$\oint \mathrm{d}\mathbf{S} \cdot \mathbf{E} = \frac{Q_{\text{enc}}}{\epsilon_0} \,. \tag{1.10}$$

1.2 Divergence Theorem and the Differential Form of Gauss's Law

The integral form of Gauss's law relates the flux over some extended surface to the enclosed charge within some region, but it turns out that we can rewrite it as a *local* relationship between the electric field and the charge *density* at a particular point in space.

Let's begin by considering some volume V, enclosed by some closed boundary that we denote ∂V. Define the usual Cartesian coordinates (x, y, z), and chop up this volume into tiny boxes, each with volume $\mathrm{d}V = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$. Let's zoom in to one of the boxes deep in the interior of V. The flux through this tiny box is given by the sum of the fluxes through each of its six faces, but you'll see immediately that the flux through one face is equal but opposite to the flux through the same face of the neighboring box. This tells us that the flux through ∂V can simply be thought of as the sum of the fluxes through all of these tiny boxes, since all of the interior faces cancel out (see Fig. 3).

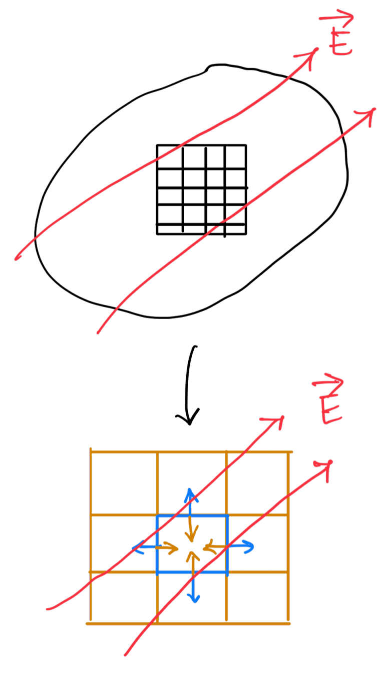

Figure 3: The flux through a small, interior box is equal and opposite to the fluxes through the neighboring boxes, due to the equal and opposite normal vectors on the shared faces.

Now, what is the flux through one of these boxes? The net flux through the two faces perpendicular to the x-axis is given by (see Fig. 4)

$$\begin{aligned} \mathrm{d}\Phi_x &= E_x|_{(x+dx,y,z)} \cdot \mathrm{d}y\,\mathrm{d}z - E_x|_{(x,y,z)} \cdot \mathrm{d}y\,\mathrm{d}z \\ &= \frac{\partial E_x}{\partial x} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \,, \end{aligned} \tag{1.11}$$

where $E_x|_{(x,y,z)}$ is the x-component of the electric field at the box located at (x, y, z), and I have simply Taylor expanded the field to the lowest order.[3] You can go on to show that likewise

[3] We are dealing with infinitesimal boxes here, and so we only need to keep the lowest order contribution.

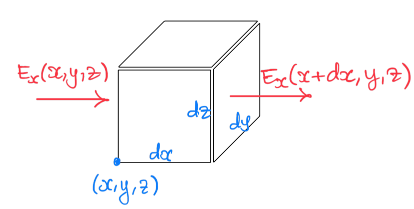

Figure 4: Flux along the two faces perpendicular to the x-axis.

$$\mathrm{d}\Phi_y = \frac{\partial E_y}{\partial y} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \,, \quad \mathrm{d}\Phi_z = \frac{\partial E_z}{\partial z} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \,. \tag{1.12}$$

Putting everything together, the total flux through the tiny box is

$$\mathrm{d}\Phi = (\nabla \cdot \mathbf{E})\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \,, \tag{1.13}$$

where I have defined the **divergence** of the electric field as

$$\nabla \cdot \mathbf{E} \equiv \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} \,. \tag{1.14}$$

End of Lecture: Wednesday, Jan 21 2026

Hence, using the argument above, we can see that the total flux through the closed surface ∂V is given by summing up the fluxes through all of these tiny boxes, i.e. we have the following result:

> The **divergence theorem** states that the total flux through a closed surface ∂V is equal to the volume integral of the divergence of the field over the enclosed volume V, i.e.
>
> $$\oint_{\partial V} \mathrm{d}\mathbf{S} \cdot \mathbf{E} = \int_V \mathrm{d}\Phi = \int_V \mathrm{d}^3\boldsymbol{r}\, (\nabla \cdot \mathbf{E}) \,. \tag{1.15}$$

Cool, let's apply this to the integral form of Gauss's law. Applying the divergence theorem, we find

$$\int_V \mathrm{d}^3\boldsymbol{r}\, (\nabla \cdot \mathbf{E}) = \frac{Q_{\mathrm{enc}}}{\epsilon_0} = \frac{1}{\epsilon_0} \int_V \mathrm{d}^3\boldsymbol{r}\, \rho \,, \tag{1.16}$$

where ρ is the **charge density**. Since the integral form of Gauss's law is true for *any* enclosing surface, the relation above must be true for any V, which means that the integrands on the LHS and RHS themselves must be equal. We therefore arrive at the following:

> The **differential form of Gauss's law**, given by
>
> $$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \,. \tag{1.17}$$

Note that this is a *local* relationship between the divergence of the electric field at some point in space, and the charge density at that point in space. It is also a *partial differential equation*, and solutions to it can only be obtained after specifying appropriate boundary conditions.

1.3 Index Notation Part I: Dot Products

We now take an important detour here to introduce a bit of notation, known as **index notation**, that we'll find very useful throughout the course.[4] We'll denote the three spatial components (x, y, z) as (x^1, x^2, x^3) respectively, and likewise denote the components of the 3D spatial vector $\mathbf{A}$ by A^i or A_i, where $i = 1, 2, 3$ corresponds to the x, y, z components respectively. Then we can write a dot product between two vectors $\mathbf{A} \cdot \mathbf{B}$ as

$$\mathbf{A} \cdot \mathbf{B} = A^1 B_1 + A^2 B_2 + A^3 B_3 = \sum_{i=1}^{3} A^i B_i = \sum_{j=1}^{3} A^j B_j \,. \tag{1.18}$$

Aside from the upper and lower index placements, this shouldn't be particularly surprising; we'll come back to upper and lower indices later. I've written the last expression by relabeling $i \to j$, as a reminder that the index is completely

[4] Our main textbook, Griffiths, decided against using this notation, but I'm not a fan of this decision for several reasons. First, it's really not so hard. Second, index notation makes many derivations much clearer, once you learn a few tricks. Finally, it is simply indispensable in relativity, and since we're going to cover that in this class, there simply is no point avoiding it.

arbitrary: they are **dummy indices**, and you can rename them to anything you want.

The next sleight-of-hand we'll introduce is known as the **Einstein summation convention**. In almost all physics equations, you will find that any sum over indices involves summing over all possible values of the index, and pairs of quantities with the same index. As such, we can drastically simplify our notation by *simply dropping the summation sign*, and remembering that *any repeated indices should be summed over all components*.[5] With this, the dot product is simply written as

$$\mathbf{A} \cdot \mathbf{B} = A^j B_j \, . \tag{1.19}$$

[5] We also frequently say that the repeated indices are **contracted**.

While we are on the topic of the dot product, let me throw some more notation at you just to get it out of the way:

> The **Kronecker delta** δ^i_j is defined as
>
> $$\delta^i_j = \begin{cases} 1 & \text{if } i = j \, , \\ 0 & \text{if } i \neq j \, . \end{cases} \tag{1.20}$$

This is often a useful object to have around. When it is contracted with another vector, $\delta^i_j A^j = A^i$, and you can think of the Kronecker delta as saying "everywhere you see a j, you can replace it with an i". In 3D Euclidean space, you can also write this as δ^{ij} or δ_{ij}, which all have the same meaning. As such, the dot product can be written as

$$\mathbf{A} \cdot \mathbf{B} = \delta_{ij} A^i B^j \, , \tag{1.21}$$

and you can think of δ_{ij} as an object that dots $\mathbf{A}$ into $\mathbf{B}$, or in physics jargon, contracts $\mathbf{A}$ with $\mathbf{B}$. Notice that $\delta_{ij} = \delta_{ji}$, i.e. the Kronecker delta is **symmetric** with respect to its indices. This is very important in manipulating expressions involving the Kronecker delta later on.

I'll gradually introduce this notation over the next few lectures, and give you a detailed set of rules for how to manipulate indices after a few examples. For now, let me just make a remark about the placement of indices: when dealing with 3D vectors in regular Euclidean space, *the placement of indices doesn't matter*, and this is the situation we'll be in for the first part of this course. However, to get into the habit for when we need it, *I will always write repeated indices with one upper and one lower index*, and you should try to get into this habit too.[6]

[6] It is extremely common in the literature, however, to write contracted indices as all subscripts, but you should learn the rules well before you can understand when they can be broken.

There is something very important about the dot product that I want to highlight here, which you probably already know, but it's worth stating clearly; remembering this will be key to helping you understand relativity later on. The components of a vector $\mathbf{A}$, A^i, necessarily depends on your choice of coordinate system: your x-axis could be someone else's y-axis, for example. But as you already know, the dot product of a vector with itself gives a length, and (without knowing anything about relativity at this point!), different observers with coordinate systems differing by a rotation or translation will all agree on the length of a vector. The same is true for the result of any dot product.

> A more jargony way of saying this is that the contraction of two vectors like $A_i B^i$ produces a **scalar**, a quantity that is invariant under any change in coordinate systems.

Finally, we will also define the notation $\partial_i \equiv \partial/\partial x^i$. With this, we can write

$$\nabla \cdot \mathbf{E} = \partial_i E^i \,. \tag{1.22}$$

1.4 Ampere's Law in Magnetostatics, Stokes' Theorem and the Lorentz Force Law

While charges source electric fields, moving charges or currents source magnetic fields $\mathbf{B}$. In the *magnetostatic* regime where we only deal with time-independent currents, the end result bears a strong resemblance to Gauss's law itself:

> The **integral form of Ampere's law under magnetostatic conditions** states that the line integral of the magnetic field $\mathbf{B}$ around a closed loop is proportional to the total current I_{enc} passing through any surface bounded by that loop, i.e. (see Fig. 5)
>
> $$\oint \mathrm{d}\mathbf{l} \cdot \mathbf{B} = \mu_0 I_{\text{enc}} \,. \tag{1.23}$$

The constant μ_0 is[7]

$$\mu_0 = 1.256\,637\,061\,27(20) \times 10^{-6}\,\mathrm{N\,A^{-2}} \,, \tag{1.24}$$

and the magnetic field is measured in teslas, $1\,\mathrm{T} = 1\,\mathrm{N\,A^{-1}\,m^{-1}}$, in the SI system. Comparing this with Gauss's law, you can see the resemblance: in Gauss's law, a closed surface integral of the electric field is proportional to the enclosed charge, while in Ampere's law, a closed line integral of the magnetic field is proportional to the enclosed current.

Once again, we would like to transform this into a differential form which is local, relating the magnetic field at a particular point in space to the sources at that same point. As before, let's take any surface bounded by the closed loop that we're considering, and chop it up into tiny surface elements, each with area $\mathrm{d}S$. If we now sum over the integral of $\mathbf{B}$ around each of these tiny surface elements, we can see that once again, the contributions from the interior edges cancel out, leaving only the contribution from the outer boundary (see Fig. 6).

Let's consider one of these tiny surface elements located at coordinates (x, y, z), and choose a coordinate system such that the normal vector $\hat{\mathbf{n}}$ of this surface element points along the z-axis, and the sides are aligned with the x and y axes. The line integral of $\mathbf{B}$ around this tiny surface element $\mathrm{d}C$ is then given by

$$\begin{aligned} \mathrm{d}C &= B_x|_{(x,y,z)} \cdot \mathrm{d}x - B_x|_{(x,y+\mathrm{d}y,z)} \cdot \mathrm{d}x - B_y|_{(x,y,z)} \cdot \mathrm{d}y + B_y|_{(x+\mathrm{d}x,y,z)} \cdot \mathrm{d}y \\ &= \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y} \right) \mathrm{d}x\,\mathrm{d}y \\ &= (\nabla \times \mathbf{B}) \cdot \mathrm{d}\mathbf{S} \,, \end{aligned} \tag{1.25}$$

where $\mathrm{d}\mathbf{S}$ is the area element vector for this tiny surface element, and the **curl** of the magnetic field is defined as

$$\nabla \times \mathbf{B} = \begin{pmatrix} \partial/\partial x \\ \partial/\partial y \\ \partial/\partial z \end{pmatrix} \times \begin{pmatrix} B_x \\ B_y \\ B_z \end{pmatrix} = \begin{pmatrix} \partial B_z/\partial y - \partial B_y/\partial z \\ \partial B_x/\partial z - \partial B_z/\partial x \\ \partial B_y/\partial x - \partial B_x/\partial y \end{pmatrix} . \tag{1.26}$$

Since we have written $\mathrm{d}C$ as a dot product, $\mathrm{d}C$ is a scalar which has the same value in every coordinate system, and is always given by the dot product $(\nabla \times \mathbf{B}) \cdot \mathrm{d}\mathbf{S}$ *regardless of the orientation of the surface element.*[8] We can therefore sum over all of these tiny surface elements to find the following result:

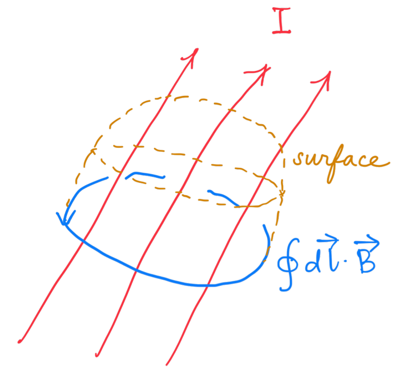

Figure 5: Illustration of Ampere's law.

[7] μ_0 used to be defined exactly as $4\pi \times 10^{-7}\,\mathrm{N\,A^{-2}}$, but since 2019, this has now become a measured quantity. Again, you should not be mystified by this: it's just a proportionality constant to ensure the right units, i.e. so that the magnetic field comes out as teslas.

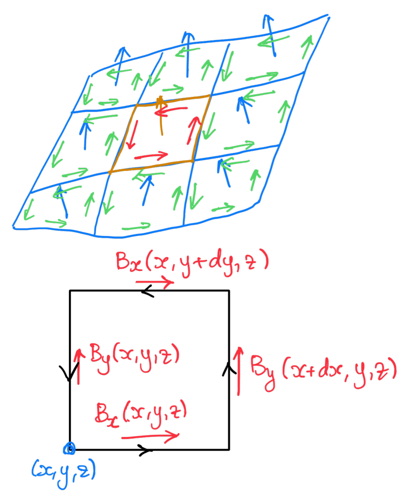

Figure 6: Cancellation of line integrals in the interior, and the magnetic field lines in a loop.

[8] This type of argument is tremendously powerful once you understand it, and crops up very frequently in physics.

Stokes' theorem states that the line integral of a vector field $\mathbf{B}$ around a closed loop ∂S is equal to the surface integral of the curl of that vector field over any surface S bounded by that loop, i.e.

$$\oint_{\partial S} \mathrm{d}\mathbf{l} \cdot \mathbf{B} = \int_S \mathrm{d}C = \int_S \mathrm{d}\mathbf{S} \cdot (\nabla \times \mathbf{B}) \,. \tag{1.27}$$

We can now apply this to the integral form of Ampere's law. Using Stokes' theorem applied to any closed loop ∂S and the surface S bounded by the loop, we have

$$\oint_{\partial S} \mathrm{d}\mathbf{l} \cdot \mathbf{B} = \int_S \mathrm{d}\mathbf{S} \cdot (\nabla \times \mathbf{B}) = \mu_0 I_{\mathrm{enc}} \,. \tag{1.28}$$

Writing I_{enc} as the surface integral of the current density $\mathbf{J}$ through the surface S, we find

$$\int_S \mathrm{d}\mathbf{S} \cdot (\nabla \times \mathbf{B}) = \mu_0 \int_S \mathrm{d}\mathbf{S} \cdot \mathbf{J} \,. \tag{1.29}$$

Since this is true for any loop, we obtain the **differential form of Ampere's law under magnetostatic conditions**,

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} \,. \tag{1.30}$$

Once again, this is a local relationship between the curl of the magnetic field at some point in space, and the current density at that same point in space. It is a partial differential equation, and can be solved once appropriate boundary conditions are specified.

Finally, magnetic fields also exert a force on particles with charge q, which when combined with the electric field gives the **Lorentz force law** for the force acting on a charged particle,

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \,, \tag{1.31}$$

where $\mathbf{v}$ is the velocity of the charged particle.

1.5 Index Notation Part II: Cross Products

We had previously seen that we could write $\nabla \cdot \mathbf{E} = \partial_i E^i$, and now we want to extend index notation to write cross products too. To do this, we need to introduce the following nifty object:

The 3D **Levi-Civita symbol** ϵ_{ijk} or ϵ^{ijk} is defined as

$$\epsilon_{ijk} = \begin{cases} +1 & \text{if } (i,j,k) \text{ is an even permutation of } (1,2,3) \,, \\ -1 & \text{if } (i,j,k) \text{ is an odd permutation of } (1,2,3) \,, \\ 0 & \text{if any two indices are equal} \,. \end{cases} \tag{1.32}$$

A permutation just means a swap between two indices, and so an even permutation of $(1,2,3)$ is any arrangement obtained by swapping an even number of times, and similarly for odd permutations.

Based on the definition, there are only 6 combinations of (i, j, k) that give non-zero values for ϵ_{ijk}, which are:

$$\begin{aligned}\epsilon_{123} &= \epsilon_{231} = \epsilon_{312} = +1\,,\\ \epsilon_{213} &= \epsilon_{132} = \epsilon_{321} = -1\,.\end{aligned} \tag{1.33}$$

We say that the Levi-Civita symbol is **antisymmetric**, because switching any two indices results in the minus sign. This will be a very important fact when we perform manipulations on the Levi-Civita symbol later on.

How does this help with the cross product? Let's consider two vectors $\mathbf{A}$ and $\mathbf{B}$: I claim that the i-th component of their cross product $\mathbf{A} \times \mathbf{B}$ can be written as

$$(\mathbf{A} \times \mathbf{B})_i = \sum_{j=1}^{3} \sum_{k=1}^{3} \epsilon_{ijk} A^j B^k = \epsilon_{ijk} A^j B^k\,, \tag{1.34}$$

where in the last step I've just dropped the summation signs by adopting the Einstein summation convention, where every repeated index is summed over all possible values. Let's check that this is true. Let's choose $i = 1$ to start with: even though the sum runs through many different indices, most of the entries are zero because of the properties of the Levi-Civita symbol, leaving only terms with $(j, k) = (2, 3)$ and $(3, 2)$ contributing. However, they contribute with opposite signs, and so we find

$$\epsilon_{ijk} A^j B^k = \epsilon_{123} A^2 B^3 + \epsilon_{132} A^3 B^2 = A^2 B^3 - A^3 B^2\,, \tag{1.35}$$

which is exactly $(\mathbf{A} \times \mathbf{B})_1$. You can check the other components yourself to see that it does indeed work out. We can thereefore also write Ampere's law in the magnetostatic limit as

$$\epsilon_{ijk} \partial^j B^k = \mu_0 J_i\,, \tag{1.36}$$

with ∂^j acting on B^k.

Example 1.1

This notation is extremely powerful, and you'll start to get a sense of why very soon, but let's start with something simple: consider the cross product of a vector $\mathbf{A}$ with itself, i.e. $\mathbf{A} \times \mathbf{A}$. Using index notation, we have

$$(\mathbf{A} \times \mathbf{A})_i = \epsilon_{ijk} A^j A^k\,. \tag{1.37}$$

Can we see that this is zero?
SOLUTION:
Indeed we can, by remembering that *repeated indices are part of a sum, and can be relabeled at will*! Relabeling $j \to k$ and $k \to j$, we find

$$\epsilon_{ijk} A^j A^k = \epsilon_{ikj} A^k A^j = -\epsilon_{ijk} A^k A^j\,, \tag{1.38}$$

where in the last line we've used the fact that the Levi-Civita symbol is antisymmetric. But since $A^j A^k = A^k A^j$, we see that

$$\epsilon_{ijk} A^j A^k = -\epsilon_{ijk} A^j A^k\,, \tag{1.39}$$

which is only possible if $\epsilon_{ijk} A^j A^k$ is zero. A quick way to understand this is that $A^j A^k$ is **symmetric** under the swap $j \leftrightarrow k$, while ϵ_{ijk} is antisymmetric; and the contraction of a symmetric and antisymmetric object is always zero.

Before moving on, I want to point out and summarize several incredibly important things about index notation here. *Understanding the following is crucial to using index notation effectively*!

1. Once you've rewritten a vector equation in index notation, **every quantity other than derivatives is just a number**, and not a vector anymore. So you can shuffle them up at will:

$$\epsilon_{ijk} A^j B^k = \epsilon_{ijk} B^k A^j = B^k \epsilon_{ijk} A^j = \cdots \tag{1.40}$$

   Derivatives of course have to follow the object that they are acting on, so e.g. $\partial_j E^j$ has to be kept together, but you can move the whole thing around as a unit: it's just a number!

2. In any equation with index notation, an index that occurs by itself is what we call a **free index**. *Free indices must match on both sides of an equation*: after all, if you want the LHS to equal the RHS, then both sides must agree component by component.

3. Contracted indices **can only occur in pairs**. *If you ever have more than two of the same index, something has gone horribly wrong, and you need to start over*. This is almost always because you used the same index for two different purposes, and you need to relabel some of them to avoid confusion.

4. **You can always relabel indices**. For free indices, you can relabel them consistently on both the LHS and the RHS. For contracted indices, you can relabel the pair with any letter you'd like, since they are dummy indices.

5. **Get into the habit of writing repeated indices with one upper and one lower index**. This is not strictly necessary in Euclidean space, but it will become necessary in relativity, and so it's best to get into the habit now. *You can also always take a pair of contracted indices, and swap their positions, i.e. turn $A_i B^i$ into $A^i B_i$ and vice versa.*

6. Pay attention to objects that have symmetric indices, and objects that have antisymmetric indices, since you can often simplify expressions by swapping indices around. Contracting a symmetric object with an antisymmetric object, for example, always gives zero.

7. Although the placement of indices doesn't matter in Euclidean space, you should get into the habit of trying to align the indices correctly, i.e. if it is an upper index on the LHS, it should be an upper index on the RHS too.

Example 1.2

Show the following second-derivative identities:

$$\nabla \times (\nabla T) = 0 \,, \qquad \nabla \cdot (\nabla \times \mathbf{v}) = 0 \,. \tag{1.41}$$

SOLUTION:
You've seen these identities in PY405, but now let's write things out slowly in index notation. For the first identity, ∇T is a vector, whose i-th compo-

nent we can write as $(\nabla T)_i = \partial_i T$. The curl can then be written as

$$\begin{aligned}[\nabla \times (\nabla T)]_i &= \epsilon_{ijk}\partial^j (\nabla T)^k \\ &= \epsilon_{ijk}\partial^j \partial^k T\,, \end{aligned} \tag{1.42}$$

with both partials acting on T. But now let's take a closer look at $\partial^j\partial^k$: except for pathological functions usually of no interest in physics, partial derivatives commute, and hence the object $\partial^j\partial^k = \partial^k\partial^j$ is symmetric with respect to $i \leftrightarrow j$. If this wasn't clear to you, you can go back to Example 1.1 and perform the same manipulation with the indices yourself. Since this is contracted with the antisymmetric Levi-Civita symbol, we conclude that $\epsilon_{ijk}\partial^j\partial^k T = 0$, and hence $\nabla \times (\nabla T) = 0$ as required.
The second identity is also likewise very straightforward: we can write it as

$$\nabla \cdot (\nabla \times \mathbf{v}) = \partial^i(\epsilon_{ijk}\partial^j v^k) = \epsilon_{ijk}\partial^i\partial^j v^k\,, \tag{1.43}$$

and for exactly the same reason above, this is zero as well. Note that ϵ_{ijk} is *just a number*: you can pull it out past the derivative as I have done here.

1.6 Index Notation Part III: Triple Products

Something else that comes up pretty often is the triple product, $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$, or in vector calculus, $\nabla \times (\nabla \times \mathbf{v})$. Let's try to write this in index notation:

$$\begin{aligned}[\nabla \times (\nabla \times \mathbf{v})]_i &= \epsilon_{ijk}\partial^j (\nabla \times \mathbf{v})^k \\ &= \epsilon_{ijk}\partial^j (\epsilon^{klm}\partial_l v_m) \\ &= \epsilon_{ijk}\epsilon^{klm}\partial^j\partial_l v_m \\ &= \epsilon_{kij}\epsilon^{klm}\partial^j\partial_l v_m\,. \end{aligned} \tag{1.44}$$

In the last line, I performed two swaps of the indices ($ijk \to ikj \to kij$) to bring k to the front. Again, let me remind you that as long as you are careful with which quantity the derivatives act on, you can shuffle around all of the quantities here as you like, since they are all just numbers. It's also good practice to check that the free indices on both sides agree at the end, and that all other indices are contracted in pairs.

Let's take a closer look at $\epsilon_{kij}\epsilon^{klm}$. Suppose $(i, j) = (1, 2)$: then when we perform the sum over k, only $k = 3$ leads to a nonzero result, and furthermore, there are only two nonzero combinations for (l, m), $(1, 2)$ or $(2, 1)$, with the first combination leading to $+1$ and the second to -1. A very compact way to write this result is as follows:

$$\epsilon_{kij}\epsilon^{klm} = \delta_i^l\delta_j^m - \delta_i^m\delta_j^l\,, \tag{1.45}$$

a result well worth keeping in your back pocket. Applying this to the triple product, and remembering that contracting with the Kronecker delta δ_i^l just tells you to replace i with l (or l with i, whichever you prefer), we get

$$\begin{aligned}[\nabla \times (\nabla \times \mathbf{v})]_i &= \epsilon_{kij}\epsilon^{klm}\partial^j\partial_l v_m \\ &= (\delta_i^l\delta_j^m - \delta_i^m\delta_j^l)\partial^j\partial_l v_m \\ &= \partial^j\partial_i v_j - \partial^j\partial_j v_i\,. \end{aligned} \tag{1.46}$$

Let's take a look at each term separately. For the first term, partial derivatives commute, and so we can write $\partial^j \partial_i v_j = \partial_i \partial^j v_j$, which is just $\partial_i (\nabla \cdot \mathbf{v}) = [\nabla(\nabla \cdot \mathbf{v})]_i$. For the second term, we see that

$$\partial^j \partial_j = \partial^1 \partial_1 + \partial^2 \partial_2 + \partial^3 \partial_3 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \equiv \nabla^2 \,, \tag{1.47}$$

and so, $\partial^j \partial_j v_i = [\nabla^2 \mathbf{v}]_i$. Putting everything together, we recover the identity

$$\nabla \times (\nabla \times \mathbf{v}) = \nabla(\nabla \cdot \mathbf{v}) - \nabla^2 \mathbf{v} \,. \tag{1.48}$$

As you can see, index notation makes this derivation very straightforward, once you're familiar with the identity given in Eq. (1.45). You almost never need any other identity beyond this one in 3D, and if you do, the Wikipedia article on the Levi-Civita symbol has a comprehensive list. This saves you from having to remember all kinds of vector calculus identities, and makes derivations and manipulations much more transparent, once you get used to it!

1.7 Gauss's Law for Magnetism and Faraday's Law

The mathematical structure for the two remaining Maxwell's equations that we haven't discussed so far are very similar to what we've seen for Gauss's law and Ampere's law, and so I'll just go through them rapidly.

First, we have Gauss's law for magnetism, which states that there are no sources or sinks for magnetic fields, i.e. there is no equivalent of "magnetic charge" in classical electromagnetism.[9]

[9] This is not to say that they're mathematical impossible, but if they did exist, then we would need to study a slightly different field theory than the one that we're studying here. This is not a disturbing possibility at all: classical electromagnetism is already incomplete, having no quantum effects at all, and so you're already studying an approximation here.

With no magnetic charges, the **integral form for Gauss's law for magnetism** reads

$$\oint \mathrm{d}\mathbf{S} \cdot \mathbf{B} = 0 \,, \tag{1.49}$$

for any surface enclosing a finite volume.

Applying the divergence theorem to an arbitrary volume V with enclosing surface ∂V,

$$\oint_{\partial V} \mathrm{d}\mathbf{S} \cdot \mathbf{B} = \int_V \mathrm{d}^3\boldsymbol{r} \, (\nabla \cdot \mathbf{B}) = 0 \,, \tag{1.50}$$

leading to the **differential form of Gauss's law for magnetism**,

$$\nabla \cdot \mathbf{B} = 0 \,. \tag{1.51}$$

Finally, toward the end of PY405, you started thinking about *dynamics*, where quantities started to have time-dependence. One of the key results was that electric fields are not just sourced by charges, but can also be sourced by magnetic fields.

This is encapsulated in the **integral form of Faraday's law**, which states that the line integral of the electric field $\mathbf{E}$ around a close loop is equal to the *negative* of the rate of change in the **magnetic flux** through any surface bounded by the loop, i.e.

$$\oint \mathrm{d}\mathbf{l} \cdot \mathbf{E} = -\frac{\mathrm{d}}{\mathrm{d}t} \int \mathrm{d}\mathbf{S} \cdot \mathbf{B} \,. \tag{1.52}$$

Applying Stokes' theorem to the integral on the left around some arbitrary closed loop ∂S of a surface S bounded by the loop, we have

$$\oint_{\partial S} \mathrm{d}\mathbf{l} \cdot \mathbf{E} = \int_S \mathrm{d}\mathbf{S} \cdot (\nabla \times \mathbf{E}) = -\frac{\mathrm{d}}{\mathrm{d}t} \int_S \mathrm{d}\mathbf{S} \cdot \mathbf{B} = -\int_S \mathrm{d}\mathbf{S} \cdot \frac{\partial \mathbf{B}}{\partial t}, \tag{1.53}$$

leading us to the **differential form of Faraday's law**,

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}. \tag{1.54}$$

1.8 Charge Conservation and the Ampere-Maxwell Law

Charge is a conserved quantity: within any volume V, any change in the total charge within the volume must be carried away by currents flowing across the boundary of the volume, ∂V. Mathematically, this is expressed as

$$\frac{\mathrm{d}Q_{\mathrm{enc}}}{\mathrm{d}t} + \oint_{\partial V} \mathrm{d}\mathbf{A} \cdot \mathbf{J} = 0 \tag{1.55}$$

where Q_{enc} is the total charge enclosed within the volume V. Writing Q_{enc} as the volume integral of the charge density ρ, we have

$$\frac{\mathrm{d}Q_{\mathrm{enc}}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t} \int_V \mathrm{d}^3\boldsymbol{r}\, \rho = \int_V \mathrm{d}^3\boldsymbol{r}\, \frac{\partial \rho}{\partial t}. \tag{1.56}$$

On the other hand, applying the divergence theorem to the surface integral of the current density, we have

$$\oint_{\partial V} \mathrm{d}\mathbf{S} \cdot \mathbf{J} = \int_V \mathrm{d}^3\boldsymbol{r}\, (\nabla \cdot \mathbf{J}). \tag{1.57}$$

Putting the last three equations together, we have

$$\int_V \mathrm{d}^3\boldsymbol{r} \left(\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} \right) = 0, \tag{1.58}$$

which is true for any arbitrary volume V.

Therefore, the integrand itself must be zero, which gives us the **continuity equation** expressing local charge conservation,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0, \tag{1.59}$$

Or, in index notation,

$$\frac{\partial \rho}{\partial t} + \partial_i J^i = 0. \tag{1.60}$$

This will be the last time I will write both vector notation and index notation out explicitly; you should feel comfortable moving between them from now on.[10]

Let's take a closer look at Ampere's law in the magnetostatic limit, $\epsilon_{ijk}\partial^j B^k = \mu_0 J_i$. We know that taking the divergence of the LHS gives

$$\partial^i(\epsilon_{ijk}\partial^j B^k) = \epsilon_{ijk}\partial^i\partial^j B^k = 0, \tag{1.61}$$

since we are contracting a symmetric object $\partial^i\partial^j$ with the antisymmetric Levi-Civita symbol (see Ex. 1.2 if this is still confusing). On the RHS, however, we have

$$\mu_0 \partial^i J_i = -\mu_0 \partial_t \rho = -\mu_0 \partial_t(\epsilon_0 \partial^i E_i) = -\mu_0 \partial^i(\epsilon_0 \partial_t E_i), \tag{1.62}$$

[10] Sometimes, I will also write ∂_t for the partial derivative with respect to time.

where in the second last step we've used Gauss's law. The RHS is clearly nonzero in general, and so we have a problem: Ampere's law in the magnetostatic limit is clearly incompatible with local charge conservation, once quantities are allowed to vary in time. But the result in the last equation tells us how to fix it: we should replace[11]

$$\mu_0 J_i \to \mu_0(J_i + \epsilon_0 \partial_t E_i) \,. \tag{1.63}$$

so that now, the divergence of $\mu_0(J_i + \epsilon_0 \partial_t E_i)$ is zero identically, in agreement with the LHS.

[11] While this argument is much cleaner, this was not how Maxwell arrived at his correction to Ampere's law historically.

Our new and improved Ampere's law, known as the **Ampere-Maxwell law** is

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \,. \tag{1.64}$$

1.9 Summary

Let's summarize what we've just recapped.

The electric and magnetic fields are governed by the following four equations, known as **Maxwell's equations**:

$$\begin{aligned} \nabla \cdot \mathbf{E} &= \frac{\rho}{\epsilon_0} \,, & \nabla \cdot \mathbf{B} &= 0 \,, \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \,, & \nabla \times \mathbf{B} &= \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \,. \end{aligned} \tag{1.65}$$

These equations tell you how the two fields, $\mathbf{E}$ and $\mathbf{B}$, are sourced by charges *and* other fields.

On the other hand, the **Lorentz force law** tells you how fields act on charged particles with charge q and velocity $\mathbf{v}$,

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \,. \tag{1.66}$$

or in terms of force per unit volume $\mathbf{f}$ on an arbitrary distribution of charge and current densities ρ and $\mathbf{J}$,

$$\mathbf{f} = \rho \mathbf{E} + \mathbf{J} \times \mathbf{B} \,. \tag{1.67}$$

Together, Maxwell's equations and the Lorentz force law tell you everything you really need to know about classical electromagnetism! Everything else is built on top of these equations. You can now solve them, with boundary conditions, to obtain the behavior of fields and charged particles as a function of space and time.

End of Lecture: Friday, Jan 23 2026

2 Conservation Laws

Momentum (both linear and angular) and energy conservation are not only extremely important principles in physics, but also incredibly powerful in mechanics. In electromagnetism, momentum and energy are also conserved, *but only if we recognize that the fields themselves carry momentum and energy.*[12] As you have already seen, the electric and magnetic fields do carry energy, but now we'll also see that they carry momentum too.

[12] The ultimate reason for why momentum and energy are conserved is the same as in mechanics: the laws of physics are invariant under spatial and time translations, and **Noether's theorem**, one of the most remarkable results in physics, tells us that such *symmetries* lead to conserved quantities that we call momentum and energy respectively.

2.1 What is a Conservation Law?

But before we look at these conservation laws, let's take a step back and look at what a conservation law looks like. Let's consider the case of **charge conservation** as a concrete example, although you should keep in mind that vector quantities can also be conserved. When we say Q is conserved, we mean that in some system, the total charge doesn't change with time, i.e.

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = 0\,. \tag{2.1}$$

This is an example of a **global** conservation law, which is obtained by integrating over the entire system.

But the statement of charge conservation is actually stronger than this. Let's look at any arbitrary volume V within the system. The total charge Q inside this volume V can change with time, but only by charge flow, i.e. currents, across the boundary of the volume, ∂V. Mathematically, we say that[13]

[13] For this chapter, I will use $\mathrm{d}\mathbf{A}$ for the infinitesimal area vector, since the Poynting vector is almost always written as $\mathbf{S}$.

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = -\oint_{\partial V} \mathrm{d}\mathbf{A}\cdot\mathbf{J}\,, \tag{2.2}$$

where the surface integral is simply the total current flowing out of the volume V. By the divergence theorem, however, we know that

$$\oint_{\partial V} \mathrm{d}\mathbf{A}\cdot\mathbf{J} = \int_V \mathrm{d}^3\boldsymbol{r}\,(\nabla\cdot\mathbf{J})\,. \tag{2.3}$$

Furthermore, we can write

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\int \mathrm{d}^3\boldsymbol{r}\,\rho = \int \mathrm{d}^3\boldsymbol{r}\,\frac{\partial\rho}{\partial t}\,. \tag{2.4}$$

Putting everything together, we find that for any arbitrary volume V,

$$\int \mathrm{d}^3\boldsymbol{r}\,\frac{\partial\rho}{\partial t} = -\int \mathrm{d}^3\boldsymbol{r}\,(\nabla\cdot\mathbf{J})\,, \tag{2.5}$$

but since this applies for any volume, we must always have

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{J} = 0\,. \tag{2.6}$$

This is an example of a **local** conservation law, expressing the idea that any change in charge at some point must be due to some inflow or outflow from the surroundings. But you can see that there is nothing particularly special about charge. Any locally conserved quantity must always have the following structure: the partial derivative with respect to a density, plus the divergence of a current density, being equal to zero. *Every local conservation law follows this structure.*

2.2 Conservation of Energy

We'll begin our study of conservation laws in electromagnetism by examining energy conservation. In PY405, you had already seen that the electromagnetic fields carry energy density: for some region of space with electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$, the energy density $u_{\rm EM}$ of the fields is

$$u_{\rm EM} = \frac{1}{2}\epsilon_0 \mathbf{E}^2 + \frac{1}{2\mu_0}\mathbf{B}^2\,. \tag{2.7}$$

If energy is conserved locally, we would expect to be able to write

$$\frac{\partial u_{\rm EM}}{\partial t} + \nabla\cdot\mathbf{S} = 0\,, \tag{2.8}$$

for some energy current density $\mathbf{S}$. Let's go ahead and take the derivative of $u_{\rm EM}$ in index notation:[14]

$$\frac{\partial u_{\rm EM}}{\partial t} = \epsilon_0 E^i \partial_t E_i + \frac{1}{\mu_0} B^i \partial_t B_i\,. \tag{2.9}$$

[14] In case this is not clear, $\partial_t \mathbf{E}^2 = \partial_t(E^i E_i) = 2E^i\partial_t E_i$, and similarly for $\partial_t \mathbf{B}^2$.

Now, we can use Faraday's law and the Ampere-Maxwell law to rewrite the time derivatives of the fields:

$$\begin{aligned}\frac{\partial u_{\rm EM}}{\partial t} &= \epsilon_0 E^i \frac{1}{\mu_0\epsilon_0}\left(\epsilon_{ijk}\partial^j B^k - \mu_0 J_i\right) + \frac{1}{\mu_0}B^i(-\epsilon_{ijk}\partial^j E^k)\\ &= \frac{1}{\mu_0}\left(E^i\epsilon_{ijk}\partial^j B^k - B^i\epsilon_{ijk}\partial^j E^k\right) - E^i J_i\\ &= \frac{\epsilon_{ijk}}{\mu_0}\left(E^i\partial^j B^k - B^i\partial^j E^k\right) - E^i J_i\,. \end{aligned} \tag{2.10}$$

To make further progress, we're going to do some index relabeling. Remember that contracted indices can always be renamed, since these indices are just dummy indices to be summed over anyway. With this in mind, let's take the term $\epsilon_{ijk}B^i\partial^j E^k$: relabeling $i\to k$, $k\to i$, giving

$$\epsilon_{ijk}B^i\partial^j E^k = \epsilon_{kji}B^k\partial^j E^i = -\epsilon_{ijk}B^k\partial^j E^i\,, \tag{2.11}$$

where in the last step I've used the fact that $\epsilon_{kji} = -\epsilon_{ijk}$, since we need three swaps to go from one to the other. With this swap, both terms in the parenthesis in the expression before this have the same index on E and B, i.e.

$$\begin{aligned}\frac{\partial u_{\rm EM}}{\partial t} &= \frac{\epsilon_{ijk}}{\mu_0}\left(E^i\partial^j B^k + B^k\partial^j E^i\right) - E^i J_i\\ &= \frac{1}{\mu_0}\epsilon_{ijk}\partial^j(E^i B^k) - E^i J_i\,, \end{aligned} \tag{2.12}$$

where in the last line I have used the product rule. At this point, we can move the Levi-Civita symbol into the parenthesis (the derivative doesn't act on it, since it is just a constant), and noting that $\epsilon_{ijk}E^iB^k = -\epsilon_{jik}E^iB^k = -(\mathbf{E}\times\mathbf{B})_j$, we finally see that

$$\frac{\partial u_{\rm EM}}{\partial t} = -\frac{1}{\mu_0}\partial^j(\mathbf{E}\times\mathbf{B})_j - E^i J_i\,, \tag{2.13}$$

or in vector notation,

$$\frac{\partial u_{\rm EM}}{\partial t} + \nabla\cdot\mathbf{S} + \mathbf{E}\cdot\mathbf{J} = 0\,, \tag{2.14}$$

which is known as **Poynting's theorem**, with

$$\mathbf{S} \equiv \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B} \tag{2.15}$$

the **Poynting vector**.

Let's examine Poynting's theorem a little more closely. In vacuum, with no currents present, we obtain

$$\frac{\partial u_{\rm EM}}{\partial t} + \nabla \cdot \mathbf{S} = 0\,, \qquad \text{(vacuum)}\,, \tag{2.16}$$

which precisely as the form of a conservation law, while the Poynting vector $\mathbf{S}$ is clearly playing the role of an **energy flux density**. But what happens when there is a medium? What is the meaning of $\mathbf{E} \cdot \mathbf{J}$? Recall that the force per unit volume acting on a distribution of charge in the presence of an EM field is

$$\mathbf{f} = \rho \mathbf{E} + \mathbf{J} \times \mathbf{B}\,. \tag{2.17}$$

Taking the dot product of this with $\mathbf{v}$, we get on the LHS a quantity that is power per unit volume transferred to the charge, while on the RHS, we obtain

$$\mathbf{v} \cdot \mathbf{f} = \rho \mathbf{v} \cdot \mathbf{E} = \mathbf{J} \cdot \mathbf{E}\,, \tag{2.18}$$

with the term with the cross product vanishing since $\mathbf{J} = \rho \mathbf{v}$ is parallel to $\mathbf{v}$. Clearly then, $\mathbf{J} \cdot \mathbf{E}$ is the *rate of energy density transferred to the charges*, and

$$\frac{\partial u_{\rm mech}}{\partial t} = \mathbf{J} \cdot \mathbf{E}\,, \tag{2.19}$$

i.e. it is the time derivative of the mechanical energy density of the charges. Poynting's theorem can therefore be rewritten as

$$\frac{\partial}{\partial t} \left(u_{\rm EM} + u_{\rm mech} \right) + \nabla \cdot \mathbf{S} = 0\,, \tag{2.20}$$

which is now a true local conservation law of energy, including both those stored in the fields and in the mechanical energy of the charges!

By now, you should know how to go from a conservation-law type equation to a global conservation law, by performing a volume integral, and obtaining an integral form of Poynting's theorem. For any arbitrary volume V with enclosing surface ∂V, we have

$$\frac{\rm d}{{\rm d}t} \int_V {\rm d}V \, \left(u_{\rm EM} + u_{\rm mech} \right) = - \oint_{\partial V} {\rm d}\mathbf{A} \cdot \mathbf{S}\,, \tag{2.21}$$

after applying the divergence theorem. The LHS is simply the rate of change of the total energy in V, while the RHS is the net rate of energy density flowing out of the volume through its boundary.

End of Lecture: Monday, Jan 26 2026

Example 2.1

Perhaps one reason why energy conservation is so much less useful in electromagnetism is that energy flow is extremely unintuitive. Here is one example illustrating this point: there is a nice discussion in Feynman's Lectures on Physics Vol. II (27.3) on this.
Consider a straight wire with finite conductivity of length L with a circular cross section of radius a, and a constant current I driven through the wire by a uniform electric field $\mathbf{E}$ inside and parallel to the wire (see Fig. 7). We know that energy is constantly being dissipated as heat in the resistor; otherwise, the charges would accelerate through the wire and the current would not be constant. But where does this energy ultimately come from? Let's use Poynting's theorem to find out.
First, the potential difference across this wire is simply $V = EL$. The magnetic field B on the other hand is given by Ampere's law. Choosing a circular loop at radius a,

$$\oint \mathrm{d}\mathbf{l} \cdot \mathbf{B} = \mu_0 I \implies B = \frac{\mu_0 I}{2\pi a}\,, \tag{2.22}$$

with the direction of B in the azimuthal direction around the wire, wrapping counterclockwise around the current. The Poynting vector $\mathbf{S}$ then points radially inward toward the wire, and since the electric and magnetic fields are orthogonal, it has magnitude

$$S = \frac{1}{\mu_0}\frac{V}{L}\frac{\mu_0 I}{2\pi a} = \frac{VI}{2\pi a L}\,. \tag{2.23}$$

The total energy flux into the wire is then obtained by integrating over the cylindrical area of the wire,

$$\int \mathrm{d}\mathbf{A} \cdot \mathbf{S} = 2\pi a L \cdot \frac{VI}{2\pi a L} = VI\,, \tag{2.24}$$

which is what you expect from standard circuit theory!
It's great that we recovered the right thing, but the picture is somewhat unexpected: the energy is flowing in due to the fields *within the wire*, which is a highly confusing idea. And yet, this is right: energy is flowing radially into the wire, and being ultimately converted into heat.

2.3 Scalars, Vectors and Tensors

We now take a little bit of a detour to discuss some mathematical concepts that we'll find extremely useful when we talking about conservation of momentum, but then more generally as well when we talk about relativity later on (and indeed, in all of physics). Please pay attention, because this is the source of a lot of unnecessary mystery!

In your physics career, virtually everything you have encountered so far can be labeled as either a **scalar** or a **vector**. The handwavy definition you are aware of is that scalars are quantities with only magnitude, while vectors have both magnitude and direction. But now, as things get more complicated, we're going to have to be a bit more precise. Ultimately, scalars and vectors are abstract objects that one would study in a course on **linear algebra**, but for our purposes here, we won't need the full machinery of that.

Let's think about how a physics calculation usually proceeds. We're always trying to describe some physical situation mathematically, and so invariably the

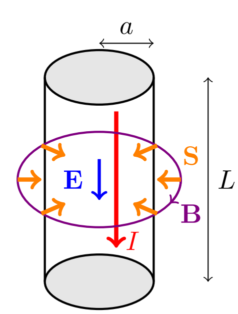

Figure 7: A cylindrical wire carrying current I with electric field $\mathbf{E}$ parallel to the wire, magnetic field $\mathbf{B}$ in the azimuthal direction, and Poynting vector $\mathbf{S}$ pointing radially inward.

first thing that we do is choose a **coordinate system**, i.e. a way to assign a number to every point in space.[15] A typical choice in 3D space that we make is the Cartesian coordinate system, with three orthogonal axes x, y and z, so that every point is (x, y, z), and we usually align these axes to make our lives as easy as possible. But the choice of coordinate system is ultimately arbitrary: we could have chosen to rotate our axes, or even use a completely different coordinate system altogether, such as spherical or cylindrical coordinates.

[15] and later, in relativity, in *spacetime*.

> But one thing that we fully expect is that *physics does not depend on the coordinate system you choose*, since it's really just an arbitrary way of labeling things. In fact, we might even expect that the equations describing the laws of physics should have the same form in any coordinate system. This is known as the **principle of relativity**.

You're already aware of this principle to some extent: Newton's laws, for example, work in any *inertial* reference frame moving at constant velocity with respect to one another. However, we know that it doesn't work in noninertial frames, and so Newtonian mechanics doesn't appear compatible with the principle of relativity in its full generality. Expecting this principle to be true is a very powerful idea, and ultimately leads to the formulation of general relativity, but at the moment, this is too many steps ahead of us. We'll come back to this when we discuss special relativity.

For now, back to scalars and vectors.

> A scalar is a number that always has the same value no matter what coordinate system you choose.

Examples in 3D Euclidean space include quantities such as charge, mass and energy.[16,17] A vector, on the other hand, is more complicated. As a mental image, you should think of a vector as an arrow pointing in some direction in space with some length, that exists *before* you even talk about coordinates. But depending on how you choose your coordinate system, clearly the **components** of the vector will change. A vector with only an x-component in one Cartesian coordinate system can clearly have nonzero y or z components in virtually any other choice of Cartesian coordinate systems, let alone in other non-Cartesian systems where the components aren't even x, y and z.

[16] In electromagnetism, scalars are always real, but they can be *complex* even in physics, especially in quantum mechanics.

[17] Later on, when we talk about relativity, you'll find that some things that were scalars in 3D are no longer scalars when you include relativity. If that's confusing, don't worry about it for now!

> What should be intuitively clear, however, is that under a change of coordinates, since the vector itself is unchanged (it's just an arrow with some length pointing in some fixed direction), the components of every vector in the system must change in the same way under a change of coordinates.

The only type of 3D coordinate transformation that will really be important to us in this course will be **rotations**, so let's just specialize to that case for the rest of this discussion. Rotations are transformations that *preserve both lengths of vectors and angles between vectors*. Since lengths and angles are computed using the dot product, this means that under a rotation, the dot product between any two vectors must remain unchanged. This is a fact that we will come back to later.

Under a rotation of the coordinate system, every vector—displacement, velocity, electric field, magnetic field[18] etc.—must transform in the same way. For example, let's consider a rotation about the z-axis by an angle θ. Then all vectors $\mathbf{v}$ with components (v^1, v^2, v^3) in the original coordinate system, then

[18] Without restricting to rotations, e.g. if we include *reflections*, this is actually not true for magnetic fields, which is why you might have heard of the statement that magnetic fields are *pseudovectors*, but we won't concern ourselves with this for this course.

in the rotated coordinate system, the components (v'^1, v'^2, v'^3) are given by

$$\begin{pmatrix} v'^1 \\ v'^2 \\ v'^3 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} , \tag{2.25}$$

Using index notation, we would write this as

$$v'^i = R^i{}_j v^j , \tag{2.26}$$

where $R^1{}_1 = \cos\theta$, $R^1{}_2 = \sin\theta$, $R^2{}_1 = -\sin\theta$, $R^2{}_2 = \cos\theta$, $R^3{}_3 = 1$, and all other components of $R^i{}_j$ are zero. Notice that I have been very careful with putting a space between the upper and lower indices of $R^i{}_j$, with the first index should denote the row, and the second index denoting the column; $R^1{}_2 \neq R^2{}_1$ here, for example, so it would be confusing if we don't put a space.[19] You can check for yourself that the expression above is equivalent to matrix multiplication. *In fact, it is very helpful to remember that the structure on the RHS is how you would express matrix multiplication in index notation, with the contracted indices being the second index of the matrix with the vector index.*

[19] The only exception to this rule is the Kronecker delta δ^i_j, which is symmetric in its indices, and so we often write it without a space.

Again, the position of the index is actually completely unimportant in 3D Euclidean space, and it is very common to see all of the indices being written as lower indices in the literature. However, just so that you get used to the notation in relativity, coordinate transformations (and indeed any linear transformation that maps vectors to vectors) should be written with one upper and one lower index, as I've done here.

End of Lecture: Wednesday, Jan 28 2026

One important example of scalars, which is obtained by taking the **scalar product** or dot product of two vectors,

$$\mathbf{A} \cdot \mathbf{B} = A^i B_i . \tag{2.27}$$

As we mentioned earlier, the dot product should be invariant under rotations. Let's see how this works out. Under a rotation, we have

$$A'^i B'_i = (R^i{}_j A^j)(R_i{}^k B_k) = R^i{}_j R_i{}^k A^j B_k , \tag{2.28}$$

where in the last step I have just rearranged the terms. Again, let me remind you that in index notation, all quantities are simply numbers, and so you can simply shuffle them around as you like. For the dot product to be invariant, we must have $A'^i B'_i = A^j B_j = \delta^k_j A^j B_k$ for any vectors $\mathbf{A}$ and $\mathbf{B}$, and so it must be true that all rotations satisfy

$$R^i{}_j R_i{}^k = \delta^k_j . \tag{2.29}$$

While this is indeed true, it is more illuminating if we write the first term $R^i{}_j$ in terms of the transpose R^T, with

$$R^i{}_j = (R^T)_j{}^i . \tag{2.30}$$

and so $(R^T)_j{}^i R_i{}^k = \delta^k_j$, or in matrix notation,

$$R^T R = \mathbb{I} , \tag{2.31}$$

where $\mathbb{I}$ is the identity matrix. This shows that $R^T = R^{-1}$, i.e. the transpose of a rotation matrix is equal to its inverse. These kinds of matrices are called **orthogonal matrices**.

In 3D Euclidean space, **tensors** are objects formed by combining multiple vectors together. In index notation, they're really not that mysterious: you've actually already seen many examples of tensors without realizing it! Simply take two vectors A^i and B^j, and put them together to form the object $T^{ij} = A^i B^j$: this is a tensor.[20] We say that T^{ij} has **rank** 2, since it has two indices, with $3\times 3 = 9$ components in 3D space. Taking a derivative of a vector field $\partial_i V^j$ also gives a rank-2 tensor. Linear transformations as represented by a matrix are also rank-2 tensors. You can continue to put vectors together, e.g. $T^{ijk} = A^i B^j C^k$ to get a rank-3 tensor, and so on.[21]

In general, the rank-2 tensor components $T^{ij} \neq T^{ji}$, i.e. the order of the indices matters, which is why if we want to write them with mixed upper and lower indices, we need to be careful about their order, i.e. we should be careful to write ${T^i}_j$ and not T^i_j.[22] Rank-2 tensors that *do* satisfy $T^{ij} = T^{ji}$ are called **symmetric tensors**. An example of such a tensor would be the Kronecker delta, δ^{ij}; because it is symmetric, it is almost invariably written as δ^i_j without spaces, which is a convention I will use, but not for any other tensor. On the other hand, rank-2 tensors that satisfy $T^{ij} = -T^{ji}$ are called **antisymmetric tensors**.

Like vectors, all tensors transform in the same way under rotations. In particular, if we look at the tensor formed by two vectors $T^{ij} = A^i B^j$, the tensor transforms as

$$T'^{ij} = A'^i B'^j = {R^i}_k A^k {R^j}_l B^l = {R^i}_k {R^j}_l T^{kl}\,. \tag{2.32}$$

Therefore, unlike a vector, a rank-2 tensor transforms under a rotation by *two* factors of ${R^i}_j$, one for each index. This generalizes straightforwardly to tensors of any rank.

Tensors have a reputation for being very hard to understand, but once you're comfortable with index notation, they're really not that bad!

[20] Formally, this is called taking the **tensor product** of two vectors, but in practice it's nothing more complicated than just putting them together like this.

[21] Once again, the upper and lower positions do not matter in 3D Euclidean space, but all the rules of index notation still apply, and we'll keep the positions of the indices as they are for consistency, so that you get used to it for when they become important in the context of relativity.

[22] A tensor formed from two vectors $A^i B^j$ clearly has $A^i B^j \neq A^j B^i$ in general, for example.

2.4 Conservation of Momentum

Based on our expression with Poynting's theorem, we now want to guess what the local conservation law for momentum should look like. In Poynting's theorem, we related the time derivative of the energy density (both in the fields and in the mechanical energy of the charges) to the divergence of the Poynting vector, which is the energy flux density. By analogy, we would expect that the time derivative of the *momentum density* (both in the fields $\wp_{\mathrm{EM}}$ and in the mechanical momentum $\wp_{\mathrm{mech}}$ of the charges) should be related to the divergence of some *momentum flux density* T,[23] i.e.

[23] Griffiths uses the notation $\overset{\leftrightarrow}{\;}$.

$$\frac{\partial}{\partial t}(\wp_{\mathrm{EM}} + \wp_{\mathrm{mech}}) - \nabla\cdot\mathsf{T} = 0\,. \tag{2.33}$$

Don't be confused by the negative sign in front of the divergence term: this is just a convention related to how we define the sign of T later on. We could have left it as a plus sign, but then T would have ended up with a relative minus sign compared to the standard convention.

In index notation, $\wp_{\mathrm{EM}}$ is a vector with components $\wp^i_{\mathrm{EM}}$; for consistency, the divergence of T must also be a vector, but ∂_j from the divergence must contract with an index of T, leaving one index free. The only way this is possible is if T is a rank-2 tensor, with components T^{ij}, so that the conservation of momentum is expressed as

$$\frac{\partial}{\partial t}(\wp^i_{\rm EM} + \wp^i_{\rm mech}) - \partial_j T^{ij} = 0\,. \tag{2.34}$$

While we can certainly obtain the full expression for each of these terms from Maxwell's equations and the Lorentz force law, we won't do it in lecture: I think it is a valuable exercise in index notation, and so I'll leave it to you to try it out in the problem set. Griffiths 8.2 shows the derivation using vector notation, which you can use as a guide. Instead, what we'll do here is that I'll simply state the final result for each of these terms, and discuss physics.

The most straightforward quantity is $\wp^i_{\rm mech}$, which is the mechanical momentum density of the charges. We know that the time derivative of momentum is force, and hence clearly

$$\frac{\partial \boldsymbol{\wp}_{\rm mech}}{\partial t} \equiv \mathbf{f} = \rho\mathbf{E} + \mathbf{J}\times\mathbf{B}\,, \tag{2.35}$$

where $\mathbf{f}$ is the force per unit volume acting on the charges due to the fields, given by the Lorentz force law.

The momentum density in the electromagnetic fields $\boldsymbol{\wp}_{\rm EM}$ is less obvious. You might guess, however, that a momentum density pointing in a particular direction should be tied to the energy flux density in that direction, i.e. the Poynting vector, $\mathbf{S}$. In fact, they (perhaps somewhat remarkably) turn out to be exactly the same, up to pesky ϵ_0 and μ_0 factors:

$$\boldsymbol{\wp}_{\rm EM} \equiv \mu_0\epsilon_0\mathbf{S} = \epsilon_0(\mathbf{E}\times\mathbf{B})\,. \tag{2.36}$$

Hence electromagnetic fields carry momentum as well, just like they carry energy. The momentum carried by fields must be included in order for momentum to be conserved overall.

Finally,

T^{ij} is known as the **Maxwell stress tensor**,

$$T^{ij} \equiv \epsilon_0\left(E^iE^j - \frac{1}{2}\delta^{ij}E^2\right) + \frac{1}{\mu_0}\left(B^iB^j - \frac{1}{2}\delta^{ij}B^2\right)\,. \tag{2.37}$$

This seems somewhat magical, and really the right way to understand where this comes from is to start from the Lagrangian formulation of electromagnetism, which is beyond the scope of this course. However, let me try to give you some physical intuition. First, it should be clear to you that this is indeed a rank-2 tensor: E^iE^j and δ^{ij} are both obviously rank-2 tensors, and similarly for the B-field terms. You can also check that this tensor is symmetric, i.e. $T^{ij} = T^{ji}$. Second, it's got a very similar structure to the energy density expression, $u_{\rm EM}$, which is also somewhat remarkable. This and the fact that $\boldsymbol{\wp}_{\rm EM} = \mathbf{S}/c^2$ are suggestive of some deeper connection between energy and momentum, which again you'll see more of when we discuss relativity.

What is the physical meaning of the Maxwell stress tensor? Think about local charge conservation, shown in Eq. (1.59),

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{J} = 0\,. \tag{2.38}$$

Here, $\mathbf{J} = \rho\mathbf{v}$ tells you something about the flow of charge. Similarly, in the momentum conservation equation, $-T^{ij}$ encodes the flow of the i-th component of the momentum density in the j-direction.

We can of course rewrite the differential version of the conservation of momentum into an integral form. Plugging in all of our expressions for the various terms, we find in index notation

$$f^i + \epsilon_0\mu_0 \frac{\partial S^i}{\partial t} - \partial_j T^{ij} = 0\,, \tag{2.39}$$

In index notation, the divergence theorem for a vector v^i reads

$$\int_V \mathrm{d}^3\boldsymbol{r}\,\partial_i v^i = \int_{\partial V} \mathrm{d}A^i\, v_i\,, \tag{2.40}$$

and indeed nothing surprising happens for a tensor:

$$\int_V \mathrm{d}^3\boldsymbol{r}\,\partial_j T^{ij} = \int_{\partial V} \mathrm{d}A_j\, T^{ij}\,. \tag{2.41}$$

Integrating the conservation equation over volume therefore gives

$$F^i + \epsilon_0\mu_0 \frac{\mathrm{d}}{\mathrm{d}t}\int_V \mathrm{d}^3\boldsymbol{r}\,S^i - \int_{\partial V} \mathrm{d}A_j\, T^{ij} = 0\,. \tag{2.42}$$

Let's take a few interesting limits of these equations. First, in the limit where fields are static, i.e. $\partial_t\mathbf{E} = \partial_t\mathbf{B} = 0$, the Poynting vector $\mathbf{S} = 0$, and so the conservation of momentum reduces to

$$F^i = \int_{\partial V} \mathrm{d}A_j\, T^{ij} \qquad \text{(static)}\,. \tag{2.43}$$

This equation says that if you have a region containing some charges, you can also interpret T^{ij} as the i-th component of the force per unit area on a small surface with normal in the j-direction. Concretely, T^{xx} is the x-component of the force per unit area on a small surface with normal in the x-direction, i.e. what we would identify as **pressure**. On the other hand, T^{xy} is the x-component of the force per unit area on a small surface with normal in the y-direction, i.e. what is called **shear stress** in fluid mechanics, i.e. forces on surface due to fluid flow parallel to the surface. Since a flow of momentum out from a region will lead to a change in momentum density in the region, i.e. a force, this makes sense. The total force on a volume is simply obtained by integrating the stress tensor over the surface, dotted into the orientation of the surface element.[24]

On the other hand, in vacuum with no charges or currents, and so $f^i = 0$, and we are left with

$$\frac{\partial}{\partial t}\left(\mu_0\epsilon_0 S^i\right) - \partial_j T^{ij} = 0\,. \tag{2.44}$$

This is the momentum conservation law for electromagnetic fields, with $-T^{ij}$ describing how momentum flows *out* of a certain volume.[25]

[24] Incidentally, this also explains the sign convention that we adopted in Eq. (2.33). T^{ij} is defined such that it tells you about the force exerted on the volume enclosed, whereas if we had chosen the plus sign, we would it to be the force exerted by the volume on the outside world, which has a relative minus sign.

[25] See previous sidenote about sign conventions.

Example 2.2

Consider an infinite parallel-plate capacitor, with the lower plate (at $z = -d/2$) carrying surface charge density $-\sigma$, and the upper plate (at $z = +d/2$) carrying surface charge density $+\sigma$.

Let's work out the Maxwell stress tensor in the region between the plates. We know that the electric field in such a system is

$$\mathbf{E} = -\frac{\sigma}{\epsilon_0}\hat{z}\,, \tag{2.45}$$

with no magnetic fields present. The Maxwell stress tensor is therefore

$$T^{ij} = \epsilon_0\left(E^iE^j - \frac{1}{2}\delta^{ij}E^2\right)\,. \tag{2.46}$$

Let's consider all elements where $i \neq j$. The Kronecker delta term vanishes in this case, but so does E^iE^j, since the only nonzero component of $\mathbf{E}$ is in the z-direction. Therefore, all off-diagonal elements of T^{ij} vanish. For the diagonal elements, we have

$$\begin{aligned} T^{11} &= \epsilon_0\left(E^1E^1 - \frac{1}{2}E^2\right) = -\frac{\sigma^2}{2\epsilon_0}\,,\\ T^{22} &= \epsilon_0\left(E^2E^2 - \frac{1}{2}E^2\right) = -\frac{\sigma^2}{2\epsilon_0}\,,\\ T^{33} &= \epsilon_0\left(E^3E^3 - \frac{1}{2}E^2\right) = +\frac{\sigma^2}{2\epsilon_0}\,. \end{aligned} \tag{2.47}$$

To compute the force, use a pillbox volume enclosing some area A with axis parallel to the z-axis. By symmetry, the force must act along the z-axis, and so we can simply compute

$$F^3 = \oint dA_j T^{3j} = \oint dA_3 T^{33}\,. \tag{2.48}$$

The top of the pillbox is outside the capacitor, and so there is no electric field on it. For the cylindrical surface of the pillbox, the normal vector is perpendicular to the z-axis, and so $dA_3 = 0$ there as well. Finally, on the bottom of the pillbox, the normal vector is pointing in the $-\hat{\mathbf{z}}$ direction, and we obtain

$$F^3 = -AT^{33} = -A\frac{\sigma^2}{2\epsilon_0}\,. \tag{2.49}$$

The force per unit area on the top plate is therefore

$$\frac{\mathbf{F}}{A} = -\frac{\sigma^2}{2\epsilon_0}\hat{\mathbf{z}}\,, \tag{2.50}$$

i.e. the force is pointing downward with magnitude $\sigma^2/(2\epsilon_0)$, which is exactly what you would expect from standard electrostatics.

End of Lecture: Friday, Jan 30 2026

2.5 Angular Momentum in Electromagnetism

Just as electromagnetic fields carry linear momentum, they also carry angular momentum. The angular momentum density of the electromagnetic field is given by

$$\boldsymbol{\ell}_{\mathrm{EM}} = \boldsymbol{r} \times \boldsymbol{\wp}_{\mathrm{EM}} = \epsilon_0\, \boldsymbol{r} \times (\mathbf{E} \times \mathbf{B})\,, \tag{2.51}$$

where $\boldsymbol{r}$ is the position vector from the origin. The total angular momentum stored in the electromagnetic field is obtained by integrating over all space:

$$\mathbf{L}_{\mathrm{EM}} = \epsilon_0 \int \mathrm{d}^3\boldsymbol{r}\, \boldsymbol{r} \times (\mathbf{E} \times \mathbf{B})\,. \tag{2.52}$$

The torque per unit volume exerted by the electromagnetic field on charges is

$$\boldsymbol{\tau} = \boldsymbol{r} \times \mathbf{f} = \boldsymbol{r} \times (\rho\mathbf{E} + \mathbf{J} \times \mathbf{B})\,. \tag{2.53}$$

Integrating this over a volume gives the total electromagnetic torque on the charges within that volume.[26]

[26] You may also have guessed that there is an angular momentum flux density tensor, similar to the Maxwell stress tensor for linear momentum, given by $\epsilon_{ijk} r^j T^{kl}$, but we won't discuss this here.

Example 2.3

We'll work through a variant of a famous example known as the *Feynman disk paradox*. Consider a very long solenoid with radius R with n turns per unit length and current I. Coaxial with the solenoid are two long cylindrical, nonconducting shells of length l—one, inside the solenoid at radius a, carrying a charge $+Q$, uniformly distributed over its surface; the other, outside the solenoid at radius b, carrying a charge $-Q$, also uniformly distributed over its surface. Show that if the current is reduced, a torque is applied on both cylinders. Compute the total angular momentum in the cylinders if the current goes to zero, and show that the total angular momentum is conserved.

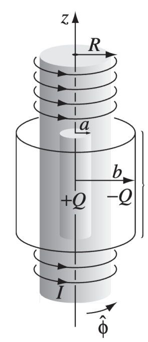

Figure 8: A solenoid with a current I, with inner cylinder of radius a, charge $+Q$ uniformly distributed, and outer cylinder of radius b, charge $-Q$, also uniformly distributed.

SOLUTION:

First, let's recall the fields produced in this set-up. The solenoid produces a magnetic field $\mathbf{B} = \mu_0 n I \hat{\mathbf{z}}$ inside, but no magnetic field outside. Meanwhile, the charge distributions produce an electric field

$$\mathbf{E}_{\mathrm{cyl}} = \frac{Q/l}{2\pi\epsilon_0 r}\hat{\boldsymbol{r}} \tag{2.54}$$

in the region between the two cylinders, i.e. for $a < r < b$. Here, Q/l is the charge per unit length, and the field points radially inward since the positively charged cylinder is inside.

If the magnetic field changes with time, then by Faraday's law, there is an induced electric field in the azimuthal direction, and that's going to turn the cylinders. Specifically, if we choose a loop around each cylinder, Faraday's law gives

$$\oint \mathrm{d}\mathbf{l} \cdot \mathbf{E} = -\frac{\mathrm{d}\Phi_B}{\mathrm{d}t}\,, \tag{2.55}$$

where Φ_B is the magnetic flux through the loop. For the inside cylinder at radius a, we have an electric field of magnitude E_{in} in the ϕ-direction,

$$2\pi a E_{\mathrm{in}} = -\pi a^2 \frac{\mathrm{d}B}{\mathrm{d}t} \implies E_{\mathrm{in}} = -\frac{a}{2}\frac{\mathrm{d}B}{\mathrm{d}t} = -\frac{a}{2}\mu_0 n \frac{\mathrm{d}I}{\mathrm{d}t}\,. \tag{2.56}$$

On the outside cylinder at radius b, the electric field has magnitude E_{out} in the ϕ-direction again, with

$$2\pi b E_{\mathrm{out}} = -\pi R^2 \frac{\mathrm{d}B}{\mathrm{d}t} \implies E_{\mathrm{out}} = -\frac{R^2}{2b}\frac{\mathrm{d}B}{\mathrm{d}t} = -\frac{R^2}{2b}\mu_0 n \frac{\mathrm{d}I}{\mathrm{d}t}\,. \tag{2.57}$$

The torque on the inner cylinder is therefore

$$\boldsymbol{\tau}_{\rm in} = \boldsymbol{r} \times Q\mathbf{E}_{\rm in} = -\frac{a^2}{2}Q\mu_0 n\frac{{\rm d}I}{{\rm d}t}\hat{\mathbf{z}}\,, \tag{2.58}$$

and on the outer cylinder,

$$\boldsymbol{\tau}_{\rm out} = \boldsymbol{r} \times (-Q)\mathbf{E}_{\rm out} = +\frac{R^2}{2}Q\mu_0 n\frac{dI}{dt}\hat{\mathbf{z}}\,. \tag{2.59}$$

If the current goes to zero, the total angular momentum of the inner cylinder is

$$\mathbf{L}_{\rm in} = \int {\rm d}t\,\boldsymbol{\tau}_{\rm in} = \frac{-a^2}{2}Q\mu_0 n\Delta I\hat{\mathbf{z}} = \frac{a^2}{2}Q\mu_0 nI\hat{\mathbf{z}}\,, \tag{2.60}$$

where $\Delta I = -I$ is the change in the current (final minus initial), and

$$\mathbf{L}_{\rm out} = \int {\rm d}t\,\boldsymbol{\tau}_{\rm out} = \frac{R^2}{2}Q\mu_0 n\Delta I\hat{\mathbf{z}} = -\frac{R^2}{2}Q\mu_0 nI\hat{\mathbf{z}}\,. \tag{2.61}$$

The total angular momentum in the cylinders is therefore

$$\mathbf{L} = \mathbf{L}_{\rm in} + \mathbf{L}_{\rm out} = \frac{1}{2}(a^2 - R^2)Q\mu_0 nI\hat{\mathbf{z}}\,. \tag{2.62}$$

Where this angular momentum come from? Well, it must have come from the electromagnetic fields! Let's compute the angular momentum in fields initially, before the current decreases. The magnetic field and electric field both only exist between the inner cylinder and the solenoid, i.e. for $a < r < R$. The angular momentum density in this region is

$$\boldsymbol{\ell} = \epsilon_0\boldsymbol{r} \times (\mathbf{E}_{\rm cyl} \times \mathbf{B}) = -\epsilon_0\boldsymbol{r} \times \frac{Q}{2\pi\epsilon_0 lr}\mu_0 nI\hat{\boldsymbol{\phi}} = -\frac{Q}{2\pi l}\mu_0 nI\hat{\mathbf{z}}\,. \tag{2.63}$$

Integrating this over the volume between $r = a$ and $r = R$, we find

$$\mathbf{L}_{\rm EM} = \pi(R^2 - a^2)l \cdot -\frac{Q}{2\pi l}\mu_0 nI\hat{\mathbf{z}} = \frac{1}{2}(a^2 - R^2)Q\mu_0 nI\hat{\mathbf{z}}\,, \tag{2.64}$$

which is exactly the angular momentum transferred to the cylinders!

3 Electromagnetic Waves

In the previous chapter, we learned that electric and magnetic fields carry energy and momentum. In this chapter, we'll take a much more in-depth look at this idea, and see how energy and momentum can be transported across space in the form of **electromagnetic waves**.

3.1 The Wave Equation

What is a wave? Intuitively, it is some disturbance in the value of some quantity in space away from some equilibrium value propagating through space. Sound waves, for example, are propagating disturbances in the pressure and density of air; waves on a string are propagating disturbances in the displacement of the string from its equilibrium position.

Let's think in one dimension for a moment, e.g. consider a string that can be displaced upward or downward along the x-axis. Give the string a shake at the origin at $t = 0$. We can think of a wave as some function $f(x, t)$. When we shake the string at the origin at $t = 0$, this creates a disturbance that propagates down the string, and so at a later time t, there is some displacement of the string at position x far away from the origin. Many functional forms for $f(x, t)$ are possible, but the perhaps the simplest functional form that looks like what we think of as a wave is

$$f(x,t) = f(x - vt)\,, \tag{3.1}$$

where f only depends on the combination $x - vt$, with v being some constant with units of velocity. For a wave that is given by $f(x - vt)$, the shape of the wave at $t = 0$ is given by $f(x)$, and at a later time t, the shape of the wave is given by $f(x - vt)$, i.e. the same shape as before, but shifted to the right by a distance vt (see Fig. 9). Evidently v is the **speed of propagation** of the wave.[27]

[27] You can set $v < 0$, in which case the wave propagates to the left instead.

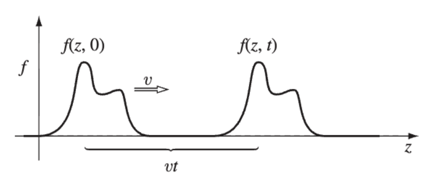

Figure 9: An example of a wave of the form $f(x - vt)$.

Functions of the form $f(x - vt)$ are solutions of a partial differential equation known as the **wave equation**:

$$\frac{\partial^2 f}{\partial x^2} - \frac{1}{v^2}\frac{\partial^2 f}{\partial t^2} = 0 \tag{3.2}$$

To see this, let's compute the partial derivatives explicitly. Define $u \equiv x - vt$, so that $f(x, t) = f(u)$. Then, by the chain rule,

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\mathrm{d}f}{\mathrm{d}u}\frac{\partial u}{\partial x}\right) = \frac{\mathrm{d}}{\mathrm{d}u}\left(\frac{\mathrm{d}f}{\mathrm{d}u}\right)\frac{\partial u}{\partial x} = \frac{\mathrm{d}^2 f}{\mathrm{d}u^2}\,, \tag{3.3}$$

since $\partial y/\partial x = 1$. On the other hand,

$$\frac{1}{v^2}\frac{\partial^2 f}{\partial t^2} = \frac{1}{v^2}\frac{\partial}{\partial t}\left(\frac{\mathrm{d}f}{\mathrm{d}u}\frac{\partial u}{\partial t}\right) = \frac{1}{v^2}\frac{\mathrm{d}}{\mathrm{d}u}\left(-v\frac{\mathrm{d}f}{\mathrm{d}u}\right)\frac{\partial u}{\partial t} = \frac{1}{v^2}\frac{\mathrm{d}^2 f}{\mathrm{d}u^2}(-v)^2 = \frac{\mathrm{d}^2 f}{\mathrm{d}u^2}\,, \tag{3.4}$$

and so clearly the wave equation is satisfied! It should be clear to you as well that for any two solutions f and g to the wave equation, $f + g$ is also a solution, since the wave equation is linear in f (i.e. it doesn't depend on f^2, or $(\partial f/\partial x)^2$, etc.).[28]

[28] In fact, the most general solution to the 1D wave equation is $f(x-vt)+g(x+vt)$, i.e. a superposition of a wave traveling to the right and a wave traveling to the left.

End of Lecture: Monday, Feb 2 2026

The wave equation can be generalized to three dimensions. Instead of waves propagating along the x-axis, we now consider waves that can propagate in any direction in 3D space.

The **3D wave equation** is given by:

$$\nabla^2 f - \frac{1}{v^2}\frac{\partial^2 f}{\partial t^2} = 0\,, \tag{3.5}$$

where $\nabla^2 \equiv \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2 = \partial^i\partial_i$ is the Laplacian operator. The general solution to the 3D wave equation is:

$$f(\boldsymbol{r}, t) = f(\mathbf{k}\cdot\boldsymbol{r} - \omega t)\,, \tag{3.6}$$

where f is an arbitrary function of only $\mathbf{k}\cdot\boldsymbol{r} - \omega t$, $\mathbf{k}$ is the **wavevector**, with $k = |\mathbf{k}|$ the **wavenumber**, and ω is the **angular frequency**. $\mathbf{k}$ and ω are constants in space and time, related by the **dispersion relation** $\omega/k = v$.

Let's think a bit about what kind of solution $f(\mathbf{k}\cdot\boldsymbol{r} - \omega t)$ represents. At $t = 0$, all points $\boldsymbol{r}$ satisfying $\mathbf{k}\cdot\boldsymbol{r} =$ constant have the same value of f. You can convince yourself that $\mathbf{k}\cdot\boldsymbol{r} =$ constant describes a plane perpendicular to $\mathbf{k}$, and therefore $f(\mathbf{k}\cdot\boldsymbol{r} - \omega t)$ describes a wave whose wavefronts (surfaces of constant values of f) are planes perpendicular to $\mathbf{k}$ (see Fig. 10). As t increases, you can see that these wavefronts move in the direction of $\mathbf{k}$, at a speed along k given by $v = \omega/k$. This is called a **plane wave**.

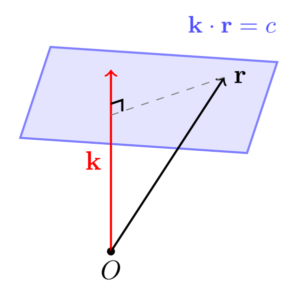

Figure 10: A plane wave has surfaces of constant phase (wavefronts) that are planes perpendicular to the wavevector $\mathbf{k}$. All position vectors $\mathbf{r}$ pointing to the plane satisfy $\mathbf{k}\cdot\mathbf{r} = c$ for some constant c.

Let's verify that this is indeed a solution to the 3D wave equation. Define $u = \mathbf{k}\cdot\boldsymbol{r} - \omega t$, so that $f = f(u)$. In index notation, we can write:

$$u = k_i r^i - \omega t\,, \tag{3.7}$$

where repeated indices (one upper, one lower) are summed over. Taking partial derivatives with respect to the spatial coordinates using the chain rule:

$$\partial_j f = \frac{\mathrm{d}f}{\mathrm{d}u}\partial_j u = f' k_j\,, \tag{3.8}$$

since $\partial_j(k_i r^i) = k_i \delta^i{}_j = k_j$. Taking a second derivative:

$$\partial^j\partial_j f = \partial^j(f' k_j) = f'' k^j k_j = f'' k^2\,, \tag{3.9}$$

where $k^2 = k^j k_j = |\mathbf{k}|^2$. Thus the Laplacian is $\nabla^2 f = f'' k^2$.

For the time derivative:

$$\frac{\partial f}{\partial t} = f'\frac{\partial u}{\partial t} = -\omega f'\,, \qquad \frac{\partial^2 f}{\partial t^2} = \omega^2 f''\,. \tag{3.10}$$

Substituting into the 3D wave equation and using the dispersion relation $\omega = vk$:

$$\nabla^2 f - \frac{1}{v^2}\frac{\partial^2 f}{\partial t^2} = f'' k^2 - \frac{\omega^2}{v^2} f'' = f'' k^2 - \frac{v^2 k^2}{v^2} f'' = 0\,. \tag{3.11}$$

Thus, $f(\boldsymbol{r}, t) = f(\mathbf{k}\cdot\boldsymbol{r} - \omega t)$ satisfies the 3D wave equation for any arbitrary function f and any wavevector $\mathbf{k}$, provided that ω and $\mathbf{k}$ satisfy the dispersion relation $\omega/k = v$.

3.2 Electromagnetic Waves in Vacuum

With all these preliminaries about the wave equation of the way, let's return to electromagnetism.

3.2.1 The Wave Equation from Maxwell's Equations

It should come as no surprise that what we want to show is that electromagnetic fields $\mathbf{E}$ and $\mathbf{B}$ satisfy the wave equation as well. In vacuum (no charges or currents), Maxwell's equations are:

$$\partial_i E^i = 0\,, \qquad \partial_i B^i = 0\,, \qquad \epsilon^{ijk}\partial_j E_k = -\frac{\partial B^i}{\partial t}\,, \qquad \epsilon^{ijk}\partial_j B_k = \mu_0\epsilon_0 \frac{\partial E^i}{\partial t}\,. \tag{3.12}$$

To derive the wave equation for $\mathbf{E}$, we take the curl of Faraday's law. We can of course do this in vector notation (see Griffiths for this), but we'll do this with index notation here. On the LHS of Faraday's law after taking the curl, the i-th component of $\nabla \times (\nabla \times \mathbf{E})$ is:

$$\epsilon^{ijk}\partial_j(\nabla \times \mathbf{E})_k = \epsilon^{ijk}\partial_j(\epsilon_{klm}\partial^l E^m) = \epsilon^{ijk}\epsilon_{klm}\partial_j\partial^l E^m\,. \tag{3.13}$$

But we have the identity that we derived in Eq. (1.45):

$$\epsilon^{ijk}\epsilon_{klm} = \epsilon^{kij}\epsilon_{klm} = \delta^i_l\delta^j_m - \delta^i_m\delta^j_l\,. \tag{3.14}$$

Substituting this identity, we obtain

$$\begin{aligned}\epsilon^{ijk}\epsilon_{klm}\partial_j\partial^l E^m &= (\delta^i_l\delta^j_m - \delta^i_m\delta^j_l)\partial_j\partial^l E^m \\ &= \partial^i\partial_j E^j - \partial^j\partial_j E^i \\ &= -\partial^j\partial_j E^i\,,\end{aligned} \tag{3.15}$$

where in the last line, we use Gauss's law in vacuum. On the RHS, we have instead

$$\begin{aligned}-\epsilon^{ijk}\partial_j(\partial_t B_k) &= -\partial_t(\epsilon^{ijk}\partial_j B_k) \\ &= -\partial_t(\mu_0\epsilon_0\partial_t E^i)\,,\end{aligned} \tag{3.16}$$

where in the last line we used the Ampere-Maxwell law in vacuum. Comparing both sides now, we see that

$$-\partial^j\partial_j E^i = -\mu_0\epsilon_0\frac{\partial^2}{\partial t^2}E^i\,, \tag{3.17}$$

or in vector notation,

$$\nabla^2\mathbf{E} - \mu_0\epsilon_0\frac{\partial^2\mathbf{E}}{\partial t^2} = 0\,, \tag{3.18}$$

i.e. the electric field satisfies the wave equation in vacuum.

What about the magnetic field? We can do the same procedure, but starting from the Ampere-Maxwell law instead. Taking the curl of the Ampere-Maxwell law, the LHS is

$$\begin{aligned}\epsilon^{ijk}\partial_j(\nabla \times \mathbf{B})_k &= \epsilon^{ijk}\partial_j\left(\epsilon_{klm}\partial^l B^m\right) \\ &= (\delta^i_l\delta^j_m - \delta^i_m\delta^j_l)\partial_j\partial^l B^m \\ &= \partial^i\partial_j B^j - \partial^j\partial_j B^i \\ &= -\partial^j\partial_j B^i\,,\end{aligned} \tag{3.19}$$

where in the second last line we used Gauss's law for magnetism. On the RHS,

$$\begin{aligned}\epsilon^{ijk}\partial_j(\mu_0\epsilon_0\partial_t E_k) &= \mu_0\epsilon_0\partial_t(\epsilon^{ijk}\partial_j E_k) \\ &= \mu_0\epsilon_0\partial_t(-\partial_t B^i)\,,\end{aligned} \tag{3.20}$$

and so once again in vector notation,

$$\nabla^2 \mathbf{B} - \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{B}}{\partial t^2} = 0\,, \tag{3.21}$$

i.e. the magnetic field also satisfies the wave equation in vacuum.

Let's take stock of what just happened:

We have shown that any electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$ in vacuum—which must be solutions to Maxwell's equations—must also satisfy the wave equation.

Note the direction of implication here: Maxwell's equations $\Longrightarrow$ wave equation, but not necessarily the other way around. This means that any solution to Maxwell's equations in vacuum is also a solution to the wave equation, but not all solutions to the wave equation are solutions to Maxwell's equations.

From the wave equation for $\mathbf{E}$ and $\mathbf{B}$ in vacuum, we can read off the speed of propagation of electromagnetic waves:

$$v = \frac{1}{\sqrt{\mu_0 \epsilon_0}} = 2.997\,924\,58 \times 10^8\,\mathrm{m\,s^{-1}}\,, \tag{3.22}$$

which turns out to be precisely the **speed of light** in vacuum, c.

This may not seem surprising to you anymore, but it is a tremendous discovery, stated most clearly in this way by James Clerk Maxwell in 1862. Of course, you as an experienced physics student knows that light is electromagnetic in nature, but this was certainly not clear for a long time. In the 19th century, μ_0 and ϵ_0 were simply constants that relate electric and magnetic fields to their sources, i.e. charges and currents, which had no obvious connection to light. The fact that $1/\sqrt{\mu_0 \epsilon_0}$ corresponded to a wave propagation speed, and the fact that this wave propagation speed was equal to the speed of light strongly suggested, simply from theoretical arguments, that light was an electromagnetic wave!

3.2.2 Galilean Invariance

Let's take a closer look at Maxwell's equations in vacuum, and the wave equation satisfied by $\mathbf{E}$ and $\mathbf{B}$. In physics, we build our equations with scalars, vectors and tensors, because they all transform nicely under a rotation of our coordinate system. So for example, Gauss's law states that $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$, which is true in Cartesian coordinate system, regardless of how we orient our axes, because both $\nabla \cdot \mathbf{E}$ and ρ are scalars that every observer agrees on. Similarly, Faraday's law $\nabla \times \mathbf{E} = -\partial \mathbf{B}/\partial t$ is true in any Cartesian coordinate system, because both sides are vectors: in a new coordinate system that is rotated with respect to the old one, both sides transform in the same way under the rotation, and so the equation remains valid. The same is true in Newtonian mechanics: $\mathbf{F} = m\mathbf{a}$ is true in any Cartesian coordinate system, because both sides are vectors.

In addition to this though, in Newtonian mechanics, one of things that you learned was how to change between frames of reference traveling at constant velocity relative to each other, also known as **inertial frames**. For example, you have probably done problems where you go from the lab frame to the center-of-mass frame of a system of particles. To go between different frames in Newtonian mechanics, we perform what we call a **Galilean transformation**, which simply consists of subtracting out the relative velocity between the two frames. Velocities in one frame $\mathbf{v}'$ and in another frame $\mathbf{v}$ are related by

$$\mathbf{v}' = \mathbf{v} + \mathbf{v}_0 \,, \tag{3.23}$$

where $\mathbf{v}_0$ is the constant velocity taking you from one frame relative to the other. And that's pretty much it! You don't have to do anything else.

Now, when you perform a Galilean transformation to go from one frame to another, one of the things you also know is that you can use Newton's laws in any inertial frame of your choice. So just like Newton's laws are **invariant** under rotations, they are also invariant under Galilean transformations. The reason this is true is that although inertial observers don't agree on the components of velocity vectors, they all do agree on $\mathbf{a}$, since in aanother frame, the acceleration $\mathbf{a}'$ is given by

$$\mathbf{a}' = \frac{\mathrm{d}\mathbf{v}'}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{v} - \mathbf{v}_0) = \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \mathbf{a} \,. \tag{3.24}$$

That's a neat property, and a powerful one, enabling you to change inertial frames at will and continue to use Newton's laws in whatever frame of your choice.[29] [30]

End of Lecture: Wednesday, Feb 4 2026

Now, are Maxwell's equations and the wave equation invariant under Galilean transformations? You can quickly see that the answer is *absolutely not*. The quickest way to see this is to look at the wave equation:

$$\nabla^2 \mathbf{E} - \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0 \,. \tag{3.25}$$

This says that electric fields (and also magnetic fields) propagate at a speed $v = 1/\sqrt{\mu_0 \epsilon_0} = c$, *which is a constant*. But with respect to what? Under a Galilean transformation, if I am an observer moving at a very large velocity v along with the electromagnetic wave, I would measure a very different wave speed than another observer moving in the opposite direction at speed v. But the wave equation only gives you one unique, constant speed c.

So you can't perform a Galilean transformation into another frame, and continue using Maxwell's equations, like you did in Newtonian mechanics. In fact, you should have already realized that Maxwell's equations themselves cannot be invariant under Galilean transformations. In a frame where a charge density distribution is stationary, there is only an electric field $\mathbf{E}$ produced by the charges, with $\mathbf{B} = 0$. But go into another frame where the charge density distribution is moving: the moving charges now look like a current density distribution, which surely has to source a magnetic field, $\mathbf{B}' \neq 0$. Somehow, complicated magnetic field configurations can appear and disappear as we go from one inertial frame to another: what could possibly be the transformation rule here? Certainly, it is not just a matter of transforming velocities. Indeed,

> Maxwell's equations are *not* Galilean invariant, even though Newtonian mechanics is. If there were a frame where Maxwell's equations hold, then under a Galilean transformation into another frame moving at constant velocity relative to the first, Maxwell's equations would not hold.

But Maxwell's equations *do* work very well on Earth, so clearly the Earth is, or is at least very close to, a frame in which Maxwell's equations hold. For a

[29] Although, if you thought about this some more, this transformation seems somewhat unsatisfactory. You probably already know that scalars are *not* invariant under Galilean transformations, e.g. kinetic energy certainly changes from one inertial frame to another! Also, different types of vectors transform differently as well (see the example of acceleration vs. velocities below). The Galilean transformation is pretty badly behaved! And if you had thought about it this way, you could already see the seeds of its destruction by Einstein.

[30] But you also know that you can't change into *any* frame, just inertial ones: in accelerating frames, for example, Newton's laws do *not* hold, and you can only make progress by introducing fictitious forces like centrifugal forces and Coriolis forces, etc. Trying to understand how we could write down physical laws that are true in *any* frame is the start of the road to general relativity!

very long time, physicists thought that there indeed was a special frame in which Maxwell's equations held, that they called the **aether**. This was a hypothetical medium that permeated all of space, in which electromagnetic fields existed, satisfying Maxwell's equations in the rest frame of the medium, but not in any other frames. If you're not willing to give up on Galilean invariance, this is really your best shot.

This was a huge dead end in physics for many years, and the history of it is absolutely fascinating, and we'll discuss some of it when we reach special relativity. But when we finally emerged from this confusion, the punchline of what we learned is as follows:

> There is no aether, and Maxwell's equations (and hence also the wave equation) hold in all inertial frames of reference.

This is because the Galilean transformation turns out not to be the correct way to relate quantities in different inertial frames.[31] A different transformation, known as the **Lorentz transformation**, is required instead, one that breaks Newtonian mechanics for objects traveling close to the speed of light. We'll come back to this later on.

[31] Indeed, who told us that Galilean transformations were the right thing to do anyway? It is certainly *intuitive*, but there had never been any guarantee that it was right, and modern physics in the last 120 years has been an exercise in tearing down our basic physical intuition. Always check your assumptions!

3.2.3 Sinusoidal Waves

The most basic one-dimensional wave that is a solution to the 1D wave equation is the **sinusoidal wave**, i.e. a wave whose functional form is given by

$$f(z,t) = A\cos[k(z - vt) + \delta] = A\cos[kz - \omega t + \delta]\,, \tag{3.26}$$

where A is the **amplitude** of the wave, k is the **wavenumber**, v is the **speed of propagation**, and δ is a **phase**. It should be clear to you that this is indeed a solution to the wave equation, since it is of the form $f(z - vt)$. We also define the **wavelength** λ as $2\pi/k$, the **angular frequency** ω of the wave as $\omega = kv$, and the **frequency** ν as $\nu = \omega/(2\pi) = v/\lambda$. Fig. 11 shows these quantities and how they relate to the sine wave; in particular, the phase controls the position of the maximum of the wave.

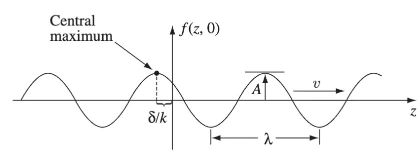

Figure 11: A sine wave and its properties.

The sinusoidal wave defined above with $k, \omega > 0$ is a wave traveling to the right. You can see this by looking at a peak located at $z = z_0$ at $t = 0$, such that $f(z_0, 0) = A$; as t increases, we see that $f(z_0 + vt, t) = A$, since this keeps the value of $z - vt$ constant, and so after some time t has elapsed, the peak is now at $z_0 + vt$. To get a wave traveling to the left would have functional form

$$f(z,t) = A\cos[kz + \omega t + \delta]\,, \tag{3.27}$$

or equivalently $f(z,t) = A\cos[-kz - \omega t + \delta]$, since

$$A\cos[-kz - \omega t + \delta] = -A\cos[kz + \omega t - \delta]\,, \tag{3.28}$$

as cosine is an even function.

Sinusoidal waves are much more commonly written in terms of complex exponentials, using Euler's formula,

$$e^{i\theta} = \cos\theta + i\sin\theta\,, \tag{3.29}$$

so that we can also write

$$f(z,t) = A\cos[kz - \omega t + \delta] = \mathrm{Re}\left[Ae^{i(kz - \omega t + \delta)}\right]\,. \tag{3.30}$$

Because exponentials are generally much easier to manipulate, this is often preferable (but equivalent!) to working with sines and cosines directly. Furthermore, if we also allow amplitudes to be *complex*, and define

$$B = Ae^{i\delta}\,, \tag{3.31}$$

every sinusoidal wave can be written as

$$f(z,t) = \mathrm{Re}\left[Be^{i(kz-\omega t)}\right]\,, \tag{3.32}$$

with the phase absorbed into the complex amplitude. In practice, we drop the $\mathrm{Re}[\cdots]$ and simply write the complex exponential $Be^{i(kz-\omega t)}$ directly, implicitly assuming that the physical wave is the real part of this expression. So you can pretend every sinusoidal wave is of the form $B\exp[i(kz-\omega t)]$, perform all your calculations with that, and then at the end of the day, simply take the real part if you want to know the physical value associated with the wave.

Example 3.1

If you don't believe me that complex exponentials are easier, consider taking the sum of two sinusoidal waves with the same wavenumber k and frequency ω, but different amplitudes A_1, A_2 and phases δ_1, δ_2. With sines and cosines, the sum would be:

$$f(z,t) = A_1\cos(kz-\omega t+\delta_1) + A_2\cos(kz-\omega t+\delta_2)\,. \tag{3.33}$$

and it's just not very obvious how to proceed (you could of course use trigonometric identities, but it's going to get messy). On the other hand, with complex exponentials, we have:

$$\begin{aligned} f(z,t) &= A_1e^{i\delta_1}e^{i(kz-\omega t)} + A_2e^{i\delta_2}e^{i(kz-\omega t)} \\ &= \left(A_1e^{i\delta_1} + A_2e^{i\delta_2}\right)e^{i(kz-\omega t)}\,, \end{aligned} \tag{3.34}$$

and so the final wave has complex amplitude $A_3 = A_1e^{i\delta_1} + A_2e^{i\delta_2}$. One thing that is clear immediately from this is that the sum of two sinusoidal waves with the same k and ω is also a sinusoidal wave with the same k and ω, but with a different amplitude and phase; this wasn't so obvious with sines and cosines.

In practice, complex exponentials work beautifully when you're adding them up, or taking derivatives/integrals. But a word of warning: *complex exponentials don't work out of the box when you're multiplying two waves together*, which can occur when you're computing things like the Poynting vector, $\mathbf{S} = (\mathbf{E}\times\mathbf{B})/\mu_0$. This is because why you multiply two complex exponentials together, the imaginary part of one multiplied by the imaginary part of the other gives a real contribution, even though only the real part of the wave was physical. We'll come back to this again.

Why are sinusoidal waves so important? You can show that under very general conditions, any function that solves the wave equation $f(z-vt)$ can be written as a superposition (i.e. an integral or sum) of (typically uncountably many) sinusoidal waves with different amplitudes, wavenumbers (or equivalently frequencies) and phases. Since the wave equation is linear, you can simply study the behavior of a sinusoidal waves of a certain wavenumber k. This fundamental idea is studied in **Fourier analysis**, which is important in all branches of physics, but we'll be able to get through this class without it.

3.2.4 Monochromatic Electromagnetic Plane Waves in Vacuum

We've been thinking about 1D solutions in the previous section, but now we're going to start thinking about waves of *3D vector fields* in *3D space* that are solutions to the wave equation that we derived earlier,

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\mathbf{E} = 0\,. \tag{3.35}$$

Again, we're only going to focus on the simplest case, which is a **monochromatic plane wave**, i.e. a wave whose functional form is given by

$$\mathbf{E}(\boldsymbol{r},t) = \mathbf{E}_0\cos(\mathbf{k}\cdot\boldsymbol{r} - \omega t + \delta)\,. \tag{3.36}$$

for some real amplitude vector $\mathbf{E}_0$. Monochromatic just means a wave having a single frequency/wavenumber. You can check for yourself that this satisfies the wave equation as before. Now, it should come as no surprise to you that we prefer to work with complex exponentials, and so you'll never see a sine or a cosine, but instead only

$$\mathbf{E}(\boldsymbol{r},t) = \mathbf{E}_0 e^{i(\mathbf{k}\cdot\boldsymbol{r} - \omega t)}\,, \tag{3.37}$$

where it is important to remember that now $\mathbf{E}_0$ is a *complex, constant vector*, encoding both the amplitude and phase of the wave. Now, even though I've only written down an electric field monochromatic plane wave, Maxwell's equations already tell us what the corresponding magnetic field must be, and this is generally the case: Maxwell's equations do not allow electric fields to propagate willy-nilly without also producing a corresponding magnetic field!

Let's take a look at what Maxwell's equations tell us about the $\mathbf{B}$-field if there is a monochromatic plane wave $\mathbf{E}$-field as above, and let's stick to the simplest case of propagation in vacuum, so no charges or currents. Applying Gauss's law, we find (using index notation and proceeding slowly)

$$\begin{aligned}
\nabla\cdot\mathbf{E} &= E_0^i\partial_i e^{i(k^j r_j - \omega t)} \\
&= E_0^i e^{i(k^j r_j - \omega t)}\partial_i[i(k^j r_j - \omega t)] \\
&= iE_0^i e^{i(k^j r_j - \omega t)}k^j\delta_{ij} \\
&= iE_0^i k_i e^{i(k^j r_j - \omega t)} \\
&= i\mathbf{k}\cdot\mathbf{E} \\
&= 0\,,
\end{aligned} \tag{3.38}$$

Let's define a the unit vector $\hat{\mathbf{n}}$ that points in the direction of the electric field, i.e.

$$\mathbf{E}_0 = E_0\hat{\mathbf{n}}e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)}\,, \tag{3.39}$$

We call $\hat{\mathbf{n}}$ the **polarization** direction of the wave.

where in the last line we used Gauss's law $\nabla\cdot\mathbf{E} = 0$ in the absence of charges.[32] This tells us the first important point:

> For a monochromatic plane wave in vacuum, we have
>
> $$\mathbf{k}\cdot\mathbf{E} = 0\,, \tag{3.40}$$
>
> i.e. the electric field is always *orthogonal* to the direction of propagation of the wave.

[32] Notice how easy it was to take derivatives of the exponential, instead of having to deal with cosines and sines.

End of Lecture: Friday, Feb 6 2026

Next, let's look at Faraday's law:

$$\begin{aligned}(\nabla\times\mathbf{E})_i &= \epsilon_{ijk}E_0^k\partial^j e^{i(k^l r_l-\omega t)}\\ &= i\epsilon_{ijk}E_0^k k^j e^{i(k^l r_l-\omega t)}\\ &= i(\mathbf{k}\times\mathbf{E})_i\\ &= -\frac{\partial B_i}{\partial t}\,, \end{aligned} \tag{3.41}$$

where I've just used the result from above to skip a few steps, and then applied Faraday's law in the last line.[33] This shows that the solution has to be[34]

$$\mathbf{B}=\mathbf{B}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \tag{3.42}$$

with

$$-\frac{\partial\mathbf{B}}{\partial t}=i\omega\mathbf{B}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}=i\mathbf{k}\times\mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}\,, \tag{3.43}$$

which means that $\mathbf{B}_0=E_0(\hat{\mathbf{k}}\times\hat{\mathbf{n}})/c$.

To summarize, a monochromatic electromagnetic plane wave in vacuum has the form

$$\begin{aligned}\mathbf{E}(\boldsymbol{r},t)&=\mathbf{E}_0 e^{i(\mathbf{k}\cdot\boldsymbol{r}-\omega t)}=E_0\hat{\mathbf{n}}e^{i(\mathbf{k}\cdot\boldsymbol{r}-\omega t)}\,,\\ \mathbf{B}(\boldsymbol{r},t)&=\frac{1}{c}\hat{\mathbf{k}}\times\mathbf{E}=\frac{1}{c}(\hat{\mathbf{k}}\times\mathbf{E}_0)e^{i(\mathbf{k}\cdot\boldsymbol{r}-\omega t)}=\frac{E_0}{c}(\hat{\mathbf{k}}\times\hat{\mathbf{n}})e^{i(\mathbf{k}\cdot\boldsymbol{r}-\omega t)}\,,\end{aligned} \tag{3.44}$$

Furthermore, the relation between $\mathbf{E}$, $\mathbf{B}$ and $\mathbf{k}$ is such that

For a monochromatic electromagnetic plane wave in vacuum, we must have

$$\hat{\mathbf{k}}\cdot\mathbf{E}=\hat{\mathbf{k}}\cdot\mathbf{B}=0\,. \tag{3.45}$$

Such waves are known as **transverse waves**, with the perturbations (electric and magnetic fields) being in a direction perpendicular to the wavevector. Furthermore, $\mathbf{E}\cdot\mathbf{B}=0$, i.e. the $\mathbf{E}$ and $\mathbf{B}$ fields are orthogonal to each other.

If the wave is traveling along the z-direction, and we align polarization vector along the x-axis, then the expressions simplify to

$$\mathbf{E}=E_0\hat{\mathbf{x}}e^{i(kz-\omega t)}\,,\qquad \mathbf{B}=\frac{E_0}{c}\hat{\mathbf{y}}e^{i(kz-\omega t)}\,. \tag{3.46}$$

Fig. 12 shows an illustration of what this wave looks like. Notice that the solution has $\mathbf{E}$ and $\mathbf{B}$ in phase as well.

3.2.5 The Electromagnetic Spectrum

Electromagnetic waves are one of the most well-studied physical phenomena in all of physics, and they span an enormous range of frequencies and wavelengths. The entire range of electromagnetic waves is known as the **electromagnetic**

[33] You will notice that for plane waves, I could have gotten the answer of $\nabla\cdot$ and $\nabla\times$ by simply replacing ∇ with $i\mathbf{k}$, which is a very useful shortcut.

[34] Technically, we have only shown that $\mathbf{B}=\mathbf{B}_0\exp(i(\mathbf{k}\cdot\boldsymbol{r}-\omega t))+\mathbf{f}(\boldsymbol{r})$, but the fact that $\mathbf{B}$ also has to satisfy its own wave equation and the other Maxwell's equations would show that $f(\boldsymbol{r})=0$.

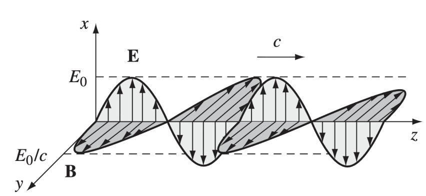

Figure 12: A plane electromagnetic wave moving in the z-direction. For a general wavevector $\mathbf{k}$, the direction of wave propagation will be $\hat{\mathbf{k}}$, with the planes of constant field vectors orthogonal to $\hat{\mathbf{k}}$.

Wavelength λ (m)	Frequency ν (GHz)	Photon Energy (eV)	Name	Applications
> 1	< 0.3	$< 10^{-6}$	Radio waves	AM/FM radio, TV broadcasting
$10^{-3} - 1$	$0.3 - 300$	$10^{-6} - 10^{-3}$	Microwaves	Radar, WiFi, microwave ovens
$7 \times 10^{-7} - 10^{-3}$	$300 - 4 \times 10^5$	$10^{-3} - 2$	Infrared	Thermal imaging, remote controls
$4 \times 10^{-7} - 7 \times 10^{-7}$	$4 \times 10^5 - 8 \times 10^5$	$2 - 3$	Visible light	Human vision, optical microscopy
$10^{-8} - 4 \times 10^{-7}$	$8 \times 10^5 - 3 \times 10^7$	$3 - 100$	Ultraviolet	Sterilization, fluorescence
$10^{-11} - 10^{-8}$	$3 \times 10^7 - 3 \times 10^{10}$	$100 - 10^5$	X-rays	Medical imaging, crystallography
$< 10^{-11}$	$> 3 \times 10^{10}$	$> 10^5$	Gamma rays	Cancer treatment, nuclear physics

Table 1: The electromagnetic spectrum.

spectrum. Table 1 shows how we broadly categorize different frequency ranges of EM waves, although the boundaries are not really sharply defined. There are also more specialized names for sub-bands (including, of course, the colors of the rainbow!), most of which I'm unfortunately not very familiar with myself. Note that I've also listed the *photon energy* associated with these waves. From quantum mechanics, you know that electromagnetic waves can be thought of as comprising quantized excitations of the electromagnetic field, each with energy given by $E = h\nu$, where h is Planck's constant. There are, of course, no such thing as photons in classical electromagnetism, but being familiar with these energies is so important in many fields of physics that it would feel somewhat ridiculous not to put it down in this table. We'll never discuss photons again in this class, but this is another reminder that you're studying a classical field theory that is still not quite the full story.

3.2.6 Energy and Momentum

Electromagnetic waves in vacuum also carry energy and momentum, as you might expect. In your problem set, you will work out the energy density, Poynting vector and momentum density of monochromatic plane waves in vacuum.

First, let me just write down the electric and magnetic fields, but using real notation here: you should have gone through the same exercise using complex notation on the problem set, and you can go through how to multiply complex exponentials in Appendix A. We have

$$\mathbf{E} = E_0 \hat{\mathbf{n}} \cos(\mathbf{k} \cdot \mathbf{r} - \omega t + \delta)\,, \tag{3.47}$$

$$\mathbf{B} = \frac{E_0}{c} (\hat{\mathbf{k}} \times \hat{\mathbf{n}}) \cos(\mathbf{k} \cdot \mathbf{r} - \omega t + \delta)\,. \tag{3.48}$$

The energy density of the wave is

$$\begin{aligned} u &= \frac{1}{2}\epsilon_0 \mathbf{E}^2 + \frac{1}{2\mu_0}\mathbf{B}^2 \\ &= \left(\frac{1}{2}\epsilon_0 E_0^2 + \frac{1}{2\mu_0 c} E_0^2\right) \cos^2(\mathbf{k} \cdot \boldsymbol{r} - \omega t + \delta) \\ &= \epsilon_0 E_0^2 \cos^2(\mathbf{k} \cdot \boldsymbol{r} - \omega t + \delta)\,. \end{aligned} \tag{3.49}$$

Not very surprisingly, the EM wave has an energy density that oscillates in space and time (and of course, it is always positive).

Next up, we have the Poynting vector, which gives the energy flux (energy per unit area per unit time) of the wave:

$$\begin{aligned} \mathbf{S} &= \frac{\mathbf{E} \times \mathbf{B}}{\mu_0} \\ &= \epsilon_0 c E_0^2 \hat{\mathbf{k}} \cos^2(\mathbf{k} \cdot \mathbf{r} - \omega t + \delta) \\ &= cu\hat{\mathbf{k}}\,. \end{aligned} \tag{3.50}$$

where I've computed the cross product of the normal vectors just by the right-hand rule, and used the fact that $1/(\mu_0 c) = \epsilon_0 c$. One way to understand this expression is that the electromagnetic wave is constantly propagating the energy density in the $\hat{\mathbf{k}}$-direction at speed c, and so the energy flux is just the energy density multiplied by the speed of propagation.

Finally, we have the momentum density of the wave, which is given by (see Eq. (2.36))

$$\wp = \mu_0 \epsilon_0 \mathbf{S} = \frac{u}{c} \hat{\mathbf{k}} \,. \tag{3.51}$$

One interesting consequence of this is that $u = |\wp| c$, i.e. the energy density of the wave is equal to the magnitude of the momentum density multiplied by the speed of light. More on this later.

End of Lecture: Monday, Feb 9 2026

All of these quantities are oscillatory, but the frequency at which these oscillations proceed is very large (or, equivalently, the wavelength over which these quantities change is very small). If you're trying to measure these quantities in an experiment, unless your detector has a very high time resolution, or equivalently a very small spatial resolution, you will only be able to measure the *time-averaged* or *space-averaged* values of these quantities. So let's compute the time-average of the energy density, which we'll denote $\langle u \rangle$. This is defined as integrating u over one period of the wave T, and then dividing that result by T:

$$\begin{aligned} \langle u \rangle &= \frac{1}{T} \int_0^T \mathrm{d}t \, \epsilon_0 E_0^2 \cos^2(\mathbf{k} \cdot \mathbf{r} - \omega t + \delta) \\ &= \frac{1}{T} \epsilon_0 E_0^2 \int_0^T \mathrm{d}t \, \frac{1}{2} \left(\cos[2(\mathbf{k} \cdot \mathbf{r} - \omega t)] + 1 \right) \,. \end{aligned} \tag{3.52}$$

The first integral is zero, since we are integrating a cosine with angular frequency 2ω over two periods (or you can just work it out to convince yourself that this is zero). The second integral just gives $T/2$, and so we have

$$\langle u \rangle = \frac{1}{2} \epsilon_0 E_0^2 \,. \tag{3.53}$$

You can also show that this is equivalently

$$\langle u \rangle = \frac{1}{2\mu_0} B_0^2 \,. \tag{3.54}$$

With this result, since all the quantities above are proportional,

$$\begin{aligned} \langle \mathbf{S} \rangle &= \frac{1}{2} c \epsilon_0 E_0^2 \hat{\mathbf{k}} \\ \langle \wp \rangle &= \frac{1}{2c} \epsilon_0 E_0^2 \hat{\mathbf{k}} \,. \end{aligned} \tag{3.55}$$

The quantity $|\langle \mathbf{S} \rangle|$ is also known as the **intensity** of the wave, and is often denoted I, which is the average power transmitted per unit area, with SI units $\mathrm{W\,m^{-2}}$.

3.3 Electromagnetic Waves in Matter

We've now seen that electromagnetic waves can propagate in vacuum with speed $c = 1/\sqrt{\mu_0\epsilon_0}$. We've taken a look at monochromatic plane waves, found that they are transverse, with electric and magnetic fields orthogonal to each other as well. We'll now move on to studying how electromagnetic waves propagate in other media, i.e. when there are materials present. This is going to get really interesting, because now we can look at waves passing from one medium to another, which leads to very rich physics.

3.3.1 Linear Media

We'll first consider linear media, where the polarization $\mathbf{P}$ and magnetization $\mathbf{M}$ are linearly related to the electric and magnetic fields, respectively. In this case, the fields $\mathbf{D}$ and $\mathbf{H}$ are proportional to $\mathbf{E}$ and $\mathbf{B}$, with $\mathbf{D} = \epsilon\mathbf{E}$ and $\mathbf{B} = \mu\mathbf{H}$. Let me remind you that Maxwell's equations in a medium is given by:

$$\begin{aligned} \nabla \cdot \mathbf{D} &= \rho_f \,, & \nabla \cdot \mathbf{B} &= 0 \,, \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \,, & \nabla \times \mathbf{H} &= \mathbf{J}_f + \frac{\partial \mathbf{D}}{\partial t} \,. \end{aligned} \tag{3.56}$$

Without free charges and currents, we have $\rho_f = 0$ and $\mathbf{J}_f = 0$. Writing everything in terms of $\mathbf{E}$ and $\mathbf{B}$, Ampere-Maxwell's law becomes

$$\nabla \times \frac{\mathbf{B}}{\mu} = \epsilon \frac{\partial \mathbf{E}}{\partial t} \implies \nabla \times \mathbf{B} = \mu\epsilon \frac{\partial \mathbf{E}}{\partial t} \,, \tag{3.57}$$

with remaining laws being

$$\nabla \cdot \mathbf{E} = 0 \,, \qquad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \,, \qquad \nabla \cdot \mathbf{B} = 0 \,. \tag{3.58}$$

Comparing this with Maxwell's equations in vacuum, we see that the only difference is that $\mu_0\epsilon_0 \to \mu\epsilon$, with everything else being the same. In particular, the fields still satisfy a wave equation:

$$\left(\nabla^2 - \frac{1}{\mu\epsilon} \frac{\partial^2}{\partial t^2} \right) \mathbf{E} = 0 \,, \tag{3.59}$$

and likewise for $\mathbf{B}$. This is pretty remarkable! Remember we said that waves satisfying the wave equation preserve their shape as they propagate. Here, we see that in linear media, waves also satisfy the wave equation, and so they *also* preserve their shape as they propagate. This means that waves don't get distorted when they propagate through linear media, and is responsible for the phenomenon of *transparency*.

> The **wave speed** in a linear medium is given by
>
> $$v = \frac{1}{\sqrt{\mu\epsilon}} \,, \tag{3.60}$$
>
> which in linear media should satisfy $v < c$. We define the **refractive index** of the medium as $n \equiv c/v$, given by
>
> $$n \equiv \sqrt{\frac{\mu\epsilon}{\mu_0\epsilon_0}} \tag{3.61}$$

Most of the time for linear materials, $\mu = \mu_0$, and so the refractive index is just $n = \sqrt{\epsilon/\epsilon_0}$. Typical values of the refractive index for common linear materials are $n \approx 1.3$ for water, $n \approx 1.5$ for glass, and $n \approx 2.4$ for diamond.

For other expressions, you can simply take the expressions for vacuum, and replace $\epsilon_0 \to \epsilon$, $\mu_0 \to \mu$, and $c \to v$.

3.3.2 Reflection and Transmission at Normal Incidence

What happens when an electromagnetic wave traveling in one medium hits the boundary of another medium? You've already seen the idea of reflection and refraction in earlier classes, but now that we have Maxwell's equations under our belt, let's take a close look at how this works. Let's start by looking at **normal incidence**, where the wave is incident on the boundary along the normal direction. Again, we'll look only at sinusoidal waves, since any arbitrary wave can be thought of as a superposition of sinusoidal waves.

Before we dive in, let me remind you how fields behave at the boundary between two linear media with no free charges (which is what we'll assume for the rest of this chapter):[35]

[35] If this is not familiar with you, please go back and review this in PY405. They can be derived by applying Maxwell's equations in an infinitesimal pillbox or loop that straddles the boundary, and then taking the limit as the size of the pillbox or loop goes to zero.

> The **boundary conditions** at the interface between two media that set how $\mathbf{E}$ and $\mathbf{B}$ are determined on either side of the boundary are given by
>
> $$\epsilon_1 E_1^\perp = D_1 = D_2 = \epsilon_2 E_2^\perp\,, \qquad B_1^\perp = B_2^\perp\,,$$
> $$\mathbf{E}_1^\parallel = \mathbf{E}_2^\parallel\,, \qquad \frac{1}{\mu_1}\mathbf{B}_1^\parallel = \mathbf{H}_1^\parallel = \mathbf{H}_2^\parallel = \frac{1}{\mu_2}\mathbf{B}_2^\parallel\,, \qquad (3.62)$$
>
> where $\perp$ and $\parallel$ denote the components of the fields perpendicular and parallel to the boundary, respectively.

We'll be using this a lot.

First, let's set up our coordinate system: the set-up is shown in Fig. 13. Let's have an incident wave traveling in the z-direction, and let's have the boundary between the two media be the xy-plane at $z = 0$, with medium 1 occupying the region $z < 0$ and medium 2 occupying the region $z > 0$. Let the polarization vector be in the x-direction, so that the magnetic field is in the y-direction. Explicitly, for the incident wave,

$$\begin{aligned}\mathbf{E}_I(z,t) &= E_{0,I}e^{i(k_I z - \omega_I t)}\hat{\mathbf{x}}\\ \mathbf{B}_I(z,t) &= \frac{E_{0,I}}{v_1}e^{i(k_I z - \omega_I t)}\hat{\mathbf{y}}\,,\end{aligned} \qquad (3.63)$$

with the velocity v_1 being the velocity of wave in medium 1, and so $\omega_I/k_I = v_1$. Let me remind you that $E_{0,I}$ is complex, and includes the phase of the sinusoidal wave.

What happens when the incident wave hits the boundary at $z = 0$? Some of the wave will be reflected back into medium 1, and some of the wave will be transmitted into medium 2. The **transmitted wave** has to travel in the same direction as the incident wave, just by symmetry, so let's write it as

$$\begin{aligned}\mathbf{E}_T(z,t) &= E_{0,T}e^{i(k_T z - \omega_T t)}\hat{\mathbf{x}}\\ \mathbf{B}_T(z,t) &= \frac{E_{0,T}}{v_2}e^{i(k_T z - \omega_T t)}\hat{\mathbf{y}}\,,\end{aligned} \qquad (3.64)$$

where v_2 is the velocity of the wave in medium 2, and so $\omega_T/k_T = v_2$. How did I know that the transmitted wave has the same polarization? Well, check the boundary condition on the parallel component of the electric field: $\mathbf{E}_I^\parallel = \mathbf{E}_T^\parallel$, which means that the transmitted wave must have the same polarization as the incident wave. This also means that both $\mathbf{B}$-fields are in the $\hat{\mathbf{y}}$ direction as well. In general, there will also be a **reflected wave**, which travels in the $-\hat{\mathbf{z}}$-direction, and so we can write it as

$$\mathbf{E}_R(z,t) = E_{0,R} e^{i(-k_R z - \omega_R t)} \hat{\mathbf{x}}$$
$$\mathbf{B}_R(z,t) = -\frac{E_{0,R}}{v_1} e^{i(-k_R z - \omega_R t)} \hat{\mathbf{y}} \, . \tag{3.65}$$

Again, the reflected wave has to have the same polarization, otherwise the boundary conditions cannot be satisfied. Notice the minus sign on the magnetic field: this is because the reflected wave is traveling in the $-\hat{\mathbf{z}}$-direction, and so the Poynting vector $(\mathbf{E} \times \mathbf{B})/\mu_0$ must point in the $-\hat{\mathbf{z}}$-direction, which means that $\mathbf{B}_R$ has to point in the $-\hat{\mathbf{y}}$-direction.

End of Lecture: Wednesday, Feb 11 2026

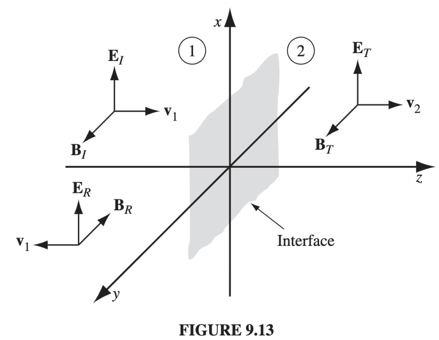

Figure 13: Coordinate system for normal incidence at a boundary.

We are now going to find the relationship between the wavenumbers, frequencies and amplitudes of the incident, transmitted and reflected waves by applying the boundary conditions at $z = 0$. Let's first set $z = 0$ and see what's going on. All of the $\mathbf{E}$-fields are in the $\hat{\mathbf{x}}$-direction, and given the boundary condition on $\mathbf{E}_{\parallel}$,

$$E_{0,T} e^{-i\omega_T t} = E_{0,I} e^{-i\omega_I t} + E_{0,R} e^{-i\omega_R t} \, . \tag{3.66}$$

This has to be true at all t! The only way this could be is if $\omega_T = \omega_I = \omega_R$, which means that the transmitted and reflected waves have the same frequency as the incident wave. This is an important result:

> The frequency of the incident, transmitted and reflected wave is *constant* across the interface.

Furthermore, we must also have

$$E_{0,T} = E_{0,I} + E_{0,R} \, . \tag{3.67}$$

All right, let's take stock of where we are now. We have

$$\mathbf{E}_I(z,t) = E_{0,I} e^{i(k_1 z - \omega t)} \hat{\mathbf{x}} \, ,$$
$$\mathbf{E}_T(z,t) = E_{0,T} e^{i(k_2 z - \omega t)} \hat{\mathbf{x}} \, ,$$
$$\mathbf{E}_R(z,t) = E_{0,R} e^{i(-k_1 z - \omega t)} \hat{\mathbf{x}} \, , \tag{3.68}$$

where I have defined $k_1 = \omega / v_1$ and $k_2 = \omega / v_2$, the wavenumbers in medium 1 and medium 2 respectively.

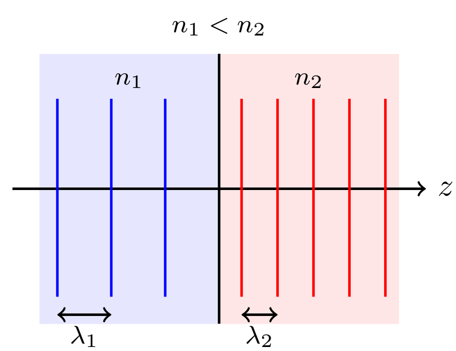

Figure 14: Incident and refracted wavefronts across an interface between two media (reflected wavefronts not shown). Note that the frequency remains constant!

> Notice that
>
> $$\frac{k_1}{k_2} = \frac{v_2}{v_1} \, , \tag{3.69}$$
>
> which is the relation between wavenumbers across the boundary of two different media. Alternatively,
>
> $$\frac{\lambda_1}{\lambda_2} = \frac{v_1}{v_2} \, , \tag{3.70}$$
>
> with a shorter wavelength in the medium with a smaller wave speed.

For the magnetic fields,

$$\mathbf{B}_I(z,t) = \frac{E_{0,I}}{v_1} e^{i(k_1 z - \omega t)} \hat{\mathbf{y}}\,,$$
$$\mathbf{B}_T(z,t) = \frac{E_{0,T}}{v_2} e^{i(k_2 z - \omega t)} \hat{\mathbf{y}}\,,$$
$$\mathbf{B}_R(z,t) = -\frac{E_{0,R}}{v_1} e^{i(-k_1 z - \omega t)} \hat{\mathbf{y}}\,, \tag{3.71}$$

with the boundary condition Eq. (3.67) still holding.

We've used the boundary condition for $\mathbf{E}_\parallel$; what about for $\mathbf{B}_\parallel$? That gives

$$\frac{1}{\mu_1}\left(\frac{E_{0,I}}{v_1} - \frac{E_{0,R}}{v_1}\right) = \frac{1}{\mu_2}\frac{1}{v_2} E_{0,T}\,, \tag{3.72}$$

or

$$E_{0,I} - E_{0,R} = \frac{\mu_1}{\mu_2}\frac{v_1}{v_2} E_{0,T}\,. \tag{3.73}$$

So we've got two equations, this one and Eq. (3.67). We can therefore relate $E_{0,R}$ and $E_{0,T}$ to $E_{0,I}$. Defining

$$\alpha = \frac{\mu_1 v_1}{\mu_2 v_2} = \frac{\mu_1}{\mu_2}\frac{n_2}{n_1}\,, \tag{3.74}$$

We find

$$E_{0,R} = \left(\frac{1-\alpha}{1+\alpha}\right) E_{0,I}\,, \qquad E_{0,T} = \left(\frac{2}{1+\alpha}\right) E_{0,I}\,. \tag{3.75}$$

Usually, in linear materials, we have $\mu_1 \approx \mu_2 \approx \mu_0$, and so $\alpha \approx n_2/n_1$. In this case, the above expressions simplify to

$$E_{0,R} = \left(\frac{v_2 - v_1}{v_2 + v_1}\right) E_{0,I}\,, \qquad E_{0,T} = \left(\frac{2v_2}{v_2 + v_1}\right) E_{0,I}\,, \tag{3.76}$$

or in terms of the refractive indices,

$$E_{0,R} = \left(\frac{n_1 - n_2}{n_1 + n_2}\right) E_{0,I}\,, \qquad E_{0,T} = \left(\frac{2n_1}{n_1 + n_2}\right) E_{0,I}\,. \tag{3.77}$$

Let me remind you that the amplitudes above are all *complex*, which means that it represents both the real amplitude and the phase as well. You can see that since $2v_2/(v_2 + v_1)$ is just a real number, we see that the transmitted wave is in phase with the incident wave. Explicitly, if we write $E_{0,I} = A_I e^{i\delta_I}$, $E_{0,R} = A_R e^{i\delta_R}$ and $E_{0,T} = A_T e^{i\delta_T}$, then we must have

$$A_T e^{i\delta_T} = \left(\frac{2v_2}{v_2 + v_1}\right) A_I e^{i\delta_I}\,, \tag{3.78}$$

which only works if $\delta_T = \delta_I$.

On the other hand, if $v_2 > v_1$, then the reflected and incident waves are in phase as well, while they are π-out of phase if $v_2 < v_1$, since a negative sign corresponds to $-1 = e^{i\pi}$.

Let's just write down the relationships between the real amplitudes; this is just

$$E_{0,R} = \left|\frac{n_1 - n_2}{n_1 + n_2}\right| E_{0,I}\,, \qquad E_{0,T} = \left(\frac{2n_1}{n_1 + n_2}\right) E_{0,I}\,. \tag{3.79}$$

The intensity of a wave in a linear medium is given by (see Eq. (3.55))

$$I = \frac{1}{2} v \epsilon E_0^2 \,, \tag{3.80}$$

and so the intensity of the incident, transmitted and reflected wave are related in the following way:

$$\begin{aligned} \frac{I_R}{I_I} &= \frac{A_R^2}{A_I^2} \,, \\ \frac{I_T}{I_I} &= \frac{v_2 \epsilon_2 A_T^2}{v_1 \epsilon_1 A_I^2} \,. \end{aligned} \tag{3.81}$$

These ratios tell you about how much energy gets reflected or transmitted at a boundary. The ratio of the reflected to incident intensity is called the **reflection coefficient**,

$$R \equiv \frac{I_R}{I_I} = \left| \frac{n_1 - n_2}{n_1 + n_2} \right|^2 \,, \tag{3.82}$$

and the ratio of the transmitted to incident intensity is called the **transmission coefficient**,

$$\begin{aligned} T \equiv \frac{I_T}{I_I} &= \frac{v_2 \epsilon_2}{v_1 \epsilon_1} \left(\frac{2n_1}{n_1 + n_2} \right)^2 \\ &= \frac{n_1}{n_2} \frac{n_2^2}{n_1^2} \left(\frac{2n_1}{n_1 + n_2} \right)^2 \\ &= \frac{4 n_1 n_2}{(n_1 + n_2)^2} \,. \end{aligned} \tag{3.83}$$

(Note again that these formulas all assume $\mu_1 \approx \mu_2 \approx \mu_0$). By energy conservation, we must have

$$R + T = 1 \,. \tag{3.84}$$

Let's take some limiting cases to see if these formulas make sense. If $n_1 = n_2$, then as far as the electromagnetic wave is concerned, the two media are the same. In that case, there should be no reflection, and indeed we get $R = 0$ and $T = 1$. On the other hand, if $n_2 \to \infty$, this is very much the same situation as a light string attached to a very heavy rope: the wave gets almost completely reflected, with

$$E_{0,R} \approx -E_{0,I} \,, \qquad E_{0,T} \approx \frac{4n_1}{n_2} E_{0,I} \,. \tag{3.85}$$

You can see that the reflected wave is completely out-of-phase, with a very small amplitude of transmission.

There was a lot of physics that just went by, so let's just recap the main points:

1. Every medium can be characterized by ϵ and μ, which leads to a different wave speed $v = 1/\sqrt{\mu\epsilon}$, and a refractive index $n = c/v$.

2. The frequency of the EM wave is constant across the boundary of two media.

3. The wavenumber or wavelength is is related by $\lambda_1/\lambda_2 = v_1/v_2$.

4. Using the boundary conditions for Maxwell's equations, you can compute the relation between the amplitudes of the incident, reflected and transmitted wave, which then gives you the reflection and transmission coefficients.

5. The transmitted wave is in phase with the incident wave, while the reflected wave is in phase with the incident wave if $v_2 > v_1$, and π-out of phase if $v_2 < v_1$.

End of Lecture: Friday, Feb 13 2026

End of Material for Midterm 1

A Multiplying Complex Exponentials

Throughout these notes, we use complex exponentials to represent sinusoidal waves, writing e.g. $\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$, with the understanding that the physical field is the real part of this expression. This works perfectly for addition, differentiation and integration, since all of these operations commute with taking the real part. However, as we mentioned earlier, things are not so simple when we need to *multiply* two such quantities together, which happens when computing things like the energy density $u \propto E^2$ or the Poynting vector $\mathbf{S} \propto \mathbf{E} \times \mathbf{B}$. This appendix explains the issue and how to handle it correctly.

A.1 The Problem with Multiplying Complex Exponentials

Let's start with a simple example. Consider two physical quantities represented by complex exponentials:

$$f(t) = \mathrm{Re}\left[Ae^{-i\omega t}\right] , \qquad g(t) = \mathrm{Re}\left[Be^{-i\omega t}\right] , \tag{A.1}$$

where A and B are complex amplitudes encoding both the real amplitude and the phase. Here I'm suppressing the spatial dependence for simplicity. The physical product of these two quantities is

$$f(t)g(t) = \mathrm{Re}\left[Ae^{-i\omega t}\right] \cdot \mathrm{Re}\left[Be^{-i\omega t}\right] . \tag{A.2}$$

You might be tempted to just multiply the two complex exponentials directly, giving $ABe^{-2i\omega t}$, and then take the real part. But this is *wrong*! The issue is that taking the real part does *not* commute with multiplication:

$$\mathrm{Re}[z_1] \cdot \mathrm{Re}[z_2] \neq \mathrm{Re}[z_1 \cdot z_2] . \tag{A.3}$$

To see why, write $z_1 = a_1 + ib_1$ and $z_2 = a_2 + ib_2$. Then $\mathrm{Re}[z_1] \cdot \mathrm{Re}[z_2] = a_1 a_2$, while $\mathrm{Re}[z_1 z_2] = a_1 a_2 - b_1 b_2$. When you multiply two complex numbers, the product of the imaginary parts gives a real contribution $-b_1 b_2$ that has no business being in the physical product. This is exactly the issue we warned about when we first introduced complex exponentials.

A.2 The Correct Way to Multiply

Let f and g be sinusoidal functions, with F and G being their complex exponential representation such that $\mathrm{Re}(F) = f$ and likewise $\mathrm{Re}(G) = g$. How do we correctly multiply two complex exponentials F and G so that the real part of the result is $f \cdot g$? Suppose $F = f + if'$ and $G = g + ig'$, with f' and g' being real. Remember that the imaginary part is unphysical, and so the only part of F and G that we care about is the real part. Notice that we can write $g = (G + G^*)/2$. With that, we can see that

$$\begin{aligned} \mathrm{Re}\left[F \cdot \frac{1}{2}(G + G^*)\right] &= \mathrm{Re}\left[(f + if') \cdot g\right] \\ &= fg \,, \end{aligned} \tag{A.4}$$

which is the the correct physical product. Hence, we have the following:

> For real functions f and g represented by the complex exponentials F and G, the product $f \cdot g$ can be obtained by taking performing the following operation on F and G:
>
> $$f \cdot g = \frac{1}{2} F(G + G^*) \,, \tag{A.5}$$
>
> where on the right-hand side we must remember to take the real part at the end in order for the two sides to truly be equal, but physicists are usually very lazy and just write it as shown above.

A.3 Time Averages

In many situations, we don't need the instantaneous value of the product, but rather the **time-averaged** value. As a reminder, the time average of a quantity $h(t)$ over one period $T = 2\pi/\omega$ is defined as

$$\langle h \rangle \equiv \frac{1}{T} \int_0^T \mathrm{d}t\, h(t) \,. \tag{A.6}$$

Let's start by collecting some useful results. First, the time average of any sinusoidal function at frequency ω (or any nonzero multiple of ω) over one full period is zero:

$$\langle \cos(n\omega t + \phi) \rangle = \langle \sin(n\omega t + \phi) \rangle = 0 \,, \qquad n = 1, 2, 3, \ldots \tag{A.7}$$

for any phase ϕ. In fact, this result follows from the definition of a period: integrating any sinusoid over one complete cycle gives zero.

From this, we can immediately derive the average of $\cos^2$ and $\sin^2$. Using the double-angle formula $\cos^2\theta = (1 + \cos 2\theta)/2$, we get

$$\langle \cos^2(\omega t + \phi) \rangle = \left\langle \frac{1 + \cos(2\omega t + 2\phi)}{2} \right\rangle = \frac{1}{2} \,, \tag{A.8}$$

since the $\cos(2\omega t + 2\phi)$ term averages to zero. By a similar argument, $\langle \sin^2(\omega t + \phi) \rangle = 1/2$ as well.

We'll now prove a very useful formula for the time average of the product of two sinusoidal quantities represented by complex exponentials at the same frequency. Consider two real, sinusoidal functions $f(t)$ and $g(t)$ represented by the complex exponentials $F(t) = Ae^{-i\omega t}$ and $G(t) = Be^{-i\omega t}$, where A and B

are complex amplitudes that are time-independent. Then the time-average of $f \cdot g$ is given by (using Eq. (A.5))

$$\langle f \cdot g \rangle = \left\langle F \cdot \frac{1}{2}(G + G^*) \right\rangle = \frac{1}{2}\langle F \cdot G \rangle + \frac{1}{2}\langle F \cdot G^* \rangle . \tag{A.9}$$

Notice that $F \cdot G = ABe^{-2i\omega t}$, which is a sinusoid at frequency 2ω, and so $\langle F \cdot G \rangle = 0$.

The time-average of the product of two sinusoidal functions represented by complex exponentials at the same frequency is given by

$$\langle f \cdot g \rangle = \frac{1}{2}\langle F \cdot G^* \rangle , \tag{A.10}$$

where once again we have a complex exponential on the right-hand side, and we must remember to take the real part at the end in order for the two sides to truly be equal, although the notation shown here is very common.

Example A.1

Let's use this to compute the time-averaged Poynting vector of a monochromatic plane wave in vacuum. Recall that the electric and magnetic fields of such a wave are given by

$$\begin{aligned} \mathbf{E}(\boldsymbol{r}, t) &= E_0 \hat{\mathbf{n}} e^{i(\mathbf{k}\cdot\boldsymbol{r} - \omega t)} , \\ \mathbf{B}(\boldsymbol{r}, t) &= \frac{E_0}{c}(\hat{\mathbf{k}} \times \hat{\mathbf{n}}) e^{i(\mathbf{k}\cdot\boldsymbol{r} - \omega t)} , \end{aligned} \tag{A.11}$$

where E_0 is the complex amplitude (encoding both the real amplitude and the phase), $\hat{\mathbf{n}}$ is the polarization direction, and $\hat{\mathbf{k}}$ is the direction of propagation. The Poynting vector is $\mathbf{S} = (\mathbf{E} \times \mathbf{B})/\mu_0$, which involves a product of two sinusoidal quantities at the same frequency. Using Eq. (A.10), the time-averaged Poynting vector is

$$\langle \mathbf{S} \rangle = \frac{1}{\mu_0}\langle \mathbf{E} \times \mathbf{B} \rangle = \frac{1}{2\mu_0}\mathbf{E}_0 \times \mathbf{B}_0^* , \tag{A.12}$$

where $\mathbf{E}_0 = E_0\hat{\mathbf{n}}$ and $\mathbf{B}_0 = (E_0/c)(\hat{\mathbf{k}} \times \hat{\mathbf{n}})$ are the complex amplitudes of the electric and magnetic fields respectively. Note again that the real part needs to be taken in the final expression, but nobody ever writes it explicitly. Computing the cross product, we get

$$\begin{aligned} \mathbf{E}_0 \times \mathbf{B}_0^* &= E_0 \hat{\mathbf{n}} \times \frac{E_0^*}{c}(\hat{\mathbf{k}} \times \hat{\mathbf{n}}) \\ &= \frac{|E_0|^2}{c}\hat{\mathbf{k}} , \end{aligned} \tag{A.13}$$

where we simply have to remember that $\hat{\mathbf{k}}$, $\hat{\mathbf{n}}$, and $\hat{\mathbf{k}} \times \hat{\mathbf{n}}$ form a right-handed orthonormal basis. This expression is already real, so we finally obtain

$$\langle \mathbf{S} \rangle = \frac{|E_0|^2}{2\mu_0 c}\hat{\mathbf{k}} = \frac{1}{2}\epsilon_0 c |E_0|^2 \hat{\mathbf{k}} , \tag{A.14}$$

where in the last step I used $1/(\mu_0 c) = \epsilon_0 c$. This is consistent with Eq. (3.55), where I used the *real* amplitude instead, which is $|E_0|$. If this is confusing, note that

$$E_0 \hat{\mathbf{n}} e^{i(\mathbf{k}\cdot\boldsymbol{r} - \omega t)} = |E_0| \hat{\mathbf{n}} e^{i(\mathbf{k}\cdot\boldsymbol{r} - \omega t + \delta)} \tag{A.15}$$

for the phase δ with $E_0 = |E_0|e^{i\delta}$, and taking the real part of this gives $|E_0|\hat{\mathbf{n}}\cos(\mathbf{k}\cdot\boldsymbol{r} - \omega t + \delta)$, which is a sinusoidal wave with real amplitude $|E_0|$.

Note that the time-averaging of a product works for a cross product too. You can see this by using index notation to write $(\mathbf{E}\times\mathbf{B})_i = \epsilon_{ijk}E_jB_k$, so that

$$\begin{aligned}\langle(\mathbf{E}\times\mathbf{B})_i\rangle &= \langle\epsilon_{ijk}E^jB^k\rangle \\ &= \epsilon_{ijk}\langle E^jB^k\rangle \\ &= \epsilon_{ijk}\frac{1}{2}E^j(B^k)^* \\ &= \frac{1}{2}\mathbf{E}\times\mathbf{B}^*\,. \end{aligned} \tag{A.16}$$