

PHY381 / PHY481

Advanced Electrodynamics

Pieter Kok, The University of Sheffield.

August 2015

PHY381 / PHY481: Lecture Topics

1. Geometrical vectors and vector fields
2. Calculating with the div, grad and curl
3. Index notation, unit vectors and coordinate systems
4. Maxwell's equations and the Lorentz Force
5. Gauge transformations and a particle in an electromagnetic field
6. The Laplace equation and the method of images
7. Separation of variables and Legendre polynomials
8. Electric and magnetic multipole moments
9. The wave equation, polarisation, phase and group velocities
10. Energy and momentum in electromagnetic field
11. Electric and magnetic dipole radiation
12. Radiation from accelerated charges
13. Macroscopic media
14. Waves in dielectrics, conductors and plasmas
15. Frequency-dependent refractivity and anomalous dispersion
16. Reflection and transmission of waves
17. Waveguides and coaxial cables
18. Cavities
19. Revision of special relativity
20. Covariant Maxwell's equations and the transformation of fields
21. Revision lecture (week before the exam)

Assessed Homework on vector calculus due at the end of week 2

Contents

1	Vector Calculus and Field Theories	7
1.1	Vectors	7
1.2	A geometrical representation of vector fields	10
1.3	Div, grad and curl	13
1.4	Second derivatives	17
1.5	Helmholtz' theorem	18
1.6	Index notation	18
1.7	Unit vectors and coordinate systems	21
2	Maxwell's Equations and the Lorentz Force	22
2.1	Electrostatic forces and potentials	22
2.2	Magnetostatic forces	23
2.3	Faraday's law	24
2.4	Charge conservation	25
2.5	The Maxwell-Ampère law	25
2.6	The Lorentz force	26
3	Scalar and Vector Potentials	28
3.1	The vector potential	28
3.2	Gauge transformations	29
3.3	A particle in an electromagnetic field	30
4	Solving Maxwell's Equations: Electro- and Magnetostatics	33
4.1	The Poisson and Laplace equations	33
4.2	Charge distributions near conducting surfaces	35
4.3	Separation of variables	36
4.4	Electric dipoles, quadrupoles and higher multipoles	38
4.5	Magnetic multipoles	41
5	Solving Maxwell's Equations: Electromagnetic Waves	46
5.1	Maxwell's equations in vacuum: the wave equation	46
5.2	The dispersion relation and group velocity	47
5.3	Polarization of electromagnetic waves	48
5.4	The wave equation with sources	49
6	Energy and Momentum of Electromagnetic Fields	51
6.1	Poynting's theorem and energy conservation	51
6.2	Momentum and stress in the electromagnetic field	52
7	Radiation Sources and Antennas	55
7.1	Electric dipole radiation	55
7.2	Magnetic dipole radiation	57
7.3	Radiation from accelerated charges	57

8	Electrodynamics in Macroscopic Media	61
8.1	Polarization and Displacement fields	61
8.2	Magnetization and Magnetic induction	62
8.3	Waves in macroscopic media	64
8.4	Frequency dependence of the index of refraction	66
9	Surfaces, Wave Guides and Cavities	69
9.1	Boundary conditions for fields at a surface	69
9.2	Reflection and transmission of waves at a surface	70
9.3	Wave guides	73
9.4	Resonant cavities	76
10	Relativistic formulation of electrodynamics	80
10.1	Four-vectors and transformations in Minkowski space	80
10.2	Covariant Maxwell equations	82
10.3	Invariant quantities	85
A	Special coordinates	86
B	Vector identities	88
C	Units in electrodynamics	89

Suggested Further Reading

1. **Introduction to Electrodynamics, 3rd Edition**, by David J. Griffiths, Pearson (2008). This is an excellent textbook, and the basis of most of these notes. This should be your first choice of background reading.
2. **Modern Electrodynamics**, by Andrew Zangwill, Cambridge University Press (2013). This is a very new, high-level textbook that will answer most of your advanced questions.
3. **Classical Electrodynamics, 3rd Edition**, by J. D. Jackson, Wiley (1998). This is the standard work on electrodynamics and considered the authority in the field. The level of this book is very high, but it is worth browsing through it for an indication of the breath of the topic, and for many advanced worked examples.

1 Vector Calculus and Field Theories

Electrodynamics is a theory of *fields*, and all matter enters the theory in the form of *densities*. All modern physical theories are field theories, from general relativity to the quantum fields in the Standard Model and string theory.

You will be familiar with particle theories such as classical mechanics, where the fundamental object is characterised by a position vector and a momentum vector. Ignoring the possible internal structure of the particles, they have six degrees of freedom (three position and three momentum components). Fields, on the other hand, are characterised by an *infinite* number of degrees of freedom. Let's look at some examples:

A vibrating string: Every point x along the string has a displacement y , which is a degree of freedom. Since there are an infinite number of points along the string, the displacement $y(x)$ is a field. The argument x denotes a location on a line, so we call the field *one-dimensional*.

Landscape altitude: With every point on a surface (x, y) , we can associate a number that denotes the altitude h . The altitude $h(x, y)$ is a two-dimensional field. Since the altitude is a scalar, we call this a *scalar* field.

Temperature in a volume: At every point (x, y, z) in the volume we can measure the temperature T , which gives rise to the three-dimensional scalar field $T(x, y, z)$.

In this course we are interested in *vector* fields. In the remainder of this chapter we will construct the theory of vector calculus from the various different notions of a vector.

1.1 Vectors

What is a vector? We often say that a vector is a geometrical object with a magnitude and direction, and we represent this graphically using an arrow. However, the arrow is not the only geometrical object with a magnitude and direction, as Fig. 1 shows. The *stack* has a direction (the grey arrow) and a magnitude given by the density of sheets. The magnitude of the *thumbtack* is given by the area of the flat surface, and the magnitude of the *sheaf* is given by the density of the lines. Often we do want to use arrows to represent vectors, but other times the stack, thumbtack or sheaf are more natural pictures of what is going on.

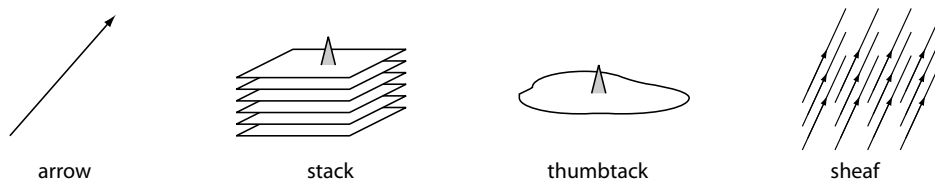


Figure 1: Geometrical objects with a magnitude and direction.

As an example, consider the capacitor C shown in Fig. 2a. The charge Q on the plates produces an electric field inside the capacitor. The direction of the vector \mathbf{d} gives the orientation of the capacitor, and its magnitude the plate separation. Similarly, the vector \mathbf{E} denotes the electric field inside the capacitor, which has both a magnitude and direction.

When we increase the distance between the capacitor plates from \mathbf{d} to \mathbf{d}' while keeping the potential V constant, the electric field inside the capacitor changes from \mathbf{E} to \mathbf{E}' . Graphically, this is represented in Fig. 2b as a longer vector \mathbf{d}' and a shorter vector \mathbf{E}' . Therefore, the arrow

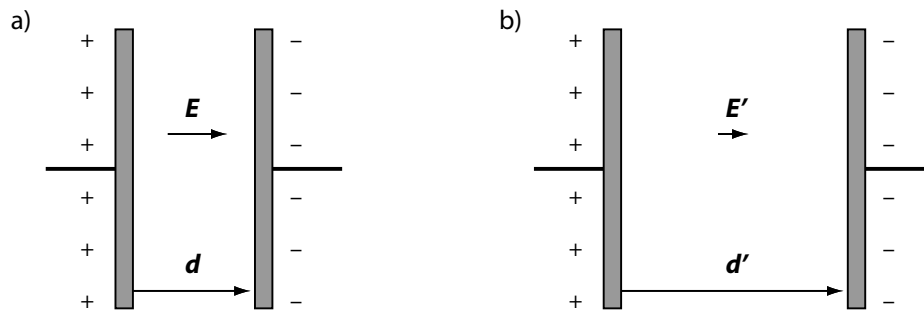


Figure 2: A capacitor with different distances between the two charges plates.

is a natural pictorial representation of the plate separation (because it grows along with the separation). However, the vector \mathbf{E} *shrinks* as \mathbf{d} grows (since the electric field decreases), and the stack is a more natural representation of the vector \mathbf{E} . As the plates are pulled apart, the sheets of the stack are pulled apart also, lowering the sheet density of the stack, and hence lowering the magnitude of the vector, as required.

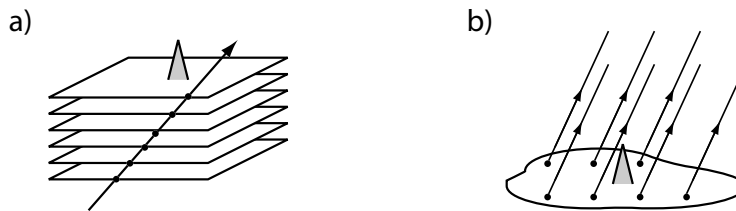


Figure 3: The inner product between vectors.

Next, consider the inner product (also known as dot product or scalar product) between the vectors \mathbf{d} and \mathbf{E} . From the physics of capacitors, we know that this is a scalar (namely the potential difference V), regardless of the separation between the plates:

$$\mathbf{E} \cdot \mathbf{d} = \mathbf{E}' \cdot \mathbf{d}' = V = \frac{Q}{C} = \text{constant.} \quad (1.1)$$

This has a straightforward meaning when we consider the arrows and stacks. The value of the inner product is equal to the number of times the arrow pierces the sheets in the stack. When there is an angle θ between the directions of the arrow and the stack, the number of pierced sheets is reduced by a factor $\cos\theta$, exactly as you would expect. We therefore always have to pair up an arrow with a stack in an inner product (see Fig. 3a).

We can also take the inner product between a thumbtack and a sheaf. In this case the value of the inner product is given by the number of times the sheaf pierces the surface of the thumbtack (see Fig. 3b). Again, this value is invariant under continuous squeezing and stretching of space.

We can also give a geometrical meaning to the cross product. Consider the cross product between two arrows, shown in Fig 4. The two arrows very naturally form the sides of a parallelogram with area $ab \sin\theta$, which coincides with the magnitude of the vector $\mathbf{c} = \mathbf{a} \times \mathbf{b}$. The natural graphical representation of a cross product is therefore the thumbtack, since the direction of the thumbtack is automatically perpendicular to the plane of the parallelogram. Can you construct the cross product of two stacks?

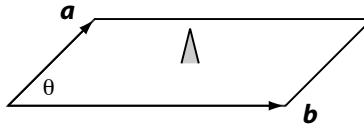


Figure 4: The cross product between vectors **a** and **b**.

In these lecture notes we denote vectors by boldface characters and vector components by regular italicised characters (as usual). Sometimes it is also instructive to indicate exactly what kind of geometrical object our vector is. In those cases we follow the convention that arrow vectors have an arrow over the symbol and sheafs have a double arrow. Stacks have a horizontal line under the vector symbol and thumbtacks have a circle. We write

$$\vec{\mathbf{A}} \times \vec{\mathbf{B}} = \vec{\mathbf{C}} \quad \text{and} \quad \underline{\mathbf{D}} \times \underline{\mathbf{E}} = \vec{\mathbf{F}}. \quad (1.2)$$

The embellishments are not strictly necessary, and most of the time we'll leave them out; in practice we can always deduce the type of vector from the context. There is a special nomenclature for the arrow, stack, thumbtack and sheaf that predates these geometrical pictures, and we will use this towards the end of the course when we discuss special relativity:

arrow = contravariant vector

stack = covariant vector

thumbtack = covariant vector capacity

sheaf = contravariant vector density

The covariant and contravariant vectors are the most important, since they play a central role in relativity.

Exercise 1.1: Show that $(\vec{\mathbf{a}} \times \vec{\mathbf{b}}) \cdot \vec{\mathbf{c}}$ is geometrically a volume (a capacity). What is $(\underline{\mathbf{a}} \times \underline{\mathbf{b}}) \cdot \underline{\mathbf{c}}$?

In addition to multiplying arrows with stacks we multiply capacities with densities to obtain scalars. You are already familiar with this in the context of mass and charge densities: you multiply the mass density of a body with its volume (a capacity) to get its mass (a scalar). Similarly, you multiply the charge density of a line charge with the length of the line (the capacity) to get the total charge (the scalar).

This pictorial approach to vector calculus is treated in detail in a little book called Geometrical Vectors by Gabriel Weinreich (Chicago Lectures in Physics, 1998), which would make excellent supplementary reading material.

It is important to distinguish between a vector and its components. The arrow *vector* \mathbf{v} is an object that is completely independent of any coordinate system, as shown in Fig. 5a (it has a length and a direction). We can decompose \mathbf{v} into components x and y relative to some coordinate system spanned by the orthonormal basis vectors \mathbf{e}_x and \mathbf{e}_y :

$$\mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix} = x \mathbf{e}_x + y \mathbf{e}_y \quad \text{with} \quad \mathbf{e}_x = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{e}_y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad (1.3)$$

as shown in Fig. 5b. (You probably know these basis better as $\mathbf{e}_x = \hat{\mathbf{i}}$ and $\mathbf{e}_y = \hat{\mathbf{j}}$.) Alternatively, we can decompose the same vector \mathbf{v} into components x' and y' relative to some *other* coordinate

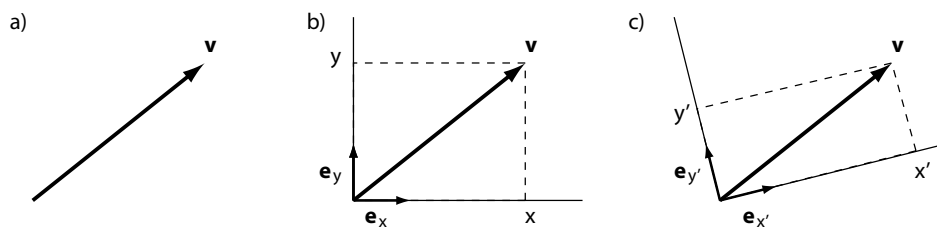


Figure 5: A vector and its components in different coordinate systems.

system spanned by the vectors $\mathbf{e}_{x'}$ and $\mathbf{e}_{y'}$:

$$\mathbf{v} = x' \mathbf{e}_{x'} + y' \mathbf{e}_{y'} = x \mathbf{e}_x + y \mathbf{e}_y. \quad (1.4)$$

The vector \mathbf{v} *does not depend on the coordinate system*, only its components do; \mathbf{v} has the same length and points in the same direction in all three instances in Fig. 5.

1.2 A geometrical representation of vector fields

Now that we are familiar with a whole family of geometrical vectors, we can construct vector *fields*. A vector field has a geometrical vector of some sort at every point in space. But before we go into more detail we first define what we mean by a *scalar field* Φ .

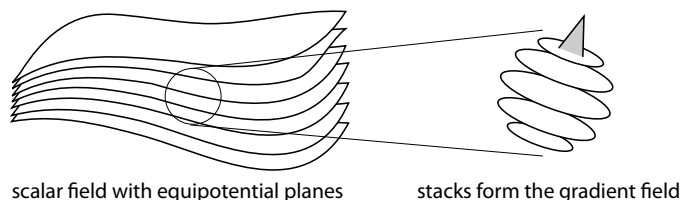


Figure 6: The scalar field.

If the scalar field is continuous, we can think of it as a collection of equipotential surfaces, as shown in Fig. 6. Each point in space lies on exactly one of the equipotential surfaces, which determines the value of the field at that point.

Next, we can take a small “melon baller” and scoop out a little three-dimensional sphere in the neighbourhood of such a point. You can see in Fig. 6 that this creates a small stack, with the direction given by the steepest ascent of the equipotential surfaces. We can repeat this procedure around every point in space, giving a stack vector at every point in space. This is of course exactly the *gradient* of the scalar field, which we can now see is a stack vector field \mathbf{F} . We therefore have

$$\underline{\mathbf{F}} = \text{grad } \Phi. \quad (1.5)$$

In physics we often add a minus sign (which is just a redefinition of \mathbf{F}), because “stuff flows downhill”. When we follow a path through the stack vector field \mathbf{F} from point 1 to point 2, the line integral along the path depends only on the end points:

$$\int_1^2 \underline{\mathbf{F}} \cdot d\vec{\mathbf{r}} = \Phi_2 - \Phi_1. \quad (1.6)$$

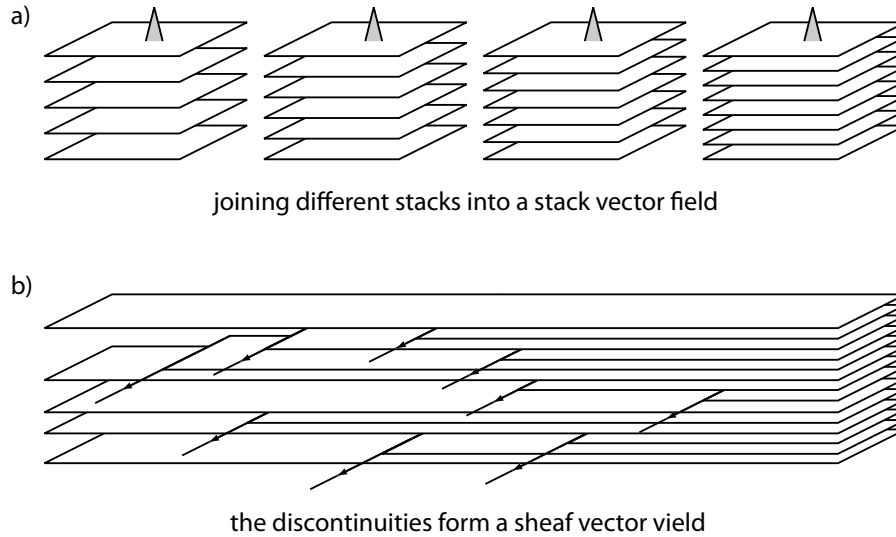


Figure 7: The stack vector field.

One possible physical interpretation of this integral is the work done moving a particle through the vector field from point 1 to point 2. If the vector field can be written as the gradient of a potential, then the work done depends only on the end points.

Next, consider a series of stack vectors as shown in Fig. 7a. If we place these stacks in the neighbourhood of each point in space, we create a stack vector field $\underline{\mathbf{A}}$. However, as you can see in Fig. 7b, they will *not* line up neatly into continuous equipotential planes. This is therefore *not* a gradient field. Rather, the discontinuities form a bundle of lines that remind us strongly of a sheaf vector. Indeed, the stack vector field that we have created has a *vector property* that is graphically represented by these sheafs. The simplest vector property of a vector field (apart from the vectors at each point in space, which we already represented by the stacks) is the *curl* of the vector field. The sheaf vectors in Fig. 7b are a graphical representation of the curl:

$$\vec{\mathbf{J}} = \text{curl} \underline{\mathbf{A}}. \quad (1.7)$$

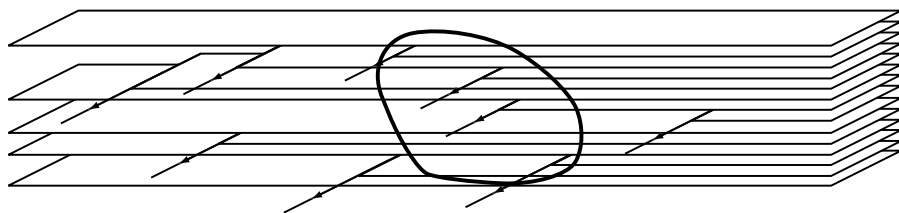
To gain some intuition that this is indeed the curl, imagine a loop around one of the lines in the sheaf. On the left, the loop will pierce fewer sheets of the stack than on the right. Since the density of the sheets is the magnitude of the vector, this means that the vector on the right will be larger than on the left. If the vectors describe a force, there will be a torque on the loop, causing a rotation. This is exactly the property of a vector field that is measured by the curl.

Now that we know that the sheaf vectors form the curl of the field $\underline{\mathbf{A}}$, we can immediately derive an important theorem. When $\underline{\mathbf{A}}$ is a gradient field ($\underline{\mathbf{A}} = \text{grad} \Phi$), there are no discontinuities (see Fig. 6), and therefore the curl of $\underline{\mathbf{A}}$ vanishes. Therefore

$$\text{curl}(\text{grad} \Phi) = 0, \quad (1.8)$$

for *any* scalar field Φ . It is kind of obvious from the pictures, but exact why this is as good a mathematical proof as any other requires some deep thinking on your part!

Another theorem that becomes almost trivial to prove is Stokes' theorem. Imagine that we create a closed loop in a stack vector field, as shown in Fig. 8. Let the loop be called C and the surface that the loop encompasses S . We can define the line integral around the loop as $\oint_C \underline{\mathbf{A}} \cdot d\mathbf{r}$,



the number of discontinuities in a loop C is the sheaf density times the surface S

Figure 8: Stokes' theorem.

which is just the number of discontinuities enclosed by C . We can calculate this differently by integrating the density of the discontinuities over the surface S . But this is nothing more than the surface integral over $\text{curl} \mathbf{A}$. We then arrive at

$$\oint_C \mathbf{A} \cdot d\mathbf{r} = \int_S \text{curl} \mathbf{A} \cdot d\mathbf{S}, \quad (1.9)$$

where the surface is naturally defined as a thumbtack. This is Stokes' theorem. You are familiar with this theorem in the form of Ampère's law, in which the magnetic field along a closed loop is proportional to the current through the surface spanned by the loop.

When we create the stack vector field \mathbf{A} , there is no beginning or end to the discontinuities that make up the sheaf $\vec{\mathbf{J}}$. This means that for a closed surface every line that goes in must come out, and the net effect is zero:

$$\oiint_S \text{curl} \mathbf{A} \cdot d\mathbf{S} = 0. \quad (1.10)$$

It is important to note that this is true for the curl of \mathbf{A} , but not necessary for \mathbf{A} itself. For example, we can create a vector field $\vec{\mathbf{H}}$ not from stacks, but from sheafs instead. When we do this, we will again encounter the possibility that line densities of the sheafs in neighbouring points do not match. In this case we have sheaf discontinuities, as shown in Fig. 9.

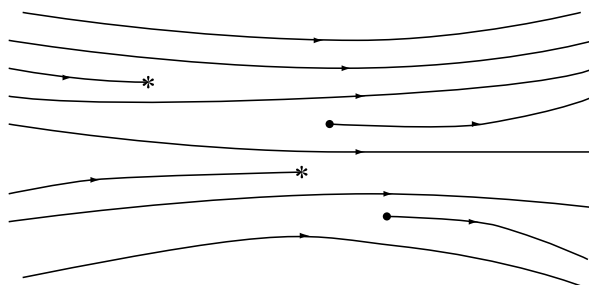


Figure 9: Sheaf discontinuities. The stars are end points and the dots are starting points.

We can call the density of these discontinuities ρ , and define this as the *divergence* of $\vec{\mathbf{H}}$:

$$\rho = \text{div} \mathbf{H}. \quad (1.11)$$

This is a scalar quantity, and behaves just as you would expect from the divergence, namely as the origin of field lines.

Since we have just seen that the curl of a vector field \mathbf{A} does not have any sheaf discontinuities, we immediately obtain the well-known theorem

$$\text{div}(\text{curl } \mathbf{A}) = 0. \quad (1.12)$$

Again, the proof is implicit in the geometrical representation.

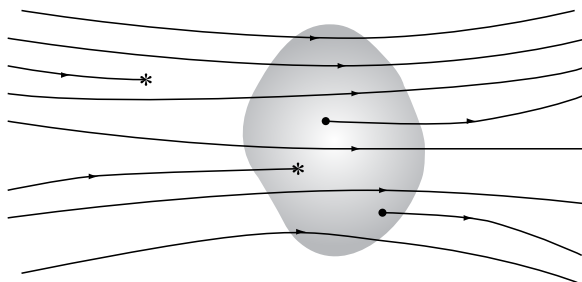


Figure 10: Gauss' theorem: the net number of lines out of the surface is given by the number of discontinuities in the volume.

Finally, we can prove Gauss' theorem for vector fields. Consider the sheaf vector field $\vec{\mathbf{H}}$ and the volume V shown in Fig. 10. The surface of the volume is denoted by S . The net number of lines coming out of the surface S must be identical to the net number of discontinuities inside the volume V (where start points are positive and end points are negative). Integrating over these two densities, we obtain

$$\oint_S \vec{\mathbf{H}} \cdot d\mathbf{S} = \int_V \text{div } \mathbf{H} dV. \quad (1.13)$$

This is Gauss' theorem. You are familiar with this theorem in electrostatics, where the flux of the electric field lines out of a closed surface is proportional to the total charge encapsulated by the surface.

Later in the course we will derive Maxwell's equations in macroscopic materials, in which the fields \mathbf{E} , \mathbf{B} , \mathbf{D} and \mathbf{H} play a crucial role. In geometric vector notation, these equations read

$$\begin{aligned} \text{curl } \underline{\mathbf{E}} &= -\frac{\partial \vec{\mathbf{B}}}{\partial t} \\ \text{div } \vec{\mathbf{D}} &= \rho \\ \text{curl } \underline{\mathbf{H}} &= \frac{\partial \vec{\mathbf{D}}}{\partial t} + \vec{\mathbf{J}} \\ \text{div } \vec{\mathbf{B}} &= 0, \end{aligned} \quad (1.14)$$

with the constitutive relations $\vec{\mathbf{D}} = \epsilon \underline{\mathbf{E}}$ and $\vec{\mathbf{B}} = \mu \underline{\mathbf{H}}$. Note how the Maxwell equations respect the geometric aspects of the vectors (e.g., where the curl of a stack is a sheaf), but the constitutive relations do not. This is because these are material properties that break the scale invariance of the geometric vectors.

1.3 Div, grad and curl

The geometrical representation is nice, and it gave us some powerful theorems, but in practice we still need to do actual calculations. In this section we will develop the mathematical tools you need to do these calculations.

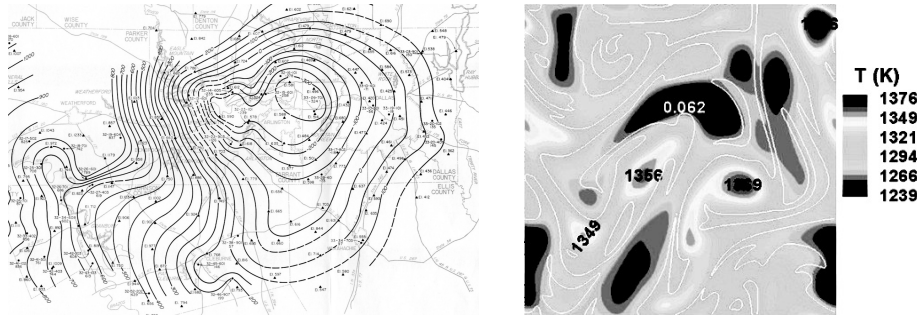


Figure 11: An altitude and temperature map with “equipotential” lines.

First, consider the scalar field. The interesting aspect of such a field is how the values of the field change when we move to neighbouring points in space, and in what direction this change is maximal. For example, in the altitude field (with constant gravity) this change determines how a ball would roll on the surface, and for the temperature field it determines how the heat flows. Note that in both these examples there are “equipotential planes” of constant altitude (where the “plane” is a line) and temperature (see Fig. 11). When we take our “melon baller” and carve out stack vectors in these maps, we obtain the steepest ascent in the altitude map, and the direction of heat flow in the temperature map. Previously we called this the *gradient*.

Let the scalar field be denoted by $f(x, y, z)$. Then the change in the x direction (denoted by $\hat{\mathbf{i}}$) is given by

$$\lim_{h \rightarrow 0} \frac{f(x+h, y, z) - f(x, y, z)}{h} \hat{\mathbf{i}} = \frac{\partial f(x, y, z)}{\partial x} \hat{\mathbf{i}}. \quad (1.15)$$

Similar expressions hold for the change in the y and z direction, and in general the spatial change of a scalar field is given by

$$\frac{\partial f}{\partial x} \hat{\mathbf{i}} + \frac{\partial f}{\partial y} \hat{\mathbf{j}} + \frac{\partial f}{\partial z} \hat{\mathbf{k}} = \nabla f \equiv \text{grad } f, \quad (1.16)$$

which we identified with the *gradient* of f . The “nabla” or “del” symbol ∇ is a differential *operator*, and it is also a (stack) vector:

$$\underline{\nabla} = \hat{\mathbf{i}} \frac{\partial}{\partial x} + \hat{\mathbf{j}} \frac{\partial}{\partial y} + \hat{\mathbf{k}} \frac{\partial}{\partial z}. \quad (1.17)$$

Clearly, this makes a stack vector field out of a scalar field.

Next, consider the curl of a vector field. In the previous section we saw that we need to construct a closed loop and see how the vectors change when we go around exactly once (see Fig. 12). Let the (stack) vector field be given by $\underline{\mathbf{A}}$ with components $A_x(x, y, z)$, $A_y(x, y, z)$ and $A_z(x, y, z)$. Note that *each* component A_x , A_y and A_z of the vector field $\underline{\mathbf{A}}$ generally depends on *all* three directions x , y and z . It is absolutely crucial that you appreciate this, because most errors in calculations tend to originate from the **false** belief that A_x depends only on x , A_y depends only on y , etc.

Consider the four-part infinitesimal closed loop in Fig. 12, starting at point $(x-l, y-l)$, going to $(x+l, y-l)$, $(x+l, y+l)$ and via $(x-l, y+l)$ back to $(x-l, y-l)$, as shown in figure 12. The area of the square is $4l^2$. The accumulated change of the vector field around this loop is given by the projection of $\underline{\mathbf{A}}$ along the line elements. We evaluate all four sides of the infinitesimal loop in

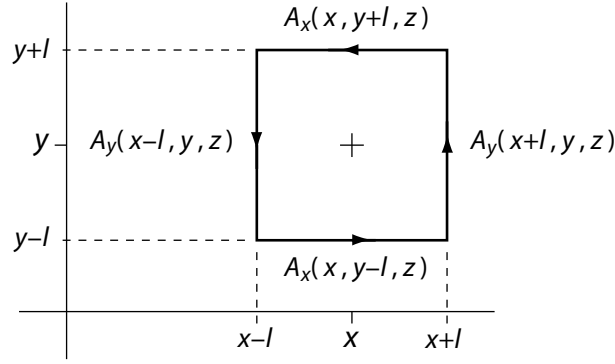


Figure 12: Calculating the *curl* of a vector field \mathbf{A} .

Fig. 12. For example, the line element that stretches from $x-l$ to $x+l$ at $y-l$ is oriented in the x direction, so \mathbf{l} directed along this part is

$$\vec{\mathbf{l}} = \begin{pmatrix} 2l \\ 0 \\ 0 \end{pmatrix}, \quad (1.18)$$

We take the inner product of \mathbf{l} with the stack vector $\underline{\mathbf{A}}$ at the point $(x, y-l, z)$:

$$\underline{\mathbf{A}} = (A_x(x, y-l, z), A_y(x, y-l, z), A_z(x, y-l, z)). \quad (1.19)$$

Note that all components of \mathbf{A} are evaluated at the *same* point. The inner product then becomes

$$\underline{\mathbf{A}} \cdot \vec{\mathbf{l}} = (A_x \ A_y \ A_z) \cdot \begin{pmatrix} 2l \\ 0 \\ 0 \end{pmatrix} = 2l A_x(x, y-l, z). \quad (1.20)$$

This is a scalar, but it changes from point to point in space.

We repeat the same procedure for the other three line elements. Make sure you get the orientation right, because the top horizontal part is directed in the $+x$ direction. The corresponding term gets a plus sign. Now add them all up and get $\mathbf{A} \cdot \mathbf{l}$ around the loop. We take the limit of $l \rightarrow 0$, so we replace every l with dl . We can then write for the infinitesimal loop

$$\oint \mathbf{A} \cdot d\mathbf{l} = 2dl [A_x(x, y-dl, z) + A_y(x+dl, y, z) - A_x(x, y+dl, z) - A_y(x-dl, y, z)]. \quad (1.21)$$

We can divide both sides by the area $4dl^2$, which gives

$$\frac{1}{4dl^2} \oint \mathbf{A} \cdot d\mathbf{l} = \left(\frac{A_x(x, y-dl, z) - A_x(x, y+dl, z)}{2dl} - \frac{A_y(x-dl, y, z) - A_y(x+dl, y, z)}{2dl} \right). \quad (1.22)$$

According to Stokes' theorem the left-hand side is proportional to the sheaf vector field $\text{curl} \mathbf{A}$ in the direction perpendicular to the surface enclosed by the loop. When we take the limit of $dl \rightarrow 0$ the vector field $\text{curl} \mathbf{A}$ becomes almost constant, and the surface area factorizes. At the same time, the right-hand side turns into partial derivatives:

$$\frac{1}{4dl^2} \int_S \text{curl} \mathbf{A} \cdot d\mathbf{S} = \frac{1}{4dl^2} \text{curl} \mathbf{A} \cdot \hat{\mathbf{k}} (2dl)^2 = (\text{curl} \mathbf{A})_z = \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right), \quad (1.23)$$

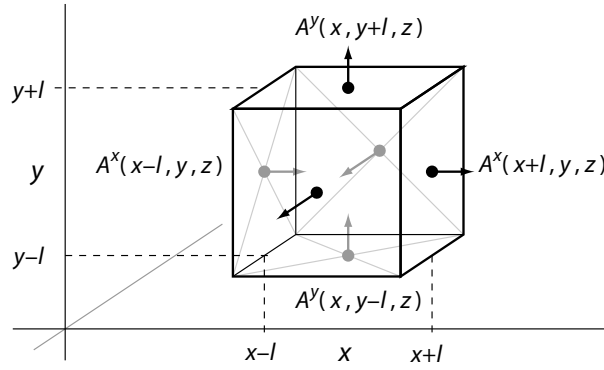


Figure 13: Calculating the *divergence* of a vector field \mathbf{A} .

where $(\text{curl } \mathbf{A})_z$ means the z component of $\text{curl } \mathbf{A}$.

We have made a loop in two dimensions (the xy plane), while our space is made of three dimensions. We can therefore make two more loops, in the xz and yz planes, respectively. The quantities $\oint_{xz} \mathbf{A} \cdot d\mathbf{l}$ and $\oint_{yz} \mathbf{A} \cdot d\mathbf{l}$ then form the y and x components of the sheaf vector field $\vec{\mathbf{J}} = \text{curl } \mathbf{A}$:

$$\vec{\mathbf{J}} = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \hat{\mathbf{i}} + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \hat{\mathbf{j}} + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \hat{\mathbf{k}}. \quad (1.24)$$

Every A_x , A_y and A_z value is taken at the same point in space (x, y, z) , which leads to the value of \mathbf{J} at that point (x, y, z) . Previously, we defined this relationship as $\mathbf{J} = \text{curl } \mathbf{A}$, so now we have an expression of this relationship in terms of differential operators. In compact matrix notation, the curl can be written as a determinant

$$\mathbf{J} = \nabla \times \mathbf{A} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \partial_x & \partial_y & \partial_z \\ A_x & A_y & A_z \end{vmatrix}. \quad (1.25)$$

Sometimes, when there are a lot of partial derivatives in an expression or derivation, it saves ink and space to write the derivative $\partial/\partial x$ as ∂_x , etc., with lower indices. Both ∂_x and A_x etc., are covariant vectors (stacks). We want to make a distinction in our notation of the components when the vectors are contravariant (the arrow and the sheaf), and we shall write these with *upper* indices.

Finally, we consider the divergence of a vector field. Consider the volume element $V = (2l)^3$ at point $\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$ and a vector field $\mathbf{A}(x, y, z)$, as shown in figure 13. We assume again that l is very small. From Gauss' theorem, the divergence is given by the number of discontinuities in the sheaf vector field inside the volume, which is directly related to the flux through the surface. All we need to do to find the divergence is calculate the net flux through the surface (and divide by the infinitesimal surface area).

To calculate the flux in the x direction we subtract the incoming and outgoing flux at the two surfaces of the cube that are perpendicular to the x direction. The incoming flux is given by the inner product of $\vec{\mathbf{A}}$ at point $(x-l, y, z)$ with the area $\mathbf{S}_0 = 4l^2 \hat{\mathbf{i}}$. For the outgoing flux we have the

same inner product, but evaluated at point $(x+l, y, z)$. The net flux in the x direction is therefore

$$\begin{aligned} \int_{S_x} \vec{\mathbf{A}} \cdot d\mathbf{S} &= 4l^2 [\mathbf{A}(x+l, y, z) \cdot \hat{\mathbf{i}} - \mathbf{A}(x-l, y, z) \cdot \hat{\mathbf{i}}] \\ &= 4l^2 [A^x(x+l, y, z) - A^x(x-l, y, z)], \end{aligned} \quad (1.26)$$

where we have written the components of \mathbf{A} with upper indices since \mathbf{A} is a sheaf vector field, and therefore contravariant. Similarly, the flux in the y and z direction is given by

$$\int_{S_y} \vec{\mathbf{A}} \cdot d\mathbf{S} = 4l^2 [A^y(x, y+l, z) - A^y(x, y-l, z)]. \quad (1.27)$$

and

$$\int_{S_z} \vec{\mathbf{A}} \cdot d\mathbf{S} = 4l^2 [A^z(x, y, z+l) - A^z(x, y, z-l)]. \quad (1.28)$$

The total flux is the sum over these. Using Gauss' theorem, we relate this to the divergence over the volume:

$$\begin{aligned} \frac{1}{(2l)^3} \int_V \text{div } \mathbf{A} dV &= \frac{A^x(x+l, y, z) - A^x(x-l, y, z)}{2l} \\ &\quad + \frac{A^y(x, y+l, z) - A^y(x, y-l, z)}{2l} \\ &\quad + \frac{A^z(x, y, z+l) - A^z(x, y, z-l)}{2l}. \end{aligned} \quad (1.29)$$

In the limit $l \rightarrow 0$, the right-hand side becomes a sum over derivatives and the left hand side is evaluated at the infinitesimal volume around a single point with constant \mathbf{A} in an infinitesimal volume $(2l)^3$. We can therefore write

$$\text{div } \mathbf{A} = \frac{\partial A^x}{\partial x} + \frac{\partial A^y}{\partial y} + \frac{\partial A^z}{\partial z} = \nabla \cdot \mathbf{A}, \quad (1.30)$$

which relates the divergence of a vector field to a differential operator that we can use in practical calculations.

1.4 Second derivatives

In physics, many properties depend on second derivatives. The most important example is Newton's second law $F = ma$, where a is the acceleration, or the second derivative of the position of a particle. It is therefore likely that we are going to encounter the second derivatives of fields as well. In fact, we are going to encounter them a lot! So what combinations can we make with the gradient, the divergence, and the curl?

The div and the curl act only on vectors, while the grad acts only on scalars. Moreover, the div produces a scalar, while the grad and the curl produce vectors. If f is a scalar field and \mathbf{A} is a vector field, you can convince yourself that the five possible combinations are

$$\begin{aligned} \nabla \cdot (\nabla f) &= \text{div grad } f, \\ \nabla(\nabla \cdot \mathbf{A}) &= \text{grad div } \mathbf{A}, \\ \nabla \times (\nabla f) &= \text{curl grad } f = 0, \\ \nabla \cdot (\nabla \times \mathbf{A}) &= \text{div curl } \mathbf{A} = 0, \\ \nabla \times (\nabla \times \mathbf{A}) &= \text{curl curl } \mathbf{A}. \end{aligned} \quad (1.31)$$

Of these, the first ($\nabla \cdot (\nabla f)$) defines a new operator $\nabla \cdot \nabla = \nabla^2$ called the Laplacian, which we can also apply to individual components of a vector field. The three non-zero derivatives are related by the vector identity

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A}, \quad (1.32)$$

which we will use very often.

Exercise 1.2: Prove equation (1.32).

1.5 Helmholtz' theorem

The curl and the divergence are in some sense complementary: the divergence measures the rate of change of a field along the direction of the field, while the curl measures the behaviour of the transverse field. If both the curl and the divergence of a vector field \mathbf{A} are known, and we also fix the boundary conditions, then this determines \mathbf{A} uniquely.

It seems extraordinary that you can determine a field (which has after all an infinite number of degrees of freedom) with only a couple of equations, so let's prove it. Suppose that we have two vector fields \mathbf{A} and \mathbf{B} with identical curls and divergences, and the same boundary conditions (for example, the field is zero at infinity). We will show that a third field $\mathbf{C} = \mathbf{A} - \mathbf{B}$ must be zero, leading to $\mathbf{A} = \mathbf{B}$. First of all, we observe that

$$\begin{aligned} \nabla \times \mathbf{C} &= \nabla \times \mathbf{A} - \nabla \times \mathbf{B} = 0 \\ \nabla \cdot \mathbf{C} &= \nabla \cdot \mathbf{A} - \nabla \cdot \mathbf{B} = 0, \end{aligned} \quad (1.33)$$

so \mathbf{C} has zero curl and divergence. We can use a vector identity of equation (1.32) to show that the second derivative of \mathbf{C} is also zero:

$$\nabla^2 \mathbf{C} = \nabla(\nabla \cdot \mathbf{C}) - \nabla \times (\nabla \times \mathbf{C}) = 0 - 0 = 0. \quad (1.34)$$

This also means that the Laplacian ∇^2 of every independent component of \mathbf{C} is zero. Since the boundary conditions for \mathbf{A} and \mathbf{B} are the same, the boundary conditions for \mathbf{C} must be zero. So with all that, can \mathbf{C} be anything other than zero? Eq. (1.34) does not permit any local minima or maxima, and it must be zero at the boundary. Therefore, it has to be zero inside the boundary as well. This proves that $\mathbf{C} = 0$, or $\mathbf{A} = \mathbf{B}$. Therefore, by determining the divergence and curl, the vector field is completely fixed, up to boundary conditions. This result is known as Helmholtz' theorem.

Looking ahead at the next section, you now know why there are four Maxwell's equations: two divergences and two curls for the electric and magnetic fields (plus their time derivatives). These four equations and the boundary conditions completely determine the fields, as they should.

1.6 Index notation

Consider again the vector \mathbf{v} from equation (1.3), but now in three dimensions:

$$\mathbf{v} = x \mathbf{e}_x + y \mathbf{e}_y + z \mathbf{e}_z. \quad (1.35)$$

Since the values x , y , and z are components of the contravariant (arrow) vector \mathbf{v} , we can write them as v^x , v^y , and v^z with upper indices. That way we have baked into the notation that we are talking about the components of \mathbf{v} . The vector can then be written as

$$\mathbf{v} = v^x \mathbf{e}_x + v^y \mathbf{e}_y + v^z \mathbf{e}_z. \quad (1.36)$$

Note that this is effectively an inner product, and the basis vectors \mathbf{e}_x , \mathbf{e}_y and \mathbf{e}_z are *covariant* vectors, or stacks. We also know that an inner product gives us something that is invariant under coordinate transformations. Indeed, it gives us the vector \mathbf{v} . If you're confused that an inner product yields a vector, remember that the vectorial character is carried by the basis vectors \mathbf{e}_x , \mathbf{e}_y and \mathbf{e}_z *inside* the inner product.

We can now introduce an enormously useful trick that will save us a lot of time. Notice that there is some redundancy in the notation of equation (1.36). We repeat, in a summation, the component multiplied by the corresponding basis vector. So we can write this straight away as a sum:

$$\mathbf{v} = \sum_{j=x,y,z} v^j \mathbf{e}_j. \quad (1.37)$$

Since we typically know (from the context) what our coordinate system is, we may refer to \mathbf{v} in terms of its components v^j . This is called *index notation*, which we will now develop for more general vector operations.

Next, suppose that $\mathbf{v} = \text{grad } f$ is a gradient field, so there is a vector \mathbf{v} at every point in space. We know that gradient fields are covariant (stack) vector fields, so the components v_j carry lower indices. We can write

$$v_x = \frac{\partial f}{\partial x} \quad v_y = \frac{\partial f}{\partial y} \quad \text{and} \quad v_z = \frac{\partial f}{\partial z}. \quad (1.38)$$

We can abbreviate the derivatives as ∂_x , ∂_y , and ∂_z (with lower indices to match the components of \mathbf{v} , since f does not carry any indices), which allows us to write

$$v_j = \partial_j f. \quad (1.39)$$

The next operator we consider is the divergence. We know that the divergence operates on a contravariant vector field (arrows or sheafs) with upper indices. This matches the del operator that, as we have seen in equation (1.39) carries lower indices. We therefore have the inner product

$$\text{div } \mathbf{A} = \nabla \cdot \mathbf{A} = \frac{\partial A^x}{\partial x} + \frac{\partial A^y}{\partial y} + \frac{\partial A^z}{\partial z}. \quad (1.40)$$

We can write the divergence as a sum over only one coordinate derivative:

$$\nabla \cdot \mathbf{A} = \sum_{j=x,y,z} \partial_j A^j. \quad (1.41)$$

So far, nothing new. However, we want to save writing even more. Notice that on the left-hand side of Eq. (1.41) there are no j 's (the divergence gives a scalar, which does not have any components), and on the right-hand side there are two j 's. Apparently, repeated indices (the j 's) imply that we need to sum over them! This allows us to be even more efficient, and also drop the summation symbol:

$$\nabla \cdot \mathbf{A} = \partial_j A^j, \quad (1.42)$$

This is called *Einstein's summation convention*: repeated indices are summed over (this is also called *contraction* of indices). A proper inner product always contracts a lower index with an upper index, because it multiplies a covariant with a contravariant vector. We will not be rigorous with the placement of indices until the last chapter on relativity, where it becomes important.

Notice that the difference between a scalar and a vector is indicated by the number of indices: f is a scalar, and $B_j \equiv \partial_j f$ is a vector (or rather the j^{th} component of a vector). We can extend this to other objects with multiple indices F_{jk} and η_{jkl} , etc. These are called *tensors*. For example, we can write a (random) mathematical equation

$$F_{jk} = \partial_j A_k + B_j(\partial_k f) + \partial_l \eta_{jk}{}^l. \quad (1.43)$$

Notice that each index on the left is matched exactly on the right, and repeated indices in the same term are summed over. If you were to write this out in long-hand you would notice that F has nine components, and η has 27 components. Index notation is not just about being lazy, it's about being practical!

The final vector operation we translate to index notation is the curl. Suppose that the curl of a covariant vector field \mathbf{A} is written as

$$\begin{aligned} \nabla \times \mathbf{A} &= \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \hat{\mathbf{e}}^x + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \hat{\mathbf{e}}^y + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \hat{\mathbf{e}}^z \\ &= (\partial_y A_z - \partial_z A_y) \hat{\mathbf{e}}^x + (\partial_z A_x - \partial_x A_z) \hat{\mathbf{e}}^y + (\partial_x A_y - \partial_y A_x) \hat{\mathbf{e}}^z. \end{aligned} \quad (1.44)$$

The curl of a covariant vector field is a contravariant field (a sheaf), so the $\hat{\mathbf{e}}^j$ carry upper indices. We will explore later how we get from lower to upper indices.

There is a pleasing symmetry (or asymmetry) in equation (1.44), which we want to exploit. First of all, notice that there are no terms in which any of the components of $\hat{\mathbf{e}}^i$, ∂_j , and A_k have identical indices. Second, if any two indices i , j , or k in $\hat{\mathbf{e}}^i$, ∂_j , and A_k are interchanged, the term changes sign (check this for yourself!). Since the curl of a vector field gives another vector, we can write component i of $\nabla \times \mathbf{A}$ as $(\nabla \times \mathbf{A})^i$. We then need to match indices, which requires a tensor of rank three (that means three indices; two to contract with the indices of ∂_j and A_k , and one for the component i). The most general choice possible is

$$(\nabla \times \mathbf{A})^i = \epsilon_{ijk} \partial_j A_k, \quad (1.45)$$

where ϵ_{ijk} is the Levi-Civita symbol¹. We have already established that

$$\epsilon_{iik} = \epsilon_{iki} = \epsilon_{iik} = 0 \quad \text{and} \quad \epsilon_{ijk} = -\epsilon_{jik} = -\epsilon_{kji} = \epsilon_{kij}, \quad (1.46)$$

and so on. The non-zero components of ϵ are either +1 or -1. A particularly useful identity is

$$\epsilon_{ijk} \epsilon_{klm} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}, \quad (1.47)$$

where δ_{ij} is the Kronecker delta:

$$\delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j. \end{cases} \quad (1.48)$$

You can always do an immediate check on vector equations with indices, because the *unpaired* indices on the left must match the *unpaired* indices on the right. The upper and lower indices must match also. When you do a calculation this is a quick sanity check: If the indices do not match you know something is wrong!

Exercise 1.3: Show that $\text{div}(\text{curl} \mathbf{A}) = 0$ using index notation.

Exercise 1.4: Prove equation (1.47).

¹Notice that we don't bother with upper and lower indices for ϵ_{ijk} . This is because the Levi-Civita symbol does not transform like normal vectors, and our geometrical picture breaks down here.

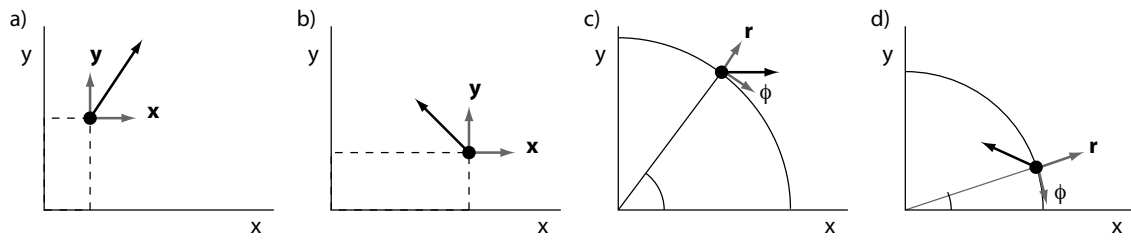


Figure 14: Unit vectors in cartesian and polar coordinates.

1.7 Unit vectors and coordinate systems

Finally, we briefly revisit the concept of unit vectors and coordinate systems. Consider a vector field in two dimensions (see Fig. 14). At any point in space, we have a vector with a magnitude and a direction, here a black arrow. We can express the vector components in cartesian coordinates, such as in Figs. 14a and 14b. In that case, for every point in the plane the (grey) cartesian unit vectors \mathbf{x} and \mathbf{y} point in the same direction, and they are perpendicular to each other.

On the other hand, if we express the vector in polar coordinates as in Figs. 14c and 14d, the (grey) polar unit vectors \mathbf{r} and ϕ point in different directions at different points! The unit vector \mathbf{r} always points radially outward, away from the origin, while ϕ always points tangential along the great circle through the point that is centred at the origin. At any point \mathbf{r} and ϕ are perpendicular. You should think of these unit vectors as a local coordinate system at the point (x, y) in the two-dimensional space.

Exercise 1.5: Express the unit vectors \mathbf{r} and ϕ in terms of \mathbf{x} and \mathbf{y} .

Summary

In this first section we have reviewed the mathematics that is necessary to master this module. In particular, you will need to be able to do the following:

1. know the difference between full and partial derivatives;
2. calculate the gradient of any scalar function $f(x, y, z)$, and calculate the divergence and curl of any vector field $\mathbf{A}(x, y, z)$;
3. use Gauss' and Stokes' theorems to manipulate integrals over vector fields;
4. evaluate line, surface, and volume integrals over vector fields.

Index notation is often a stumbling block for students, but it is nothing more than a new notation that requires some getting used to.

2 Maxwell's Equations and the Lorentz Force

In this section we review the laws of electrodynamics as you have learned them previously, and write them in the form of Maxwell's equations. We start with electrostatics and magnetostatics, and then we include general time-dependent phenomena. We also introduce the scalar and vector potential.

The theory of electrodynamics is mathematically quite involved, and it will get very technical at times. It is therefore important to know when we are being mathematically rigorous, and when we are just putting equations together to fit the observed phenomena. First, we postulate the laws. In fact, it was Coulomb, Biot and Savart, Gauss, Faraday, etc., who did measurements and formulated their observations in mathematical form. There is nothing rigorous about that (although the experiments were amazingly accomplished for the time). However, when Maxwell put all the laws together in a consistent mathematical framework the rules of the game changed: In order to find out what are the consequences of these *postulated* laws, we have to be mathematically rigorous, and see if our mathematical predictions are borne out in experiment.

2.1 Electrostatic forces and potentials

We start our journey to Maxwell's equations with the electrostatic force. You know from PHY101 that Gauss' law relates the flux of the electric field lines through a closed surface to the charge density inside the surface:

$$\oint_S \mathbf{E} \cdot d\mathbf{S} = \frac{Q_{\text{inside}}}{\epsilon_0}, \quad (2.1)$$

where Q_{inside} is the total (net) charge enclosed by the surface, and $\epsilon_0 = 8.85 \dots \times 10^{-12} \text{ F m}^{-1}$ is the electric permittivity of free space. Next, we use Gauss' theorem, which relates the flux of a vector field through a closed surface to the divergence of that vector field over the volume enclosed by the surface:

$$\oint_S \mathbf{E} \cdot d\mathbf{S} = \int_V \nabla \cdot \mathbf{E} d\mathbf{r}. \quad (2.2)$$

In addition, we can write the total charge inside the volume V as the volume integral over the charge *density* ρ :

$$Q_{\text{inside}} = \int_V \rho d\mathbf{r}, \quad (2.3)$$

We therefore have

$$\int_V \nabla \cdot \mathbf{E} d\mathbf{r} = \int_V \frac{\rho}{\epsilon_0} d\mathbf{r}, \quad (2.4)$$

Since this equation must hold for *any* volume V the integrands of both sides must be equal. This leads to our first Maxwell equation:

$$\boxed{\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}}. \quad (2.5)$$

It determines the divergence of \mathbf{E} . The physical meaning of this equation is that electric charges are the *source* of electric fields. The constant of proportionality ϵ_0 fixes the size of the electric field produced by a unit amount of charge.

You have also encountered the fact that a (static) electric field can be expressed as the gradient of a scalar function Φ , called the *scalar potential* (previously called V). This implies that the curl of \mathbf{E} is zero:

$$\mathbf{E} = -\nabla\Phi \quad \rightarrow \quad \nabla \times \mathbf{E} = 0, \quad (2.6)$$

since $\text{curl}(\text{grad } \Phi) = 0$ for any Φ . Combining Eqs. (2.5) and (2.6), we obtain the *Poisson* equation

$$\nabla^2\Phi = -\frac{\rho}{\epsilon_0}, \quad (2.7)$$

which, in vacuum ($\rho = 0$) becomes the *Laplace* equation

$$\nabla^2\Phi = 0. \quad (2.8)$$

The zero curl of \mathbf{E} , and hence the Poisson and Laplace equations are valid *only* in electrostatics. We will see later how fields that are changing over time modify these equations.

2.2 Magnetostatic forces

We have seen that the divergence of \mathbf{E} is given by the electric charge density. It is an experimental fact (at least at low energies in everyday situations) that there are no magnetic charges; only magnetic dipoles (such as electron spins and bar magnets) and higher seem to exist. This immediately allows us to postulate the next Maxwell equation:

$$\boxed{\nabla \cdot \mathbf{B} = 0}, \quad (2.9)$$

which is sometimes called *Gauss' law for magnetism*. The physical meaning of this equation is therefore that there are no magnetic monopoles. Perhaps we will find magnetic monopoles in the future², in which case this Maxwell equation needs to be modified.

For the last magnetostatic equation (the one for $\text{curl } \mathbf{B}$) we first consider Ampère's law, where the magnetic field component added along a closed loop is proportional to the total current flowing through the surface defined by the loop:

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \mu_0 I_{\text{enclosed}}, \quad (2.10)$$

where $\mu_0 = 4\pi \times 10^{-7} \text{ H m}^{-1}$ is the magnetic permeability of free space. You are familiar with this law from PHY102, where you used it to calculate the magnetic field due to an infinite current carrying wire, and similar problems.

We can apply Stokes' theorem to Eq. (2.10), which allows us to write

$$\oint_C \mathbf{B} \cdot d\mathbf{l} = \int_S (\nabla \times \mathbf{B}) \cdot d\mathbf{S}, \quad (2.11)$$

where S is any surface enclosed by the loop C . At the same time, we can write the enclosed current I_{enclosed} in terms of the current density \mathbf{J} as

$$I_{\text{enclosed}} = \int_S \mathbf{J} \cdot d\mathbf{S}. \quad (2.12)$$

²According to theoretical arguments in quantum field theory, the existence of even a single magnetic monopole in the universe would explain why charges come in discrete values of $e = 1.602176565 \times 10^{-19} \text{ C}$.

Combining these equations we can write

$$\int_S (\nabla \times \mathbf{B} - \mu_0 \mathbf{J}) \cdot d\mathbf{S} = 0. \quad (2.13)$$

Since this must be true for any surface S , the integrand of the integral must be zero:

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}. \quad (2.14)$$

This is Ampère's law, and valid *only* in magnetostatics.

2.3 Faraday's law

Now, let's consider general electromagnetic fields that may vary in time. Lenz' law states that the electromotive force \mathcal{E} on a closed conducting loop C is related to the change of the magnetic flux Φ_B through the loop shown in figure 15:

$$\mathcal{E} = \oint_C \mathbf{E} \cdot d\mathbf{l} = -\frac{\partial \Phi_B}{\partial t}, \quad (2.15)$$

with

$$\Phi_B = \int_S \mathbf{B}(\mathbf{r}, t) \cdot d\mathbf{S}. \quad (2.16)$$

Do not confuse Φ_B with the scalar potential Φ : they are two completely different things! We can use Stokes' theorem to write Eq. (2.15) as

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = \int_S (\nabla \times \mathbf{E}) \cdot d\mathbf{S} = - \int_S \frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} \cdot d\mathbf{S}. \quad (2.17)$$

Since this must be true for *any* surface S we can equate the integrands of the two integrals over S , and this gives us Maxwell's equation for the curl of \mathbf{E} :

$$\boxed{\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}}. \quad (2.18)$$

This is commonly known as *Faraday's law*. The physical meaning of this law is that a time-varying magnetic field generates vortices in the electric field (e.g., the discontinuities in figure 7). Our previous result of $\nabla \times \mathbf{E} = 0$ is true only when the magnetic field does not change over time (magnetostatics).

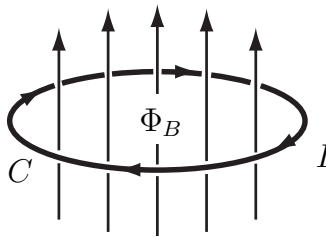


Figure 15: Lenz' law.

2.4 Charge conservation

Before we move to the last Maxwell equation, it is convenient to consider the mathematical expression of charge conservation. Charges inside a volume V bounded by the surface S can be a function of time:

$$Q(t) = \int_V \rho(\mathbf{r}, t), \quad (2.19)$$

and the current flowing *out of* the volume V bounded by the surface S is given by

$$I(t) = \oint_S \mathbf{J} \cdot d\mathbf{S} = \int_V \nabla \cdot \mathbf{J} d\mathbf{r}, \quad (2.20)$$

where \mathbf{J} is the current *density*. Conservation of charge then tells us that an increase of the charge in the volume V (a positive time derivative) must be balanced by a flow of charge into V through the surface S (a negative $I(t)$). The difference between these quantities must be zero, or

$$\frac{\partial Q(t)}{\partial t} + I(t) = 0. \quad (2.21)$$

When we substitute Eqs. (2.19) and (2.20) into Eq. (2.21), we find

$$\int_V \frac{\partial \rho}{\partial t} d\mathbf{r} + \int_V \nabla \cdot \mathbf{J} d\mathbf{r} = 0. \quad (2.22)$$

When we write this as a single integral over V , this equality can be satisfied only when the integrand is zero:

$$\int_V \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} d\mathbf{r} = 0, \quad (2.23)$$

and we find the general expression for charge conservation

$$\boxed{\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0.} \quad (2.24)$$

As far as we know today, charge conservation is strictly true in Nature.

2.5 The Maxwell-Ampère law

The fourth, and last, Maxwell equation is a modification of Ampère's law $\nabla \times \mathbf{B} = \mu_0 \mathbf{J}$. As it turns out, Ampère's law, as stated in Eq. (2.14) was wrong! Or at least, it does not have general applicability. To see this, let's take the divergence of Eq. (2.14):

$$\nabla \cdot \mathbf{J} = \frac{1}{\mu_0} \nabla \cdot (\nabla \times \mathbf{B}) = 0, \quad (2.25)$$

since the divergence of a curl is always zero. But charge conservation requires that

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t}, \quad (2.26)$$

So Ampère's law in Eq. (2.14) is valid only for static charge distributions (hence “electrostatics”). We can fix this by adding the relevant term to Ampère's law.

$$\nabla \cdot \mathbf{J} = \frac{1}{\mu_0} \nabla \cdot (\nabla \times \mathbf{B}) - \frac{\partial \rho}{\partial t}. \quad (2.27)$$

We use Gauss' law of Eq. (2.5) to write the charge density as a divergence $\rho = \epsilon_0 \nabla \cdot \mathbf{E}$, which yields

$$\nabla \cdot \mathbf{J} = \nabla \cdot \left(\frac{1}{\mu_0} (\nabla \times \mathbf{B}) - \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right). \quad (2.28)$$

The final Maxwell equation (the *Maxwell-Ampère law*) therefore has an extra term due to a time-varying electric field:

$$\boxed{\nabla \times \mathbf{B} - \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} = \mu_0 \mathbf{J}}. \quad (2.29)$$

The physical meaning of this is that closed magnetic field lines are created by currents and electric fields changing over time.

The complete set of (microscopic) Maxwell equations is

$$\nabla \cdot \mathbf{B} = 0 \quad \text{and} \quad \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \quad (2.30)$$

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \quad \text{and} \quad \nabla \times \mathbf{B} - \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} = \mu_0 \mathbf{J}. \quad (2.31)$$

The top two equations are the *homogeneous* Maxwell equations (they are equal to zero), and the bottom two are the *inhomogeneous* Maxwell equations (they are equal to a charge or current density). For every Maxwell equation we have the behaviour of the fields on the left hand side, and the source terms on the right hand side. This is generally how field equations are written. This way, it is clear how different source terms affect the fields.

Exercise 2.1: Show that charge conservation is contained in Maxwell's equations by taking the divergence of $\nabla \times \mathbf{B}$ and substituting $\nabla \cdot \mathbf{E}$.

2.6 The Lorentz force

The Maxwell equations tell us how the electric and magnetic fields are generated by charges and currents, and how they interact with each other (i.e., their dynamics). However, we still have to relate Maxwell's equations for the electromagnetic field to Newton's second law, $\mathbf{F} = m\mathbf{a}$. In other words, how are the fields related to the *force* on a particle?

A particle of charge q and velocity \mathbf{v} experiences the *Lorentz force*

$$\mathbf{F} = q\mathbf{E} + q\mathbf{v} \times \mathbf{B}, \quad (2.32)$$

and is sometimes used to *define* what we mean by \mathbf{E} and \mathbf{B} in the first place. In modern physics we say that \mathbf{E} and \mathbf{B} exist as physical objects in their own right, regardless of the presence of test charges q , or an æther, for that matter. The Lorentz force is necessary to complete the dynamical description of charged particles in an electromagnetic field.

Consider a zero \mathbf{B} field, and the electric field is created by a point particle with charge Q at the origin. We can go back to Gauss' law in integral form in Eq. (2.1) and say that the flux of electric field through a spherical surface centred around the charge Q is proportional to the charge:

$$\int_S \mathbf{E} \cdot d\mathbf{S} = \frac{Q}{\epsilon_0}. \quad (2.33)$$

By symmetry, the \mathbf{E} field points in the radial direction $\hat{\mathbf{r}}$, or $\mathbf{E} = E\hat{\mathbf{r}}$ with E the magnitude of \mathbf{E} . This is aligned perfectly with the infinitesimal surface elements $d\mathbf{S}$, or $\hat{\mathbf{r}} \cdot d\mathbf{S} = dS$ (since $\hat{\mathbf{r}}$ has

length 1). Moreover, the magnitude E is constant over the surface, also by symmetry. The integral therefore becomes

$$E \int_S dS = \frac{Q}{\epsilon_0}. \quad (2.34)$$

The radius of the sphere is $|\mathbf{r}|$, where \mathbf{r} is a point on the surface of the sphere. We can therefore evaluate the integral as the surface area of the sphere:

$$E \int_S dS = 4\pi |\mathbf{r}|^2 E = \frac{Q}{\epsilon_0}. \quad (2.35)$$

Putting back the direction of the \mathbf{E} field as $\hat{\mathbf{r}}$, we find that the electric field of a point charge at the origin is given by

$$\mathbf{E} = \frac{Q}{4\pi\epsilon_0 |\mathbf{r}|^2} \hat{\mathbf{r}}. \quad (2.36)$$

Substituting this into the Lorentz force with zero \mathbf{B} field yields

$$\mathbf{F} = \frac{qQ}{4\pi\epsilon_0 |\mathbf{r}|^2} \hat{\mathbf{r}}, \quad (2.37)$$

which you recognise as the *Coulomb force*. The force on the charge q at position \mathbf{r} points *away* from the charge Q (in the direction $\hat{\mathbf{r}}$), which means it is a *repulsive* force. When q and Q have opposite signs \mathbf{F} gains a minus sign and becomes *attractive*.

Summary

This section is a revision of last year's electromagnetism. We derived the Maxwell equations from experimental laws, and thus found a unified theory of electricity and magnetism. To relate the electric and magnetic fields to mechanical forces, we introduce the Lorentz force. From this section, you need to master the following techniques:

1. derive the Poisson and Laplace equations;
2. calculate the enclosed charge in a volume;
3. calculate the enclosed current and magnetic flux through a surface;
4. derive the law of charge conservation from Maxwell's equations;
5. calculate the Lorentz force on a particle given electric and magnetic fields.

Together with the previous section, this section forms the basis of the module. Make sure you fully understand its contents.

3 Scalar and Vector Potentials

You are familiar with the concept of a potential difference between two plates of a capacitor, or a potential produced by a battery. You also know that such a potential produces an electric field that is a gradient of this potential. In this section, we extend the concept of a potential not only to this familiar one (which we will denote by Φ), but also to a *vector* potential \mathbf{A} .

3.1 The vector potential

Since the electric field is the gradient of a scalar field (the *scalar* potential Φ), it is natural to ask whether the magnetic field is also some function of a potential. Indeed, this is possible, but rather than a scalar potential, we need a vector potential to make the magnetic field:

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (3.1)$$

The proof is immediate:

$$\nabla \cdot \mathbf{B} = \nabla \cdot (\nabla \times \mathbf{A}) = 0, \quad (3.2)$$

since the divergence of a curl is always zero. We can consider this the *definition* of the vector potential \mathbf{A} . At this point, the vector potential \mathbf{A} , as well as the scalar potential Φ , are strictly mathematical constructs. The physical fields are \mathbf{E} and \mathbf{B} . However, you can see immediately that whereas \mathbf{E} and \mathbf{B} have a total of six components (E_x, E_y, E_z, B_x, B_y , and B_z), the vector and scalar potentials have only four independent components (Φ, A_x, A_y , and A_z). This means that we can get a more economical description of electrodynamics phenomena using the scalar and vector potential.

Before we embark on this more compact description, we note that the electric field is also affected by the vector potential \mathbf{A} . The Maxwell equation

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (3.3)$$

implies that \mathbf{E} can't be a pure gradient, or we would have $\nabla \times \mathbf{E} = 0$. Instead, we use this equation and substitute Eq. (3.1) to give us

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = \nabla \times \mathbf{E} + \frac{\partial}{\partial t} \nabla \times \mathbf{A} = 0 \quad \text{or} \quad \nabla \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \right) = 0. \quad (3.4)$$

Therefore $\mathbf{E} + \dot{\mathbf{A}}$ is a gradient field:

$$\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} = -\nabla \Phi \quad \text{or} \quad \mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \nabla \Phi. \quad (3.5)$$

This is the proper expression of the electric field \mathbf{E} in terms of potentials. When \mathbf{A} does not depend on time, this reduces to the electrostatic case where the electric field is simply the (negative) gradient of the familiar scalar potential.

We should re-derive Poisson's equation (2.7) from $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$ using this form of \mathbf{E} :

$$\boxed{\nabla^2 \Phi + \frac{\partial}{\partial t} (\nabla \cdot \mathbf{A}) = -\frac{\rho}{\epsilon_0}}. \quad (3.6)$$

Similarly, we can write

$$\nabla \times \mathbf{B} = \nabla \times (\nabla \times \mathbf{A}) = \nabla (\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial}{\partial t} \left(-\nabla \Phi - \frac{\partial \mathbf{A}}{\partial t} \right). \quad (3.7)$$

We can rewrite the last equality as

$$\boxed{\nabla^2 \mathbf{A} - \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{A}}{\partial t^2} - \mu_0 \epsilon_0 \nabla \left(\frac{\partial \Phi}{\partial t} \right) - \nabla (\nabla \cdot \mathbf{A}) = -\mu_0 \mathbf{J}}. \quad (3.8)$$

Eqs. (3.6) and (3.8) are Maxwell's equations in terms of the scalar and vector potential, and amount to two second-order differential equations in Φ and \mathbf{A} . This looks a lot worse than the original Maxwell equations! Can we simplify these equations so that they look a bit less complicated? The answer is yes, and involves so-called *gauge transformations*.

3.2 Gauge transformations

As we have seen, the electric and magnetic fields can be written as derivative functions of a scalar potential Φ and a vector potential \mathbf{A} . You are familiar with the fact that adding a constant to a function will not change the derivative of that function. The same is true for the electric and magnetic fields: they will not change if we add constants to the potentials (you already know this about the scalar potential). However, since we are dealing with *three-dimensional* derivatives, it is not just adding constants to the potentials that do not change the fields; we can do a bit more!

Recall that $\mathbf{B} = \nabla \times \mathbf{A}$. Since $\nabla \times (\nabla \Lambda) = 0$ for all possible scalar functions Λ , we see that we can always add a gradient field $\nabla \Lambda$ to the vector potential. However, we established that the electric field \mathbf{E} depends on the time derivative of the vector potential. If \mathbf{E} is to remain invariant under these changes of the potentials, we need to add the time derivative of Λ to the scalar potential Φ :

$$\mathbf{E}' = -\nabla \Phi - \frac{\partial \mathbf{A}}{\partial t} - \frac{\partial \nabla \Lambda}{\partial t} + \nabla \frac{\partial \Lambda}{\partial t} = \mathbf{E}, \quad (3.9)$$

since $\nabla(\partial_t \Lambda) = \partial_t(\nabla \Lambda)$. It is clear that the full transformation is

$$\begin{aligned} \Phi(\mathbf{r}, t) &\rightarrow \Phi'(\mathbf{r}, t) = \Phi(\mathbf{r}, t) - \frac{\partial \Lambda(\mathbf{r}, t)}{\partial t}, \\ \mathbf{A}(\mathbf{r}, t) &\rightarrow \mathbf{A}'(\mathbf{r}, t) = \mathbf{A}(\mathbf{r}, t) + \nabla \Lambda(\mathbf{r}, t). \end{aligned} \quad (3.10)$$

This is called a *gauge* transformation. Let's give a formal definition:

Definition: A *gauge transformation* $\Lambda(\mathbf{r}, t)$ is a transformation (a change) of the potentials that leave the fields derived from these potentials unchanged.

On the one hand, you may think that it is rather inelegant to have non-physical degrees of freedom, namely the function $\Lambda(\mathbf{r}, t)$, because it indicates some kind of redundancy in the theory. However, it turns out that this gauge freedom is extremely useful, because it allows us to simplify our equations, just by choosing the right *gauge* Λ .

One possible choice is to set $\nabla \cdot \mathbf{A} = 0$, which is called the *Coulomb gauge* (or *radiation gauge*). We will see why this is a very useful choice when we discuss electromagnetic waves. Another useful gauge is the *Lorenz gauge*, in which we set

$$\nabla \cdot \mathbf{A} + \mu_0 \epsilon_0 \frac{\partial \Phi}{\partial t} = 0. \quad (3.11)$$

This leads to the following Maxwell equations:

$$\left(\nabla^2 - \mu_0 \epsilon_0 \frac{\partial^2}{\partial t^2} \right) \Phi = -\frac{\rho}{\epsilon_0}, \quad (3.12)$$

$$\left(\nabla^2 - \mu_0 \epsilon_0 \frac{\partial^2}{\partial t^2} \right) \mathbf{A} = -\mu_0 \mathbf{J}. \quad (3.13)$$

Now you see that Maxwell's equations have simplified a tremendous amount! Gauge transformations are important in field theories, particularly in modern *quantum* field theories. The differential operator in brackets in Eqs. (3.12) and (3.13) is called the *d'Alembertian*, after Jean le Rond d'Alembert (1717–1783), and is sometimes denoted by the symbol \square . This makes for a super-compact expression of Maxwell's equations:

$$\square \Phi = -\frac{\rho}{\epsilon_0} \quad \text{and} \quad \square \mathbf{A} = -\mu_0 \mathbf{J}. \quad (3.14)$$

The differential equations (3.12) and (3.13) are called d'Alembert equations.

Exercise 3.1: Show that the Lorenz gauge leads to Maxwell's equations of the form given in Eqs. (3.12) and (3.13).

3.3 A particle in an electromagnetic field

Often we want to find the equations of motion for a particle in an electromagnetic field, as given by the scalar and vector potentials. There are several ways of doing this, one of which involves the *Hamiltonian* $H(\mathbf{r}, \mathbf{p})$, where \mathbf{r} and \mathbf{p} are the position and momentum of the particle, respectively. You know the Hamiltonian from quantum mechanics, where it is the *energy operator*, and it is used in the Schrödinger equation to find the dynamics of the wave function of a quantum particle. However, the Hamiltonian was first constructed for classical mechanics as a regular function of position and momentum, where it also completely determines the dynamics of a classical particle. The idea behind this is that the force on an object is the spatial derivative of some potential function. Once the Hamiltonian is known, the equations of motion become

$$\frac{dr_i}{dt} = \frac{\partial H}{\partial p_i} \quad \text{and} \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial r_i}. \quad (3.15)$$

These are called *Hamilton's equations*, and it is clear that the second equation relates the force (namely the change of momentum) to a spatial derivative of the Hamiltonian.

The derivation of the Hamiltonian of a particle with mass m and charge q in an electromagnetic field is beyond the scope of this course, so we will just state it here:

$$H(\mathbf{r}, \mathbf{p}) = \frac{1}{2m} [\mathbf{p} - q\mathbf{A}(\mathbf{r}, t)]^2 + q\Phi(\mathbf{r}, t), \quad (3.16)$$

where \mathbf{r} is the position and \mathbf{p} the momentum of the particle. In a homework exercise you will be asked to show that this very general Hamiltonian leads directly to the Lorentz force on the particle. The Hamiltonian can be further simplified for the specific problem at hand by choosing the most suitable gauge. In addition, the field is often weak enough that we can set $\mathbf{A}^2 \approx 0$. In this case the Hamiltonian becomes that of a free particle with a coupling term $\frac{q}{m}\mathbf{p} \cdot \mathbf{A}$, also known as *dipole coupling*.

As a simple example, consider a particle with mass m and charge q in a field with scalar potential $\Phi = 0$ and vector potential $\mathbf{A} = A_0 \cos(kz - \omega t)\hat{\mathbf{x}}$. What are the equations of motion if the particle is at rest at time $t = 0$? Since \mathbf{A} has only an x -component, we can write the Hamiltonian as

$$\begin{aligned} H &= \frac{1}{2m} (\mathbf{p} \cdot \mathbf{p} - 2q\mathbf{p} \cdot \mathbf{A} + q^2 \mathbf{A} \cdot \mathbf{A}) + q\Phi \\ &= \frac{1}{2m} [p_x^2 + p_y^2 + p_z^2 - 2qp_x A_0 \cos(kz - \omega t) + q^2 A_0^2 \cos^2(kz - \omega t)]. \end{aligned} \quad (3.17)$$

the first set of Hamilton's equations then become

$$\dot{x} = \frac{dx}{dt} = \frac{\partial H}{\partial p_x} = \frac{p_x}{m} - \frac{qA_0}{m} \cos(kx - \omega t); \quad \dot{y} = \frac{dy}{dt} = \frac{\partial H}{\partial p_y} = \frac{p_y}{m}; \quad \dot{z} = \frac{dz}{dt} = \frac{\partial H}{\partial p_z} = \frac{p_z}{m},$$

and the second set become

$$\begin{aligned} \dot{p}_x &= \frac{dp_x}{dt} = -\frac{\partial H}{\partial x} = 0, \\ \dot{p}_y &= \frac{dp_y}{dt} = -\frac{\partial H}{\partial y} = 0, \\ \dot{p}_z &= \frac{dp_z}{dt} = -\frac{\partial H}{\partial z} = \frac{qp_x A_0 k}{m} \sin(kz - \omega t) + \frac{q^2 A_0^2 k}{m} \sin(2kz - 2\omega t). \end{aligned}$$

The boundary condition states, among others, that $p_x = 0$ at time $t = 0$. This will remain the case at all times, since $\dot{p}_x = 0$. The first term in \dot{p}_z is therefore zero, and the force on the particle is given by

$$\mathbf{F} = \frac{dp_x}{dt} \hat{\mathbf{x}} + \frac{dp_y}{dt} \hat{\mathbf{y}} + \frac{dp_z}{dt} \hat{\mathbf{z}} = \frac{q^2 A_0^2 k}{m} \sin(2kz - 2\omega t) \hat{\mathbf{z}}. \quad (3.18)$$

The momentum in the z -direction will change accordingly. This is a periodic force on the charged particle due to an oscillating potential. Note the doubling of the frequency.

This example elucidates the most basic operation of a receiver antenna. A wave with frequency ω produces a periodic force on a charge q in the $\hat{\mathbf{z}}$ -direction. If we align a thin conducting wire to the $\hat{\mathbf{z}}$ -direction, the moving charges create a current that can be measured. We will discuss the antenna as a source of radiation in Section 7.

As mentioned, the Hamiltonian completely determines the dynamics of a system, and as such is a very useful quantity to know. In quantum mechanics, we replace the position vector \mathbf{r} by the operator $\hat{\mathbf{r}}$ and the momentum vector \mathbf{p} by the operator $\hat{\mathbf{p}}$ (note that the hat now denotes an operator, rather than a unit vector. In the rest of the lecture notes the hat denotes unit vectors). This makes the Hamiltonian an operator as well. The Schrödinger equation for a particle in an electromagnetic field is therefore

$$i\hbar \frac{d}{dt} |\Psi\rangle = \left[\frac{1}{2m} (\hat{\mathbf{p}} - q\mathbf{A})^2 + q\Phi \right] |\Psi\rangle. \quad (3.19)$$

The vector potential \mathbf{A} and the scalar potential Φ are classical fields (not operators). The full theory of the quantized electromagnetic field is *quantum electrodynamics*.

The electric and magnetic fields are the physical objects that we usually say “exist” in Nature, independent of observers or test charges. On the other hand, the potentials are purely mathematical constructs in the classical theory of electrodynamics. The fact that we can change them using gauge transformations without changing any physically measurable quantities suggests that they are merely convenient mathematical tools. However, in quantum mechanics things are more complicated (as always). Since the Hamiltonian of a charged particle couples to the potentials, rather than the fields, there are instances where a change in the potentials leads to observable phenomena, e.g., the Aharonov-Bohm effect.

Summary

In this section, we introduced the electric potential (familiar from previous years) and the vector potential. These give an *alternative* description of electromagnetism to electric and magnetic fields, and describe *the same* phenomena. However, the potentials have a certain freedom in their mathematical description that has no observable consequences. This is called a *gauge* freedom. Changing the gauge of the potentials is called a gauge transformation. For a charged particle in an electromagnetic field we can set up a Hamiltonian, which via Hamilton's equations leads to the equations of motion of the particle.

You should master the following techniques:

1. calculate the fields \mathbf{E} and \mathbf{B} from the potentials Φ and \mathbf{A} ;
2. apply gauge transformations to the potentials, in other words: calculate

$$\Phi'(\mathbf{r}, t) = \Phi(\mathbf{r}, t) - \frac{\partial \Lambda(\mathbf{r}, t)}{\partial t} \quad \text{and} \quad \mathbf{A}'(\mathbf{r}, t) = \mathbf{A}(\mathbf{r}, t) + \nabla \Lambda(\mathbf{r}, t)$$

for some scalar function $\Lambda(\mathbf{r}, t)$;

3. find a $\Lambda(\mathbf{r}, t)$ to simplify the potentials, e.g., when we require $\mathbf{A} = 0$ or $\Phi = 0$ in a problem;
4. calculate Hamilton's equations for a given Hamiltonian.

4 Solving Maxwell's Equations: Electro- and Magnetostatics

We have established that Maxwell's equations (together with the Lorentz force) describe all electrodynamics phenomena. However, they are *differential equations* in the electric and magnetic fields, and we need to establish how to solve them given some initial or boundary conditions. This is usually a very complicated task, and we can only give a few methods for simple problems here. More complicated problems typically require numerical methods on computers. I will be following chapter 3 in Griffiths (2008) for this section.

4.1 The Poisson and Laplace equations

First we consider electrostatic situations in which the magnetic field plays no role. We have seen that this leads to a zero curl of the electric field $\nabla \times \mathbf{E} = 0$, and \mathbf{E} is fully determined by the scalar potential $\mathbf{E} = -\nabla\Phi$. This leads to the Poisson equation

$$\nabla^2\Phi = -\frac{\rho}{\epsilon_0}, \quad (4.1)$$

or, in regions of space without charge, the Laplace equation

$$\nabla^2\Phi = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \Phi = 0. \quad (4.2)$$

Electrostatics is all about solving these equations. In particular, we will be looking at solving the Laplace equation in some region of space V where the scalar potential is due to some distant charges outside the area of our interest.

The Laplace equation in one dimension

In one dimension, the potential Φ depends only on a single variable, and we can write

$$\frac{d^2\Phi}{dx^2} = 0. \quad (4.3)$$

You can immediately solve this very easy differential equation analytically by writing

$$\Phi(x) = ax + b, \quad (4.4)$$

where the boundary conditions determine a and b .

This is pretty much all there is to it in one dimension, but it will be instructive to note two properties of this general solution:

1. At point x , the scalar potential $\Phi(x)$ can be seen as the *average* of the two potentials on equidistant points on either side of x :

$$\Phi(x) = \frac{1}{2} [\Phi(x-u) + \Phi(x+u)] \quad \text{for any } u \text{ in the domain of } \Phi.$$

2. There are no local maxima or minima, as this would invalidate point 1.

The Laplace equation in two dimensions

Now we will consider the more interesting problem

$$\frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} = 0. \quad (4.5)$$

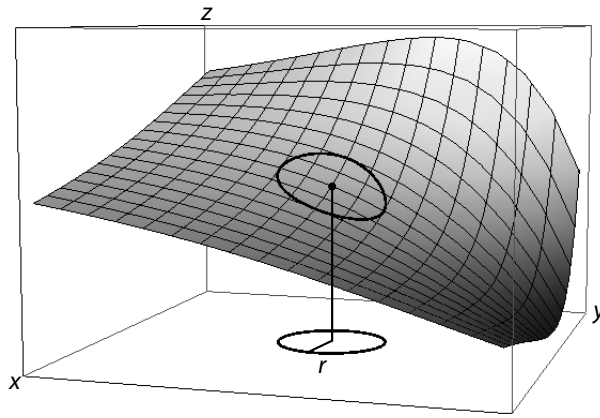


Figure 16: A solution to Laplace's equation in two dimensions.

This is a *partial* differential equation, and it will have much more complicated solutions than just flat sloped surfaces, as you may have gathered from the Laplace equation in one dimension. However, the two properties listed for the 1D equation are still valid in two dimensions, and will give us some valuable clues when we try to solve the equation.

1. At the point (x, y) , the scalar potential $\Phi(x, y)$ can be seen as the average of the potentials on equidistant points around (x, y) , i.e., a circle C of radius r :

$$\Phi(x, y) = \frac{1}{2\pi r} \oint_C \Phi dl \quad \text{for any } r \text{ in the domain of } \Phi.$$

An example of such a circle is given in Fig. 16.

2. There are again no local maxima or minima (as this would invalidate point 1), and there are no points where a charge can sit in an equilibrium, stable or unstable.

A consequence of point 1 (the averaging method) is that the surface defined by $\Phi(x, y)$ is pretty featureless. A physical picture of what comes close to $\Phi(x, y)$ is that of an uneven rim with a rubber membrane stretched over it. The surface of the membrane will generally be curved because of the uneven height of the rim (see for example Fig 16), but it also wants to be as flat as possible.

The Laplace equation in three dimensions

In three dimensions it is impossible to draw intuitive pictures such as Fig 16, but the two listed properties still remain true:

1. At the point \mathbf{r} , the scalar potential $\Phi(\mathbf{r})$ can be seen as the average of the potentials on equidistant points around \mathbf{r} , i.e., a sphere S of radius r :

$$\Phi(\mathbf{r}) = \frac{1}{4\pi r^2} \oint_S \Phi dA \quad \text{for any } r \text{ in the domain of } \Phi.$$

2. There are again no local maxima or minima (as this would invalidate point 1), and there are no points where a charge can sit in an equilibrium, stable or unstable.

For the proof of point 1, see Griffiths (2008), page 114.

Boundary conditions for Laplace's equation

Every differential equation needs boundary conditions to pick out the relevant solution to your problem. But how many boundary conditions are sufficient to determine the solution, and when do we specify too much? This is an important question, because specifying too many boundary conditions will likely result in contradictory requirements, and you won't be able to find a solution at all (not even the trivial one, $\Phi = 0$). To help determine what are necessary and sufficient boundary conditions, we state two "uniqueness" theorems:

First Uniqueness Theorem: The Laplace equation for Φ in a volume V with surface S has a unique solution when Φ is given on the boundary S .

This theorem requires knowledge of the potential at the boundary. However, in practice we may be able to specify only the charges on various conducting surfaces. In this case the second uniqueness theorem helps:

Second Uniqueness Theorem: The electric field in a volume V is uniquely specified by the *total* charge on each conductor surrounding V and the charge density inside V .

For the proofs of these theorems, see Griffiths (2008), pages 116–121.

4.2 Charge distributions near conducting surfaces

The uniqueness theorems have an interesting consequence, which we will explore by considering the following example. Consider a grounded conducting plate parallel to the xy -plane at $z = 0$, and a point charge q a distance d above the plane. Since the conducting plate is grounded, the potential Φ is zero. We can consider this a boundary condition for working out the potential above the plate: $\Phi = 0$ for $z = 0$. At infinity the potential also drops off to zero, so we have a completely specified the potential on the boundary of the volume we are interested in (the half-space of the positive z -axis). The first uniqueness theorem then tells us that we can uniquely determine Φ in this half-space.

However, since the boundary potential determines Φ , *any* physical situation that produces these potentials at the boundary will necessarily produce the right potential in the volume of interest. We can consider the different physical situation where the conducting plate is replaced with a point charge $-q$ on the z -axis at the position $-d$. We can superpose the potentials of the two charges and obtain

$$\Phi(x, y, z) = \frac{1}{4\pi\epsilon_0} \left(\frac{q}{\sqrt{x^2 + y^2 + (z - d)^2}} - \frac{q}{\sqrt{x^2 + y^2 + (z + d)^2}} \right). \quad (4.6)$$

It is easy to see that $\Phi = 0$ for $z = 0$, and Φ vanishes at infinity. So this situation gives the same boundary conditions as the charge above a conductor. By the first uniqueness theorem, the scalar potential in Eq. (4.6) is the correct description of the potential above the conducting plate. This is called the *method of images*, since the conductor effectively acts as a mirror. The image charge must have the opposite sign to the original charge.

It also means that the charge is attracted to the plate with a force

$$\mathbf{F} = -\frac{q^2 \hat{\mathbf{k}}}{4\pi\epsilon_0 (2d)^2}. \quad (4.7)$$

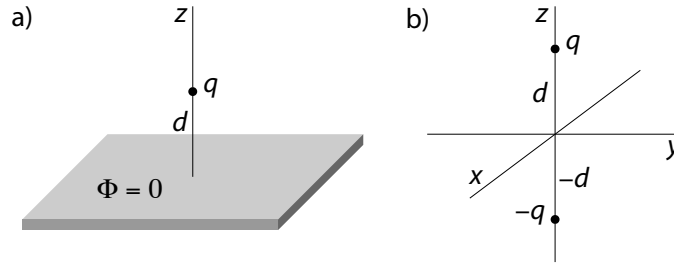


Figure 17: The method of images.

In turn, this must be due to negative charges in the conductor rushing towards the point closest to the point charge. This is the *induced* surface charge σ . The electric field at the surface of the conductor is perpendicular to the surface, and from Gauss theorem we have

$$\mathbf{E} = \frac{\sigma}{\epsilon_0} \hat{\mathbf{k}} \quad \text{at } z = 0. \quad (4.8)$$

From this we find that the induced surface charge is

$$\sigma(x, y) = -\epsilon_0 \left. \frac{\partial \Phi}{\partial z} \right|_{z=0} = -\frac{qd}{2\pi(x^2 + y^2 + d^2)^{3/2}}. \quad (4.9)$$

When we integrate over the surface of the conductor, we find that the total induced charge is $-q$. Did you expect this?

Exercise 4.1: Derive Eq. (4.7) from the Coulomb force and by calculating the field directly from Φ .

4.3 Separation of variables

One of the most powerful methods for solving the Laplace equation is separation of variables, in which we *assume* that Φ is a product of functions that only depend on x and y , respectively. For example, in two dimensions we can choose

$$\Phi(x, y) = X(x)Y(y). \quad (4.10)$$

Ignoring complications where $X(x) = 0$ or $Y(y) = 0$, we can divide the Laplace equation by $X(x)Y(y)$ to obtain

$$\frac{1}{X} \frac{d^2 X}{dx^2} + \frac{1}{Y} \frac{d^2 Y}{dy^2} = 0. \quad (4.11)$$

Note that these are no longer partial derivatives. For this equation to hold, both terms must be proportional to a constant, C_1 and C_2 with $C_1 + C_2 = 0$:

$$\frac{1}{X} \frac{d^2 X}{dx^2} = C_1 \quad \text{and} \quad \frac{1}{Y} \frac{d^2 Y}{dy^2} = C_2. \quad (4.12)$$

If either term were to depend on their corresponding variable x or y we could keep one constant fixed while changing the other, thereby breaking the requirement $C_1 + C_2 = 0$. The Laplace equation in two dimensions then becomes two separate linear differential equations

$$\frac{d^2 X}{dx^2} = k^2 X \quad \text{and} \quad \frac{d^2 Y}{dy^2} = -k^2 Y, \quad (4.13)$$

where we chose $C_1 = -C_2 \equiv k^2$. The solutions to these equations are

$$X(x) = Ae^{kx} + Be^{-kx} \quad \text{and} \quad Y(y) = C \sin ky + D \cos ky, \quad (4.14)$$

and the potential becomes

$$\Phi_k(x, y) = X(x)Y(y) = (Ae^{kx} + Be^{-kx})(C \sin ky + D \cos ky), \quad (4.15)$$

for some choice of k . Since the Laplace equation is linear, we can choose any superposition of Φ_k with different values of k to fit the boundary conditions.

In three dimensions, the separation of variables method assumes that $\Phi(x, y, z) = X(x)Y(y)Z(z)$, which leads to

$$\frac{1}{X} \frac{d^2 X}{dx^2} + \frac{1}{Y} \frac{d^2 Y}{dy^2} + \frac{1}{Z} \frac{d^2 Z}{dz^2} = 0. \quad (4.16)$$

We then require that each term is constant, and $C_x + C_y + C_z = 0$.

The separation of variables does not have to happen in Cartesian coordinates; we can use spherical coordinates as well. These are often better suited to our problem (e.g., when we have a spherical or cylindrical symmetry in our system). The Laplace equation in spherical coordinates is given by

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \Phi}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Phi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \Phi}{\partial \phi^2} = 0. \quad (4.17)$$

For the special case of cylindrical symmetry Φ does not depend on ϕ , and we have effectively a two-variable equation. Using separation of variables and setting $\Phi(r, \theta) = R(r)\Theta(\theta)$ then yields

$$\frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{1}{\Theta \sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) = 0. \quad (4.18)$$

This leads to two ordinary differential equations

$$\frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) = l(l+1)R \quad \text{and} \quad \frac{1}{\sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) = -l(l+1)\Theta, \quad (4.19)$$

where the constants $\pm l(l+1)$ are chosen such that the solutions to the differential equation for R are given by

$$R(r) = Ar^l + \frac{B}{r^{l+1}}, \quad (4.20)$$

and the solutions to the differential equation for Θ are given by the Legendre polynomials

$$\Theta(\theta) = P_l(\cos \theta). \quad (4.21)$$

These polynomials are defined by *Rodrigues'* formula

$$P_l(x) = \frac{1}{2^l l!} \left(\frac{d}{dx} \right)^l (x^2 - 1)^l. \quad (4.22)$$

The first few Legendre polynomials are

$$P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = \frac{1}{2}(3x^2 - 1), \quad P_3(x) = \frac{1}{2}(5x^3 - 3x). \quad (4.23)$$

These polynomials are orthogonal in the sense that

$$\int_{-1}^1 P_l(x) P_{l'}(x) dx = \begin{cases} 0 & \text{if } l' \neq l, \\ \frac{2}{2l+1} & \text{if } l' = l. \end{cases} \quad (4.24)$$

Exercise 4.2: Derive a formula for $P_{l+1}(x)$ given $P_l(x)$, $P_{l-1}(x)$, $P_{l-2}(x)$, \dots

The general solution for the scalar potential in cylindrically symmetric problems is now given by

$$\Phi(r, \theta) = \frac{1}{\epsilon_0} \sum_{l=0}^{\infty} \left(A_l r^l + \frac{B_l}{r^{l+1}} \right) P_l(\cos \theta). \quad (4.25)$$

As an example, we calculate the potential inside a hollow sphere of radius R whose shell is held at a potential $\Phi_0(\theta)$ that may depend on the angle θ . First, we note that $B_l = 0$ for all l , otherwise the potential becomes singular at $r = 0$, the centre of the sphere. We therefore have

$$\Phi(r, \theta) = \frac{1}{\epsilon_0} \sum_{l=0}^{\infty} A_l r^l P_l(\cos \theta). \quad (4.26)$$

We can see that if the potential is constant only the $l = 0$ term contributes, for if we include $l = 1$ the $\cos \theta$ term must be cancelled by an $l = 3$ term, which in turn introduces a higher power in $\cos \theta$, and so on. This means that the potential throughout the sphere is constant. In general, we can calculate the coefficients using the orthogonality relation for the Legendre polynomials, and we find

$$A_l = \frac{2l+1}{2R^l} \int_0^\pi \Phi_0(\theta) P_l(\cos \theta) \sin \theta d\theta. \quad (4.27)$$

You should prove this relationship. When we choose $\Phi_0(\theta) = \Psi \sin^2(\theta/2)$ for some real constant Ψ , we can write this in terms of Legendre polynomials as

$$\Phi_0(\theta) = \Psi \sin^2(\theta/2) = \frac{1}{2}(1 - \cos \theta) = \frac{\Psi}{2} [P_0(\cos \theta) - P_1(\cos \theta)]. \quad (4.28)$$

We then find that $A_0 = \Psi/2$ and $A_1 = -\Psi/2R$.

Exercise 4.3: Let $\Phi_0(\theta) = +\Psi$ for $0 \leq \theta < \pi/2$ and $\Phi_0(\theta) = -\Psi$ for $\pi/2 \leq \theta \leq \pi$. Which l contribute to the potential in Eq. (4.26)?

4.4 Electric dipoles, quadrupoles and higher multipoles

What if we do not have any special symmetries in the problem? In that case the potential will depend on all three coordinates: $\Phi = \Phi(r, \theta, \phi)$. We will now consider the potential of an arbitrary localised charge distribution. Moreover, we are reasonably far away from the charge distribution, so we can develop Φ in powers of $1/r$. From the cylindrically symmetric potential in Eq. (4.25) you see that all $A_l = 0$, otherwise the potential would blow up towards infinity (what about A_0 ?). We now include the ϕ dependence, so the potential can be written as

$$\Phi(r, \theta, \phi) = \frac{1}{\epsilon_0} \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} \frac{Q_{lm}}{2l+1} \frac{Y_{lm}(\theta, \phi)}{r^{l+1}}, \quad (4.29)$$

where the $Y_{lm}(\theta, \phi)$ are the *spherical harmonics*. These are defined in terms of the *generalised* Legendre polynomials $P_l^m(\cos \theta)$,

$$Y_{lm}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos \theta) e^{im\phi} \quad \text{for } m \geq 0, \quad (4.30)$$

and $Y_{l,-m}(\theta, \phi) = (-1)^m Y_{lm}(\theta, \phi)$, with

$$P_l^m(x) = \frac{(-1)^m}{2^l l!} (1-x^2)^{m/2} \left(\frac{d}{dx} \right)^{l+m} (x^2-1)^l. \quad (4.31)$$

The spherical harmonics also obey an orthogonality relation:

$$\int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta Y_{lm}(\theta, \phi) Y_{l'm'}^*(\theta, \phi) = \delta_{ll'} \delta_{mm'}, \quad (4.32)$$

which will come in handy next, when we want to calculate the coefficients Q_{lm} in Eq. (4.29). The spherical harmonics are also *complete*, in the sense that any potential Φ outside the charge distribution can be written in terms of these polynomials, without needing extra functions of θ and ϕ . The first few spherical harmonics are given by

$$\begin{aligned} Y_{00}(\theta, \phi) &= \frac{1}{\sqrt{4\pi}} \\ Y_{10}(\theta, \phi) &= \sqrt{\frac{3}{4\pi}} \cos \theta = \sqrt{\frac{3}{4\pi}} \frac{z}{r} \\ Y_{1,\pm 1}(\theta, \phi) &= \mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\phi} = \mp \sqrt{\frac{3}{8\pi}} \frac{x \pm iy}{r} \\ Y_{20}(\theta, \phi) &= \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1) = \sqrt{\frac{5}{16\pi}} \frac{3z^2 - r^2}{r^2} \\ Y_{2,\pm 1}(\theta, \phi) &= \mp \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{\pm i\phi} = \mp \sqrt{\frac{15}{8\pi}} \frac{z(x \pm iy)}{r^2} \\ Y_{2,\pm 2}(\theta, \phi) &= \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{\pm 2i\phi} = \sqrt{\frac{15}{32\pi}} \frac{(x \pm iy)^2}{r^2} \\ Y_{30}(\theta, \phi) &= \sqrt{\frac{7}{16\pi}} (5 \cos^3 \theta - 3 \cos \theta) = \sqrt{\frac{7}{16\pi}} \frac{5z^3 - 3zr^2}{r^3}. \end{aligned} \quad (4.33)$$

You see that these are *complex* polynomials.

Exercise 4.4: Verify the $Y_{lm}(\theta, \phi)$ in Eq. (4.33).

You know from the first year that the potential at a point $\mathbf{r} = (r, \theta, \phi)$ due to a point charge q at the origin is given by

$$\Phi(r, \theta, \phi) = \frac{q}{4\pi\epsilon_0 r}. \quad (4.34)$$

If the charge is not at the origin, but instead at a point $\mathbf{r}' = (r', \theta', \phi')$, we have to replace r in the expression for Φ with the distance between the charge and the point where we evaluate the

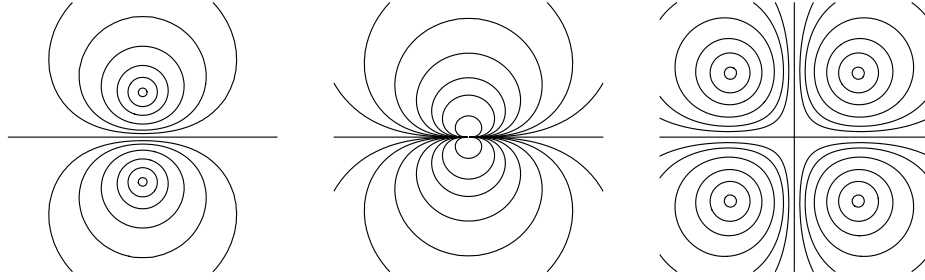


Figure 18: The potential of a physical dipole, a “perfect” dipole and a quadrupole.

potential: $r \rightarrow |\mathbf{r} - \mathbf{r}'|$ In addition, we can think of the point charge at \mathbf{r}' as the charge density $\rho(\mathbf{r}')$ at \mathbf{r}' multiplied by an infinitesimal volume around \mathbf{r}' , denoted by $d\mathbf{r}'$. The superposition principle will now let us sum over all charges in those infinitesimal volumes (which is an integral over all space), and we obtain

$$\Phi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}', \quad (4.35)$$

where the integral over space is given by

$$\int d\mathbf{r}' = \int_0^\infty r'^2 dr' \int_0^\pi \sin\theta' d\theta' \int_0^{2\pi} d\phi', \quad (4.36)$$

and $\rho(\mathbf{r}')$ drops off quickly as $r' \rightarrow \infty$. We now have two completely general expressions for the potential outside the charge distribution, namely Eqs. (4.35) and (4.29), so we can compare the two. In particular, we can solve for the Q_{lm} by multiplying both versions of Φ with $Y_{lm}^*(\theta, \phi)$ and integrate of θ and ϕ . This yields

$$Q_{lm} = \int Y_{lm}^*(\theta, \phi) r^l \rho(\mathbf{r}) d\mathbf{r}. \quad (4.37)$$

These are the so-called *multipole moments* of the charge distribution. The first couple of multipole moments are the *monopole*, the *dipole*, and the *quadrupole*. You may think that the dipole moment Q_{10} is all you need to describe the physical dipole $p = qd$ of two equal and opposite charges q a distance d apart. However, this is not true. The dipole moment here describes a “perfect” dipole, where $d \rightarrow 0$ and $q \rightarrow \infty$ while keeping p constant (see Fig. 18). A similar argument holds for higher-order multipoles.

Let us calculate the Q_{lm} for a physical dipole with a positive point charge q at the position $(0, 0, d/2)$ and a negative point charge $-q$ at the position $(0, 0, -d/2)$. The charge density for such a configuration is given by

$$\rho(x, y, z) = q \delta(x) \delta(y) \delta\left(z - \frac{d}{2}\right) - q \delta(x) \delta(y) \delta\left(z + \frac{d}{2}\right), \quad (4.38)$$

where $\delta(x)$, etc., are the Dirac delta functions. We substitute this charge density into the expression for Q_{lm} , and we find that

$$Q_{00} = \int Y_{00}^*(\theta, \phi) \rho(x, y, z) d\mathbf{r} = \frac{1}{\sqrt{4\pi}} \int \rho(x, y, z) d\mathbf{r} = 0, \quad (4.39)$$

since the integral over ρ gives us the *total* charge, which is $q - q = 0$. If the two charges are unequal, you see that the potential has a net monopole contribution.

The dipole moments are calculated as

$$Q_{10} = \int Y_{10}^*(\theta, \phi) r \rho(x, y, z) d\mathbf{r} = q \sqrt{\frac{3}{4\pi}} \int z \left[\delta\left(z - \frac{d}{2}\right) - \delta\left(z + \frac{d}{2}\right) \right] dz = \sqrt{\frac{3}{4\pi}} qd, \quad (4.40)$$

and

$$Q_{11} = \int Y_{11}^*(\theta, \phi) r \rho(x, y, z) d\mathbf{r} = -\sqrt{\frac{3}{8\pi}} \int (x + iy) [q\delta(x)\delta(y) - q\delta(x)\delta(y)] dx dy = 0. \quad (4.41)$$

A perfect dipole would be completely characterised by Q_{10} and Q_{11} , but a physical dipole (with finite d) has higher order moments as well. You can show that the quadrupole moments are all zero because the Y_{lm} are either even powers in z or contain x and y , but Q_{30} will not:

$$\begin{aligned} Q_{30} &= \sqrt{\frac{7}{16\pi}} \int (5z^3 - 3zr^2) q\delta(x)\delta(y) \left[\delta\left(z - \frac{d}{2}\right) - \delta\left(z + \frac{d}{2}\right) \right] dx dy dz \\ &= \sqrt{\frac{7}{16\pi}} q \int 2z^3 \left[\delta\left(z - \frac{d}{2}\right) - \delta\left(z + \frac{d}{2}\right) \right] dz = \sqrt{\frac{7}{64\pi}} qd^3. \end{aligned} \quad (4.42)$$

When $d \rightarrow 0$ for a perfect dipole, you see that $Q_{30} \rightarrow 0$, as expected.

Exercise 4.5: Construct an octupole (8 charges in a cubic arrangement) to cancel the Q_{30} of the physical dipole, and make it more like a pure dipole.

4.5 Magnetic multipoles

Analogous to the scalar potential Φ in Eq. (4.35), we can express the vector potential \mathbf{A} in terms of the current distribution \mathbf{J} . At the point \mathbf{r} the vector potential $\mathbf{A}(\mathbf{r})$ is proportional to the current distribution $\mathbf{J}(\mathbf{r}')$ at point \mathbf{r}' , and the strength is inverse proportional to the distance between the two points $|\mathbf{r} - \mathbf{r}'|$. The total vector potential is then the superposition of all these contributions:

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int \frac{\mathbf{J}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}', \quad (4.43)$$

and the magnetic field is obtained as usual via $\mathbf{B} = \nabla \times \mathbf{A}$. To find the multipole expansion of \mathbf{A} we assume that the current density is nonzero in a small region so that $\mathbf{r} \gg \mathbf{r}'$. Generally, the function $1/|\mathbf{r} - \mathbf{r}'|$ can be expressed in terms of Legendre polynomials and spherical harmonics (to all orders, and with $r' < r$) as

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \sum_{l=0}^{\infty} \frac{r'^l}{r^{l+1}} P_l(\cos \gamma) \quad (4.44)$$

$$= 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^l \frac{1}{2l+1} \frac{r'^l}{r^{l+1}} Y_{lm}^*(\theta', \phi') Y_{lm}(\theta, \phi), \quad (4.45)$$

where $\mathbf{r} = (r, \theta, \phi)$ and $\mathbf{r}' = (r', \theta', \phi')$, and the angle γ between \mathbf{r} and \mathbf{r}' is given by

$$\cos \gamma \equiv \cos \theta \cos \theta' + \sin \theta \sin \theta' \cos(\phi - \phi'). \quad (4.46)$$

Since $1/|\mathbf{r} - \mathbf{r}'|$ is such an important function in electrodynamics, Eqs. (4.44) and (4.45) are extremely useful expansions of this function.

Exercise 4.6: Show that $|\mathbf{r} - \mathbf{r}'| = \sqrt{r^2 + r'^2 - 2rr' \cos \gamma}$.

The general expansion of the vector potential in spherical harmonics is obtained by substituting Eq. (4.45) into Eq. (4.43)

$$\begin{aligned} \mathbf{A}(\mathbf{r}) &= \frac{\mu_0}{4\pi} \int \frac{\mathbf{J}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \\ &= \mu_0 \int \left[\sum_{l=0}^{\infty} \sum_{m=-l}^l Y_{lm}^*(\theta', \phi') \frac{1}{2l+1} \frac{r'^l}{r^{l+1}} Y_{lm}(\theta, \phi) \right] \mathbf{J}(\mathbf{r}') d\mathbf{r}'. \end{aligned} \quad (4.47)$$

We can collect all the terms with dependence on r' , θ' and ϕ' into a single integral over \mathbf{r}' , which becomes a vector \mathbf{M}_{lm} depending on l and m :

$$\mathbf{M}_{lm} = \int \left[Y_{lm}^*(\theta', \phi') r'^l \mathbf{J}(r', \theta', \phi') \right] r'^2 \sin \theta' dr' d\theta' d\phi', \quad (4.48)$$

where we included the Jacobian $r'^2 \sin \theta'$ for integration over spherical coordinates. These so-called *magnetic moments* play the same role as the Q_{lm} for the scalar potential in Eq. (4.37), and the expansion of the vector potential in magnetic moments is given by

$$\mathbf{A}(\mathbf{r}) = \mu_0 \sum_{l=0}^{\infty} \sum_{m=-l}^{+l} \frac{\mathbf{M}_{lm}}{2l+1} \frac{Y_{lm}(\mathbf{r})}{r^{l+1}}. \quad (4.49)$$

Finding $\mathbf{A}(\mathbf{r})$ thus reduces to finding the magnetic moments \mathbf{M}_{lm} , and for $r \gg r'$ generally only the lowest orders of l contribute.

We can substitute specific current densities to calculate the vector potential directly using Eq. (4.43), but this is often a very difficult integral. So instead we calculate the vector potential indirectly via the magnetic moments in Eq. (4.48). Since in physics we always deal with finite precision, this is often all we need. We consider two examples.

As our first example, let the current density given by an infinitely thin circular wire of radius a carrying a current I in the xy -plane:

$$\mathbf{J}(\mathbf{r}) = \frac{I}{a^2} \delta\left(\theta - \frac{1}{2}\pi\right) \delta(r - a) \hat{\phi}, \quad (4.50)$$

We substitute this form of \mathbf{J} into the expression for \mathbf{M}_{lm} to obtain

$$\mathbf{M}_{lm} = \frac{I}{a^2} \int Y_{lm}^*(\theta', \phi') r'^l \delta\left(\theta' - \frac{1}{2}\pi\right) \delta(r' - a) \hat{\phi}' r'^2 \sin \theta' dr' d\theta' d\phi'. \quad (4.51)$$

We immediately evaluate the Dirac delta functions, which leads to

$$\mathbf{M}_{lm} = I a^l \int_0^{2\pi} Y_{lm}^*\left(\frac{1}{2}\pi, \phi'\right) \hat{\phi}' d\phi'. \quad (4.52)$$

We can't go much further with this general solution, so we evaluate the lowest-order moments. For \mathbf{M}_{00} this integral reduces to

$$\mathbf{M}_{00} = \frac{I}{\sqrt{4\pi}} \int_0^{2\pi} \hat{\phi}' d\phi'. \quad (4.53)$$

The unit vector $\hat{\phi}'$ actually depends on ϕ , so we cannot take it out of the integral (remember exercise 1.5 on page 21). Instead, we need to decompose it into $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$ components:

$$\mathbf{M}_{00} = \frac{I}{\sqrt{4\pi}} \int_0^{2\pi} (-\sin\phi' \hat{\mathbf{i}} + \cos\phi' \hat{\mathbf{j}}) d\phi'. \quad (4.54)$$

The unit vectors $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$ are independent of ϕ' and can be taken out of the integral. Now we use

$$\int_0^{2\pi} \cos\phi' d\phi' = 0 \quad \text{and} \quad \int_0^{2\pi} \sin\phi' d\phi' = 0, \quad (4.55)$$

and we find that $\mathbf{M}_{00} = 0$. The circular ring current does not have a monopole. Similarly you can show that $\mathbf{M}_{10} = 0$, and we calculate $\mathbf{M}_{1,\pm 1}$ next:

$$\begin{aligned} \mathbf{M}_{11} &= Ia \int_0^{2\pi} Y_{11}^*\left(\frac{1}{2}\pi, \phi'\right) \hat{\phi} d\phi' = -Ia \sqrt{\frac{3}{8\pi}} \int_0^{2\pi} e^{-i\phi'} \hat{\phi} d\phi' \\ &= -Ia \sqrt{\frac{3}{8\pi}} \int_0^{2\pi} (\cos\phi' - i\sin\phi') (-\sin\phi' \hat{\mathbf{i}} + \cos\phi' \hat{\mathbf{j}}) d\phi' \\ &= -Ia \sqrt{\frac{3}{8\pi}} \int_0^{2\pi} (-\cos\phi' \sin\phi' \hat{\mathbf{i}} + \cos^2\phi' \hat{\mathbf{j}} + i\sin^2\phi' \hat{\mathbf{i}} - i\cos\phi' \sin\phi' \hat{\mathbf{j}}) d\phi' \\ &= -Ia \sqrt{\frac{3}{32\pi}} (i\hat{\mathbf{i}} + \hat{\mathbf{j}}), \end{aligned} \quad (4.56)$$

$$\mathbf{M}_{1,-1} = Ia \sqrt{\frac{3}{32\pi}} (-i\hat{\mathbf{i}} + \hat{\mathbf{j}}). \quad (4.57)$$

Substituting these values of $\mathbf{M}_{1,\pm 1}$ into \mathbf{A} in Eq. (4.49) yields, to lowest order

$$\begin{aligned} \mathbf{A}(\mathbf{r}) &\approx \frac{\mu_0}{3} \left(\mathbf{M}_{11} \frac{Y_{11}(\theta, \phi)}{r^2} + \mathbf{M}_{1,-1} \frac{Y_{1,-1}(\theta, \phi)}{r^2} \right) \\ &= \frac{\mu_0 Ia \sin\theta}{16\pi r^2} \left[e^{i\phi} (i\hat{\mathbf{i}} + \hat{\mathbf{j}}) + e^{-i\phi} (-i\hat{\mathbf{i}} + \hat{\mathbf{j}}) \right] \\ &= \frac{\mu_0 Ia^2 \sin\theta}{8\pi r^2} \hat{\phi}. \end{aligned} \quad (4.58)$$

Since $\mathbf{A} = A_\phi \hat{\phi}$, the curl of \mathbf{A} in spherical coordinates simplifies and the magnetic field becomes

$$\mathbf{B}(r, \theta, \phi) = \frac{1}{r \sin\theta} \left(\frac{\partial A_\phi \sin\theta}{\partial \theta} \right) \hat{r} - \frac{1}{r} \left(\frac{\partial r A_\phi}{\partial r} \right) \hat{\theta} = \frac{\mu_0 Ia^2}{8\pi r^3} (2\cos\theta \hat{r} + \sin\theta \hat{\theta}). \quad (4.59)$$

The vector potential and the field are shown in Fig. 19. The central region is excluded, because our solutions hold only for $r > r'$. To find the field inside the circular loop the fraction of radii in Eq. (4.43) must be modified as

$$\frac{r'^l}{r^{l+1}} \rightarrow \frac{r^l}{r'^{l+1}}. \quad (4.60)$$

The solution is then found by following the exact same steps, but the multipole expansion may not converge as fast as the one in the far field.

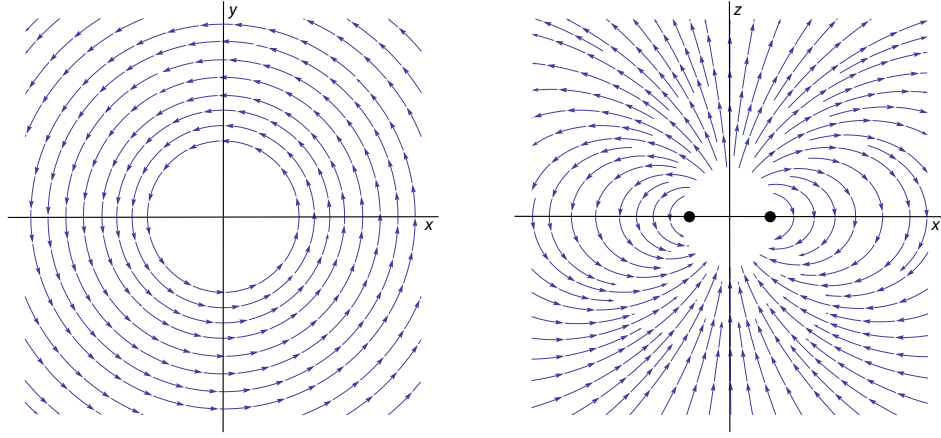


Figure 19: The vector potential \mathbf{A} at $z = 0$ of a circular current loop (left), and the corresponding \mathbf{B} field at $\phi = 0$ (right) to lowest order in l . The region $r < r'$ is excluded, since the solutions are not valid there.

The second example we consider here is the field due to a rotating charged spherical shell of radius a . The charge density σ given by

$$\rho = \frac{Q}{4\pi a^2} \delta(r - a), \quad (4.61)$$

where Q is the total charge on the sphere. You can see that charge is distributed uniformly over the surface, because ρ does not depend on θ and ϕ . The current density can be found from $\mathbf{J} = \rho \boldsymbol{\omega} \times \mathbf{r}$, where $\boldsymbol{\omega}$ is the angular velocity of the sphere, pointing in the z -direction. This can be written as

$$\mathbf{J} = \rho \boldsymbol{\omega} \times \mathbf{r} = \frac{Q\omega}{4\pi a^2} r \sin\theta \delta(r - a) \hat{\phi}. \quad (4.62)$$

We again calculate the magnetic moments \mathbf{M}_{lm} . We see that $\mathbf{M}_{00} = \mathbf{M}_{10} = 0$ due to the absence of ϕ in the corresponding spherical harmonics, so we calculate the $l = 1$ moments with $m = \pm 1$:

$$\begin{aligned} \mathbf{M}_{11} &= \int Y_{11}(\theta, \phi) r \mathbf{J}(r, \theta, \phi) d\mathbf{r} \\ &= -\sqrt{\frac{3}{8\pi}} \frac{Q\omega}{4\pi a^2} \int \sin\theta e^{-i\phi} r^2 \sin\theta \delta(r - a) \hat{\phi} d\mathbf{r} \\ &= -\sqrt{\frac{3}{8\pi}} \frac{Q\omega a^2}{4\pi} \left(\int_0^\pi \sin^3\theta d\theta \right) \left(\int_0^{2\pi} e^{-i\phi} \hat{\phi} d\phi \right) \\ &= -\sqrt{\frac{3}{8\pi}} \frac{Q\omega a^2}{3\pi} \frac{i\hat{\mathbf{i}} + \hat{\mathbf{j}}}{2}, \end{aligned} \quad (4.63)$$

$$\mathbf{M}_{1,-1} = \sqrt{\frac{3}{8\pi}} \frac{Q\omega a^2}{3\pi} \frac{-i\hat{\mathbf{i}} + \hat{\mathbf{j}}}{2}, \quad (4.64)$$

where we solved the integral over ϕ in the same way as in Eqs. (4.56) and (4.57). The vector

potential then becomes

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{3} \left(\mathbf{M}_{11} \frac{Y_{11}(\theta, \phi)}{r^2} + \mathbf{M}_{1,-1} \frac{Y_{1,-1}(\theta, \phi)}{r^2} \right) = \frac{\mu_0 Q \omega a^2}{12\pi^2} \frac{\sin \theta}{r^2} \hat{\phi}, \quad (4.65)$$

and the magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$ is

$$\mathbf{B}(r, \theta, \phi) = \frac{\mu_0 Q \omega a^2}{12\pi^2 r^3} (2 \cos \theta \hat{r} + \sin \theta \hat{\theta}). \quad (4.66)$$

The dependence on r , θ , and ϕ of \mathbf{A} and \mathbf{B} is identical to that of the circular current loop, since we included exactly the same magnetic moments $\mathbf{M}_{1,\pm 1}$. So in order to simulate the field of the rotating spherical shell with a current loop, we require

$$I = \frac{2}{3\pi} Q \omega. \quad (4.67)$$

While it is no surprise that $I \propto Q \omega$, the numerical factor $2/3\pi$ is much harder to guess.

Summary

In this section, we studied electrostatic and magnetostatic problems. We presented a way to find the solutions to the Laplace equation using separation of variables, and used some of the insights from that to develop the method of images. We gave polynomial of the scalar and vector potential in terms of Legendre polynomials and spherical harmonics. The first few terms of these expansions can be used to approximate the potential at large distances.

You should master the following techniques:

1. use the method of images to find the electric potential Φ for charges near grounded conductors;
2. apply the separation of variables to the Laplace equation, and use the boundary conditions to find the solution for Φ .
3. calculate the electric and magnetic multipole moments Q_{lm} and \mathbf{M}_{lm} for Φ and \mathbf{A} , respectively.

5 Solving Maxwell's Equations: Electromagnetic Waves

In the previous section we explored solutions of Maxwell's equations in electro- and magnetostatic situations. Here we will take into account the situation of time-varying fields.

5.1 Maxwell's equations in vacuum: the wave equation

In field theories we can often find nontrivial solutions to the field equations (meaning that the field is not zero everywhere), even when there are no sources present. For electrodynamics, in the absence of charges and currents we can still have interesting effects. So interesting in fact, that it covers a whole discipline in physics, namely optics.

We obtain Maxwell's equations in vacuum by setting $\rho = \mathbf{J} = 0$ in Eq. (2.30):

$$\nabla \cdot \mathbf{B} = 0 \quad \text{and} \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (5.1)$$

$$\nabla \cdot \mathbf{E} = 0 \quad \text{and} \quad \nabla \times \mathbf{B} = \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}. \quad (5.2)$$

Taking the curl of $\nabla \times \mathbf{E}$ and substituting $\nabla \times \mathbf{B} = \mu_0 \epsilon_0 \partial_t \mathbf{E}$ we find

$$\nabla \times \left(\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} \right) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} + \frac{\partial}{\partial t} \left(\mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) = 0, \quad (5.3)$$

which, with $\nabla \cdot \mathbf{E} = 0$, leads to the wave equation

$$\left(\nabla^2 - \mu_0 \epsilon_0 \frac{\partial^2}{\partial t^2} \right) \mathbf{E}(\mathbf{r}, t) = 0. \quad (5.4)$$

This differential equation has the well-known real and complex plane wave solutions

$$\mathbf{E} = \mathbf{E}_A \cos(\mathbf{k} \cdot \mathbf{r} - \omega t) + \mathbf{E}_B \sin(\mathbf{k} \cdot \mathbf{r} - \omega t) \quad \text{and} \quad \mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}, \quad (5.5)$$

where \mathbf{E}_A , \mathbf{E}_B , and \mathbf{E}_0 are constant vectors, \mathbf{k} is the wave vector pointing in the propagation direction, and ω is the frequency. By virtue of Eq. (5.4), the wave vector and frequency obey the relation ($k^2 = \mathbf{k} \cdot \mathbf{k}$):

$$k^2 - \mu_0 \epsilon_0 \omega^2 = 0. \quad (5.6)$$

Clearly, in SI units (see Appendix B) $\mu_0 \epsilon_0$ has the dimension of inverse velocity-squared, and substituting the numerical values for μ_0 and ϵ_0 reveals that this velocity is c , the speed of light. Another great triumph of Maxwell's theory was the identification of light as electromagnetic waves. For the right frequencies, the above equation is therefore the dispersion relation of light propagating through vacuum:

$$k^2 = \frac{\omega^2}{c^2}. \quad (5.7)$$

So not only does Maxwell's theory unify electric and magnetic phenomena, it also encompasses the whole of classical optics. On top of that, it predicts a range of new types of radiation from radio waves and infrared in the long wavelengths, to ultraviolet, X-rays (or Röntgen rays), and gamma rays in the short wavelength. Prior to Maxwell's discovery of electromagnetic waves, the known types of radiation other than light were infrared, discovered by William Herschel in 1800, and ultraviolet, discovered in 1801 by Johann Wilhelm Ritter. After Maxwell, Heinrich Hertz

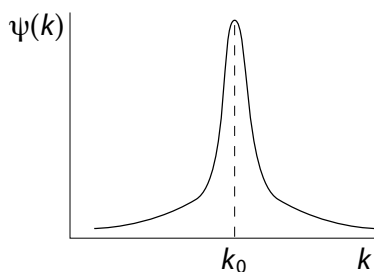


Figure 20: The amplitude $\psi(k)$ of the components in the wave packet $E(x, t)$.

discovered radiowaves and microwaves in 1887 and 1888, respectively. He also made important contributions to the understanding of electromagnetic waves. Wilhelm Conrad Röntgen discovered X-rays in 1895, and Paul Ulrich Villard discovered gamma rays in 1900. However, it was not until the work of Ernest Rutherford and Edward Andrade in 1914 that gamma rays were understood as electromagnetic waves.

Similarly, Maxwell's equations in vacuum give rise to a wave equation for the magnetic field, yielding the real and complex solutions

$$\mathbf{B} = \mathbf{B}_A \cos(\mathbf{k} \cdot \mathbf{r} - \omega t) + \mathbf{B}_B \sin(\mathbf{k} \cdot \mathbf{r} - \omega t) \quad \text{and} \quad \mathbf{B} = \mathbf{B}_0 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}, \quad (5.8)$$

Since $\nabla \cdot \mathbf{E} = \nabla \cdot \mathbf{B} = 0$, the electric and magnetic fields are perpendicular to the direction of propagation, \mathbf{k} . Electromagnetic waves are thus *transverse* waves. It is left as a homework question to prove that $\mathbf{E} \perp \mathbf{B}$.

Exercise 5.1: Derive the wave equation for \mathbf{B} by taking the curl of $\nabla \times \mathbf{B}$.

You may have noticed that we also gave a *complex* solution to the wave equation in Eq. (5.5). However, the electric field must be a real quantity, otherwise we would get complex forces, momenta, velocities, etc. This is clearly nonsense. The way out of this is that we take the *real* part of Eq. (5.5) as the physical solution, and we discard the imaginary part. We may calculate with the complex solution because it is often easier to deal with exponentials than with trigonometric functions.

5.2 The dispersion relation and group velocity

The dispersion relation for a wave with definite frequency ω and wave number k defines a propagation velocity. In vacuum, this is $c = 299792458 \text{ ms}^{-1}$. In simple dielectric media, ϵ_0 and μ_0 are modified to $\epsilon = \epsilon_r \epsilon_0$ and $\mu = \mu_r \mu_0$, where $\epsilon_r, \mu_r > 1$. This has the effect that the speed of the wave with frequency ω becomes smaller. The propagation velocity of a (plane) wave with a single frequency is called the *phase velocity*.

In many materials, the phase velocity is different for different frequencies. Suppose that we create a *wave packet* of light by superposing plane waves of different frequencies. For simplicity we consider a wave in one dimension. The wave packet is given by

$$E(x, t) = E_0 \int dk \psi(k) e^{ikx - i\omega(k)t}, \quad (5.9)$$

where the function $\psi(k)$ gives the amplitude of the plane wave with wave vector k . At $t = 0$ the peak of the wave function is positioned at $x = 0$.

Exercise 5.2: Show that $E(x, t)$ in Eq. (5.9) satisfies the wave equation.

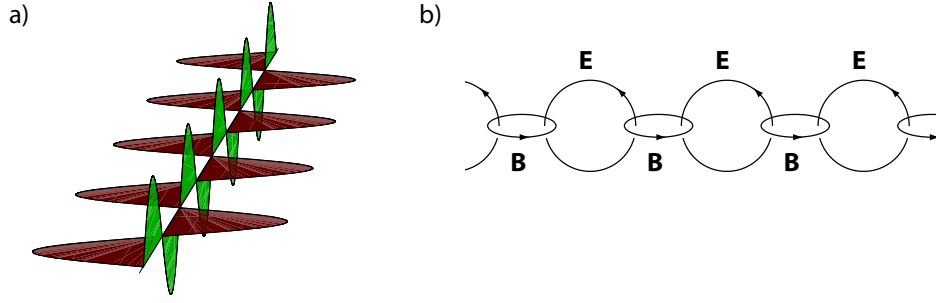


Figure 21: An electromagnetic wave with linear polarization. a) The \mathbf{E} field is pointed in the x direction, while the \mathbf{B} field points in the y direction. The orientation of the \mathbf{E} and \mathbf{B} fields determine the polarization of the wave. b) The field lines are closed and form interlocking loops.

Suppose that $\psi(k)$ is smooth and has a single peak at the value k_0 , as shown in Fig. 20. We can then develop $\omega(k)$ into a first-order Taylor expansion around k_0 :

$$\omega(k) = \omega(k_0) + (k - k_0) \left. \frac{d\omega}{dk} \right|_{k=k_0}, \quad (5.10)$$

which gives to a good approximation

$$E(x, t) \simeq E_0 e^{ik_0 v_g t - i\omega_0 t} \int dk \psi(k) e^{ik(x - v_g t)}, \quad (5.11)$$

where $\omega_0 \equiv \omega(k_0)$, and we defined

$$v_g \equiv \left. \frac{d\omega}{dk} \right|_{k=k_0} \quad (5.12)$$

as the *group velocity*. The reason for this name is that the peak at time t will be concentrated near $x = v_g t$, so the velocity of the wave packet (a *group* of plane waves) is given by v_g . In free space the group velocity is the same as the phase velocity c . By contrast, waves in general dielectrics typically have a frequency dependency in the phase velocity, and as a result the group velocity can be different from the phase velocity. In specially prepared (absorptive) materials, the group velocity can even be greater than the speed of light in vacuum. However, the speed of the actual *signal* will never exceed c . The higher-order terms in the Taylor expansion of $\omega(k)$ that are not included in Eq. (5.10) cause a *distortion* of the wave packet, and the approximation here works only when the second derivative is small compared to the first derivative. The second derivative of $\omega(k)$ is called the *group velocity dispersion*.

5.3 Polarization of electromagnetic waves

The electric field \mathbf{E} is a vector, and a natural question is: what is the direction of \mathbf{E} in an electromagnetic wave? First of all, we have seen that \mathbf{E} must be perpendicular to the direction of propagation \mathbf{k} . That still leaves two transverse directions. Suppose (without loss of generality) that the wave vector is oriented in the z direction: $\mathbf{k} = (0, 0, k)$. The electric (and magnetic) field then has two components, in the x - and y -directions. The direction of the \mathbf{E} field is called the *polarization*.

The field must satisfy the wave equation *in each direction*, so we have (for a monochromatic wave with frequency $\omega = ck$)

$$\mathbf{E} = (E_x e^{i\mathbf{k}\cdot\mathbf{r}-i\omega t}, E_y e^{i\mathbf{k}\cdot\mathbf{r}-i\omega t+i\phi}, 0), \quad (5.13)$$

and a similar expression for the magnetic field. Notice the relative phase ϕ in the y -component: When $\phi = 0$ or π radians (that is zero degrees or 180 degrees), the expression reduces to $\mathbf{E} = (E_x, \pm E_y, 0)e^{i\mathbf{k}\cdot\mathbf{r}-i\omega t}$, which is an oscillating wave in the direction $(E_x, \pm E_y, 0)$. This is called *linear* polarization. When the relative phase ϕ is not a multiple of π , we speak of *elliptical* polarization. In the special case of $\phi = \pi/2$ or $3\pi/2$, the polarization is called *circular* (and the two possible phases correspond to left-handed and right-handed polarization).

To see what this means physically, consider circularly polarized light (and $E_x = E_y = E$):

$$\mathbf{E}(\mathbf{r}, t) = E(\cos(\mathbf{k} \cdot \mathbf{r} - \omega t), \mp \sin(\mathbf{k} \cdot \mathbf{r} - \omega t)), \quad (5.14)$$

where we have taken the real (physical) part of the field. You may recognize this as a typical rotated vector $(\cos\theta, \sin\theta)$ in a plane. If we keep \mathbf{r} fixed, the \mathbf{E} field therefore rotates as time passes. The sign in E_y determines whether the rotation is clockwise or counter-clockwise. Similarly, we can keep t fixed and vary \mathbf{r} along the direction of propagation. This shows that \mathbf{E} traces out a corkscrew motion in space.

A couple of things are worth mentioning: There is a difference between changing the values of E_x and E_y , and changing the relative phase ϕ . For waves with linear polarization we can always find a reference frame in which either E_x or E_y are zero, but this is not possible for elliptical polarization. Second, the polarization is determined completely by the \mathbf{E} field (we do not rely on the specific form of the \mathbf{B} field at all). Nevertheless, the \mathbf{B} field must be perpendicular to both \mathbf{E} and \mathbf{k} . In addition, the fields must satisfy Maxwell's equations, which means that for plane waves

$$\mathbf{B} = \frac{1}{c} \frac{\mathbf{k} \times \mathbf{E}}{k}. \quad (5.15)$$

We could have derived the properties of polarization from the \mathbf{B} field, however. In that case the \mathbf{E} field would follow from the expression of the \mathbf{B} field.

5.4 The wave equation with sources

So far we have set $\rho = \mathbf{J} = 0$, but in general this is not the case. Following the steps at the beginning of this section with arbitrary ρ and \mathbf{J} gives a new set of equations. We start again with

$$\nabla \times \left(\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} \right) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} + \frac{\partial}{\partial t} \left(\mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \mathbf{J} \right) = 0. \quad (5.16)$$

Substituting $\nabla \cdot \mathbf{E} = \rho/\epsilon_0$, this leads to

$$\left(\nabla^2 - \mu_0 \epsilon_0 \frac{\partial^2}{\partial t^2} \right) \mathbf{E} = \frac{1}{\epsilon_0} \nabla \rho + \mu_0 \frac{\partial \mathbf{J}}{\partial t}. \quad (5.17)$$

This is a second-order differential equation for \mathbf{E} , and does not depend on the magnetic field \mathbf{B} . Can you reduce this to the Poisson equation for the electrostatic case? (Hint: all the time derivatives will be zero.)

Similarly, we can take the curl of $\nabla \times \mathbf{B}$, which gives

$$\nabla \times \left(\nabla \times \mathbf{B} - \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} - \mu_0 \mathbf{J} \right) = \nabla(\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B} + \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{B}}{\partial t^2} + \mu_0 \nabla \times \mathbf{J} = 0. \quad (5.18)$$

Using $\nabla \cdot \mathbf{B} = 0$ this gives rise to the second-order differential in \mathbf{B} :

$$\left(\nabla^2 - \mu_0 \epsilon_0 \frac{\partial^2}{\partial t^2} \right) \mathbf{B} = \mu_0 \nabla \times \mathbf{J} . \quad (5.19)$$

This differential equation does not contain the charge density ρ , since static charges do not generate magnetic fields, and we notice that the magnetic field depends on the curl of \mathbf{J} , as expected.

The two differential equations (5.17) and (5.19) are equivalent to the four linear first-order Maxwell equations, but these have the advantage that we have separate equations for the electric and magnetic fields. Being second order equations, they need an extra boundary condition to specify a particular solution.

Summary

In this section, we solved Maxwell's equations for time-varying (dynamical) problems. We derived the wave equation, which is a linear (second order) differential equation for \mathbf{E} and \mathbf{B} that allows for superpositions of the solutions. This leads to the possibility of wave packets with a range of frequencies, and these solutions obey a dispersion relation. In turn, this can be used to define the phase and group velocities of waves. Since the electric and magnetic fields are vectors the EM waves have polarisation, which is defined by the direction of \mathbf{E} and \mathbf{B} relative to the direction of the wave \mathbf{k} . The wave equations for \mathbf{E} and \mathbf{B} that include the sources ρ and \mathbf{J} are equivalent to Maxwell's equations.

You should master the following techniques:

1. derive the wave equation and construct its solutions;
2. find the dispersion relation from the wave equation;
3. calculate phase and group velocities;
4. give a vector description of linear, circular and elliptical polarisation.

6 Energy and Momentum of Electromagnetic Fields

The electric and magnetic fields are very much physical objects in their own right. This means that they not only have energy, but also momentum, stress and angular momentum.

6.1 Poynting's theorem and energy conservation

One of the most fundamental principles of physics is energy conservation. It is exact, as far as we know, and the question is therefore how it is manifested in electrodynamics. To find out we start with a simple vector identity:

$$\nabla \cdot (\mathbf{E} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{B}). \quad (6.1)$$

We can substitute the Maxwell equations for the curl of \mathbf{E} and \mathbf{B} to obtain

$$\nabla \cdot (\mathbf{E} \times \mathbf{B}) = -\mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial t} - \mu_0 \epsilon_0 \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} - \mu_0 \mathbf{E} \cdot \mathbf{J}. \quad (6.2)$$

Moreover, the cross product of \mathbf{E} and \mathbf{B} is related to the Poynting vector

$$\mathbf{S} = \frac{\mathbf{E} \times \mathbf{B}}{\mu_0}. \quad (6.3)$$

Putting this all together and simplifying the time derivative we have

$$\nabla \cdot \mathbf{S} + \mathbf{E} \cdot \mathbf{J} + \frac{\partial}{\partial t} \left(\frac{\mathbf{B} \cdot \mathbf{B}}{2\mu_0} + \frac{\epsilon_0 \mathbf{E} \cdot \mathbf{E}}{2} \right) = 0. \quad (6.4)$$

The question is now: what does this mathematical theorem mean?

Exercise 6.1: Show that the terms in Eq. (6.4) have units of *power density*.

Since these are densities, to interpret this theorem properly, let's integrate Eq. (6.4) over an arbitrary volume V bounded by a surface S :

$$\oint_S \mathbf{S} \cdot \mathbf{n} dS + \int_V \mathbf{J} \cdot \mathbf{E} dV + \frac{d}{dt} \int_V U dV = 0, \quad (6.5)$$

where we abbreviated

$$U = \frac{\mathbf{B} \cdot \mathbf{B}}{2\mu_0} + \frac{\epsilon_0 \mathbf{E} \cdot \mathbf{E}}{2} \equiv \frac{B^2}{2\mu_0} + \frac{\epsilon_0 E^2}{2}. \quad (6.6)$$

We again used Gauss' theorem to relate a volume integral to a surface integral. The first term is the rate of the energy flow out of the surface S . In the second term we calculate how much charges move along the direction of the electric field, which is a measure of the rate of work done on the charges by the electric field \mathbf{E} . We need the current density \mathbf{J} , because the charges have to be moving³. The final term in Eq. (6.4) is the remaining part of the puzzle, namely the change in energy contained in the fields \mathbf{E} and \mathbf{B} . The quantity U is therefore the *energy density* of the electromagnetic field.

This is called Poynting's theorem of conservation of energy. The energy can leave the fields in a region of space, but it must then be either transported through the surface (measured by the Poynting vector), or it must be used to do work on the charges in the volume. When we make the volume infinitesimally small, you see that energy conservation is not only true globally, but it is true *locally*, at every point in space.

³The work done on a charge in a time interval Δt is given by $W = \mathbf{d} \cdot \mathbf{F} = q\mathbf{v} \cdot \mathbf{E}\Delta t = \mathbf{J} \cdot \mathbf{E}\Delta t$.

6.2 Momentum and stress in the electromagnetic field

It is not surprising that the electric and magnetic fields carry energy. After all, you're quite familiar with the notion that it is the potential energy in the fields that make charges move. The general behaviour of the energy in the fields is described by Poynting's theorem, discussed in the previous lecture.

The electromagnetic field also has *momentum*. There are several immediate reasons why this must be the case:

1. The theory of electromagnetism is a relativistic theory, and we know that energy in one frame of reference gives rise to momentum in another reference frame. Therefore, to have energy is to have momentum⁴.
2. Last year you learned about *radiation pressure*, so you know that electromagnetic waves can exert a force on objects, and therefore transfer momentum.
3. When two charged particles fly close past each other with small relative velocity ($v \ll c$), Newton's third law holds, and the momentum of the two particles is conserved. However, when the relative velocity becomes large, the particles will each experience a \mathbf{B} field due to the other particle's motion. We can configure the situation such that the total momentum of the two charges is not conserved (see exercise), and in order to save Newton's third law, the field must be imbued with momentum.

Now let's study the properties of the momentum of the electromagnetic field. Since momentum is closely related to force, we first consider the Lorentz force.

We can write the Lorentz force $\Delta \mathbf{F}$ on a small volume ΔV as

$$\Delta \mathbf{F} = (\rho \mathbf{E} + \mathbf{J} \times \mathbf{B}) \Delta V. \quad (6.7)$$

we now define the force \mathbf{f} on a unit volume exerted by the electromagnetic field as $\mathbf{f} = \Delta \mathbf{F} / \Delta V = \rho \mathbf{E} + \mathbf{J} \times \mathbf{B}$. Since we are interested in the fields, and not the charge or current densities, we want to eliminate ρ and \mathbf{J} from the expression for \mathbf{f} . We use the inhomogeneous Maxwell equations for this:

$$\mathbf{f} = \epsilon_0 (\nabla \cdot \mathbf{E}) \mathbf{E} + \left(\frac{1}{\mu_0} \nabla \times \mathbf{B} - \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) \times \mathbf{B}. \quad (6.8)$$

Next, we wish to rewrite the term $\dot{\mathbf{E}} \times \mathbf{B}$ using the chain rule:

$$\frac{\partial \mathbf{E}}{\partial t} \times \mathbf{B} = \frac{\partial (\mathbf{E} \times \mathbf{B})}{\partial t} - \mathbf{E} \times \frac{\partial \mathbf{B}}{\partial t}. \quad (6.9)$$

Also, the term $(\nabla \times \mathbf{B}) \times \mathbf{B}$ can be rewritten as

$$\frac{1}{\mu_0} (\nabla \times \mathbf{B}) \times \mathbf{B} = \frac{1}{\mu_0} (\mathbf{B} \cdot \nabla) \mathbf{B} - \frac{1}{2\mu_0} \nabla B^2 = \frac{1}{\mu_0} (\mathbf{B} \cdot \nabla) \mathbf{B} + \frac{1}{\mu_0} (\nabla \cdot \mathbf{B}) \mathbf{B} - \frac{1}{2\mu_0} \nabla B^2, \quad (6.10)$$

where we used the fact that $\nabla \cdot \mathbf{B} = 0$, always. We put this term in to make the expression more symmetric. This leads to a lengthy expression:

$$\mathbf{f} = \epsilon_0 [(\nabla \cdot \mathbf{E}) \mathbf{E} + (\mathbf{E} \cdot \nabla) \mathbf{E}] + \frac{1}{\mu_0} [(\nabla \cdot \mathbf{B}) \mathbf{B} + (\mathbf{B} \cdot \nabla) \mathbf{B}] - \nabla U - \frac{1}{c^2} \frac{\partial \mathbf{S}}{\partial t}, \quad (6.11)$$

⁴Even though the actual *value* of the momentum in specific frames may be zero.

or in index notation

$$f_j = \epsilon_0 \left[(\partial_i E_i) E_j + (E_i \partial_i) E_j - \frac{1}{2} \partial_j (E_i E_i) \right] + \frac{1}{\mu_0} \left[(\partial_i B_i) B_j + (B_i \partial_i) B_j - \frac{1}{2} \partial_j (B_i B_i) \right] - \frac{1}{c^2} \frac{\partial S_j}{\partial t},$$

where U is the energy density per unit volume, and \mathbf{S} is the Poynting vector. The first two brackets can be written as a divergence:

$$f_j = \partial_i \left[\epsilon_0 \left(E_i E_j - \frac{\delta_{ij}}{2} E^2 \right) + \frac{1}{\mu_0} \left(B_i B_j - \frac{\delta_{ij}}{2} B^2 \right) \right] - \frac{1}{c^2} \frac{\partial S_j}{\partial t}. \quad (6.12)$$

We define the quantity in square brackets as the *Maxwell stress tensor* T_{ij} :

$$T_{ij} = \epsilon_0 \left(E_i E_j - \frac{\delta_{ij}}{2} E^2 \right) + \frac{1}{\mu_0} \left(B_i B_j - \frac{\delta_{ij}}{2} B^2 \right). \quad (6.13)$$

It is a symmetric rank two tensor and for all practical purposes behaves like a matrix. If we understand the divergence operator to act on one index to T we can write

$$\mathbf{f} = \nabla \cdot T - \frac{1}{c^2} \frac{\partial \mathbf{S}}{\partial t} \quad \text{or} \quad f_j = \partial_i T_{ij} - \frac{1}{c^2} \partial_t S_j. \quad (6.14)$$

When we integrate this over a volume, the Lorentz force on that volume becomes

$$\mathbf{F} = \int_V \left(\nabla \cdot T - \frac{1}{c^2} \frac{\partial \mathbf{S}}{\partial t} \right) d\tau = \oint_S T \cdot d\mathbf{a} - \frac{d}{dt} \int_V \frac{\mathbf{S}}{c^2} d\tau, \quad (6.15)$$

where $d\mathbf{a}$ is the surface increment on the boundary surface S of volume V . The tensor T therefore has units of force per area, or *stress*. The first term on the right-hand side is the pressure and shear on the volume, while the second term is the momentum transfer to the electromagnetic field.

There is a very intuitive picture for the meaning of the stress tensor in terms of the field lines. The stress tensor tells us that the field lines cannot be bent without any cost. They have a tension in them that must be overcome. Imagine bringing together two particles with equal charge. As you move them closer, the field lines must bend more and more, and we have to apply a force to bring them together. But this is just another way of saying that like charges repel each other! The opposite situation is also easily explained: two opposite charges are connected by curves field lines. Since the field lines want to straighten out, the two charges move toward each other to make this happen.

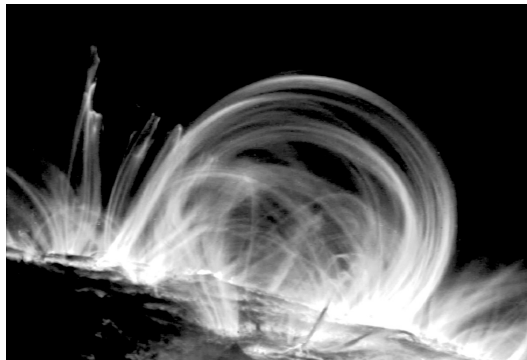


Figure 22: Arcs of magnetic field lines guiding hot plasma.

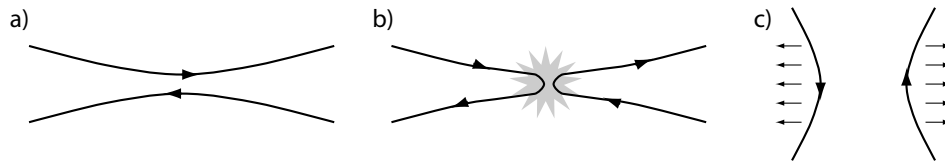


Figure 23: Origin of solar prominences.

A much more spectacular manifestation of tension in field lines occurs at the surface of the Sun. The magnetic field of the Sun is a giant tangle of field lines that get twisted and stretched due to the motion of the material inside. A hot, glowing plasma of charged particles spirals around the magnetic field lines and form the beautiful arcs on the surface shown in Fig. 22. These arcs explode into space, producing directed streams of high-energy charged particles that can knock out satellites and whole parts of the power grid. Maxwell's stress tensor explains why this is so.

The field lines in the arcs shown in Fig. 22 are not all going in the same direction. Rather, these are areas where field lines are oriented in opposite directions, creating a huge magnetic field gradient across relatively low-tension field lines (Fig. 23a). When the gradient becomes too high, the plasma causes a catastrophic breakdown of the field, and the field lines “reconnect”, as shown in Fig. 23b. However, now the field lines are bent to the extreme, and all that tension will cause the lines to move outwards, straightening. Charged particles trapped around these field lines will be accelerated to very high speeds and launched into space (Fig. 23c). These are *solar prominences*.

Summary

In this section we showed that electric and magnetic fields have energy and momentum. We defined the Poynting vector \mathbf{S} , and derived energy conservation for electromagnetic fields (Poynting's theorem). Using the Lorentz force we defined the Maxwell stress tensor, which describes the energy and the momentum in the field.

You should master the following techniques:

1. calculate the Poynting vector for given \mathbf{E} and \mathbf{B} ;
2. state Poynting's theorem, and explain the physical meaning of the three terms.

NB: You are not required to reproduce the derivation of the stress tensor, but you should be able to follow the mathematical steps in the derivation.

7 Radiation Sources and Antennas

7.1 Electric dipole radiation

So far we have looked at properties of electromagnetic waves (radiation) without asking how they are produced. Since radiation is an electromagnetic effect, we expect that it is created by electric charges. However, both static and uniformly moving charges cannot produce radiation (a uniformly moving charge is the same as a static charge in a different frame) because there is no characteristic time scale ω^{-1} in the physical system. We therefore need to look at *accelerating* charges, and to this end we construct an electric dipole in which the charge q oscillates between two points separated by a distance s (see figure 24). The charge at the top of the dipole can be written as $Q = q e^{-i\omega t}$, and we have a dipole $p = sq e^{-i\omega t}$ in the z direction. This is equivalent to a sinusoidal current $I_0 e^{-i\omega t}$.

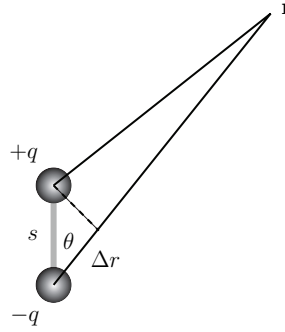


Figure 24: The Hertzian dipole: Radiation from an electric dipole current.

We now calculate the scalar and vector potentials in spherical coordinates for this situation. At some distance r the scalar potential due to the top of the dipole is given by

$$\Phi(r, t') = \frac{q e^{-i\omega t'}}{4\pi\epsilon_0 r} = \frac{q e^{-i\omega(t-r/c)}}{4\pi\epsilon_0 r} = \frac{q e^{-i\omega t + ikr}}{4\pi\epsilon_0 r} = \Phi_R(r, t), \quad (7.1)$$

where $t' = t - r/c$ is the *retarded* time: It takes a while for the change in the potential at $r = 0$ to propagate to r . To take this into account, we introduced the *retarded potential* $\Phi_R(r, t)$, and we used the dispersion relation for radiation in free space $k = \omega/c$.

Next, we add the two retarded potentials of the two charges $\Phi_R^{(+)}$ and $\Phi_R^{(-)} = -\Phi_R^{(+)}$ in the dipole. They almost cancel, but for the small separation s in the z direction, which gives rise to δr in the r direction:

$$\Phi_R(r, t) = \Phi_R^{(+)} + \Phi_R^{(-)} = \Phi_R^{(+)}(r, t) - \Phi_R^{(+)}(r + \delta r, t) \simeq -\delta r \frac{\partial \Phi_R^{(+)}}{\partial r} \quad (7.2)$$

Substituting Eq. (7.1) and $\delta r = s \cos \theta$ we obtain

$$\begin{aligned} \Phi_R(r, t) &= -s \cos \theta \frac{\partial}{\partial r} \left(\frac{q e^{i(kr - \omega t)}}{4\pi\epsilon_0 r} \right) \\ &= \frac{q s e^{i(kr - \omega t)} k \cos \theta}{4\pi\epsilon_0 r} \left(\frac{1}{kr} - i \right) \\ &= \frac{p_R k \cos \theta}{4\pi\epsilon_0 r} \left(\frac{1}{kr} - i \right), \end{aligned} \quad (7.3)$$

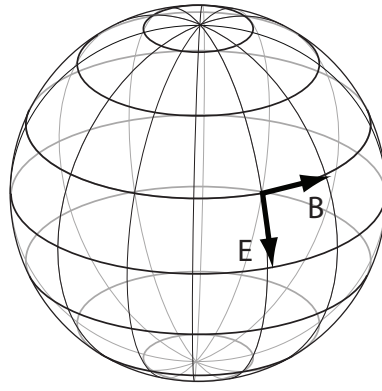


Figure 25: The directions of the \mathbf{E} and \mathbf{B} fields of a spherical wave radiated by a dipole in the centre. After half the oscillation period the fields point in the opposite direction.

where p_R is the retarded dipole moment.

The vector potential of this oscillating dipole in spherical coordinates is

$$\mathbf{A}_R(\mathbf{r}, t) = \frac{\mu_0}{4\pi} \int \frac{\mathbf{J}_R(\mathbf{r}', t)}{r} dV = \frac{1}{4\pi\epsilon_0 c^2 r} I_0 e^{i(kr - \omega t)} \mathbf{s} = \frac{1}{4\pi\epsilon_0 c^2 r} I_R \mathbf{s}, \quad (7.4)$$

where \mathbf{s} is the dipole vector $s(\cos\theta \hat{\mathbf{r}} - \sin\theta \hat{\theta})$, parallel to the z axis, and I_R is the retarded current in the dipole. Using the relations $Q = q e^{-i\omega t}$ and $I = dQ/dt$, we can write the dipole moment p_R in the scalar potential in terms of I_R :

$$p_R = \frac{is}{\omega} I_R. \quad (7.5)$$

We therefore have

$$\Phi_R(\mathbf{r}, t) = \frac{I_R s \cos\theta}{4\pi\epsilon_0 cr} \left(1 + \frac{i}{kr}\right) \quad \text{and} \quad \mathbf{A}_R(\mathbf{r}, t) = \frac{I_R s (\cos\theta \hat{\mathbf{r}} - \sin\theta \hat{\theta})}{4\pi\epsilon_0 c^2 r}. \quad (7.6)$$

The electric field \mathbf{E} is given by $\mathbf{E} = -\nabla\Phi_R - \dot{\mathbf{A}}_R$. Using the derivatives in spherical coordinates we find

$$\begin{aligned} E_r(\mathbf{r}, t) &= -\frac{\partial\Phi_R}{\partial r} - (\dot{\mathbf{A}}_R)_r = \frac{k I_R s \cos\theta}{4\pi\epsilon_0 cr} \left(\frac{2}{kr} + \frac{2i}{(kr)^2} \right), \\ E_\theta(\mathbf{r}, t) &= -\frac{1}{r} \frac{\partial\Phi_R}{\partial\theta} - (\dot{\mathbf{A}}_R)_\theta = \frac{k I_R s \sin\theta}{4\pi\epsilon_0 cr} \left(\frac{1}{kr} + \frac{i}{(kr)^2} - i \right). \end{aligned} \quad (7.7)$$

We already know that radiation carries energy, so the total energy flux through a closed surface around the source is constant. If we take the surface to be that of a sphere with radius r , the energy flux $4\pi r^2(\epsilon|\mathbf{E}|^2/2 + |\mathbf{B}|^2/2\mu_0)$ must be constant for all r . The radiating part of the \mathbf{E} and \mathbf{B} fields are therefore proportional to $1/r$. Looking at Eq. (7.7), we see that the only contribution to the radiation is due to the i term in E_θ . Similarly, the only contributing term of the \mathbf{B} field is (see exercise)

$$B_\phi(\mathbf{r}, t) = -i \frac{k I_R s \sin\theta}{4\pi\epsilon_0 c^2 r} = \frac{E_\theta(\mathbf{r}, t)}{c}, \quad (7.8)$$

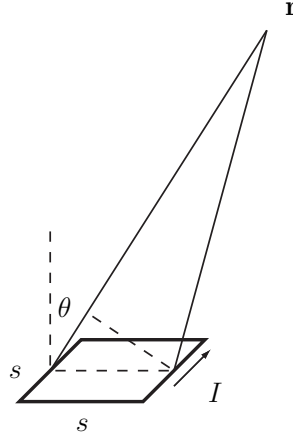


Figure 26: Radiation from a magnetic dipole current.

where $E_\theta(\mathbf{r}, t)$ now denotes the radiating part of the \mathbf{E} field. Poynting's vector is given by

$$\mathbf{S}_e = \frac{\text{Re} E_\theta \times \text{Re} B_\phi}{\mu_0} = \text{Re} \frac{\mathbf{E} \times \mathbf{B}^*}{\mu_0} = \frac{1}{16\pi^2 \epsilon_0 c} \left(\frac{ksI \sin \theta}{r} \right)^2 \hat{\mathbf{r}}. \quad (7.9)$$

Note that we can calculate the Poynting vector by taking the complex conjugate of the \mathbf{B} field and take the real part of \mathbf{S} afterwards. Again, it is often much easier to calculate with complex phase factors of the form $e^{i\varphi}$ than with the trigonometric counterparts $\sin \varphi$, $\cos \varphi$ and $\tan \varphi$.

7.2 Magnetic dipole radiation

Now suppose we have a small current going in a square loop of area s^2 , leading to a magnetic dipole moment $m = s^2 I_0 e^{-i\omega t}$. The current is changing in time in a periodic way. Again, we can define the retarded magnetic dipole by substituting $t \rightarrow t - r/c$. If the current has a low impedance the scalar potential is zero, while the vector potential is

$$\mathbf{A}_R(\mathbf{r}, t) = \frac{\mu_0 m_R k \sin \theta}{4\pi r} \left(-\frac{1}{kr} + i \right) \hat{\phi}. \quad (7.10)$$

We can again find the radiating parts of the \mathbf{E} and \mathbf{B} fields via $\mathbf{E} = -\dot{\mathbf{A}}$ and $\mathbf{B} = \nabla \times \mathbf{A}$:

$$E_\phi = \frac{\mu_0 m_R k^2 c \sin \theta}{4\pi r} = -c B_\theta, \quad (7.11)$$

and the Poynting vector is

$$\mathbf{S}_m = \text{Re} \frac{\mathbf{E} \times \mathbf{B}^*}{2\mu_0} = \frac{1}{16\pi^2 \epsilon_0 c} \left[\frac{(ks)^2 I \sin \theta}{r} \right]^2 \hat{\mathbf{r}}. \quad (7.12)$$

The difference between \mathbf{S}_e and \mathbf{S}_m is a factor $(ks)^2$.

7.3 Radiation from accelerated charges

Clearly, radiation is generated by accelerated charges and changing currents. In the previous sections we have looked at periodically changing charges and currents, but in many situations

the charges and currents change in an aperiodic way. For example, when a charge is kicked, it will emit a burst of radiation. We will now consider the question of how much power P is radiated when a charge is accelerated by an amount a . This will lead to Larmor's formula.

Suppose a particle with charge q follows a trajectory $\mathbf{r}_q(t)$. We can formally associate a dipole moment \mathbf{p} to this particle:

$$\mathbf{p}(t) = q \mathbf{r}_q(t). \quad (7.13)$$

The current due to the motion of the charge is clearly related to the velocity \mathbf{v}_q , so we can write

$$I = \frac{d\mathbf{p}}{dt} = q \frac{d\mathbf{r}_q}{dt} = q \mathbf{v}_q(t). \quad (7.14)$$

The acceleration \mathbf{a}_q of the particle is proportional to the second derivative of the dipole moment:

$$\frac{d^2\mathbf{p}}{dt^2} = q \mathbf{a}_q(t). \quad (7.15)$$

Next, we calculate the vector potential and the magnetic field \mathbf{B} of the moving charge.

First, we note that the current I is the spatial integration over the current density, so we have

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0}{4\pi r} \int \mathbf{J}(\mathbf{r}', t - r/c) d\mathbf{r}' = \frac{\mu_0}{4\pi r} I_R = \frac{\mu_0}{4\pi r} \frac{d\mathbf{p}(t - r/c)}{dt}, \quad (7.16)$$

evaluated at the *retarded* time, because the field needs time to propagate from the charge to the point where the vector potential is calculated.

The magnetic field is found using $\mathbf{B} = \nabla \times \mathbf{A}$. We need to keep in mind two things: (1) We can ignore all the contributions that do not scale like $1/r$, because they do not propagate. (2) We must remember that the time variable depends on r because we consider the retarded time. The magnetic field is then

$$\mathbf{B}(\mathbf{r}, t) = \nabla \times \mathbf{A}(\mathbf{r}, t) = \frac{\mu_0}{4\pi} \left[\nabla \times \left(\frac{1}{r} \right) \frac{d\mathbf{p}}{dt} + \frac{1}{r} \nabla \times \frac{d\mathbf{p}}{dt} \right] = \frac{\mu_0}{4\pi r} \nabla \times \frac{d\mathbf{p}}{dt}, \quad (7.17)$$

where the last equality holds because $\nabla \times (1/r)$ leads to a non-propagating contribution to the field. The calculation of the curl of $d\mathbf{p}/dt$ is a bit tricky, so we'll do it here in some detail. In particular, it is convenient to use index notation:

$$\left(\nabla \times \frac{d\mathbf{p}}{dt} \right)_i = \epsilon_{ijk} \partial_j \left(\frac{d\mathbf{p}}{dt} \right)_k. \quad (7.18)$$

Using the chain rule, we can write ($t = t_q - r/c$ with)

$$\partial_j = \frac{\partial t}{\partial x_j} \frac{d}{dt} = -\frac{1}{c} \frac{x_j}{r} \frac{d}{dt}, \quad (7.19)$$

which allows us to rewrite the cross product:

$$\left(\nabla \times \frac{d\mathbf{p}}{dt} \right)_i = -\frac{1}{c} \epsilon_{ijk} \frac{x_j}{r} \left(\frac{d^2\mathbf{p}}{dt^2} \right)_k. \quad (7.20)$$

In vector notation, this becomes

$$\nabla \times \frac{d\mathbf{p}}{dt} = -\frac{1}{cr} \mathbf{r} \times \frac{d^2\mathbf{p}}{dt^2} = -\frac{1}{c} \hat{\mathbf{r}} \times \frac{d^2\mathbf{p}}{dt^2} = -\frac{q}{c} \hat{\mathbf{r}} \times \mathbf{a}_q, \quad (7.21)$$

and therefore the \mathbf{B} field can be written as

$$\mathbf{B}(\mathbf{r}, t) = \frac{\mu_0}{4\pi r} \nabla \times \frac{d\mathbf{p}}{dt} = -\frac{\mu_0}{4\pi cr} \hat{\mathbf{r}} \times \frac{d^2\mathbf{p}}{dt^2} = -\frac{q\mu_0}{4\pi cr} \hat{\mathbf{r}} \times \mathbf{a}_q, \quad (7.22)$$

evaluated at time t , and due to an accelerated charge at the retarded time $t - r/c$.

The radiating part of the electric field \mathbf{E} must be perpendicular to the radiating part of the magnetic field that we just calculated, and they differ by a factor c . Keeping the correct handedness, the radiating \mathbf{E} field can be written as a cross product between \mathbf{B} and $\hat{\mathbf{r}}$:

$$\mathbf{E}(\mathbf{r}, t) = c\mathbf{B} \times \hat{\mathbf{r}} = -\frac{q\mu_0}{4\pi r} (\hat{\mathbf{r}} \times \mathbf{a}_q) \times \hat{\mathbf{r}}. \quad (7.23)$$

Now that we know the \mathbf{E} and \mathbf{B} fields, we can calculate the Poynting vector \mathbf{S} , which gives the energy flux through a surface. If we take the surface to be a sphere with radius R surrounding the accelerating charge, we can calculate the total radiated power

$$P = \oint_S \mathbf{S} \cdot \hat{\mathbf{R}} R^2 d\Omega, \quad (7.24)$$

where $d\Omega$ is a solid angle.

The Poynting vector is straightforward to calculate:

$$\mathbf{S} = \frac{\mathbf{E} \times \mathbf{B}}{\mu_0} = \frac{\mu_0 q^2}{16\pi^2 c r^2} [(\hat{\mathbf{r}} \times \mathbf{a}_q) \times \hat{\mathbf{r}}] \times (\hat{\mathbf{r}} \times \mathbf{a}_q). \quad (7.25)$$

Using the vector identity $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})$ we rewrite this as

$$\begin{aligned} \mathbf{S} &= \frac{q^2}{16\pi^2 \epsilon_0 c^3 r^2} [\mathbf{a}_q - \hat{\mathbf{r}}(\mathbf{a}_q \cdot \hat{\mathbf{r}})] \times (\hat{\mathbf{r}} \times \mathbf{a}_q) \\ &= \frac{q^2}{16\pi^2 \epsilon_0 c^3 r^2} \{ \hat{\mathbf{r}} [|\mathbf{a}_q|^2 - (\mathbf{a}_q \cdot \hat{\mathbf{r}})^2] - \mathbf{a}_q (\mathbf{a}_q \cdot \hat{\mathbf{r}} - \mathbf{a}_q \cdot \hat{\mathbf{r}}) \} \\ &= \frac{q^2}{16\pi^2 \epsilon_0 c^3 r^2} [a_q^2 - (\mathbf{a}_q \cdot \hat{\mathbf{r}})^2] \hat{\mathbf{r}} \\ &= \frac{q^2 a_q^2 \sin^2 \theta}{16\pi^2 \epsilon_0 c^3 r^2} \hat{\mathbf{r}}, \end{aligned} \quad (7.26)$$

where we defined $a_q = |\mathbf{a}_q|$ and we oriented our coordinate system such that the acceleration is along the z direction. The inner product $\mathbf{a}_q \cdot \hat{\mathbf{r}}$ is therefore equal to $a_q \cos \theta$.

The radiated power through a sphere surrounding the accelerating charge is now straightforward to calculate:

$$P = \oint_S \mathbf{S} \cdot \hat{\mathbf{R}} R^2 d\Omega = \oint_S \frac{q^2 a_q^2 \sin^2 \theta}{16\pi^2 \epsilon_0 c^3 R^2} R^2 d\Omega = \frac{q^2 a_q^2}{16\pi^2 \epsilon_0 c^3} \int_0^\pi \int_0^{2\pi} \sin^2 \theta d\theta d\phi. \quad (7.27)$$

The integral over $\sin^2 \theta$ is equal to $8\pi/3$, so we arrive at

$$P = \frac{1}{4\pi \epsilon_0} \frac{2q^2 a_q^2}{3c^3}. \quad (7.28)$$

This is the celebrated *Larmor formula*, derived by Joseph Larmor in 1897.

Summary

In this section we started with a calculation of the fields produced by an oscillating dipole, and showed that this gives rise to *electromagnetic radiation*. The radiation from an electric dipole is generally much stronger than that of a magnetic dipole. We then derived the power radiated by an accelerated charge.

You should master the following techniques:

1. given time-varying scalar and vector potentials, calculate the electromagnetic radiation;
2. use Poynting's theorem to set up an energy balance for radiating charges and currents;
3. calculate the radiated power by an accelerating charge.

NB: You do not need to learn the derivation of Larmor's formula, but you should know the final formula.

8 Electrodynamics in Macroscopic Media

So far, we have considered the fields of (moving) charges and the interactions between the charges. In principle, this covers all classical electromagnetic phenomena, since all matter is made up of charged particles that interact with each other. However, if we are dealing with 10^{23} atoms we *don't want to* solve the Maxwell equations for every individual atom. Especially when the time and length scales relevant to our problem are large compared to that of the atoms, we much rather average over the atomic fields and have *effective fields* in the media.

The electric and magnetic fields in Maxwell's equations are now average fields $\langle \mathbf{E} \rangle$ and $\langle \mathbf{B} \rangle$, and the charge and current densities $\langle \rho \rangle$ and $\langle \mathbf{J} \rangle$ now consist of free charges and currents as well as *bound* charges and currents. Maxwell's equations in macroscopic media then become

$$\nabla \cdot \langle \mathbf{B} \rangle = 0 \quad \text{and} \quad \nabla \times \langle \mathbf{E} \rangle + \frac{\partial \langle \mathbf{B} \rangle}{\partial t} = 0 \quad (8.1)$$

$$\nabla \cdot \langle \mathbf{E} \rangle = \frac{\langle \rho \rangle}{\epsilon_0} \quad \text{and} \quad \nabla \times \langle \mathbf{B} \rangle - \mu_0 \epsilon_0 \frac{\partial \langle \mathbf{E} \rangle}{\partial t} = \mu_0 \langle \mathbf{J} \rangle. \quad (8.2)$$

The charge and current densities are

$$\langle \rho \rangle = \rho_f + \rho_b \quad \text{and} \quad \langle \mathbf{J} \rangle = \mathbf{J}_f + \mathbf{J}_b \quad (8.3)$$

where the subscripts f and b denote “free” and “bound”, respectively.

The fields in the materials are no longer the bare \mathbf{E} and \mathbf{B} fields due to freely moving charges and currents (as was the case for the microscopic Maxwell equations), but they are modified by the presence of the bound charges and currents. We want to relate the modified fields to material properties (the bound charges and currents) and modify Maxwell's equations such that they involve the macroscopic fields and the free charges and currents.

8.1 Polarization and Displacement fields

First, we consider a material that has bound charges. These are, for example, electrons that cannot move freely through the material, but that are bound to the parent molecule. When an external electric field is applied, the negative electrons move with respect to the positive nuclei, creating a small dipole moment. This dipole moment is aligned *against* the external field, because the external field pulls the charges such that the local field becomes smaller. All these little dipole moments create a macroscopic field \mathbf{P} , which is related to the bound charge density via

$$\rho_b = -\nabla \cdot \mathbf{P}. \quad (8.4)$$

This is in some sense the *definition* of \mathbf{P} . We must now relate the \mathbf{P} to the dipole moments of the molecules.

Let V be the volume of the medium, and the dipole density in the i direction is $\rho_b r_i$. Then

$$\int_V r_i \rho_b d\mathbf{r} = - \sum_{j=1}^3 \int_V r_i \frac{\partial P_j}{\partial r_j} d\mathbf{r} = - \sum_{j=1}^3 \int_S r_i P_j d^2 r_{k \neq j} \Big|_{r_j \text{ on } S}^{r_j \text{ on } S} + \int_V P_i d\mathbf{r}, \quad (8.5)$$

where the first equality comes from Eq. (8.4), and the second comes from partial intergration. The surface term is zero when we take the surface just outside the medium. In vector notation, we therefore have

$$\int_V \mathbf{r} \rho_b d\mathbf{r} = \int_V \mathbf{P} d\mathbf{r} = \sum \mathbf{p}_n, \quad (8.6)$$

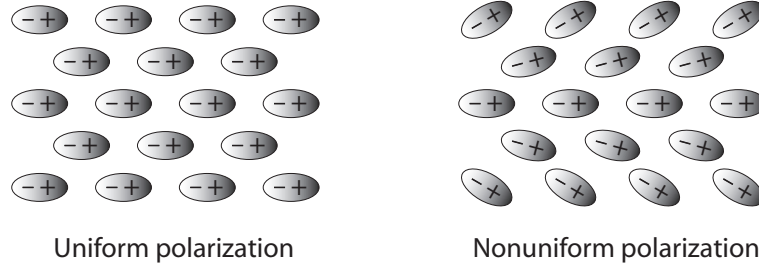


Figure 27: Polarization in a medium

where the last term is the sum of all microscopic dipole moments in the volume V . Outside V we have $\mathbf{P} = 0$.

Now that we know what \mathbf{P} means, we can rewrite the Maxwell equation $\nabla \cdot \langle \mathbf{E} \rangle = \langle \rho \rangle / \epsilon_0$ using Eqs. (8.3) and (8.4) as

$$\nabla \cdot \langle \mathbf{E} \rangle = \frac{\rho_f + \rho_b}{\epsilon_0} = \frac{\rho_f}{\epsilon_0} - \frac{1}{\epsilon_0} \nabla \cdot \mathbf{P}. \quad (8.7)$$

Considering we want the macroscopic field in terms of the free charges, we can *define* the macroscopic field as the displacement field \mathbf{D} :

$$\mathbf{D} = \epsilon_0 \langle \mathbf{E} \rangle + \mathbf{P} \quad (8.8)$$

All the material properties are determined by \mathbf{P} . The corresponding macroscopic Maxwell equation is then

$$\boxed{\nabla \cdot \mathbf{D} = \rho_f} \quad (8.9)$$

Often the polarization field is directly proportional to \mathbf{E} , and we write $\mathbf{P} = \chi \epsilon_0 \mathbf{E}$, where χ is the *susceptibility* of the medium. The displacement field then becomes

$$\mathbf{D} = \epsilon_0(1 + \chi)\mathbf{E} \equiv \epsilon_0 \epsilon_r \mathbf{E} \equiv \epsilon \mathbf{E}. \quad (8.10)$$

Assuming for the moment that the permeability μ_0 does not change, one can see immediately that the wave equation yields a propagation velocity c^2/ϵ_r . In ordinary materials, we therefore have $\epsilon_r \geq 1$. As we will see in the next section, this is directly related to the index of refraction.

8.2 Magnetization and Magnetic induction

Now we will construct the macroscopic magnetic field using the free and bound currents. Apart from the free currents in the material, there may be bound currents that are confined to the molecules. In quantum mechanics these bound currents can be the spins of the electrons and nuclei, but we'll keep things classical.

The bound currents may be due to a change of bound charges, which can be written as $\partial \mathbf{P} / \partial t$. Alternatively, they may correspond to little current loops, which give rise to a microscopic magnetic dipole moment. Under the influence of an external field \mathbf{B} , these little magnets can line up to form a macroscopic field that we shall denote by $\nabla \times \mathbf{M}$. We then have the following expression for the bound current:

$$\mathbf{J}_b = \frac{\partial \mathbf{P}}{\partial t} + \nabla \times \mathbf{M} \quad (8.11)$$

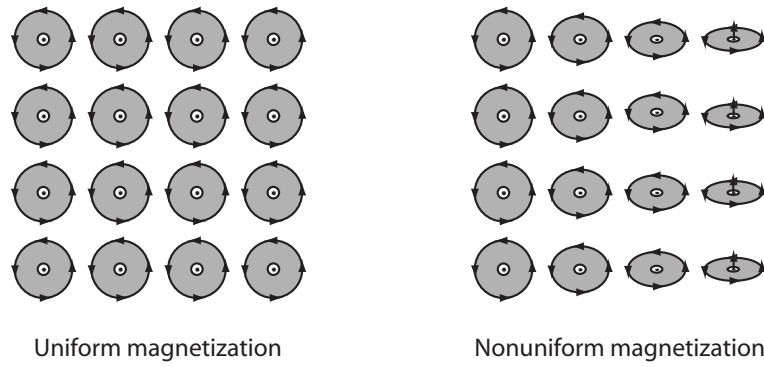


Figure 28: Magnetization in a medium.

The divergence of the bound charge is

$$\nabla \cdot \mathbf{J}_b = \frac{\partial \nabla \cdot \mathbf{P}}{\partial t} + \nabla \cdot (\nabla \times \mathbf{M}) = -\frac{\partial \rho_b}{\partial t} = 0, \quad (8.12)$$

since the bound charge is conserved.

Now let's consider the second inhomogeneous Maxwell equation

$$\nabla \times \langle \mathbf{B} \rangle - \mu_0 \epsilon_0 \frac{\partial \langle \mathbf{E} \rangle}{\partial t} = \nabla \times \langle \mathbf{B} \rangle - \mu_0 \epsilon_0 \frac{\partial}{\partial t} \left(\frac{\mathbf{D}}{\epsilon_0} - \frac{\mathbf{P}}{\epsilon_0} \right) = \mu_0 \langle \mathbf{J} \rangle = \mu_0 (\mathbf{J}_b + \mathbf{J}_f). \quad (8.13)$$

This leads to

$$\nabla \times \langle \mathbf{B} \rangle - \mu_0 \frac{\partial \mathbf{D}}{\partial t} + \mu_0 \frac{\partial \mathbf{P}}{\partial t} = \mu_0 \frac{\partial \mathbf{P}}{\partial t} + \mu_0 \nabla \times \mathbf{M} + \mu_0 \mathbf{J}_f. \quad (8.14)$$

We can now define the *magnetic field* \mathbf{H} in terms of the magnetic *induction* \mathbf{B} and the macroscopic magnetization \mathbf{M} as

$$\mathbf{H} = \frac{1}{\mu_0} \langle \mathbf{B} \rangle - \mathbf{M}, \quad (8.15)$$

yielding the Maxwell equation

$$\boxed{\nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{J}_f} \quad (8.16)$$

The homogeneous Maxwell equations remain unchanged, since here \mathbf{E} and \mathbf{B} are fundamental (that is, no material properties are present in these equations). Also, for the macroscopic case it is understood that all fields are averages, and we drop the brackets $\langle \cdot \rangle$. The macroscopic Maxwell equations then become

$$\nabla \cdot \mathbf{B} = 0 \quad \text{and} \quad \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \quad (8.17)$$

$$\nabla \cdot \mathbf{D} = \rho_f \quad \text{and} \quad \nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{J}_f. \quad (8.18)$$

8.3 Waves in macroscopic media

When waves penetrate a medium, the molecules in the medium respond to the electric field and the magnetic induction to form polarization and magnetization fields. Let's assume that the medium is electrically neutral, so that there are no free charges ($\rho = 0$), but there may be electrons in a conduction band that can produce free currents \mathbf{J} . These electrons presumably respond directly to the electric field, such that the free current is proportional to \mathbf{E} :

$$\mathbf{J} = \sigma \mathbf{E} \quad (\text{Ohm's law}). \quad (8.19)$$

This is Ohm's law, and the constant of proportionality σ is the *conductance*. This law has a wide range of applicability, but you should be aware that there are many interesting situations where Ohm's law does not hold.

There are essentially three regimes of interest for σ : (1) σ is small, and the medium is a dielectric; (2) σ is large, and the medium is a conductor; or (3) σ is imaginary, and the medium is a plasma. We will derive the index of refraction for dielectrics, the skin depth for conductors, and the plasma frequency for plasmas. We will also derive the reflectivity of these three types of materials.

Waves in dielectrics

First, we rederive the wave equations from the Maxwell equations, but this time we keep the \mathbf{J} term. We find the following differential equations for \mathbf{E} and \mathbf{B} :

$$\nabla^2 \mathbf{E} - \mu\epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} = \mu \frac{\partial \mathbf{J}}{\partial t}, \quad (8.20)$$

$$\nabla^2 \mathbf{B} - \mu\epsilon \frac{\partial^2 \mathbf{B}}{\partial t^2} = -\mu \nabla \times \mathbf{J}, \quad (8.21)$$

where $\epsilon = \epsilon_r \epsilon_0$ and $\mu = \mu_r \mu_0$. Using Ohm's law, this leads to

$$\nabla^2 \mathbf{E} - \mu\epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} - \mu\sigma \frac{\partial \mathbf{E}}{\partial t} = 0, \quad (8.22)$$

$$\nabla^2 \mathbf{B} - \mu\epsilon \frac{\partial^2 \mathbf{B}}{\partial t^2} - \mu\sigma \frac{\partial \mathbf{B}}{\partial t} = 0. \quad (8.23)$$

We substitute again the plane wave solutions for waves in the positive z direction $\mathbf{E} = \mathbf{E}_0 e^{i(kz - \omega t)}$ and $\mathbf{B} = \mathbf{B}_0 e^{i(kz - \omega t)}$, and we immediately find that

$$k = \sqrt{\epsilon\mu\omega^2 + i\omega\mu\sigma} = k_0 \sqrt{\mu_r \epsilon_r} \sqrt{1 + \frac{i\sigma}{\omega\epsilon}}, \quad (8.24)$$

where $k_0 = \omega/c$. You can see that the wave vector k is complex!

Now we can look at the three different regimes for σ . If the conductance is small ($\sigma/\omega\epsilon \ll 1$), the medium is a dielectric and the imaginary root is approximately equal to 1. The wave vector is real, so the waves propagate freely. However, the value of the wave vector is modified by the factor $\sqrt{\epsilon_r \mu_r}$. It turns out that this is indeed the index of refraction, at least for apolar materials⁵.

⁵In apolar media the distribution of the electrons in the molecules is fast enough to follow the oscillating fields. In polar materials, however, the response of the medium is slow since the polar molecules have to re-orient themselves. This is the principle behind the microwave: The water molecules try to realign themselves with an oscillating microwave field, and as a consequence the water gets heated.

For ordinary dielectrics where $\mathbf{D} = \epsilon \mathbf{E}$ and $\mathbf{B} = \mu \mathbf{H}$, the ratio between the electric field and the magnetic induction is given by the velocity of the wave in the medium:

$$\frac{E}{B} = \frac{c}{\sqrt{\epsilon_r \mu_r}}. \quad (8.25)$$

In air, this leads to approximately $E = cB$, and the (average) Poynting vector in a medium is then

$$\mathbf{S}_{\text{av}} = n \epsilon_0 c E_{\text{rms}}^2, \quad (8.26)$$

where n is the index of refraction, and E_{rms}^2 is the average strength to the electric field squared.

Waves in conductors

In a conducting medium σ is very large, and we can ignore the $+1$ term in Eq. (8.24). The wave vector k becomes

$$k = k_0 \sqrt{\frac{i \mu_r \sigma}{\omega \epsilon_0}} = \pm \sqrt{\frac{\omega \sigma \mu}{2}} (1 + i) \equiv \pm \left(\frac{1}{\delta} + \frac{i}{\delta} \right). \quad (8.27)$$

Since the wave propagates in the positive z direction, we choose the positive sign. We defined the *skin depth* δ as

$$\delta = \sqrt{\frac{2}{\omega \sigma \mu}}. \quad (8.28)$$

When we substitute this value for k into the solution for the \mathbf{E} field we obtain

$$\mathbf{E}(z, t) = \mathbf{E}_0 \exp \left[i \left(\frac{z}{\delta} - \omega t \right) \right] \exp \left(-\frac{z}{\delta} \right), \quad (8.29)$$

that is, there is a propagating term and an exponentially decaying term, which falls off to e^{-1} when the wave has reached the skin depth.

The magnetic field is given by $B = kE/\omega$, where now k is complex. With the time dependence made explicit, we find that

$$B = \sqrt{\frac{\sigma \mu}{\omega}} E_0 e^{i(z/\delta - \omega t + \pi/4)} e^{-z/\delta}. \quad (8.30)$$

Note the $\pi/4$ phase lag of B with respect to E . Inside the metal we have

$$B_0 = \sqrt{\frac{\sigma \mu}{\omega}} E_0 e^{i\pi/4} \quad (8.31)$$

and the Poynting vector is

$$\mathbf{S}_{\text{av}} = \frac{1}{2} \sqrt{\frac{\sigma}{2\omega \mu}} E_0^2 e^{-2z/\delta} \hat{\mathbf{k}}. \quad (8.32)$$

This shows that the power falls off twice as fast as the fields.

Waves in plasmas

Finally, consider a plasma in which the electrons are disassociated from the atomic nuclei. When an electric field is applied the electrons accelerate. The current density \mathbf{J} therefore *lags* the electric field, and the phase happens to be $e^{i\pi/2} = i$. The nuclei are much heavier than the electrons,

so they do not accelerate as fast and make a negligible contribution to the conductivity. We can find the velocity \mathbf{v} of the electrons by integrating $\mathbf{F} = m\mathbf{a} = e\mathbf{E}$, where all vector quantities are in the z direction.

$$\mathbf{v} = \frac{e}{m} \int dt \mathbf{E}_0 e^{-i\omega t} = \frac{ie}{m\omega} \mathbf{E}, \quad (8.33)$$

where we explicitly wrote the time dependence of \mathbf{E} . Using the electron density Ne , the current density is

$$\mathbf{J} = Ne\mathbf{v} = Ne \frac{ie\mathbf{E}}{\omega m}, \quad \text{and} \quad \sigma = \frac{\mathbf{J}}{\mathbf{E}} = \frac{ie^2 N}{\omega m}. \quad (8.34)$$

We substitute this into our expression for the complex wave vector k

$$k = k_0 \sqrt{1 - \frac{e^2 N}{\omega^2 m \epsilon_0}} = k_0 \sqrt{1 - \left(\frac{\omega_p}{\omega}\right)^2}, \quad (8.35)$$

where we defined the plasma frequency

$$\omega_p = \sqrt{\frac{e^2 N}{m \epsilon_0}}. \quad (8.36)$$

Clearly, when the frequency of the wave ω is larger than the plasma frequency, k is real and the plasma acts as a dielectric. When $\omega < \omega_p$ the plasma acts as a conductor.

8.4 Frequency dependence of the index of refraction

The very simple description of waves in dielectric media with constant ϵ and μ cannot be the whole story. We know that in real materials the index of refraction has some frequency dependence, or the phenomena of prisms and rainbows would not exist. We will now give a simple model of dielectric materials that leads to such a frequency dependence.

Imagine that the material is made up of atoms whose bound electrons respond to the electromagnetic wave. The electrons will follow the electric field and will oscillate in, say, the x -direction. Since the electrons are bound (and the wave intensity is relatively low), we can model the electrons as sitting in a harmonic potential. The equation of motion for the position x of each electron is then given by

$$m \frac{d^2 x}{dt^2} + m\gamma \frac{dx}{dt} + m\omega_0^2 x = qE_0 e^{-i\omega t}, \quad (8.37)$$

where m is the mass of the electron, q its charge, ω_0 is the resonance frequency of the electron, and γ is a damping factor that will siphon energy from the wave in a process generally called *dissipation*. The driving force on the electron is given by the electromagnetic wave with frequency ω and amplitude E_0 . Note that we are treating the atoms in a purely classical way, and not quantum mechanically. This is sufficient for our purposes here, but it is of course not a true description of the dielectric material.

It is easy to verify that the solution to this differential equation is given by the *complex* position function

$$x(t) = \frac{q}{m} \frac{E_0 e^{-i\omega t}}{\omega_0^2 - \omega^2 - i\gamma\omega}. \quad (8.38)$$

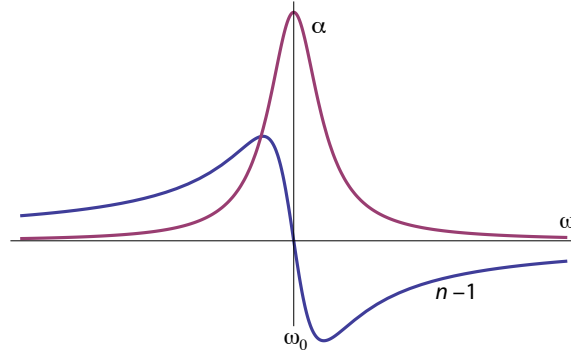


Figure 29: Index of refraction $n - 1$ and absorption coefficient α near the resonance frequency ω_0 of a simple dielectric material.

The dipole moment of the electron then becomes

$$p(t) = qx(t) = \frac{q^2}{m} \frac{E_0 e^{-i\omega t}}{\omega_0^2 - \omega^2 - i\gamma\omega}. \quad (8.39)$$

Since this is an oscillating dipole it will generate radiation at the driving frequency ω . The driving wave is propagating in one direction and the dipole emits in a radial direction, so this amounts to losses in the energy of the wave. The polarisation field of the dielectric medium is given by the density of dipoles in the material, and we can define the polarisation field as

$$\mathbf{P} = \frac{Nq^2}{m} \frac{1}{\omega_0^2 - \omega^2 - i\gamma\omega} \mathbf{E}, \quad (8.40)$$

where N is the density of electrons with resonance frequency ω_0 . If there are multiple resonance frequencies in the material we add them to the polarisation field with the corresponding density. In this case, both \mathbf{P} and \mathbf{E} are complex, and they obey the relationship

$$\mathbf{P} = \chi\epsilon_0\mathbf{E}, \quad (8.41)$$

as before. The *complex* dielectric constant is then given by

$$\epsilon = \epsilon_0(1 + \chi) = \epsilon_0 \left(1 + \frac{Nq^2}{m} \frac{1}{\omega_0^2 - \omega^2 - i\gamma\omega} \right). \quad (8.42)$$

The wave in the dielectric medium therefore has a complex wave vector $k = \omega\sqrt{\epsilon\mu}$. If we assume that $\mu = \mu_0$ and that N is relatively small, we can use the approximation $\sqrt{1+x} \approx 1 + \frac{1}{2}x$, and we find

$$k = \frac{\omega}{c} \left(1 + \frac{Nq^2}{2m\epsilon_0} \frac{1}{\omega_0^2 - \omega^2 - i\gamma\omega} \right). \quad (8.43)$$

In turn, the index of refraction is given by the real part of ck/ω , which is given by

$$n = \frac{c\text{Re}(k)}{\omega} = 1 + \frac{Nq^2}{2m\epsilon_0} \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}. \quad (8.44)$$

The imaginary part of the wave vector causes an exponential decay of the wave amplitude, and is therefore a measure of the *loss* of energy in the wave. Since the intensity of the wave is proportional to E^2 , the absorption coefficient α is given by

$$\alpha = 2\text{Im}(k) = \frac{Nq^2\omega^2}{mc\epsilon_0} \frac{\gamma}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}. \quad (8.45)$$

The index of refraction and the absorption coefficient are shown in Fig. 29.

You see that the index of refraction n dips below 1 at frequencies higher than the resonance frequency. This means that the phase velocity will be larger than c ! However, it is the group velocity that measures the velocity of the energy in the wave, and this will still be lower than c . In the limit of negligible damping, far away from the resonance frequency, the group velocity is

$$v_g = \frac{d\omega}{dk} = \frac{c}{1 + \frac{Nq^2}{2m\epsilon_0} \frac{\omega^2 + \omega_0^2}{(\omega^2 - \omega_0^2)^2}}. \quad (8.46)$$

The region where the index of refraction drops below one is called *anomalous* dispersion.

Exercise 8.1: Verify the formula for the group velocity in Eq. (8.46).

Summary

In this section we derived Maxwell's equations for macroscopic media. In addition to the standard fields \mathbf{E} and \mathbf{B} , we introduced the displacement field \mathbf{D} and the magnetic field \mathbf{H} (\mathbf{B} is now called the magnetic *induction*). The new fields are related to \mathbf{E} and \mathbf{B} via the constitutive relations $\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P}$ and $\mathbf{H} = \mu_0^{-1}\mathbf{B} - \mathbf{M}$, where the polarisation field \mathbf{P} and the magnetisation field \mathbf{M} describe the material. We then describe waves in dielectrics and conductors, leading to the index of refraction and the skin depth, respectively. Using a simple model for a dielectric, we explain the frequency dependence of the index of refraction.

You should master the following techniques:

1. calculate \mathbf{P} and \mathbf{M} for simple charge and current distributions;
2. calculate \mathbf{D} and \mathbf{H} from the constitutive relations;
3. calculate the skin depth of waves in conductors;
4. calculate the plasma frequency in plasmas;
5. explain frequency dependence of the index of refraction and anomalous dispersion.

9 Surfaces, Wave Guides and Cavities

Electric and magnetic fields are often confined in waveguides or cavities. In this section we will determine the boundary conditions of the fields at the surface between two materials. We then consider waves confined in one dimension, which leads to the concept of reflection and transmission. Waves confined in two dimensions create waveguides, and three-dimensional confinement of waves leads to the concept of cavities.

9.1 Boundary conditions for fields at a surface

The boundary conditions for the fields \mathbf{E} , \mathbf{D} , \mathbf{B} and \mathbf{H} are entirely determined by Maxwell's equations in macroscopic media. In particular, each Maxwell equation gives a boundary condition. Let the surface between two materials (1) and (2) be parallel to the xy -plane at $z = 0$ (with material 1 in the region $z < 0$). We start with the equation for the displacement field

$$\nabla \cdot \mathbf{D} = \rho_f. \quad (9.1)$$

The component \mathbf{D}_\perp that is *normal* to the surface then obeys $\partial_z D_z = \sigma_f$, where σ_f is the free surface charge. We can compare \mathbf{D}_\perp on either side of the surface at $z = \pm\epsilon$, such that

$$\partial_z D_z \hat{\mathbf{k}} \simeq \left[D_z^{(2)}(\epsilon) - D_z^{(1)}(-\epsilon) \right] \hat{\mathbf{k}} \equiv \Delta \mathbf{D}_\perp. \quad (9.2)$$

Note that \mathbf{D}_\perp is a vector, while D_z is just a component. When there is no free surface charge ($\sigma_f = 0$) the normal part of \mathbf{D} is continuous ($\Delta \mathbf{D}_\perp = 0$). In general, the boundary condition for \mathbf{D} is that

$$\Delta \mathbf{D}_\perp = \sigma_f \hat{\mathbf{k}}. \quad (9.3)$$

This boundary condition determines $\Delta \mathbf{E}_\perp$ via the constitutive relation between \mathbf{D} and \mathbf{E} .

The boundary condition for \mathbf{E}_\parallel is determined by Maxwell's equation for $\nabla \times \mathbf{E}$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}. \quad (9.4)$$

We can integrate this equation over a surface S that intersects the boundary between the two materials, and use Stokes' theorem to turn the surface integrals into line integrals:

$$\oint \mathbf{E} \cdot d\mathbf{l} = - \int \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{a}. \quad (9.5)$$

Next, we take the contour as a rectangle perpendicular to the surface and sticking half-way into the surface (see Fig. 30). The parallel sides of the loop have length l , and the sides perpendicular to the surface have length Δz . We therefore have

$$l(\mathbf{E}_\parallel^{(1)} - \mathbf{E}_\parallel^{(2)}) = -l\Delta z \frac{\partial B}{\partial t} = 0, \quad (9.6)$$

where the last equality follows from taking the limit $\Delta z \rightarrow 0$. We therefore have the second boundary condition

$$\Delta \mathbf{E}_\parallel \equiv \mathbf{E}_\parallel^{(1)} - \mathbf{E}_\parallel^{(2)} = 0, \quad (9.7)$$

which determines $\Delta \mathbf{D}_\parallel$ via the constitutive relation between \mathbf{D} and \mathbf{E} .

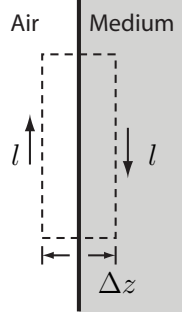


Figure 30: Transverse boundary conditions on a surface.

The boundary condition from the equation $\nabla \cdot \mathbf{B} = 0$ follows along the same lines as that for $\nabla \cdot \mathbf{D}$, but this time there is never a surface charge (since there are no magnetic charges). We immediately deduce that

$$\Delta \mathbf{B}_{\perp} = 0. \quad (9.8)$$

This determines $\Delta \mathbf{H}_{\perp}$ via the constitutive relation between \mathbf{B} and \mathbf{H} .

Finally, we consider the Maxwell equation

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J}_f. \quad (9.9)$$

Integrating over a surface intersecting the two materials and using Stokes' theorem then yields

$$\oint \mathbf{H} \cdot d\mathbf{l} = \int \frac{\partial \mathbf{D}}{\partial t} \cdot d\mathbf{a} + K_f, \quad (9.10)$$

where K_f is the free surface current. Taking again the contour shown in Fig. (30) we find that

$$\Delta \mathbf{H}_{\parallel} = K_f. \quad (9.11)$$

Again, $\Delta \mathbf{B}_{\parallel}$ is determined via the constitutive relation between \mathbf{B} and \mathbf{H} . In summary, the general boundary conditions for the electromagnetic fields at a surface are given by

$$\begin{aligned} \Delta \mathbf{E}_{\parallel} &= 0 & \text{and} & & \Delta \mathbf{D}_{\perp} &= \sigma_f, \\ \Delta \mathbf{B}_{\perp} &= 0 & \text{and} & & \Delta \mathbf{H}_{\parallel} &= K_f. \end{aligned} \quad (9.12)$$

9.2 Reflection and transmission of waves at a surface

One of the most common phenomena is the reflection and transmission of waves at a surface, for example light travelling through glass or reflecting off a mirror. We will now calculate the reflection and transmission coefficients for electromagnetic waves hitting dielectric and conducting surfaces. The fields of the incoming waves carry subscript I , the transmitted wave fields carry subscript T , and the reflected waves carry subscript R .

Dielectrics

For the reflection and transmission of waves at the boundary between two simple dielectrics we use the constitutive relations

$$\mathbf{D} = \epsilon^{(j)} \mathbf{E} \quad \text{and} \quad \mathbf{H} = \frac{1}{\mu^{(j)}} \mathbf{B}, \quad (9.13)$$

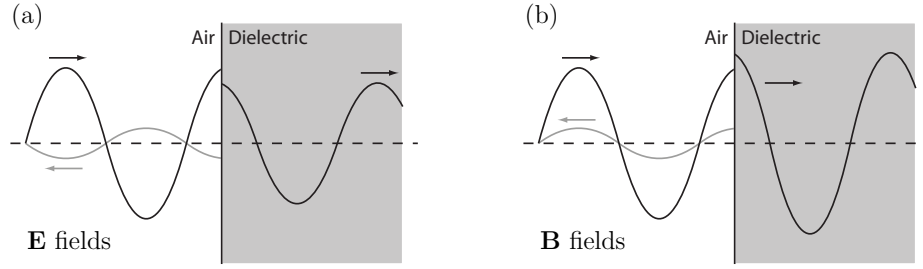


Figure 31: Waves incident on a dielectric. The waves propagate in the dielectric. (a) The electric field is mostly transmitted, and the reflected wave experiences a phase shift. (b) The magnetic field is mostly transmitted, and the reflected wave does not experience a phase shift.

where $j = 1, 2$ denotes the medium. In addition, we can determine the direction of the \mathbf{B} field by constructing $\mathbf{B} = \mathbf{k} \times \mathbf{E}/\omega$ and noting that the frequency is the same in the two media. We can furthermore select the normal and transverse components by taking the dot product and the cross product with the normal vector $\hat{\mathbf{n}}$. The boundary conditions in the absence of surface charges and currents then become

$$\begin{aligned}
 \Delta \mathbf{D}_n : & \quad \left[\epsilon^{(1)}(\mathbf{E}_I + \mathbf{E}_R) - \epsilon^{(2)}\mathbf{E}_T \right] \cdot \hat{\mathbf{n}} = 0 \\
 \Delta \mathbf{B}_n : & \quad [\mathbf{k}_I \times \mathbf{E}_I + \mathbf{k}_R \times \mathbf{E}_R - \mathbf{k}_T \times \mathbf{E}_T] \cdot \hat{\mathbf{n}} = 0 \\
 \Delta \mathbf{E}_t : & \quad [\mathbf{E}_I + \mathbf{E}_R - \mathbf{E}_T] \times \hat{\mathbf{n}} = 0 \\
 \Delta \mathbf{H}_t : & \quad \left[\frac{1}{\mu^{(1)}}\mathbf{k}_I \times \mathbf{E}_I + \frac{1}{\mu^{(1)}}\mathbf{k}_R \times \mathbf{E}_R - \frac{1}{\mu^{(2)}}\mathbf{k}_T \times \mathbf{E}_T \right] \times \hat{\mathbf{n}} = 0.
 \end{aligned} \tag{9.14}$$

When a beam is hitting the surface at a 90 degree angle, the normal vector $\hat{\mathbf{n}}$ is always perpendicular to the fields (since electromagnetic waves are transverse waves). In this special case we only deal with $\Delta \mathbf{E}_t = 0$ and $\Delta \mathbf{H}_t = 0$. For a dielectric with ϵ_r and $\mu_r = 1$, we write $\mathbf{k}_I = -\mathbf{k}_R = \sqrt{\epsilon_r}\mathbf{k}_T$, and the boundary conditions become

$$\mathbf{E}_I + \mathbf{E}_R = \mathbf{E}_T \quad \text{from} \quad \Delta \mathbf{E}_t = 0$$

$$\mathbf{E}_I - \mathbf{E}_R = \frac{\mathbf{E}_T}{\sqrt{\epsilon_r}} \quad \text{from} \quad \Delta \mathbf{H}_t = 0. \tag{9.15}$$

Eliminating \mathbf{E}_T , we have

$$\frac{E_R}{E_I} = \frac{\sqrt{\epsilon_r} - 1}{\sqrt{\epsilon_r} + 1} = \frac{n - 1}{n + 1}. \tag{9.16}$$

The reflectivity R of a surface can be defined as the ratio of the reflective and incident Poynting vectors. Since the Poynting vector is proportional to E^2 in a dielectric medium ($B = nkE/\omega$), the reflectivity becomes

$$R = \left(\frac{n - 1}{n + 1} \right)^2. \tag{9.17}$$

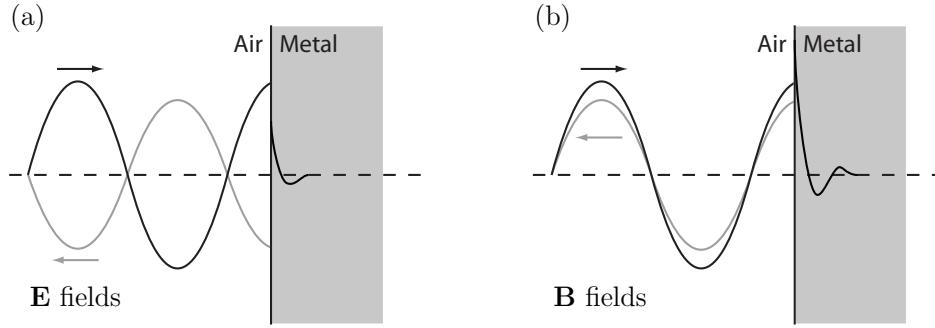


Figure 32: Waves incident on a metal. (a) The electric field is mostly reflected with a phase shift. (b) The magnetic field is mostly reflected, and does not have a phase shift. There is no propagation of the wave in the metal, and the exponential fall-off of the amplitudes is characterized by the *skin depth*.

For glass, the index of refraction is $n = 1.6$, so the normal reflection is about ten percent.

Note that the reflected \mathbf{B} field picks up a relative minus sign. This is because the Poynting vector must point in the opposite way to the incident wave, and from Eq. (8.25) we know that the magnetic induction must gain in relative strength in the medium.

Exercise 9.1: Calculate the reflectivity for a wave coming in at an angle θ to the normal of the surface.

Conductors

The reflectivity of metals is high, and consequently the transmittance T is low ($R + T = 1$ from energy conservation). Using again $B = kE/\omega$, and the boundary conditions of the previous section, we once more use

$$E_I - E_R = E_T \quad \text{and} \quad B_I + B_R = \frac{B_T}{\mu_r}, \quad (9.18)$$

and find

$$\frac{E_T}{E_I} = \frac{2}{1+n} \simeq \frac{2}{n}, \quad (9.19)$$

where n is now a large complex index of refraction

$$n = \sqrt{\frac{i\mu_r\sigma}{\omega\epsilon_0}}. \quad (9.20)$$

In addition, we can solve Eq. (9.18) for the magnetic induction, which yields

$$\frac{B_T}{B_I} = \frac{2}{1+1/n} \simeq 2. \quad (9.21)$$

The ratio of the Poynting vectors for the transmitted and incident waves is therefore

$$\frac{S_T}{S_I} = \text{Re} \frac{4}{n}. \quad (9.22)$$

This leads to the transmittance

$$T = \sqrt{\frac{8\omega\epsilon_0}{\sigma\mu_r}}. \quad (9.23)$$

Usually, the transmitted waves will be absorbed completely unless the metal is sufficiently thin. When $\sigma \rightarrow \infty$ such as in a superconductor, the transmittance drops to zero and everything is reflected.

9.3 Wave guides

In this section, we consider waveguides with conduction walls. Since there are no free charges or currents inside the waveguide, Maxwell's equations can be written as

$$\nabla \cdot \mathbf{E} = 0 \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad \nabla \cdot \mathbf{B} = 0 \quad \nabla \times \mathbf{B} = \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t}. \quad (9.24)$$

We are interested in the behaviour of monochromatic waves of frequency ω and wave number k inside the waveguide, so we assume that the fields take the generic form

$$\mathbf{E} = \mathbf{E}_0 e^{ikz - i\omega t} \quad \text{and} \quad \mathbf{B} = \mathbf{B}_0 e^{ikz - i\omega t}. \quad (9.25)$$

In terms of the components E_x, E_y, E_z of \mathbf{E}_0 and B_x, B_y, B_z of \mathbf{B}_0 , the two equations for the curl become

$$\begin{aligned} \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} &= i\omega B_z & \frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y} &= -\frac{i\omega}{c^2} E_z, \\ \frac{\partial E_z}{\partial y} - ikE_y &= i\omega B_x & \frac{\partial B_z}{\partial y} - ikB_y &= -\frac{i\omega}{c^2} E_x, \\ ikE_x - \frac{\partial E_z}{\partial x} &= i\omega B_y & ikB_x - \frac{\partial B_z}{\partial x} &= -\frac{i\omega}{c^2} E_y. \end{aligned} \quad (9.26)$$

Note that we do not set $\omega = ck$, since the walls of the waveguide may impose a different dispersion relation. We can express E_x, E_y, B_x , and B_y in terms of E_z and B_z , yielding

$$\begin{aligned} E_x &= \frac{i}{(\omega/c)^2 - k^2} \left(k \frac{\partial E_z}{\partial x} + \omega \frac{\partial B_z}{\partial y} \right), \\ E_y &= \frac{i}{(\omega/c)^2 - k^2} \left(k \frac{\partial E_z}{\partial y} - \omega \frac{\partial B_z}{\partial x} \right), \\ B_x &= \frac{i}{(\omega/c)^2 - k^2} \left(k \frac{\partial B_z}{\partial x} - \frac{\omega}{c^2} \frac{\partial E_z}{\partial y} \right), \\ B_y &= \frac{i}{(\omega/c)^2 - k^2} \left(k \frac{\partial B_z}{\partial y} + \frac{\omega}{c^2} \frac{\partial E_z}{\partial x} \right). \end{aligned} \quad (9.27)$$

Substituting these equations into $\nabla \cdot \mathbf{E} = 0$ and $\nabla \cdot \mathbf{B} = 0$ then gives the differential equations

$$\begin{aligned} \left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \left(\frac{\omega}{c} \right)^2 - k^2 \right] E_z &= 0, \\ \left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \left(\frac{\omega}{c} \right)^2 - k^2 \right] B_z &= 0. \end{aligned} \quad (9.28)$$

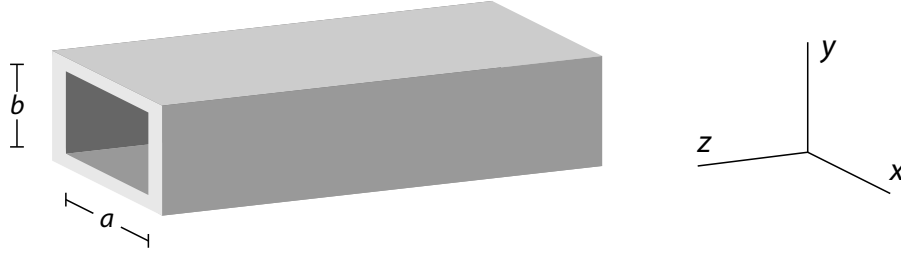


Figure 33: A rectangular hollow wave guide with conducting walls.

The other components of \mathbf{E}_0 and \mathbf{B}_0 can then be obtained from E_z and B_z by differentiating according to Eq. (9.27).

Exercise 9.2: Derive Eq. (9.28).

Since the waveguide is oriented in the z -direction, one would ordinarily expect the longitudinal E_z and B_z components to vanish. However, inside a waveguide we can have the situation that the wave is not travelling straight, but bounces off the surfaces at an angle. In this case E_z and B_z are not necessarily zero. We classify electromagnetic waves in a waveguide by their transverse properties. A transverse electric (TE) wave has no E_z component, and a transverse magnetic (TM) wave has no B_z component. If both $E_z = 0$ and $B_z = 0$, we speak of TEM waves. These cannot occur in a hollow waveguide, but they can occur in waveguides with a conducting core, such as a coaxial cable.

The rectangular hollow waveguide

As an example, consider TE waves in a rectangular waveguide of sides a and $b < a$ shown in Fig. 33 with perfectly conducting walls. This means that the fields inside the walls are zero, and the boundary conditions $\Delta \mathbf{E}_{||} = 0$ and $\Delta \mathbf{B}_{\perp} = 0$ become

$$\mathbf{E}_{||} = 0 \quad \text{and} \quad \mathbf{B}_{\perp} = 0. \quad (9.29)$$

There will be induced charges and currents on the surface that are hard to calculate, so we see how far we can get with these two boundary conditions. We have to solve Eq. (9.28) with $E_z = 0$, and we proceed by separation of variables with the *Ansatz* $B_z(x, y) = X(x)Y(y)$, such that

$$Y \frac{d^2 X}{dx^2} + X \frac{d^2 Y}{dy^2} + \left(\frac{\omega^2}{c^2} - k^2 \right) XY = 0. \quad (9.30)$$

As we have seen, this is equivalent to solving the two ordinary differential equations

$$\frac{1}{X} \frac{d^2 X}{dx^2} = -k_x^2 \quad \text{and} \quad \frac{1}{Y} \frac{d^2 Y}{dy^2} = -k_y^2, \quad (9.31)$$

subject to the constraint

$$-k_x^2 - k_y^2 + \frac{\omega^2}{c^2} - k^2 = 0. \quad (9.32)$$

The standard solution is

$$X(x) = A \sin(k_x x) + B \cos(k_x x) \quad \text{and} \quad Y(y) = C \sin(k_y y) + D \cos(k_y y). \quad (9.33)$$

The boundary conditions require that $B_x = 0$ at $x = 0$ and $x = a$, and $B_y = 0$ at $y = 0$ and $y = b$. To find the effect of the boundary conditions on A and B , we substitute Eq. (9.33) into Eq. (9.27) and set $B_x = 0$. We find that $dX/dx = 0$, and therefore $A = 0$ in Eq. (9.33). This leads to specific values of k_x and k_y :

$$k_x = \frac{m\pi}{a} \quad \text{and} \quad k_y = \frac{n\pi}{b}, \quad (9.34)$$

where m and n are integers. The solution for B_z then becomes

$$B_z = B_0 \cos\left(\frac{m\pi x}{a}\right) \cos\left(\frac{n\pi y}{b}\right). \quad (9.35)$$

Each pair (m, n) forms a particular solution of the field in a waveguide, and these are called the TE_{mn} modes. By convention, the first index refers to the largest side of the rectangle (here a).

The wavenumber of the waves travelling inside the waveguide are found by substituting Eq. (9.34) into Eq. (9.32):

$$k = \sqrt{\frac{\omega^2}{c^2} - \frac{m^2\pi^2}{a^2} - \frac{n^2\pi^2}{b^2}}. \quad (9.36)$$

If the frequency of the wave is smaller than a certain threshold, the wave number k becomes imaginary ($k = i\kappa$), and the solution does not propagate throughout the waveguide:

$$\mathbf{B} = \mathbf{B}_0 e^{-\kappa z} e^{-i\omega t}. \quad (9.37)$$

This happens at the cut-off frequencies

$$\omega_{mn} = c\pi \sqrt{\frac{m^2}{a^2} + \frac{n^2}{b^2}}. \quad (9.38)$$

The lowest cut-off frequency is $\omega_{10} = c\pi/a$. No frequencies smaller than this value can propagate through the waveguide. The wave number for the TE_{mn} mode can then be written as

$$k = \frac{1}{c} \sqrt{\omega^2 - \omega_{mn}^2}, \quad (9.39)$$

and the phase velocity v_p is calculated as

$$v_p = \frac{\omega}{k} = \frac{c}{\sqrt{1 - (\omega_{mn}/\omega)^2}}. \quad (9.40)$$

This is greater than c , since $\omega > \omega_{mn}$ for propagating waves! Fortunately, the group velocity v_g is well-behaved:

$$v_g = \frac{d\omega}{dk} = c \sqrt{1 - (\omega_{mn}/\omega)^2} < c. \quad (9.41)$$

This is the velocity of the energy in the wave.

Exercise 9.3: Calculate the surface charges and currents from Eq. (9.35).

The coaxial transmission line

As a second example, consider a coaxial cable. The first thing to note is that such waveguides do admit TEM waves, since the core of the waveguide can be held at a different potential from

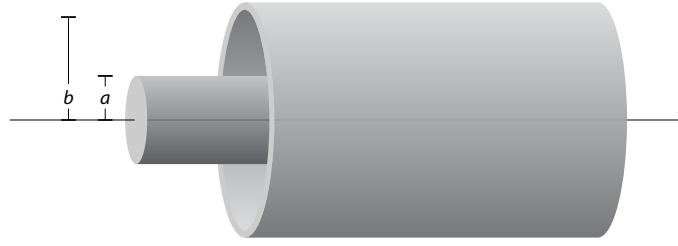


Figure 34: The coaxial transmission cable.

the outer wall. In this case, the two Maxwell's equations in Eq. (9.26) lead to the ordinary wave number

$$k = \frac{\omega}{c}, \quad (9.42)$$

so the waves behave just as free-space waves: there is no dispersion and they travel at the speed of light in the medium between the core and the outer wall (c in the case of vacuum). Specifically, the coaxial cable does not act as a filter with some cut-off frequency, and is therefore particularly suited for the transmission of a wide variety of wave signals.

The fields inside the coaxial cable obey the standard relations $cB_x = -E_y$ and $cB_y = E_x$, and Maxwell's equations become

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} = 0 \quad \text{and} \quad \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = 0. \quad (9.43)$$

In cylindrical coordinates (ρ, ϕ, z) this leads to the fields

$$\begin{aligned} \mathbf{E}(\rho, \phi, z, t) &= \frac{E_0}{\rho} \cos(kz - \omega t) \hat{\rho}, \\ \mathbf{B}(\rho, \phi, z, t) &= \frac{E_0}{c\rho} \cos(kz - \omega t) \hat{\phi}. \end{aligned} \quad (9.44)$$

A coaxial cable has surface currents and charges in the core and the wall, so it will have both a capacitance C and inductance L . These can be calculated as

$$C = \frac{2\pi\epsilon}{\ln(b/a)} \quad \text{and} \quad L = \frac{\mu}{2\pi} \ln\left(\frac{b}{a}\right). \quad (9.45)$$

We can use these to define the *impedance* Z as

$$Z = \sqrt{\frac{L}{C}} = 2\pi \sqrt{\frac{\mu}{\epsilon}} \ln\left(\frac{b}{a}\right). \quad (9.46)$$

If the coaxial cable is plugged into a receiver that has a different impedance, part of the signal will be reflected back into the cable. This is called *impedance mismatch*. The units of impedance are Ohms (Ω).

9.4 Resonant cavities

Imagine that we again have the rectangular waveguide with width a and height b , but this time we close off the waveguide with a perfectly conducting wall at $z = 0$ and $z = d$. What are the TE modes?

To answer this question, we note that we can use the solutions for the rectangular waveguide from the previous section, with the TE_{mn} mode

$$B_z = B_0 \cos\left(\frac{m\pi x}{a}\right) \cos\left(\frac{n\pi y}{b}\right) e^{ikz - i\omega t}. \quad (9.47)$$

We now have to impose additional boundary conditions in the z -direction, noting that $\mathbf{B}_\perp = 0$ at the boundary. This means that the magnetic field becomes

$$B_z \rightarrow B'_z = B_0 \cos\left(\frac{m\pi x}{a}\right) \cos\left(\frac{n\pi y}{b}\right) \sin\left(\frac{p\pi z}{d}\right) e^{-i\omega t}, \quad (9.48)$$

and the cavity modes are standing waves with frequencies

$$\omega_{mnp} = c\pi \sqrt{\frac{m^2}{a^2} + \frac{n^2}{b^2} + \frac{p^2}{d^2}}. \quad (9.49)$$

These results are valid for perfectly conducting walls that do not allow fields to penetrate.

In real cavities, the walls do not have perfect conductivity, and fields will propagate into the walls. The amplitude of the fields drops off with a length scale given by the skin depth, and the fields can leak out of the cavity into the walls (and possibly beyond). One of the most important characteristics of a cavity is therefore how well the cavity retains the fields inside. Since the fields carry energy, one way to define the quality of a cavity is to set up an energy balance. For a wave with frequency ω_0 , the quality factor, or Q factor of the cavity is given by

$$Q = \omega_0 \frac{\text{EM energy in the cavity}}{\text{Energy loss per cycle}}. \quad (9.50)$$

Since the energy loss per cycle is just the negative time derivative of the energy we can write

$$\frac{dU}{dt} = -\frac{\omega_0}{Q} U, \quad (9.51)$$

where U is the energy in the cavity. We can solve this straight away:

$$U(t) = U_0 e^{-\omega_0 t/Q}, \quad (9.52)$$

which is an exponential damping of the waves in the cavity. The higher the value of Q , the smaller the damping. The electric field in the cavity will then have the form

$$\mathbf{E}(t) = \mathbf{E}_0 e^{-\omega_0 t/2Q} e^{-i(\omega_0 + \Delta\omega)t}, \quad (9.53)$$

where we included a shift in the frequency $\Delta\omega$ that typically accompanies damping. We want to look at the spectrum of the electric field $\mathbf{E}(\omega)$, so we take the Fourier transform of $\mathbf{E}(t)$:

$$\mathbf{E}(\omega) = \frac{\mathbf{E}_0}{\sqrt{2\pi}} \int_0^\infty e^{-\omega_0 t/2Q} e^{-i(\omega_0 + \Delta\omega - \omega)t} dt. \quad (9.54)$$

The electrical energy density in the cavity is equal to $\frac{1}{2}\epsilon_0 |\mathbf{E}(\omega)|^2$, which can now be calculated as

$$\frac{1}{2}\epsilon_0 |\mathbf{E}(\omega)|^2 = \frac{\epsilon_0 |\mathbf{E}_0|^2}{4\pi} \frac{1}{(\omega - \omega_0 - \Delta\omega)^2 + (\omega_0/2Q)^2}, \quad (9.55)$$

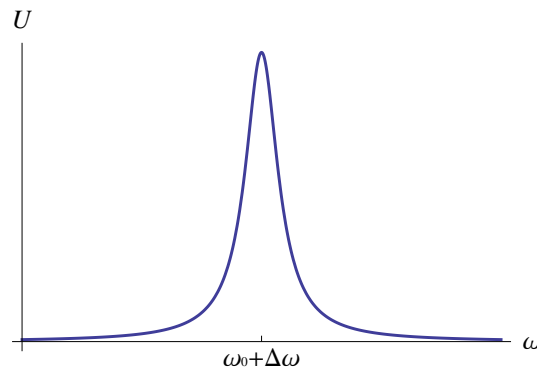


Figure 35: The spectrum of a cavity with finite Q factor.

This spectrum is shown in Fig. (35). Narrower shapes indicate better cavities. You can determine Q from the plot by dividing ω_0 with the full-width at half maximum Γ .

The higher the Q , the longer the field remains in the cavity, and the stronger the interaction between the field and an atom, ion or quantum dot placed in the cavity. Values of $Q = 10^9$ have been achieved in microwave cavities with superconducting walls. These cavities are good enough to make single photons interact very strongly with Rydberg atoms. In optical cavities, values on the order of $Q = 10^3$ to 10^4 can be realistically achieved.

The Q factor depends on the shape of the cavity, the permeability of the interior cavity μ_c and the walls μ , and the skin depth δ . The Q factor is proportional to the relative strength of the permeabilities μ/μ_c . From Eq. (9.50) the energy stored in the cavity is proportional to the volume, while the energy loss per cycle is proportional to the surface (though which the energy must flow) multiplied by the skin depth. The Q factor is therefore also proportional to the ratio $V/S\delta$. Finally, there is a form factor f that depends on the specific geometry of the cavity:

$$Q = f \frac{\mu}{\mu_c} \frac{V}{S\delta}, \quad (9.56)$$

Smaller skin depth (higher conductivity) and maximised volume to surface ratio (i.e., more spherical cavities) will help obtain a higher Q .

An interesting application of cavity theory is the Earth's ionosphere. The ionosphere is a plasma, and consequently for low enough frequencies (long wavelengths) it acts as a mirror. Since the earth is also a conductor, low frequency radio waves (AM) can bounce between the ionosphere and the Earth to reach much further than high frequency (FM) radiowaves. The plasma frequency of the ionosphere is about 3 MHz.

Summary

In this section we studied the effects of surface boundaries on electromagnetic fields. We found that waves are reflected and transmitted at boundaries, and used this to construct wave guides and cavities. Maxwell's equations determine the boundary conditions of the waves at the surfaces, and this allowed us to calculate reflection and transmission coefficients. Inside waveguides the transversality of electromagnetic waves no longer holds necessarily, and we defined TE, TM, and TEM waves. We also found that waveguides only transmit certain frequencies. Finally, we showed how we can determine the quality Q of an electromagnetic cavity.

You should master the following techniques:

1. apply the boundary conditions to work out the reflected and transmitted parts of the electric and magnetic fields;
2. solve Maxwell's equations for waveguides using the boundary conditions, and determine the cutoff frequency of the waveguide;
3. explain the concept of impedance matching;
4. estimate the Q factor for cavities from the cavity materials and from the cavity spectrum.

10 Relativistic formulation of electrodynamics

One of the remarkable things about the Maxwell equations is that they give rise to a correct description of all sorts of radiation, most notably light. In particular, it predicts the correct velocity of light in vacuum via the relation $c^2 = (\mu_0 \epsilon_0)^{-1}$, as we have seen in Section 5. Experimentally it was found by Michelson and Morley in 1887 that the velocity of light is independent of the velocity of the source. This caused Lorentz quite a headache, which ultimately resulted in the Lorentz transformations for moving bodies: Objects that move with respect to an inertial observer are seen to experience a length contraction in the direction of their motion, and their clocks seem to run slower.

There are similar rules for the electromagnetic field, meaning that what one inertial observer calls the \mathbf{E} field, a second observer may describe in terms of both \mathbf{E} and \mathbf{B} fields. We therefore want to know two things: (1) How do the fields transform from one coordinate system to another, and (2) What do Maxwell's equations look like when written in covariant form, that is, independent of the coordinate system? We will first revise the basics of special relativity and Minkowski space, and then we will construct Maxwell's equations from the vector and scalar potentials, and the fields. This leads naturally to the transformation laws for the \mathbf{E} and \mathbf{B} fields. We end with a look at some invariant properties. This section will rely heavily on the transformation properties of vectors and tensors, and everything will be in index notation.

10.1 Four-vectors and transformations in Minkowski space

Remember that covariant descriptions take place in four-dimensional Minkowski space. There is the position four-vector x^μ , where $\mu = 0, 1, 2, 3$, $\mathbf{r} = (x^1, x^2, x^3)$ or $\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$, and $x^0 = ct$, and the momentum four-vector p^μ with $\mathbf{p} = (p^1, p^2, p^3)$ and $p^0 = U/c$. Minkowski space is a strange place in that it has a non-trivial *metric*: Ordinarily, if you want to find the length of a (short) interval, you add all the components squared, according to Pythagoras:

$$ds^2 = dx^2 + dy^2 + dz^2. \quad (10.1)$$

However, in relativity the length changes in different reference frames, and we need to include the change in the time coordinate as well. It is tempting to just add $c^2 dt^2$ to Eq. (10.1), but that would be wrong! The correct distance between two events is

$$ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2, \quad (10.2)$$

with a *minus* sign in front of $c^2 dt^2$! This is the *definition* of the distance in Minkowski space, and we can rewrite this as a vector equation (and also using Einstein's summation convention):

$$ds^2 = (ct, x, y, z) \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \equiv x^\mu g_{\mu\nu} x^\nu = x^\mu x_\mu = x_\mu x^\mu, \quad (10.3)$$

where the four-vector x^μ can be written as (ct, x, y, z) . Notice that the matrix $g_{\mu\nu}$ has a “-1” on the diagonal, which is responsible for the minus sign in the metric. You also see that there is a difference between upper and lower indices. The four-vectors with upper indices are called *contravariant*, and the four-vectors with lower indices are called *covariant*. You can verify from Eq. (10.3) that the covariant four-vector has a relative minus sign in the first component:

$$x_\mu = (-ct, x, y, z). \quad (10.4)$$

We are cheating a bit here, because the x_μ on the left is a *component*, while the quantity on the right is a proper vector. However, this expression is to emphasize the effect of the vertical position of the index; it is not a proper equation. As you can see from Eq. (10.3), the metric raises or lowers the index, and $g_{\mu\nu} = g^{\mu\nu}$ in the case of special relativity.

We now amend the Einstein summation convention: the sum is implied over two repeating indices, one of which is upper, and the other is lower. Carrying out the sum is called *contraction*. We can construct invariant quantities (scalars) by contracting contravariant four-vectors with covariant four-vectors. In general we write the components of a four-vector a^μ as (a^0, a^1, a^2, a^3) , and we have:

$$\begin{aligned} a^\mu b_\mu &= a^\mu g_{\mu\nu} b^\nu = \sum_{\mu,\nu} a^\mu g_{\mu\nu} b^\nu \\ &= -a^0 b^0 + \sum_{j=1}^3 a^j b^j = -a^0 b^0 + \mathbf{a} \cdot \mathbf{b}, \end{aligned} \quad (10.5)$$

where greek indices sum over all four components, while roman indices sum only over the spatial part⁶.

As an example, let's look at the four-momentum of a particle: $p = (U/c, p^1, p^2, p^3)$, and $p^2 = p^\mu p_\mu$ can be written as $p^2 = -U^2/c^2 + |\mathbf{p}|^2$. Since this is true for all inertial frames, it is true for the frame where the particle is at rest, so $|\mathbf{p}|^2 = 0$. Using $U = mc^2$ with m the rest mass of the particle, we find $p^2 = -m^2 c^2$ (yes, in Minkowski space a square can be negative without complex numbers!), and the energy of a particle is therefore

$$U = \sqrt{m^2 c^4 + |\mathbf{p}|^2 c^2}. \quad (10.6)$$

This is one of the most important formulas in physics.

Any four-vector a^μ is transformed into $a^{\mu'}$ due to a Lorentz transformation $\Lambda^{\mu'}_\mu$ via

$$a^{\mu'} = \Lambda^{\mu'}_\mu a^\mu = \sum_\mu \Lambda^{\mu'}_\mu a^\mu. \quad (10.7)$$

Note that the primed frame of reference is indicated by a primed index μ' . The Lorentz transformation involves a contraction over the old (unprimed) coordinates in order to remove them from the equation. In general, a tensor transforms as

$$T^{\mu'\nu'} = \Lambda^{\mu'}_\mu \Lambda^{\nu'}_\nu T^{\mu\nu}. \quad (10.8)$$

Every component is transformed with a separate Lorentz transformation $\Lambda^{\mu'}_\mu$. The actual form of the Lorentz transformation, for example for a boost in the z direction, is

$$\Lambda^{\mu'}_\mu = \begin{pmatrix} \gamma & 0 & 0 & -\gamma\beta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\gamma\beta & 0 & 0 & \gamma \end{pmatrix}, \quad (10.9)$$

⁶You should be aware of the fact that the metric $g_{\mu\nu}$ can also be written as a diagonal matrix with three -1 entries for the spatial part and one $+1$ for the temporal part. This does not lead to observable differences in the theory, so it is a convention. Unfortunately, both conventions are used regularly, so make sure you know what the metric is before you copy a relativistic formula from a book!

with $\gamma = 1/\sqrt{1-\beta^2}$ and $\beta = v/c$. In component notation the Lorentz transformation $x^{\mu'} = \Lambda^{\mu'}_{\mu} x^{\mu}$ becomes

$$\begin{aligned} x^{0'} &= \gamma x^0 - \gamma \beta x^3 = \frac{1}{\sqrt{1-v^2/c^2}} \left(ct - \frac{vz}{c} \right), \\ x^{1'} &= x^1 = x, \\ x^{2'} &= x^2 = y, \\ x^{3'} &= -\gamma \beta x^0 + \gamma x^3 = \frac{1}{\sqrt{1-v^2/c^2}} (z - vt). \end{aligned} \quad (10.10)$$

Note that the Lorentz transformation is a proper coordinate transformation:

$$\Lambda^{\mu'}_{\mu} = \frac{\partial x^{\mu'}}{\partial x^{\mu}}. \quad (10.11)$$

Using these rules it is straightforward (but somewhat lengthy) to find the transformation rule for any tensor.

10.2 Covariant Maxwell equations

Just like the position and momentum four-vectors, we would like to construct electric and magnetic four-vectors. Unfortunately, it is not that simple: There are six field components (three for \mathbf{E} and three for \mathbf{B}), and we cannot force them into four-vectors. On the other hand, we *can* combine the vector potential and the scalar potential into a fourvector A^{μ} :

$$A^{\mu}(x) = \left(\frac{\Phi(x)}{c}, \mathbf{A}(x) \right), \quad (10.12)$$

where x is again the position four-vector. The second four-vector that we can construct straight away is the current density j^{μ} :

$$j^{\mu}(x) = (c\rho(x), \mathbf{J}(x)). \quad (10.13)$$

Since the Maxwell equations are differential equations, we also need to construct a four-vector out of the differential operators. Again, they come in two variations, covariant ∂_{μ} and contravariant ∂^{μ} :

$$\partial^{\mu} = \frac{\partial}{\partial x_{\mu}} = \left(-\frac{1}{c} \frac{\partial}{\partial t}, \nabla \right) \quad \text{and} \quad \partial_{\mu} = \frac{\partial}{\partial x^{\mu}} = \left(\frac{1}{c} \frac{\partial}{\partial t}, \nabla \right). \quad (10.14)$$

These operators transform just like ordinary vectors (but note the vertical position of the indices). We can construct an invariant with these operators if we contract them:

$$\partial_{\mu} \partial^{\mu} = \partial_x^2 + \partial_y^2 + \partial_z^2 - \frac{1}{c^2} \partial_t^2 = \square. \quad (10.15)$$

This is the d'Alembertian, which we first encountered in Section 3.

The continuity equation (conservation of charge) then becomes particularly compact:

$$\partial_{\mu} j^{\mu} = 0. \quad (10.16)$$

Compare this with Eq. (2.21); the two equations say exactly the same thing! The Maxwell equations in Eqs. (3.12) can be written as

$$\partial_\mu \partial^\mu A^\nu = -\mu_0 j^\nu \quad (10.17)$$

Often people say that this encompasses all of electrodynamics. However, you should remember that we also need to know how charges respond to the fields, that is, we need to know the Lorentz force. Since this force depends explicitly on the velocity of the particle, it is not in covariant form. We will consider the covariant Lorentz force later. We also see explicitly that the Maxwell equations are already relativistically correct: We did not change anything in Eqs. (3.12) and (3.13), we just rewrote everything. The compactness of Eq. (10.17) indicates that Lorentz invariance is a *symmetry* of electrodynamics (The more symmetric the object, the more compact we can make its description: For example, compare the description of a perfect sphere with the description of a sponge). Lorentz invariance is one of the most important symmetries in Nature.

Maxwell thought that the electric fields were carried by some substance that pervades all space, called the *æther*. Such a substance defines a natural reference frame, which would break Lorentz invariance. It was Einstein who disposed of the *æther* and made all inertial reference frames equivalent. Thus he uncovered that electrodynamics was a relativistic theory all along. In fact, his famous paper of 1905 on special relativity was titled “On the electrodynamics of moving bodies”.

In addition to the Maxwell equations in Eq. (10.17), we need to specify a particular gauge for the four-vector potential A^μ if we actually want to calculate anything. The relativistically invariant gauge is the Lorenz⁷ gauge

$$\partial_\mu A^\mu = 0. \quad (10.18)$$

Now that we have constructed the four-vector potential for the electromagnetic field, we can ask what the \mathbf{E} and \mathbf{B} fields look like in a relativistic setting. We can use the relations

$$\mathbf{E} = -\nabla\Phi - \frac{\partial\mathbf{A}}{\partial t} \quad \text{and} \quad \mathbf{B} = \nabla \times \mathbf{A}. \quad (10.19)$$

If we write out a few components we get, for example,

$$\begin{aligned} E_x &= -\frac{\partial\Phi}{\partial x} - \frac{\partial A_x}{\partial t} = -c(\partial^1 A^0 - \partial^0 A^1), \\ B_x &= \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} = -(\partial^2 A^3 - \partial^3 A^2). \end{aligned} \quad (10.20)$$

Obviously we can only fit all six components of the electric and magnetic field into a relativistic object if we have something with *two* indices. We call this the (anti-symmetric) field-strength tensor F :

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & -B_z & B_y \\ -E_y/c & B_z & 0 & -B_x \\ -E_z/c & -B_y & B_x & 0 \end{pmatrix}. \quad (10.21)$$

The relativistic transformation of the fields then becomes

$$F^{\mu'\nu'} = \Lambda^{\mu'}_{\mu} \Lambda^{\nu'}_{\nu} F^{\mu\nu}. \quad (10.22)$$

⁷This is not a typo: Lorenz and Lorentz were different guys.

For a boost v in the z direction given by Eq. (10.9) we find

$$E'_x = \gamma(E_x + vB_y), \quad E'_y = \gamma(E_y - vB_x), \quad E'_z = E_z \quad (10.23)$$

$$B'_x = \gamma\left(B_x - \frac{v}{c^2}E_y\right), \quad B'_y = \gamma\left(B_y + \frac{v}{c^2}E_x\right), \quad B'_z = B_z. \quad (10.24)$$

In terms of the parallel and perpendicular components of the fields with respect to the direction of motion this becomes

$$\begin{aligned} \mathbf{E}'_{\parallel} &= \mathbf{E}_{\parallel} & \mathbf{E}'_{\perp} &= \gamma\mathbf{E}_{\perp} + \gamma\mathbf{v} \times \mathbf{B} \\ \mathbf{B}'_{\parallel} &= \mathbf{B}_{\parallel} & \mathbf{B}'_{\perp} &= \gamma\mathbf{B}_{\perp} - \gamma\frac{\mathbf{v} \times \mathbf{E}}{c^2}. \end{aligned} \quad (10.25)$$

A physical situation with only a static charge distribution and no currents obviously does not involve \mathbf{B} fields. The magnetic fields arise when the charge distribution is viewed by a moving observer. In particular, the electrostatic force $\mathbf{F} = q\mathbf{E}$ acquires a term proportional to $\mathbf{v} \times \mathbf{B}$, so that $\mathbf{F}' = q\mathbf{E} + q\mathbf{v} \times \mathbf{B}$ (verify this!). This is of course the familiar Lorentz force. In other words, when an observer sees a moving electric charge, the \mathbf{E} field is not spherically symmetric due to Lorentz contraction in the direction of motion. As a result, there is a component of the electric force that acts only on moving charges. This is the magnetic field. This argument shows that you can think of the magnetic field as just the relativistic part of the electric field of moving charges. Whenever there are \mathbf{B} fields, there must be moving charges, whether it be macroscopic currents or spinning charges in atoms.

Finally the Lorentz force must be formulated in a covariant way. The force experienced by a particle with charge q , as viewed from the lab frame, is proportional to the velocity, so we need to define a four-velocity u . We can use the relation

$$p^{\mu} \equiv mu^{\mu} = m\gamma(c, \mathbf{v}) \quad (10.26)$$

such that $u_{\mu}u^{\mu} = -c^2$ (verify this!). Rather than deriving the relativistic Lorentz force, we will state it here and confirm that it behaves the way it should. Suppose that τ is the proper time of a particle, as recorded by a co-moving clock. The force will cause a momentum transfer $dp^{\mu}/d\tau$. This must be proportional to $\mathbf{E} + \mathbf{v} \times \mathbf{B}$. The cross product is built in the field strength tensor $F^{\mu\nu}$, so we guess that the covariant form of the Lorentz force is

$$m\frac{d^2x^{\mu}}{d\tau^2} = \frac{dp^{\mu}}{d\tau} = qu_{\nu}F^{\mu\nu} = qF^{\mu\nu}\frac{dx^{\nu}}{d\tau}. \quad (10.27)$$

These are the equations of motion for a relativistic particle in an electromagnetic field. Let's see what the different components of \mathbf{f} look like:

$$\begin{aligned} f_x &= \frac{dp^1}{d\tau} = qF^{1\nu}u_{\nu} = \gamma q \left(c\frac{E_x}{c} + v_x \cdot 0 + v_y B_z - v_z B_y \right) \\ f_y &= \frac{dp^2}{d\tau} = qF^{2\nu}u_{\nu} = \gamma q \left(c\frac{E_y}{c} - v_x B_z + v_y \cdot 0 + v_z B_x \right) \\ f_z &= \frac{dp^3}{d\tau} = qF^{3\nu}u_{\nu} = \gamma q \left(c\frac{E_z}{c} + v_x B_y - v_y B_x - v_z \cdot 0 \right) \end{aligned} \quad (10.28)$$

These are indeed the three components of the Lorentz force, with an extra factor γ . This is correct, because the term $dp^{\mu}/d\tau$ also has an implicit γ factor in the relativistic mass. Remains the question of what is f_t :

$$cf_t = c\frac{dp^0}{d\tau} = qcF^{0\nu}u_{\nu} = \gamma q (v_x \cdot E_x + v_y E_y + v_z E_z) = \gamma q \mathbf{v} \cdot \mathbf{E}. \quad (10.29)$$

This is the relativistic work done on the particle by the fields. So you see that it all works out.

10.3 Invariant quantities

When the fields and potentials change when viewed in different inertial frames, it is important to know what the *invariant* quantities of a theory are. These quantities are the same in all reference frames, so you can evaluate them in whichever frame makes the calculation easiest. There are many invariants, most notably the magnitude of four-vectors $a^2 = a_\mu a^\mu$. You can use the differential operator as well in the construction of invariants (see Eq. (10.13) for an important example). A quantity is invariant when it does not depend on the coordinates, so in index notation the quantity has *no* indices.

The quantity $\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ is invariant (why?), and can be evaluated as

$$\frac{1}{4}F_{\mu\nu}F^{\mu\nu} = \frac{1}{4}g_{\rho\mu}g_{\sigma\nu}F^{\rho\sigma}F^{\mu\nu} = \frac{B^2}{2} - \frac{E^2}{2c^2}. \quad (10.30)$$

This is the Lagrangian (density) of the electromagnetic field, which plays a pivotal role in particle physics.

We can also construct the *dual* of the field-strength tensor

$$\mathcal{F}^{\mu\nu} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}F_{\rho\sigma} = \begin{pmatrix} 0 & -B_x & -B_y & -B_z \\ B_x & 0 & E_z/c & -E_y/c \\ B_y & -E_z/c & 0 & E_x/c \\ B_z & E_y/c & -E_x/c & 0 \end{pmatrix}. \quad (10.31)$$

which is a pseudo-tensor. Here we used the Levi-Civita (pseudo) tensor $\epsilon^{\mu\nu\rho\sigma}$, which returns 1 if the indices make an even permutation, -1 if they make an odd permutation, and 0 otherwise. The invariant product $F_{\mu\nu}\mathcal{F}^{\mu\nu}$ is

$$F_{\mu\nu}\mathcal{F}^{\mu\nu} = 4\mathbf{E} \cdot \mathbf{B}/c. \quad (10.32)$$

Finally, the microscopic Maxwell equations can be written in terms of the fields as

$$\boxed{\partial_\mu F^{\mu\nu} = -\mu_0 j^\nu \quad \text{and} \quad \partial_\mu \mathcal{F}^{\mu\nu} = 0}. \quad (10.33)$$

These equations plus the Lorentz force encompass *all* of electrodynamics.

Summary

In this section we presented the relativistic description of electromagnetic fields. After a brief revision of special relativity, we gave the relativistically invariant form of Maxwell's equations and showed how \mathbf{E} and \mathbf{B} change under Lorentz transformations. We defined the field stress tensor $F^{\mu\nu}$, and its dual $\mathcal{F}^{\mu\nu}$, and gave the covariant form of the Lorentz force.

You should master the following techniques:

1. relativistically transform \mathbf{E} and \mathbf{B} ;
2. calculate the field stress tensor $F^{\mu\nu}$, and its dual $\mathcal{F}^{\mu\nu}$;
3. calculate the Lorentz force in different frames of reference.

NB: This section relies heavily on index notation. If you are not yet comfortable with this, go through this section by writing out every equation in terms of all its components.

A Special coordinates

Often it is much more convenient to use cylindrical or spherical coordinates, rather than cartesian coordinates. The differential operators change accordingly. For cartesian coordinates:

$$\mathbf{A} \times \mathbf{B} = (A_y B_z - A_z B_y) \hat{x} + (A_z B_x - A_x B_z) \hat{y} + (A_x B_y - A_y B_x) \hat{z} \quad (\text{A.1})$$

$$\nabla f = \frac{\partial f}{\partial x} \hat{x} + \frac{\partial f}{\partial y} \hat{y} + \frac{\partial f}{\partial z} \hat{z} \quad (\text{A.2})$$

$$\nabla \cdot \mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \quad (\text{A.3})$$

$$\nabla \times \mathbf{A} = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \hat{x} + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \hat{y} + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \hat{z} \quad (\text{A.4})$$

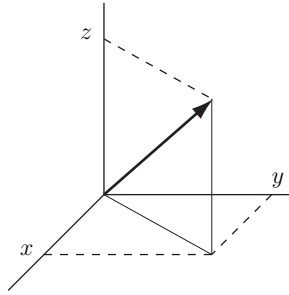


Figure 36: Cartesian coordinates (x, y, z) .

Cylindrical coordinates:

$$\nabla f = \frac{\partial f}{\partial \rho} \hat{\rho} + \frac{1}{\rho} \frac{\partial f}{\partial \phi} \hat{\phi} + \frac{\partial f}{\partial z} \hat{z} \quad (\text{A.5})$$

$$\nabla \cdot \mathbf{A} = \frac{1}{\rho} \frac{\partial(\rho A_\rho)}{\partial \rho} + \frac{1}{\rho} \frac{\partial A_\phi}{\partial \phi} + \frac{\partial A_z}{\partial z} \quad (\text{A.6})$$

$$\nabla \times \mathbf{A} = \left(\frac{1}{\rho} \frac{\partial A_z}{\partial \phi} - \frac{\partial A_\phi}{\partial z} \right) \hat{\rho} + \left(\frac{\partial A_\rho}{\partial z} - \frac{\partial A_z}{\partial \rho} \right) \hat{\phi} + \frac{1}{\rho} \left(\frac{\partial(\rho A_\phi)}{\partial \rho} - \frac{\partial A_\rho}{\partial \phi} \right) \hat{z} \quad (\text{A.7})$$

$$\nabla^2 f = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial f}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 f}{\partial \phi^2} + \frac{\partial^2 f}{\partial z^2} \quad (\text{A.8})$$

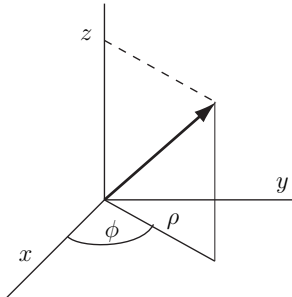


Figure 37: Cylindrical coordinates (ρ, ϕ, z) .

Spherical coordinates:

$$\nabla f = \frac{\partial f}{\partial r} \hat{r} + \frac{1}{r} \frac{\partial f}{\partial \theta} \hat{\theta} + \frac{1}{r \sin \theta} \frac{\partial f}{\partial \phi} \hat{\phi} \quad (\text{A.9})$$

$$\nabla \cdot \mathbf{A} = \frac{1}{r^2} \frac{\partial(r^2 A_r)}{\partial r} + \frac{1}{r \sin \theta} \frac{\partial(A_\theta \sin \theta)}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial A_\phi}{\partial \phi} \quad (\text{A.10})$$

$$\nabla \times \mathbf{A} = \frac{1}{r \sin \theta} \left(\frac{\partial(A_\phi \sin \theta)}{\partial \theta} - \frac{\partial A_\theta}{\partial \phi} \right) \hat{r} \quad (\text{A.11})$$

$$+ \frac{1}{r} \left(\frac{1}{\sin \theta} \frac{\partial A_r}{\partial \phi} - \frac{\partial(r A_\phi)}{\partial r} \right) \hat{\theta} + \frac{1}{r} \left(\frac{\partial(r A_\theta)}{\partial r} - \frac{\partial A_r}{\partial \theta} \right) \hat{\phi} \quad (\text{A.12})$$

$$\nabla^2 f = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 f}{\partial \phi^2} \quad (\text{A.13})$$

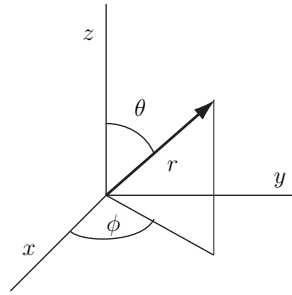


Figure 38: Spherical coordinates (r, ϕ, θ) .

B Vector identities

Having defined the divergence and the curl of vector fields, we can state the important integral theorems (without proof):

$$\int_V \nabla \cdot \mathbf{A} dV = \oint_S \mathbf{A} \cdot d\mathbf{S} \text{ (Gauss' theorem)} \quad (\text{B.1})$$

$$\int_S (\nabla \times \mathbf{A}) \cdot d\mathbf{S} = \oint_C \mathbf{A} \cdot d\mathbf{l} \text{ (Stokes' theorem)} \quad (\text{B.2})$$

$$(\text{B.3})$$

and some important vector identities, such as relations between 3-vectors:

$$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) = \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A}) = \mathbf{C} \cdot (\mathbf{A} \times \mathbf{B}) \quad (\text{B.4})$$

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C} \quad (\text{B.5})$$

$$(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) = (\mathbf{A} \cdot \mathbf{C})(\mathbf{B} \cdot \mathbf{D}) - (\mathbf{A} \cdot \mathbf{D})(\mathbf{B} \cdot \mathbf{C}) \quad (\text{B.6})$$

first derivatives:

$$\nabla \cdot (\phi \mathbf{A}) = \phi(\nabla \cdot \mathbf{A}) + \mathbf{A} \cdot (\nabla \phi) \quad (\text{B.7})$$

$$\nabla \times (\phi \mathbf{A}) = \phi(\nabla \times \mathbf{A}) + (\nabla \phi) \times \mathbf{A} \quad (\text{B.8})$$

$$\nabla(\mathbf{A} \cdot \mathbf{B}) = \mathbf{A} \times (\nabla \times \mathbf{B}) + \mathbf{B} \times (\nabla \times \mathbf{A}) + (\mathbf{A} \cdot \nabla)\mathbf{B} + (\mathbf{B} \cdot \nabla)\mathbf{A} \quad (\text{B.9})$$

$$\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{A}) - \mathbf{A} \cdot (\nabla \times \mathbf{B}) \quad (\text{B.10})$$

$$\nabla \times (\mathbf{A} \times \mathbf{B}) = (\mathbf{B} \cdot \nabla)\mathbf{A} - (\mathbf{A} \cdot \nabla)\mathbf{B} + \mathbf{A}(\nabla \cdot \mathbf{B}) - \mathbf{B}(\nabla \cdot \mathbf{A}) \quad (\text{B.11})$$

and second derivatives:

$$\nabla \cdot (\nabla \times \mathbf{A}) = 0 \quad (\text{B.12})$$

$$\nabla \times (\nabla \phi) = 0 \quad (\text{B.13})$$

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2 \mathbf{A} \quad (\text{B.14})$$

Kronecker deltas and Levi-Civita symbols:

$$\epsilon_{ijk}\epsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{il}\delta_{lm}. \quad (\text{B.15})$$

C Units in electrodynamics

Traditionally, electrodynamics is plagued by the use of many different sets of units. In these lecture notes we use SI units (*Système International*) exclusively. However, many people also use *Gaussian* units, so it is worthwhile to spend some time exploring the difference between these units. In particular, you should be able to read the set of units used off the Maxwell equations. The rule of thumb is: if they involve factors of 4π , then they most likely use Gaussian units.

Carl Friedrich Gauss was a German mathematician who in the 1830s did (among *very* many other things) worldwide measurements of the Earth's magnetic field. Not only are the Gaussian units named after him, the unit of magnetic induction in Gaussian units is also called the *Gauss* (G). In SI units the unit of magnetic induction is the *Tesla* (T), after the Serbian inventor Nikola Tesla (1853-1943).

There is a certain arbitrariness in Maxwell's equations, the Lorentz force, and the continuity equation, in the sense that we can add constants k_1, k_2, \dots to the equations that affect only the units, and not the physics:

$$k_1 \nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = 0 \quad \text{and} \quad \mathbf{F} = k_2 \rho \mathbf{E} + k_3 \mathbf{J} \times \mathbf{B} \quad (\text{C.1})$$

$$\nabla \cdot \mathbf{E} = k_4 \rho \quad \text{and} \quad \nabla \cdot \mathbf{B} = 0 \quad (\text{C.2})$$

$$\nabla \times \mathbf{E} + k_5 \frac{\partial \mathbf{B}}{\partial t} = 0 \quad \text{and} \quad \nabla \times \mathbf{B} = k_6 \mathbf{J} + k_7 \frac{\partial \mathbf{E}}{\partial t} \quad (\text{C.3})$$

Not all the constants k_1 to k_7 are independent. Deriving the continuity equation from Maxwell's equations will give the identity

$$k_6 \nabla \cdot \mathbf{J} + k_4 k_7 \frac{\partial \rho}{\partial t} = 0 \quad (\text{C.4})$$

so we have the restriction that $k_6 = k_1 k_4 k_7$. Similarly, by rederiving the wave equations we find that $k_5 k_7 = c^{-2}$.

More confusion arises when we consider macroscopic media. Gauss' law in terms of the polarization field becomes

$$\nabla \cdot (\mathbf{E} + k_4 \mathbf{P}) = k_4 \rho_{\text{free}}. \quad (\text{C.5})$$

which defines the displacement field

$$\mathbf{D} = k_8 (\mathbf{E} + k_4 \mathbf{P}). \quad (\text{C.6})$$

Similarly, with a bound current density $\mathbf{J}_{\text{bound}} = \nabla \times \mathbf{M} + \dot{\mathbf{P}}$ the Maxwell-Ampère law becomes

$$\nabla \times (\mathbf{B} - k_6 \mathbf{M}) = k_6 \mathbf{J}_{\text{free}} + k_7 \frac{\partial}{\partial t} (\mathbf{E} + k_4 \mathbf{P}). \quad (\text{C.7})$$

The magnetic field is then defined as

$$\mathbf{H} = k_9 (\mathbf{B} - k_6 \mathbf{M}). \quad (\text{C.8})$$

The different constants $k_1 \dots k_9$ in both SI and Gaussian units are given by

	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	k_9
SI units	1	1	1	ϵ_0^{-1}	1	μ_0	c^{-2}	ϵ_0	μ_0^{-1}
Gaussian units	1	1	c^{-1}	4π	c^{-1}	$4\pi c^{-1}$	c^{-1}	1	1

As you know, in the SI system the fundamental units are meter (m), kilogram (kg), second (s), and ampère (A). The unit of force is the Newton (N), which is of course kg m/s^2 . Just like the second and the kilogram, the ampère is defined operationally. One ampère is the amount of current needed to create a force of $2 \cdot 10^{-7}$ N per meter length between two infinitely long parallel wires separated by a distance of one meter.

In the Gaussian system, the fundamental units of mass, length, and time are the gram (g), centimeter (cm), and again the second (s). Instead of the ampère, the other fundamental unit is the statcoulomb (statC) for the unit of charge. The statvolt (statV) is the unit of potential, and obeys $1 \text{ statV} = 1 \text{ g cm/statC s}^2$.

	charge	current	E	B	μ_0	ϵ_0
SI	C = As	A	N/C = V/m	T = N/Am	Tm/A	C ² /Nm ²
Gauss	statC	statA = statC/s	statV/cm	G(statV/cm)	–	–

In Gaussian units the permittivity and the permeability are dimensionless.