Lecture notes

PHY3211–Electromagnetic Theory

Hubert de Guise
Department of Physics,
Lakehead University,
Thunder Bay, ON, P7B 5E1

September 27, 2007

Contents

1 Electrostatic Fields

The operative word here is static, meaning time-independent. The title of this chapter could equally be "Time-independent electric fields". The objective of the chapter is to learn about electrostatic fields in vacuum: electric fields in materials is the topic of the next chapter.

The source of an electrostatic field is a static (i.e. non-moving) charge or a static charge distribution. The electrostatic field is a useful (albeit abstract) concept because, given a charge distribution, one can compute various useful quantities knowing the fields.

For instance, the energy stored in a capacitor is computed once the electrostatic field between the plates of the capacitor is known.

The calculation of the electric field produced by many point charges or by a charge distribution depends critically on what is known as the superposition principle.

When charge distributions exhibit a particular symmetry (v.g. they are uniform over a plane, a cylinder or a sphere), it is possible to considerably simplify the calculation of an electric field by using the so–called Gauss's flux theorem. This theorem is closely related to one of Maxwell's equations and it is always true. Unfortunately, it is only useful for the calculation of electric fields in the limited cases where charge distribution are symmetrical.

1.1 Charges

Electricity is associated with electric charges, and it is known that there are two types of electric charges. For convenience, they are called positive and negative charges. They could have been called "male" and "female" charges, but the use of "positive" and "negative" comes with some definite mathematical undertones, and it is the long–standing rule to use the adjective "positive" and "negative".

Any amount of charge on a body is the integer multiple of an elementary charge, which is denoted by e. Under normal circumstances, a body will carry a huge number of elementary charges, so the discrete nature of the net charge is not immediately apparent.

This situation is similar to, say, water. A bucket of water contains, ultimately, a huge integer multiple of the basic "unit" of water: the water molecule. For most practical purposes, we can consider water as a continuous fluid without worrying about the individual water molecules.

To measure the amount of charges, we use the unit of the **Coulomb**, with symbol "C". One Coulomb contains

$$1\mathrm{C} \sim 6 \times 10^{18} e. \tag{1.1}$$

More accurately,

$$e = 1.6019 \times 10^{-19} C. \tag{1.2}$$

The elementary unit of electric charge is an extremely small fraction of a Coulomb, but it is nevertheless discrete. In particular, an electron carries a charge $-e$, while a proton carries a charge of $+e$.

Quite generally, we will say that a body carries a charge of Q C. Most of the time, it is not necessary to know precisely how many multiples of e does Q represent, no more that it is necessary to know precisely how many molecules of water are found in a bucket.

1.2 Force between two isolated charges: Coulomb's law

It is found experimentally that when two sufficiently isolated bodies carrying charges Q_1 and Q_2 respectively are brought in relative proximity, they will either repel or attract one another. If both bodies carry like charges (either positive or negative), they will repel one another. If one body carries a positive charge while the other carries a negative charge, they will attract.

1.2.1 Magnitude of the force

If the first body, carrying charge Q_1, is located at position $\vec{r}_1$ in space, while the second body, carrying charge Q_2, is located at position $\vec{r}_2$ in space, the force of attraction or repulsion is a vector with magnitude

$$F = k\frac{Q_1 Q_2}{|\vec{r}_2 - \vec{r}_1|^2}. \tag{1.3}$$

In this expression, k is a constant which will be specified later, and $|\vec{r}_2 - \vec{r}_1|$ is the distance between the two bodies. Recall that this distance is the length of $|\vec{r}_2 - \vec{r}_1|$, and is calculated from the positive square root of the scalar product:

$$|\vec{r}_2 - \vec{r}_1| = \sqrt{(\vec{r}_2 - \vec{r}_1) \cdot (\vec{r}_2 - \vec{r}_1)}. \tag{1.4}$$

We can make this more explicit. If a Cartesian system of coordinates is used, then the vector $\vec{r}_1$ and $\vec{r}_2$ have, respectively, components

$$\begin{aligned} \vec{r}_1 &= x_1\hat{x} + y_1\hat{y} + z_1\hat{z}, & (1.5)\\ \vec{r}_2 &= x_2\hat{x} + y_2\hat{y} + z_2\hat{z}. & (1.6) \end{aligned}$$

Then,

$$\begin{aligned} \vec{r}_2 - \vec{r}_1 &= (x_2 - x_1)\,\hat{x} + (y_2 - y_1)\,\hat{y} + (z_2 - z_1)\,\hat{z}, & (1.7)\\ (\vec{r}_2 - \vec{r}_1) \cdot (\vec{r}_2 - \vec{r}_1) &= (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2\,, & (1.8)\\ |\vec{r}_2 - \vec{r}_1| &= \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} & (1.9) \end{aligned}$$

Note carefully how Eqn.(1.3) is displayed. On the left, we have the *magnitude* of the force. This is a scalar. On the left, we have a combination of scalars, so that both sides are scalars.

The constant k is usually written in terms of other constants. Numerically, we have

$$k = \frac{1}{4\pi\varepsilon_0}. \tag{1.10}$$

The constant

$$\varepsilon_0 = 8.854187817... \times 10^{-12}\mathrm{Fm}^{-1}, \tag{1.11}$$

where F is the Farad, the unit of capacitance, and m is the unit of length, the meter. ε_0 is called the permittivity of free space.

The combination of units of charge, permittivity of space and distances in Eqn.(1.3) produces units of physical force (i.e. Newton, or N), so that both sides of the equation have identical units.

A convenient numerical shortcut is to use $k \sim 9 \times 10^9$.

1.2.2 Direction of the force

We now need to establish the vectorial nature of Eqn.(1.3), i.e. make the vector signs appear correctly.

The force is found to act along the line joining the two charges, i.e. parallel to the vector $\vec{r}_2 - \vec{r}_1$. Since the magnitude of the force is already established, we thus have

$$\vec{F}_{12} = \pm \hat{r}_{12} \frac{1}{4\pi\varepsilon_0} \frac{Q_1\, Q_2}{|\vec{r}_2 - \vec{r}_1|^2}\,, \tag{1.12}$$

where $\hat{r}_{12}$ is a unit vector in the direction of $\vec{r}_2 - \vec{r}_1$:

$$\hat{r}_{12} = \frac{\vec{r}_2 - \vec{r}_1}{|\vec{r}_2 - \vec{r}_1|}\,. \tag{1.13}$$

To determine if one must keep the "+" or the "-" sign, we look at the special case where the first charge is located at the origin: $\vec{r}_1 = \vec{0}$. Then, $\vec{r}_2 - \vec{r}_1$ points away from the first charge. If we assume Q_1 and Q_2 are positive, for simplicity, and choose the + sign, the force $\vec{F}$ will be pointing away from Q_1, indicating that Q_2 will be repelled by Q_1, which is something experimentally verified. If, on the other hand, Q_1 is positive but Q_2 is negative, their product will be $Q_1 Q_2 = -|Q_1 Q_2|$ so the + sign would produce a force directed towards Q_1, indicating that the second charge would be attracted by Q_1, which is also experimentally correct. Thus, we conclude that we must keep the + sign and write the vectorial form of Coulomb's law as

$$\vec{F}_{12} = \hat{r}_{12} \frac{1}{4\pi\varepsilon_0} \frac{Q_1\, Q_2}{|\vec{r}_2 - \vec{r}_1|^2}\,, \tag{1.14}$$

$$= \frac{Q_1\, Q_2}{4\pi\varepsilon_0} \frac{(\vec{r}_2 - \vec{r}_1)}{|\vec{r}_2 - \vec{r}_1|^3}\,, \tag{1.15}$$

where Eqn.(1.13) has been used. This is Eqn.(4.6b) of Sadiku.

Eqn.(1.15) describes the force on the second charge produces by the first charge. The force on the first charge produced by the second charge is obtained by replacing $\hat{r}_{12}$ by $\hat{r}_{21}$, with the appropriate definition

$$\hat{r}_{21} = \frac{\vec{r}_1 - \vec{r}_2}{|\vec{r}_1 - \vec{r}_2|}\,. \tag{1.16}$$

Since $|\vec{r}_1 - \vec{r}_2| = |\vec{r}_2 - \vec{r}_1|$, we see that

$$\hat{r}_{21} = -\hat{r}_{12}\,, \tag{1.17}$$

so that

$$\vec{F}_{21} = -\vec{F}_{12}\,, \tag{1.18}$$

in accordance with Newton's law of action–reaction.

1.2.3 Comments

It is not really surprising that the direction of the force should be along the line joining the two charges. If you have two isolated charges in space, the only preferred direction is the direction joining them. There is nothing special about any other direction, so it would have been disconcerting to find the force acting at some angle from the direction joining the two charges.

As soon as there is a force on, say, the second charge, it will be set in motion according to Newton's law $\vec{F} = m\vec{a}$. Conversely, the first charge will also start accelerated motion, albeit in the opposite direction. Thus, one never strictly fulfills the condition that the charges are static (i.e. not moving). However, it is found experimentally that Coulomb's law holds in those circumstances where the charges are almost static, i.e. move with "small" velocities. Here, "small" velocities means velocities much smaller than the speed of light $c \sim 3 \times 10^8$m/s. When the velocities are large, Coulomb's law breaks down and one must use Einstein's special theory of relativity to properly compute forces.

1.3 Net force from multiple charges: the superposition principle

We now generalize the problem as follows. Let us place a charge Q at some location $\vec{r}_p$. The subscript p refers to this point. Suppose there are two charges Q_1 and Q_2, located respectively at $\vec{r}_1$ and $\vec{r}_2$. What is the net force $\vec{F}$ on Q? The answer is thankfully very simple: compute the force $\vec{F}_1$ on Q due to Q_1 alone, compute the force $\vec{F}_2$ on Q due to Q_2 alone, and simply add, i.e.

$$\vec{F} = \vec{F}_1 + \vec{F}_2. \tag{1.19}$$

Note the slight difference in notation with Eqn.(1.15) : since the charge Q is not "numbered", we have

$$\begin{aligned}
\vec{F}_1 &= \frac{QQ_1}{4\pi\varepsilon_0}\frac{(\vec{r}_p - \vec{r}_1)}{|\vec{r}_p - \vec{r}_1|^3}, && (1.20)\\
\vec{F}_2 &= \frac{QQ_2}{4\pi\varepsilon_0}\frac{(\vec{r}_p - \vec{r}_2)}{|\vec{r}_p - \vec{r}_2|^3}, && (1.21)\\
\vec{F} &= Q\left[\frac{1}{4\pi\varepsilon_0}\left(Q_1\frac{(\vec{r}_p - \vec{r}_1)}{|\vec{r}_p - \vec{r}_1|^3} + Q_2\frac{(\vec{r}_p - \vec{r}_2)}{|\vec{r}_p - \vec{r}_2|^3}\right)\right]. && (1.22)
\end{aligned}$$

The net force is the simple sum, or simple superposition, of the individual forces dues to each charge Q_1, Q_2. The presence of Q_2 does not affect in any way the force on Q due to Q_1. Of course, the net force on Q depends on both Q_1 and Q_2, but not the individual contributions.

This generalizes to any number of charges: the net force on a charge Q located at $\vec{r}_p$ dues to charges $Q_1, Q_2, ..., Q_N$, located respectively at $\vec{r}_1, \vec{r}_2, ..., \vec{r}_N$, is the simple sum, or simple superpositoin, of the individual forces due to $Q_1, Q_2, ..., Q_N$ separately. Mathematically, we can write the generalization of Eqn. (1.22) to N charges as

$$\vec{F} = Q\left[\frac{1}{4\pi\varepsilon_0}\left(Q_1\frac{(\vec{r}_p - \vec{r}_1)}{|\vec{r}_p - \vec{r}_1|^3} + Q_2\frac{(\vec{r}_p - \vec{r}_2)}{|\vec{r}_p - \vec{r}_2|^3} + ... + Q_N\frac{(\vec{r}_p - \vec{r}_N)}{|\vec{r}_p - \vec{r}_N|^3}\right)\right], \tag{1.23}$$

$$= Q\left[\frac{1}{4\pi\varepsilon_0}\sum_{k=1}^{N} Q_k \frac{(\vec{r}_p - \vec{r}_k)}{|\vec{r}_p - \vec{r}_k|^3}\right]. \tag{1.24}$$

Eqn.(1.24) is **fundamental**. It is known as the superposition principle for forces. It is far from obvious that Nature would work in a way that would allow us to simply "superpose" the results of numerous individual calculations.

1.4 The electrostatic field: Definition

Eqn.(1.24) clearly states that the net force on charge Q is proportional to Q itself. Thus, is is convenient to define the electric field $\vec{E}$ at point $\vec{r}_p$ due to the N charges as

$$\vec{E}(\vec{r}_p) \equiv \lim_{Q\to 0}\frac{\vec{F}}{Q} = \frac{1}{4\pi\varepsilon_0}\sum_{k=1}^{N} Q_k \frac{(\vec{r}_p - \vec{r}_k)}{|\vec{r}_p - \vec{r}_k|^3}. \tag{1.25}$$

Essentially, one imagines that a smaller and smaller charge Q is placed somewhere, that one then calculates the force on this charge, and divides by Q . Since Q is a scalar and $\vec{F}$ a vector, the result, the electric field, is a vector. One could, in theory, repeat the calculation for every single possible $\vec{r}_p$, and thus obtain the electric field $\vec{E}$ at any point $\vec{r}_p$.

If we know the electric field at $\vec{r}_p$, we can calculate immediately what *would be* the force on a charge Q located there: it is simply

$$\vec{F} = Q\vec{E}(\vec{r}_p). \tag{1.26}$$

Physicists consider the electric field to be more fundamental than the force. This is because, as we will see, it is possible to store energy in the electric field, without reference to any force whatsoever. Furthermore, it is often easier to compute fields than forces, especially in these cases (not covered in this class) where the *sources* of the field, i.e. the charges $Q_1, ..., Q_N$ that occur on the right hand side of Eqn.(1.25) move at velocities close to the speed of light.

It is clear from the definition that, since the electric field is defined via the net force on a charge, and since the superposition principle holds for forces, the superposition holds for the calculation of electric fields. In other words, the net field at some point $\vec{r}_p$ is the sum of all the fields calculated individually for all charges acting as the source of field.

1.5 Worked out examples: point charges and dipoles

1.5.1 Point charge at the origin

Let us obtain an expression for the electric field in the simplest possible case: a point charge $+Q$ located at the origin.

Here, we have a single source located at $\vec{r}_1 = \vec{0}$ so, at any point $\vec{r}_p$, application of Eqn.(1.25) gives

$$\vec{E}(\vec{r}_p) = \frac{Q}{4\pi\varepsilon_0}\frac{\vec{r}_p}{|\vec{r}_p|^3} = \frac{Q}{4\pi\varepsilon_0}\frac{\hat{r}_p}{|\vec{r}_p|^2}\,, \tag{1.27}$$

where $\hat{r}_p$ is a unit vector point radially from the origin toward $\vec{r}_p$.

The field is spherically symmetric: if $\vec{r}_{p_1}$ and $\vec{r}_{p_2}$ are two points such that $|r_{p_1}| = |r_{p_2}|$, the electric field at $\vec{r}_1$ has the same magnitude as the electric field at $\vec{r}_2$: the fields only differ in their orientation.

Graphically, we can represent the field on a piece of paper by drawing a representative sets of electric field vectors in the $z = 0$ plane. The result is presented in Fig.2.

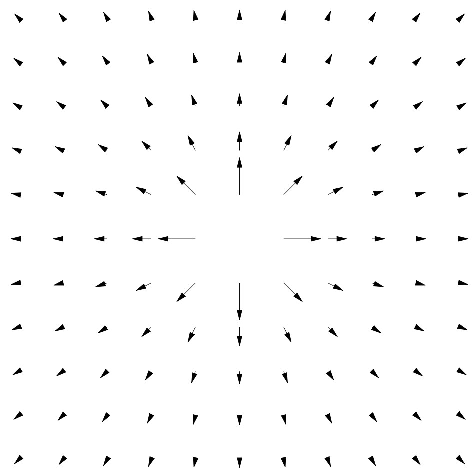

Figure 1: The electric field in the $z = 0$ plane of a positive point charge located at the origin.

Similarly, the electric field of a point charge $-Q$ located at the origin is given by

$$\vec{E}(\vec{r}_p) = -\frac{Q}{4\pi\varepsilon_0}\frac{\vec{r}_p}{|\vec{r}_p|^3}\,, \tag{1.28}$$

$$= -\frac{Q}{4\pi\varepsilon_0}\frac{\hat{r}_p}{|\vec{r}_p|^2}\,, \tag{1.29}$$

Some representative $\vec{E}$ vectors located in the $z = 0$ plane are illustrated in Fig.2. We see that the direction of the electric field is now always towards the origin. Thus, if a positive charge is located anywhere in space, it will feel a force of magnitude $+QE(\vec{r}_p)$ in the direction of $\vec{E}(\vec{r}_p)$, i.e. in the direction of the origin. Thus, this positive charge will be attracted towards the (negative) charge at the origin.

1.5.2 Dipole in the near field

We next consider a more complicated configuration. Imagine we have a positive charge $+Q$ located at $\vec{r}_1 = (d, 0, 0)$, i.e. at some distance d from the origin along the $\hat{x}$ axis. Another charge $-Q$ is located at $\vec{r}_2 = (-d, 0, 0)$, also on the $\hat{x}$ axis. This configuration is known as an electric dipole.

Let us compute the net field at a point $\vec{r}_p = (-2d, 0, 0)$. We have:

$$\vec{E}(-2d, 0, 0) = \frac{Q}{4\pi\varepsilon_0}\frac{\vec{r}_p - \vec{r}_1}{|\vec{r}_p - \vec{r}_1|^3} + \frac{-Q}{4\pi\varepsilon_0}\frac{\vec{r}_p - \vec{r}_2}{|\vec{r}_p - \vec{r}_2|^3}\,. \tag{1.30}$$

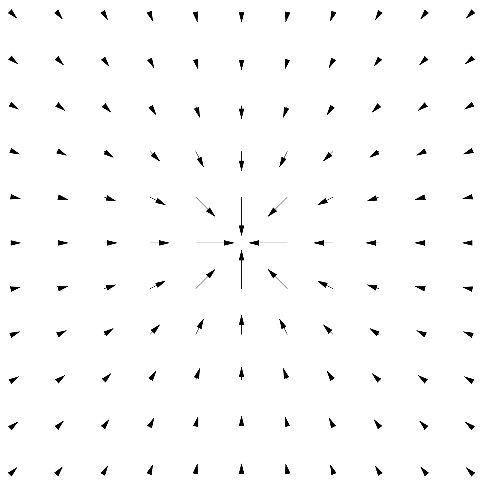

Figure 2: The electric field in the $z = 0$ plane of a negative point charge located at the origin.

Using

$$\vec{r}_p - \vec{r}_1 = -2d\,\hat{x} - (d\hat{x}) = -3d\,\hat{x}\,, \qquad \vec{r}_p - \vec{r}_2 = -2d\hat{x} - (-d\hat{x}) = -d\,\hat{x}\,, \tag{1.31}$$

we have

$$|\vec{r}_p - \vec{r}_1|^3 = 27d^3\,, \qquad |\vec{r}_p - \vec{r}_2|^3 = d^3\,, \tag{1.32}$$

and

$$\vec{E}(-2d,0,0) = \frac{Q}{4\pi\varepsilon_0}\left(\frac{-\hat{x}\,3d}{27d^3}\right) + \frac{-Q}{4\pi\varepsilon_0}\left(\frac{-\hat{x}\,d}{d^3}\right)\,, \tag{1.33}$$

$$= \hat{x}\,\frac{Q}{4\pi\varepsilon_0}\left(-\frac{1}{9d^2} + \frac{1}{d^2}\right) = \hat{x}\,\frac{Q}{4\pi\varepsilon_0}\,\frac{8}{9d^2}\,. \tag{1.34}$$

We can understand qualitatively this results as follows. The point $\vec{r}_p$ is to the left of the negative charge, at a distance d from this large. Thus, the field due to this charge will be in the $+\hat{x}$ direction. The positive charge located at $d\,\hat{x}$ produces at $\vec{r}_p = (-2d, 0, 0)$ a field in the $-\hat{x}$ direction. However, this field is smaller in magnitude than the field of the $-Q$ charge because the latter is closer to $\vec{r}_p$ than the charge $+Q$. Thus, the net field is obtained by superposing a large positive contribution from $-Q$ to a small negative contribution from $+Q$. The result is in the $\hat{x}$ direction.

Quite generally, the field lines in the $z = 0$ plane near the two charges are reproduced in Fig.3

Let us now compute the $\vec{E}$ at $\vec{r}_p = (0, y_p, 0)$, i.e. at an arbitrary point along the $\hat{y}$ axis. We now have:

$$\vec{E}(0, y_p, 0) = \frac{Q}{4\pi\varepsilon_0}\,\frac{\vec{r}_p - \vec{r}_1}{|\vec{r}_p - \vec{r}_1|^3} + \frac{-Q}{4\pi\varepsilon_0}\,\frac{\vec{r}_p - \vec{r}_2}{|\vec{r}_p - \vec{r}_2|^3}\,, \tag{1.35}$$

where

$$\vec{r}_p - \vec{r}_1 = y_0\,\hat{y} - (d\hat{x}) = (-d, y_p, 0), \qquad |\vec{r}_p - \vec{r}_1| = \sqrt{d^2 + y_p^2}\,, \tag{1.36}$$

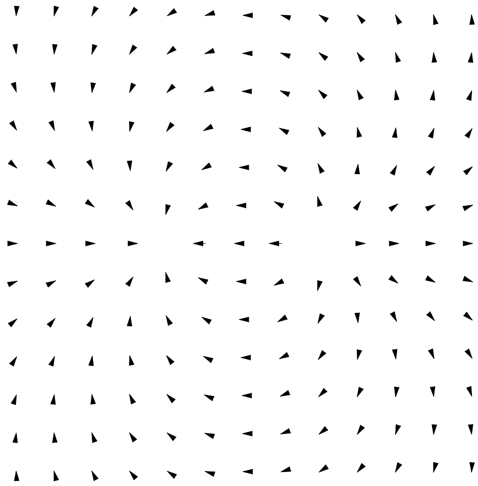

Figure 3: The electric field in the $z = 0$ plane of an electric dipole located at the origin.

and

$$\vec{r}_p - \vec{r}_2 = y_0\,\hat{y} - (-d\hat{x}) = (d, y_p, 0), \qquad |\vec{r}_p - \vec{r}_2| = \sqrt{d^2 + y_p^2}\,. \tag{1.37}$$

Using this in Eqn.(1.25), we find

$$\vec{E}(0, y_p, 0) = \frac{Q}{4\pi\varepsilon_0}\left(\frac{-d\,\hat{x} + y_p\hat{y}}{(d^2 + y_p^2)^{3/2}}\right) + \frac{-Q}{4\pi\varepsilon_0}\left(\frac{\hat{x}\,d + y_p\hat{y}}{(d^2 + y_p^2)^{3/2}}\right), \tag{1.38}$$

$$= \hat{x}\,\frac{Q}{4\pi\varepsilon_0}\frac{-2d}{(d^2 + y_p^2)^{3/2}}\,. \tag{1.39}$$

Note the components along $\hat{y}$ have canceled out. Again, we can understand this qualitatively. Pick a point y_p on the $\hat{y}$ axis, and draw the $\vec{E}$ field at that point due to the positive charge. Since the $\hat{y}$ axis is on the left of the positive charge, the $\vec{E}$ vector will be pointing towards the "left–up" position (assuming you have chosen $y_p > 0$.) The field due to the negative charge points in the "left–down" direction, and has the same magnitude as the field due to the $+Q$ charge since both charges are at the same distance from y_p. The resulting vector is completely towards the left, since the "up" part of the one field is canceled by the "down" part of the other. Again, this is visible on Fig.3. Although there is no vector drawn exactly on the $\hat{y}$ axis, it is clear that the resulting field is in the $-\hat{x}$ direction.

1.5.3 Dipole in the far field

Finally, we consider one last case, that we will rediscover when we deal with antennae. Let us compute the $\vec{E}$ field for any point $\vec{r}_p$ such that $|\vec{r}_p| >> |\vec{r}_1|$, i.e. when we are very far from both sources.

The trick here is to use

$$|\vec{r}_p - \vec{r}_1|^2 = (\vec{r}_p - \vec{r}_1)\cdot(\vec{r}_p - \vec{r}_1)\,, \tag{1.40}$$

$$= \vec{r}_p\cdot\vec{r}_p + \vec{r}_1\cdot\vec{r}_1 - 2\vec{r}_p\cdot\vec{r}_1\,\cos\theta\,, \tag{1.41}$$

where θ is the angle between $\vec{r}_p$ and $\vec{r}_1$. Let us denote by r_p the quantity $|\vec{r}_p|$ and recall that $|\vec{r}_1| = d$. Thus,

$$|\vec{r}_p - \vec{r}_1|^2 = r_p^2 + d^2 - 2r_p\, d\cos\theta\,. \tag{1.42}$$

Likewise, since $\vec{r}_2 = -\vec{r}_1$, we have

$$|\vec{r}_p - \vec{r}_2|^2 = r_p^2 + d^2 + 2r_p\, d\cos\theta\,. \tag{1.43}$$

Now,

$$\frac{1}{|\vec{r}_p - \vec{r}_1|^3} = \frac{1}{(r_p^2 + d^2 - 2r_p\, d\cos\theta)^{3/2}}\,, \tag{1.44}$$

$$= \frac{1}{r_p^3(1 + \frac{d^2}{r_p^2} - 2\frac{d}{r_p}\cos\theta)^{3/2}}\,. \tag{1.45}$$

We now use the fact that $r_p >> d$ and recall the approximation

$$\frac{1}{(1-x)^n} \approx 1 + nx + \frac{1}{2}n(n-1)x^2 + \ldots\,, \tag{1.46}$$

which is valid when $|x| <<< 1$. Here, we can identify

$$x \leftrightarrow 2\frac{d}{r_p}\cos\theta - \frac{d^2}{r_p^2}\,, \qquad n = \frac{3}{2}\,, \tag{1.47}$$

so

$$\frac{1}{|\vec{r}_p - \vec{r}_1|^3} \approx \frac{1}{r_p^3}\left(1 + \frac{3}{2} \times 2\frac{d}{r_p}\cos\theta\right)\,. \tag{1.48}$$

Likewise,

$$\frac{1}{|\vec{r}_p - \vec{r}_2|^3} \approx \frac{1}{r_p^3}\left(1 - \frac{3}{2} \times 2\frac{d}{r_p}\cos\theta\right)\,. \tag{1.49}$$

The terms d^2/r_p^2 have been dropped because, if d/r_p is small, then d^2/r_p^2 is even smaller and can be neglected. This will work so long as $\cos\theta$ is not too small.

Hence, we have, approximately,

$$\vec{E}(\vec{r}_p) \approx \frac{Q}{4\pi\varepsilon_0\, r_p^3}\,(\vec{r}_p - \vec{r}_1)\left(1 + 3\frac{d}{r_p}\cos\theta\right) - \frac{Q}{4\pi\varepsilon_0\, r_p^3}\,(\vec{r}_p + \vec{r}_1)\left(1 - 3\frac{d}{r_p}\cos\theta\right)\,, \tag{1.50}$$

where $\vec{r}_2 = -\vec{r}_1$ has been used. This can be cleaned up somewhat:

$$\vec{E}(\vec{r}_p) = \frac{Q\, \times 2d}{4\pi\varepsilon_0\, r_p^3}\,(3\cos\theta\,\hat{r}_p - \,\hat{r}_1) \tag{1.51}$$

The combination $2dQ$ is called the dipole moment of the configuration. The "far–field" of a dipole is characterized by the $1/r_p^3$ dependence.

2 The electrostatic field of continuous charge distributions

In most practical applications, the actual number of charges in an electrical device is so large that it is not convenient to keep track of them individually.

The situation is similar to, say, a bar of metal. Even though the bar is, at the microscopic level, a collection of individual atoms, such a detailed description is not very convenient and many calculations are simplified by using macroscopic quantities like the mass density.

Thus, we will consider the electric field produced by electric charge distributions. We will encounter three types of distributions: linear, surface or volume. Although the types of distributions are geometrically different, the computational approach to obtaining $\vec{E}$ is always based on the same principle: the principle of superposition.

2.1 Linear charge distributions

2.1.1 General expression

A linear charge distribution usually occurs when only one of the dimension of the problem is relevant. For instance, when dealing with a wire of negligible cross-section, one may often consider the wire as a one–dimensional object carry a linear charge density λ. The linear charge density may or may not depend on the position along the wire.

To compute the electric field of a wire, we divide it in little pieces of length dl_s. (Notice how dl_s, which is a small length, has no direction.) The little piece located at position $\vec{r}_s$ will carry a small amount of charge

$$dq_s = \lambda(r_s) dl_s. \tag{2.1}$$

(Notice how dq_s, which is a small charge, is not a vector, and how the right hand side is a product of two numbers.) We imagine dl_s sufficiently small so that we treat the little piece as a lump of charge dq_s. Thus, this little piece at $\vec{r}_s$ will produce at $\vec{r}_p$ a small electric field $d\vec{E}$, of magnitude and direction given by

$$d\vec{E} = \frac{dq_s}{4\pi\varepsilon_0} \frac{(\vec{r}_p - \vec{r}_s)}{|\vec{r}_p - \vec{r}_s|^3} = \frac{dq_s}{4\pi\varepsilon_0} \frac{(\vec{r}_p - \vec{r}_s)}{|\vec{r}_p - \vec{r}_s|^3}. \tag{2.2}$$

The total electric field is computed by the addition (or superposition) of the small electric field. As the addition of infinitesimal quantities is an integral, we thus obtain

$$\begin{aligned} \vec{E}(\vec{r}_p) &= \int_{\vec{r}_s} d\vec{E} = \int_{r_s} \frac{dq_s}{4\pi\varepsilon_0} \frac{(\vec{r}_p - \vec{r}_s)}{|\vec{r}_p - \vec{r}_s|^3} \\ &= \int_{\vec{r}_s} \frac{\lambda(r_s) dl_s}{4\pi\varepsilon_0} \frac{(\vec{r}_p - \vec{r}_s)}{|\vec{r}_p - \vec{r}_s|^3}. \end{aligned} \tag{2.3}$$

As we are summing over all the sources, the integration region contains all the sources, i.e. it covers all the wire.

2.1.2 Worked out example 1: constant linear charge density

Consider an infinitely long wire, stretched along $\hat{z}$ and going through the origin. The wire carries a constant linear charge density λ. To find the electric field $\vec{E}(\vec{r}_p)$ at any point $\vec{r}_p = (x_p, 0, 0)$ not on the wire, consider first a small piece of the wire of length dl_s located at $(0, 0, z_s)$. By (2.1), it will hold a small amount of charge $\lambda\, dl_s$. Thus, this small pieces produces at $\vec{r}_p$ the small field

$$
\begin{aligned}
d\vec{E} &= \frac{\lambda\, dl_s}{4\pi\varepsilon_0} \frac{(x_p, 0, -z_s)}{\left|x_p^2 + z_s^2\right|^{3/2}}, &(2.4)\\
&= \frac{\lambda\, dl_s}{4\pi\varepsilon_0} \frac{(x_p\hat{x} - z_s\hat{z})}{\left|x_p^2 + z_s^2\right|^{3/2}}. &(2.5)
\end{aligned}
$$

Next, we note that, since the wire is along $\hat{z}$, the length dl_s is really dz_s. The region of integration is from $z_s = -\infty$ to $z_s = \infty$ since the wire extends between those limits along $\hat{z}$. Thus, Eq.(2.3) now reads, for this problem:

$$
\vec{E}(x_p, 0, 0) = \int_{-\infty}^{\infty} \frac{dz_s\lambda}{4\pi\varepsilon_0} \frac{(x_p\hat{x} - z_s\hat{z})}{\left|x_p^2 + z_s^2\right|^{3/2}}. \tag{2.6}
$$

By inspection, there is no $\hat{y}$ component to $\vec{E}(x_p, 0, 0)$: nothing in Eqn.(2.6) is along $\hat{y}$. The $\hat{x}$ component is given by

$$
E_x(x_p, 0, 0) = \int_{-\infty}^{\infty} \frac{dz_s\lambda}{4\pi\varepsilon_0} \frac{x_p}{\left|x_p^2 + z_s^2\right|^{3/2}}. \tag{2.7}
$$

The integration is over z, so we can pull out a bunch of constants:

$$
E_x(x_p, 0, 0) = \frac{\lambda\, x_p}{4\pi\varepsilon_0} \int_{-\infty}^{\infty} \frac{dz_s}{\left|x_p^2 + z_s^2\right|^{3/2}}. \tag{2.8}
$$

A primitive for

$$
\int \frac{dz_s}{\left|x_p^2 + z_s^2\right|^{3/2}} \tag{2.9}
$$

is found by trigonometric substitution. Let $z_s = x_p \tan\theta$, so that

$$
\begin{aligned}
dz_s &= x_p(1 + \tan^2\theta)d\theta = \frac{x_p}{\cos^2\theta} d\theta, &(2.10)\\
x_p^2 + z_s^2 &= x_p^2(1 + \tan^2\theta) = \frac{x_p^2}{\cos^2\theta}, &(2.11)\\
\int \frac{dz_s}{\left|x_p^2 + z_s^2\right|^{3/2}} &= \int \frac{x_p}{\cos^2\theta} d\theta \frac{1}{\left(\frac{x_p^2}{\cos^2\theta}\right)^{3/2}} &(2.12)\\
&= \frac{1}{x_p^2} \int d\theta \cos\theta = \frac{\sin\theta}{x_p^2}. &(2.13)
\end{aligned}
$$

To complete the calculation, we need to return to the original variables by constructing an auxiliary triangle of slope $\tan\theta = z_s/x_p$. We obtain from this triangle the relation

$$\sin\theta = \frac{z_s}{\sqrt{x_p^2 + z_s^2}} \tag{2.14}$$

so that

$$\int_{-\infty}^{\infty} \frac{dz_s}{\left|x_p^2 + z_s^2\right|^{3/2}} = \frac{1}{x_p^2}\left.\frac{z_s}{\sqrt{x_p^2 + z_s^2}}\right|_{-\infty}^{\infty} = \frac{1}{x_p^2} - \left(\frac{-1}{x_p^2}\right) = \frac{2}{x_p^2}. \tag{2.15}$$

Thus, we have, from Eq.(2.8)

$$E_x(x_p, 0, 0) = \frac{\lambda x_p}{4\pi\varepsilon_0}\frac{2}{x_p^2} = \frac{\lambda}{2\pi\varepsilon_0 x_p}. \tag{2.16}$$

The $\hat{z}$ component is given by

$$E_z(x_p, 0, 0) = \int_{-\infty}^{\infty} \frac{dz_s\lambda}{4\pi\varepsilon_0}\frac{z_s}{\left|x_p^2 + z_s^2\right|^{3/2}}. \tag{2.17}$$

$$= \frac{\lambda}{4\pi\varepsilon_0}\int_{-\infty}^{\infty} \frac{z_s dz_s}{\left|x_p^2 + z_s^2\right|^{3/2}}. \tag{2.18}$$

Here, the direct substitution $Z = x_p^2 + z_s^2$ produces

$$dZ = 2z_s dz_s. \tag{2.19}$$

Thus a primitive for

$$\int \frac{z_s dz_s}{\left|x_p^2 + z_s^2\right|^{3/2}} = \frac{1}{2}\int \frac{dZ}{Z^{3/2}} = \frac{-2}{2Z^{1/2}} = -\frac{1}{Z^{1/2}} \tag{2.20}$$

$$= -\frac{1}{\sqrt{x_p^2 + z_s^2}} \tag{2.21}$$

so that

$$\int_{-\infty}^{\infty} \frac{z_s dz_s}{\left|x_p^2 + z_s^2\right|^{3/2}} = -\left.\frac{1}{\sqrt{x_p^2 + z_s^2}}\right|_{-\infty}^{\infty} = -0 - 0 = 0\,. \tag{2.22}$$

Thus,

$$E_z(x_p, 0, 0) = 0\,. \tag{2.23}$$

Hence, the net electric field is

$$\vec{E}(x_p, 0, 0) = \hat{x}E_x = \hat{x}\frac{\lambda}{2\pi\varepsilon_0 x_p}. \tag{2.24}$$

Please make good notes of the following observations.

The result of Eqn.(2.23) is actually *obvious* by the symmetry of the problem. To see this, consider two pieces of the wire, located respectively at $(-z_s, 0, 0)$ and $(z_s, 0, 0)$. Each piece is of the same length dz_s, and thus carry the same charge $dq_s = \lambda\, dz_s$. By construction, both pieces are at the same distance from $\vec{r}_p$; this distance is $\sqrt{z_s^2 + x_p^2}$. Thus, both pieces with produce at $\vec{r}_p$ electric fields of the *same magnitude* but pointing in *different direction.* Let us add them:

$$\frac{\lambda\, dz_s}{4\pi\varepsilon_0} \frac{(x_p, 0, 0) - (0, 0, -z_s)}{(z_s^2 + x_p^2)^{3/2}} + \frac{\lambda\, dz_s}{4\pi\varepsilon_0} \frac{(x_p, 0, 0) - (0, 0, z_s)}{(z_s^2 + x_p^2)^{3/2}} = \frac{\lambda\, dz_s}{4\pi\varepsilon_0} \frac{2x_p\hat{x}}{(z_s^2 + x_p^2)^{3/2}} \,. \tag{2.25}$$

Already, there is no $\hat{z}$ component in this expression.

The *net* electric field $\vec{E}$ is the integral sum of all such pairs, so

$$\vec{E} = \int_0^\infty \frac{\lambda\, dz_s}{4\pi\varepsilon_0} \frac{2x_p\hat{x}}{(z_s^2 + x_p^2)^{3/2}} \,. \tag{2.26}$$

Here, the sum is from 0 to ∞ because we are only summing over pair members on the right of the axis. Since the integrand does not have a $\hat{z}$ component, $\vec{E}$ cannot have a $\hat{z}$ component.

Even when we calculate the $\hat{x}$ component of $\vec{E}$, we need to integrate over dz_s. We are using the principle of superposition, so we need to add the contribution of the source charges. These sources are located along $\hat{z}$. Note that, in fact, all integrals are along dz_s because this is where the sources are located. The location of the charges has nothing to do with the component of $\vec{E}$ we want to evaluate.

2.1.3 Worked out example 2: variable linear charge density

This is the previous example with a twist: instead of having a constant linear charge density on the whole wire, a linear constant charge density λ is placed on the right side of the wire while a linear constant charge density $-\lambda$ is placed on the left side of the wire.

The wire is otherwise the same: it is again infinitely long, stretched along $\hat{z}$ and going through the origin.

To find the electric field $\vec{E}(\vec{r}_p)$ at any point $\vec{r}_p = (x_p, 0, 0)$ not on the wire, we first break the wire into a very large number of small pieces of length dl_s.

Let us pick one piece, located at $(0, 0, z_s)$, and assume $z_s > 0$, i.e. the piece is on the right. By (2.1), it will hold a small amount of charge $\lambda\, dl_s$. Thus, this small pieces produces at $\vec{r}_p$ the small field

$$d\vec{E} = \frac{\lambda\, dl_s}{4\pi\varepsilon_0} \frac{(x_p, 0, -z_s)}{\left|x_p^2 + z_s^2\right|^{3/2}}, \tag{2.27}$$

$$= \frac{\lambda\, dl_s}{4\pi\varepsilon_0} \frac{(x_p\hat{x} - z_s\hat{z})}{\left|x_p^2 + z_s^2\right|^{3/2}}. \tag{2.28}$$

Let us consider now a piece in the $z_s < 0$ region. The charge density is $-\lambda$ there, so this small piece will produce a small field

$$d\vec{E} = \frac{-\lambda\, dl_s}{4\pi\varepsilon_0} \frac{(x_p, 0, -z_s)}{\left|x_p^2 + z_s^2\right|^{3/2}}, \tag{2.29}$$

$$= \frac{-\lambda\, dl_s}{4\pi\varepsilon_0} \frac{(x_p\hat{x} - z_s\hat{z})}{\left|x_p^2 + z_s^2\right|^{3/2}}. \tag{2.30}$$

The only differences between Eqn.(2.28) and Eqn.(2.30) are: the sign of the charge density is different and the expressions are valid for different ranges of values of z_s

The net field is then obtained by superposition, by integrating all the contribution from the left to all the contributions from the right. BEWARE: the result is NOT 0!

Indeed, if you correctly consider the correct range of application of Eqn.(2.28) and Eqn.(2.30) and recall that $dl_s = dz_s$, the net field is

$$\vec{E} = \int_{-\infty}^{0} \frac{-\lambda\, dz_s}{4\pi\varepsilon_0} \frac{(x_p\hat{x} - z_s\hat{z})}{\left|x_p^2 + z_s^2\right|^{3/2}} + \int_{0}^{\infty} \frac{\lambda\, dz_s}{4\pi\varepsilon_0} \frac{(x_p\hat{x} - z_s\hat{z})}{\left|x_p^2 + z_s^2\right|^{3/2}} \tag{2.31}$$

The first integral is the contribution from the left part of the wire while the second is the contribution from the right part of the wire.

We will now manipulate the first integral so that we get some (surprising) simplification. In

$$\int_{-\infty}^{0} \frac{-\lambda\, dz_s}{4\pi\varepsilon_0} \frac{(x_p\hat{x} - z_s\hat{z})}{\left|x_p^2 + z_s^2\right|^{3/2}} \tag{2.32}$$

make the change of variable

$$z_s \to -Z_s\,. \tag{2.33}$$

Then, $dz_s = -dZ_s$ and the integration range for z_s becomes $(\infty, 0)$ for Z_s. Hence,

$$\int_{\infty}^{0} \frac{\lambda\, dZ_s}{4\pi\varepsilon_0} \frac{(x_p\hat{x} + Z_s\hat{z})}{\left|x_p^2 + Z_s^2\right|^{3/2}} \tag{2.34}$$

Turn the limits of integration around:

$$\int_{\infty}^{0} \frac{\lambda\, dZ_s}{4\pi\varepsilon_0} \frac{(x_p\hat{x} + Z_s\hat{z})}{\left|x_p^2 + Z_s^2\right|^{3/2}} = -\int_{0}^{\infty} \frac{\lambda\, dZ_s}{4\pi\varepsilon_0} \frac{(x_p\hat{x} + Z_s\hat{z})}{\left|x_p^2 + Z_s^2\right|^{3/2}} \tag{2.35}$$

and do the change $Z_s \to z_s$ (the integration variable is just a dummy so I can call it what I want):

$$-\int_{0}^{\infty} \frac{\lambda\, dz_s}{4\pi\varepsilon_0} \frac{(x_p\hat{x} + z_s\hat{z})}{\left|x_p^2 + z_s^2\right|^{3/2}} \tag{2.36}$$

and put this back in Eqn.(2.31):

$$\begin{aligned} \vec{E} &= -\int_{0}^{\infty} \frac{\lambda\, dz_s}{4\pi\varepsilon_0} \frac{(x_p\hat{x} + z_s\hat{z})}{\left|x_p^2 + z_s^2\right|^{3/2}} + \int_{0}^{\infty} \frac{\lambda\, dz_s}{4\pi\varepsilon_0} \frac{(x_p\hat{x} - z_s\hat{z})}{\left|x_p^2 + z_s^2\right|^{3/2}}\,, && (2.37)\\ &= -2\int_{0}^{\infty} \frac{\lambda\, dz_s}{4\pi\varepsilon_0} \frac{z_s\hat{z}}{\left|x_p^2 + z_s^2\right|^{3/2}}\,. && (2.38) \end{aligned}$$

This makes it clear that it is the $\hat{x}$ component that has now vanished while the $\hat{z}$ component remains. To complete the calculation and find the field, one must do the integral, but let us see if it is possible to deduce by symmetry that E_x must be 0 for this setup.

As before, pick two small pieces of identical length dz_s, located at $-z_s$ and $+z_s$ respectively so that both pieces are at the same distance from $\vec{r}_p$. The contribution from the piece at $-z_s$ is

$$d\vec{E} = -\frac{\lambda\, dz_s}{4\pi\varepsilon_0} \frac{(-z_s\hat{z} - x_p\hat{x})}{|x_p^2 + z_s^2|^{3/2}} \,. \tag{2.39}$$

(Remember that dz_s, the length of the piece, is positive, despite z_s being negative.) The contribution from the piece on the right is

$$d\vec{E} = \frac{\lambda\, dz_s}{4\pi\varepsilon_0} \frac{z_s\hat{z} - x_p\hat{x}}{|x_p^2 + z_s^2|^{3/2}} \,. \tag{2.40}$$

Add them to get the contribution of the pair:

$$-\frac{\lambda\, dz_s}{4\pi\varepsilon_0} \frac{(-z_s\hat{z} - x_p\hat{x})}{|x_p^2 + z_s^2|^{3/2}} + \frac{\lambda\, dz_s}{4\pi\varepsilon_0} \frac{z_s\hat{z} - x_p\hat{x}}{|x_p^2 + z_s^2|^{3/2}} = 2\frac{\lambda\, dz_s}{4\pi\varepsilon_0} \frac{z_s\hat{z}}{|x_p^2 + z_s^2|^{3/2}}. \tag{2.41}$$

The $\hat{x}$ contribution of every such pair cancels, so the net field, when summed over such pairs, cannot have any $\hat{x}$ component either.

2.2 Surface charge distributions

2.2.1 General expression

A surface charge distribution usually occurs when only the depth of an object is so small as to be ignorable. For instance, when dealing with a small disk of negligible thickness, one may often consider the disk as a two–dimensional object carry a linear charge density σ. The surface of a conducting wire of radius r , or the interface between a conductor and an insulator are other examples of a situation where a surface charge density may occur. The surface charge density may or may not depend on the position on the surface.

To compute the electric field produced by a surface charge density, we divide the surface in little pieces of area dA_s. (Notice how dA_s, which is a small area, has no direction.) The little piece located at position $\vec{r}_s$ will carry a small amount of charge

$$dq_s = \sigma(\vec{r}_s)dS_s. \tag{2.42}$$

We imagine dS_s sufficiently small so that we treat the little area as a lump of charge dq_s. Thus, this little piece at $\vec{r}_s$ will produce at $\vec{r}_p$ a small electric field $d\vec{E}$, of magnitude and direction given by

$$d\vec{E} = \frac{dq_s}{4\pi\varepsilon_0} \frac{(\vec{r}_p - \vec{r}_s)}{|\vec{r}_p - \vec{r}_s|^3} = \frac{\sigma(\vec{r}_s)dS_s}{4\pi\varepsilon_0} \frac{(\vec{r}_p - \vec{r}_s)}{|\vec{r}_p - \vec{r}_s|^3}. \tag{2.43}$$

Note the similarities between Eqs.(2.2) and (2.43). The net field at $\vec{r}_p$ is the sum, or superposition, of all the small fields. Since the sources are now spread over an area, the net field is the double sum (*i.e.* the surface integral) of all the small fields:

$$\begin{aligned} \vec{E}(\vec{r}_p) &= \int_{A_s} d\vec{E} = \int_{r_s} \frac{dq}{4\pi\varepsilon_0} \frac{(\vec{r}_p - \vec{r}_s)}{|\vec{r}_p - \vec{r}_s|^3} \\ &= \int_{A_s} \frac{\sigma(r_s)dA_s}{4\pi\varepsilon_0} \frac{(\vec{r}_p - \vec{r}_s)}{|\vec{r}_p - \vec{r}_s|^3}. \end{aligned} \tag{2.44}$$

The integration is carried over the entire area where there are charges.

2.2.2 Worked out example 1: infinitely long strip

Consider an infinitely long strip of negligeable thickness, stretched along $\hat{y}$ from -∞ to ∞, along $\hat{x}$ between $(-a, 0, 0)$ and $(a, 0, 0)$ and going through the origin. The strip carries a surface charge density σ that varies with the position on the strip as

$$\sigma(x_s, y_s, 0) = \sigma_0 x_s. \tag{2.45}$$

To find the electric field $\vec{E}(\vec{r}_p)$ at any point $\vec{r}_p = (0, 0, z_p)$ above the middle of the strip, consider first a small piece of the strip, having area $dA_s = dx_s dy_s$, and located at $(x_s, y_s, 0)$. By (2.42), the small area will hold a small amount of charge $\sigma(x_s, y_s, 0)dA_s = \sigma_0 x_s dx_s dy_s$. Thus, this small pieces produces at $\vec{r}_p$ the small field

$$d\vec{E} = \frac{\sigma_0 x_s dx_s dy_s}{4\pi\varepsilon_0} \frac{(-x_s, -y_s, z_p)}{\left|x_s^2 + y_s^2 + z_p^2\right|^{3/2}}, \tag{2.46}$$

The region of integration is from $x_s = -a$ to $x_s = a$ and from $y_s = -\infty$ to $y_s = \infty$ since the strips extends between those limits along $\hat{x}$ and $\hat{y}$ Thus, we have, for this problem:

$$\vec{E}(0, 0, z_p) = \int_{-a}^{a} dx_s \int_{-\infty}^{\infty} dy_s \frac{\sigma_0 x_s}{4\pi\varepsilon_0} \frac{(-x_s\hat{x} - y_s\hat{y} + z_p\hat{z})}{\left|x_s^2 + y_s^2 + z_p^2\right|^{3/2}}. \tag{2.47}$$

The $\hat{x}$ component is given by

$$E_x(0, 0, z_p) = -\frac{\sigma_0}{4\pi\varepsilon_0} \int_{-a}^{a} dx_s x_s^2 \int_{-\infty}^{\infty} \frac{dy_s}{\left|x_s^2 + y_s^2 + z_p^2\right|^{3/2}}. \tag{2.48}$$

Experience shows that it is simplest to start with the integration is over y. This is a trigonometric substitution using

$$y_s = \sqrt{x_s^2 + z_p^2}\tan\theta, \tag{2.49}$$

$$dy_s = \frac{\sqrt{x_s^2 + z_p^2}}{\cos^2\theta} d\theta, \tag{2.50}$$

$$x_s^2 + y_s^2 + z_p^2 = \left(x_s^2 + z_p^2\right)\left(1 + \tan^2\theta\right), \tag{2.51}$$

$$\int \frac{dy_s}{\left|x_s^2 + y_s^2 + z_p^2\right|^{3/2}} = \int \frac{\sqrt{x_s^2 + z_p^2}}{\cos^2\theta} d\theta \frac{1}{\left(\frac{x_s^2 + z_p^2}{\cos^2\theta}\right)^{3/2}}, \tag{2.52}$$

$$= \frac{1}{x_s^2 + z_p^2} \int \cos\theta d\theta = \frac{\sin\theta}{x_s^2 + z_p^2}. \tag{2.53}$$

To convert back to the original variables, we need an auxiliary triangle of slope $\tan\theta = y_s/\sqrt{x_s^2 + z_p^2}$, so that

$$\sin\theta = \frac{y_s}{\sqrt{x_s^2 + y_s^2 + z_p^2}} \tag{2.54}$$

and thus

$$\int_{-\infty}^{\infty} \frac{dy_s}{\left|x_s^2+y_s^2+z_p^2\right|^{3/2}} = \frac{1}{x_s^2+z_p^2} \left.\frac{y_s}{\sqrt{x_s^2+y_s^2+z_p^2}}\right|_{-\infty}^{\infty} \tag{2.55}$$

$$= \frac{2}{x_s^2+z_p^2}. \tag{2.56}$$

Next, we need to integrate over dx_s :

$$E_x(0,0,z_p) = -\frac{\sigma_0}{4\pi\varepsilon_0}\int_{-a}^{a} dx_s x_s^2 \int_{-\infty}^{\infty} \frac{dy_s}{\left|x_s^2+y_s^2+z_p^2\right|^{3/2}} \tag{2.57}$$

$$= -\frac{\sigma_0}{4\pi\varepsilon_0}\int_{-a}^{a} dx_s x_s^2 \frac{2}{x_s^2+z_p^2} \tag{2.58}$$

$$= -\frac{\sigma_0}{2\pi\varepsilon_0}\int_{-a}^{a} dx_s \frac{x_s^2}{x_s^2+z_p^2} \tag{2.59}$$

First, we manipulate the integral to

$$E_x(0,0,z_p) = -\frac{\sigma_0}{2\pi\varepsilon_0}\int_{-a}^{a} dx_s \frac{x_s^2}{x_s^2+z_p^2} \tag{2.60}$$

$$= -\frac{\sigma_0}{2\pi\varepsilon_0}\int_{-a}^{a} dx_s \frac{x_s^2+z_p^2-z_p^2}{x_s^2+z_p^2} \tag{2.61}$$

$$= -\frac{\sigma_0}{2\pi\varepsilon_0}\int_{-a}^{a} dx_s \left(1-\frac{z_p^2}{x_s^2+z_p^2}\right). \tag{2.62}$$

A simple trigonometric substitution $x_s = z_p \tan\theta$ will simplify the second term in the integral to

$$\int dx_s \frac{1}{x_s^2+z_p^2} = z_p \arctan\left(\frac{x_s}{z_p}\right), \tag{2.63}$$

and the net result is

$$E_x(0,0,z_p) = -\frac{\sigma_0}{2\pi\varepsilon_0}\left(x_s - z_p \arctan\left(\frac{x_s}{z_p}\right)\right)_{-a}^{a} \tag{2.64}$$

$$= \frac{\sigma_0}{\pi\varepsilon_0}\left(z_p \arctan\left(\frac{a}{z_p}\right) - a\right). \tag{2.65}$$

The $y-$component is zero by symmetry, but we can verify this the hard way by integrating

$$E_y(0,0,z_p) = -\frac{\sigma_0}{4\pi\varepsilon_0}\int_{-a}^{a} dx_s x_s \int_{-\infty}^{\infty} \frac{y_s dy_s}{\left|x_s^2+y_s^2+z_p^2\right|^{3/2}}. \tag{2.66}$$

The integral of dy_s is done by the simple substitution $\xi = x_s^2+y_s^2+z_p^2$, which yields

$$\int \frac{y_s dy_s}{\left|x_s^2+y_s^2+z_p^2\right|^{3/2}} = \frac{1}{2}\int \frac{d\xi}{\xi^{3/2}} = -\frac{1}{\sqrt{\xi}} = -\frac{1}{\sqrt{x_s^2+y_s^2+z_p^2}}, \tag{2.67}$$

so that

$$\int_{-\infty}^{\infty} \frac{y_s dy_s}{\left|x_s^2 + y_s^2 + z_p^2\right|^{3/2}} = -\left.\frac{1}{\sqrt{x_s^2 + y_s^2 + z_p^2}}\right|_{-\infty}^{\infty} = 0, \tag{2.68}$$

as anticipated.

The $z-$component is also relatively easy:

$$\begin{aligned}
E_z(0,0,z_p) &= \frac{\sigma_0}{4\pi\varepsilon_0} \int_{-a}^{a} dx_s x_s \int_{-\infty}^{\infty} \frac{z_p\, dy_s}{\left|x_s^2 + y_s^2 + z_p^2\right|^{3/2}}, &(2.69)\\
&= \frac{\sigma_0 z_p}{4\pi\varepsilon_0} \int_{-a}^{a} dx_s x_s \int_{-\infty}^{\infty} \frac{dy_s}{\left|x_s^2 + y_s^2 + z_p^2\right|^{3/2}}, &(2.70)\\
&= \frac{\sigma_0 z_p}{4\pi\varepsilon_0} \int_{-a}^{a} dx_s x_s \frac{2}{x_s^2 + z_p^2}. &(2.71)
\end{aligned}$$

To proceed, we substitute $\eta = x_s^2 + z_p^2$ to get

$$\begin{aligned}
\int dx_s \frac{x_s}{x_s^2 + z_p^2} &= \frac{1}{2}\int \frac{d\eta}{\eta} = \frac{1}{2}\ln\eta &(2.72)\\
&= \frac{1}{2}\ln\left(x_s^2 + z_p^2\right), &(2.73)\\
E_z(0,0,z_p) &= \frac{\sigma_0 z_p}{2\pi\varepsilon_0} \int_{-a}^{a} dx_s \frac{x_s}{x_s^2 + z_p^2} &(2.74)\\
&= \frac{\sigma_0 z_p}{4\pi\varepsilon_0}\left[\ln\left(a^2 + z_p^2\right) - \ln\left(a^2 + z_p^2\right)\right] &(2.75)\\
&= 0. &(2.76)
\end{aligned}$$

Thus, the net field only has an $\hat{x}$ component, and we have

$$\vec{E}(0,0,z_p) = \hat{x}\frac{\sigma_0}{\pi\varepsilon_0}\left(z_p \arctan\left(\frac{a}{z_p}\right) - a\right). \tag{2.77}$$

2.2.3 Important Worked out example 2: punctured disk and parallel plates

Consider a punctured disk of inner radius a and outer radius b. The center of the disk is at the origin and the disk lies in the $z = 0$ plane. The disk carries a uniform surface charge density $\sigma(\vec{r}_s) = \sigma_0$. What is the electric field $\vec{E}$ at $\vec{r}_p = (0,0,z_p)$?

We proceed, as always, by superposition. Because of the circular shape of the disk, we will use a mixture of cartesian coordinates (to easily add vectors) and cylindrical coordinates (to easily locate charges). Thus,

$$x_s = r_s \cos\theta_s\,, \qquad y_x = r_s \sin\theta_s\,. \tag{2.78}$$

We divide the disk into small pieces. The piece located at $\vec{r}_s = (x_s, y_s, 0)$ has area $dx_s\, dy_s$. In cylindrical, this is just $dA_s = r_s\, dr_s\, d\theta_s$.

This small area will contain a small amount of charges

$$dq_s = \sigma(\vec{r}_s)\, dA_s = \sigma_0\, r_s\, dr_s\, d\theta_s\,. \tag{2.79}$$

Thus, the small field $d\vec{E}$ produced at $\vec{r}_p$ by this small amount of charge is

$$d\vec{E}(\vec{r}_p) = \frac{dq_s}{4\pi\varepsilon_0} \frac{(\vec{r}_p - \vec{r}_s)}{|\vec{r}_p - \vec{r}_s|^3}\,, \tag{2.80}$$

$$= \frac{\sigma_0\, r_s\, dr_s\, d\theta_s}{4\pi\varepsilon_0} \frac{(z_p\hat{z} - r_s\cos\theta_s\hat{x} - r_s\sin\theta_s\hat{y})}{(r_s^2 + z_p^2)^{3/2}} \tag{2.81}$$

and the net field is thus given by the double integral

$$\vec{E}(0,0,z_p) = \frac{\sigma_0}{4\pi\varepsilon} \int_a^b dr_s\, r_s \int_0^{2\pi} d\theta_s\, \frac{(z_p\hat{z} - r_s\cos\theta_s\hat{x} - r_s\sin\theta_s\hat{y})}{(r_s^2 + z_p^2)^{3/2}} \tag{2.82}$$

Because

$$\int_0^{2\pi} \cos\theta_s\, d\theta_s = \int_0^{2\pi} \sin\theta_s\, d\theta_s = 0\,, \tag{2.83}$$

it is easy to integrate the $\hat{x}$ and $\hat{y}$ of $\vec{E}$ and find that the net field in these directions will be 0. Thus, only E_z remains to be evaluated.

We have

$$\vec{E}(0,0,z_p) = \hat{z}\,\frac{\sigma_0\, z_p}{4\pi\varepsilon_0} \int_a^b dr_s\, r_s \int_0^{2\pi} d\theta_s\, \frac{1}{(r_s^2 + z_p^2)^{3/2}}\,, \tag{2.84}$$

$$= \hat{z}\,\frac{\sigma_0\, z_p}{4\pi\varepsilon_0}\, 2\pi \int_a^b dr_s\, r_s \frac{1}{(r_s^2 + z_p^2)^{3/2}}\,. \tag{2.85}$$

In this last integral, we could proceed by trigonometric substitution, but experience shows it is quicker to set

$$\xi = r_s^2 + z_p^2\,, \qquad \tfrac{1}{2}d\xi = r_s\, dr_s\,. \tag{2.86}$$

With this,

$$\int dr_s \frac{r_s}{(r_s^2 + z_p^2)^{3/2}} = \tfrac{1}{2} \int \frac{d\xi}{\xi^{3/2}} = -\frac{1}{\sqrt{\xi}} = -\frac{1}{\sqrt{r_s^2 + z_p^2}}\,. \tag{2.87}$$

After putting in the limits, we thus have

$$\vec{E}(0,0,z_p) = \hat{z}\,\frac{\sigma_0\, z_p}{2\varepsilon_0} \left[\frac{1}{\sqrt{a^2 + z_p^2}} - \frac{1}{\sqrt{b^2 + z_p^2}} \right]\,. \tag{2.88}$$

Note that the final expression depends only on σ_0 and $\vec{r}_p$, as it should.

Let us make $a = 0$: the disk is no longer punctured at its center. This yields

$$\vec{E}(0,0,z_p) = z\,\frac{\sigma_0}{2\varepsilon_0} \left[\frac{z_p}{\sqrt{z_p^2}} - \frac{z_p}{\sqrt{b^2 + z_p^2}} \right]\,. \tag{2.89}$$

How do we handle $z_p/\sqrt{z_p^2}$? $\sqrt{z_p}^2$ is a positive number, so clearly, if $z_p > 0$, then the ratio is $+1$. However (and be mindful of this), if $z_p < 0$ then the ratio is -1.

We now consider a **very important** special case. Take $b \to \infty$ and $a = 0$. The disk is now an infinite plane. Eq.(2.88) is now

$$\vec{E}(0,0,z_p) = \hat{z}\,\frac{\sigma_0}{2\varepsilon_0}\frac{z_p}{\sqrt{z_p^2}}. \tag{2.90}$$

Hence, we find that, if $z_p > 0$, the field is along $\hat{z}$ and constant in magnitude everywhere, while if $z_p < 0$, the field is along $-\hat{z}$ and constant in magnitude everywhere. In other words, the field of an infinite plane carrying a uniform surface charge density σ_0 is

$$\vec{E}(\vec{r}_p) = \frac{\sigma_0}{2\varepsilon_0}\begin{cases} +\hat{z} & \text{if } z_p \text{ is above the plane,} \\ -\hat{z} & \text{if } z_p \text{ is below the plane.} \end{cases} \tag{2.91}$$

2.2.4 Worked out example 3: punctured disk with variable surface charge density

Consider a disk of inner radius r_1 and outer radius r_2, centered at the origin in the $z = 0$ plane, and carrying a surface charge density $\sigma(\theta) = \sigma_0\,\cos\theta_s$, where the angle θ_s is defined via

$$x_s = r_s\cos\theta_s\,, \qquad y_s = r_s\sin\theta_s\,. \tag{2.92}$$

Let us find $\vec{E}$ at $\vec{r}_p = (0,0,z_p)$.

First, we divide the disk into a large number of small areas. The small patch located at $\vec{r}_s$ has surface $dA_s = r_s\,dr_s\,d\theta_s$ in plane polar coordinates. (The use of plane polar coordinates will simplify some integrals later on.) Thus, this small patch carries a small amount of charge

$$\begin{aligned} dq_s &= \sigma(\theta_s)dA_s\,, & (2.93)\\ &= \sigma_0\,\cos\theta_s\,r_s\,dr_s\,d\theta_s\,, & (2.94) \end{aligned}$$

The distance from this small patch located at $\vec{r}_s$ to the point of interest $\vec{r}_p$ is

$$|\vec{r}_p - \vec{r}_s| = \sqrt{z_p^2 + r_s^2}\,. \tag{2.95}$$

Thus, the small amount of charge will generate a small field

$$\begin{aligned} d\vec{E}(0,0,z_p) &= \frac{dq_s}{4\pi\varepsilon_0}\,\frac{(z_p\hat{z} - r_s\cos\theta_s\hat{x} - r_s\sin\theta_s\hat{y})}{(z_p^2+r_s^2)^{3/2}}\,, & (2.96)\\ &= \frac{\sigma_0\,\cos\theta_s\,r_s\,dr_s\,d\theta_s}{4\pi\varepsilon_0}\,\frac{(z_p\hat{z} - r_s\cos\theta_s\hat{x} - r_s\sin\theta_s\hat{y})}{(z_p^2+r_s^2)^{3/2}}\,, & (2.97) \end{aligned}$$

From this, we rapidly conclude that

$$\begin{aligned} E_z(0,0,z_p) &= \frac{z_p\sigma_0}{4\pi\varepsilon_0}\int_{r_1}^{r_2} dr_s\,\frac{r_s}{(z_p^2+r_s^2)^{3/2}}\int_0^{2\pi} d\theta_s\,\cos\theta_s = 0\,, & (2.98)\\ E_y(0,0,z_p) &= -\frac{\sigma_0}{4\pi\varepsilon_0}\int_{r_1}^{r_2} dr_s\,\frac{r_s^2}{(z_p^2+r_s^2)^{3/2}}\int_0^{2\pi} d\theta_s\,\cos\theta_s\,\sin\theta_s = 0\,. & (2.99) \end{aligned}$$

There remains the $\hat{x}$ component, which is not 0. It is given by

$$E_x = -\frac{\sigma_0}{4\pi\varepsilon_0}\int_{r_1}^{r_2} dr_s\,\frac{r_s^2}{(z_p^2+r_s^2)^{3/2}}\int_0^{2\pi} d\theta_s\,\cos^2\theta_s\,, \tag{2.100}$$

$$= -\frac{\sigma_0}{4\pi\varepsilon_0}\,\frac{1}{2}\int_{r_1}^{r_2} dr_s\,\frac{r_s^2}{(z_p^2+r_s^2)^{3/2}}\,. \tag{2.101}$$

It is possible to integrate by first making a trigonometric substitution of the form $r_s = z_p \tan\xi$. However, this does not appreciably simplify the situation so we will proceed by a different route.

First, we integrate by parts. Recall that

$$\int u\,dv = uv - \int v\,du\,. \tag{2.102}$$

Set

$$u = r_s\,, \qquad du = dr_s\,, \qquad dv = \frac{r_s\,dr_s}{(z_p^2+r_s^2)^{3/2}} \Rightarrow v = -\frac{1}{\sqrt{z_p^2+r_s^2}}\,. \tag{2.103}$$

Hence,

$$\int dr_s\,\frac{r_s^2}{(z_p^2+r_s^2)^{3/2}} = -\frac{r_s}{\sqrt{r_s^2+z_p^2}} + \int \frac{dr_s}{\sqrt{z_p^2+r_s^2}}\,. \tag{2.104}$$

A trigonometric substitution on the rightmost integral does not simplify much, so we use a trick known as Euler's first substitution. Introduce the new variable t via the substitution

$$\sqrt{z_p^2+r_s^2} + r_s = t\,. \tag{2.105}$$

We can turn this around to find

$$r_s = \frac{t^2-z_p^2}{2t} = \tfrac{1}{2}t - \tfrac{z_p^2}{2t}\,, \tag{2.106}$$

which in turn implies

$$dr_s = \left(\tfrac{1}{2} + \tfrac{z_p^2}{2t^2}\right)\,dt = \frac{(t^2+z_p^2)}{2t^2}\,dt\,. \tag{2.107}$$

Furthermore,

$$\sqrt{z_p^2+r_s^2} = t - r_s = t - \tfrac{1}{2}t + \tfrac{z_p^2}{2t} = \tfrac{t^2+z_p^2}{2t} \tag{2.108}$$

so our integral becomes

$$\int \frac{dr_s}{\sqrt{z_p^2+r_s^2}} = \frac{t^2+z_p^2}{2t^2}\,dt\,\frac{2t}{t^2+z_p^2} = \frac{dt}{t} = \ln(t) = \ln\left(\sqrt{z_p^2+r_s^2}+r_s\right)\,. \tag{2.109}$$

Thus,

$$E_x = \left(-\frac{\sigma_0}{8\pi\varepsilon_0}\right)\left(-\frac{r_s}{\sqrt{z_p^2+r_s^2}} + \ln\left(\sqrt{z_p^2+r_s^2}+r_s\right)\right)\Bigg|_{r_1}^{r_2}\,, \tag{2.110}$$

$$= \left(-\frac{\sigma_0}{8\pi\varepsilon_0}\right)\left(\frac{r_1}{\sqrt{z_p^2+r_1^2}} - \frac{r_2}{\sqrt{z_p^2+r_2^2}} + \ln\left(\frac{\sqrt{z_p^2+r_2^2}+r_2}{\sqrt{z_p^2+r_1^2}+r_1}\right)\right)\,. \tag{2.111}$$

3 Gauss' law

We have seen how symmetry considerations can help in determining when some components of $\vec{E}$ are 0. If a component is non–zero, we must apply the superposition principle to complete the calculation.

Are there situation so symmetric so that we can evaluate every component of $\vec{E}$ without recourse to the superposition principle? The answer is yes. This is what we will be looking at.

3.1 Preamble: two spherically symmetric charge distributions

Before we state and study Gauss' law *per se*, we will calculate using the superposition principle the electric field of two very special charge distributions.

3.1.1 A spherical thin shell of radius r.

Consider a *spherical* shell of negligible thickness, centered on the origin, having radius r_s and carrying a *constant* surface charge density σ_0. We will calculate the electric field $\vec{E}$ at a point $\vec{r}_p = (0, 0, z_p)$.

Note that, by symmetry, if we have the magnitude of $\vec{E}$ at $(0, 0, z_p)$, we also have the magnitude of $\vec{E}$ at any point on a sphere or radius $r_p = z_p$. If we pick another point $\vec{r}_{p'}(x_p, 0, 0)$ with $|x_p| = |z_p|$, *i.e.* if the point $\vec{r}_{p'}(x_p, 0, 0)$ is at the same distance from the origin as the point $\vec{r}_p = (0, 0, z_p)$, then $|\vec{E}(x_p, 0, 0)| = |\vec{E}(0, 0, z_p)|$. The *direction* of $\vec{E}$ will be different at the two points, but the *magnitude* of $\vec{E}$ will be the same.

In other words, because the charge distribution is spherically symmetric and constant, there is nothing special about z_p: it might as well be any point on the sphere containing z_p. The choice of z_p is convenient for calculations, nothing more.

Imagine dividing the shell small areas. Because the charge distribution is spherically symmetric, it is convenient to use spherical coordinates where

$$z_s = r_s \cos\theta_s\,, \qquad x_s = r_s \sin\theta_s \cos\phi_s\,, \qquad y_s = r_s \sin\theta_s \sin\phi_s\,. \tag{3.1}$$

Note that r_s is constant: it is the radius of the shell. The area of a small piece of the shell is

$$dA_s = r_s^2 \sin\theta_s\, d\theta_s\, d\phi_s\,. \tag{3.2}$$

For later convenience, we note that the total area of the sphere is

$$A = r_s^2 \int_\pi^0 d\theta_s\, \sin\theta_s \int_0^{2\pi} d\phi = 4\pi\, r_s^2\,, \tag{3.3}$$

where the limits of integration have been adjusted to that the area is positive.

Our small patch of shell contains a small amount of charge

$$dq_s = \sigma_0\, dA_s = \sigma_0 r_s^2 \sin\theta_s\, d\theta_s\, d\phi_s\,, \tag{3.4}$$

and will produce at $\vec{r}_p$ a small field

$$\vec{E} = \frac{r_s^2\sigma_0}{4\pi\varepsilon_0}\int_\pi^0 d\theta_s \sin\theta_s \int_0^{2\pi} d\phi_s \, \frac{(z_p - r_s\cos\theta_s)\hat{z} - r_s\sin\theta_s\cos\phi_s\hat{x} - r_s\sin\theta_s\sin\phi_s\hat{y}}{(r_s^2\sin^2\theta_s + (z_p - r_s\cos\theta_s)^2)^{3/2}} \,. \quad (3.5)$$

Simple integration over ϕ_s shows that $E_x = E_y = 0$. Thus, the field is in the direction of $\vec{r}_p$. The remaining component is obtained from

$$E_z(0,0,z_p) = \frac{r_s^2\sigma_0}{4\pi\varepsilon_0}\, 2\pi \int_0^\pi d\theta_s \sin\theta_s \frac{(z_p - r_s\cos\theta_s)}{(r_s^2\sin^2\theta_s + (z_p - r_s\cos\theta_s)^2)^{3/2}} \,. \quad (3.6)$$

This last integral is moderately difficult. To unravel it, we first set

$$\xi = -r_s\cos\theta_s \;\Rightarrow\; d\xi = r_s\sin\theta_s\, d\theta_s \,, \qquad r_s^2\sin^2\theta_s = r_s^2 - \xi^2 \,. \quad (3.7)$$

Inserting this in Eq.(3.6), we have

$$E_z(0,0,z_p) = \frac{r_s\,\sigma_0}{2\varepsilon_0}\int_{-r_s}^{r_s} d\xi\, \frac{(z_p + \xi)}{(r_s^2 + 2\xi\, z_p + z_p^2)^{3/2}} \,, \quad (3.8)$$

Note that I have (exceptionally) found it easier to directly replace the θ_s limits of integration by the ξ limits.

A little cleaning up of the denominator yields

$$\begin{aligned} E_z(0,0,z_p) &= \frac{r_s\,\sigma_0}{2\varepsilon_0}\int_{-r_s}^{r_s} d\xi\, \frac{z_p}{(r_s^2 + 2\xi\, z_p + z_p^2)^{3/2}} + \frac{r_s\,\sigma_0}{2\varepsilon_0}\int_{-r_s}^{r_s} d\xi\, \frac{\xi}{(r_s^2 + 2\xi\, z_p + z_p^2)^{3/2}} \,, & (3.9)\\ &= \frac{r_s\,\sigma_0}{2\varepsilon_0}\frac{1}{\sqrt{r_s^2 + z_p^2 + 2z_p\, r_s}}\Bigg|_{-r_s}^{r_s} + \frac{r_s\,\sigma_0}{2\varepsilon_0}\int_{-r_s}^{r_s} d\xi\, \frac{\xi}{(r_s^2 + 2\xi\, z_p + z_p^2)^{3/2}} \,, & (3.10)\\ & & (3.11) \end{aligned}$$

The rightmost integral is easily simplified using integration by parts:

$$u = \xi \,, \qquad du = d\xi \,, \qquad dv = \frac{d\xi}{(r_s^2 + 2\xi\, z_p + z_p^2)^{3/2}} \,, \qquad v = -\frac{1}{z_p}\frac{1}{\sqrt{r_s^2 + z_p^2 + 2z_p\xi}} \,, \quad (3.12)$$

so that

$$\begin{aligned} \int d\xi \frac{\xi}{(r_s^2 + 2\xi\, z_p + z_p^2)^{3/2}} &= -\frac{\xi}{z_p}\frac{1}{\sqrt{r_s^2 + z_p^2 + 2z_p\xi}} + \frac{1}{z_p}\int d\xi\, \frac{1}{\sqrt{r_s^2 + z_p^2 + 2z_p\xi}} \,, & (3.13)\\ &= -\frac{\xi}{z_p}\frac{1}{\sqrt{r_s^2 + z_p^2 + 2z_p\xi}} + \frac{1}{z_p^2}\sqrt{r_s^2 + z_p^2 + 2z_p\, r_s} \,, & (3.14)\\ &= \frac{r_s^2 + z_p^2 + z_p\, r_s}{z_p^2\sqrt{r_s^2 + z_p^2 + 2z_p\, r_s}} \,, & (3.15) \end{aligned}$$

Inserting the limits and straightforward manipulations eventually yield

$$\begin{aligned} \int_{-r_s}^{r_s} d\xi \frac{(z_p + \xi)}{(r_s^2 + 2\xi\, z_p + z_p^2)^{3/2}} &= \left[-\frac{1}{\sqrt{r_s^2 + z_p^2 + 2z_p\,\xi}} + \frac{(r_s^2 + z_p^2 + z_p\xi)}{z_p^2\sqrt{r_s^2 + 2\xi\, z_p + z_p^2}} \right]_{-r_s}^{r_s} , & (3.16)\\ &= \frac{r_s}{z_p^2}\left[1 + \frac{(z_p - r_s)}{\sqrt{(r_s - z_p)^2}} \right] . & (3.17) \end{aligned}$$

This last result must be simplified very carefully because we are taking the square root of a square, which should always turn out to be positive. In other words,

$$\frac{(z_p - r_s)}{\sqrt{(r_s - z_p)}} = \begin{cases} -1 & \text{if } z_p < r_s\,, \\ +1 & \text{if } z_p > r_s\,, \end{cases} \tag{3.18}$$

so that

$$\int_{-r_s}^{r_s} d\xi \frac{(z_p + \xi)}{(r_s^2 + 2\xi\, z_p + z_p^2)^{3/2}} = \begin{cases} 0 & \text{if } z_p < r_s\,, \\ \frac{2\,r_s}{z_p^2} & \text{if } z_p > r_s\,. \end{cases} \tag{3.19}$$

Eq.(3.8) then becomes

$$E_z(0,0,z_p) = \begin{cases} 0 & \text{if } z_p < r_s\,, \\ \frac{4\pi\, r_s^2\,\sigma_0}{4\pi\varepsilon_0\, z_p^2} & \text{if } z_p > r_s\,. \end{cases} \tag{3.20}$$

Note that $4\pi\, r_s^2\,\sigma_0$ is the total area of the sphere multiplied by the surface charge density. Thus, we can write $q_0 = 4\pi\, r_s^2\,\sigma_0$ as the net charge on the sphere and simplify our final answer to

$$E_z(0,0,z_p) = \begin{cases} 0 & \text{if } z_p < r_s\,, \\ \frac{q_0}{4\pi\varepsilon_0\, z_p^2} & \text{if } z_p > r_s\,. \end{cases} \tag{3.21}$$

In summary, the field is 0 if $\vec{r}_p$ is inside the shell. If $\vec{r}_p$ is outside the shell, then the field at $\vec{r}_p$ is just the field of a charge q_0 located at the origin. By symmetry, this conclusion holds for any point. Thus, we can write

$$\vec{E}(\vec{r}_p) = \begin{cases} 0 & \text{if } r_p < r_s\,, \\ \frac{q_0}{4\pi\varepsilon_0\, r_p^2}\,\hat{r}_p & \text{if } r_p > r_s\,, \end{cases} \tag{3.22}$$

where $\hat{r}_p$ is a unit vector along the line connecting the origin to $\vec{r}_p$.

3.1.2 A spherical shell of thickness dr and radius r

Let us now consider a slightly more complicated case. Suppose we are given a solid sphere of radius r, with a spherically symmetric volume charge density $\rho(r_s)$ which depends only on the radial distance to the origin. Let us compute the field at $\vec{r}_p = \hat{z}\, z_p$.

We begin by dividing our sphere in shells of thickness dr_s. The volume element for such a shell is

$$dV_s = r_s^2\, \sin\theta_s\, dr_s\, d\theta_s\, d\phi_s\,. \tag{3.23}$$

You can verify for yourself that, upon integration, this produces the volume of a sphere:

$$V = \int_0^r \int_\pi^0 \int_0^{2\pi} r_s^2\, \sin\theta_s\, dr_s\, d\theta_s\, d\phi_s = \frac{4}{3}\pi\, r^3. \tag{3.24}$$

The small piece of volume located (in spherical) at (r_s, θ_s, ϕ_s) contains a small amount of charge

$$dq_s = \rho(r_s) r_s^2\, \sin\theta_s\, dr_s\, d\theta_s\, d\phi_s \tag{3.25}$$

and thus creates a small field

$$d\vec{E}(0,0,z_p) = \frac{\sigma_0}{4\pi\varepsilon_0}\int_0^r dr_s\, r_s^2 \int_\pi^0 d\theta_s \sin\theta_s \int_0^{2\pi} d\phi_s \,\frac{(z_p - r_s\cos\theta_s)\hat{z} - r_s\sin\theta_s\cos\phi_s\hat{x} - r_s\sin\theta_s\sin\phi_s\hat{y}}{(r_s^2\sin^2\theta_s + (z_p - r_s\cos\theta_s)^2)^{3/2}}\,. \tag{3.26}$$

Compare this last expression with Eq.(3.5). In this problem we have an extra integration over dr_s but the integration over the angles are otherwise identical. Thus, we don't have to redo the angular integration: we can simply pinch the result from Eq.(3.20) to obtain

$$E_z(0,0,z_p) = \frac{4\pi}{4\pi\varepsilon_0\, z_p^2}\int_0^{r_0} dr_s\, \rho(r_s)\, r_s^2\,. \tag{3.27}$$

Here, r_0 is defined as follows. We know from Eq.(3.22) that, if the radius r_s of the our shell of thickness dr_s is greater than z_p, then this shell will contribute nothing to the electric field. If, on the other hand, the radius r_s is smaller than z_p, then the total charge contained on the our shell will contribute to the field. Thus, if $z_p < r$, the radius of our sphere, only those thin spherical shells with radius $r_s < z_p$ will contribute. If $z_p > r$, all the thin spherical shells will have a non–zero contribution. Thus, we define

$$r_0 = \begin{cases} z_p & \text{if } z_p < r\,, \\ r & \text{if } z_p \geq r\,. \end{cases} \tag{3.28}$$

Having clarified this, we see that

$$4\pi\int_0^{r_0} dr_s\, \rho(r_s)\, r_s^2 \tag{3.29}$$

is simply the net amount of charge enclosed in the sphere of radius r_0. Thus, we can write

$$E_z(0,0,z_p) = \frac{1}{4\pi\varepsilon_0\, z_p^2}\, q_{encl.} \tag{3.30}$$

where $q_{encl.}$ is the amount of charge enclosed in the sphere of radius r_0. More generally, we write

$$\vec{E}(\vec{r}_p) = \frac{1}{4\pi\varepsilon_0\, r_p^2}\, q_{encl.}\,\hat{r}_p\,. \tag{3.31}$$

3.2 Gauss' law

4 Dielectrics

It will be sufficient for our purposes to (somewhat naively) think of a dielectric as not a conductor. Whereas a conductor is characterized, at the atomic level, by essentially free electrons that can move large distances about the physical bulk of the conductor, dielectrics (or insulators) have much more tightly bound electrons which must remain in the immediate vicinity of the parent atom or molecule.

One should properly distinguish between two types of dielectric: polar and non-polar.

4.1 Definition the electric dipole.

A dipole is an electrical system made from one positive and one negative charge, separated by a distance $2d$. Between the charges, the electric field E always points away from the positive towards the negative charge.

Imagine we have a positive charge $+Q$ located at $\vec{r}_1 = (d, 0, 0)$, *i.e.* at a distance d from the origin along the $\hat{x}$ axis. Another charge $-Q$ is located at $\vec{r}_2 = (-d, 0, 0)$, also on the $\hat{x}$ axis. Let us compute the net field at a point $\vec{r}_p = (x_p, 0, 0)$ to the right of the positive charge, *i.e.* supposing $x_p > d$ for convenience. The system is illustrated in Fig.4.

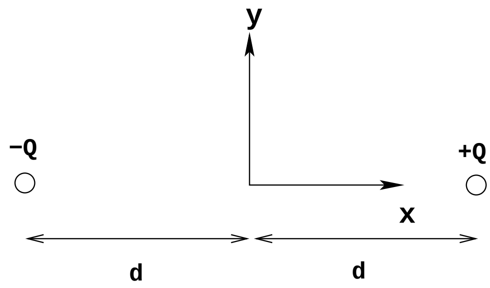

Figure 4: The charge configuration of an electric dipole oriented along $+\hat{x}$.

We have:

$$\vec{E}(x_p, 0, 0) = \frac{Q}{4\pi\varepsilon_0} \frac{\vec{r}_p - \vec{r}_1}{|\vec{r}_p - \vec{r}_1|^3} + \frac{-Q}{4\pi\varepsilon_0} \frac{\vec{r}_p - \vec{r}_2}{|\vec{r}_p - \vec{r}_2|^3} . \tag{4.1}$$

Using

$$\vec{r}_p - \vec{r}_1 = x_p \, \hat{x} - (d\hat{x}) = (x_p - d) \, \hat{x} \, , \qquad \vec{r}_p - \vec{r}_2 = x_p \hat{x} - (-d\hat{x}) = (x_p + d) \, \hat{x} \, , \tag{4.2}$$

we have (remembering that $x_p - d > 0$)

$$|\vec{r}_p - \vec{r}_1|^3 = (x_p - d)^3 \, , \qquad |\vec{r}_p - \vec{r}_2|^3 = (d + x_p)^3 \, , \tag{4.3}$$

and

$$\begin{aligned} \vec{E}(x_p, 0, 0) &= \frac{Q}{4\pi\varepsilon_0} \left(\frac{\hat{x}}{(x_p - d)^2} \right) + \frac{-Q}{4\pi\varepsilon_0} \left(\frac{\hat{x}}{(x_p + d)^2} \right) , & (4.4) \\ &= \hat{x} \, \frac{Q}{4\pi\varepsilon_0} \left(\frac{1}{(x_p - d)^2} - \frac{1}{(x_p + d)^2} \right) & (4.5) \\ &= \hat{x} \, \frac{Q}{4\pi\varepsilon_0} \, \frac{2x_p \times 2d}{(d^2 - x_p^2)^2} . & (4.6) \end{aligned}$$

Quite generally, the field lines in the $z = 0$ plane near the two charges are reproduced in Fig.3.

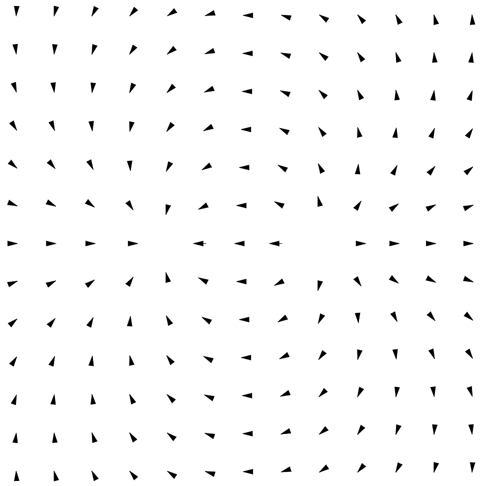

Figure 5: The electric field in the $z = 0$ plane of an electric dipole located at the origin.

To continue, we observe that, with x_p to the right of the positive charge, the electric field of the dipole is directed along $+\hat{x}$. Thus, we *define* the dipole moment $\vec{p}$ of the arrangement as the product

$$\vec{p} \equiv Q \times (\vec{r}_1 - \vec{r}_2) = Q \times (2d)\hat{x} \,. \tag{4.7}$$

Note that the vector $\vec{r}_1 - \vec{r}_2$ point from the negative to the positive charge, so the dipole vector points from the negative charge to the positive charge. In particular, if the field point $x_p\,\hat{x}$ is not in between the charges, the $\vec{E}$-field is parallel to $\vec{p}$.

It is important to note, for later discussion, that if the field point $x_p\,\hat{x}$ is between the charges, the electric field is antiparallel to $\vec{p}$.

Finally, if we assume that x_p is much larger than d, we have

$$(d^2 - x_p^2)^2 \approx x_p^4 \,, \tag{4.8}$$

so that

$$\begin{aligned} \vec{E} &\approx \hat{x}\,\frac{p}{4\pi\varepsilon_0}\frac{2x_p}{x_p^4}\,, & (4.9)\\ &= \frac{1}{4\pi\varepsilon_0 r_p^3} 2\,(\vec{p}\cdot\hat{r}_p)\,\hat{r}_p\,, & (4.10) \end{aligned}$$

where $\vec{r}_p = x_p\hat{x}$ and $\hat{r}_p = \hat{x}_p$ have been used.

Although this has been derived for a specific $\vec{r}_p$, the results remains true for any $\vec{r}_p$ provided that the point $\vec{r}_p$ is not on the axis perpendicular to the line joining the charges.

4.2 Response of a permanent dipole to an external field

Suppose a molecule of some compound is such that the electronic charges are not evenly distributed around the molecular nuclei. This can be the case when the atoms of the molecules have different physical size, as for instance in water. This uneven distribution will result in a local excess of negative charge somewhere around the molecule, to be compensated by a deficit of negative charge (*i.e.* a surplus of positive charge) somewhere else. This type of situation can be modeled, in a first approximation, as a permanent dipole. The magnitude of the dipole moment $\vec{p}$ of water, for instance, is $p = 6.2 \times 10^{-30}$mC.

Suppose now the dipole is plunged in a constant external $\vec{E}$ field of magnitude E_0 (supplied, for instance, by two parallel plates), in such a way that $\vec{p}$ is not initially parallel to $\vec{E}$.

The net force on the dipole is the sum of the electric forces on the positive and negative charges, respectively. Thus,

$$\vec{F} = +Q\vec{E} + (-Q)\vec{E} = 0\,, \tag{4.11}$$

i.e. there is no net force on the dipole and thus no displacement of its center of mass.

There is, however, a net torque about the midpoint of the dipole. This torque is

$$\vec{\tau} = \vec{r}_1 \times \vec{F}_+ + \vec{r}_2 \times \vec{F}_-\,, \tag{4.12}$$

where $\vec{F}_\pm$ is the force on the positive or negative charge, respectively. We can simplify Eq.(4.12) if we note that, by the geometry of the system,

$$\vec{F}_- = -\vec{F}_+\,, \qquad \vec{r}_2 = -\vec{r}_1\,. \tag{4.13}$$

Hence,

$$\vec{\tau} = \vec{r}_1 \times \vec{F}_+ + (-\vec{r}_1) \times (-\vec{F}_+) = 2\vec{r}_1 \times \vec{F}_+\,. \tag{4.14}$$

The magnitude of this torque is

$$\tau = r_1\, Q\, E_0 \sin\theta\,, \tag{4.15}$$

where θ is the angle between the dipole vector $\vec{p}$ and the electric field $\vec{E}$. The torque vanishes when the angle is 0 or π. However, when the angle is π, the dipole is antialigned with the $\vec{E}$ field, like a child on a swing in the upright position. The smallest little perturbation will send it swinging towards the $\theta = 0$ position.

Hence, we see that the effect of the external electric field is to orient the permanent dipole so that it becomes aligned with the external field $\vec{E}$, with $\vec{p}$ parallel to $\vec{E}$.

Obviously, if there is no external field, there will be no alignment. Thus, when one considers a macroscopic material under normal conditions, it is found that the permanent molecular dipoles have random alignments, so that the bulk remains largely non-polar. In other words, although each molecule of water has a permanent dipole, two neighboring dipoles will typically be aligned in different directions (because of thermal energy, collisions, and other such factors). Thus, a bucket of water does not have a measurable dipole moment unless it is plunges into an external field.

4.3 Response of a non-polar molecule to an external field

Suppose a molecule does not have a permanent dipole moment. If we plunge it in a constant electric field, the $\vec{E}$ field will *distort* the electronic distribution around the molecule and will *induce* a dipole moment.

To gain some insight into the deformed charge distribution, imagine once again a uniform external field produced by two parallel plates placed vertically. On the left plate, we have (let us say) a positive surface charge distribution whereas a negative charge distribution is to be found on the right plate. Thus, the external field between the plates is perpendicular to the plates. At every point, it has some given constant magnitude and is pointing towards the negatively charged plate.

In response to this, the electronic cloud around the atoms or molecules of a non-polar dielectric substance will be slightly displaced towards the bottom of the left (towards the positive charges), leaving a small deficit of negative charges (or a small surplus of positive charges) towards the right of the molecule.

If we recall that $\vec{p}$ goes from the positive to the negative charges in a dipole, we can, in a first approximation, model this situation by a small induced dipole parallel to the external $\vec{E}$ field.

4.4 Response of a dielectric to an external field

Whether the molecules of a dielectric have a permanent dipole moment or not, the previous analysis shows that, when plunged in an external $\vec{E}$ field, the dielectric will develop a net dipole moment parallel to the $\vec{E}$ field. The effect on each molecule, *i.e* the resulting final dipole moment $\vec{p}$ of each molecule, is proportional to the net field existing inside the dielectric. A very naive but useful picture is obtained given in Fig.6(a). This figure shows a small volume of a dielectric with dipoles aligned with an external field. The net dipole moment is the sum of the individual dipole, and the dipole moment density $\vec{P}$ is just the net dipole moment divided by the small volume under consideration.

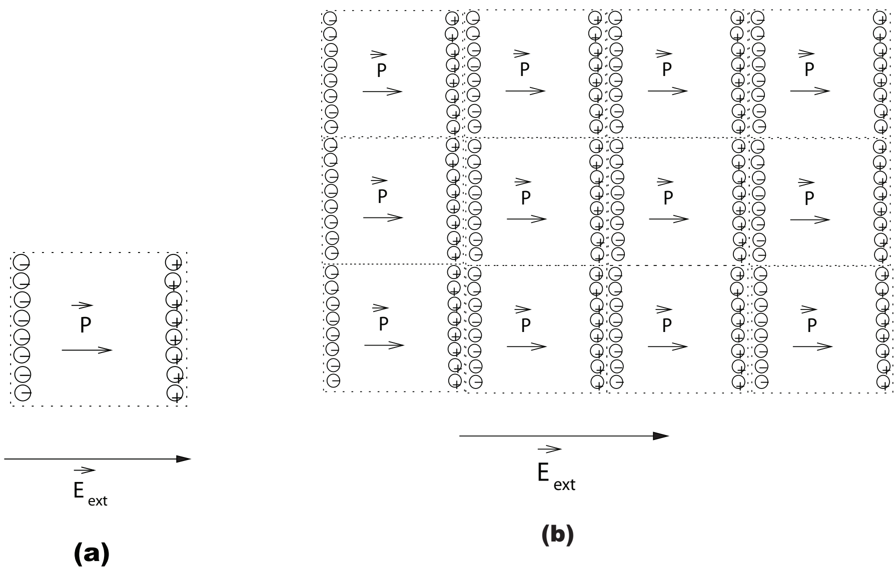

Figure 6: (a) One small volume element containing aligned dipoles. (b) A “macroscopic” piece of dielectric, containing a large number of polarized volume elements.

The bulk dielectric is made by putting together all the small volumes. A cartoon of this is given in Fig.6(b). It has been assumed that the polarization $\vec{P}$ is everywhere constant, *i.e.* all the small volumes have identical $\vec{P}$. Inside the dielectric, one can see that there is net charge, since the layers of surplus negative charges are "cancelled" by the layers of negative charge deficits. However, there is a net effect at the surface of the dielectric.

This net effect is the resulting macroscopic dipole of the sample. It is clear that, although $\vec{P}$ is aligned with the external field, the net electric field inside the sample is *smaller* than the external field because, between the positive and negative charges of a dipole, the $\vec{E}$ field of the dipole is antiparallel to the external field.

To calculate $\vec{E}_{in}$, the electric field inside the sample, we start with

$$\vec{E}_{in} = \vec{E} + \frac{\vec{P}}{\varepsilon_0}\,, \tag{4.16}$$

where the extra ε_0 needs to be included because of the way $\vec{P}$ is defined.

Normally, $\vec{P}$ not only increases with $\vec{E}$ (this is intuitively obvious), but also increases with $\vec{E}_{in}$. In fact, it makes perfect sense to think of the polarization $\vec{P}$ at a point inside the dielectric to be linked to the net electric field inside the material. We will limit ourselves to material and regimes where this increase is proportional to $\vec{E}_{in}$, so

$$\vec{P} = \chi_e \varepsilon_0 \vec{E}_{in}\,, \tag{4.17}$$

where χ_e is the proportionality constant, known as the electric susceptibility of the material. For all known dielectrics, the value of χ_e is greater than 0. Thus, we can rearrange Eq.(4.16) into

$$\vec{E} = \vec{E}_{in} - \frac{\vec{P}}{\varepsilon_0}\,. \tag{4.18}$$

Now, the $\vec{P}$ is antiparallel to $\vec{E}_{in}$ between the polarized charges so we have, as far as magnitude go:

$$|\vec{E}_{in}|\,(1+\chi_e) = |\vec{E}|\,, \tag{4.19}$$

showing the field inside the sample $\vec{E}_{in}$ to be smaller in magnitude than the external field $\vec{E}$, as commanded by our analysis.

For historical reasons, one consider not $\vec{E}$ but rather the material $\vec{D}$ field inside the dielectric. It is given by

$$\vec{D}_{in} = \varepsilon \vec{E}_{in} \tag{4.20}$$

where the electric permittivity ε has been introduced:

$$\varepsilon \equiv (1+\chi_e)\varepsilon_0\,. \tag{4.21}$$

In practice, one usually quotes the *relative permittivity* ε_r, defined by

$$\varepsilon_r \equiv \frac{\varepsilon}{\varepsilon_0} = 1 + \chi_e \tag{4.22}$$

so that $\varepsilon = \varepsilon_r\,\varepsilon_0$. Since values of ε_r are always greater than 1, we see that the electric flux density $\vec{D}$ inside a dielectric is always smaller than the electric flux density in vacuum.

4.5 Bound charges

Before we consider further the response of a dielectric plunged in an external field, we will introduce an analogy.

Consider the population of a high-rise building. Here, the gravitational field of the Earth plays the role of the $\vec{E}$ field. The role of negative charges is played by human heads, while positive charges correspond to pairs human feet. Thus, the high-rise is "neutral", in the sense that the number of heads is equal to the number of pairs of feet.

The gravitational field is oriented downwards, so we expect to see an "alignment" of bodies in such a way that most feet will be pointing down while most heads will be pointing up.

This analogy is useful because it is clear that, in outer space (where there is no gravitational field), there is no preferred orientation to each body: astronauts need not sleep horizontally or walk standing up vertically.

The constant gravitational merely reorients bodies in such a way that, on a given layer, all feet are at the bottom and all heads are near the ceiling. There is an apparent "surplus" of heads at the top of each layers, and an apparent "surplus" of feet (or "deficit" of heads), on the floor of every layer. The constant gravitational field does not add any heads or feet to the astronauts.

In the electric case, the external field reorients the dipoles to create an apparent "surplus" of negative charge near the positively–charged plate and a "surplus" of positive charge near the negatively–charged plate. These "apparent" charges are called bound charges.

If we go back to Fig.6(b), the bound charges are those that are on the physical surface of the dielectric.

If $\hat{n}$ is a unit vector perpendicular to the surface of the dielectric at some point $\vec{r}$, then the surface bound charge density ρ_{ps} and the volume bound charge density ρ_{pv} work out to

$$\rho_{bs} = \vec{P} \cdot \hat{n}\,, \qquad \rho_{bv} = -\vec{\nabla} \cdot \vec{P}\,. \tag{4.23}$$

*(**Notational Quirck**: Sadiku uses the subscript p rather than my b. Furthermore, it is usual to denote surface charge densities by σ rather than ρ. Thus, I would have written σ_b and ρ_b for the bound surface charge density and bound volume charge density, respectively.)* These definitions are best understood in integral form.

The bound charge Q_b on a surface is given by the flux of $\vec{P}$ through that surface:

$$Q_b = \oint_S \vec{P} \cdot d\vec{S} = \int \rho_{bs}\, dS\,. \tag{4.24}$$

As the dielectric is globally electrically neutral, there is an apparent charge that locally remains "not on the surface", *i.e.* inside the dielectric. It is defined to that the net charge of the dielectric is 0. Hence,

$$-Q_b = \int_V \rho_{bv} dV\,. \tag{4.25}$$

Using the divergence theorem to transform

$$\oint \vec{P} \cdot d\vec{S} = \int_V \vec{\nabla} \cdot P\, dv \tag{4.26}$$

yields

$$-Q_b = \int_V \rho_{bv} dV = \int_V \vec{\nabla} \cdot P \, dv \,, \tag{4.27}$$

from which the second of Eq.(4.23) follows.

4.6 Gauss' law for dielectric: free charges

Gauss' law in vacuum reads

$$\oint_S \vec{D} \cdot d\vec{S} = q_{encl.} \,. \tag{4.28}$$

We must now be careful to distinguish between various types of charges in the problem.

Coming back to our working idea of plunging a dielectric in a constant external $\vec{E}$ field, we see that, to produce this external field, we need some charges that are distinct in nature from the bound or "apparent" charges that result from the polarization of the material.

The charges that create the external field are called *free charges*. The free charges are the charges we control. Unlike the bound charges, we can add or take away free charges at will. In Eq.(4.28), q_{encl} refers to *free* charges.

In the presence of a dielectric, what shall become of the RHS of Eq.(4.28)? The answer is: **nothing**. Basically, the polarization effects of the induced dipoles is already included in the definition of $\vec{D}$: in a polarized dielectric, the $\vec{D}$ field is already smaller than it would be if there were not polarization. Including the polarization again would be double-counting the effect of the bound charges.

Thus, we can simply rewrite

$$\oint_S \vec{D} \cdot d\vec{S} = q_{free} \,. \tag{4.29}$$

We can recover the $\vec{E}$ field from $\vec{D}$ using the relation $\vec{D} = \epsilon\vec{E}$ of Eq.(4.20), being careful to use the correct electrical permittivity for our problem.

5 Capacitor problems and Laplace's equation.

In this chapter, we will study capacitors and dielectrics from two opposite premices: in one case, we will assume some charge distribution is specified; in the other, the potential will be specified in some region of space.

We assume we are given two conductors with opposite net charges. Recall that the charges on a conductor are necessarily on the physical surface of the conducting body. When the conductors are brought close together, they will form a capacitors. The capacitance of the device is *defined* to be the ratio

$$C = \frac{Q}{V}, \tag{5.1}$$

where Q is the charge on one of the conductors and V is the potential difference between the conductors. Note that, since the capacitance is a positive number, we don't have to worry about the signs of the charge or the potential.

5.1 Laplace's and Poisson's equation

Thus far, we have proceeded using some charge distribution as initial data. From a distribution, we determine the resulting electric field (usually by integration) and from the field we obtain the potential difference. In other words, from a charge distribution we eventually infer the potential difference between two points. Once Q and V are known, the capacitance can be determined.

Although using charge distribution as a starting point is very natural, one must admit (alas!) that, in applications, one is often presented with a case where some device is operated with some potential difference between two points in a circuit. Thus, it is often more practical to start with V and eventually infer Q.

Whereas the sequence $Q \to \vec{E} \to V$ requires two integrations, the sequence $V \to \vec{E} \to Q$ requires two differentiation: the equation connecting Q and V is a second order differential equation for V. It is known as the Poisson or the Laplace equation, depending on the presence or absence of charge in the neighborhood where we are computing the potential.

To make this explicit, recall that

$$\vec{E} = -\vec{\nabla} V \tag{5.2}$$

Gauss' flux theorem turned around using the divergence theorem yields

$$\oint \vec{D} \cdot d\vec{S} = \int \left(\vec{\nabla} \cdot \vec{D} \right) dV = \int \rho_V dV \Rightarrow \vec{\nabla} \cdot \vec{D} = \rho_V . \tag{5.3}$$

Since, quite generally,

$$\vec{D} = \varepsilon \vec{E}, \tag{5.4}$$

we find

$$\vec{\nabla} \cdot \vec{E} = -\vec{\nabla} \cdot \vec{\nabla} V = \frac{\rho_V}{\varepsilon}. \tag{5.5}$$

This is usually written

$$\nabla^2 V = -\frac{\rho_V}{\varepsilon}, \tag{5.6}$$

where ∇^2 is the Laplacian operator. Eq.(5.6) is called the Poisson equation. It relates the volume charge density ρ_V at some point to the Laplacian of the potential evaluated at that point. When $\rho_V = 0$, we have the Laplace equation.

In the three familiar coordinate system, the Laplacian is

$$\begin{aligned}\nabla^2 V &= \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2}, && (5.7)\\ &= \frac{1}{\rho}\frac{\partial}{\partial \rho}\left(\rho \frac{\partial V}{\partial \rho}\right) + \frac{1}{\rho^2}\frac{\partial^2 V}{\partial \phi^2} + \frac{\partial^2 V}{\partial z^2}, && (5.8)\\ &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial V}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial V}{\partial \theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 V}{\partial \phi^2} && (5.9)\end{aligned}$$

showing the "double–derivative" nature of the resulting equation.

Poisson's equation is an example of a *partial* differential equation, and there are known techniques to solve those. We will only be concerned with cases where the symmetry of the potential is such that V depends on only *one* variable, so that all but one of the derivatives will vanish. This will, in effect, turn Poisson's equation into an *ordinary* second order differential equation.

When the right hand side of Eq.(5.6) is 0, a solution to a second order differential equation is completely determined once two integration constants are specified so that the solution is compatible with the initial conditions of the problem. Once these two integration constants are found, the theory of ordinary differential equation tells us that we are, basically, done. This is the essence of the *uniqueness theorem*: if you have a solution that satisfies the correct boundary solution, then you have *the* solution. If the right hand side of Eq.(5.6) is not 0, then one requires a special solution.

In this course, we will deal exclusively with zero right hand side:

$$\nabla^2 V = 0. \tag{5.10}$$

The full Poisson equation is useful, for instance, in calculations involving semi–conductor pn junctions. In these cases, there is a complicated non–zero charge density in the problem.

5.2 The parallel plate capacitor

Consider two thick perfectly conducting plates, infinite in extent, placed parallel one to the other. For definiteness, assume they are placed in the xy plane, with the top of the lower plate at $z = 0$ and the bottom of the upper plate at $z = h$. Let us assume the space between them is filled with some generic dielectric with dielectric constant $\varepsilon = \varepsilon_r \varepsilon_0$. The dielectric is electrically neutral.

5.2.1 Starting with a charge distribution.

Suppose some constant surface charge density σ_s is placed on the bottom plate. By drawing a Gaussian cube with its bottom inside the thick conducting lower plate and top inside the thick upper plate, one easily deduces from the condition of zero–static field inside a perfect conductor, that a surface charge density $-\sigma_s$ must be present at bottom of the upper plate.

To calculate the field $\vec{E}$ between the plates, we use Gauss's flux theorem again. By symmetry, it must be that the field is entirely in the $\hat{z}$ direction. The presence of charges on the bottom of the upper plate is due to the charges on the lower plate, so we cannot use the charges on the upper plate to calculate the field. So, we choose as our Gaussian surface a pillbox with square cross–section ΔS, with four sides perpendicular to $\hat{z}$, having its lower end inside the bottom plate and having height so that the point $\vec{r}_p$ inside the dielectric lies on its top face. There is no flux on the sides of the box, because the field is always along $\hat{z}$ whereas $d\vec{S}$ is always perpendicular to $\hat{z}$ on the sides of the box. There is no flux from the bottom of the box, because the bottom is in a perfect conductor, in which there is no $\vec{E}$. Thus, the only contribution to the flux is from the top face of the box.

By symmetry, the field is constant in magnitude on the top face, so

$$\oint_S \vec{D}\cdot d\vec{S} = \int_{\text{top}} D(z_p)dS = D(z_p)\Delta S = \sigma_s \Delta S, \tag{5.11}$$

from which we conclude

$$\vec{E} = \frac{\sigma_s}{\varepsilon}\hat{z}. \tag{5.12}$$

You can check this against the result of Sec.4.6C, being careful to take note that, in 4.6C, the infinite sheet has negligible thickness. In 4.6C, it is impossible to have one end of the Gaussian surface "inside" the sheet, as we've done here. This accounts for the factor $\frac{1}{2}$ present in Sec.4.6C but absent from Eq.(5.12).

Knowing the field, we can find the potential difference between the plates. Since the distance between the plates is h, we find, by simple integration, the potential difference ΔV to be

$$\Delta V = V(h) - V(0) = -\int_h^0 \vec{E}\cdot d\vec{l} = \int_0^h \frac{\sigma_s}{\varepsilon}dz = -\frac{\sigma_s}{\varepsilon}h. \tag{5.13}$$

[**Note:** For convenience, we will use $V(0)$ (on the bottom plate) to be the reference and set $V(0) = 0$. Thus, the negative sign in Eq.(5.13) indicates that the charge density on the upper plate is negative, in accordance with what we found before.] If S is the surface area of the plates, the capacitance is therefore given by

$$C = \frac{Q}{\Delta V} = \frac{\rho_s S}{\frac{\sigma_s}{\varepsilon}h} = \frac{\varepsilon S}{h} = \varepsilon_r \frac{\varepsilon_0 S}{h}, \tag{5.14}$$

Note how the charge and potential difference have disappeared from the result. In Eq.(5.14), all negative signs have been eliminated to make sure that the capacitance comes out as a positive number.

This calculation is *not* self-consistent. We have assumed the plates are infinite in extent in order to use Gauss' flux theorem to calculate the field, so S should properly be infinite. In practical situations, of course, the plates are finite in area. However, if the distance h between the plate is small compared to the area S of the plates, and we calculate the field for a point $\vec{r}_p$ located not too close to the edges of the plates, then the plates are effectively infinite and our calculation holds. The assumptions used here to obtain Eq.(5.14) amount to neglecting the fringing field outside the parallel plates.

Eq.(5.14) captures the essential properties of a parallel plate capacitor. If the surface area S of the plates is increased, the capacitance is increased. If the distance between the plates is decreased, the capacitance is increased. Finally, if we use a "good" dielectric, *i.e.* one with a "high" value of ε_r, we increase the capacitance.

5.2.2 Starting from the potential.

Suppose instead we are given as initial data the potential difference between the plates. There is no charge between the plates since the dielectric is electrically neutral, so we have to solve

$$\nabla^2 V = 0. \tag{5.15}$$

It is clear from the setup that the potential can only depend on z, as this is the only physically relevant dimension to our problem. Thus, we have

$$\nabla^2 V \to \frac{d^2V}{dz^2} = 0. \tag{5.16}$$

The solution to this is clearly

$$V(z) = az + b, \tag{5.17}$$

where a and b are two constants to be determined. Let us declare the bottom plate to be at potential 0. Setting $z = 0$ in Eq.(5.17), we find that we need $b = 0$. If the potential at $z = h$ is $\Delta V = V(h) - V(0)$, then we find, with $z = h$ the distance between the plates,

$$\Delta V = ah \Rightarrow a = \frac{\Delta V}{h} \tag{5.18}$$

so that V is given everywhere between the plates by

$$V(z) = \frac{\Delta V}{h} z. \tag{5.19}$$

Note that the electric field is constant and along $\hat{z}$.

$$\vec{E} = -\vec{\nabla} V = -\frac{\Delta V}{h}\hat{z}. \tag{5.20}$$

[**Note:** if you're worried about the negative sign, don't be. We will find, in agreement with Eq.(5.13), that ΔV is actually negative. Thus the field is really along $+\hat{z}$, as expected.] It is possible to infer the charge on the bottom or the bottom plate using the electric boundary conditions:

$$D_n = \sigma_s. \tag{5.21}$$

In the dielectric, there is a $\vec{D}$ field but no surface charge density; in the conductor there is no $\vec{D}$ but a surface charge σ_s. At $z = h$, we have

$$D_n = \lambda_s = \varepsilon E_n(h) = -\varepsilon \frac{\Delta V}{h}, \tag{5.22}$$

which is simply another form of Eq.(5.13).

The capacitance then then be determined since $Q = \sigma_s S$. Hence, we reobtain Eq.(5.14).

5.3 The cylindrical capacitor

Consider one perfectly conducting cylinder of radius ρ_1, surrounded by another hollow cylinder, coaxial with the first and also perfectly conducting. The inner radius of the hollow cylinder is ρ_2; its outer radius is ρ_3. Both cylinders have infinite length.

This is the geometry of a coaxial cable. The space between them is filled with some generic dielectric with dielectric constant $\varepsilon = \varepsilon_r \varepsilon_0$. The dielectric is electrically neutral.

5.3.1 Starting with a charge distribution.

Suppose some constant surface charge density σ_1 is placed on the inner cable. By drawing a Gaussian cylinder with radius ρ greater than ρ_2 but smaller than ρ_3, one easily deduces from the condition of zero–static field inside a perfect conductor, that a surface charge density

$$\sigma_2 = -\frac{\rho_1}{\rho_2}\sigma_1 \tag{5.23}$$

must be present on the inner surface of the outer cylinder. The smaller surface charge density arises because the inner surface of the outer tube has a greater area than the surface of the inner cable.

To calculate the field $\vec{E}$ between the inner cable and inner surface of the outer cable, we use Gauss's flux theorem. The charge distribution does not depend on ϕ or z. Thus, by symmetry, it must be that the field is entirely in the $\hat{\rho}$ direction. We choose as our Gaussian surface a cylinder with radius ρ_p such that $\rho_1 < \rho_p < \rho_2$ so as to be recover the field at a point between the cables. The length of the Gaussian cylinder is ℓ.

There is no flux on front and back of the cylinder, because the field is always along $\hat{\rho}$ but $d\vec{S}$ is along $\hat{z}$ on those surfaces. Thus, the only contribution to the flux is from circular side of the cylinder.

By symmetry, the field is constant in magnitude on the side of the cylinder

$$\oint_S \vec{D}\cdot d\vec{S} = \int_{\text{side}} D(\rho_p)dS = D(\rho_p)\times 2\pi\rho_p \times \ell = \sigma_s \times 2\pi\rho_1 \times \ell, \tag{5.24}$$

from which we conclude

$$\vec{E}(\rho_p) = \frac{\sigma_s\rho_1}{\varepsilon}\frac{1}{\rho_p}\hat{\rho}. \tag{5.25}$$

Knowing the field, we can find the potential difference between the cylinders. Using a radial path along $\hat{\rho}$ such that $d\vec{l} = \hat{\rho}d\rho$, we find, by simple integration, the potential difference ΔV to be

$$\Delta V = V(\rho_1) - V(\rho_2) = -\int_{\rho_2}^{\rho_1}\vec{E}\cdot d\vec{l} = \int_{\rho_1}^{\rho_2}\vec{E}\cdot d\vec{l} = \frac{\sigma_s\rho_1}{\varepsilon}\ln\frac{\rho_2}{\rho_1} \tag{5.26}$$

The capacitance per unit length is therefore given by

$$\mathcal{C} = \frac{C}{\ell} = \frac{Q/\ell}{\Delta V} = \frac{\sigma_s\rho_1}{\frac{\sigma_s\rho_1}{\varepsilon}\ln\frac{\rho_2}{\rho_1}} = \frac{\varepsilon_r\varepsilon_0}{\ln\frac{\rho_2}{\rho_1}}, \tag{5.27}$$

Once again, the charge and voltage have disappeared from this (simple) result.

Eq.(5.27) displays the essential properties of the cylindrical capacitor. The voltage difference is a function of the ratio ρ_2/ρ_1 The $\vec{E}$ field varies radially as $1/\rho$, so the field can reach quite large magnitudes near the inner cable if the radius ρ_1 of the inner cable is small.

5.3.2 Starting with a voltage difference.

Suppose a potential difference ΔV is maintained between the inner cable and the inner surface of the outer cylinder. The geometry of the problem indicates that V will be a function of ρ only, since the surface of a conducting cylinder, irrespective of the angle ϕ, is an equipotential. Thus, we set $V = V(\rho)$ and solve

$$\frac{1}{\rho}\frac{d}{d\rho}\left(\rho\frac{dV}{d\rho}\right) = 0, \tag{5.28}$$

as there is no charge between the cylinders.

Since $\rho = 0$ is excluded on the grounds that we are looking for V in the region $\rho_1 < \rho < \rho_2$, we can multiply Eq.(5.28) by $\rho \neq 0$ to obtain

$$\frac{d}{d\rho}\left(\rho\frac{dV}{d\rho}\right) = 0 \Rightarrow \rho\frac{dV}{d\rho} = C_1, \tag{5.29}$$

where C_1 is an integration constant. Dividing by $\rho \neq 0$, we now obtain

$$\frac{dV}{d\rho} = \frac{C_1}{\rho} \Rightarrow V(\rho) = C_1 \ln\rho + C_2, \tag{5.30}$$

with C_2 another integrating constant.

To obtain the correct potential difference $V(\rho_1) - V(\rho_2)$, we need

$$\begin{aligned} V(\rho_1) - V(\rho_2) &= C_1 \ln\rho_1 + C_2 - (C_1 \ln\rho_2 + C_2) && (5.31)\\ &= C_1 \ln\frac{\rho_1}{\rho_2}, && (5.32) \end{aligned}$$

which implies

$$\begin{aligned} C_1 &= \frac{\Delta V}{\ln\frac{\rho_1}{\rho_2}}, && (5.33)\\ V(\rho) &= \left(\frac{\Delta V}{\ln\frac{\rho_1}{\rho_2}}\right)\ln\rho + C_2. && (5.34) \end{aligned}$$

At ρ_1, we have

$$V(\rho_1) - \left(\frac{V(\rho_1) - V(\rho_2)}{\ln\frac{\rho_1}{\rho_2}}\right)\ln\rho_1 = C_2. \tag{5.35}$$

Thus,

$$\begin{aligned} V(\rho) &= \left(\frac{\Delta V}{\ln\frac{\rho_1}{\rho_2}}\right)\ln\rho + V(\rho_1) - \left(\frac{\Delta V}{\ln\frac{\rho_1}{\rho_2}}\right)\ln\rho_1, && (5.36)\\ &= -\left(\frac{\Delta V}{\ln\frac{\rho_2}{\rho_1}}\right)\ln\frac{\rho}{\rho_1} + V(\rho_1). && (5.37) \end{aligned}$$

Now, we can recover the field via

$$\vec{E}(\rho) = -\vec{\nabla}V(\rho), \tag{5.38}$$

$$= -\hat{\rho}\frac{dV}{d\rho} = \hat{\rho}\left(\frac{\Delta V}{\ln\frac{\rho_2}{\rho_1}}\right)\frac{1}{\rho}. \tag{5.39}$$

Likewise, the charge density on the inner cylinder

$$D_n = \sigma_s = \varepsilon\left(\frac{\Delta V}{\ln\frac{\rho_2}{\rho_1}}\right)\frac{1}{\rho_1}, \tag{5.40}$$

in agreement with Eq.(5.26).

5.4 The spherical capacitor

Consider one perfectly conducting sphere of radius r_1, surrounded by another hollow concentric sphere, also perfectly conducting. The inner radius of the hollow sphere is r_2; its outer radius is r_3.

The space between them is filled with some generic dielectric with dielectric constant $\varepsilon = \varepsilon_r\varepsilon_0$. The dielectric is electrically neutral.

5.4.1 Starting with a charge distribution.

Suppose some constant surface charge density σ_1 is placed on the inner sphere. By drawing a Gaussian sphere with radius r greater than r_2 but smaller than r_3, one easily deduces from the condition of zero–static field inside a perfect conductor, that a surface charge density

$$\sigma_2 = -\frac{r_1^2}{r_2^2}\sigma_1 \tag{5.41}$$

must be present on the inner surface of the outer sphere. The smaller surface charge density arises because the inner surface of the outer sphere has a greater area than the surface of the inner cable.

To calculate the field $\vec{E}$ between the inner sphere and outer surface of the outer sphere, we use Gauss's flux theorem. The charge distribution does not depend on ϕ or θ. Thus, by symmetry, it must be that the field is entirely in the $\hat{r}$ direction. We choose as our Gaussian surface a sphere with radius r_p such that $r_1 < r_p < r_2$ so as to be recover the field at a point between the spheres.

By symmetry, the field is constant in magnitude on the surface of the Gaussian sphere, so

$$\oint_S \vec{D}\cdot d\vec{S} = \int D(r_p)dS = D(r_p)\times 4\pi r_p^2 = \sigma_s \times 4\pi r_1^2, \tag{5.42}$$

from which we conclude

$$\vec{E}(r_p) = \frac{\sigma_s r_1^2}{\varepsilon}\frac{1}{r_p^2}\hat{r}. \tag{5.43}$$

Knowing the field, we can find the potential difference between the spheres. Using a radial path along $\hat{r}$ such that $d\vec{l} = \hat{r}dr$, we find, by simple integration, the potential difference ΔV to be

$$\begin{aligned}\Delta V &= V(r_1) - V(r_2) = -\int_{r_2}^{r_1} \vec{E}\cdot d\vec{l} = \int_{r_1}^{r_2} \vec{E}\cdot d\vec{l}, \\ &= \frac{\sigma_s r_1^2}{\varepsilon}\left(\frac{1}{r_1} - \frac{1}{r_2}\right) = \frac{\sigma_s r_1}{\varepsilon r_2}(r_2 - r_1) \qquad (5.44)\end{aligned}$$

The capacitance is therefore given by

$$C = \frac{Q}{\Delta V} = \frac{\sigma_s 4\pi r_1^2}{\frac{\sigma_s r_1}{\varepsilon r_2}(r_2 - r_1)} = \frac{4\pi\varepsilon r_1 r_2}{(r_2 - r_1)}, \qquad (5.45)$$

Once again, the charge and the voltage disappear from the result.

Eq.(5.45) displays the essential properties of a parallel plate capacitor. The voltage difference is a function of the difference $r_2 - r_1$. The $\vec{E}$ field varies radially as $1/r_p^2$, as if the entire charge of the inner sphere was concentrated at the origin.

5.4.2 Starting with a voltage difference.

Suppose a potential difference ΔV is maintained between the inner sphere and the inner surface of the outer sphere. The geometry of the problem indicates that V will be a function of r only, since the surface of any one of the conducting spheres, irrespective of the angles θ and ϕ, is an equipotential. Thus, we set $V = V(r)$ and solve

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial V}{\partial r}\right) = 0, \qquad (5.46)$$

as there is no charge between the inner and outer spheres.

Since $r = 0$ is excluded on the grounds that we are looking for V in the region $r_1 < r < r_2$, we can multiply Eq.(5.46) by $r \neq 0$ to obtain

$$\frac{d}{dr}\left(r^2\frac{dV}{dr}\right) = 0 \Rightarrow r^2\frac{dV}{dr} = -C_1, \qquad (5.47)$$

where C_1 is an integration constant. The negative sign is for convenience. Dividing by $r^2 \neq 0$, we now obtain

$$\frac{dV}{dr} = -\frac{C_1}{r^2} \Rightarrow V(r) = \frac{C_1}{r} + C_2, \qquad (5.48)$$

with C_2 another integrating constant.

To obtain the correct potential difference $V(r_1) - V(r_2)$, we need

$$\begin{aligned}V(r_1) - V(r_2) &= \frac{C_1}{r_1} + C_2 - \left(\frac{C_1}{r_2} + C_2\right) \qquad (5.49)\\ &= C_1\left(\frac{1}{r_1} - \frac{1}{r_2}\right) = C_1\frac{(r_2 - r_1)}{r_1 r_2}, \qquad (5.50)\end{aligned}$$

which implies

$$C_1 = \frac{\Delta V}{\frac{r_2 - r_1}{r_2 r_1}} = \frac{\Delta V}{(r_2 - r_1)} r_1 r_2, \tag{5.51}$$

$$V(r) = \frac{\Delta V}{(r_2 - r_1)} r_1 r_2 \frac{1}{r} + C_2. \tag{5.52}$$

At r_1, we have

$$C_2 = V(r_1) - \frac{\Delta V}{(r_2 - r_1)} r_1 r_2 \frac{1}{r_1}, \tag{5.53}$$

$$= V(r_1) - \frac{V(r_1) - V(r_2)}{(r_2 - r_1)} r_2 \tag{5.54}$$

Thus,

$$V(r) = \frac{\Delta V}{(r_2 - r_1)} r_1 r_2 \frac{1}{r} + V(r_1) - \frac{\Delta V}{(r_2 - r_1)} r_2, \tag{5.55}$$

$$= -\frac{\Delta V}{(r_2 - r_1)} r_1 r_2 (\frac{1}{r_1} - \frac{1}{r}) + V(r_1). \tag{5.56}$$

We can verify this works by setting $r = r_2$ to find

$$V(r_2) = -\frac{\Delta V}{(r_2 - r_1)} r_1 r_2 (\frac{1}{r_1} - \frac{1}{r_2}) + V(r_1), \tag{5.57}$$

$$= -\frac{\Delta V}{(r_2 - r_1)} r_1 r_2 \frac{(r_2 - r_1)}{r_1 r_2} + V(r_1), \tag{5.58}$$

$$= -(V(r_1) - V(r_2)) + V(r_1) = V(r_2)\,. \tag{5.59}$$

Now, we can recover the field via

$$\vec{E} = -\vec{\nabla} V(r), \tag{5.60}$$

$$= -\hat{r}\frac{dV}{dr} = -\hat{r}\left(\frac{\Delta V}{(r_2 - r_1)} r_1 r_2\right)\left(-\frac{1}{r^2}\right), \tag{5.61}$$

$$= \hat{r}\left(\frac{\Delta V}{(r_2 - r_1)} r_1 r_2\right)\left(\frac{1}{r^2}\right) \tag{5.62}$$

Likewise, the charge density on the inner sphere

$$D_n = \sigma_s = \varepsilon\left(\frac{\Delta V}{(r_2 - r_1)} r_1 r_2\right)\frac{1}{r_1^2}, \tag{5.63}$$

$$= \frac{\varepsilon \Delta V}{(r_2 - r_1)} \frac{r_2}{r_1} \tag{5.64}$$

in agreement with Eq.(5.44).

6 Magnetostatics

6.1 Definition of $\vec{H}$

In magnetostatics, the source of the magnetic field is not a charge, but a time–independent current. If a constant current I flows in a small piece of wire oriented along $\vec{dl_s}$, the product $I\vec{dl_s}$ functions as a small source producing a small magnetic field defined by

$$d\vec{H} = \frac{1}{4\pi} I\vec{dl_s} \times \frac{(\vec{r}_p - \vec{r}_s)}{|\vec{r}_p - \vec{r}_s|^3}. \tag{6.1}$$

This is know as the law of Biot-Savart.

The net magnetic field is thus

$$\vec{H} = \int_L \frac{1}{4\pi} I\vec{dl_s} \times \frac{(\vec{r}_p - \vec{r}_s)}{|\vec{r}_p - \vec{r}_s|^3}, \tag{6.2}$$

where the integral extends over the region where the current is flowing.

Compare the situation to electrostatics, where a small stationary linear charge density λ_s produces in vacuum a small electric field

$$d\vec{E} = \frac{1}{4\pi\varepsilon_0} \lambda_s dl_s \frac{(\vec{r}_p - \vec{r}_s)}{|\vec{r}_p - \vec{r}_s|^3}. \tag{6.3}$$

Notice how Eq.(6.3) and Eq.(6.1) are similar, but also *how they are different.* In the expression for the electric field, the source term $\lambda_s dl_s$ is a scalar, *i.e.* a quantity that has no orientation. In the expression for the magnetic field, the source term $I\vec{dl_s}$ is a vector, i.e. an oriented quantity. This is because one must specify the direction of flow of the current. The vector $\vec{dl_s}$ required in the calculation of the $\vec{H}$-field is always tangent to the current-carrying wire because the current always flows tangentially to the wire. Alternatively, $\vec{dl_s}$ always points in the direction of local current flow. Notice also the cross product in the construction of $\vec{H}$.

In electrostatics, we have, besides linear charges densities, surface and volume charge densities. Likewise in magnetostatics, where a surface current density $\vec{K}_s$ and a volume current density $\vec{J}_s$ (this is, basically, the current density of Ohm's law $\vec{J} = \sigma\vec{E}$, with σ the conductivity) produce magnetic fields

$$\vec{H} = \int_S \frac{1}{4\pi} \left(\vec{K}_s dS\right) \times \frac{(\vec{r}_p - \vec{r}_s)}{|\vec{r}_p - \vec{r}_s|^3}, \tag{6.4}$$

$$\vec{H} = \int_v \frac{1}{4\pi} \left(\vec{J}_s dv\right) \times \frac{(\vec{r}_p - \vec{r}_s)}{|\vec{r}_p - \vec{r}_s|^3}, \tag{6.5}$$

respectively.

Here, note how the vectorial nature of the current is described by the vector $\vec{K}_s$ and the scalar dS, or the vector $\vec{J}_s$ and the scalar dv. This should be constrasted with the scalar I and the vector $\vec{dl_s}$. If History were completely consistent, one should probably have $\vec{I}dl_s$ rather than $I\vec{dl_s}$, but the long-time convention of using $I\vec{dl_s}$, $\vec{K}_s dS$ and $\vec{J}_s dv$ remains.

6.2 Example: the infinite straight line

As an example, consider the problem of finding the $\vec{H}$-field generated by an infinitely long, straight wire, carrying a constant current I. Suppose for simplicity that the wire is stretched along $\hat{z}$, and let us pick a field point $\vec{r}_p = (x_p, 0, 0)$.

To use Eq.(6.1), we need to deduce that $d\vec{l}_s = \hat{z} dz_s$, and that the integral will be from $z_s = -\infty$ to $z_s = +\infty$, since this is the range of z over which the sources are distributed.

Then, we have

$$\begin{aligned} d\vec{H} &= \frac{1}{4\pi} I dz_s \hat{z} \times \frac{(x_p \hat{x} - z_s \hat{z})}{\left|x_p^2 + z_s^2\right|^{3/2}}, && (6.6) \\ &= \frac{1}{4\pi} \frac{I dz_s}{\left|x_p^2 + z_s^2\right|^{3/2}} \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ 0 & 0 & 1 \\ x_p & 0 & -z_s \end{vmatrix}, && (6.7) \\ &= \frac{1}{4\pi} \frac{I dz_s}{\left|x_p^2 + z_s^2\right|^{3/2}} \left(x_p \hat{y}\right). && (6.8) \end{aligned}$$

It is easy to see that the direction is qualitatively correct, because

$$\hat{z} \times \hat{x} = \hat{y}, \qquad \hat{z} \times \hat{z} = 0. \tag{6.9}$$

There remains to integrate:

$$\vec{H} = \frac{1}{4\pi} I x_p \hat{y} \int_{-\infty}^{\infty} \frac{dz_s}{\left|x_p^2 + z_s^2\right|^{3/2}}. \tag{6.10}$$

This is an integral of a type that we have seen before. Using the trigonometric substitution

$$z_s = x_p \tan\theta, \qquad dz_s = \frac{x_p}{\cos^2\theta} d\theta, \tag{6.11}$$

and the equality

$$1 + \tan^2\theta = \frac{1}{\cos^2\theta} \tag{6.12}$$

we transform

$$\begin{aligned} \int \frac{dz_s}{\left|x_p^2 + z_s^2\right|^{3/2}} &= \int \frac{1}{\left|x_p^2 \left(1 + \tan^2\theta\right)\right|^{3/2}} \frac{x_p}{\cos^2\theta} d\theta, && (6.13) \\ &= \frac{x_p}{|x_p|^3} \int \cos\theta d\theta = \frac{x_p}{|x_p|^3} \sin\theta. && (6.14) \end{aligned}$$

We need to convert back to the original variables, so we construct an auxiliary triangle of slope z_s/x_p from Eq.(6.11), and find

$$\sin\theta = \frac{z_s}{\sqrt{x_p^2 + z_s^2}}. \tag{6.15}$$

Hence,

$$\begin{aligned}
\vec{H} &= \frac{1}{4\pi} I x_p \hat{y} \int_{-\infty}^{\infty} \frac{dz_s}{\left|x_p^2 + z_s^2\right|^{3/2}} && (6.16)\\
&= \frac{1}{4\pi} I x_p \hat{y} \frac{x_p}{|x_p|^3} \frac{z_s}{\sqrt{x_p^2 + z_s^2}} && (6.17)\\
&= \frac{I}{2\pi} \frac{\hat{y}}{|x_p|}. && (6.18)
\end{aligned}$$

In obtaining this last equation, we have implicitly assumed that $x_p > 0$: the slope of our auxiliary triangle was taken as positive, so that the angle θ in Eq.(6.15) was assumed positive. If $x_p < 0$, we have a negative slope to our triangle, and so a negative angle θ. In summary, we have the result:

$$\vec{H} = \begin{cases} +\hat{y} \frac{I}{2\pi |x_p|} & if \quad x_p > 0, \\ -\hat{y} \frac{I}{2\pi |x_p|} & if \quad x_p < 0. \end{cases} \tag{6.19}$$

A more general result, valid for $\vec{r}_p = (r_p \cos\phi_p, r_p \sin\phi_p, 0)$, gives

$$\vec{H} = \frac{I}{2\pi r_p} \hat{\phi}, \tag{6.20}$$

where $\hat{\phi}$ is a the unit vector at $\vec{r}_p$. This should be compared with the electric field produced by a infinite wire with linear charge density, which we recall to be

$$\vec{E} = \frac{\lambda_s}{2\pi r_p} \hat{\rho} \tag{6.21}$$

6.3 Example: the current loop

Consider a loop of radius r, located in the xy plane. The loop carries a current I. The line element $d\vec{l}_s$ tangent to the point $\vec{r}_s = (r\cos\theta_s, r\sin\theta_s, 0)$ on the loop is simply

$$d\vec{l}_s = (-r\sin\theta_s d\theta_s, r\cos\theta_s d\theta_s, 0) = (-r\sin\theta_s, r\cos\theta_s, 0) d\theta_s \tag{6.22}$$

To find the $\vec{H}$ field at a point $\vec{r}_p = (0, 0, z_p)$, we first construct

$$\begin{aligned}
d\vec{H} &= \frac{1}{4\pi} I d\vec{l}_s \times \frac{(\vec{r}_p - \vec{r}_s)}{|\vec{r}_p - \vec{r}_s|^3}, && (6.23)\\
&= \frac{1}{4\pi} \frac{I}{\left|r^2 + z_p^2\right|^{3/2}} \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ -r\sin\theta_s d\theta_s & r\cos\theta_s d\theta_s & 0 \\ -r\cos\theta_s & -r\sin\theta_s & z_p \end{vmatrix}, && (6.24)\\
&= \frac{1}{4\pi} \frac{I}{\left|r^2 + z_p^2\right|^{3/2}} \left(\hat{x} z_p r \cos\theta_s d\theta_s - \hat{y}\hat{x} z_p r \sin\theta_s d\theta_s + \hat{z} r^2 d\theta_s \right). && (6.25)
\end{aligned}$$

We will look at the each component of $\vec{H}$ in turn.

For H_x, we have

$$H_x = \frac{1}{4\pi}\frac{I}{\left|r^2+z_p^2\right|^{3/2}}\int_0^{2\pi} z_p r\cos\theta_s d\theta_s = 0 \tag{6.26}$$

since $\int_0^{2\pi}\cos\theta d\theta = 0$. The same result holds for H_y, this time due to $\int_0^{2\pi}\sin\theta d\theta = 0$. There remains H_z :

$$\begin{aligned} H_z &= \frac{1}{4\pi}\frac{I}{\left|r^2+z_p^2\right|^{3/2}}\int_0^{2\pi} r^2 d\theta_s, & (6.27)\\ &= \frac{1}{4\pi}\frac{Ir^2 2\pi}{\left|r^2+z_p^2\right|^{3/2}} = \frac{1}{2}\frac{Ir^2}{\left|r^2+z_p^2\right|^{3/2}} & (6.28)\end{aligned}$$

so that our final answer is

$$\vec{H} = \frac{1}{2}\frac{Ir^2}{\left|r^2+z_p^2\right|^{3/2}}\hat{z}. \tag{6.29}$$

Note that, at large distances $z_p \gg r$, we have the approximate expression

$$\vec{H} \approx \frac{1}{2}\frac{Ir^2}{z_p^3}\hat{z}. \tag{6.30}$$

The $1/z_p^3$ dependence for large distance is similar to the large distance dependence of E field of an electric dipole. In fact, a current loop is the prototype of a *magnetic* dipole.

6.4 The infinite sheet

Consider finally an infinite sheet, placed in the xy plane, and carrying a surface current density $\vec{K}_s = K_0\hat{x}$. A little piece of this sheet, located at $r_s = (x_s, y_s, 0)$ and having area $dS = dx_s dy_s$, will produce at a point $\vec{r}_p = (0, 0, z_p)$ a small field given by the infinitesimal version of Eq.(6.4):

$$\begin{aligned} d\vec{H} &= \frac{1}{4\pi}\left(K_0\hat{x}dx_s dy_s\right)\times\frac{\left(z_p\hat{z}-x_s\hat{x}-y_s\hat{y}\right)}{\left|x_s^2+y_s^2+z_p^2\right|^{3/2}} & (6.31)\\ &= \frac{K_0}{4\pi}\frac{dx_s dy_s}{\left|x_s^2+y_s^2+z_p^2\right|^{3/2}}\begin{vmatrix}\hat{x} & \hat{y} & \hat{z}\\ 1 & 0 & 0\\ -x_s & -y_s & z_p\end{vmatrix}, & (6.32)\\ &= \frac{K_0}{4\pi}\frac{dx_s dy_s}{\left|x_s^2+y_s^2+z_p^2\right|^{3/2}}\left(-\hat{y}z_p - \hat{z}y_s\right). & (6.33)\end{aligned}$$

Calculations are simplified by introducing polar coordinates in the plane,

$$\begin{aligned} x_s &= r_s\cos\phi_s, & (6.34)\\ y_s &= r_s\sin\phi_s, & (6.35)\end{aligned}$$

so that, under this change of variables,

$$dx_s dy_s = r_s dr_s d\phi_s. \tag{6.36}$$

Hence,

$$d\vec{H} = \frac{K_0}{4\pi} \frac{r_s dr_s d\phi_s}{\left|r_s^2 + z_p^2\right|^{3/2}} \left(-\hat{y} z_p - \hat{z} r_s \sin\phi_s\right). \tag{6.37}$$

Integrating first the z-component:

$$H_z = -\frac{K_0}{4\pi} \int_0^\infty \frac{r_s^2 dr_s}{\left|r_s^2 + z_p^2\right|^{3/2}} \int_0^{2\pi} d\phi_s \sin\phi_s = 0. \tag{6.38}$$

The only component is therefore along $\hat{y}$:

$$\begin{aligned} H_y &= -\frac{K_0 z_p}{4\pi} \int_0^\infty \frac{r_s dr_s}{\left|r_s^2 + z_p^2\right|^{3/2}} \int_0^{2\pi} d\phi_s && (6.39) \\ &= -\frac{K_0 z_p}{4\pi} 2\pi \int_0^\infty \frac{r_s dr_s}{\left|r_s^2 + z_p^2\right|^{3/2}}, && (6.40) \\ &= -\frac{K_0 z_p}{2} \frac{-1}{\sqrt{r_s^2 + z_p^2}} && (6.41) \\ &= -\frac{K_0 z_p}{2\left|z_p\right|}. && (6.42) \end{aligned}$$

Thus, we have

$$\vec{H} = \frac{K_0}{2} \times \begin{cases} -\hat{y} & if \quad z_p > 0, \\ +\hat{y} & if \quad z_p < 0. \end{cases} \tag{6.43}$$

This ought to be compare with the expression for the electric field of an infinite sheet of surface charge density σ_0, which we recall to be

$$\vec{E} = \frac{\sigma_0}{2} \times \begin{cases} +\hat{z} & if \quad z_p > 0, \\ -\hat{z} & if \quad z_p < 0. \end{cases} \tag{6.44}$$

7 The displacement current

We have seen that, through induction,

$$V_{emf}^{tr} = \oint_C \vec{E} \cdot d\vec{\ell} = -\int_S \frac{\partial \vec{B}}{\partial t} \cdot d\vec{S} \tag{7.1}$$

leads to

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}. \tag{7.2}$$

In other words, a changing magnetic field generates an electric field. Is it possible for a changing electric field to generate a magnetic field? The following argument illustrates how the answer to the previous question must be positive.

Consider a piece of circuit containing a capacitor, and assume that the plates of the capacitor are separated by the vacuum. Everyday experience shows that a very legitimate circuit can be constructed even though both sides formally form an open circuit (there is no conductor connecting the plates). Thus, something must "flow" between the plates of the capacitor, although by construction it cannot be a conduction current.

To make the argument quantitative, consider the problem of calculating the B field generated by the current feeding the capacitor. Suppose we want to apply Ampère's law to this problem, so imagine an Amperian loop around the wire.

The following problem then arises: how are we to compute the current enclosed by the loop? This concerns arises because there are two ways of constructing a surface around our loop. The first way is straightforward: simply stretch an imaginary cellophane sheet immediately between the loop. Is this case, the enclose current is the (conduction) current carried by the wire. The second way is valid only because the plates of the capacitor are not physically connected: imagine that you place some soap film between your loop and that this soap film is deformed so that it forms a cylinder with the back of the cylinder going precisely between the capacitor plates. For this type of construction, there is never any physical current that "punctures" the surface defined by the Amperian loop!

Clearly, there is only one unique magnetic field, so both construction should be equivalent: what is going on?

Now, it is true that the second construction does not enclose any physical conduction current, but, contrary to the first case, there is an electric flux through this second surface. If we assume the capacitor is perfect so we can neglect fringing, then the electric field between the plates is

$$|\vec{D}| = \frac{Q}{A}, \tag{7.3}$$

and the electric field outside is 0, so the flux of the electric field through this second surface is

$$\Phi_D = |\vec{D}|A = Q. \tag{7.4}$$

The change in flux is then

$$\frac{d\Phi_D}{dt} = \frac{dQ}{dt}, \tag{7.5}$$

and

$$\frac{dQ}{dt} = I \tag{7.6}$$

is the conduction current through the wire feeding the capacitor. Thus, if we use our first surface to calculate the magnetic field,

$$\begin{aligned}
\oint_C \vec{H} \cdot d\vec{\ell} &= I = \frac{d\Phi_D}{dt} && (7.7) \\
&= \oint_S \frac{\partial \vec{D}}{\partial t} \cdot d\vec{S}, && (7.8) \\
&= \oint_S \left(\vec{\nabla} \times \vec{H} \right) \cdot d\vec{S} && (7.9)
\end{aligned}$$

from which we conclude that

$$\vec{\nabla} \times \vec{H} = \frac{\partial \vec{D}}{\partial t}. \tag{7.10}$$

This is valid clearly "inside" the capacitor, where there is no conduction current. The term $\frac{\partial \vec{D}}{\partial t}$ is called the displacement current density $\vec{J}_D$. In a region where there is a conduction current, we have, as before,

$$\vec{\nabla} \times \vec{H} = \vec{J}_c, \tag{7.11}$$

and in a region where there is a conduction current and a changing $\vec{D}$ field, we have the most general form of Ampère's law:

$$\vec{\nabla} \times \vec{H} = \vec{J}_c + \frac{\partial \vec{D}}{\partial t}. \tag{7.12}$$

In most practical case, one does not deal with $\vec{D}$ but rather with $\vec{E}$ which is tied to the voltage difference, so, remembering that

$$\begin{aligned}
\vec{D} &= \varepsilon \vec{E}, && (7.13) \\
\vec{J}_c &= \sigma \vec{E}, && (7.14)
\end{aligned}$$

and taking, typically, $\varepsilon \sim 10^{-12}$, we see that, unless $\vec{J}$is very small (i.e. the medium is a good insulator) or $\frac{\partial \vec{D}}{\partial t}$ very large, the conduction current is typically negligible. In a conductor, for instance, taking $\sigma \sim 10^7$, and $|\vec{E}| = E_0 \cos \omega t$, we have

$$\left| \frac{J_c}{J_D} \right| = \frac{10^7}{\omega \times 10^{-12}} = \frac{10^{19}}{\omega}. \tag{7.15}$$

8 The time-dependent Maxwell's equations

To summarize, we have:

$$\begin{array}{ll}
\vec{\nabla} \cdot \vec{D} = \rho_V & \oint_S \vec{D} \cdot d\vec{S} = q_{encl.} \\
\vec{\nabla} \times \vec{E} = -\frac{\partial B}{\partial t} & \oint \vec{E} \cdot d\ell = -\frac{\partial}{\partial t} \int_S \vec{B} \cdot d\vec{S} \\
\vec{\nabla} \cdot \vec{B} = 0 & \oint_S \vec{B} \cdot d\vec{S} = 0 \\
\vec{\nabla} \times \vec{H} = \vec{J}_c + \frac{\partial \vec{D}}{\partial t} & \oint \vec{H} \cdot d\vec{\ell} = \int_S \left(\vec{J}_c + \frac{\partial \vec{D}}{\partial t} \right) \cdot d\vec{S}
\end{array} \tag{8.1}$$

9 Plane wave propagation

9.1 Maxwell's equation in phasor form: time-harmonic fields

We consider the problem of solving Maxwell's equation under the assumption that the source distributions are time-harmonic. This allows us to use phasors and ignore the explicit harmonic time dependence and replace time derivative by a complex factor, when applicable. Thus, the four Maxwell equations become, in phasor form:

$$\begin{aligned}
&\vec{\nabla}\cdot\vec{D}=\rho_V && \Longrightarrow && \vec{\nabla}\cdot\vec{E}_s=\frac{\rho_{Vs}}{\varepsilon}\\
&\vec{\nabla}\times\vec{E}=-\frac{\partial\vec{B}}{\partial t} && \Longrightarrow && \vec{\nabla}\times\vec{E}_s=-j\omega\mu\vec{H}_s\\
&\vec{\nabla}\cdot\vec{H}=0 && \Longrightarrow && \vec{\nabla}\cdot\vec{H}_s=0\\
&\vec{\nabla}\times\vec{H}=\vec{J}_c+\frac{\partial\vec{D}}{\partial t} && \Longrightarrow && \vec{\nabla}\times\vec{H}_s=\vec{J}_{cs}+j\omega\varepsilon\vec{E}_s
\end{aligned}\tag{9.1}$$

We can manipulate these into a more convenient form. Using $\vec{J}_{cs}=\sigma\vec{E}_s$, we can rewrite

$$\vec{\nabla}\times\vec{H}_s = \sigma\vec{E}_s+j\omega\varepsilon\vec{E}_s \tag{9.2}$$

$$= j\omega\varepsilon_c\vec{E}_s, \tag{9.3}$$

where the complex permittivity

$$\varepsilon_c=\varepsilon-j\frac{\sigma}{\omega}=\varepsilon\left(1-j\frac{\sigma}{\omega\varepsilon}\right) \tag{9.4}$$

has been introduced.

9.2 Waves in a charge-free medium

Suppose $\rho_V=0$. Then, Maxwell's equations simplify to

$$\vec{\nabla}\cdot\vec{E}_s = 0, \tag{9.5}$$

$$\vec{\nabla}\times\vec{E}_s = -j\omega\mu\vec{H}_s, \tag{9.6}$$

$$\vec{\nabla}\cdot\vec{H}_s = 0, \tag{9.7}$$

$$\vec{\nabla}\times\vec{H}_s = j\omega\varepsilon_c\vec{E}_s. \tag{9.8}$$

Take the curl of $\vec{\nabla}\times\vec{E}_s$:

$$\vec{\nabla}\times\left(\vec{\nabla}\times\vec{E}_s\right) = -j\omega\mu\left(\vec{\nabla}\times\vec{H}_s\right)=-j\omega\mu\left(j\omega\varepsilon_c\vec{E}_s\right), \tag{9.9}$$

$$= \vec{\nabla}\left(\vec{\nabla}\cdot\vec{E}_s\right)-\nabla^2\vec{E}_s, \tag{9.10}$$

$$= -\nabla^2\vec{E}_s \tag{9.11}$$

using vector calculus identities and Maxwell's equations. This produces the final equation

$$-\nabla^2\vec{E}_s=\omega^2\mu\varepsilon_c\vec{E}_s. \tag{9.12}$$

Introducing the complex quantity

$$\gamma^2 = -\omega^2 \mu \varepsilon_c = -\omega^2 \mu \left(\varepsilon - j\frac{\sigma}{\omega} \right), \tag{9.13}$$

we finally obtain the (complex) phasor wave equation

$$\nabla^2 \vec{E}_s - \gamma^2 \vec{E}_s = 0. \tag{9.14}$$

Similarly, from $\vec{\nabla} \times \left(\vec{\nabla} \times \vec{H}_s \right)$ and Maxwell's equations, one obtains the wave equation for $\vec{H}_s$:

$$\nabla^2 \vec{H}_s - \gamma^2 \vec{H}_s = 0. \tag{9.15}$$

Both equations are obviously of the same general form. Note that Eqn.(9.14) is really three equations, as there is one equation for each component of $\vec{E}_s$, *i.e.* we really have

$$\begin{aligned} \nabla^2 E_{sx} - \gamma^2 E_{sx} &= 0, & (9.16) \\ \nabla^2 E_{sy} - \gamma^2 E_{sy} &= 0, & (9.17) \\ \nabla^2 E_{sz} - \gamma^2 E_{sz} &= 0, & (9.18) \end{aligned}$$

and similarly for $\vec{H}_s$. Each of the field components E_{sx}, E_{sy}, E_{sz} could, in principle, be a complex function.

By direct substitution, one can verify that, for some complex constant vectors $\vec{E}_0$, $\vec{H}_0$, the solutions to Eqns.(9.14) and (9.15) are

$$\begin{aligned} \vec{E}_s &= \vec{E}_0 e^{\pm\gamma \hat{n}\cdot\vec{r}}, & (9.19) \\ \vec{H}_s &= \vec{H}_0 e^{\pm\gamma \hat{n}\cdot\vec{r}} & (9.20) \end{aligned}$$

where

$$\begin{aligned} \hat{n}\cdot\vec{r} &= n_x x + n_y y + n_z z, & (9.21) \\ n_x^2 + n_y^2 + n_z^2 &= 1. & (9.22) \end{aligned}$$

At this stage, either choice of $+\gamma$ or $-\gamma$ is possible. To verify, take the Laplacian in cartesian coordinates:

$$\begin{aligned} \nabla^2 \vec{E}_s &= \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \vec{E}_0 e^{\pm\gamma(n_x x + n_y y + n_z z)}, & (9.23) \\ &= \vec{E}_0 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) e^{\pm\gamma(n_x x + n_y y + n_z z)}, & (9.24) \\ &= \vec{E}_0 \left(\gamma^2 n_x^2 + \gamma^2 n_y^2 + \gamma^2 n_z^2 \right) e^{\pm\gamma(n_x x + n_y y + n_z z)}, & (9.25) \\ &= \vec{E}_0 \gamma^2 e^{\pm\gamma(n_x x + n_y y + n_z z)} = \gamma^2 \vec{E}_0 e^{\pm\gamma \hat{n}\cdot\vec{r}}. & (9.26) \end{aligned}$$

Identical manipulations show that Eqn.(9.20) is also a solution to Eqn.(9.15).

The solution of Eqn.(9.19) is called the *plane* wave solution. This is because, on a fixed *plane* determined by $\hat{n}\cdot\vec{r}$ =constant, $\vec{E}_s$ and $\vec{H}_s$ have constant amplitude.

For simplicity, let us suppose for the moment $\hat{n}\cdot\vec{r}=z$ and explore in more details some of the properties of Eqn.(9.19). First, write $\gamma=\alpha+j\beta$. The physical fields are then given by

$$\vec{E}(z,t) = \Re\left(\vec{E}_0 e^{-\alpha z}e^{-j\beta z}e^{j\omega t}\right) = \vec{E}_0 e^{-\alpha z}\cos(\omega t-\beta z), \tag{9.27}$$

$$\vec{H}(z,t) = \Re\left(\vec{H}_0 e^{-\alpha z}e^{-j\beta z}e^{j\omega t}\right) = \vec{H}_0 e^{-\alpha z}\cos(\omega t-\beta z), \tag{9.28}$$

Some temporary assumptions about the sign of γ has been made, and we have assumed $\vec{E}_0$ and $\vec{H}_0$ real for simplicity. This shows that fields are a wave travelling in the $\hat{z}$ direction, a consequence of choosing the direction vector $\hat{n}$ to be in the $\hat{z}$ direction.

Let us now show that E_{sz} must be zero. To this end, use $\vec{\nabla}\times\vec{H}_s = j\omega\varepsilon_c\vec{E}_s$. The $\hat{z}$-component of this is

$$\left(\vec{\nabla}\times\vec{H}_s\right)_z = j\omega\varepsilon_c E_{sz} = \frac{\partial}{\partial x}H_{sy} - \frac{\partial}{\partial y}H_{sx} = 0 \tag{9.29}$$

since $\vec{H}_0$ is constant and thus $\vec{H}_0 e^{\pm\gamma z}$ does not depend on x or y. Likewise, from $\vec{\nabla}\times\vec{E}_s = -j\omega\mu\vec{H}_s$, one rapidly shows that H_{sz}.

Thus, for the plane wave solution, the electric and magnetic fields are completely perpendicular to the direction of propagation.

Finally, suppose

$$\vec{E}_0 = \hat{x}E_{0x}+\hat{y}E_{0y}. \tag{9.30}$$

Then, from

$$\vec{\nabla}\times\vec{E}_s = -j\omega\mu\vec{H}_s = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ E_{0x}e^{\pm\gamma z} & E_{0y}e^{\pm\gamma z} & 0 \end{vmatrix} \tag{9.31}$$

$$= -\hat{x}\left(\pm\gamma\right)E_{0y}e^{\pm\gamma z} + \hat{y}\left(\pm\gamma\right)E_{0x}e^{\pm\gamma z}, \tag{9.32}$$

$$= \left(\pm\gamma\right)\left(-\hat{x}E_{0y}+\hat{y}E\right)_{0x}e^{\pm\gamma z}. \tag{9.33}$$

Hence,

$$\vec{H}_s = j\frac{(\pm\gamma)}{\omega\mu}\left(-\hat{x}E_{0y}+\hat{y}E_{0x}\right)e^{\pm\gamma z} \tag{9.34}$$

so that

$$H_{0x} = -j\frac{(\pm\gamma)}{\omega\mu}E_{0y}, \tag{9.35}$$

$$H_{0y} = j\frac{(\pm\gamma)}{\omega\mu}E_{0x}. \tag{9.36}$$

In particular,

$$\vec{E}_0\cdot\vec{H}_0 = 0. \tag{9.37}$$

For the plane wave solution, $\vec{E}(z,t)$ is perpendicular to $\vec{H}(z,t)$, and both are perpendicular to the direction of propagation.

9.3 Plane waves in lossy media

In this section, we find the connection between γ and the constitutive parameters ε, μ, σ of the material in which the wave propagates.

9.3.1 General form

Recall

$$\gamma^2 = -\omega^2\mu\left(\varepsilon - j\frac{\sigma}{\omega}\right) = -\omega^2\mu\varepsilon\left(1 - j\frac{\sigma}{\omega\varepsilon}\right), \tag{9.38}$$

$$\gamma = -j\sqrt{\omega^2\mu\varepsilon}\sqrt{1 - j\frac{\sigma}{\omega\varepsilon}} \tag{9.39}$$

Write $\gamma = \alpha + j\beta$. To determine α, β,

$$\gamma^2 = \alpha^2 - \beta^2 + 2j\alpha\beta \tag{9.40}$$

$$= -\omega^2\mu\varepsilon + j\omega\varepsilon\sigma. \tag{9.41}$$

Equating the real and imaginary parts separately:

$$2\alpha\beta = \omega\varepsilon\sigma \Longrightarrow \beta = \frac{\omega\varepsilon\sigma}{2\alpha}, \tag{9.42}$$

$$\alpha^2 - \beta^2 = \alpha^2 - \left(\frac{\omega\varepsilon\sigma}{2\alpha}\right)^2 = -\omega^2\mu\varepsilon. \tag{9.43}$$

This last equation can be rewritten, after obvious manipulations, as

$$\alpha^4 + \alpha^2\omega^2\mu\varepsilon - \frac{1}{4}\omega^2\mu^2\sigma^2 = 0, \tag{9.44}$$

with solution

$$\alpha^2 = \frac{-\omega^2\mu\varepsilon \pm \sqrt{(\omega^2\mu\varepsilon)^2 + \omega^2\mu^2\sigma^2}}{2}. \tag{9.45}$$

Since α is assumed real, we must keep the positive root to guarantee that $\alpha^2 \geq 0$. Thus,

$$\alpha^2 = \frac{-\omega^2\mu\varepsilon + \sqrt{(\omega^2\mu\varepsilon)^2 + \omega^2\mu^2\sigma^2}}{2}, \tag{9.46}$$

and, after rearrangements

$$\alpha = \sqrt{\frac{\omega^2\mu\varepsilon}{2}}\left[\left(1 + \frac{\sigma^2}{\omega^2\varepsilon^2}\right)^{1/2} - 1\right]^{1/2}. \tag{9.47}$$

At this point, we have kept the positive root of the quadratic, thus making sure that α is positive. One should keep in mind that the negative root is also possible, albeit our choice will be justified later.

Using now Eqn.(9.45), we can solve for β^2 in Eqn.(9.42). This rapidly yields

$$\begin{aligned}
\beta^2 &= \alpha^2 + \omega^2\mu\varepsilon && (9.48)\\
&= \frac{1}{2}\omega^2\mu\varepsilon + \frac{1}{2}\sqrt{(\omega^2\mu\varepsilon)^2 + \omega^2\mu^2\sigma^2}, && (9.49)\\
&= \frac{1}{2}\omega^2\mu\varepsilon\left[\left(1+\frac{\sigma^2}{\omega^2\varepsilon^2}\right)^{1/2}+1\right], && (9.50)\\
\beta &= \sqrt{\frac{\omega^2\mu\varepsilon}{2}}\left[\left(1+\frac{\sigma^2}{\omega^2\varepsilon^2}\right)^{1/2}+1\right]^{1/2}. && (9.51)
\end{aligned}$$

The same comment applies to the choice of positive root for β.

Observe that the physical fields have the general form

$$\begin{aligned}
\vec{E}(z,t) &= \Re\left(\vec{E}_0 e^{\pm\alpha z} e^{\pm j\beta z} e^{j\omega t}\right), && (9.52)\\
\vec{H}(z,t) &= \Re\left(\vec{H}_0 e^{\pm\alpha z} e^{\pm j\beta z} e^{j\omega t}\right). && (9.53)
\end{aligned}$$

The constant vectors $\vec{E}_0$ and $\vec{H}_0$ have no $\hat{z}$ component, and in general, they will have both an $\hat{x}$ and $\hat{y}$ component, as per Eqn.(9.30). To continue and facilitate interpretation, we restrict to

$$\begin{aligned}
\vec{E}_0 &= \hat{x}\,|E_{0x}|\,e^{j\varphi_x}, && (9.54)\\
\vec{H}_0 &= \hat{y}\,|H_{0y}|\,e^{j\varphi_y}. && (9.55)
\end{aligned}$$

It is then clear that

$$\begin{aligned}
\vec{E}(z,t) &= \Re\left(\hat{x}\,|E_{0x}|\,e^{j\varphi_x} e^{\pm\alpha z} e^{\pm j\beta z} e^{j\omega t}\right), && (9.56)\\
&= \hat{x}\,|E_{0x}|\,e^{\pm\alpha z}\cos\left(\omega t \pm \beta z + \varphi_x\right). && (9.57)
\end{aligned}$$

If we keep the $\cos(\omega t + \beta z + \varphi_x)$, we have a wave travelling to from the right to the left. Indeed, suppose we have $t = t_0$ and $z = z_0$ such that $\omega t_0 + \beta z_0 + \varphi_x = 0$: the amplitude of the electric field is maximal for those values of t and z. If we are to stay at this maximum of amplitude, it must be that at some *later time* $t = t_1$, we have

$$\omega t_0 + \beta z_0 = \omega t_1 + \beta z_1 \Longrightarrow \omega(t_1 - t_0) = \beta(z_0 - z_1). \quad (9.58)$$

This indicates that $z_0 - z_1 > 0$, i.e the value of z decreases with increasing time: this properly describes a wave traveling from the right to the left. In the same spirit, one show that $\cos(\omega t - \beta z + \varphi_x)$ describes a wave traveling from the left to the right.

Simple physics dictates the choice between $e^{\alpha z}$ and $e^{-\alpha z}$.. Recall that have chosen $\alpha \geq 0$. The wave cannot gain amplitude as it travels away from the source. Thus, for a right-traveling wave, we keep $e^{-\alpha z}$ so that, with increasing z and later times, the amplitude decreases. For a left-traveling wave, we must keep $e^{\alpha z}$. Hence, our physical solutions are:

$$\vec{E}(z,t) = \begin{cases} \hat{x}\,|E_{0x}|\,e^{-\alpha z}\cos\left(\omega t - \beta z + \varphi_x\right): & \text{right-travelling wave,}\\ \hat{x}\,|E_{0x}|\,e^{\alpha z}\cos\left(\omega t + \beta z + \varphi_x\right): & \text{left-travelling wave.}\end{cases} \quad (9.59)$$

In other words, the right- and left-travelling waves correspond to $-\gamma$ and $+\gamma$, respectively.

The waves travel with phase velocity $u = \pm\omega/\beta$ (depending on the direction). This is the velocity at which one must travel to keep the phase of the cosine constant. The amplitude peaks are reproduced at every z_n such that $\beta z_n = 2n\pi$. Consecutive peaks are separated by one wavelength λ :

$$\lambda = z_{n+1} - z_n = \frac{2\pi}{\beta}. \tag{9.60}$$

9.3.2 Amplitude relations

For a right-travelling wave, we have, from Eqn. (9.36),

$$H_{0y} = j\frac{\gamma}{\omega\mu}E_{0x} = \frac{\beta - j\alpha}{\omega\mu}E_{0x}, \tag{9.61}$$

Thus, the amplitude of the magnetic field is related to the amplitude of the electric field by

$$H_{0y} = \frac{1}{\eta_c}E_{0x}, \tag{9.62}$$

where the complex impedence

$$\frac{1}{\eta_c} = j\frac{\gamma}{\omega\mu} = \frac{\beta - j\alpha}{\omega\mu} = \frac{1}{|\eta_c|}e^{-j\phi_\eta} \tag{9.63}$$

has been introduced. Using this, we find, for the physical magnetic field propagating to the right,

$$\begin{aligned} \vec{H}(z,t) &= \Re\left(\hat{y}\frac{|E_{0x}|}{|\eta_c|}e^{-j\phi_\eta}e^{j\varphi_x}e^{-\alpha z}e^{-j\beta z}e^{j\omega t}\right), && (9.64)\\ &= \hat{y}\frac{|E_{0x}|}{|\eta_c|}e^{-\alpha z}\cos\left(\omega t - \beta z + \varphi_x - \phi_\eta\right). && (9.65) \end{aligned}$$

Hence, there is, in general, a phase mismatch between the magnetic and electric fields, i.e. both do not reach their maxima (or minima) at the same time. The two fields are in phase when the impedence is real, i.e. when $\alpha = 0$. This can only occur in media where $\sigma = 0$, *i.e.* in perfectly non-conducting media. An identical conclusion is reached for left–traveling waves.

9.4 Three special cases.

It is clear, from Eqns.(9.4),(9.47) and (9.51) that the ratio $\frac{\sigma}{\omega\varepsilon}$ is of special importance. This ratio determines the imaginary part of the complex permittivity. We will distinguish three cases.

9.4.1 Lossless dielectric: $\sigma = 0$

If $\sigma = 0$, then we have

$$\begin{aligned}
\alpha &= \sqrt{\frac{\omega^2\mu\varepsilon}{2}}\left[(1+0)^{1/2}-1\right]^{1/2} = 0, && (9.66)\\
\beta &= \sqrt{\frac{\omega^2\mu\varepsilon}{2}}\left[(1+0)^{1/2}+1\right]^{1/2} = \sqrt{\omega^2\mu\varepsilon}, && (9.67)\\
\eta &= \frac{\omega\mu}{\beta} = \sqrt{\frac{\mu}{\varepsilon}}. && (9.68)
\end{aligned}$$

There is no attenuation. The special case of vacuum, where $\mu = \mu_0$ and $\varepsilon = \varepsilon_0$, produces the phase velocity $c = 1/\sqrt{\mu_0\varepsilon_0}$.

9.4.2 Low-loss dielectric: $\frac{\sigma}{\omega\varepsilon} \ll 1$.

In this case, we can use the binomial theorem:

$$(1+x)^n \approx 1 + nx + ... \tag{9.69}$$

valid for small x to approximate

$$\begin{aligned}
\alpha &= \sqrt{\frac{\omega^2\mu\varepsilon}{2}}\left[\left(1+\frac{\sigma^2}{\omega^2\varepsilon^2}\right)^{1/2}-1\right]^{1/2} \approx \sqrt{\frac{\omega^2\mu\varepsilon}{2}}\left[\left(1+\frac{\sigma^2}{2\omega^2\varepsilon^2}\right)-1\right]^{1/2}, \\
&\approx \sqrt{\frac{\omega^2\mu\varepsilon}{2}}\left[\frac{\sigma^2}{2\omega^2\varepsilon^2}\right]^{1/2} = \frac{\sigma}{2}\sqrt{\frac{\mu}{\varepsilon}}, && (9.70)\\
\beta &= \sqrt{\frac{\omega^2\mu\varepsilon}{2}}\left[\left(1+\frac{\sigma^2}{2\omega^2\varepsilon^2}\right)+1\right]^{1/2} \approx \omega\sqrt{\mu\varepsilon}, && (9.71)\\
\eta &= -j\frac{\omega\mu}{\gamma} = \frac{\omega\mu}{\sqrt{\omega^2\mu\varepsilon}}\left(1-j\frac{\sigma}{\omega\varepsilon}\right)^{-1/2} = \sqrt{\frac{\mu}{\varepsilon}}\left(1+j\frac{\sigma}{2\omega\varepsilon}\right) && (9.72)\\
&\approx \sqrt{\frac{\mu}{\varepsilon}}. && (9.73)
\end{aligned}$$

9.4.3 Good conductor: $\frac{\sigma}{\omega\varepsilon} \gg 1$.

In this case, we find, again using the binomial expansion,

$$\begin{aligned}
\alpha &= \sqrt{\frac{\omega^2\mu\varepsilon}{2}}\left[\left(1+\frac{\sigma^2}{\omega^2\varepsilon^2}\right)^{1/2}-1\right]^{1/2}, && (9.74)\\
&= \sqrt{\frac{\omega^2\mu\varepsilon}{2}}\left[\frac{\sigma}{\omega\epsilon}\left(1+\frac{\omega^2\varepsilon^2}{\sigma^2}\right)^{1/2}-1\right]^{1/2}, \\
&\approx \sqrt{\frac{\omega^2\mu\varepsilon}{2}}\left[\frac{\sigma}{\omega\epsilon}\left(1+\frac{\omega^2\varepsilon^2}{2\sigma^2}\right)-1\right]^{1/2} = \sqrt{\frac{\omega^2\mu\varepsilon}{2}}\sqrt{\frac{\sigma}{\omega\varepsilon}} = \sqrt{\frac{\omega\mu\sigma}{2}}, && (9.75)
\end{aligned}$$

$$
\begin{aligned}
\beta &= \sqrt{\frac{\omega^2\mu\varepsilon}{2}}\left[\left(1+\frac{\sigma^2}{\omega^2\varepsilon^2}\right)^{1/2}+1\right]^{1/2}, && (9.76)\\
&= \sqrt{\frac{\omega^2\mu\varepsilon}{2}}\left[\frac{\sigma}{\omega\epsilon}\left(1+\frac{\omega^2\varepsilon^2}{\sigma^2}\right)^{1/2}+1\right]^{1/2}, \\
&\approx \sqrt{\frac{\omega^2\mu\varepsilon}{2}}\left[\frac{\sigma}{\omega\epsilon}\left(1+\frac{\omega^2\varepsilon^2}{2\sigma^2}\right)+1\right]^{1/2} = \sqrt{\frac{\omega^2\mu\varepsilon}{2}}\sqrt{\frac{\sigma}{\omega\varepsilon}}, && (9.77)\\
&\approx \sqrt{\frac{\omega\mu\sigma}{2}} = \alpha\,, && (9.78)\\
\eta_c &= \sqrt{\frac{\mu}{\varepsilon}}\left(1-j\frac{\sigma}{\omega\varepsilon}\right)^{-1/2} = \sqrt{\frac{\mu}{\varepsilon}}\left(-j\frac{\sigma}{\omega\varepsilon}\right)^{-1/2}\left(1+j\frac{\omega\varepsilon}{\sigma}\right)^{-1/2}, && (9.79)\\
&\approx \sqrt{\frac{\mu\omega}{\sigma}}(-j)^{-1/2} = \sqrt{\frac{\mu\omega}{2\sigma}}\,(1+j)\ , && (9.80)
\end{aligned}
$$

as

$$
-j = \mathrm{e}^{3j\pi/2}\,, \quad -\frac{1}{j} = \mathrm{e}^{-3j\pi/2}\,, \quad \frac{1}{\sqrt{-j}} = \mathrm{e}^{-3j\pi/4} = \frac{1}{\sqrt{2}}\,(1+j)\ . \qquad (9.81)
$$

10 Multiple interfaces at normal incidence

We now consider the problem of reflection and transmission at multiple interfaces by normally incident plane waves. A typical problem would be one where a wave hits a film $(\epsilon_2, \sigma_2, \mu_2)$ of thickness d_2. The wave transmitted in the film will then hit another interface with $(\epsilon_3, \sigma_3, \mu_3)$ of infinite extent. Applications include anti-reflecting coatings and shielding.

In the three-interface problem described above, the situation can be worded as follows. An incident wave from $z = -\infty$, traveling in medium 1, strikes the surface of medium 2. Some of the wave will be reflected back to medium 1, some of the wave will be transmitted in medium 2. The transmitted wave in medium 2 then hits the 2-3 boundary: some of the wave will be transmitted to medium 3 and some will be reflected to medium 2. The backscattered wave in medium 2 will eventually reach the 1-2 interface: some of it will be transmitted back to medium 1, propagating toward $z = -\infty$ along with the wave reflected at the 1-2 interface. Some of the backscattered wave will bounce back in medium 2, reaching again the 2-3 boundary where some of it will be transmitted to medium 3 and some will be rebackscattered in medium 2. This bouncing clearly goes *ad infinitum.*

It would be possible but extremely annoying to actually sum all the partial amplitudes in various media, but there is a much more direct approach.

Considering the speed of the signal in each medium will not be significantly different from the speed of light for ordinary material (remember that, if $\sigma = 0$, $u_p = 1/\sqrt{\epsilon\mu} = c/\sqrt{\mu_r\epsilon_r}$, where $c \sim 3 \times 10^8$ m/s), the situation within each medium will rapidly evolve to a steady state: we will work from this steady state.

In this regime, we have the following:

In medium 1: one incident wave with total (complex) amplitude E_1^+, described by the phasor $\vec{E}_{s1}^+ = \hat{x}E_1^+ e^{-\alpha_1 z} e^{-j\beta_1 z}$, and one reflected wave with net amplitude E_1^- and phasor $\vec{E}_{s1}^- = \hat{x}E_1^- e^{\alpha_1 z} e^{j\beta_1 z}$. (I have chosen the electric field to be along $\hat{x}$ for convenience. This is not essential to the argument.) The net amplitude E_1^r is the sum of all the waves emerging from medium 2 to medium 1 from the multiple reflections in medium 2. The superscript on the phasor indicates the direction of propagation, and the subscript indicates the medium of propagation. The net electric field phasor in medium 1 is therefore

$$\vec{E}_{s1} = \hat{x}\left(E_1^+ e^{-\alpha_1 z} e^{-j\beta_1 z} + E_1^- e^{\alpha_1 z} e^{j\beta_1 z}\right) \tag{10.1}$$

The net magnetic field is obtained by observing

$$\begin{aligned}
\tilde{H}_1^+ &= \frac{1}{\eta_{c1}}(+\hat{z}) \times \left(\hat{x}E_1^+ e^{-\alpha_1 z} e^{-j\beta_1 z}\right) = \frac{1}{\eta_{c1}}\hat{y}E_1^+ e^{-\alpha_1 z} e^{-j\beta_1 z}, && (10.2)\\
\tilde{H}_1^- &= \frac{1}{\eta_{c1}}(-\hat{z}) \times \left(\hat{x}E_1^- e^{\alpha_1 z} e^{j\beta_1 z}\right) = \frac{1}{\eta_{c1}}(-\hat{y})\, E_1^- e^{\alpha_1 z} e^{j\beta_1 z}, && (10.3)\\
\tilde{H}_1 &= \frac{1}{\eta_{c1}}\hat{y}\left(E_1^+ e^{-\alpha_1 z} e^{-j\beta_1 z} - E_1^- e^{\alpha_1 z} e^{j\beta_1 z}\right). && (10.4)
\end{aligned}$$

In medium 2: one wave traveling to the right, which is a sum of the incident wave transmitted in 2 plus the multiplicity of waves reflected back in medium 2 at the 1-2 interface. The net amplitude of this wave is $\tilde{E}_2^+ = \hat{x}E_2^+ e^{-\alpha_2 z} e^{-j\beta_2 z}$. The wave traveling to the left is likewise given by $\vec{E}_{s2}^- = \hat{x}E_2^- e^{\alpha_2 z} e^{j\beta_2 z}$. Thus, the net electric field phasor in medium 2 is:

$$\vec{E}_{s2} = \hat{x}\left(E_2^+ e^{-\alpha_2 z} e^{-j\beta_2 z} + E_2^- e^{\alpha_2 z} e^{j\beta_2 z}\right). \tag{10.5}$$

The net magnetic field in medium to is also readily obtained:

$$\tilde{H}_2 = \frac{1}{\eta_{c2}}\hat{y}\left(E_2^+ e^{-\alpha_2 z} e^{-j\beta_2 z} - E_2^- e^{\alpha_2 z} e^{j\beta_2 z}\right). \tag{10.6}$$

In medium 3, there is only a transmitted wave, but it is the sum of numerous transmissions originating from repeated bounces at the 2-3 interface. Thus:

$$\vec{E}_{s3} = \hat{x}E_3^+ e^{-\alpha_3 z} e^{j\beta_3 z}, \tag{10.7}$$

$$\tilde{H}_3 = \frac{1}{\eta_{c3}}\hat{y}E_3^+ e^{-\alpha_3 z} e^{-j\beta_3 z}. \tag{10.8}$$

We now apply the boundary conditions at the various interfaces. At the 1-2 interface, we have, by continuity of the tangential component of the net electric field and continuity of the tangential component of the net magnetic field:

$$E_1^+ + E_1^- = E_2^+ + E_2^-, \tag{10.9}$$

$$\frac{1}{\eta_{c1}}\left(E_1^+ - E_1^-\right) = \frac{1}{\eta_{c2}}\left(E_2^+ - E_2^-\right). \tag{10.10}$$

These are obtained by equating Eqns. (10.1) and (10.5), and equating Eqns. (10.2) with (10.6) at $z = 0$, where the interface is located.

Applying now the electric and magnetic boundary conditions at the 2-3 interface, which is located at $z = d_2$, we obtain

$$E_2^+ e^{-\alpha_2 d_2} e^{-j\beta_2 d_2} + E_2^- e^{\alpha_2 d_2} e^{j\beta_2 d_2} = E_3^+ e^{-\alpha_3 d_2} e^{-j\beta_3 d_2}, \tag{10.11}$$

$$\frac{1}{\eta_{c2}}\left(E_2^+ e^{-\alpha_2 d_2} e^{-j\beta_2 d_2} - E_2^- e^{\alpha_2 d_2} e^{j\beta_2 d_2}\right) = \frac{1}{\eta_{c3}}E_3^+ e^{-\alpha_3 d_2} e^{-j\beta_3 d_2}. \tag{10.12}$$

We therefore have a system of four equations with five unknowns. It is usual to solve in terms of the incident amplitude E_1^+. After some straightforward but otherwise pretty tedious manipulations, we obtain the solution as:

$$\frac{E_1^-}{E_1^+} = \frac{(\eta_{c2} - \eta_{c1})(\eta_{c2} + \eta_{c3}) + \mathrm{e}^{-2\gamma_2 d_2}(\eta_{c1} + \eta_{c2})(\eta_{c3} - \eta_{c2})}{(\eta_{c1} + \eta_{c2})(\eta_{c2} + \eta_{c3}) + \mathrm{e}^{-2\gamma_2 d_2}(\eta_{c1} - \eta_{c2})(\eta_{c2} - \eta_{c3})}, \tag{10.13}$$

$$\frac{E_2^+}{E_1^+} = \frac{2\eta_{c2}(\eta_{c2} + \eta_{c3})}{(\eta_{c1} + \eta_{c2})(\eta_{c2} + \eta_{c3}) + \mathrm{e}^{-2\gamma_2 d_2}(\eta_{c1} - \eta_{c2})(\eta_{c2} - \eta_{c3})}, \tag{10.14}$$

$$\frac{E_2^-}{E_1^+} = \frac{2\eta_{c2}(\eta_{c3} - \eta_{c2})\,\mathrm{e}^{-2\gamma_2 d_2}}{(\eta_{c1} + \eta_{c2})(\eta_{c2} + \eta_{c3}) + e^{-2\gamma_2 d_2}(\eta_{c1} - \eta_{c2})(\eta_{c2} - \eta_{c3})}, \tag{10.15}$$

$$\frac{E_3^+}{E_1^+} = \frac{4\eta_{2c}\eta_{3c}\mathrm{e}^{-\gamma_2 d_2}}{(\eta_{c1} + \eta_{c2})(\eta_{c2} + \eta_{c3}) + \mathrm{e}^{-2\gamma_2 d_2}(\eta_{c1} - \eta_{c2})(\eta_{c2} - \eta_{c3})}. \tag{10.16}$$

These rather formidable expressions considerably simplify in one important case.

For three perfect dielectrics, we have $\alpha_1 = \alpha_2 = \alpha_3 = 0$ and all impedances are real. Thus, in this case,

$$E_1^- \rightarrow E_1^+ \frac{(\eta_2 - \eta_1)(\eta_2 + \eta_3) + e^{-2j\beta_2 d_2}(\eta_1 + \eta_2)(\eta_3 - \eta_2)}{(\eta_1 + \eta_2)(\eta_2 + \eta_3) + e^{-2j\beta_2 d_2}(\eta_1 - \eta_2)(\eta_2 - \eta_3)}, \tag{10.17}$$

$$
\begin{aligned}
E_2^+ &= E_1^+ \frac{2\eta_2\left(\eta_2+\eta_3\right)}{\left(\eta_1+\eta_2\right)\left(\eta_2+\eta_3\right)+e^{-2j\beta_2 d_2}\left(\eta_1-\eta_2\right)\left(\eta_2-\eta_3\right)}, &(10.18)\\
E_2^- &= E_1^+ \frac{2\eta_2\left(\eta_3-\eta_2\right)e^{-2j\beta_2 d_2}}{\left(\eta_1+\eta_2\right)\left(\eta_2+\eta_3\right)+e^{-2j\beta_2 d_2}\left(\eta_1-\eta_2\right)\left(\eta_2-\eta_3\right)}, &(10.19)\\
E_3^+ &= E_1^+ \frac{4\eta_2\eta_3\mathrm{e}^{-j\beta_2 d_2}}{\left(\eta_1+\eta_2\right)\left(\eta_2+\eta_3\right)+e^{-2j\beta_2 d_2}\left(\eta_1-\eta_2\right)\left(\eta_2-\eta_3\right)}. &(10.20)
\end{aligned}
$$

One application of (nearly perfect) dielectrics is for anti-coating purposes. If we demand that E_1^- be 0, i.e. no wave reflected back in medium 1, then we obtain the condition

$$
\begin{aligned}
E_1^- &= 0 \Rightarrow \left(\eta_2-\eta_1\right)\left(\eta_2+\eta_3\right)+\mathrm{e}^{-2j\beta_2 d_2}\left(\eta_1+\eta_2\right)\left(\eta_3-\eta_2\right)=0, &(10.21)\\
&\Rightarrow \left(\eta_2-\eta_1\right)\left(\eta_2+\eta_3\right) = -\mathrm{e}^{-2j\beta_2 d_2}\left(\eta_1+\eta_2\right)\left(\eta_3-\eta_2\right) &(10.22)
\end{aligned}
$$

Clearly, this can only happen when $e^{j\beta_2 d_2}$ is a real number since the left hand side of Eqn.(10.22) is real while the right hand side can be complex. Thus, the thickness of medium 2 must be such that $\mathrm{e}^{-2j\beta_2 d_2} = \pm 1$.

This does not completely determine the second medium. Assuming $\mathrm{e}^{-2j\beta_2 d_2} = +1$, the impedance of medium 2 must now be chosen so that

$$\left(\eta_1-\eta_2\right)\left(\eta_2+\eta_3\right)+\left(\eta_1+\eta_2\right)\left(\eta_2-\eta_3\right)=0, \tag{10.23}$$

which produces the solution $\eta_2 = 0$, obviously unacceptable. Assuming now $\mathrm{e}^{-2j\beta_2 d_2} = -1$, we obtain

$$\left(\eta_1-\eta_2\right)\left(\eta_2+\eta_3\right)-\left(\eta_1+\eta_2\right)\left(\eta_2-\eta_3\right)=0, \tag{10.24}$$

with solution $\eta_2 = \sqrt{\eta_1\eta_3}$. The thickness must then be adjusted so that

$$2\beta_2 d_2 = (2n+1)\,\pi. \tag{10.25}$$

The thinnest possible layers are therefore obtained for $2\beta_2 d_2 = \pi, 3\pi, 5\pi$ etc.

Because the numerator in Eqn.(10.13) depends on $\mathrm{e}^{-2\alpha_2 d_2}\mathrm{e}^{-2j\beta_2 d_2}$ but not on $\alpha_1, \beta_1, \alpha_3$ or β_3, the previous analysis for perfect dielectrics can be extended and remains valid is medium 2 is a low-loss dielectric, sufficiently thin so that $e^{-2\alpha_2 d_2} \approx 1$. It is then possible to obtained no reflection in the more general situation where media 1 and 3 are not low-loss dielectrics. The condition for zero reflection is then

$$\left(\eta_{c1}-\eta_2\right)\left(\eta_2+\eta_{c3}\right)+\mathrm{e}^{-2j\beta_2 d_2}\left(\eta_1+\eta_{c2}\right)\left(\eta_2-\eta_{c3}\right)=0, \tag{10.26}$$

and must be solved by breaking the equation into its real and imaginary parts. One then obtains two equations to determine η_2 and the thickness d_2 of the coating.

11 Propagation in the atmosphere: plasma frequency

In the upper atmosphere, one can find a significant quantity of ionized gas (or plasma), i.e. gas molecule in which some of the electrons are separated from the core of the molecule. These charged particles then react to the presence of an electric field and disturb the propagation of a signal.

One important parameter in the propagation of signals in the atmosphere is the so–called plasma frequency ω_p,defined by

$$\omega_p = \sqrt{\frac{N_e q_e^2}{m_e \varepsilon_0}}, \tag{11.1}$$

where N_e is the density of free electrons in the upper atmosphere, $q_e = 1.6 \times 10^{-19} C$ is the electronic charge and m_e is the electron mass. The effects of these free electrons is captured by the effective dielectric permittivity

$$\varepsilon_{\text{eff}} = \varepsilon_0 \left(1 - \frac{\omega_p^2}{\omega^2}\right) \tag{11.2}$$

where ω is the angular frequency of the signal. Note that ε_{eff} is real but could be negative if the angular frequency of the signal is below the plasma frequency.

In general, the electric field plane wave will therefore have the form (assuming propagation along $\hat{z}$)

$$\vec{E}_x(z) = \hat{x} E_{s0} e^{-j\beta z}, \tag{11.3}$$

where

$$\beta = \omega \sqrt{\mu_0 \varepsilon_{\text{eff}}}. \tag{11.4}$$

Thus, if $\varepsilon_{\text{eff}} > 0$, the electric field propagates without attenuation. If $\varepsilon_{\text{eff}} < 0$,the propagation constant β becomes imaginary and the wave attenuates as it propagates. Note that this attenuation is of a different physical origin than the attenuation due to finite conductivity, which is captured by the parameter α : in plasma, there can be attenuation even is $\sigma = 0$ provided that the $\omega < \omega_p$.

12 Oblique incidence for lossless dielectrics

12.1 Snell's law for lossless dielectrics

We have thus far considered the problem of reflection and transmission of a wave propagating along $\hat{z}$ and hitting an interface "head-on".

Geometrically, the direction of propagation is always parallel or antiparallel to a vector normal to the interface. This is known as normal incidence. Since the electric and magnetic fields of a plane wave are always perpendicular to the direction of propagation, it therefore follows that $\vec{E}_s$ and $\vec{H}_s$ were always completely perpendicular to the interface.

We now move on to oblique incidence, where the wave hits the interface at some slanted angle. In this situation, the $\vec{E}_s$ and $\vec{H}_s$ are not necessarily completely perpendicular to the normal at the interface, and we must distinguish two cases.

The section is organized as follows: first, we establish results and nomenclature which is common to every case of oblique incidence. Then we consider in turn the problems of reflection and transmission of a plane wave with parallel polarization and perpendicular polarization.

All media are considered lossless, meaning $\sigma_1 = \sigma_2 = \sigma_3 = 0$. It is easily verified that this implies that the attenuation constant in medium a :

$$\alpha_a = \omega_a \sqrt{\frac{\mu_a \varepsilon_a}{2}} \left[\sqrt{1 + \left(\frac{\sigma_a}{\omega_a \varepsilon_a} \right)^2} - 1 \right]^{1/2} = 0, \tag{12.1}$$

$$\beta_a = \omega_a \sqrt{\frac{\mu_a \varepsilon_a}{2}} \left[\sqrt{1 + \left(\frac{\sigma_a}{\omega_a \varepsilon_a} \right)^2} + 1 \right]^{1/2} = \omega_a \sqrt{\mu_a \varepsilon_a}, \tag{12.2}$$

$$\eta_a = \sqrt{\frac{\mu_a}{\varepsilon_a}}. \tag{12.3}$$

Note that these are written under the most general assumption that the angular frequencies ω_a might be different from one medium to the next (this assumption will be shown to be false: the angular frequency in all media will be the same).

The most general expression for the electric field phasor in medium a is thus

$$\vec{E}_{as} = \vec{E}_{ao} e^{-j\beta_a \hat{n}_a \cdot \vec{r}_a}. \tag{12.4}$$

In lossless media, it is conventional to use the notation:

$$\beta_a = k_a = \omega_a \sqrt{\mu_a \varepsilon_a}, \tag{12.5}$$

$$\beta_a \hat{n}_a = \vec{k}_a = (k_{ax}, k_{ay}, k_{az}). \tag{12.6}$$

Here, $\hat{n}_a$ is a unit vector in the direction of propagation, and is often denoted by $\hat{k}_a$ and denoted by $\vec{a}_k$ by Sadiku. Hence, $\vec{k}_a$ is a vector of magnitude β_a and points in the direction of propagation. In particular, we have

$$\left| \vec{k}_a \right|^2 = k_{ax}^2 + k_{ay}^2 + k_{az}^2 = \omega_a^2 \mu_a \varepsilon_a, \tag{12.7}$$

$$\vec{k}_a = k_a \hat{k}_a. \tag{12.8}$$

and the "nice" relations, valid for electromagnetic plane waves in any lossless medium,

$$\vec{k}_a \times \vec{E}_a = \omega_a \mu_a \vec{H}_a, \tag{12.9}$$
$$\vec{H}_a \times \vec{k}_a = \omega_a \varepsilon_a \vec{E}_a, \tag{12.10}$$
$$\vec{k}_a \cdot \vec{E}_a = \vec{k}_a \cdot \vec{H}_a = \vec{E}_a \cdot \vec{H}_a = 0. \tag{12.11}$$

There are applicable to the physical fields as well as the phasors.

From Eqns. (12.5), (12.8) and (12.9), one also derives the handy relation

$$\omega_a \mu_a \vec{H}_a = \vec{k}_a \times \vec{E}_a = \omega_a \sqrt{\mu_a \varepsilon_a} \left(\hat{k}_a \times \vec{E}_a \right), \tag{12.12}$$
$$\vec{H}_a = \frac{\omega_a \sqrt{\mu_a \varepsilon_a}}{\omega_a \mu_a} \left(\hat{k}_a \times \vec{E}_a \right) = \frac{1}{\sqrt{\frac{\mu_a}{\varepsilon_a}}} \left(\hat{k}_a \times \vec{E}_a \right), \tag{12.13}$$
$$= \frac{1}{\eta_a} \hat{k}_a \times \vec{E}_a, \tag{12.14}$$

showing (once more) that the amplitude of the magnetic field is smaller than the amplitude of the electric field by a factor of $1/\eta_a$.

If we want to argue that the angular frequencies are the same in all media, we must revert to the physical fields, as the ω-dependence is explicitly evacuated from the phasor forms.

If the incident and reflected fields propagate in medium 1, and the transmitted field in medium 2, the full time-dependent forms of the fields are:

$$\vec{E}_i(\vec{r}, t) = \vec{E}_{io}(\vec{r}) \cos(\omega_i t - \beta_i \hat{n}_i \cdot \vec{r}) = \vec{E}_{io}(\vec{r}) \cos\left(\omega_i t - \vec{k}_i \cdot \vec{r}\right) \tag{12.15}$$
$$= \vec{E}_{io}(\vec{r}) \cos(\omega_i t - k_{ix} x - k_{iy} y - k_{iz} z), \tag{12.16}$$
$$\vec{E}_r(r, t) = \vec{E}_{ro}(r) \cos(\omega_r t - k_{rx} x - k_{ry} y - k_{rz} z), \tag{12.17}$$
$$\vec{E}_t(r, t) = \vec{E}_{to}(r) \cos(\omega_t t - k_{tx} x - k_{ty} y - k_{tz} z). \tag{12.18}$$

Let us now apply the electric boundary conditions at the interface between media 1 and 2. We will assume for simplicity that the interface is located in the z=0 plane.

Note that, although the boundary conditions were derived for *static* fields, they remain valid for time-dependent fields. For instance, an expression like

$$\oint \vec{E} \cdot d\vec{l} = 0 \tag{12.19}$$

is not valid anymore with dynamic fields because of induction. However, the boundary conditions were obtained for an *infinitesimal* contour, so the induced EMF, which will depend on the time-variation of the flux over an infinitesimal area, is effectively nil. The other boundary condition depends on

$$\oint \vec{D} \cdot d\vec{S} = 0 \tag{12.20}$$

if there is no charge at the interface (as is the case here), and this equation is not modified by the presence of time-dependent fields.

Thus, at $z = 0$, we have, for all values of time and for all points $\vec{r}$ on the interface.

$$\vec{E}_{io}(\vec{r}) \cos(\omega_i t - k_{ix}x - k_{iy}y) + \vec{E}_{ro}(r) \cos(\omega_r t - k_{rx}x - k_{ry}y) \tag{12.21}$$

$$= \vec{E}_{to}(r) \cos(\omega_t t - k_{tx}x - k_{ty}y). \tag{12.22}$$

As this must hold for all t and all x, y, it follows that we must have

$$\omega_i t - k_{ix}x - k_{iy}y = \omega_r t - k_{rx}x - k_{ry}y = \omega_t t - k_{tx}x - k_{ty}y. \tag{12.23}$$

In particular, this must hold at $x = y = 0$, so we obtain

$$\omega_i = \omega_r = \omega_t = \omega. \tag{12.24}$$

In other words, the angular frequency of the signal is unchanged by the interface. This makes intuitive sense: the interface represents a discontinuity in the *spatial* properties of the media, so we expect the *temporal* part of the wave to be unaffected. Selecting now $t = 0$ and $x = 0$, we find

$$k_{iy} = k_{ry} = k_{ty}. \tag{12.25}$$

A similar argument at $t = 0$ and $y = 0$ produces

$$k_{ix} = k_{rx} = k_{tx}. \tag{12.26}$$

It is *essential* to observe that, since the interface is at $z = 0$, we cannot obtain a condition on the z-components on $\vec{k}_i$, $\vec{k}_r$ and $\vec{k}_t$.

Using this, and noting that both incident and reflected wave propagate in medium 1, we thus have

$$\beta_i = k_i = \omega\sqrt{\mu_1\varepsilon_1} = k_r. \tag{12.27}$$

In other words, the propagation vectors of the incident and reflected waves have the same magnitude, although the vectors k_i and k_r are expected to have different directions as the wave themselves propagate in different direction.

Since $\beta_t = k_t = \omega\sqrt{\mu_2\varepsilon_2} \neq \beta_i$, and $k_{ix} = k_{tx}$, $k_{iy} = k_{ty}$, it follows that

$$k_{iz} \neq k_{tz}. \tag{12.28}$$

To examine the consequences of these equations, suppose for simplicity that we orient our axes so that

$$\hat{n}_i = \hat{x}\sin\theta_i + \hat{z}\cos\theta_i \tag{12.29}$$

$$\vec{k}_i = k_i\hat{n}_i \tag{12.30}$$

It is important to realize that the planar symmetry of the interface *always* allows us to do this simplification without loss of generality. With this choice and Eqn.(12.25), it is clear that the incident, reflected and transmitted $\vec{k}$-vectors will always lie in the xz-plane.

This observation, taken with the setup of our interface as a plane located at $z = 0$, makes it natural to decompose the vectors $\vec{k}_i, \vec{k}_r$ and $\vec{k}_t$ in terms of components along a vector normal to the interface, i.e. a vector along $\hat{z}$, and components perpendicular to $\hat{z}$. Thus, we write

$$\begin{aligned} \vec{k}_i &= k_i\left(\hat{x}\sin\theta_i + \hat{z}\cos\theta_i\right), && (12.31)\\ \vec{k}_r &= k_i\left(\hat{x}\sin\theta_r - \hat{z}\cos\theta_r\right), && (12.32)\\ \vec{k}_t &= k_t\left(\hat{x}\sin\theta_t + \hat{z}\cos\theta_t\right). && (12.33) \end{aligned}$$

The angle θ_i is known as the angle of incidence, θ_r is the angle of reflection, and the angle θ_t is known as the angle of refraction. These angles are computed from the normal to the interface. The case of normal incidence, discussed in the previous sections, is the case where $\theta_i = 0$.

Notice how, in the expression of Eqn.(12.32), the sign of the z-component has been reversed, in accordance with the geometrical interpretation of a reflection. In this equation, $k_r = k_i$ has also been used. Comparing the $\hat{x}$-components of Eqns.(12.31) and (12.32) shows that

$$\sin\theta_i = \sin\theta_r, \tag{12.34}$$

or, in words, the angle of incidence equals the angle of reflection. The plane containing $\vec{k}_i, \vec{k}_r$ and $\vec{k}_t$ is called the plane of incidence. Here, the plane of incidence is the xz plane.

Comparing the x-components of Eqns.(12.31) and (12.33) yields

$$\begin{aligned} k_i\sin\theta_i &= k_t\sin\theta_t, && (12.35)\\ \omega\sqrt{\mu_1\varepsilon_1}\sin\theta_i &= \omega\sqrt{\mu_2\varepsilon_2}\sin\theta_t. && (12.36) \end{aligned}$$

The combination

$$\sqrt{\mu_1\varepsilon_1} = \sqrt{\mu_{r1}\varepsilon_{r1}}\sqrt{\mu_0\varepsilon_0} = \frac{1}{c}\sqrt{\mu_{r1}\varepsilon_{r1}} = \frac{n_1}{c}, \tag{12.37}$$

where

$$n_1 = \sqrt{\mu_{r1}\varepsilon_{r1}} = \sqrt{\frac{\mu_1\varepsilon_1}{\mu_0\varepsilon_0}} \tag{12.38}$$

is the index of refraction of the medium. Note that, for non-magnetic materials, $n_1 > 1$ since ε_{r1} is always > 1. Furthermore, ε is experimentally found to be dependent on ω, so n is usually itself dependent on ω. Thus, Eqn.(12.36) takes a form known as Snell's law for lossless dielectrics:

$$n_1\sin\theta_i = n_2\sin\theta_t. \tag{12.39}$$

Let us briefly explore this last relation through a simple example. By the geometry of the problem, the angle of incidence θ_i is limited in range to lie between 0 and π. Now, if a plane wave from air ($n_1 = 1$) hits the surface of a fresh water lake ($n_2 = 4/3$), at an angle of incidence $\theta_i = \theta_{air}$, then the wave will be transmitted with an angle $\theta_t = \theta_{water}$ such that

$$\sin\theta_{water} = \frac{n_{air}}{n_{water}}\sin\theta_{air} = \frac{3}{4}\sin\theta_{air}. \tag{12.40}$$

It is clear there is always an angle θ_{water} that will satisfy this, and that the angle θ_{water} is shallower than the angle of incidence θ_{air} since the sine function is increasing in the interval form 0 to π.

Suppose we go instead from water to air. Then, we have $\theta_i = \theta_{water}$ and $\theta_t = \theta_{air}$. Snell's law then gives

$$\sin\theta_{air} = \frac{4}{3}\sin\theta_{water}. \tag{12.41}$$

It is now clear that, if the angle of incidence θ_{water} is such that $\sin\theta_{water} > \frac{3}{4}$, there will be no angle θ_{air} that will be solution. In this case, there is no transmission of the wave and we have a phenomenon called *total internal reflection.* The general analysis shows without difficulty that total internal reflection can happen only if $n_2 < n_1$. The maximum angle of incidence is called the critical angle θ_c. When the wave is incident at the critical angle θ_c, the transmitted wave propagates at an angle $\theta_t = \pi/2$, i.e. parallel to the interface. If the angle of incidence is above the critical angle, we have total internal reflection. Thus, one easily finds

$$\sin\theta_{crit.} = \frac{n_2}{n_1}\sin\left(\frac{\pi}{2}\right) = \frac{n_2}{n_1}. \tag{12.42}$$

The phenomenon of total internal reflection is quite practical, as it allows transmitters to send signal "bouncing" off surfaces. Total internal reflection is used to propagate radiowaves over large distances by bouncing on the ionosphere, and also used to propagate light in optical fibers.

12.2 Reflections and transmission at oblique incidence.

Although the vectors $\vec{k}_i, \vec{k}_r$ and $\vec{k}_t$ are always in the same plane (the plane of incidence), the electric and magnetic fields need not be: they are limited to be perpendicular to propagation vectors $\vec{k}_i, \vec{k}_r$ and $\vec{k}_t$. Using Eqn.(12.29), we have

$$\begin{aligned}
\beta_1 \hat{n}_i \cdot \vec{r} &= \vec{k}_i \cdot \vec{r}, && (12.43)\\
&= \beta_1 \left(\hat{x}\sin\theta_i + \hat{z}\cos\theta_i\right)\cdot\left(\hat{x}x + \hat{y}y + \hat{z}z\right) && (12.44)\\
&= \beta_1 \left(x\sin\theta_i + z\cos\theta_i\right), && (12.45)
\end{aligned}$$

it is convenient to write the incident electric field phasor in the form

$$\vec{E}_{is}(\vec{r}) = \left(\vec{E}_{i\parallel} + \vec{E}_{i\perp}\right) e^{-j\beta_1(x\sin\theta_i + z\cos\theta_i)}, \tag{12.46}$$

where

$$\begin{aligned}
\vec{E}_{i\parallel} &= \hat{x}E_{ix} + \hat{z}E_{iz}, && (12.47)\\
\vec{E}_{i\perp} &= \hat{y}E_{iy}. && (12.48)
\end{aligned}$$

The condition of orthogonality for plane waves now reads

$$\begin{aligned}
0 &= \hat{n}_i \cdot \vec{E}_{is}(\vec{r}), && (12.49)\\
&= \left(\hat{n}_i \cdot \vec{E}_{i\parallel} + \hat{n}_i \cdot \vec{E}_{i\perp}\right) e^{-j\beta_1(x\sin\theta_i + z\cos\theta_i)}, && (12.50)\\
&= \left(\hat{n}_i \cdot \vec{E}_{i\parallel}\right) e^{-j\beta_1(x\sin\theta_i + z\cos\theta_i)}, && (12.51)
\end{aligned}$$

by the geometrical setup of $\vec{E}_{i\|}$ and $\vec{E}_{i\perp}$.

The vector $\vec{E}_{i\perp} = \hat{y}E_{iy}$ is known as the perpendicular component of the electric field, as this component lies entirely in a direction perpendicular to the plane of incidence.

Equation (12.51) forces

$$\vec{E}_{i\|} = E_{i\|}\left(\hat{x}\cos\theta_i - \hat{z}\sin\theta_i\right). \tag{12.52}$$

This is known as the parallel component of the electric field, in contrast with the perpendicular component. The parallel component lies entirely in the plane of incidence.

An identical decomposition holds for the reflected and transmitted waves. In the corresponding cases, we have

$$\begin{aligned} \vec{E}_{r\|} &= E_{r\|}\left(\hat{x}\cos\theta_i + \hat{z}\sin\theta_i\right), & (12.53)\\ \vec{E}_{t\|} &= E_{t\|}\left(\hat{x}\cos\theta_t - \hat{z}\sin\theta_t\right). & (12.54) \end{aligned}$$

It is clear that the most general signal is a superposition of a parallel and perpendicular signal. However, both components must be handled separately for transmission and reflection properties.

12.2.1 Transmission and reflections of TE waves.

Consider first the case of an incident electric field with only perpendicular component:

$$\begin{aligned} \vec{E}_{is}(\vec{r}) &= \vec{E}_{i\perp}e^{-j\beta_1(x\sin\theta_i + z\cos\theta_i)}, & (12.55)\\ &= \hat{y}E_{i\perp}e^{-j\beta_1(x\sin\theta_i + z\cos\theta_i)}. & (12.56) \end{aligned}$$

This is known as a TE wave, for Transverse Electric, indicating the electric field is transverse to the plane of incidence. The reflected and transmitted phasors have the form

$$\begin{aligned} \vec{E}_{rs}(\vec{r}) &= \hat{y}E_{r\perp}e^{-j\beta_1(x\sin\theta_i - z\cos\theta_i)}, & (12.57)\\ \vec{E}_{ts}(\vec{r}) &= \hat{y}E_{t\perp}e^{-j\beta_2(x\sin\theta_t + z\cos\theta_t)}. & (12.58) \end{aligned}$$

The corresponding magnetic fields are

$$\begin{aligned} \vec{H}_{is}(\vec{r}) &= \frac{1}{\eta_1}\left(\hat{x}\sin\theta_i + \hat{z}\cos\theta_i\right)\times\vec{E}_{is} & (12.59)\\ &= \frac{1}{\eta_1}\begin{vmatrix} \hat{x} & \hat{y} & \hat{z}\\ \sin\theta_i & 0 & \cos\theta_i\\ 0 & E_{i\perp}e^{-j\beta_1(x\sin\theta_i+z\cos\theta_i)} & 0 \end{vmatrix}, & (12.60)\\ &= \frac{1}{\eta_1}E_{i\perp}e^{-j\beta_1(x\sin\theta_i+z\cos\theta_i)}\left(-\hat{x}\cos\theta_i + \hat{z}\sin\theta_i\right), & (12.61)\\ \vec{H}_{rs}(\vec{r}) &= \frac{1}{\eta_1}E_{r\perp}e^{-j\beta_1(x\sin\theta_t-z\cos\theta_t)}\left(\hat{x}\cos\theta_i + \hat{z}\sin\theta_i\right), & (12.62)\\ \vec{H}_{ts}(\vec{r}) &= \frac{1}{\eta_2}E_{t\perp}e^{-j\beta_2(x\sin\theta_t+z\cos\theta_t)}\left(-\hat{x}\cos\theta_t + \hat{z}\sin\theta_t\right). & (12.63) \end{aligned}$$

(Note the change in sign of the x-component of $\vec{H}_{rs}(\vec{r})$, something which results from the reversed sign of z in the direction of propagation of the reflected wave.) At the interface

$z = 0$, we have, from the boundary conditions on tangential electric and magnetic fields, the equations

$$\begin{aligned} E_{i\perp} + E_{r\perp} &= E_{t\perp}, &(12.64)\\ \frac{\cos\theta_i}{\eta_1}\left(-E_{i\perp} + E_{r\perp}\right) &= -\frac{\cos\theta_t}{\eta_2}E_{t\perp}. &(12.65)\end{aligned}$$

We solve for $E_{r\perp}$ and $E_{t\perp}$ in terms of the incident field $E_{i\perp}$:

$$\begin{aligned} \Gamma_\perp &\equiv \frac{E_{r\perp}}{E_{i\perp}} = \frac{\eta_2\cos\theta_i - \eta_1\cos\theta_t}{\eta_2\cos\theta_i + \eta_1\cos\theta_t}, &(12.66)\\ \tau_\perp &\equiv \frac{E_{t\perp}}{E_{i\perp}} = \frac{2\eta_2\cos\theta_i}{\eta_2\cos\theta_i + \eta_1\cos\theta_t}. &(12.67)\end{aligned}$$

Note that these expressions reduce to those for normal incidence when we set $\theta_i = 0$.

Because the index of refraction is often a tabulated quantity, it is often convenient to manipulate $\Gamma_\perp$ and $\tau_\perp$ so as to make the index of refraction of the media appear, especially in the case of non–magnetic media where $\mu = \mu_0$. Thus

$$\begin{aligned} \Gamma_\perp &= \frac{\eta_2\cos\theta_i - \eta_1\cos\theta_t}{\eta_2\cos\theta_i + \eta_1\cos\theta_t}\,, &(12.68)\\ &= \frac{\sqrt{\frac{\mu_0}{\varepsilon_2}}\cos\theta_i - \sqrt{\frac{\mu_0}{\varepsilon_1}}\cos\theta_t}{\sqrt{\frac{\mu_0}{\varepsilon_2}}\cos\theta_i + \sqrt{\frac{\mu_0}{\varepsilon_1}}\cos\theta_t}\,, &(12.69)\\ &= \frac{\sqrt{\frac{1}{\mu_0\varepsilon_2}}\cos\theta_i - \sqrt{\frac{1}{\mu_1\varepsilon_1}}\cos\theta_t}{\sqrt{\frac{1}{\mu_0\varepsilon_2}}\cos\theta_i + \sqrt{\frac{1}{\mu\varepsilon_1}}\cos\theta_t}\,, &(12.70)\\ &= \frac{\frac{c}{n_2}\cos\theta_i - \frac{c}{n_1}\cos\theta_t}{\frac{1}{n_2}\cos\theta_i + \frac{c}{n_1}\cos\theta_t}\,, &(12.71)\\ &= \frac{n_1\cos\theta_i - n_2\cos\theta_t}{n_1\cos\theta_i + n_2\cos\theta_t}\,. &(12.72)\end{aligned}$$

From this, it can be shown that $\Gamma_\perp$ is never 0, unless both media are identical. Indeed, $\Gamma_\perp = 0$ requires

$$\begin{aligned} n_1\cos\theta_i &= n_2\cos\theta_t\,, &(12.73)\\ n_1^2(1-\sin^2\theta_i) &= n_2^2(1-\sin^2\theta_t)\,, &(12.74)\\ n_1^2 - n_1^2\sin^2\theta_i &= n_2^2\left(1-\frac{n_1^2}{n_2^2}\sin^2\theta_i\right), &(12.75)\\ n_1^2 - n_1^2\sin^2\theta_i &= n_2^2 - n_1^2\sin^2\theta_i\,, &(12.76)\end{aligned}$$

which implies $n_1 = n_2$, *i.e.* both media are the same.

Of course, $\tau_\perp$ can also be expressed in terms of n_1, n_2:

$$\begin{aligned} \tau_\perp &= \frac{\frac{2c}{n_2}\cos\theta_i}{\frac{c}{n_2}\cos\theta_i + \frac{c}{n_1}\cos\theta_t}\,, &(12.77)\\ &= \frac{2n_1\cos\theta_i}{n_1\cos\theta_i + n_2\cos\theta_t}\,. &(12.78)\end{aligned}$$

12.2.2 Transmission and reflection of TM waves

Consider next the case of an incident, reflected and transmitted electric fields of the form

$$\vec{E}_{is}(\vec{r}) = E_{i\parallel}\left(\hat{x}\cos\theta_i - \hat{z}\sin\theta_i\right)e^{-j\beta_1(x\sin\theta_i + z\cos\theta_i)}, \tag{12.79}$$
$$\vec{E}_{rs}(\vec{r}) = E_{r\parallel}\left(\hat{x}\cos\theta_i + \hat{z}\sin\theta_i\right)e^{-j\beta_1(x\sin\theta_i - z\cos\theta_i)}, \tag{12.80}$$
$$\vec{E}_{ts}(\vec{r}) = E_{t\parallel}\left(\hat{x}\cos\theta_t - \hat{z}\sin\theta_t\right)e^{-j\beta_2(x\sin\theta_t - z\cos\theta_t)} \tag{12.81}$$

The corresponding magnetic fields are

$$\vec{H}_{is}(\vec{r}) = \frac{1}{\eta_1}\left(\hat{x}\sin\theta_i + \hat{z}\cos\theta_i\right)\times\vec{E}_{is} \tag{12.82}$$
$$= \frac{1}{\eta_1}\begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \sin\theta_i & 0 & \cos\theta_i \\ E_{i\parallel}\cos\theta_i & 0 & -E_{i\parallel}\sin\theta_i \end{vmatrix} e^{-j\beta_1(x\sin\theta_i + z\cos\theta_i)} \tag{12.83}$$
$$= \frac{1}{\eta_1}\left(\hat{y}E_{i\parallel}\sin^2\theta_i + \hat{y}E_{i\parallel}\cos^2\theta_i\right)e^{-j\beta_1(x\sin\theta_i + z\cos\theta_i)}, \tag{12.84}$$
$$= \hat{y}\frac{1}{\eta_1}E_{i\parallel}e^{-j\beta_1(x\sin\theta_i + z\cos\theta_i)}, \tag{12.85}$$
$$\vec{H}_{rs}(\vec{r}) = -\hat{y}\frac{1}{\eta_1}E_{r\parallel}e^{-j\beta_1(x\sin\theta_i - z\cos\theta_i)}, \tag{12.86}$$
$$\vec{H}_{ts}(\vec{r}) = \hat{y}\frac{1}{\eta_2}E_{t\parallel}e^{-j\beta_2(x\sin\theta_t + z\cos\theta_t)}. \tag{12.87}$$

It is now the magnetic field that is transverse to the plane of incidence, so this situation is referred to as a TM, or Transverse Magnetic, wave.

At the interface $z = 0$, we have, from the boundary conditions on tangential electric and tangential magnetic fields, the equations

$$\left(E_{i\parallel} + E_{r\parallel}\right)\cos\theta_i = E_{t\parallel}\cos\theta_t, \tag{12.88}$$
$$\frac{1}{\eta_1}\left(E_{i\parallel} - E_{r\parallel}\right) = \frac{1}{\eta_2}E_{t\parallel}, \tag{12.89}$$

with solution

$$\Gamma_\parallel = \frac{E_{r\parallel}}{E_{i\parallel}} = \frac{\eta_2\cos\theta_t - \eta_1\cos\theta_i}{\eta_2\cos\theta_t + \eta_1\cos\theta_i}, \tag{12.90}$$
$$\tau_\parallel = \frac{E_{t\parallel}}{E_{i\parallel}} = \frac{2\eta_2\cos\theta_i}{\eta_2\cos\theta_t + \eta_1\cos\theta_i}. \tag{12.91}$$

Again, these expressions reduce to those found for normal incidence when we set $\theta_i = 0$.

In terms of n_1, n_2, one rapidly finds

$$\tau_\parallel = \frac{2n_1\cos\theta_i}{n_1\cos\theta_t + n_2\cos\theta_i}\,. \tag{12.92}$$

The expression for $\Gamma_\parallel$ is much more interesting. Indeed we get

$$\Gamma_\parallel = \frac{n_1\cos\theta_t - n_2\cos\theta_i}{n_1\cos\theta_t + n_2\cos\theta_i}\,. \tag{12.93}$$

If we now look to see if an angle θ_B will make $\Gamma_\parallel = 0$, we have

$$\begin{aligned}
n_1 \cos\theta_t &= n_2 \cos\theta_B\,, && (12.94)\\
n_1^2(1-\sin^2\theta_t) &= n_2^2(1-\sin^2\theta_B)\,, && (12.95)\\
n_1^2(1-\frac{n_1^2}{n_2^2}\sin^2\theta_B) &= n_2^2 - n_2^2\sin^2\theta_B\,, && (12.96)\\
n_2^2 n_1^2 - n_1^4\sin^2\theta_B &= n_2^4 - n_2^4\sin^2\theta_B\,, && (12.97)\\
(n_2^4 - n_1^4)\sin^2\theta_B &= n_2^4 - n_2^2 n_1^2 = n_2^2(n_2^2 - n_1^2)\,, && (12.98)\\
(n_2^2 + n_1^2)(n_2^2 - n_1^2)\sin^2\theta_B &= n_2^2(n_2^2 - n_1^2)\,, && (12.99)\\
\sin^2\theta_B &= \frac{n_2^2}{n_2^2 + n_1^2} = \frac{1}{1+\frac{\varepsilon_1}{\varepsilon_2}}\,, && (12.100)\\
\sin\theta_B &= \frac{n_2}{\sqrt{n_1^2+n_2^2}} = \sqrt{\frac{\varepsilon_2}{\varepsilon_1+\varepsilon_2}} && (12.101)
\end{aligned}$$

The solution to this is more cleanly expressed using a trigonometric identity. From Eqn.(12.101), we find

$$\begin{aligned}
\cos\theta_B &= \sqrt{1-\sin_B^\theta}\,, && (12.102)\\
&= \sqrt{1-\frac{n_2^2}{n_1^2+n_2^2}} = \sqrt{\frac{n_1^2}{n_1^2+n_2^2}}\,, && (12.103)\\
&= \frac{n_1}{\sqrt{n_1^2+n_2^2}}\,, && (12.104)
\end{aligned}$$

so that

$$\tan\theta_B = \frac{n_2}{n_1} = \sqrt{\frac{\varepsilon_2}{\varepsilon_1}}\,. \qquad (12.105)$$

The angle θ_B which makes $\Gamma_\parallel = 0$ and does not exists for the perpendicular waves is called the Brewster angle.

The Brewster angle is useful when high intensity beams of polarized lights are required. Many polarizers function by dissipating one of the two components of the electric field; in high intensity applications, the energy dissipated would physically burn the polarizer.

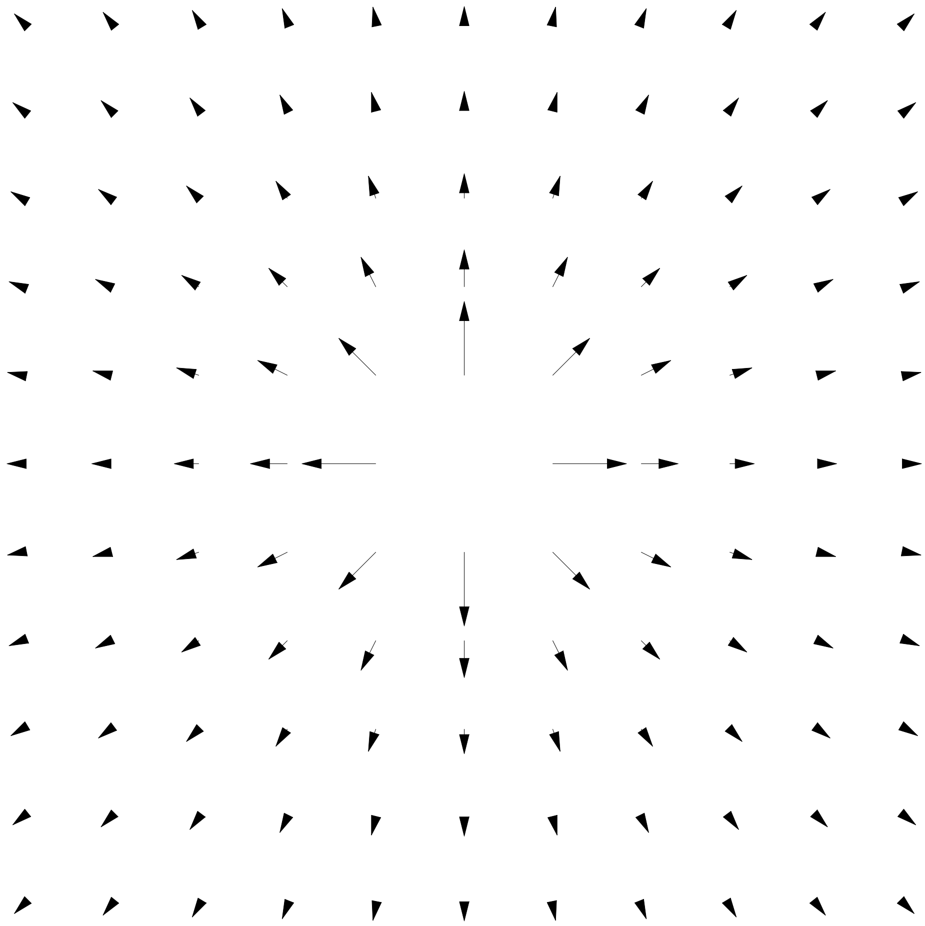

Figure 1: The electric field in the $z = 0$ plane of a positive point charge located at the origin.

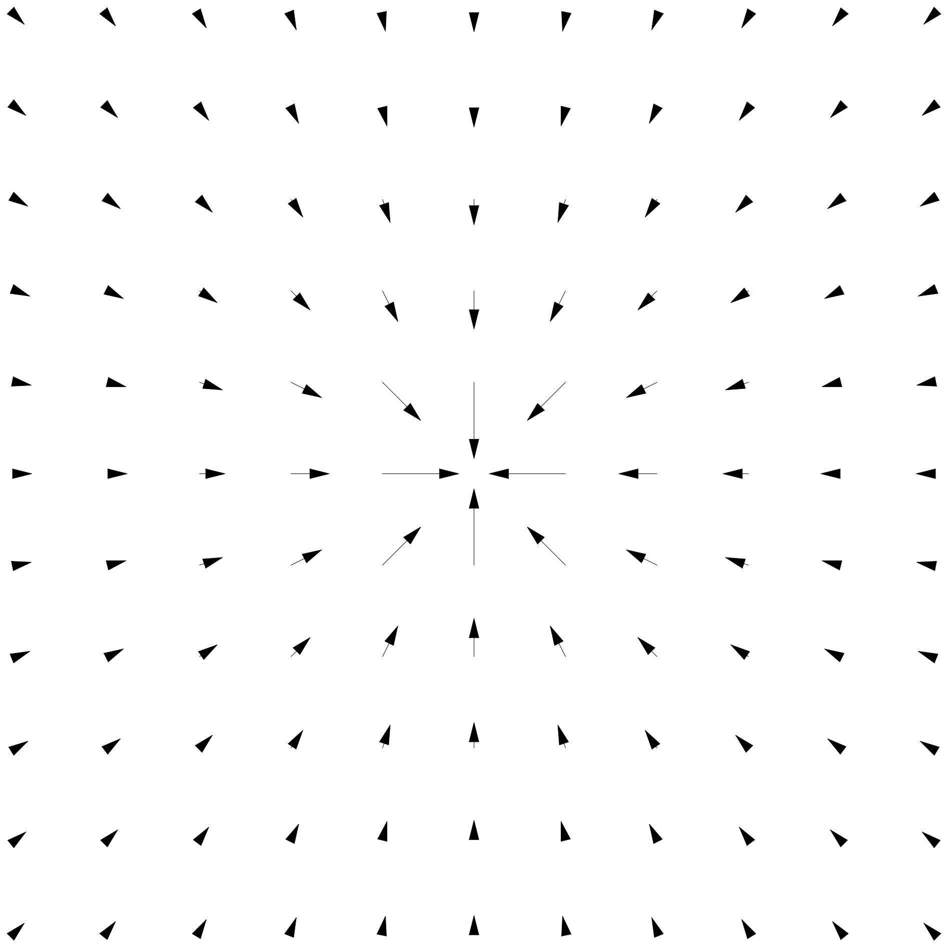

Figure 2: The electric field in the $z = 0$ plane of a negative point charge located at the origin.

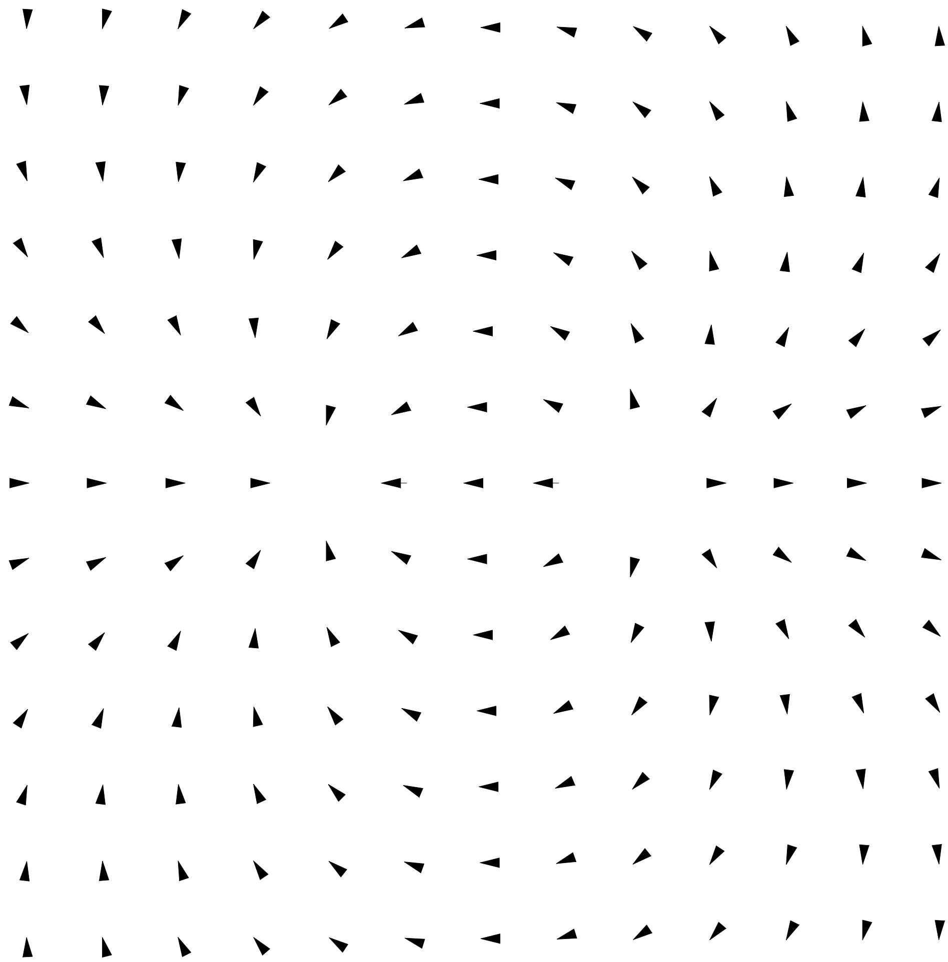

Figure 3: The electric field in the $z = 0$ plane of an electric dipole located at the origin.