Preprint typeset in JHEP style - PAPER VERSION

Lecture notes for General Physics 219

Martin Kruczenski

Department of Physics, Purdue University, 525 Northwestern Avenue,
W. Lafayette, IN 47907-2036. E-mail: markru@purdue.edu

Abstract: These are the notes for the lectures. They contain what is explained in class and can be used to refresh your memory or to stay up to date if you miss a class. They do *not* replace the book since they have much less information. Also take into account that the actual lectures might run a little behind schedule.

Keywords: Introductory physics.

Contents

---

1. Lecture 1

1.1 Introduction

One of the things that science does is to use the tools and ideas from our everyday experience and extrapolates them to other realms. If you look around you will see that one of the first things we notice and are able to determined about object is its size, perhaps inherited from our ancestor for whom a big or small animal meant the difference between predator or food. For that reason, when we explore a new area of science, from galaxies to atoms one of the first things we need to ask to get a grasp on the new subject is what is the typical size of the objects that we are going to deal with. We say that we determine the scale or order of magnitude of the systems we analyze.

As a way to fix ideas then let us revise the size of some systems in figure 1.

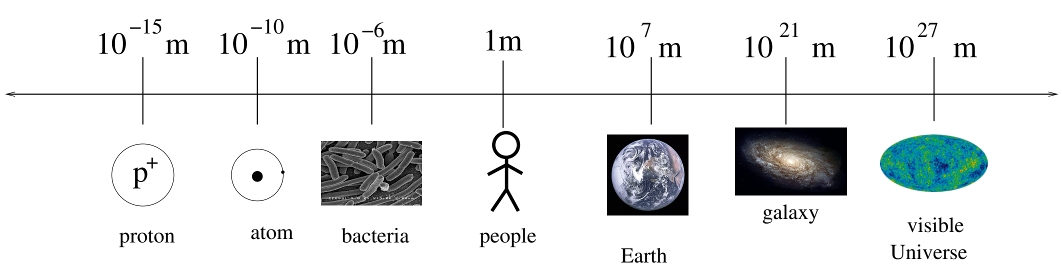


Figure 1: When studying a new phenomenon the first question is at which length scale it occurs. Typical examples are shown.

We see that in physics we have to deal with objects of very different size. For that reason it is convenient to use scientific notation where we write for example 10^3 for a thousand (also we use the prefix kilo) and 10^{-3} for a thousandth (prefix milli). Some commonly used prefixes

Factor	prefix	abbr.
10^{-15}	femto	f
10^{-12}	pico	p
10^{-9}	nano	n
10^{-6}	micro	μ
10^{-3}	milli	m
1		
10^{3}	kilo	K
10^{6}	mega	M
10^{9}	giga	G
10^{12}	tera	T

Not all of them are always used, for example Megameter is not used although we are familiar with MegaByte or GigaByte.

The other important thing to notice is what are the forces acting on objects at different scales. At the planetary scale and larger the dominant force is gravity. Not because it is particularly strong but because most macroscopic objects are neutral under the other forces. At our scale and down to the atomic scale electromagnetism (electricity and magnetism) is the most important one. When we think about electromagnetism we first think of light-bulbs, phones, computers etc. but we should remind ourselves that everything around us works through the electromagnetic force. Solids are solids because electric charges keep the atoms together, chemical reactions occurs as a consequence of transfer of charged electrons between atoms making and destroying molecular bonds. Only inside the atomic nucleus, namely scales of $10^{-15}m$ do we find new forces, the strong force that keep the nucleus together and the weak force that induces certain radioactive decays. So if we understand gravity and electromagnetism we pretty much understand everything that surrounds us, at least in principle!. A very notable exception is the Sun, only after the discovery of the strong force it became apparent that the source of energy for the Sun is nuclear reactions.

Finally, once we understand the forces we need to know how they modify the objects that interact with them. At distances much larger that the atomic nucleus this is given by Newton's famous third law:

$$\vec{F} = m\vec{a} \tag{1.1}$$

namely a force acting on an object produces an acceleration, a change in velocity, proportional to the force. If we explore distances of 10^{-10} meters an below (atomic scale) then Newton's law is replaced by quantum mechanics as we will find out later in the course. Furthermore, if objects move at speeds close to the speed of light then Einstein's theory of relativity should be used.

Before we start with electromagnetism we can make a quick review of gravity. Newton's law of gravity is one of the greatest achievements of mankind and made clear what Galileo and others had expressed, namely that, through the use of reasoning and mathematics we can gain insight into Nature at a depth that was previously unimaginable. It is not clear why this is so but it has been proved right until now, the more we explore Nature the more amazing phenomena we discover and the more interesting the mathematical constructions that are needed to describe them.

In any case, going back to the Law of Gravity, it simply states that two massive bodies attract each other with a force proportional to the mass and inversely proportional to the square of the distance separating them:

$$|\vec{F}| = G\frac{M_1 M_2}{r^2} \tag{1.2}$$

Remember that the force is a vector, it has a magnitude that we just gave and a direction which is toward the other body. The constant G is called Newton's constant. Its value is $G = 6.67 \times 10^{-11} m^2 kg^{-1} s^{-2} = 6.67 \times 10^{-11} N (m/kg)^2$. Please take your time to see that you understand the units. Units are fundamental in physics since a number without a unit has no meaning.

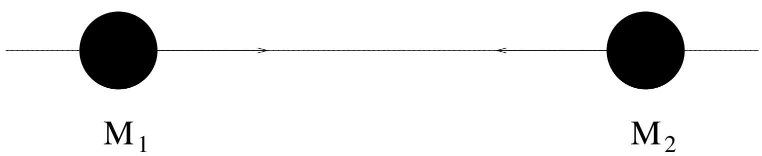


Figure 2: Two masses attract each other due to gravity.

Problem 1: Using that the radius of the Earth is $r_E \simeq 6000 Km$, the acceleration of gravity on the surface $g = 10 m/s^2$ and the density five times that of water, give an estimate of G and compare with the value given.

Solution: The gravitational force on an object of mass m at the surface of the Earth is given by

$$|\vec{F}| = G\frac{M_E m}{R_E^2} \tag{1.3}$$

where M_E is the mass of the Earth and R_E is its radius. According to Newton's law the acceleration is

$$g = \frac{|\vec{F}|}{m} = G\frac{M_E}{R_E^2} \tag{1.4}$$

The mass of the Earth is given by

$$M_E = \frac{4}{3}\pi R_E^3 \rho_E \tag{1.5}$$

with $\rho_E = 5\times 10^3 \frac{Kg}{m^3}$ its density. Replacing in our previous equation and after some algebra we find

$$G = \frac{gR_E^2}{\frac{4}{3}\pi R_E^3 \rho_E} = \frac{3}{4\pi}\frac{g}{R_E \rho_E} \tag{1.6}$$

replacing the values $R_E = 6,000 Km$, $g = 10m/s^2$, $\rho_E = 5\times 10^3 \frac{Kg}{m^3}$ we find

$$G \simeq 8\times 10^{-11}\frac{m^3}{Kgs^2} \tag{1.7}$$

a good estimate of the actual measured value $G = 6.7\times 10^{-11}\frac{N\,m^2}{Kg^2}$. Notice that the units are the same since $1N = 1Kg\,m/s^2$. A final observation is that historically this was done the other way around, namely by measuring G the density of the Earth was determined.

Problem 2: Using that the period of the Moon orbit is around a month, estimate the distance of the Earth to the Moon. How can you use that to know the size of the Moon? Hint: Remember that the centripetal acceleration if $a_r = v^2/r$ and $v = \omega r$ where ω is the angular velocity.

Solution: Now we equate the gravitational force with the mass of the Moon times the centripetal acceleration given by $a_r = \frac{v^2}{r}$. We have:

$$|\vec{F}| = G\frac{M_E M_m}{R_0^2} = M_m\frac{v^2}{R_0} \tag{1.8}$$

where R_0 is the radius of the orbit. The velocity v is given by $v = \omega R_0$ where ω is the angular velocity given by $\omega = \frac{2\pi}{T}$, with T the period of the orbit. Furthermore we can use that

$$g = \frac{GM_E}{R_E^2} \tag{1.9}$$

as we had before. This allows us to do the calculation without using G, we simple need g the acceleration of gravity on the surface of the Earth. Putting everything together we find

$$R_0^3 = g\frac{T^2}{4\pi^2}R_E^2 \tag{1.10}$$

Replacing the numbers we find an estimate

$$R_0 = 4.5\times 10^8 m = 450,000 Km \tag{1.11}$$

This is the right order of magnitude but slightly larger than the actual value. You can try to see why we got a larger value. In any case try to do the calculation yourself and be sure that everything is clear. Finally knowing the distance to the Moon is should be simple to look at the Moon and figure out how big it is.

2. Lecture 2

2.1 Electric charge

In the same way that gravity describes the interaction of masses, electrostatics describes the interaction of electric charges. Notice that we say electrostatics because this applies to static charges. If charges move they generate a magnetic field. This can be ignored if the charges move slowly compared with the speed of light except if we have a large number of charges moving together in the same direction as in an electric cable. We will look at that later, for the moment we concentrate in static charges (or moving slowly). Before continuing however one might wonder if the law of gravity might not need to be amended and perhaps masses also interact differently if they move fast. This is actually true and it is described by Einstein's theory of general relativity.

Going back then to electric charges, a difference with gravity is that in this case the force can be repulsive as well as attractive. In fact electric charge can be positive or negative, charges of opposite sign attract and those of the same sign repel. Other than that, the law for the force is similar as established by Coulomb's law:

$$|\vec{F}| = \frac{1}{4\pi\varepsilon_0}\frac{q_1 q_2}{r^2} \tag{2.1}$$

Namely the magnitude of the force is proportional to the product of the charges and inversely proportional to the square of the distance.

The constant that replaces Newton's constant is $\frac{1}{4\pi\varepsilon_0} = 9\times 10^9 N\frac{m^2}{C^2}$. This values assumes that we measure the electric charge in Coulombs (C). In atomic and nuclear physics sometimes other units of charge are used so that the constant is just one. To get an idea of how much a Coulomb is we can consider the minimal unit of charge which is the charge of the electron $e = -1.6\times 10^{-19}C$. The proton, has the same but opposite charge. It is still a mystery why the charge of the proton and electron are exactly opposite but that implies that atoms are exactly neutral since they have the same number of protons (which are in the nucleus) and electrons (which orbit the nucleus). Perhaps it should be pointed out that the proton is made out of quarks called u and d that have charges $+2/3e$ and $-1/3e$ respectively. In any case we see that a large number of electrons are needed to make a charge of a Coulomb. However, the Avogadro number $N_A = 6\times 10^{23}$ is much larger. Since the Avogadro number is the number of atoms in a mol of a substance the means that we have available that number of electrons. For example one gram of Hydrogen (which is one mol) has $N_A = 6\times 10^{23}$ atoms each of which has one proton and one electron. However large charges are not usually obtained because, as mentioned before atoms are neutral. This is because the electric interaction is large and it would cost a large amount of energy to split positive and negative charges.

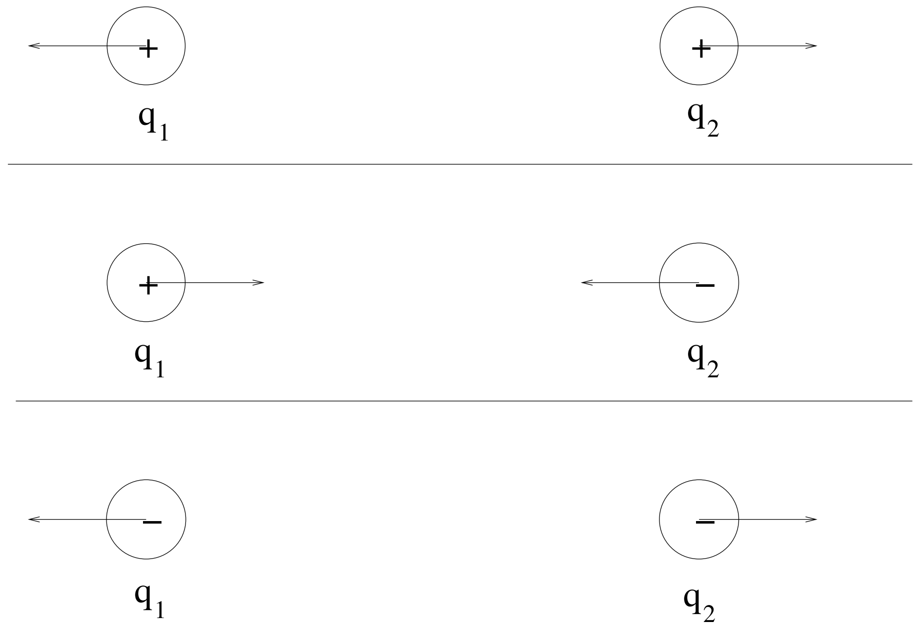


Figure 3: Two charges can attract or repel each other depending on their sign.

Nevertheless, a small amount of charge can be created, one example is by rubbing two materials. Sometimes one of them gets charged and allows us to check the laws of attraction and repulsion.

To understand more about charge we need to know that it is conserved. Experimentally it has been observed that the total charge of an isolated system is always the same. One can create charge but only of opposite signs in such a way that the total is always the same. So if you rub two objects and one is charged positive the other will be charged negative. Another important property is that of materials. Certain material such as metals are conductors, which means that electric charges can move freely inside them. Insulators on the other hand do not allow the motion of charge. These are the most common types, other materials such as semiconductors, superconductors etc. are more rare but very important in technological applications as we all know (electronics is based on semiconductors for example). If a metal is charged, since equal charges repel each other, the charges will try to be as far of each other as possible and will migrate to the surface. In fig.4 we see for example that if we approach a conducting sphere with a negatively charged rod, positive charge will be attracted close to the rod and negative charge well be repelled. However the total charge of the sphere should be zero if it was so initially. A different situation is if we connect the sphere to ground. This means that we run a conducting cable into the ground basically connecting the sphere to the Earth which for this purpose can be thought of as an unlimited reservoir of charge. In that case the negative charge will be repelled all the way to the ground. If we then disconnect the sphere and afterward remove the rod, the sphere will acquire a charge. The process is summarized in fig.5. This phenomenon is call induction and can be used to generate relatively large amounts of charge. An example is the machine demonstrated in class (unfortunately not very successfully) and which can be seen in fig. 6.

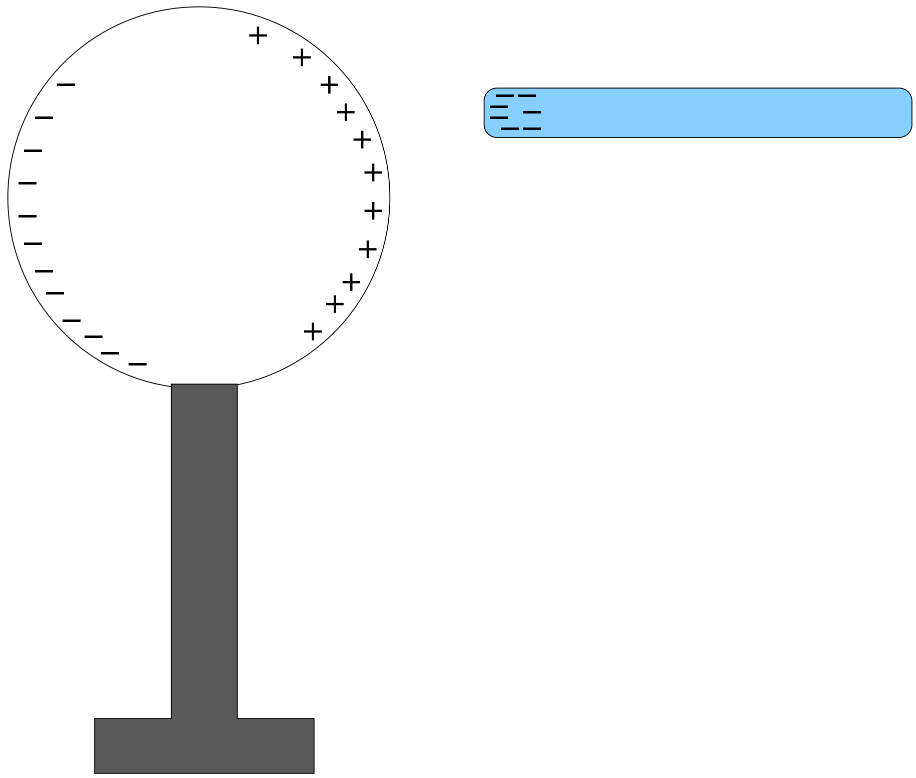


Figure 4: A negatively charged rod attracts positive charges and repels negative ones. The total charge of the sphere however remains constant since it is insulated

The idea of connecting something to ground is extremely important and when using an electrical device an important point is if it is appropriately connected to ground. One hand connecting the chassis (metallic case) to ground is a safety precaution against accidentally connecting it to a power line. It also avoids static electricity that can damage electronic circuits. Although connecting to ground literally means a connection to a conductor embedded in the soil, sometimes this is not practical (for example cell phones etc.) and then the "ground" refers to a common connection of electrical parts to the chassis. This gives a stable common reference to all circuits. This is particularly important in sensitive electronic devices since this common reference gives them stability, otherwise they can function erratically and can be a common reason for

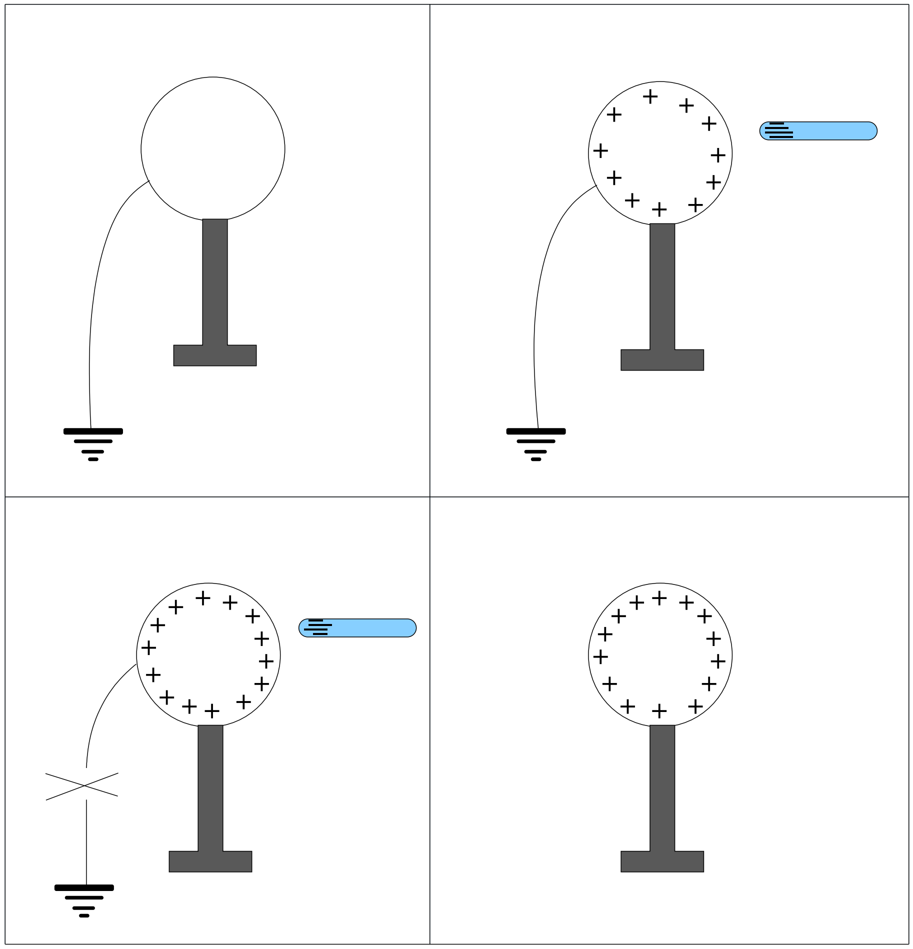


Figure 5: If the sphere is connected to ground we can induced a charge in it by using the procedure in the figure. Be sure you understand what happens in each step.

malfunctioning (for example, you create your own circuit to plug to a computer port but forget to connect the ground of your circuit to the computer ground either through the port or directly to the chassis).

Another interesting point is what happens if you have a charged conducting sphere and you touch it with another conducting one which is not charged. Since they are both conducting the charge will distribute among them. However, which one gets more charge? If the spheres have the same radius, by symmetry the total charge is distributed equally. However if one is bigger, the charges on it can be further apart which they prefer because the force between charges of the same sign is repulsive. What this means is that the sphere of larger radius gets more charge. In fact will see later that the charge it gets is proportional to its radius.

Figure 6: Machines used to generate static electricity. The first and last one are Wimshurst machines. In the middle is a small Van der Graaf generator. Van der Graaf made huge generators to accelerate atomic particles which can still be seen in the Boston museum of Science.

2.2 Electric field

Although initially one might think that electricity describes only forces between charges, simple experiments show that electromagnetic waves propagate from one place to another. Those include radio waves, light etc. A simple demonstration of this is when we create sparks in class and that was detected by the AM radio receiver which generated a noise with each spark. All this suggests that there is a form of energy that exist independently of the charges and leads us to the idea of electric field. Charges are sources for the electric field but electric fields can exist independently of charges. We denote the electric field as $\vec{E}$. It is a vector and it has the property that if you put a charge q in it, the charge experiences a force:

$$\vec{F} = q\vec{E} \tag{2.2}$$

The direction of the force is parallel to $\vec{E}$ although it can have opposite orientation if the charge is negative. It also follows that the force is proportional to the charge as can be verified experimentally. Furthermore we find that we reproduce Coulomb's law if a charge q_1 generates an electric field

$$\vec{E} = \frac{1}{4\pi\varepsilon_0}\frac{q_1}{r^2}\hat{r} \tag{2.3}$$

where $\hat{r}$ is a radial unit vector pointing away from the charge. An important principle that can be verified experimentally is the principle of superposition, if several sources produce fields $\vec{E}_1$, $\vec{E}_2$, ..., $\vec{E}_n$ then the total electric field is

$$\vec{E} = \vec{E}_1 + \vec{E}_2 + \ldots + \vec{E}_n \tag{2.4}$$

This implies the principle of superposition for the forces acting on an object namely

$$\vec{F} = q\vec{E} = q\vec{E}_1 + q\vec{E}_2 + \ldots + q\vec{E}_n = \vec{F}_1 + \vec{F}_2 + \ldots + \vec{F}_n \tag{2.5}$$

It is important that we add the forces and also the electric fields as vectors. Remember the rule of addition of vectors as illustrated in fig.7. Notice that vectors are mathematical entities with properties independent of what they represent. The easiest way to think about them is to think that they represent displacement from one place to another. Everybody is able to figure out where you are going to end up if you move lets say 30 meters in certain given direction, for example North and then 20 meters in another, for example South-West. This is addition of vectors and the same principle can be applied to the vectors represent electric fields, forces, velocities etc. But always add vectors that represent the same thing!

By the way, we emphasized that every quantity has a unit and we see that electric field should be measured in N/C, where N is Newton, a unit of force and C a unit of charge.

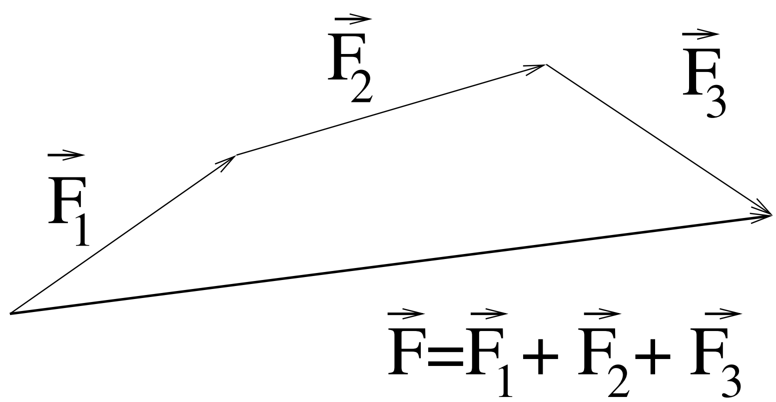


Figure 7: Vector are added in the same way you add displacements form one place to another. Be sure you have no confusion on how to add vectors

The concept of electric field is very useful but in principle it looks as a mess to draw. Indeed you have to draw an arrow at each point in space!. For that reason people found other way to represent the electric field and introduced the concept of "lines of electric field". What you do is to draw lines in such way that they are always tangent to the electric field. This gives you the direction. Then one draws more lines where the field is more intense. This is not a precise representation but gives a pictorial idea of how the electric field actually looks like. The simplest case is that of a single charge that we draw in figures 8 and 9. Notice that the Electric field points away from a positive charge and toward a negative one. This is because a small positive probe charge will be repelled by the positive charge and be attracted by the negative one.

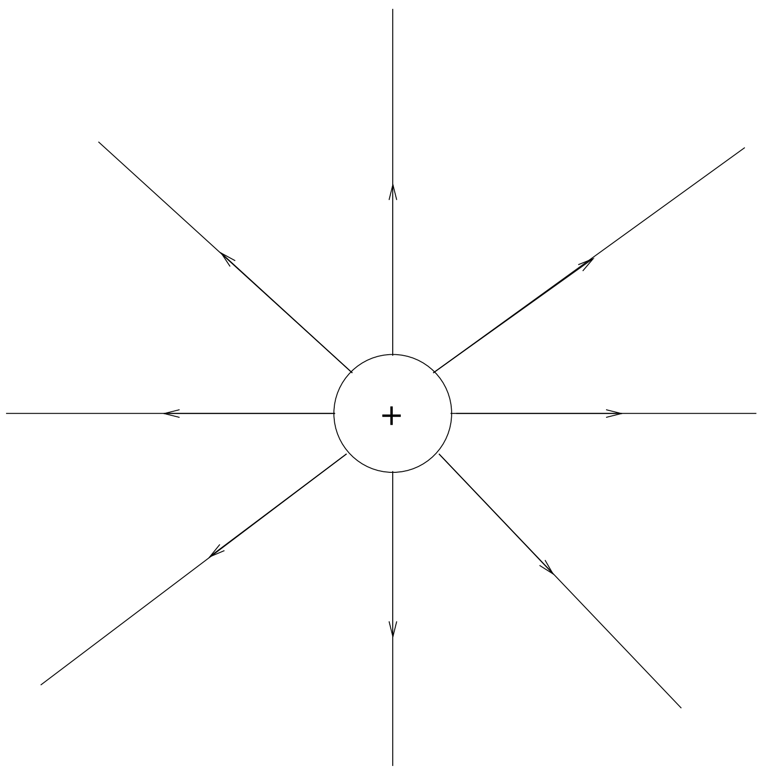


Figure 8: A positive charge generates an electric field pointing away from it.

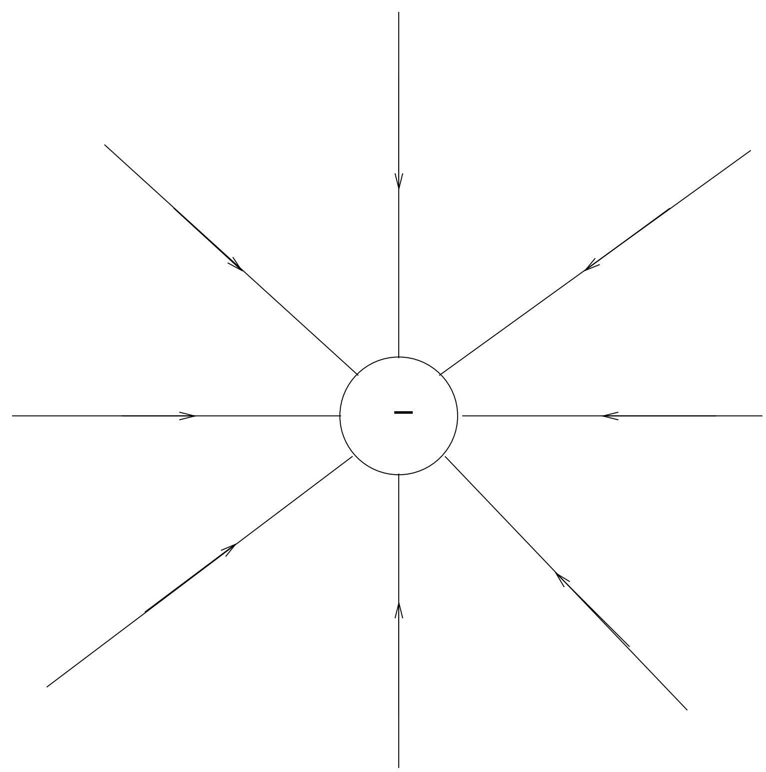


Figure 9: A negative charge generates an Electric field pointing toward it.

3. Lecture 3

3.1 Dipole and quadrupole

The law of addition of vectors allows us to find the electric field produced by two charges. If the two charges are of equal magnitude but opposite sign (and separated by a distance so that they do not cancel each other) then we have what is called a dipole. The electric lines can be found in fig.10. In the laboratory practice you will be able to use a program that draws these lines for you. Try several configurations. For example arrange four charges so that they look as two dipoles of opposite orientation and see what happens. In fig. 11 and 12 we illustrate typical outputs of the program. Besides the electric lines it also shows equipotential surfaces everywhere perpendicular to the electric field. We discuss what they mean below.

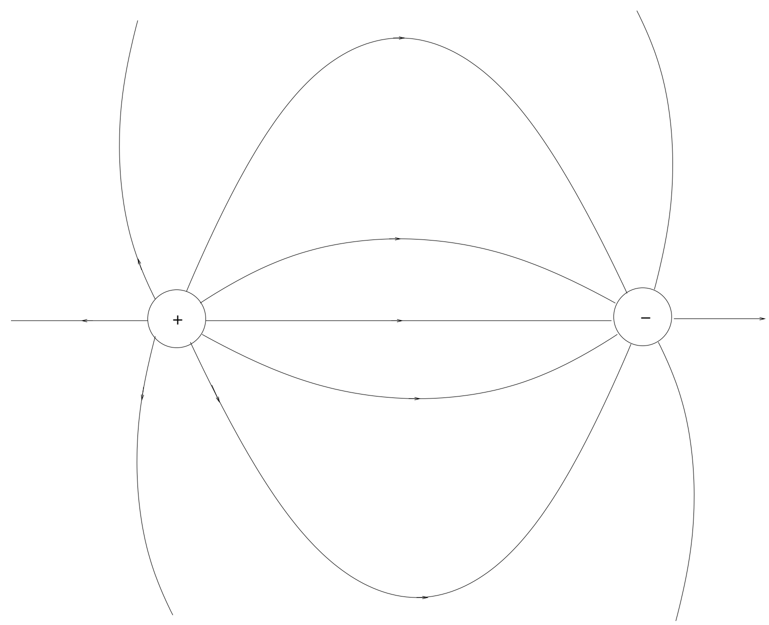


Figure 10: Electric field of a dipole obtained by adding (vectorially) the electric fields of each charge.

A simple rule for the lines of electric field is that they only start from positive charges and end in negative ones. They can never end "in the middle of the air" so to say. Before, we mentioned that electric field can exist independently of charges, in that case the lines of electric field have to close on themselves since there is nowhere for them to end. We will see how this works later.

3.2 Electrostatic energy

To motivate the idea of energy we use the Kelvin water dropper (fig.13), a simple

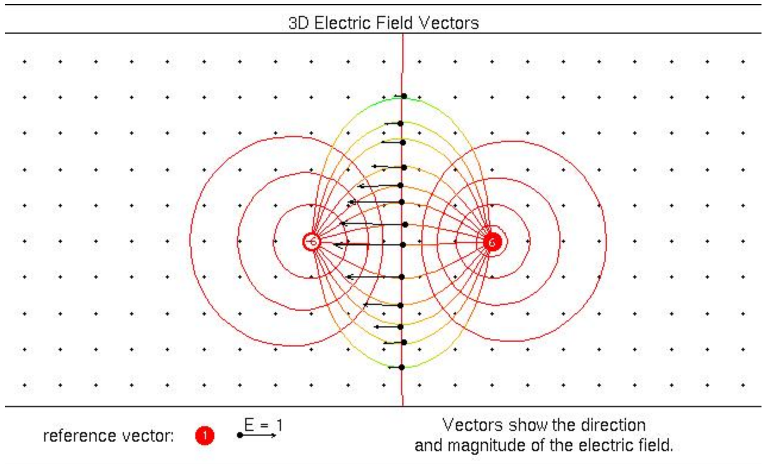


Figure 11: Typical output from the computer program you use in the lab. In this case illustrating the electric field and equipotentials of a dipole.

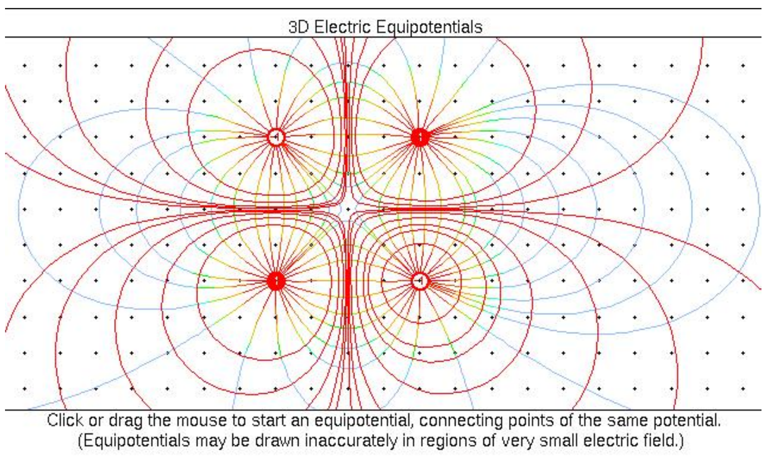


Figure 12: Illustrating the electric field and equipotentials of a quadrupole (meaning four poles).

device that generates electricity. The way it works is illustrated in fig.14. Assuming that one metallic jar is positively charged and the other negative, opposite charges will be induced in the streams of water falling through the respective rings. The water breaks in droplets which then fall into the jars accumulating charge. This is because of the cross connection between jars and rings which makes the positive droplets fall into the positively charged jar and the negative ones in the negatively charged one. If the water were not to break in droplets we would have a conducting circuit from the jars to the water tank, the charge will be repelled and would not accumulate in the jar. It is the same effect as when we discussed induction that we needed to cut the connection to ground and remove the rod. Ingeniously here the connection breaks simply because the water stream breaks into droplets. When enough charge is accumulated the electric field is strong enough to push the charge through the bulbs lighting them up which simultaneously neutralizes the system. Charge will build up again by repeating the procedure. The initial charge imbalance appears just randomly, the important effect is that any random charge fluctuation will be amplified by the device.

The idea we wanted to illustrate is also that we were able to generate electric power and so the energy has to come from somewhere. It might not be apparent initially where it is coming from. To figure that out one has to run the system and see when it will stop and what I need to replace to make it work again. Obviously many electric devices have batteries and we know they are the source of energy because when they run out we need to replace them or the device won't work anymore. Here it is clear that after a while the tank runs out of water. So we need to take the water from the jars and put it back up into the tank. To do that we need to lift water against gravity and that is where we put energy back into the system. Therefore, gravitational energy is being converted into electric energy. One way of seeing that is that, since the jars repel the water droplets, gravity is necessary to push them down into their respective jars allowing the accumulation of charge.

Having said that let us now look at how we understand energy when there are electric fields present.

From Newtonian mechanics we know that forces can be derived from potentials. Basically a particle feels a force that tries to make it move in the direction where it can decrease its potential energy faster. The same occurs in the case of electrostatics. Consider a positive charge Q which is fixed at some point and take another charge q that we are going to move. Let us say that the other charge is negative so there is an attractive force. If we want to move the charge q further apart by an amount Δr we need to do some work because there is a force that opposes us. The amount of work needed is

$$W = |\vec{F}|\Delta r = \Delta U = U_{\text{final}} - U_{\text{initial}} \tag{3.1}$$

where ΔU is the change in potential energy of the system. A couple of comments. When computing the work we should use only the component of the force in the direction of the motion. Since here we move parallel to the force the formula is OK. If not we need to multiply by $\cos\theta$ where θ is the angle between the force and the direction of motion. Second comment is that the formula is valid for constant force. Since the force depends on r we are going to consider $\Delta r \ll r$ and therefore to consider the force constant is a

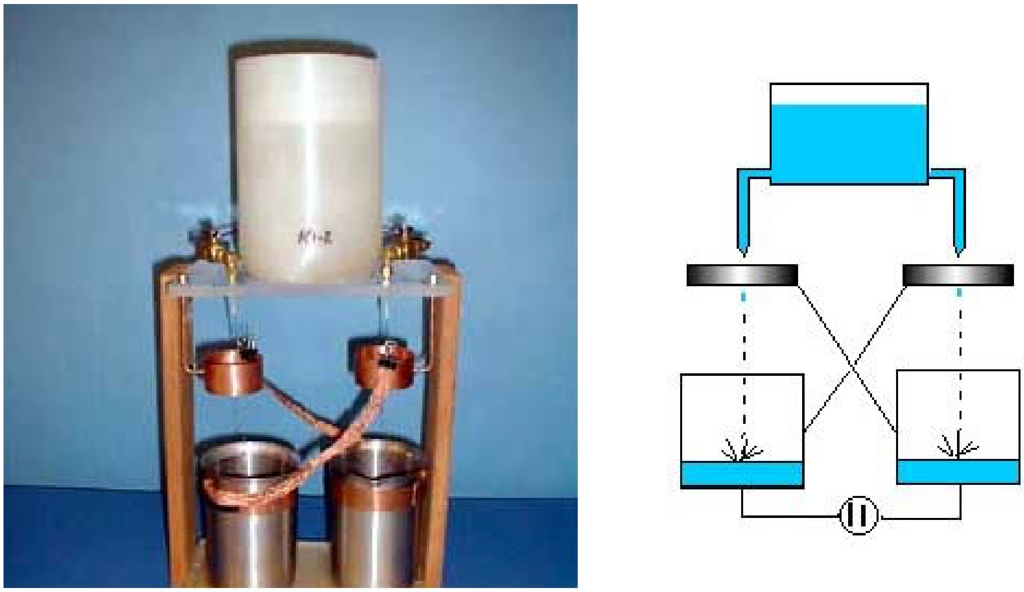


Figure 13: Kelvin invented this device to generate electricity. I find it of great interest since it is basically just water dropping from a tank and nevertheless it is able to generate electricity enough to light a few neon bulbs. Essentially illustrates how, by understanding how nature works one can create something interesting out of almost nothing.

good approximation. With this in mind we compute

$$\Delta U = U_{\text{final}} - U_{\text{initial}} = -\frac{1}{4\pi\varepsilon_0}\frac{Qq}{r^2}\delta r \tag{3.2}$$

The minus sign is because we should have $\Delta U > 0$ as discussed but the product of the charges is negative since they have opposite sign. We claim now that the potential energy is

$$U = \frac{1}{4\pi\varepsilon_0}\frac{Qq}{r} \tag{3.3}$$

Indeed the difference in potential energy between the initial and final situation is

$$\Delta U = U_{\text{final}} - U_{\text{initial}} = \frac{1}{4\pi\varepsilon_0}\frac{Qq}{r+\Delta r} - \frac{1}{4\pi\varepsilon_0}\frac{Qq}{r} = -\frac{1}{4\pi\varepsilon_0}\frac{Qq}{r(r+\Delta r)}\Delta r \tag{3.4}$$

If we use now that $\Delta r \ll r$ we find that

$$\Delta U = -\frac{1}{4\pi\varepsilon_0}\frac{Qq}{r^2}\delta r \tag{3.5}$$

as we computed from the force. This verifies our expression (3.3) for the potential energy.

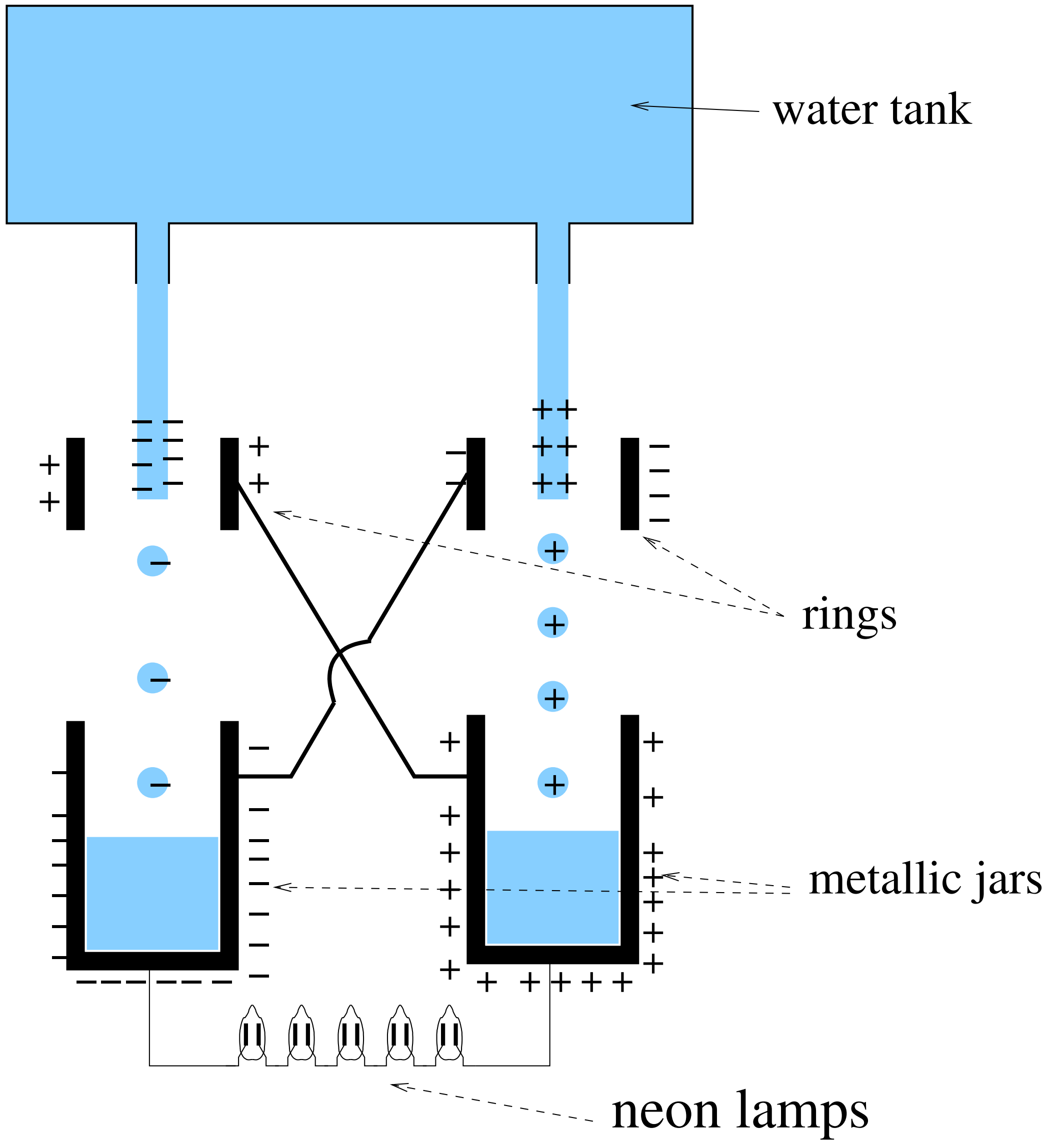


Figure 14: The key is the cross connection of the rings and the jars. This amplifies any charge difference accumulating opposite charges in the two metallic jars.

Another case in which we can compute the potential energy is that of a constant uniform electric field. Namely $\vec{E}$ is independent of the position, that is a probe charge feels the same force no matter where it is located. Of course this is an idealized situation but in many cases is a good approximation for the electric field in certain regions where it does not change very much. In such case, let us say that the electric field points in the direction $\hat{x}$. If we move a charge in directions, $\hat{y}$ or $\hat{z}$, perpendicular to $\hat{x}$, the force does not oppose or help the motion so we do not need to do any work. If we move it in direction $\hat{x}$ however, since we move in the direction of the force we extract work, or, we make negative work:

$$W = -|\vec{F}|\Delta x = -q|\vec{E}|\Delta x = \delta U = U_{\text{final}} - U_{\text{initial}} = U(x + \Delta x) - U(x) \tag{3.6}$$

The final energy of the system is therefore smaller than the initial one. We see that the equation is satisfied if the potential energy is simply given by

$$U = -q|\vec{E}|x \tag{3.7}$$

Another way to figure out the sign is to notice that the charge will move in the direction of decreasing energy. So, if $q > 0$ then U has to decrease toward the right and therefore the minus sign.

3.3 Electrostatic potential

We see, as a consequence of the force being proportional to the charge, that so is the potential energy. For that reason we define the electrostatic potential $V(x, y, z)$ such that is we put a charge at position (x, y, z) the energy of the system changes by

$$U = qV(x, y, z) \tag{3.8}$$

For a charge Q the electrostatic potential is given by

$$V = \frac{1}{4\pi\varepsilon_0}\frac{Q}{r}, \qquad \text{where, as always,} \qquad r = \sqrt{x^2 + y^2 + z^2} \tag{3.9}$$

On the other hand, for a constant electric field in direction $\hat{x}$ it is

$$V = -|\vec{E}|x \tag{3.10}$$

By analogy with the equation that determines the work done in terms of the potential energy we see that if we move by a distance Δx in a direction parallel to the electric field the change in electrostatic potential is

$$\Delta V = -|\vec{E}|\Delta x \tag{3.11}$$

Equivalently the component of $\vec{E}$ in direction x is

$$E_x = -\frac{\Delta V}{\Delta x} \tag{3.12}$$

So, the electric field can be computed from the potential by moving in a direction x by an amount Δx. The change in potential ΔV determines the electric field through eq.(3.12). This actually works if we move in any direction, it always give the component of $\vec{E}$ in that particular direction (or equivalently the projection of $\vec{E}$ along the direction of motion). For example, if we move in a direction perpendicular to $\vec{E}$ then there is no change in the potential. This gives rise to the notion of equipotential surfaces. This means surfaces of the same (equal) potential, that is surfaces where V has a constant value. From what we just said, such surfaces should be perpendicular to the electric field. For example, for a single charge they are spheres concentric with the charge. For a constant electric field they are planes perpendicular to $\vec{E}$. In other cases, for example for the dipole, they are more complicated but one can have some idea by drawing the electric field and then drawing surfaces perpendicular to it (see fig. 11).

An important property of equipotential surfaces is that if we move a charge along such surface we do not do any work.

Another important property is that to be in a static situation, namely with charges not moving, the surface of a conductor has to be an equipotential surface. This is because on the surface of the conductor the electric field is perpendicular to it, otherwise if there were a component parallel to the surface, charges would move until they cancel the electric field. Similarly, if charges are not moving then inside a conductor the electric field is zero, implying that all the conductor has the same value of the potential. We should emphasize that this refers to the static situation. If charges are moving then the electric field inside a conductor need not be zero and the conductor need not be all at the same potential.

The principle of superposition also applies to the electrostatic potential. So if we have several charges, the total electrostatic potential is the sum of the electrostatic potentials due to each of them. It becomes apparent the usefulness of the potential, since the potential add as numbers as opposed to the electric field which add as vectors.

Finally the potential is measured in Nm/C since the electric field is measured in N/C and Δx in meters. Since this is a very common unit it has received the special name of Volt (V). So we have

$$1V = 1\frac{Nm}{C} \tag{3.13}$$

The potential difference many times is referred simply as "voltage"

Another demo illustrates the existence of this potential by measuring directly the potential of a charged sphere and another by showing how putting a fluorescent light in the electric field produces a discharge because of the potential difference between the two ends of the tube (fig.15).

Figure 15: A voltmeter and a fluorescent bulb make manifest the existence of a potential difference (and electric field).

4. Lecture 4

4.1 More on electrostatic potential

From formula (3.12) we see that V cannot have abrupt jumps otherwise the electric field would be infinite. One example is the potential of a charged sphere which is constant inside and decreases as $1/r$ outside. The electric field on the other hand has a jump, it is zero inside and has a finite value on the surface. We plot this in figs.16 and 17.

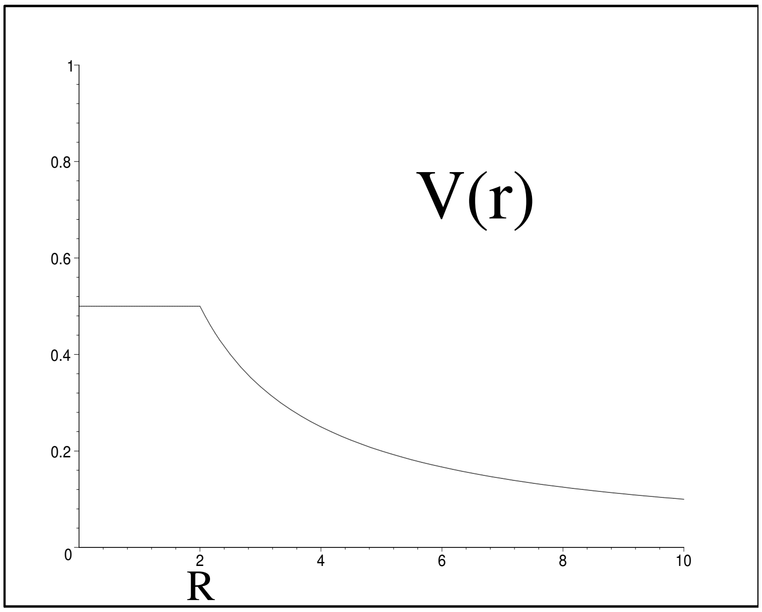


Figure 16: Dependence of the electric potential with the distance to a center of a charged sphere.

An interesting situation that we have already discussed is what happens if we touch two charged spheres of different radii, how is the total charge distributed between the two?.

Now we know that it distributes so that the potential is constant. In a first approximation we can ignore the presence of the other sphere and say that the potential on the surface of each sphere is given by

$$V_1 = \frac{1}{4\pi\varepsilon_0}\frac{Q_1}{R_1}, \qquad V_2 = \frac{1}{4\pi\varepsilon_0}\frac{Q_2}{R_2} \tag{4.1}$$

where $Q_{1,2}$ and $R_{1,2}$ are the charges and radii of each sphere respectively. If the potentials are equal ($V_1 = V_2$) then we evidently have

$$\frac{Q_1}{R_1} = \frac{Q_2}{R_2} \tag{4.2}$$

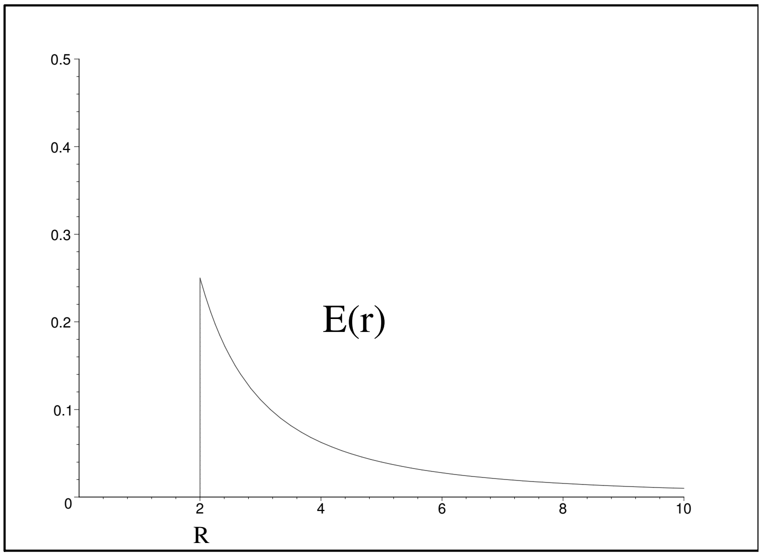


Figure 17: Dependence of the electric field with the distance to a center of a charged sphere.

implying that the smallest sphere has less charge, namely the charge is proportional to the radius as mentioned before. Naively or pictorially speaking the charges prefer to be in the large sphere since they repel each other and can then be further apart. However one important observation is that the electric field is larger in the surface of the smaller sphere. Indeed we have

$$|\vec{E}_1| = \frac{1}{4\pi\varepsilon_0}\frac{Q_1}{R_1^2}, \qquad |\vec{E}_2| = \frac{1}{4\pi\varepsilon_0}\frac{Q_2}{R_2^2}, \tag{4.3}$$

From here, and eq.(4.2) it is evident that

$$|\vec{E}_1|R_1 = |\vec{E}_2|R_2. \tag{4.4}$$

We see that if R_1 is very small then $|\vec{E}_1|$ is very large. Notice this is because, although Q_1 is smaller, we can get closer to the center of the small sphere. An extreme case is when we have a sharp point which can be thought as a sphere of almost zero size. Near it the electric field will be extremely large ionizing the air and allowing charge to flow out of the conductor. This is the principle of the lightning rod invented by Benjamin Franklin. Lightning is essentially a giant spark. The lightning rod ionizes the air around making it conductive and favoring the initiation of the spark. Therefore lightning is more likely to strike on the lightning rod than on the structure that it protects.

Another important observation is that, since the electric field inside a conductor is zero, it will still be zero if we carve a hole inside it (fig.19). This means that

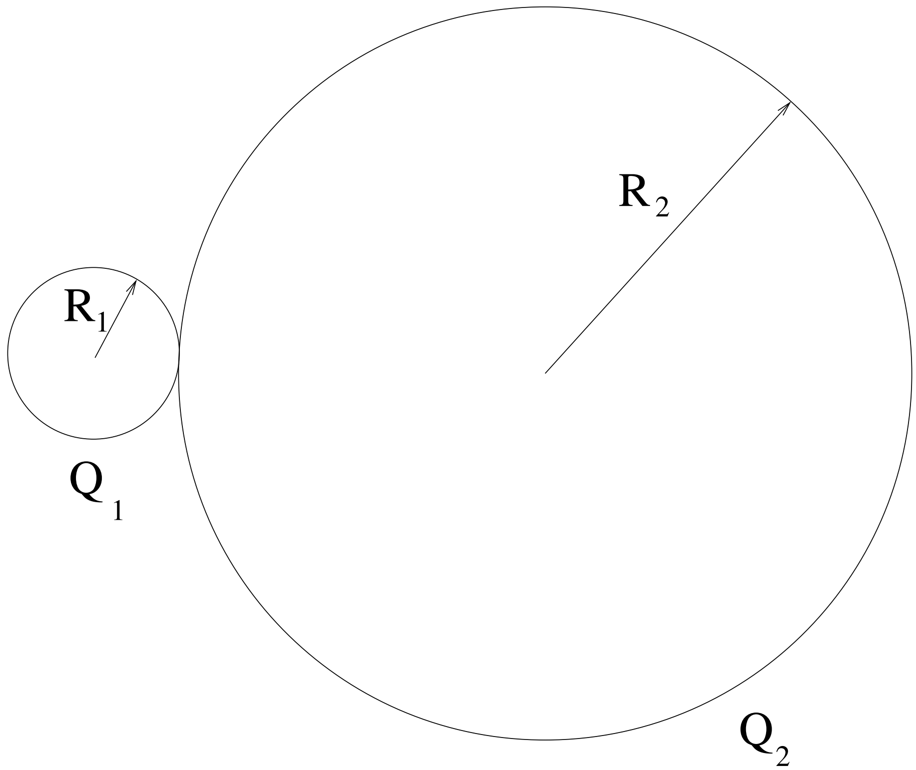


Figure 18: Distribution of charges between two conducting sphere of different radii and in contact with each other.

conductors act as shields for electric fields. An important application is the so called Faraday cage. Inside a metallic cage no electric field penetrates and therefore it is isolated from electromagnetic waves, namely radios, cell phones etc. do not work.

4.2 Electric flux

One interesting concept that can be defined for a vector field such as the electric field is that of flux through a surface. The idea originates from considering the motion of water. If water is moving at velocity $\vec{v}$ then the flow of water (usually given in liters per second, or gallons per minute) through a surface of area A is given by

$$\text{flow} = v_{\perp} A = |\vec{v}| A \cos\theta \tag{4.5}$$

where $v_{\perp} = |\vec{v}| \cos\theta$ is the component of the velocity perpendicular to the surface. Here θ is the angle between the normal and the velocity. To see why this is so consider figure 20. During a time interval Δt it is clear that all the fluid contained in the volume $A|\vec{v}| \cos\theta\, \Delta t$ will go through the area A and hence the result.

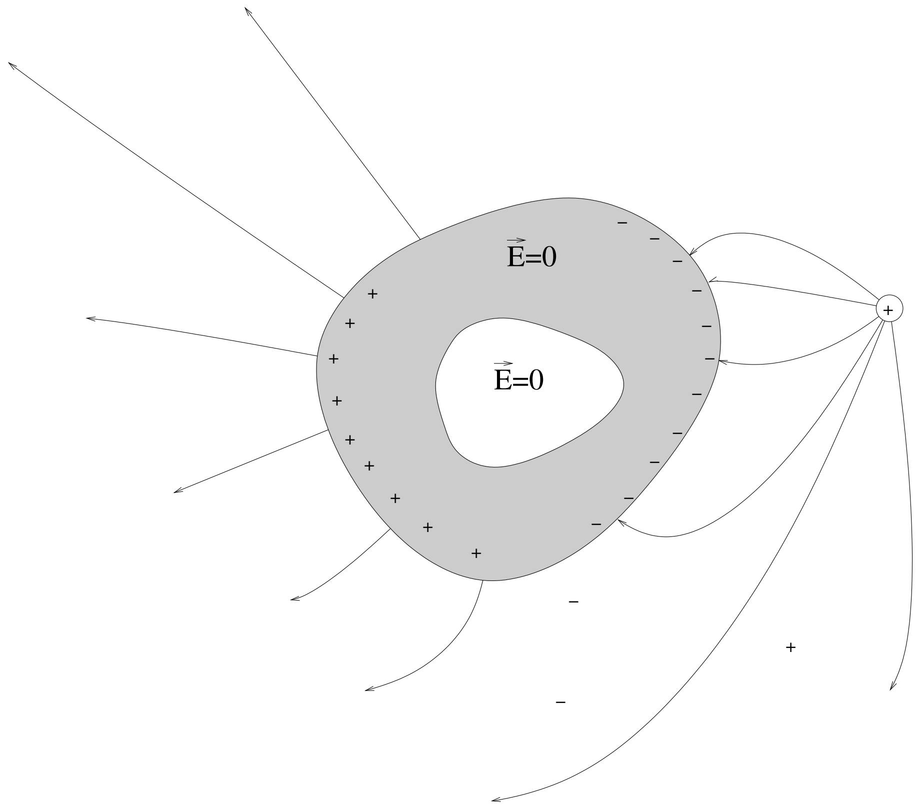


Figure 19: In a region with no charge and surrounded by a conductor the electric field is zero.

By analogy[1], given a surface we define the electric flux through it as the area of the surface times the value of the component of the electric field perpendicular to the surface. The easiest case is when the surface is everywhere perpendicular to $\vec{E}$ in which case we just multiply area times $|\vec{E}|$. For example for a single charge, we can take as a surface the sphere of radius R concentric with the charge (fig.21). In that case we have that the surface is perpendicular to $\vec{E}$ and the flux is

$$\text{flux} = 4\pi R^2 \frac{1}{4\pi\varepsilon_0} \frac{Q}{R^2} = \frac{Q}{\varepsilon_0} \tag{4.6}$$

namely it is independent of the radius R!. Moreover, consider a cone that determines a surface as the one in figure 22, namely a truncated cone where the base and the top are spherical. Notice that the areas are related by

$$\frac{A_1}{R_1^2} = \frac{A_2}{R_2^2} \tag{4.7}$$

[1]In the case of the electric field the situation is static, there no fluid moving but it is a useful analogy

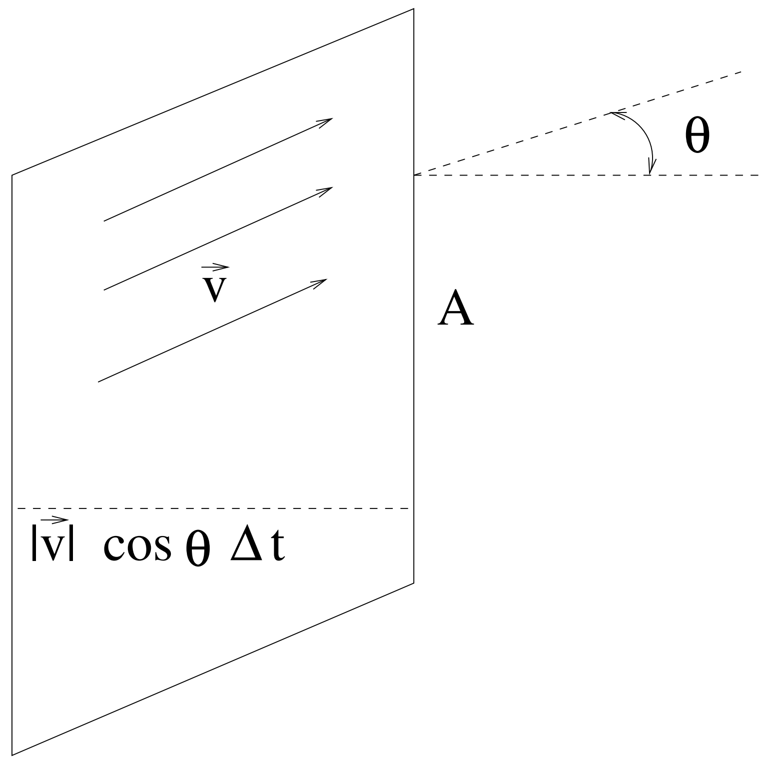


Figure 20: Flow of water through a cross section area A.

because the area of an object scales as the square of the (linear) size. Along the laterals of the truncated cone the flux is zero since it is parallel to the electric field. Through the base and the top the flux has opposite signs since it is entering through one and exiting through the other. The total flux is

$$\text{flux} = \frac{1}{4\pi\varepsilon_0}\frac{Q}{R_1^2}A_1 - \frac{1}{4\pi\varepsilon_0}\frac{Q}{R_2^2}A_2 = 0 \tag{4.8}$$

It vanishes in view of eq.(4.7). Therefore the flux through this surface which does not surround the charge is zero. With some imagination one can think of constructing any surface with very small blocks of this (truncated) conical shape and the result will be the same. If the surface does not surround the charge the flux is zero and if it does then the flux is $\frac{Q}{\varepsilon_0}$. The superposition principle also applies to the flux, so if the surface surrounds several charges the flux is given by the total charge surrounded. This is called Gauss's theorem, the flux through a closed surface is given by $\frac{Q}{\varepsilon_0}$ where Q is the total charge surrounded by the surface. It is a very powerful theorem since one can take a surface very far from the charges and knowing the electric field there is enough information to know how much charge we enclose.

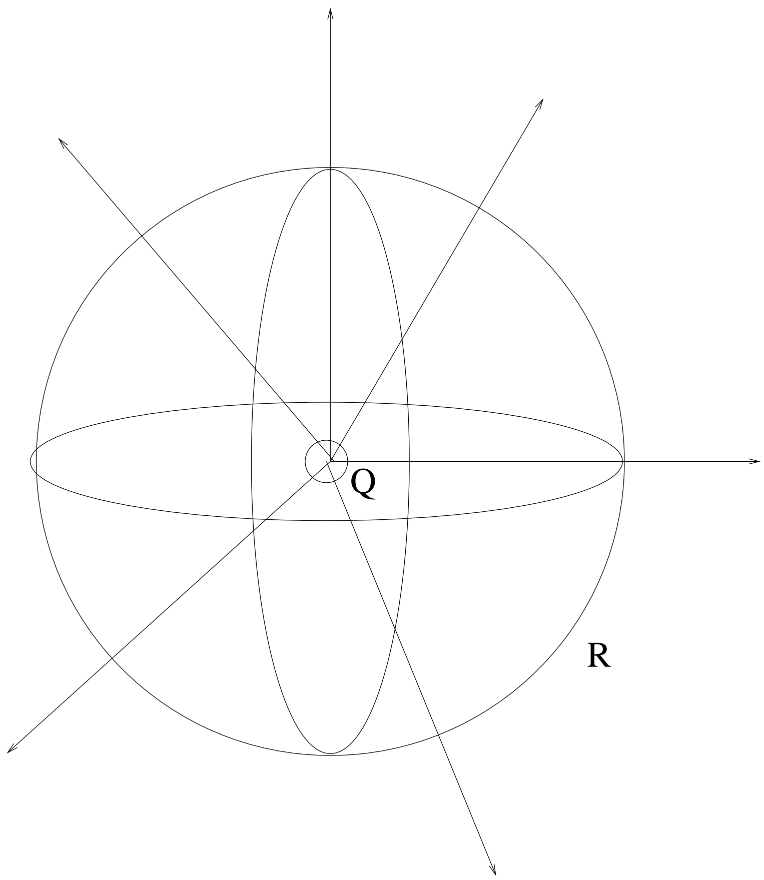


Figure 21: Computing the flux of the electric field produced by a charge through a concentric sphere of radius R.

One simple example where we can use the theorem is for a flat surface with a charge density σ. Namely, if we cut a region of area A of the surface, the charge contained is σA. Looking at fig.23 and by symmetry the electric field will be uniform and pointing away from the plane. Taking different surfaces and computing the flux we see that

$$2|\vec{E}|A = \frac{\sigma A}{\varepsilon_0} \quad \Rightarrow \quad |\vec{E}| = \frac{\sigma}{2\varepsilon_0} \tag{4.9}$$

If the plane is (x, y), that is normal to $\hat{z}$ then we have an electric field

$$\vec{E} = \begin{cases} \frac{\sigma}{2\varepsilon_0}\hat{z} & \text{if } z > 0 \\ -\frac{\sigma}{2\varepsilon_0}\hat{z} & \text{if } z < 0 \end{cases} \tag{4.10}$$

In fact you can show using Gauss theorem that the electric field has to be independent from the distance to the plane. Try to see why this is so using a cylindrical surface which does not intersect the plane.

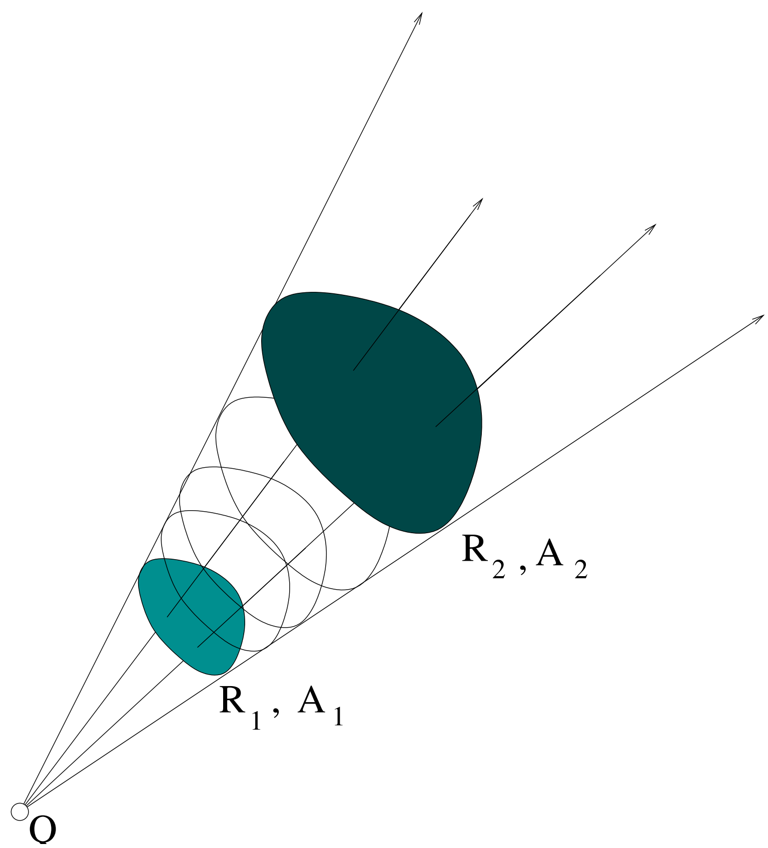


Figure 22: For the truncated cone in the figure, the total flux is zero since the flux incoming at the bottom is equal to the outgoing at the top.

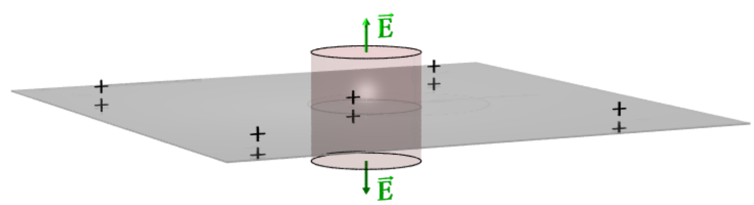


Figure 23: The electric field created by a plane can be computed using Gauss law.

4.3 Electric current

As we explained, electrostatic forces are responsible for chemistry, explain why solids are solids whereas liquids are liquids etc. It has practical applications in ink-jet printers, electrophoresis, cathode tubes, particle accelerators, etc. However, by far the most common use of electricity is through the use of electric currents and circuits.

The idea is that if we have a constant electric field in a conductor, this will produce a motion of charges until the electric field vanishes inside. However the idea emerges that if we can remove the charges from one side and put it back in the other then we can have a continuous motion of charge from one place to another. This "removal" of charge should be done against the potential and therefore requires energy. The device that produces such effect is known as a battery. Usually the energy in a battery comes from a chemical reaction inside it. Such chemical reactions are capable of moving charges across a potential difference of around 1 V (the usual battery if 1.5 V). A simple comment is how does then the 9V battery work?. Well, by opening up one it is easy to see that there are 6 "elements" inside it. That essentially means six 1.5V batteries connected in series, that is one after the other. We see that the potential just adds as actually follows from the properties of V we discussed before.

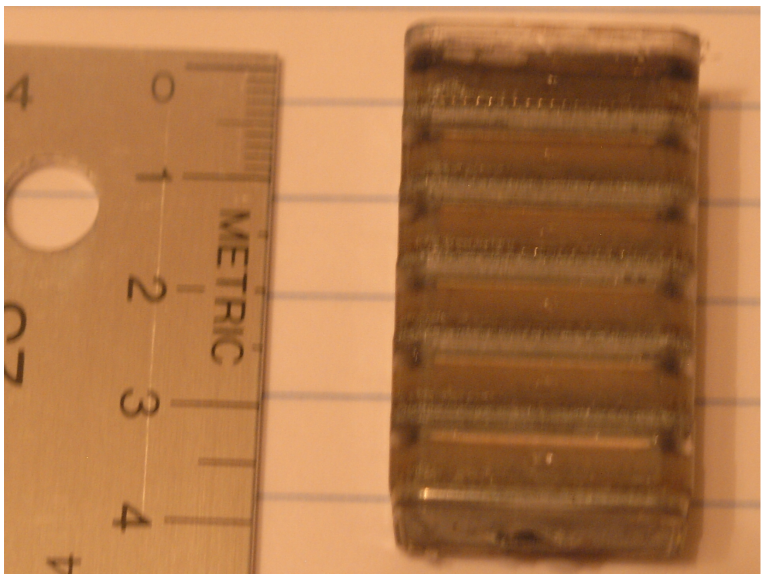


Figure 24: Typical chemical batteries have an emf of 1.5V. A 9V battery consists of 6 of those in series, that is one after the other.

If there were no resistance to the flow of charge, the electrons would accelerate indefinitely inside the conductor. As it were, there is an effective maximum velocity and therefore, after a short time, if a conductor is connected to a battery a steady flow of charge will be present. Such flow of charge is known as a current and is measured in Amperes (or Amps). A current of 1A means that one Coulomb of charge is going through a section of the conductor every second $1A = 1\frac{C}{s}$. Experimentally it turns out that the current inside a conductor is proportional to the potential difference or voltage across it. This is called Ohm's law and reads:

$$V = IR \tag{4.11}$$

where V is the voltage across a conductor, I is the current and R is the resistance. The resistance in measured in Ohms (Ω) where $1\Omega = 1\frac{V}{A}$. So a potential difference of 1V across a resistor of 1 Ω creates a current of 1A. Remember also that the potential always decreases along the direction of the current.

Examples of resistors are shown in figure 26. Typically such resistors are color coded. By reading the colors one can figure out the resistance.

digit	color	multiplier
-	silver	0.01
-	gold	0.1
0	black	1
1	brown	10
2	red	100
3	orange	1K
4	yellow	10K
5	green	100K
6	blue	1M
7	violet	10M
8	gray	
9	white	

The way it works is explained in fig.27 where we see that the value is written with the first three stripes. The first two are read as a number using the table. The third one is a multiplier whose value is also taken from the table. For example red black orange is read as $20K\Omega = 20,000\Omega$.

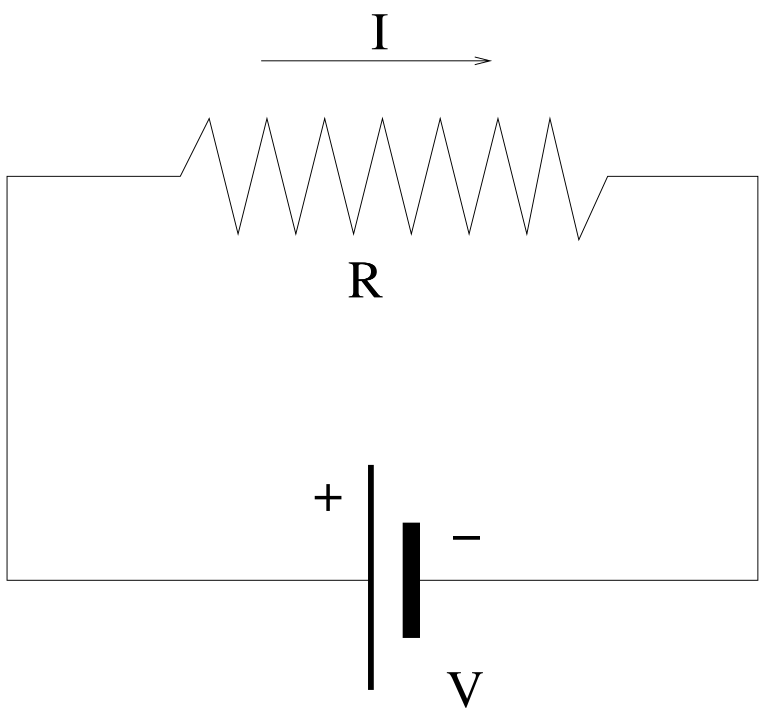


Figure 25: Simple circuit with a battery and a resistor. By Ohm's law we have $V = I\,R$.

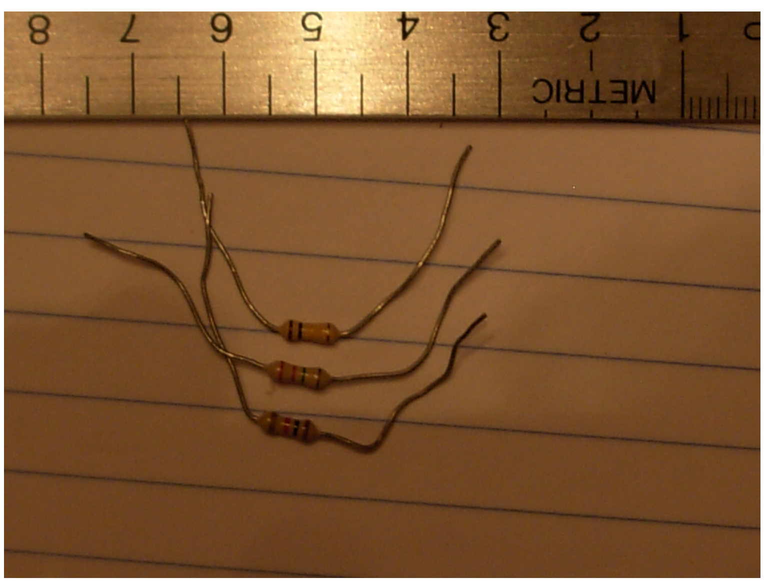


Figure 26: Examples of commonly used resistors.

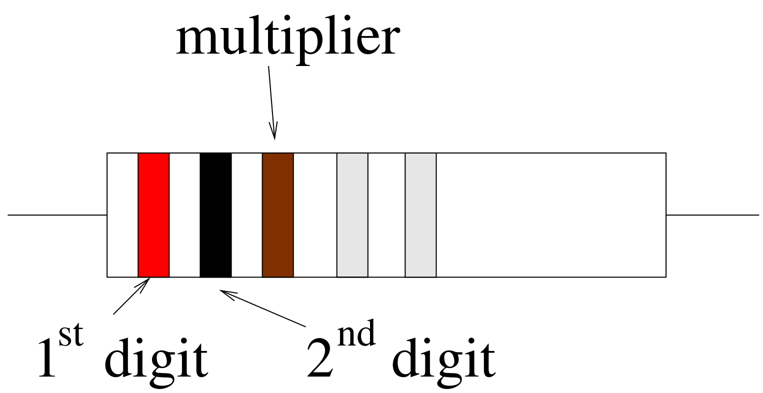


Figure 27: The first three stripes indicate the value of the resistance according to the color code table. The other ones are related to tolerance and reliability.

5. Lecture 5

5.1 Resistivity

For a conductor of uniform cross sectional area A it turns out experimentally that the resistance is given by

$$R = \rho \frac{L}{A} \tag{5.1}$$

where L is the length and ρ is the resistivity (measured in $\Omega \cdot m$), a property of the material. So, the longer the cable and the smaller the cross section the larger the resistance. It seems that the lower the resistance the better, however in electric circuits sometimes a resistance plays an important role so there are special components called resistors which, although being conductors have a relatively high resistance. We already discussed them in the previous lecture. A typical application is to limit the current in a circuit. Namely, for a given V, a very small resistance will create a large current I which could damage some component. In fact it is easy to see that energy is being dissipated in a conductor since charge is moving from a value of the potential to another (see fig.28). The energy difference ΔU gained when transferring a charge Q across a potential difference ΔV is

$$\Delta U = Q\Delta V \tag{5.2}$$

Such energy is converted into heat. In a given time interval Δt an amount of charge $Q = I\Delta t$ is transferred so the energy is

$$\Delta U = I\Delta V\Delta t \tag{5.3}$$

or

$$P = \frac{\Delta U}{\Delta t} = I\Delta V \tag{5.4}$$

Here P is the power, namely energy per unit time that is dissipated in the resistor. It is converted into heat. Such heat has to be removed by cooling the equipment either passively (letting air dissipate the heat) or actively (for example with a cooling fan). Using Ohm's law we have alternative equivalent expressions for the power

$$P = I \cdot \Delta V = I^2 R = \frac{(\Delta V)^2}{R} \tag{5.5}$$

For a given ΔV if the resistance R is very small the power dissipated is very large. This is called a short-circuit, the large amount of heat produced usually melts the insulation and even the conductor itself with the consequent fire risk. All materials conduct a small amount of electricity but an idea of the difference between an insulator and a conductor can be seen in the following table where the resistivity of various material is given.

Material	Resistivity ($\Omega \cdot m$)
Silver	1.59×10^{-8}
Copper	1.68×10^{-8}
Aluminum	2.82×10^{-8}
Tungsten	5.60×10^{-8}
Zinc	5.90×10^{-8}
Nickel	6.99×10^{-8}
Iron	1.0×10^{-7}
Germanium	4.6×10^{-1}
seawater	2×10^{-1}
Silicon	6.40×10^{2}
Glass	10^{10} to 10^{14}

Between copper and glass there is a factor of at least 10^{18}. We should also mention that resistivity depends on the temperature. The table are typical values at 20 ℃.

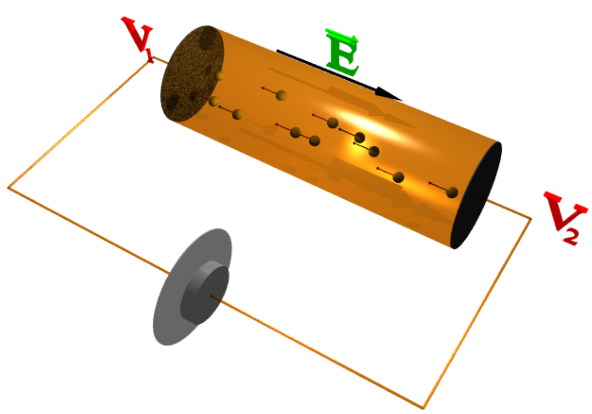


Figure 28: A battery connected to a conductor produces a steady flow of current. The electrons move opposite to the direction of the current. When going across the potential they gain energy which is converted into heat by collision with the atoms.

5.2 Capacitors

A capacitor is a device that can be used to store electric energy that can be used later on. It has numerous applications some of which we are going to discuss in the rest of this lecture or later in the course. In fact if you open any electronic device one easily finds several of them.

In its most simple form it has two parallel surfaces of area A separated by a distance d. One surface has charge Q and the other $-Q$. Since the electric field of a flat surface we already discussed, we can easily find that the electric field cancels outside the capacitor and inside it is given by $|\vec{E}| = \frac{Q}{A\varepsilon_0}$. Notice that $\sigma = \frac{Q}{A}$ and the factor of two goes away because we have two planes. Given the electric field we use the formula for the potential difference:

$$\Delta V = -|\vec{E}|\Delta x = -\frac{Qd}{A\varepsilon_0} \tag{5.6}$$

The surface with positive charge has the largest potential. We also see that the potential difference is proportional to Q. If we define the capacity C (do not confuse with the symbol for Coulomb!) through the equation

$$Q = C\,\Delta V \tag{5.7}$$

then we find the capacity equal to

$$C = \frac{A\varepsilon_0}{d} \tag{5.8}$$

From the definition (5.7) we see that the unit of capacity is C/V (here C is Coulombs!) a unit known as Faraday (F). That is $1F = 1\frac{C}{V}$. Although this is just an example most of the capacitors in practical applications are similar. Commonly the planes are very thin and an insulator is in the middle. The foils are then rolled so that capacitor occupies less space but it essentially works in the same way. If we want to have large capacity we need d very small and that's why thin insulators sheets between the plates are used and also large area which is why we need to roll it to occupy less volume. In practical circuits a Faraday is a large unit so capacity is usually measured in $\mu F = 10^{-6}F$, that is micro Faraday or even $nF = 10^{-9}F$ or $pF = 10^{-12}F$. In the picture (fig. 30) we see some examples.

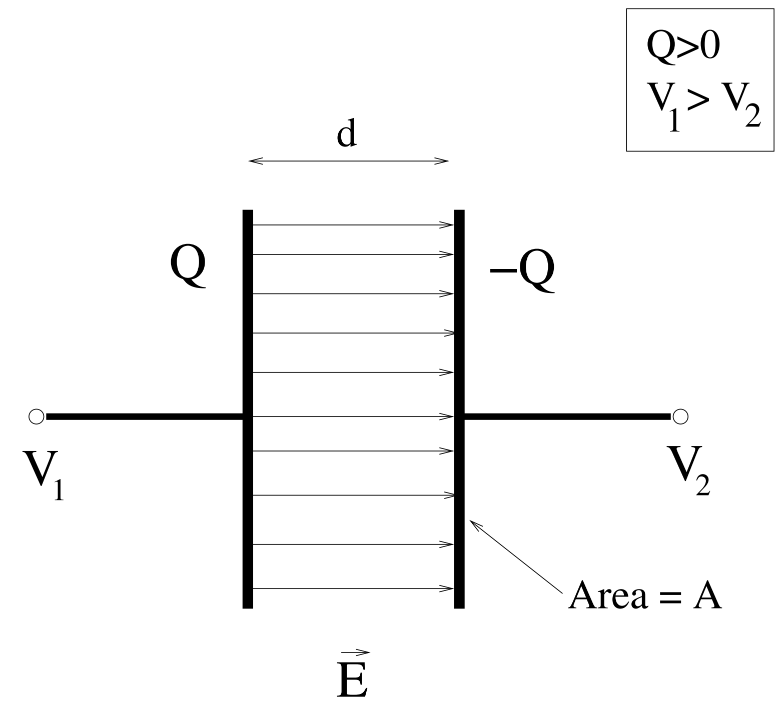


Figure 29: A simple capacitor consists of two oppositely charged surfaces. This configuration stores energy.

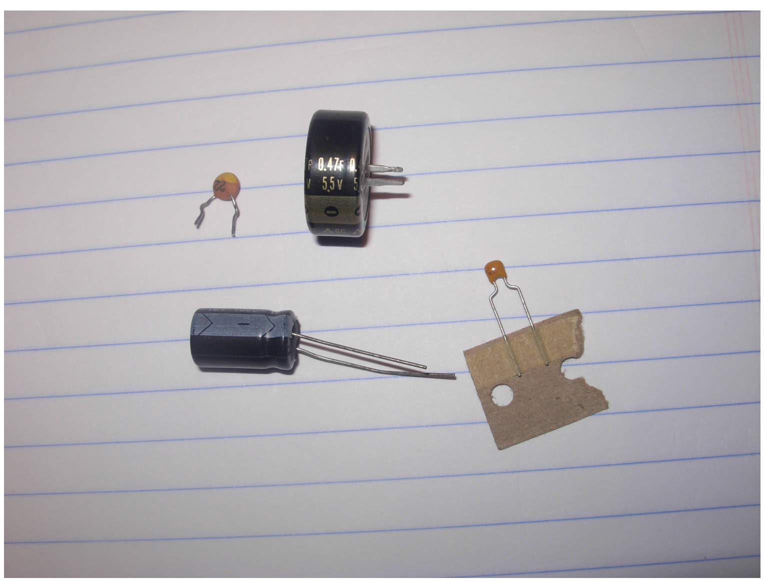


Figure 30: Examples of capacitors. The electrolytic ones have to be connected in a particular polarity. The – signs indicate that the corresponding terminal should always have lower potential, namely if connected to a battery it should be connected to the negative terminal.

6. Lecture 6

6.1 Energy contained in a capacitor

We said that a capacitor stores energy. We now compute how much energy it actually contains. A way to do that is to consider how much work we need to separate the two plates that form the capacitor. Suppose the left plate is fixed and we move the right one by a distance Δd (see fig.31). The electric field produced by the other plate is $|\vec{E}| = \frac{\sigma}{2\varepsilon_0}$ where $\sigma = \frac{Q}{A}$. The force is then

$$|\vec{F}| = \frac{\sigma}{2\varepsilon_0} Q \tag{6.1}$$

The work we need to do to separates the plates by a distance Δd is

$$W = \Delta U = U_{\text{final}} - U_{\text{initial}} = |\vec{F}|\Delta d = \frac{\sigma}{2\varepsilon_0} Q \Delta d \tag{6.2}$$

From here we find that

$$U = \frac{\sigma}{2\varepsilon_0} Q d = \frac{1}{2} Q^2 \frac{d}{A\varepsilon_0} = \frac{1}{2}\frac{Q^2}{C} \tag{6.3}$$

where we remembered that the capacity is

$$C = \frac{A\varepsilon_0}{d} \tag{6.4}$$

Using that $Q = C\,\Delta V$ we find the equivalent expressions for the energy contained in the capacitor:

$$U = \frac{1}{2}\frac{Q^2}{C} = \frac{1}{2} C\,(\Delta V)^2 = \frac{1}{2} Q \Delta V \tag{6.5}$$

6.2 Dielectrics

We mentioned before that between the plates in a capacitor we include an insulator. This can actually modify the properties of the capacitor and increase the capacity. The reason is that insulators are generically dielectric. This means that their electrical properties can be understood as if they are made out of tiny dipoles. The dipoles align themselves with the electric field decreasing its value in a phenomenon known as screening. If we consider a single charge inside a dielectric the screening phenomenon is illustrated in figure 32. Notice that outside the dielectric the field is the same from symmetry and Gauss law. Experimentally the result is that the electric field inside is

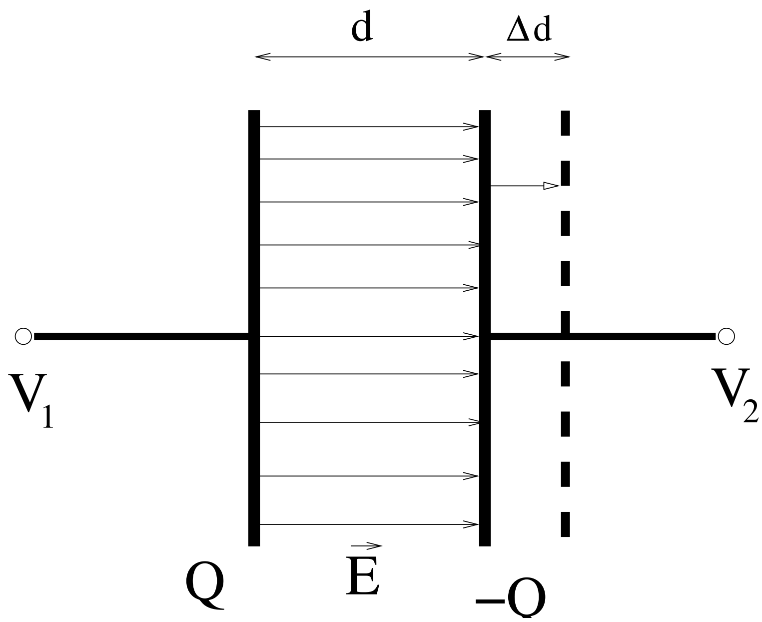


Figure 31: The energy stored in a capacitor can be computed by doing a small displacement of the plates by Δd and computing the work needed to do that.

suppressed by a factor $\varepsilon_0/\varepsilon$ where ϵ is known as the permitivity of the medium. The electric field inside the dielectric is then

$$\vec{E} = \frac{1}{4\pi\varepsilon}\frac{Q}{r^2}\hat{r} \tag{6.6}$$

outside is again

$$\vec{E} = \frac{1}{4\pi\varepsilon_0}\frac{Q}{r^2}\hat{r} \tag{6.7}$$

In the case of the capacitor it works the same, we just need to replace $\varepsilon_0 \to \varepsilon$. Therefore the capacity when there is a dielectric in between the plates is

$$C = \frac{A\varepsilon}{d} \tag{6.8}$$

Typical values for $\varepsilon/\varepsilon_0$ are

Material	$\frac{\varepsilon}{\varepsilon_0}$
Vacuum	1 (definition)
Air	1.0006
Teflon	2.1
Polyethylene	2.25
Paper	3.5
Water	80

We see that air and vacuum are very close to each other which is why we ignore the difference. It also shows that if we want to store more energy in a capacitor, at the same voltage it is convenient to use a dielectric of high dielectric constant.

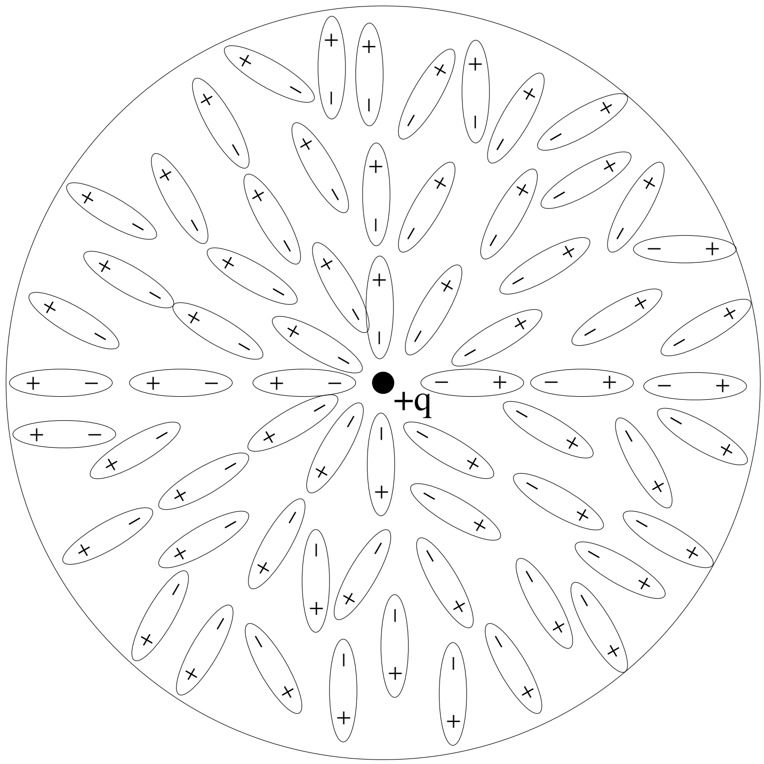


Figure 32: A charge in a dielectric is partially screened by the dipoles aligning with the electric field. Due to temperature and other variables the alignments is partially random. As a result the field inside the dielectric is smaller that without the medium. Outside the medium, in this case, the electric field is the same with or without the dielectric. This is because the configuration is spherically symmetric and using Gauss's law, the electric field is given by the total charge enclosed in a surface (the dielectric has no net charge).

6.3 RC circuits

Now that we know what a capacitor is we can see how to charge it and what applications we can find for it. The simplest way to charge it is with a battery and through a resistor. One might think the resistor is not necessary but not matter what we do there is some resistance in the circuit so we need to include it. Also, sometimes we do not want to charge the capacitor as fast s we can and, as we are going to see, a resistor makes the charging process slower. When the contact is closed, current starts to circulate but it cannot go through the capacitor so charge starts to accumulate in it. Initially the current is just $I = \frac{\Delta V}{R}$ but as the capacitor charges a potential difference appears and the current is smaller. In the final stage, the capacitor is completely charged, the potential difference between its terminals is the same as the battery (but opposes it) and no more current circulates. An estimate of the time the capacitor takes to charge is obtained by dividing the total charge by the initial current:

$$\tau = \frac{Q}{I} = \frac{C\Delta V}{\Delta V/R} = RC \tag{6.9}$$

This is a characteristic time of the circuit. The larger the resistance and the capacity the larger the time it takes to charge. We can then discharge the capacitor through a smaller resistance for example and it will release its energy in a much shorter interval. An application is the flash in a photographic camera. The battery with relatively large internal resistance charges a capacitor and then it is suddenly discharged though the flash. Another typical application is in timing circuits. For example oscillators etc. A resistor of variable resistance is inserted and the time of charge and discharge is used to determine the period of the oscillations. It is not very precise so it can be used only when the exact timing is not crucial. Another important application is in computer memories. A charged capacitor for example can represent a 1 and a discharged one a 0. In that case we want the capacity small to be able to charge it and discharge it fast. However if we do not read the memory current will leak and the capacitor will be discharged. For that reason when the computer is turned off what was in the memory disappears. When it is working a special circuit continuously reads the memory and rewrites it before the capacitors have time to discharge (these memories are called DRAM and are presently the most common ones).

We can analyze the circuit in slightly more detail by defining some points (a), (b), (c) as in figure 33 and considering the voltages at those points V_a, V_b, V_c. Using Ohm's law $\Delta V = IR$ for the resistor and $Q = C\Delta V$ for the capacitor we have

$$V_b - V_c = \Delta V_{\text{resistor}} = IR \tag{6.10}$$

$$V_a - V_b = \Delta V_{\text{capacitor}} = \frac{Q}{C} \tag{6.11}$$

$$V_a - V_c = \Delta V_{\text{battery}} \tag{6.12}$$

If we add the first two we have

$$V_b - V_c + V_a - V_b = IR + \frac{Q}{C} = V_a - V_c = \Delta V_{\text{battery}} \tag{6.13}$$

So the current circulating through the circuit is

$$I = \frac{1}{R}\left(\Delta V_{\text{battery}} - \frac{Q}{C}\right) \tag{6.14}$$

Since $\Delta V_{\text{Battery}}$ is fixed we see that initially, when the capacitor is not charged ($Q = 0$) the current is $I = \frac{1}{R}\Delta V_{\text{Battery}}$. The current starts to charge the capacitor, Q increases and I decreases but keeps charging the capacitor until $Q = C\Delta V_{\text{battery}}$ which gives $I = 0$. If we plot the voltage across the capacitor and the current we get a plot as in fig.34.

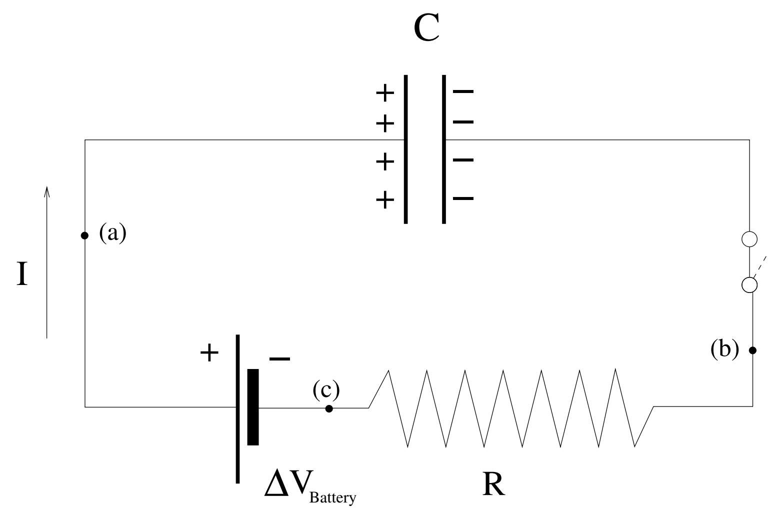


Figure 33: A capacitor is usually charged with a battery and through a resistor. The time it takes to charge depends on the capacity and the resistor. It is given by $\Delta t = RC$ and is a characteristic time of this type of circuit.

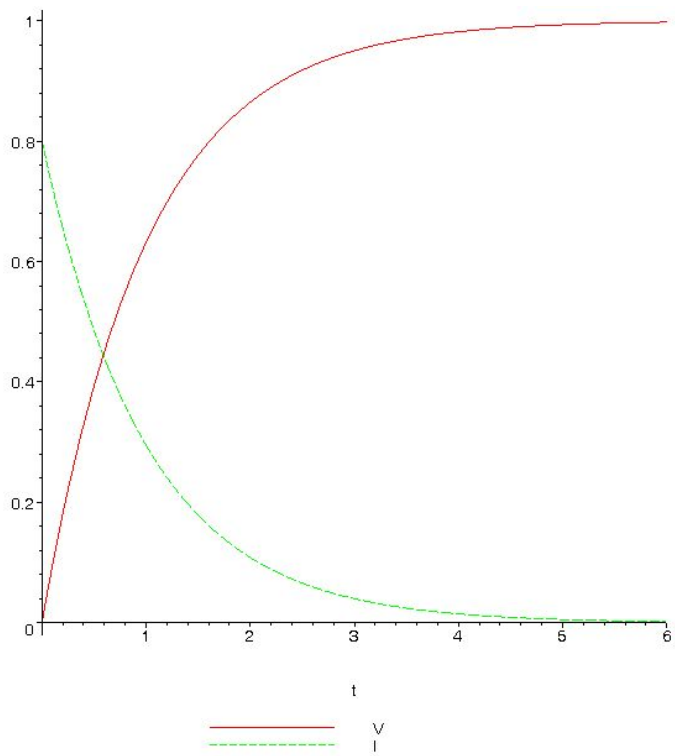


Figure 34: Time dependence of the voltage and current across the capacitor for the RC circuit in arbitrary units. It illustrates the general behavior of these quantities.

7. Lecture 7

7.1 Capacitor charge and discharge

In the same way we can charge a capacitor through a resistor we can also discharge it. In the circuit of fig.35 we have a switch with two positions. In one position the capacitor charges through resistor R_1 and in the other it discharges through resistor R_2. The characteristic times are given by $\Delta t_{\text{charge}} = R_1 C$ and $\Delta t_{\text{charge}} = R_2 C$. Notice that these are properties of the circuit. The capacitor itself has no characteristic time associated with it, it depends on both the capacity C and the resistance R. A more practical circuit substitutes the switch by an integrated circuit that does the same job. At the end of this lecture I put a description of the circuit we used in class and how you can build it yourself if you are interested.

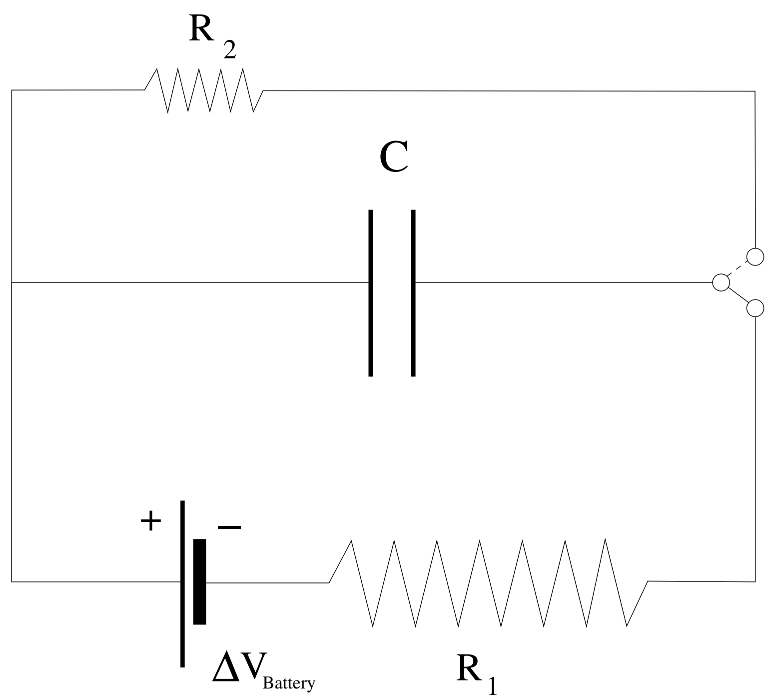


Figure 35: By flipping the switch we can charge and discharge the capacitor. The characteristic times are given by $\Delta t_{\text{charge}} = R_1 C$ and $\Delta t_{\text{charge}} = R_2 C$.

7.2 DC circuits

7.2.1 Resistors in series and parallel

More complicated circuits include many components. Let us see now what happens if you connect several resistors in series. In that case the current going through all of them is the same. On the other hand the potential difference across the system is the sum of the potential differences across each resistors. For example if we look at fig. 36 we see that the potential difference

$$\Delta V = V_a - V_d = (V_a - V_b) + (V_b - V_c) + (V_c - V_d) = \Delta V_1 + \Delta V_2 + \Delta V_3. \tag{7.1}$$

Using Ohm's law we find

$$\Delta V_1 = IR_1, \quad \Delta V_2 = IR_2, \quad \Delta V_3 = IR_3, \tag{7.2}$$

from where we obtain

$$\Delta V = IR_1 + IR_2 + IR_3 = I(R_1 + R_2 + R_3) = IR \tag{7.3}$$

So the total resistance is $R = R_1 + R_2 + R_3$ a simple rule to remember!. Notice that I, the current is the same for all resistors since charge is conserved therefore the same amount of charge has to circulate per unit time in each point of the circuit. This is not true if the circuit branches as we see in fig.37 a connection known as parallel. In this case the potential difference across all of resistors is the same as is easily seen if you remember that all points connected by a cable are at the same potential. The total current however is split between the three branches as:

$$I = I_1 + I_2 + I_3 \tag{7.4}$$

Again, this is because charge is conserved so the total charge entering the circuit is distributed among the three resistors. Again using Ohm's law we have

$$I = \frac{\Delta V}{R_1} + \frac{\Delta V}{R_2} + \frac{\Delta V}{R_3} = \Delta V \left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} \right) = \Delta V \frac{1}{R} \tag{7.5}$$

So we find

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} \tag{7.6}$$

another easy rule to remember. However it requires more computations than just adding the resistance as when they are in series. By the way, if you have the time an interest it is easy and relatively cheap to buy a multimeter and a few resistors and check these by yourself.

It is useful to remember the case where all resistors are equal $R_1 = R_2 = R_3$ where we get

$$R = \frac{R_1}{3} \tag{7.7}$$

This is the same if we out any number of resistors, for example with two resistors in parallel the resistance is half etc.

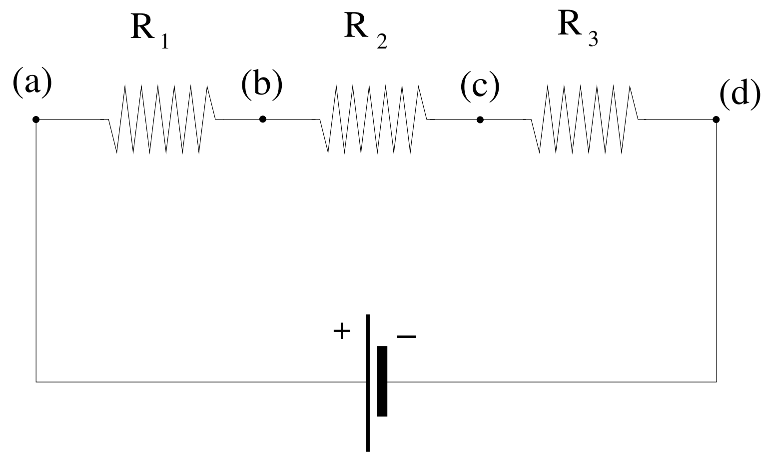


Figure 36: Resistors in series. The total resistance is the sum of the individual resistances.

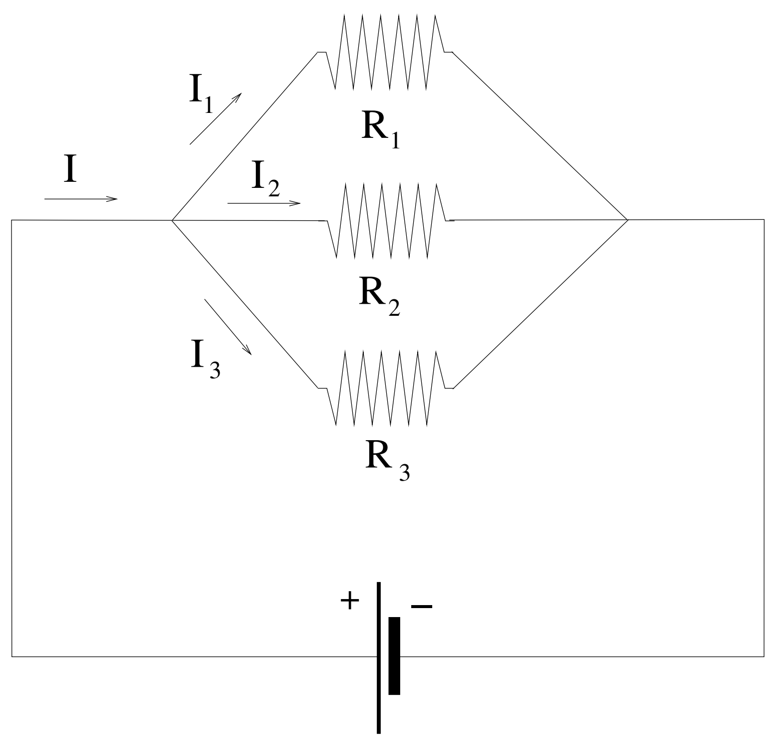


Figure 37: Resistors in parallel. The inverse of the total resistance is the sum of the inverse of the individual resistances.

7.2.2 Kirchhoff's laws

More complicated circuits can be treated similarly. Many times one can split the circuits in resistors in series and parallel. For example consider the situation in fig.38 that is self-explanatory. Another example is fig.39. Notice that we get the same total resistance. The reason is that because of the values of the resistors the points (a) and (b) have the same potential already in fig.38 so no current actually circulates through the cable in the middle. However if one resistor changes slightly its value then a current will circulate through the cable in the middle which can be used to detect those small variations.

Although this can take you a long way analyzing circuits, some circuits just cannot be analyzed in this way. Consider fig.40. In this case one cannot use the idea of resistors in parallel or series. However we can use similar rules as those used to derive what happens for resistors in series and parallel. This can be cast in a set of rules which are known as Kirchhoff's laws and read as follows:

- The total potential difference between two points connected by a path is the sum of the potential differences between the circuit components in the path. Therefore for any close loop the sum of the potential difference of its components have to add up to zero.

- Charge is conserved. Therefore in each node or connections where cables come together the total current coming in has to equal the total current coming out.

These laws allow us to write as many equations as variables we need to determine. Care should be taken to write equations which are independent namely not equivalent to each other. If the component is a battery then the resistance across it is fixed. If it is a resistor it is given by Ohm's law.

Going back to the example of the figure we can write the following equations:

$$I = I_1 + I_3 \tag{7.8}$$

$$I_1 = I_2 + I_5 \tag{7.9}$$

$$I_3 + I_5 = I_4 \tag{7.10}$$

$$I_1 R_1 + I_5 R_5 + \Delta V_{b2} - I_3 R_3 = 0 \tag{7.11}$$

$$I_2 R_2 - I_4 R_4 - \Delta V_{b2} - I_5 R_5 = 0 \tag{7.12}$$

$$I_1 R_1 + I_2 R_2 - \Delta V_{b1} = 0 \tag{7.13}$$

The first three are obtained by looking at the currents in each node. The last three by looking at the loops shown in the figure. We imagine that we go down in potential so we get for example for the loop colored pink, $I_1 R_1$ when we go through R_1 but $-I_3 R_3$ when we go through R_3 because we go against the current so the potential increases. You can imagine you are in a hill and go up and down across each resistor or battery and have to come back at the same height when you return to the same point. Be sure you understand exactly how each equation is derived. These equations are not difficult to solve. One uses each equation to eliminate one variable and replace it in the other equations. At the end we have only one equation and one variable and we can solve it. Although not difficult it is tedious, instead one can use a computer algebra program to find the solution (for example Maple, Mathematica, Matlab etc). For example, in this case the current I is:

$$I = \frac{((R_1 + R_3)(R_2 + R_4) + (R_1 + R_2 + R_3 + R_4)R_5)\Delta V_{b1} + (R_1 R_4 - R_3 R_2)\Delta V_{b2}}{R_5(r_1 + R_2)(R_3 + R_4) + R_1 R_2 R_4 + R_3 R_2 R_4 + R_3 R_2 R_1 + R_1 R_3 R_4} \tag{7.14}$$

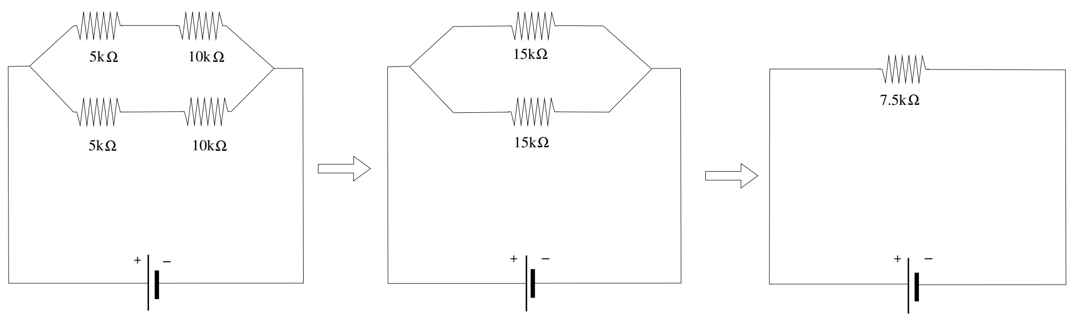


Figure 38: Resistors in series and parallel.

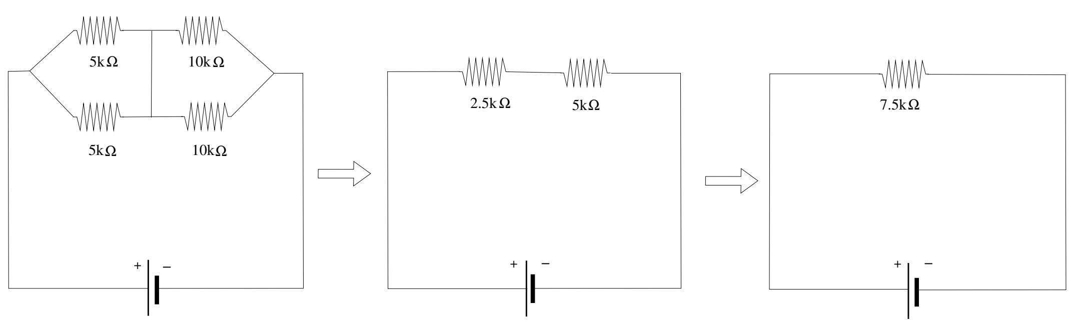


Figure 39: Resistors in series and parallel.

7.3 Example of circuit using charge and discharge of a capacitor

Here we describe how to build a simple circuit that charges and discharges a capacitor. The time constants depend on variable resistors that we can control by hand and also on the value of the capacity. Building the circuit is beyond the contents of the course but I'll briefly describe how to do it because it is easy to do and you can have fun with it if you decide to do it. It can also open the possibility to do many other small projects you can find in the Internet, books etc.

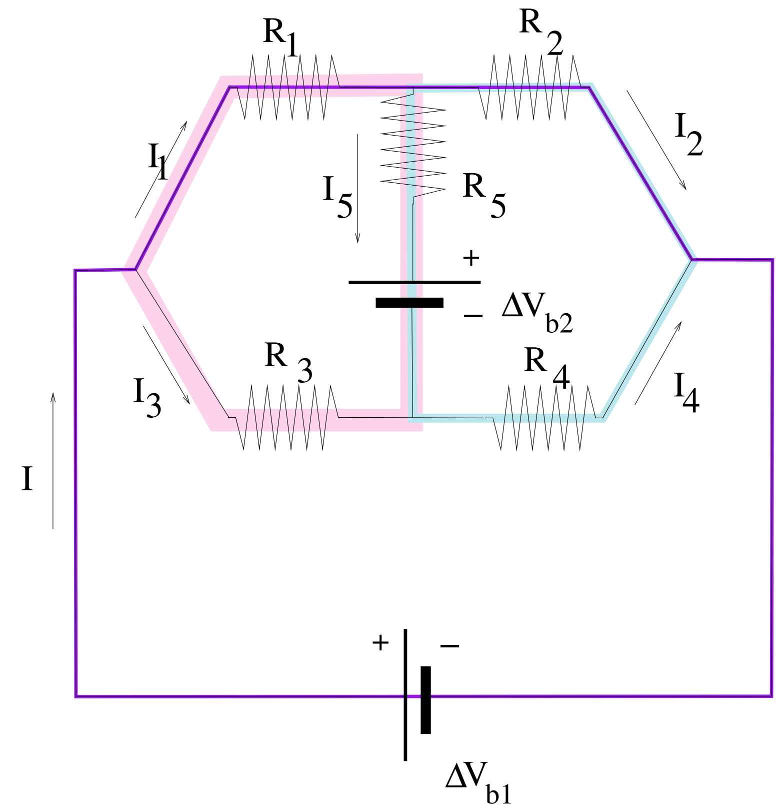


Figure 40: More complicated circuit. The closed loops used to write the equations are shown in color.

7.3.1 Oscillator using the 555 chip

The circuit is diagrammed in fig. 41. The 555 chip senses the voltage in one of its legs and when it goes above 2/3 of the supply voltage (here 5v) discharges the capacitor. When the voltage goes below 1/3 of 5 volts it charges it back again. The timing is regulated by the value of the resistors R_1 and R_2. In our case we put two variable resistors of maximum value $10k\Omega$.

7.3.2 Actual construction

To build the circuit we use a breadboard. All holes in each column are connected to each other except that the top and bottom parts are separate. The top and lower rails area also single conductors that usually are connected to positive and negative (ground) of the power supply. The supply voltage we need is 5V but the battery is 9V. So we use a regulator 78L05 which takes the 9V and supplies 5V. The legs of the 555 chip are labeled counterclockwise as in fig.41. Then it is a matter of patience and putting all components on the breadboard to have the circuit built, tested and working. One can

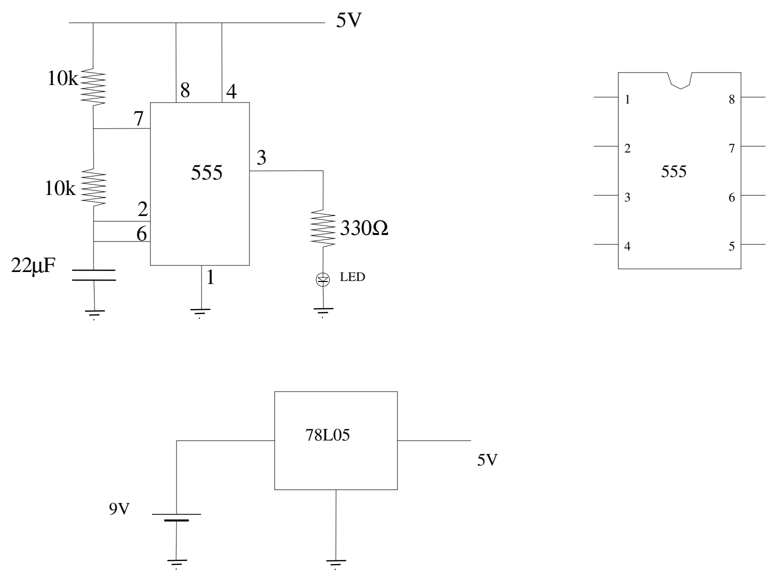


Figure 41: Oscillator circuit using the 555 chip. The legs of the 555 are numbered counter-clockwise. The 78L05 is used to get 5V supply from a 9V standard battery.

use the multimeter to check the voltage across the capacitor and see how it increases and decreases. This timing circuit can be used for many applications such and beepers, timing circuit for a microcomputer or for example for a joystick. A computer can be connected where the beeper is and by measuring the period of oscillations determine the value of the resistor. If the variable resistor is connected to the stick of a joystick, then the computer can sense its position. For more information on the 555 you can look at the specifications with circuit examples http://www.national.com/ds/LM/LM555.pdf and for the 78L05: http://www.national.com/ds/LM/LM78L05.pdf

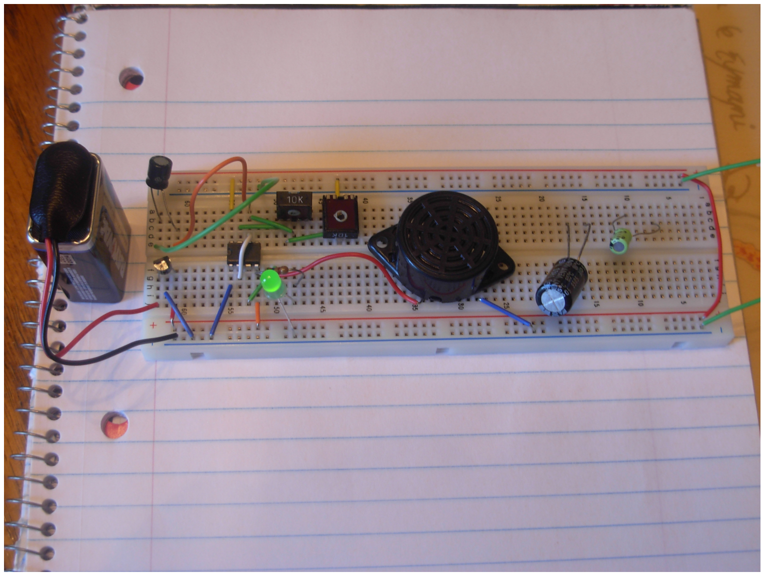


Figure 42: Oscillator circuit using the 555 chip. The two capacitors on the right are not part of the circuit, they are there to replace the working capacitor and check how the period changes.

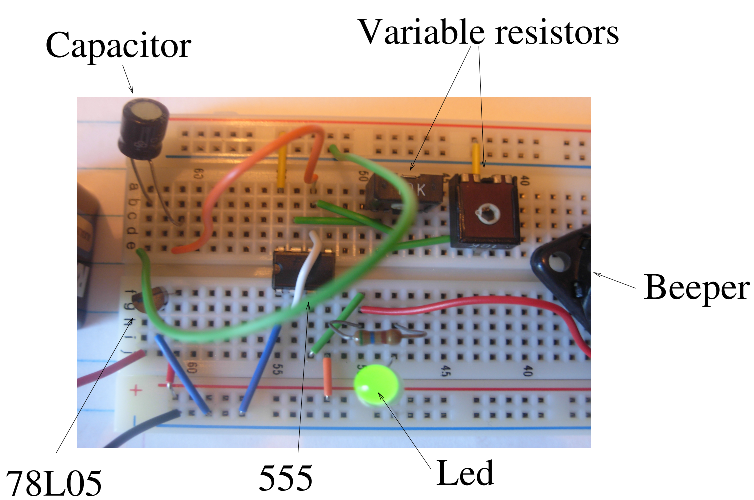


Figure 43: Oscillator circuit using the 555 chip.

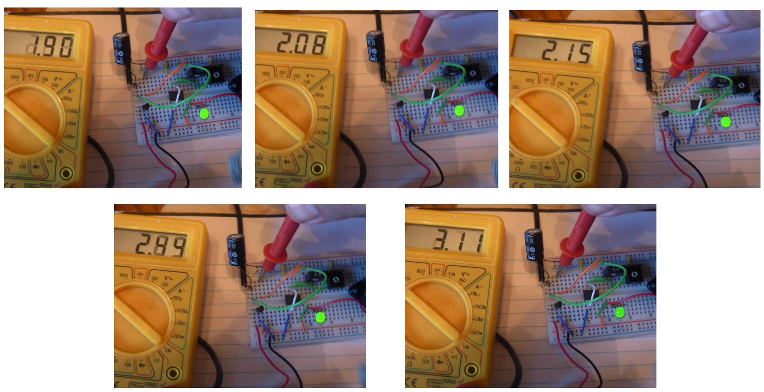


Figure 44: With a multimeter one can easily check that the capacitor charges and discharges.

8. Lecture 8

8.1 Capacitors in series and parallel

In the same way that one can analyze resistors in series and parallel one can understand what happens for capacitors in series and parallel. Notice that whereas Ohm's law is $\Delta V = IR$ we have $Q = C\Delta V$. So the potential difference or voltage ΔV is proportional to R but inversely proportional to C. As a consequence, and working with the circuits in fig.45 one finds that

Series: $\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}$

Parallel: $C = C_1 + C_2 + C_3$

Try to derive these relations as an exercise. As a hint notice that when the capacitors are in series, the conductors in the middle are neutral so the charge in each capacitor is the same (see fig.45).

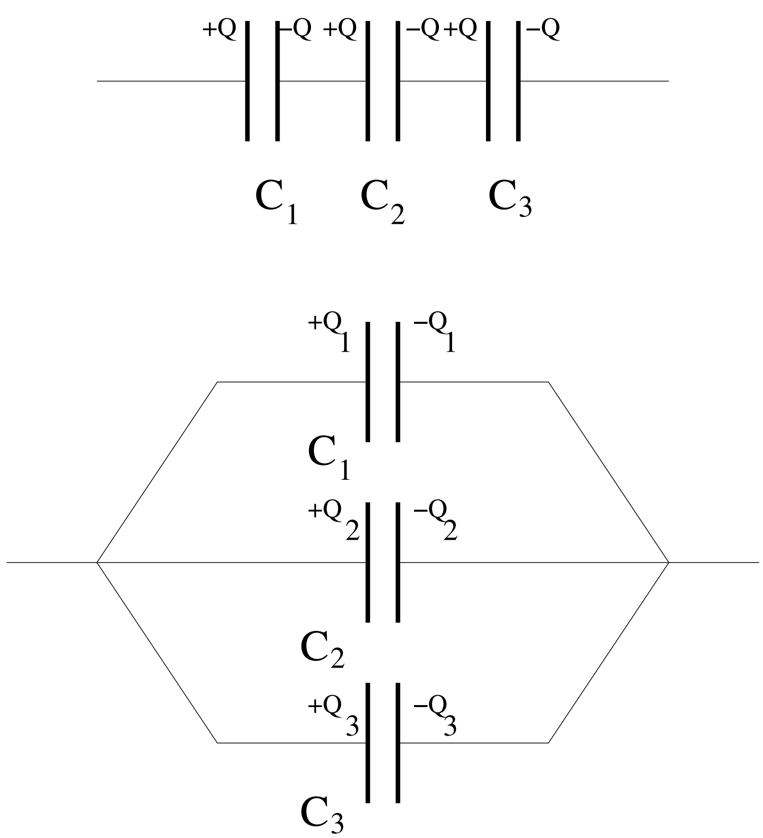


Figure 45: Capacitors in series and parallel. In parallel they add and in series their inverse add. Notice the difference with the case of resistors.

8.2 Magnetic field

Magnetism is a well-known phenomenon, for example the magnetic field of the Earth has provided orientation through the use of the compass for thousands of years. Magnetic fields can be generated by permanent magnets, electric currents and time dependent electric fields. In the case of the magnet the magnetic field looks like the one in fig.46 (see also the demo in fig. 48). Outside the magnet it is very similar to the one of a dipole electric field. But inside it is not. The lines of magnetic field close, they do not start or terminate at any point. Without knowing that, one might, at first sight, think that cutting in half a magnet will separate two opposite magnetic charges. Traditionally they are called the North and South pole. A pure North pole would have a magnetic field as the one in fig.47 However such object has never been observed in Nature. In particle physics there are theoretical indications that heavy particles with the properties of magnetic monopoles might exists. For that reason it is still an open question if monopoles (as they are called) exist or not. In any case if they do exist they are not common objects and we are not going to consider them here. The question naturally arises of what happens if we keep dividing the magnet, do we find an "elementary magnet"?. The answer is yes, in fact the elementary magnet is our old friend the electron. Electrons not only have charge but they also behave as a tiny magnet. The superposition of the magnetic fields of all electrons creates the field of the magnet. Proton and neutrons are also tiny magnets but their strength is 2000 times smaller than the one of the electrons.

An important consequence of the fact that there are no magnetic monopoles is that the flux of the magnetic field through a closed surface is always zero. This is because a theorem similar to Gauss' theorem applies. The total flux is proportional to the total magnetic charge enclosed but since there is no magnetic charge, the flux is always zero, namely for each closed surface the same amount of magnetic flux comes in as it goes out.

The magnetic field is a vector usually denoted as $\vec{B}$. It can be detected by the effect it causes on charges. When a charge moves in a magnetic field it experiences a force equal to

$$\vec{F} = q\vec{E} + q\vec{v} \times \vec{B} \tag{8.1}$$

where $\vec{v}$ is the velocity, q the charge and $\vec{B}$ the magnetic field. We included also an electric field $\vec{E}$ for completeness. We need to explain what the operation $\vec{v} \times \vec{B}$ is. It is known as a vector product and given two vectors it gives another vector:

$$\vec{v} \times \vec{B} = |\vec{v}||\vec{B}| \sin\theta \hat{n} \tag{8.2}$$

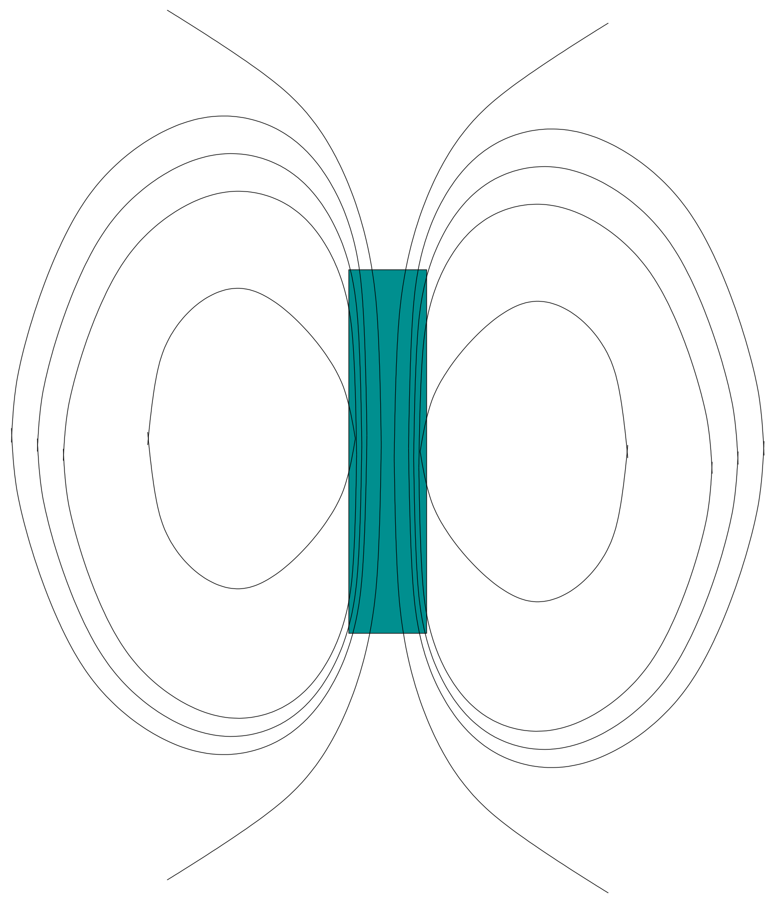


Figure 46: Magnetic field of a permanent magnet. Outside is similar to an electric dipole but inside there are no charges where the lines end. They lines of magnetic field are actually closed.

Its magnitude is the product of the modulus of $\vec{v}$, $\vec{B}$ and the sine of the angle between them that we denote as θ. Its orientation we denote by the unit vector $\hat{n}$ and is perpendicular to both $\vec{v}$ and $\vec{B}$. We write that as

$$\hat{n} \perp \vec{v}, \quad \hat{n} \perp \vec{B} \tag{8.3}$$

Its orientation is given by the right-hand rule. If we move our fingers in a screw motion going from $\vec{v}$ to $\vec{B}$ then our thumb points in the direction of the force. This is illustrated in fig.50. Some important consequences are

- If the particle is at rest then $\vec{v} = 0$ and there is no magnetic force.
- If the velocity is parallel to $\vec{B}$ then $\theta = 0$ and also there is no force.
- The magnetic force is always perpendicular to the velocity so it never does work!.

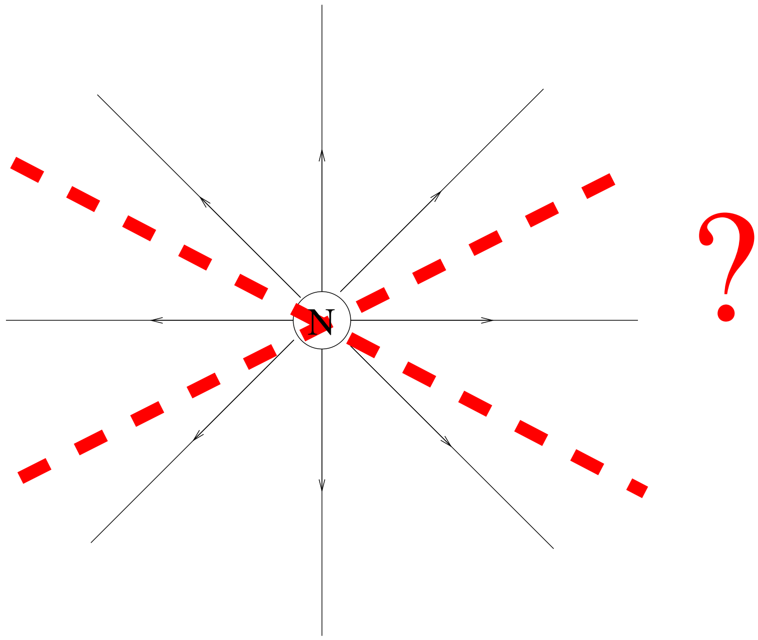


Figure 47: A conjectured magnetic monopole will have all lines of magnetic field coming out. It has not yet been observed in Nature but theoretical ideas suggest it might be an actual but exotic particle.

Figure 48: The lines of magnetic field can be made evident using iron filings which orient themselves parallel to the magnetic field.

Also form the form of the Lorentz force we see that the magnetic field is measured in units of $\frac{Ns}{Cm}$, such unit is known as a Tesla:

$$1T = 1\frac{Ns}{Cm} = 1\frac{Kg}{Cs} \tag{8.4}$$

where we used $1N = 1\frac{Kgm}{s^2}$. To have an idea the magnetic field of the Earth is around $10^{-5}T$ but in the lab fields of $10T$ can be produced.

Since the magnetic force is perpendicular to the velocity it will change its direction but not its magnitude. Therefore, if there are no other forces present, the kinetic energy

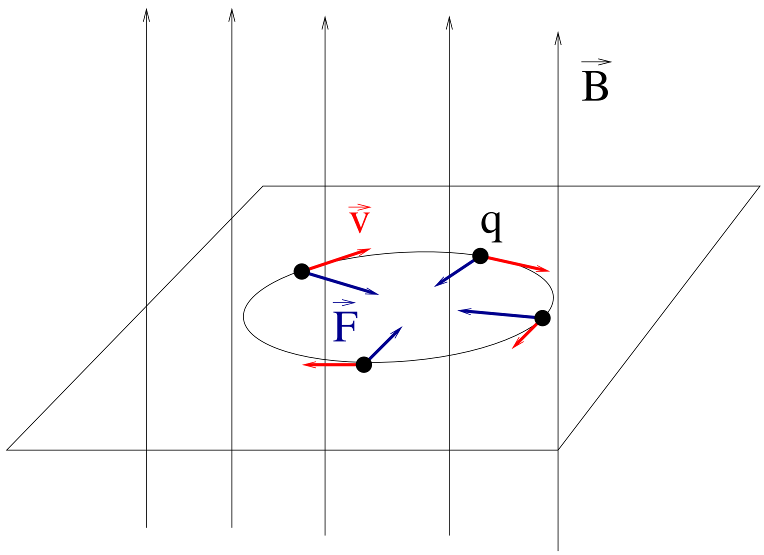


Figure 49: A particle moving in a magnetic field $\vec{B}$ experiences a force perpendicular to the velocity and to $\vec{B}$. Its orientation is given by the right-hand rule.

$E = \frac{1}{2}mv^2$ will remain constant, a consequence of the fact that the magnetic force does no work. Furthermore, if the magnetic field is constant the magnitude of the force will be constant and the trajectory of the particle will be a circle in a plane perpendicular to the magnetic field. Indeed, in fig.49 we see that if the particle moves in circles then the force points always toward the center keeping the particle in its trajectory in the same way that we can tie an object at the end of a rope and make it move in circles.

Using Newton's law

$$\vec{F} = m\vec{a} \tag{8.5}$$

we can compute the the period of motion in a similar manner that we discussed at the beginning of the course for the Moon orbiting the Earth. the acceleration is in the radial direction, namely it is centripetal acceleration whose magnitude is

$$|\vec{a}| = \frac{v^2}{r} \tag{8.6}$$

as we remember from the mechanics course. Since the force point in the same direction we only need to equate their magnitudes:

$$|\vec{F}| = qvB = m\frac{v^2}{r} \tag{8.7}$$

giving

$$v = \frac{qBr}{m} \tag{8.8}$$

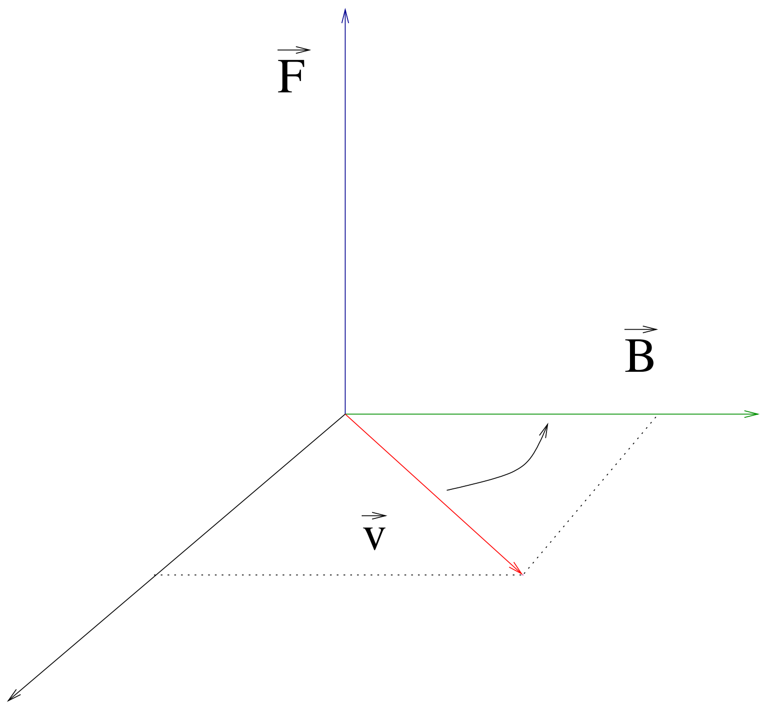


Figure 50: Example of the right-hand rule. The force is the vector product of the velocity and the magnetic field and has the orientation indicated here.

The period of motion is how long does it take for a particle to go around the circle. It is given by

$$\Delta t = \frac{2\pi r}{v} = 2\pi r \frac{m}{qBr} = \frac{2\pi m}{qB} \tag{8.9}$$

amazingly it is independent of the radius. Namely if the radius is large the particle moves faster and takes the same time to go around. Notice that a peculiar property of this motion is that the particle can move in circles only in one direction. If one tries to move the particle around the circle in opposite direction then the force will point outward, it will not stay in the trajectory. Actually will again describe a circle in the "correct" direction. On the other hand the motion in the other direction is precisely what happens for a particle of opposite charge.

Although rather simple this idea has had numerous applications. The most significant one perhaps is in particle accelerators where magnetic fields keep particles in circular trajectories while at the same time electric fields accelerate them at higher and higher energies. When these particles collide with fixed targets (or two opposite beams are brought into collision) new particles are created revealing the fundamental constituents of matter and their interactions.

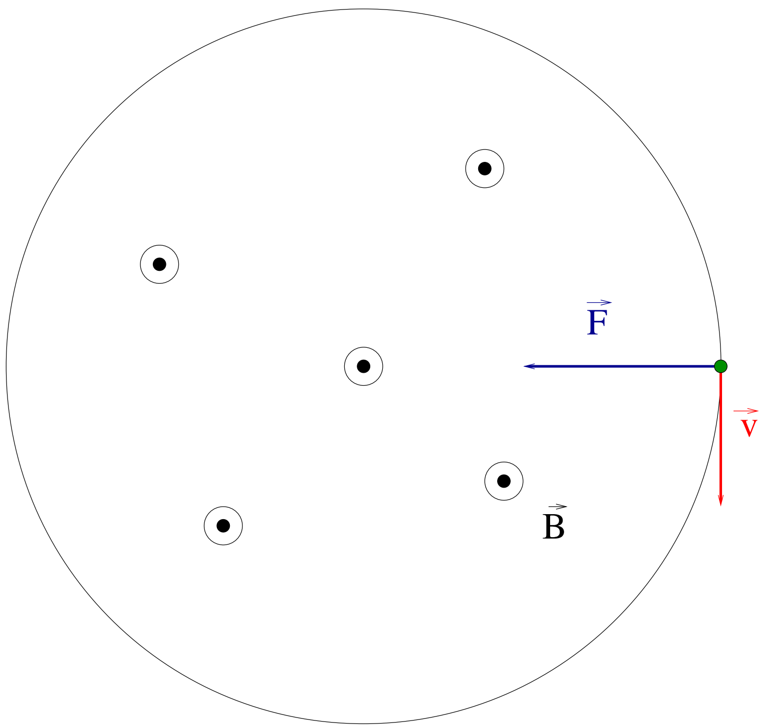


Figure 51: A free particle moving in a constant magnetic field $\vec{B}$ moves in circles since the magnetic force is oriented toward the center.

9. Lecture 9

9.1 Magnetic forces on an electric current

Since an electric current is charges in motion it follows that a wire through which a current circulates will experience a force in the presence of a magnetic field. In fact an interesting demo (fig.52) makes this effect evident. More precisely, suppose that, inside the conductor, an electric charge density (charge per unit volume) ρ is moving with velocity v. Then, during time Δt all the charge contained in a volume $v\Delta t A$ will pass through a cross section of area A as can be seen in fig.53. The electric current, namely the amount of charge going through a section of the conductor of area A per unit time is therefore

$$I = \frac{\Delta Q}{\Delta t} = \frac{\rho v \Delta t A}{\Delta t} = \rho v A \tag{9.1}$$

On the other hand the magnetic force on a moving charge is

$$|\vec{F}| = q|\vec{v}||\vec{B}|\sin\theta = \rho A L|\vec{v}||\vec{B}|\sin\theta = IL|\vec{B}|\sin\theta \tag{9.2}$$

where we used that the total charge is given by $q = \rho L A$ for a conductor of length L. Therefore the force is proportional to the current and the length, one can define a force per unit length as

$$\frac{|\vec{F}|}{L} = I|\vec{B}|\sin\theta \tag{9.3}$$

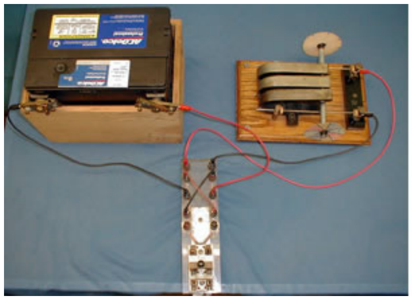
The direction and orientation of the force is given by exactly the same right-hand rule as before. Only that instead of the velocity we use the direction of the current.

A very interesting observation is that the magnetic field allows to determine the sign of the charge density ρ responsible for the current. This is illustrated in fig.54 and known as the Hall effect. The same current can be produced by positive carriers moving in one direction along a wire or by negative ones moving in the opposite direction. However, the magnetic force on the wire is the same in both cases and for that purpose we do not need to know the sign of the carriers. On the other hand it means that the carriers would like to accumulate on one side of the wire and therefore if that side becomes positively charged the carriers are positive and if it becomes negatively charged they are negative. The experiment shows that the carriers are negative as we know since they are electrons.

9.2 Magnetic field created by a current

9.2.1 Ampere's law

We mentioned that an electric current creates a magnetic field as can be seen with an electromagnet, a coil through which a current circulates behaves as a magnet. The



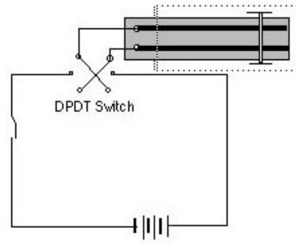


Figure 52: This interesting demo shows that a current feels a force in the presence of a magnetic field. By switching the direction of the current the direction of the force changes and the aluminum bar rolls one way and then the other.

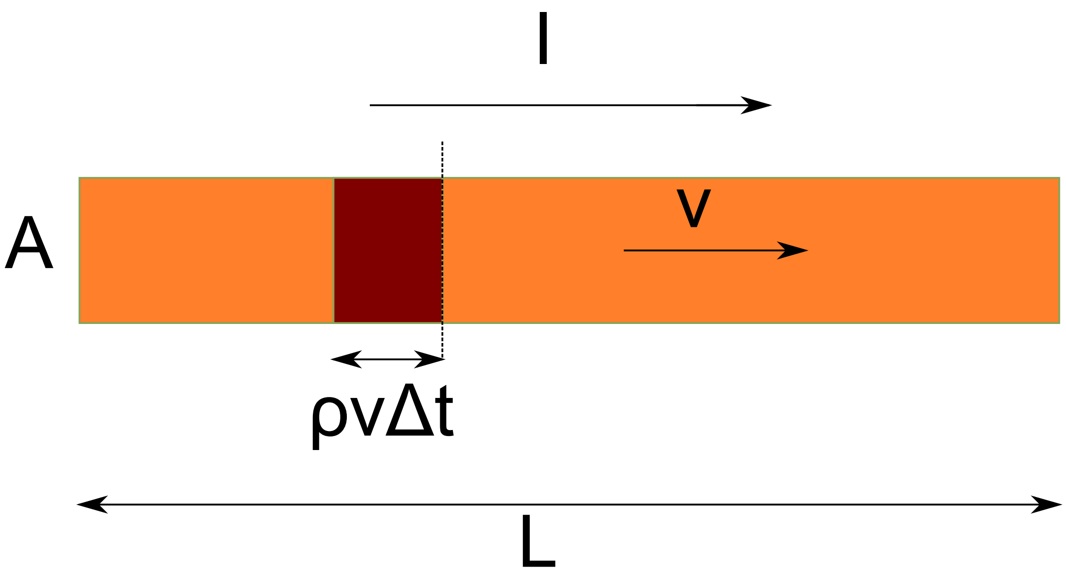


Figure 53: A current is a charge density ρ moving with velocity $\vec{v}$ through a conductor of cross section A and length L. The charge contained in a volume $|\vec{v}|\Delta t\, A$ goes through the cross section A during a time Δt. This gives $I = \rho|\vec{v}|A$.

intensity of the magnetic field is given by Ampere's law. We mentioned that the flux of the magnetic field through a surface is zero. However given a vector field there is another important quantity known as the circulation. Given a closed path one multiplies the component of the magnetic field in the direction of the path times the displacement and sums over all portions of the path:

$$\text{Circulation around a path} = \sum |\vec{B}|\Delta\ell \cos\theta \tag{9.4}$$

where $\Delta\ell$ is the displacement, θ the angle between the magnetic field and the direction of the path and the sum is over all portions in which we divide the path for convenience

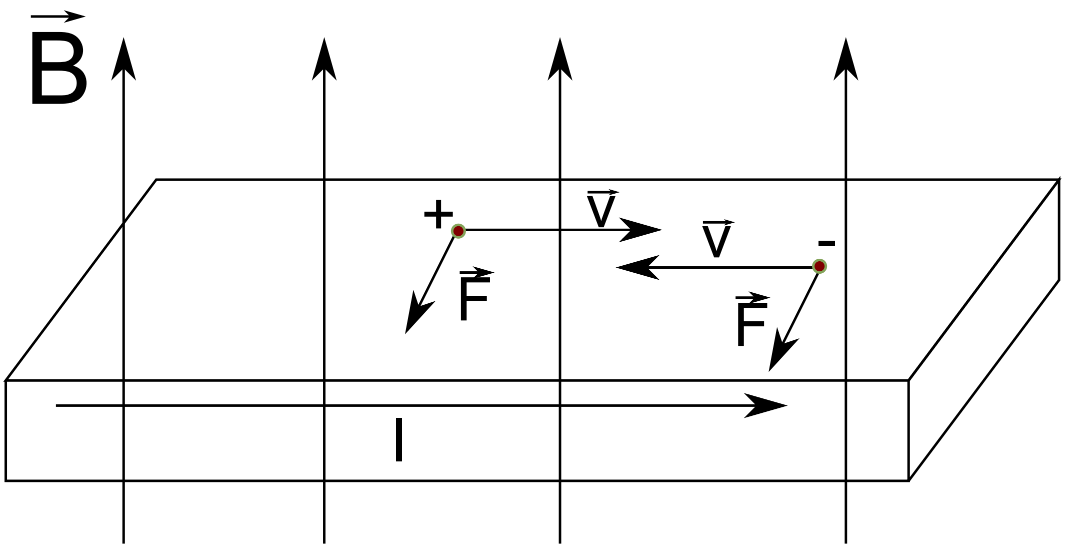


Figure 54: The same current is produced by positive charges moving to the right as by negative charges moving to the left. In a magnetic field both feel a force as indicated meaning that different charge accumulates on the sides of the conductor depending on the sign of the carriers. This is the Hall effect and allows to determine the sign of the particles responsible for the current.

of the calculation (see fig.55). It is exactly the same type of formula used to compute the work done by a force when you move and object from one place to another. Only that here we use the magnetic field and not the force in the computation (the magnetic force does no work anyway). What Ampere's law says is that the circulation around any closed path is proportional to the current that pierces through any surface whose boundary is the given path.

$$\sum |\vec{B}| \Delta \ell \cos\theta = \mu_0 I \tag{9.5}$$

where the constant μ_0 is

$$\mu_0 = 4\pi\, 10^{-7} \frac{Tm}{A} \tag{9.6}$$

9.2.2 Displacement current

Maxwell realized that Ampere's law contains an ambiguity and should actually be amended to be true in all cases. Indeed, the circulation of the current around a closed path is equal to the current piercing a surface whose boundary is the loop. However there are many such surfaces. In fig. 56 we see two surfaces ending in the circular path. One (S_1) is just a disk whereas S_2 is a dome-shaped surface. By charge conservation, if the current going through S_1 is not the same as the one going through S_2, charge has to accumulate in the volume between the two surfaces. For example, if we put

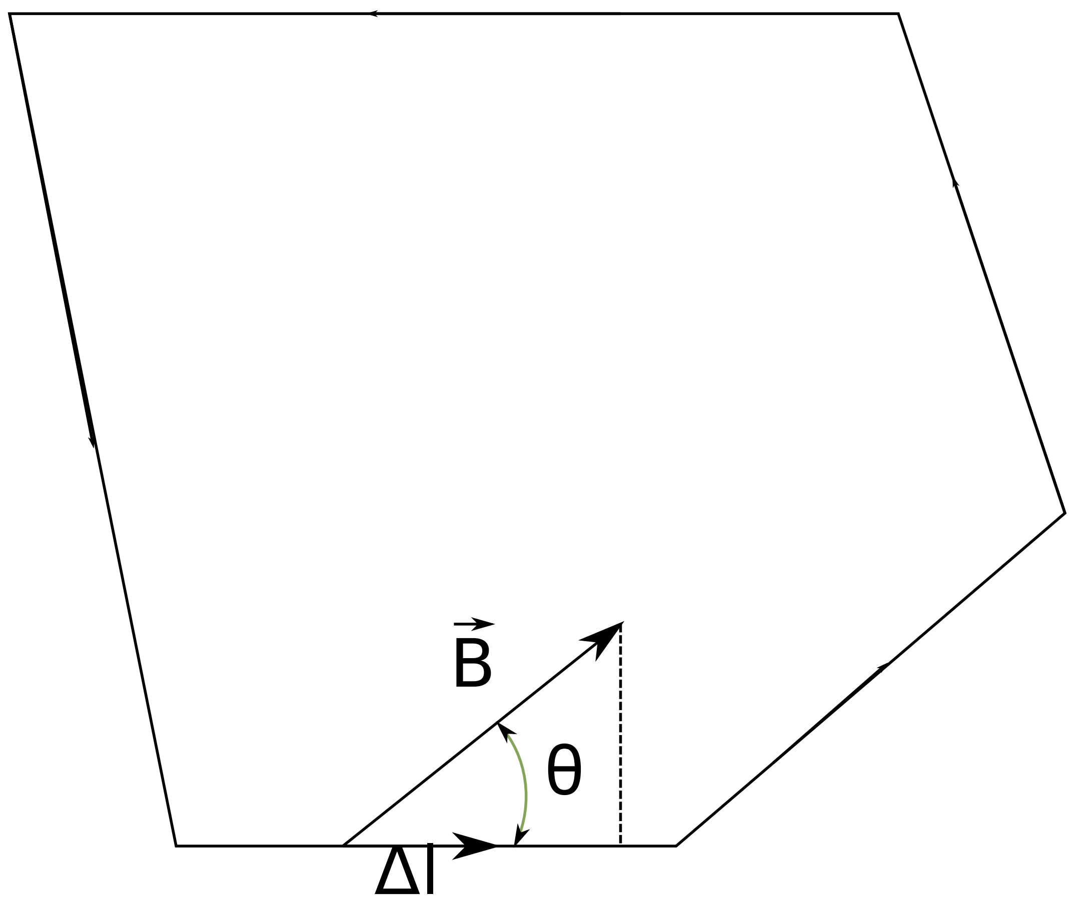


Figure 55: The circulation of the magnetic field is defined in a similar way as the work of a force along a path. In fact the circulation along a path can be defined for any vector field.

a capacitor as in the figure, we can charge the capacitor and no current would go through S_2 whereas, at the same time, current I is going through S_1. This makes Ampere's law unclear since which surface should we use?. Here Maxwell realized that a time dependent electric field would be piercing surface S_2 since the capacitor is being charged and an increasing electric field exists between the two plates of the capacitor. So he proposed that the circulation of the electric field should be equal to the current plus the variation of the electric flux through the surface. He also realized that this had profound consequences, a varying electric field creates a magnetic field. We are going to see that a varying magnetic field creates an electric field and so on. This process gives rises to a wave that propagates in space. In this way Maxwell predicted the existence of radio waves and later also suggested that light could be such an electromagnetic wave.

It is important to study this reasoning in detail. It illustrates perfectly the way that theoretical physics works. Starting from Ampere's law that had been verified experimentally, Maxwell realized an ambiguity, corrected it therefore predicting the existence of a new phenomenon, electromagnetic waves, which was later verified by Hertz. It is very instructive and a great scientific achievement.

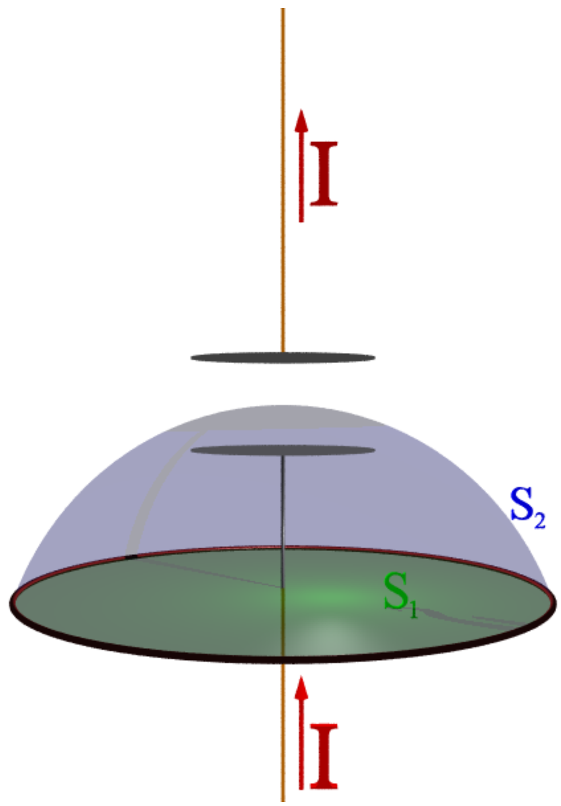


Figure 56: The current through two surfaces, S_1 and S_2 is different but through S_2 there is a time varying electric flux.

9.2.3 Magnetic field of a wire and a solenoid (coil)

We can use Ampere's law to compute the magnetic field surrounding a cable. For a straight cable the magnetic field goes around as indicated in fig.57. Using Ampere's law for a circular path around the cable we get

$$\text{circulation} = 2\pi r B = \mu_0 I \tag{9.7}$$

We then have

$$|\vec{B}| = \frac{\mu_0 I}{2\pi r} \tag{9.8}$$

We can create a stronger magnetic field by putting several cables such that their magnetic fields add up. In fact we do not need to put different wires, we can take one wire and form a coil as in fig.58. In that case the magnetic field is non-zero inside and outside is very small (if the coil is very long). Using a path as in the figure we get from Ampere's law:

$$BL = \mu_0 I N \tag{9.9}$$

where N is the number of turns of the coil and L is its length. Therefore, the magnetic field inside the coil is

$$\vec{B} = \mu_0 I \frac{N}{L} \hat{z} \tag{9.10}$$

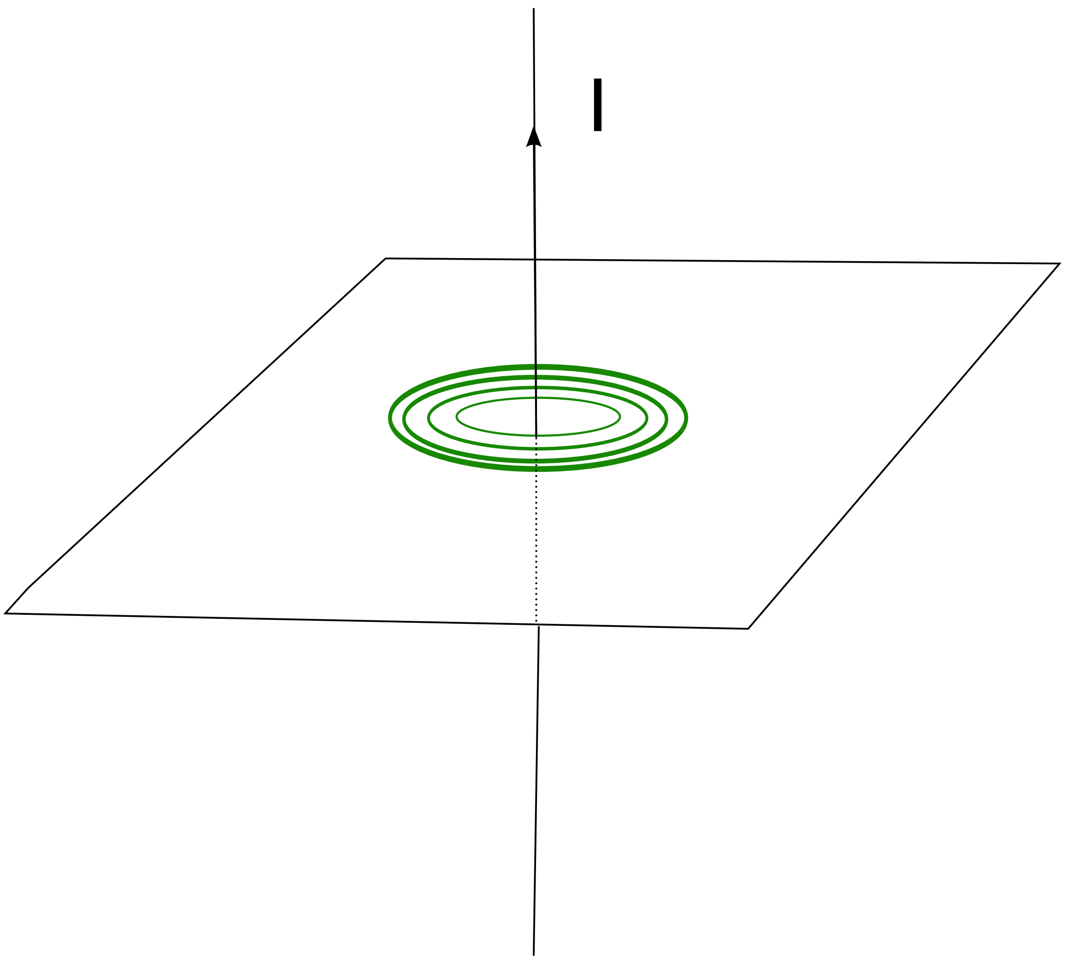


Figure 57: Magnetic field produced by straight wire.

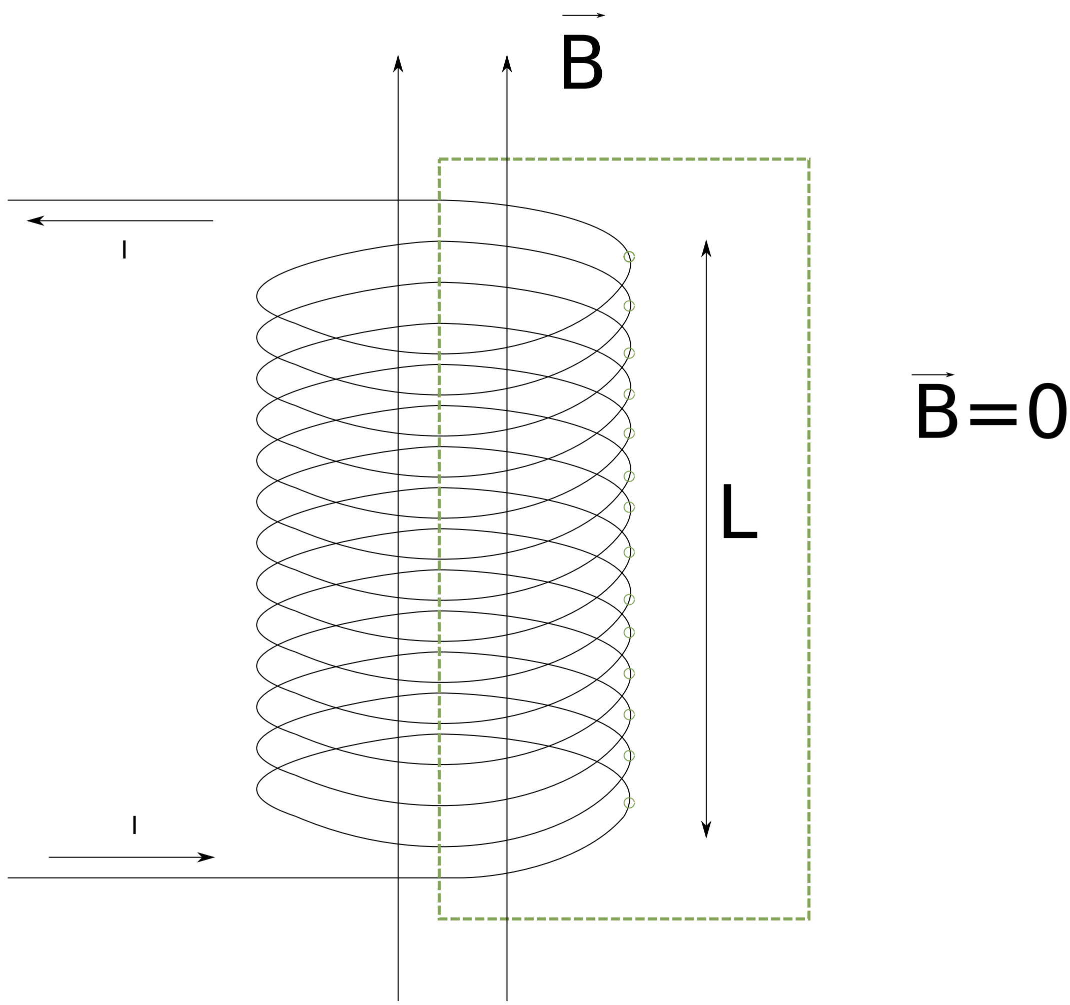


Figure 58: Using Ampere's law to find the magnetic field inside a coil.

10. Lecture 10

10.1 Force between currents

We saw that a current circulating along a wire creates a magnetic field. If another current is in its vicinity it experiences a force. In fig.59 we see a simple configuration of two parallel wires carrying currents in the same direction. As discussed in the previous lecture, the magnetic field produced by wire 1 has the direction indicated in the figure and its intensity is

$$|\vec{B}| = \frac{\mu_0}{2\pi r} I_1 \tag{10.1}$$

According to the right-hand rule, the force experienced by the second wire is directed toward the first one (attractive force) and is

$$|\vec{F}| = I_2 L |\vec{B}| = \frac{\mu_0}{2\pi} \frac{L}{r} I_1 I_2 \tag{10.2}$$

where r is the distance between the cables, L their length and $I_{1,2}$ the respective currents. Suppose they are separated by 1cm, they are $1m$ long and $I_1 = I_2 = 10A$. Using that $\frac{\mu_0}{2\pi} = 2 \times 10^{-7} \frac{mT}{A}$ we find

$$|\vec{F}| = 2 \times 10^{-7} \frac{mT}{A} \frac{1m}{1cm} 10^2 A^2 = 2 \times 10^{-3} mTA = 2 \times 10^{-3} N \tag{10.3}$$

A very small but measurable force.

10.2 Magnetic induction

If we move a cable through a magnetic field, the charges in the cable will feel a force and then, if the cable forms a closed loop, a current will circulate. Consider the example of fig.60. The force on a charge inside the cable is

$$|\vec{F}| = qvB \tag{10.4}$$

Since the same force would be produced by an electric field of magnitude $|\vec{E}| = vB$ such force can move charges against a potential difference $\Delta V = vBL$ where L is the length of the cable. Namely if we have a resistor R in the rest of the circuit a current $I = \frac{\Delta V}{R}$ will circulate. This is in fact an electric generator!. Not very practical but conceptually simple. In reality, it is better to move the cable in a circular motion so that it comes back to the original position and can keep moving. Notice that from the direction of the induced current we see that the moving cable acts as an effective battery whose positive terminal is in the lower cable.

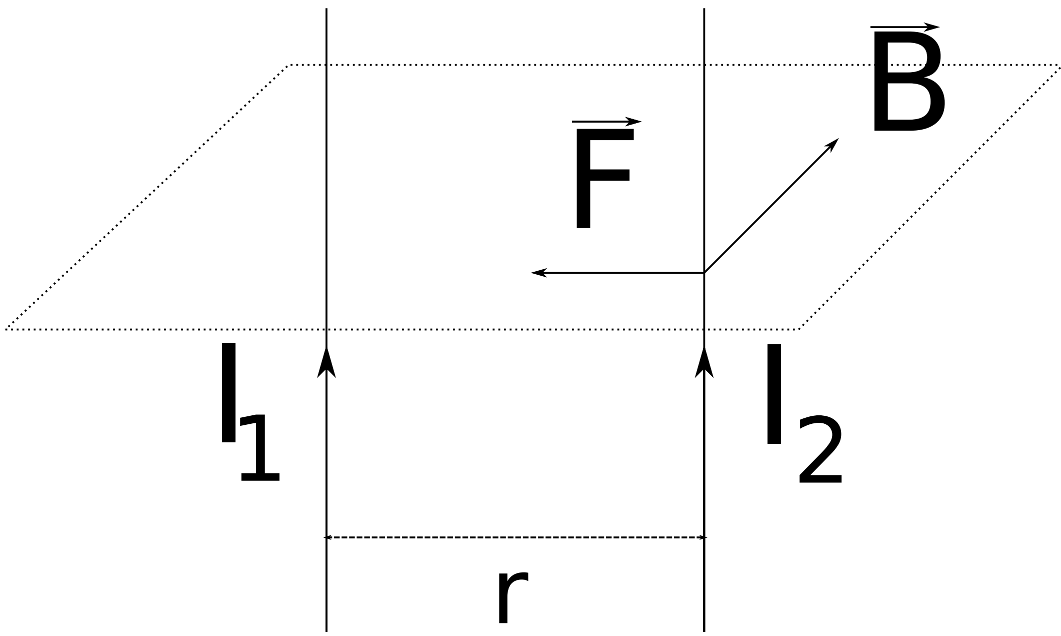


Figure 59: Two parallel wires attract if current circulate in the same direction. They repel is the current circulates in opposite direction.

A deeper insight into the problem can be obtained if we compute the magnetic flux through the loop. Since the magnetic field is constant and perpendicular to the loop, the flux is simply:

$$\text{Flux} = \Phi = |\vec{B}|Ld \tag{10.5}$$

where L is the length of the moving conductor and d is its distance to the resistor. The distance d changes in time such that

$$\frac{\Delta d}{\Delta t} = v \tag{10.6}$$

where v is the velocity of the moving wire. Since the loop becomes bigger, namely d is growing we see that the flux changes as

$$\frac{\Delta \Phi}{\Delta t} = |\vec{B}|Lv \tag{10.7}$$

which is nothing else but the emf or voltage between the end points of the moving cable. Moreover, we see that the current generated in the cable will itself create a magnetic field in the loop which is opposite to the one already present, namely it creates an extra field that opposes the increase in flux. This phenomenon can be exemplified by a demo (see fig.61 and fig.62) where a loop of wire is moved in the proximity of a magnet and a current is detected by an Ammeter. The same experiment shows that if we move the magnet instead of the loop of wire then the effect is the same. Although this appears more or less evident it is not clear in the latter case which force is moving the charges in the loop. Thinking about this problem was one of the reasons Einstein came up with the theory of relativity that we are going to discuss later in the course.

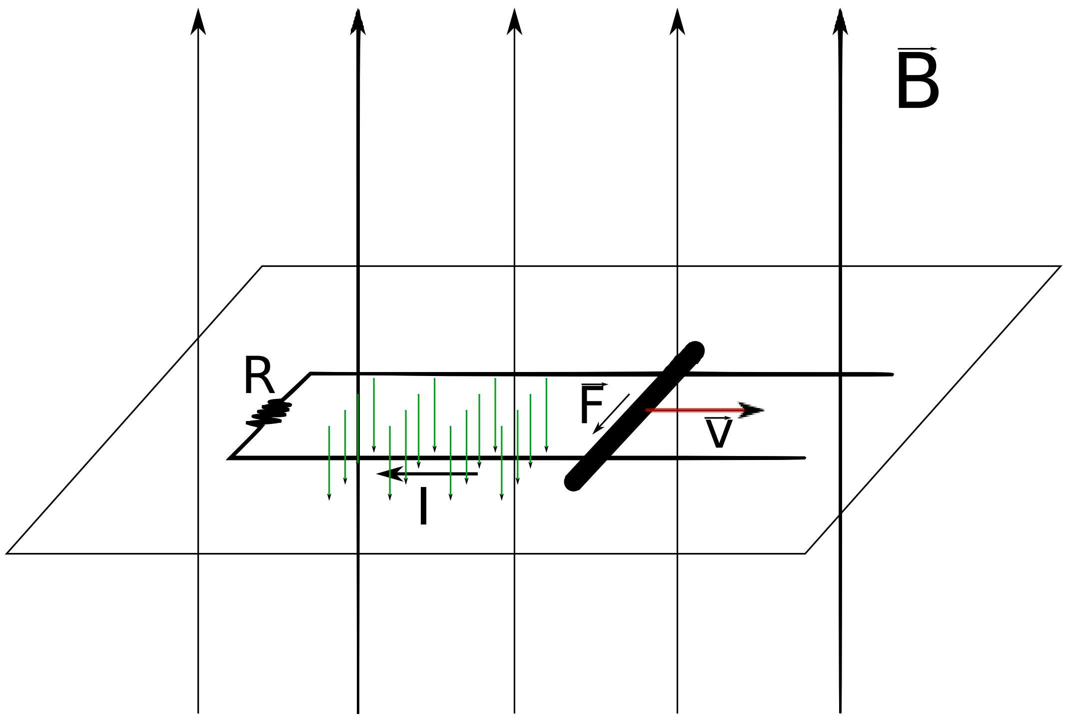


Figure 60: When a conductor moves in a magnetic field, the charges inside it feel a force that induces a current. The resulting emf can be computed from the Lorentz force or from the Faraday law with identical results. Notice that the induced current also creates a magnetic field, indicated with the green arrows, which opposes the change in flux.

For the moment we can summarize these finding in two important laws, the Faraday and Lenz law.

Faraday's law says that if the magnetic flux through a loop changes in time then an electromotive force is generated proportional to the rate of variation of the flux:

$$|\text{emf}| = \left|\frac{\Delta\,\Phi}{\Delta t}\right| \tag{10.8}$$

where the Greek letter Φ is used to denote flux. It is completely equivalent as if we put a battery with $\Delta V = \text{emf}$. It is important to know also in which direction the force goes or equivalently, in such effective battery which is the positive and which the negative pole. This is given by Lenz law. It says that the current generated in the loop creates a magnetic field that opposes the change in flux. Namely if the flux increases, the magnetic field created by the current is opposite to the magnetic field already present and vice versa, if the flux decreases, the current tend to reinforce the magnetic field. This phenomenon is known as magnetic induction, a magnetic field induces a current.

To exemplify this issues further we can use two interesting demos (see fig.63 and 64). In the first the current induced in a ring creates a repulsion that shoots the ring into the air. In the second the same repulsion is used to stop a pendulum showing how magnetic braking works.

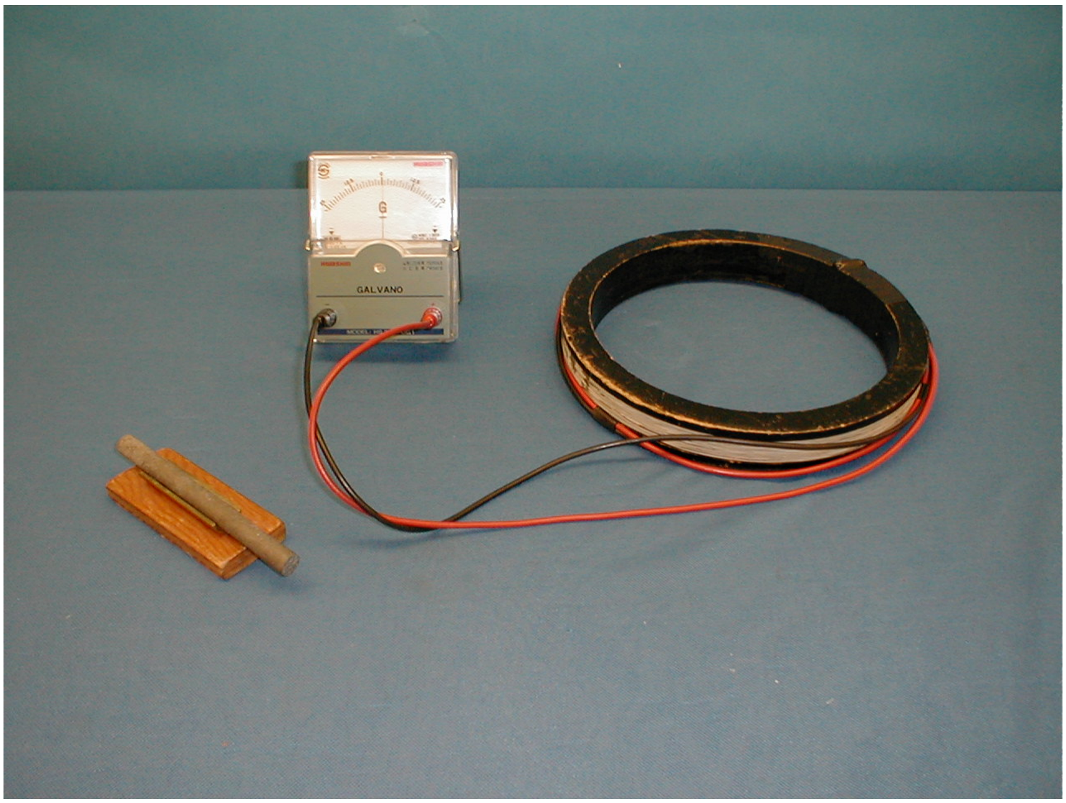


Figure 61: Demo. A current is induced in a loop if we move it in the presence of a magnetic field. See also fig. 62.

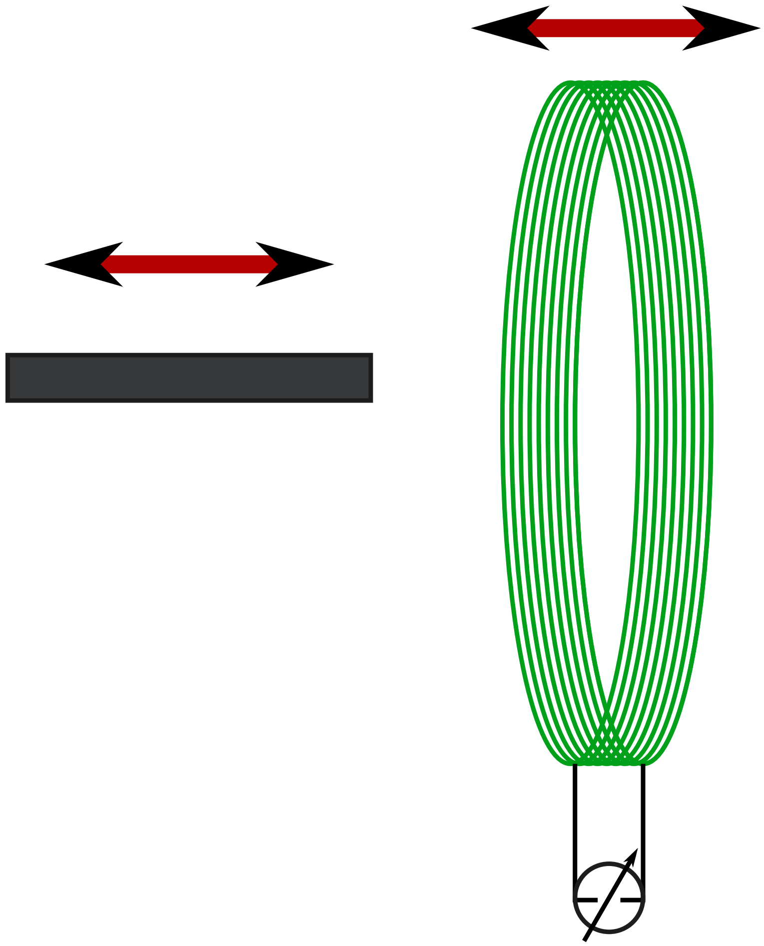


Figure 62: A current is induced in the loop if we either move the magnet or the loop. In both cases the magnetic flux through the loop changes.

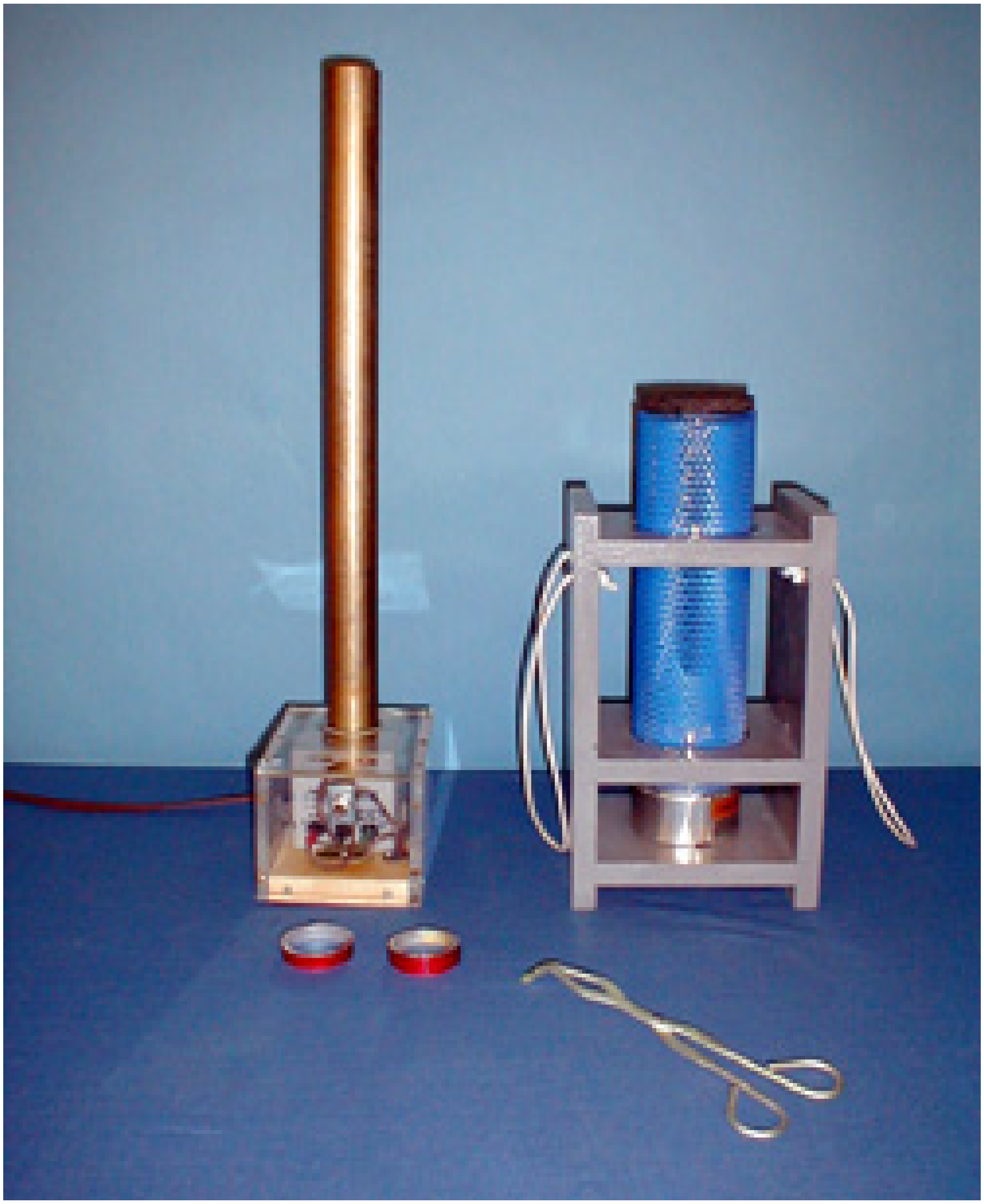


Figure 63: Demo. Starting up the current through the coil induces a current on a ring producing a repulsive force and shooting the ring in the air. Cooling the ring decreases the resistance therefore increasing the induced current and the height reached by the ring.

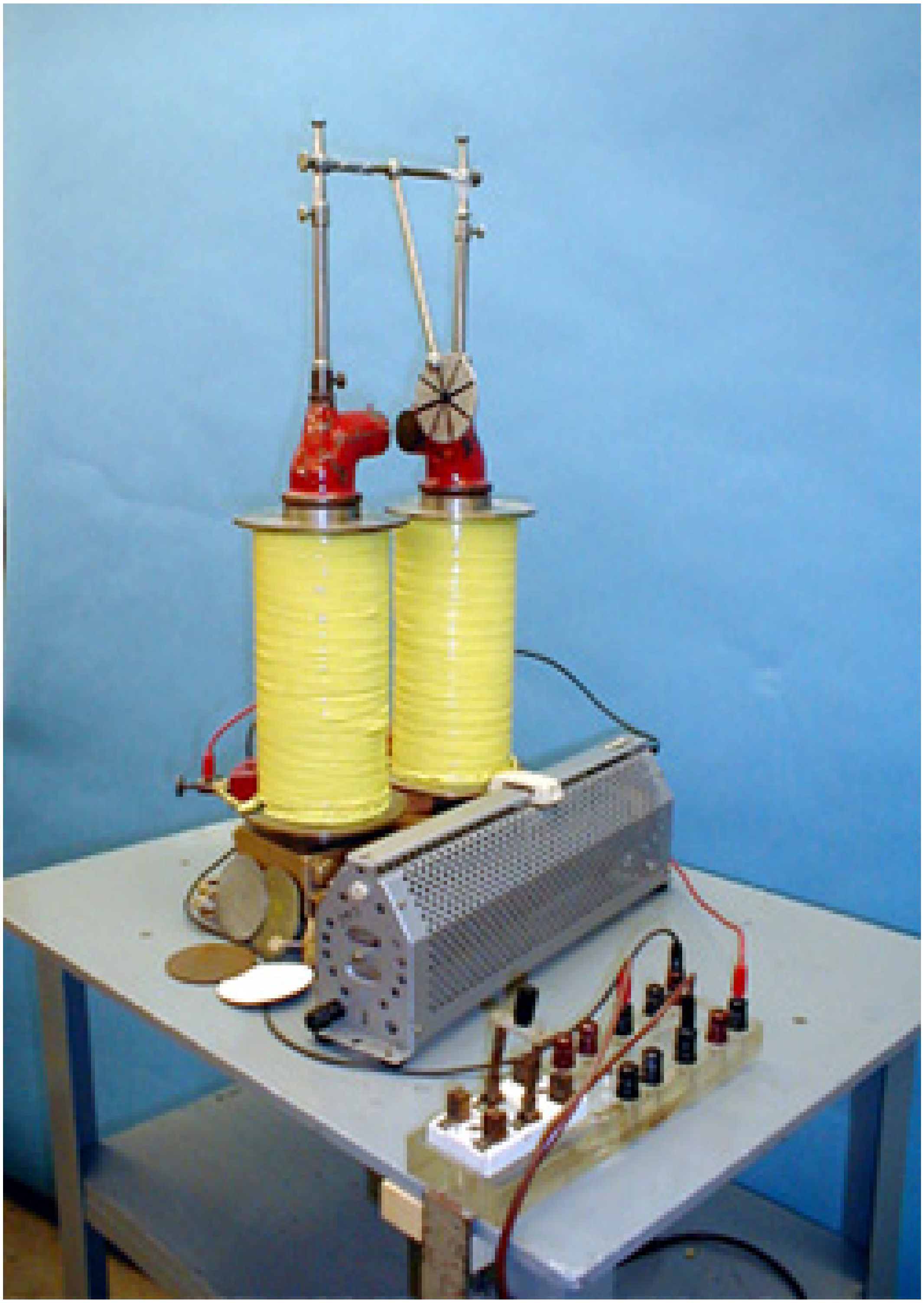


Figure 64: Demo. The current induced in the pendulum when it goes through the poles of the magnet creates a repulsive force that brakes the pendulum halting it. The same principle can be used in electric car brakes. In that case, the induced current can charge the battery and the energy reused to propel the car.

11. Lecture 11

11.1 Inductors

A solenoid a part of a circuit is also called an inductor. Its purpose is to store energy similarly to a capacitor. However it stores energy in a magnetic field as opposed to a capacitor which stores energy in an electric field. This makes its electrical characteristic different from the capacitor. Indeed, we know that the magnetic field produced inside the solenoid is given by

$$\vec{B} = \frac{\mu_0 I N}{\ell}\,\hat{z} \tag{11.1}$$

where N is the number of turns, ℓ is the length of the solenoid and I is the current. If the current changes, the magnetic field and therefore its flux changes (see fig65) producing a voltage between the terminals of the solenoid. Using Faraday's law, Lenz law and looking at the figure we conclude that

$$V_a - V_b = \frac{\Delta\Phi}{\Delta t} = NA\frac{\Delta|\vec{B}|}{\Delta t} = NA\mu_0\frac{N}{\ell}\frac{\Delta I}{\Delta t} = L\frac{\Delta I}{\Delta t} \tag{11.2}$$

where A is the cross section area of the solenoid. We defined the quantity

$$L = \mu_0\frac{N^2 A}{\ell} \tag{11.3}$$

which is given in terms of the area A, the length ℓ and the number of turns. It is a property of the solenoid and is called the inductance. It is measured in Henrys (H):

$$1H = 1\frac{V\,s}{A} \tag{11.4}$$

as can be seen from eq.(11.3). Going back to that formula we see that if we have a DC current, namely $\frac{\Delta I}{\Delta t} = 0$ then $\Delta V = 0$, that is, it essentially acts as a cable. However any change in current will generate a potential difference that opposes that change. To compare with a capacitor we recall that

$$Q = C\Delta V \tag{11.5}$$

If a current goes into a capacitor it will increase its charge so we find

$$\frac{\Delta Q}{\Delta t} = I = C\frac{\Delta V}{\Delta t} \tag{11.6}$$

Comparing with eq.(11.3) we see that the current is proportional to the change in voltage whereas for a solenoid it is the other way around. Although both capacitors and solenoids both store energy, this important difference in their voltage to current relation makes solenoids play a different role in electric circuits. An important application of solenoids is in transformers which we now study.

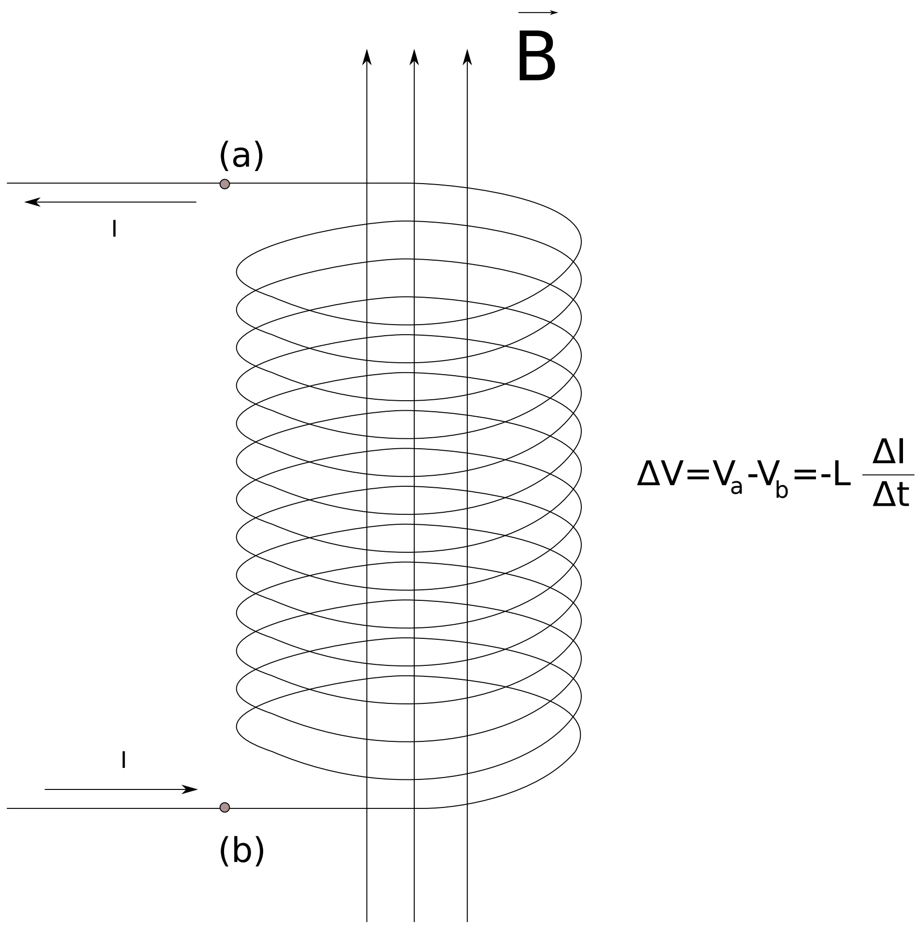


Figure 65: A coil through which a variable current circulates produces a voltage across its terminals given by Faraday's law.

11.2 Transformers

A transformer consists of two solenoids as indicates in fig.66. Through one solenoid, called the primary, we pass a time dependent current which creates a time dependent magnetic flux through the other solenoid called the secondary. This time dependent flux induces a voltage in the secondary which then acts as a battery. The interesting point however is that the voltage in the secondary is not the same as in the primary. Indeed using Faraday's law we find for the voltage in the primary and secondary:

$$V_a - V_b = N_1 A \frac{\Delta B}{\Delta t} \tag{11.7}$$

$$V_c - V_d = N_2 A \frac{\Delta B}{\Delta t} \tag{11.8}$$

Their ratio is given by

$$\frac{V_a - V_b}{V_c - V_d} = \frac{N_1}{N_2} \tag{11.9}$$

So we see that by having different number of turns in the secondary we can increase or decrease the voltage. In the demo we see a transformer that increase the voltage from $120V$ to $15,000V$ producing an interesting display. Stepping up the voltage is not just for show, it is very important in power transmission. When transmitting power through a line, we can consider that the resistance of the line is in series with the load. Therefore the current going through both is the same. To minimize the power lost in the line we need to minimize the current since the power lost is $P = I^2/R$. On the other hand we need to maintain the same power at the load. The only way is to increase the voltage of the line and for that one can use a transformer to step up the voltage to typical values of a few hundred kilovolts.

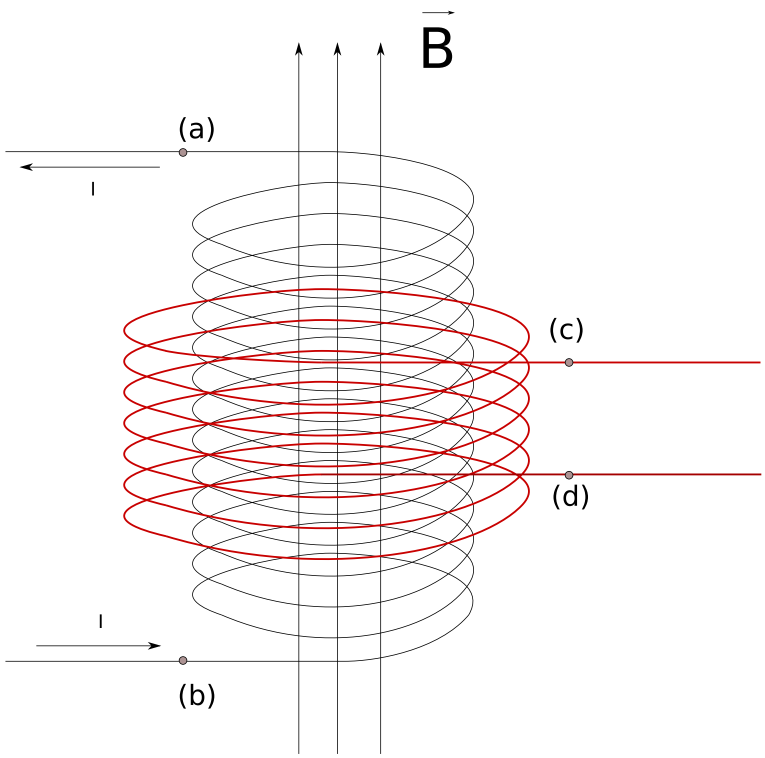


Figure 66: Two coils wrapped around each other constitute a transformer. If a variable current goes through one of them, the other acts as a battery with a variable voltage. The ratio of the voltages in the primary and secondary is given by the ration in the number of turns of each of them.

Figure 67: Demo. A transformer can be used to increase the voltage of an oscillating (alternate) current. The increased voltage can be used to create an interesting display.

12. Lecture 12

12.1 LR circuit, comparison with RC

An inductor stores energy in the form of a magnetic field similarly as a capacitor stores energy in the form of an electric field. In both cases we can "charge" the device by loading energy into it taking it it for example from a battery. In the case of a capacitor we did that through a resistor in an RC circuit. The charging time is given by $\tau = RC$. See the fig.68 to remind yourself how the current and voltage behave as a function of time in such case.

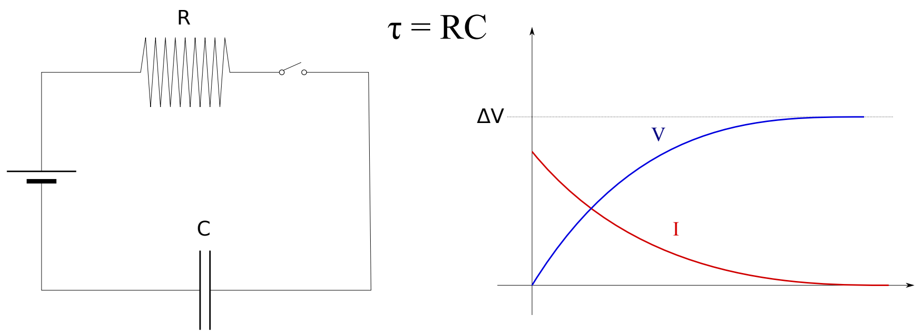


Figure 68: AS seen before, energy from a battery can be stored in a capacitor using an RC circuit. The time constant of the circuit is given by $\tau = RC$. Compare with the RL circuit below.

We can do the same for an inductor by using an RL circuit, a resistor in series with an inductor. See fig.69. Initially, when we turn on the switch, the battery tries to increase the current in the circuit and that will occur almost instantaneously if the inductor is replaced by a cable. However such large changes in current are opposed by the inductor which creates between its terminals a voltage that opposes the increase in current. For that reason the current increases at a slow rate until it reaches its final value given by:

$$I_f = \frac{1}{R}\Delta V_{\text{battery}} \tag{12.1}$$

at that time the current does no change anymore and the inductor behaves as a cable, namely no voltage cross it terminals. However the inductor is generating a magnetic field so it has energy stored in it. To see how long it takes to reach such a stationary state, consider the voltage across the inductor

$$V_a - V_b = L\frac{\Delta I}{\Delta t} \tag{12.2}$$

Initially, as we said, the inductor opposes the battery and does not allow any current to circulate. This means that $V_a - V_b = \Delta V_{\text{battery}}$ and the initial slope with which the current start increasing is then:

$$\left.\frac{\Delta I}{\Delta t}\right| = \frac{1}{L}\Delta V_{\text{battery}} \tag{12.3}$$

As the current increases, this slope decreases as can be see in the plot in fig.69. However, we get a good estimate of the time it takes to reach the steady state if we assume the slope to be constant and ask how long it would take with that slope to reach the current I_f. It is

$$\Delta t = \frac{I_f}{\frac{\Delta I}{\Delta t}} = \frac{\Delta V_{\text{battery}}}{R}\frac{1}{\frac{\Delta V_{\text{battery}}}{L}} = \frac{L}{R} \tag{12.4}$$

So we find the characteristic time for an RL circuit is

$$\tau = \frac{L}{R} \tag{12.5}$$

The larger the inductance the longer it takes to charge. If the resistance is small it also takes longer since we want to reach a larger current in the end.

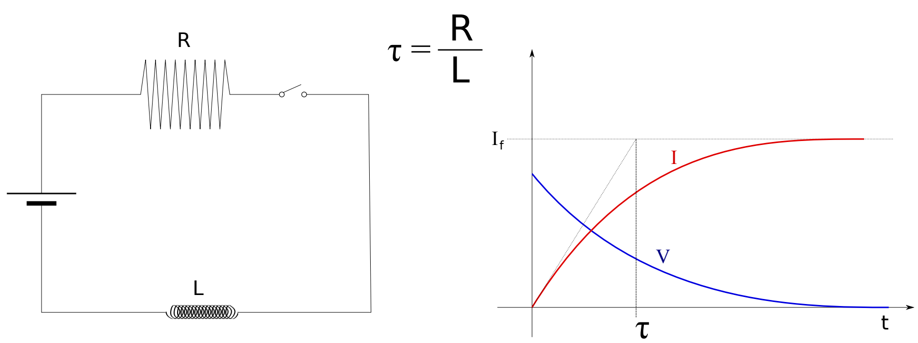


Figure 69: Energy from a battery can be loaded in an inductor using an RL circuit. The time constant of the circuit is given by $\tau = \frac{R}{L}$. Pay attention to the different voltage and current curves compared to the capacitor case.

12.2 Using the oscilloscope

A very useful tool to check what we discussed in the previous section is the oscilloscope (sometimes called simply "scope"). It measures the voltage across its terminals and plots it as a function of time. One terminal is usually grounded and the other, which is called a probe, has a conducting pin that you can touch to different parts of the circuit to measure the voltage. Digital oscilloscopes have integrated circuits called analog to digital converters which take the voltage and covert it into a number which they store in a memory. The instrument then displays the voltage as a function of time using an LCD display. On the screen the horizontal scale is time and the vertical scale is voltage. The scale can be set manually or can be left automatic (easier). In the past, and also for high frequencies oscilloscopes work with a vacuum tube where an electron beam was directed toward a screen. The path of the beam was changed by a potential applied to two parallel plates. In the horizontal direction is was moved back and forth at a constant rate (backward much faster than forward). In the vertical direction it was moved according to the input voltage. The resulting point on the screen traces the voltage as a function of time. Both type of oscilloscopes are perfectly good for looking at the RL circuit. In the demo we do that and check the voltage and current as a function of time. notice that to measure the current we can use Ohm's law:

$$\Delta V = IR \tag{12.6}$$

So we can just measure the voltage cross the resistor to know the current. The humble resistor can be though as a current to voltage converter!.

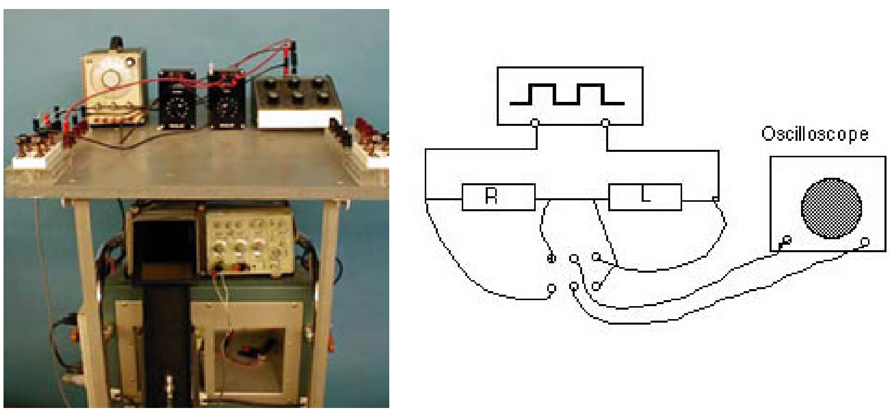


Figure 70: Demo. An RL circuit is analyzed with the aid of an oscilloscope.

12.3 Energy contained in a solenoid

We mentioned several times that a solenoid or inductor contain energy. We are going to derive that such energy is given by

$$U = \frac{1}{2}LI^2 \tag{12.7}$$

where L is the inductance and I the current circulating through the solenoid. To see how this works recall that for a solenoid of area A, length ℓ and number of turns N,

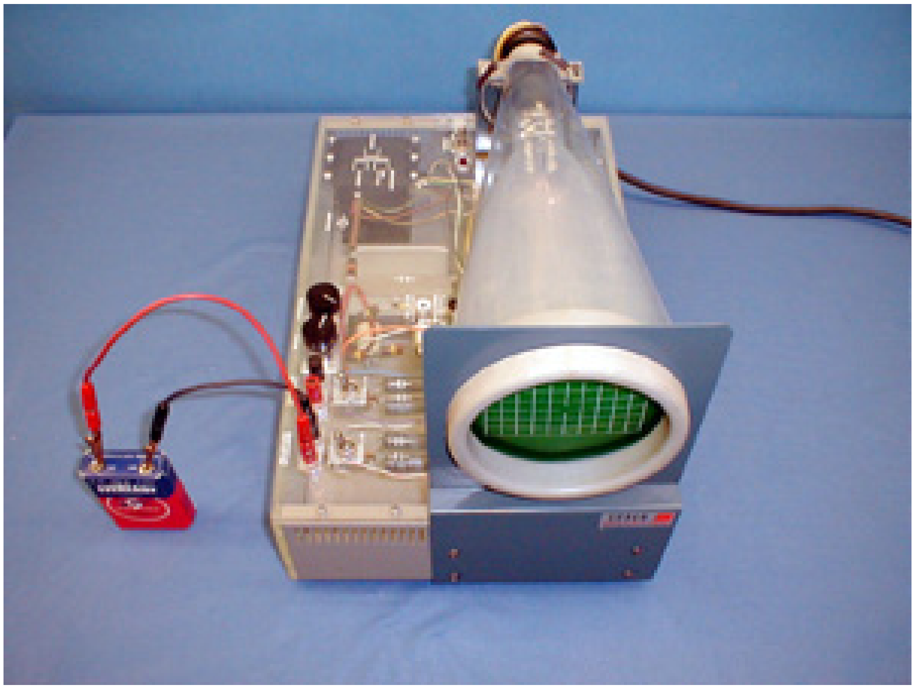


Figure 71: Demo. A vacuum tube where an electron beam is directed toward a screen producing a bright spot. The beam can be deflected by electric and magnetic fields. This allows to build an oscilloscope by using external input to deflect the beam.

the inductance is given by

$$L = \mu_0 \frac{N^2 A}{\ell} \tag{12.8}$$

The energy therefore should be given by

$$U = \frac{1}{2}\mu_0 \frac{N^2 A}{\ell} I^2 \tag{12.9}$$

To verify this result we are going to is a procedure analogous to what we did for a capacitor. In that case we computed how much energy was required to separate the parallel plates by a small distance Δd. This allowed us to compute the energy in the capacitor. Here we can see how much energy is required to expand the solenoid so that the area is increased by ΔA. The reason we do that is that, from eq.(12.9) we expect that the energy is linear in A so it should be easy to compute the energy necessary to change A. The first consideration is that if we increase A, the flux through the solenoid changes and therefore a voltage is induced:

$$V_a - V_b = \frac{\Delta\Phi}{\Delta t} = -N|\vec{B}|\frac{\Delta A}{\Delta t} = -\mu_0 \frac{N^2}{\ell} I \frac{\Delta A}{\Delta t} \tag{12.10}$$

This means that as we do this, we need to push the current against such potential difference. The power necessary to do that is

$$P = I(V_a - V_b) = \mu_0 \frac{N^2}{\ell} I^2 \frac{\Delta A}{\Delta t} \tag{12.11}$$

Since power is energy consumed in unit time, to get the total energy necessary to do this we multiply the power by Δt obtaining

$$\Delta U_1 = P\Delta t = \mu_0 \frac{N^2}{\ell} I^2 \Delta A \tag{12.12}$$

This energy we loose, so it is incorporated into the solenoid. If we compare with eq.(12.9) we see that there is a factor of two. So we must be gaining energy somewhere else because we overestimated the energy we needed to do this. The reason is that the magnetic field exerts a force on the solenoid that tends to expand it. This forces does work when we expand the solenoid and therefore we get energy form there. Let us see how much this is. The force on a small portion of the cable is radial and per unit length equal to

$$\frac{|\vec{F}|}{\lambda} = |\vec{B}'| I \tag{12.13}$$

The magnetic field is indicated as $\vec{B}'$ because we should not consider the magnetic field created by the small portion of cable itself. To see what this is we can use Ampere's law to compute the magnetic field generated by a small piece of the solenoid (see fig.72). It gives:

$$2|\vec{B}|\Delta\ell = \mu_0 \frac{N}{\ell} \Delta\ell I \tag{12.14}$$

implying that

$$|\vec{B}| = \frac{\mu_0 N}{2\ell} I \tag{12.15}$$

which is precisely half the magnetic field of the solenoid!. What that from the magnetic field acting on a small portion of the cable, half is created by the rest of the solenoid and half is created by itself and should be excluded (because the small portion of cable cannot make a force on itself). We conclude that

$$|\vec{B}'| = \frac{\mu_0 N}{2\ell} I \tag{12.16}$$

We can now compute the force

$$\frac{|\vec{F}|}{l} = \frac{\mu_0 N}{2\ell} I^2 \tag{12.17}$$

The length of each turn is $2\pi R$ where R is the radius of the coil. The displacement if we increase the radius by ΔR is precisely ΔR. We can the compute the work as

$$W = \frac{|\vec{F}|}{l} 2\pi R \, \Delta R = \frac{\mu_0 N}{2\ell} I^2 \Delta A \tag{12.18}$$

where we used that, if we change the radius by ΔR, the area changes precisely by $\Delta A = 2\pi R \Delta R$ (assuming $\Delta R \ll R$). This is work done by the solenoid, so we gain energy. Therefore the total change in the energy of the solenoid is

$$\Delta U = \Delta U_1 - W = \frac{\mu_0 N}{2\ell} I^2 \Delta A \tag{12.19}$$

If we now assume that $U = 0$ when $A = 0$ because when $A = 0$ there is no magnetic field to speak of, we obtain that

$$U = \frac{\mu_0 N}{2\ell} I^2 A = \frac{1}{2} L I^2 \tag{12.20}$$

as we wanted to show. Admittedly this derivation is a bit lengthy but is a very useful exercise if you want to understand several properties of magnetic field, work, energy, etc. If you go through it carefully, the basic idea should appear quite simple and then one has to do the calculations with great care.

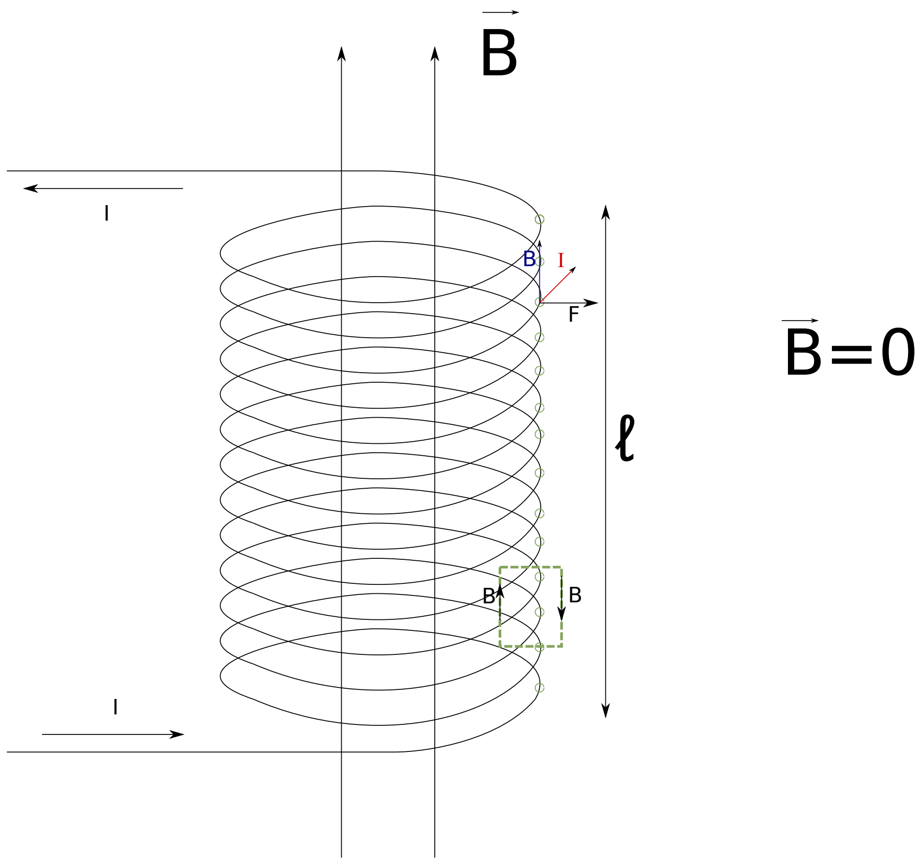


Figure 72: Energy contained in a coil. There is a force trying to expand the coil. It is created by the magnetic field acting on the current. However the magnetic field created by the small piece of coil considered should be excluded. If we expand the coil this force does work and the energy of the coil is reduced. On the other hand, a voltage is induced and we need to do work to keep the current circulating. In total the energy of the coil increases. It is given by $U = \frac{1}{2}LI^2$.

13. Lecture 13

13.1 Electric generators and alternate current

We discussed before that a cable moving in a magnetic field acts as a battery since the magnetic field creates a force that moves the charges along the cable. This is the principle of the electric generator. It is more efficient to move the cable in circles so the motion is repetitive. In figs.73, 74, 75 we have a simple generator (see fig.76 for an actual example). A rectangular loop moves inside a magnetic field an a voltage is generated between terminals (a) and (b). We can find the voltage from Faraday's law or equivalent from the Lorentz force.

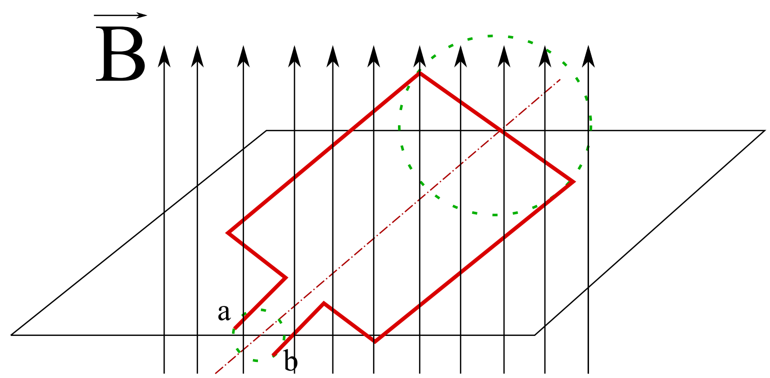


Figure 73: Schematic AC generator. A cable in the shape of a loop rotates inside a magnetic field. View in perspective.

The magnetic flux going through the loop is given by

$$\Phi = A|\vec{B}|\cos\phi \tag{13.1}$$

where A is the area of the loop, $|\vec{B}|$ is the modulus of the magnetic field (created for example by a permanent magnet) and ϕ is the angle between the normal to the loop and the magnetic field. If the loop is horizontal the flux is maximum and if it is vertical the flux is zero. If the loop rotates with angular velocity ω we have $\phi = \omega t$. The flux changes in time and, according to Faraday's law a voltage is appears between the terminals:

$$V_a - V_b = \frac{\Delta\Phi}{\Delta t} = A|\vec{B}|\frac{\Delta\cos\omega t}{\Delta t} \tag{13.2}$$

The only complication is that we need to compute the variation in time of $\cos\omega t$.

This can be done by using a mechanical analog. Indeed, if the position of an object is given by

$$x = \cos\omega t \tag{13.3}$$

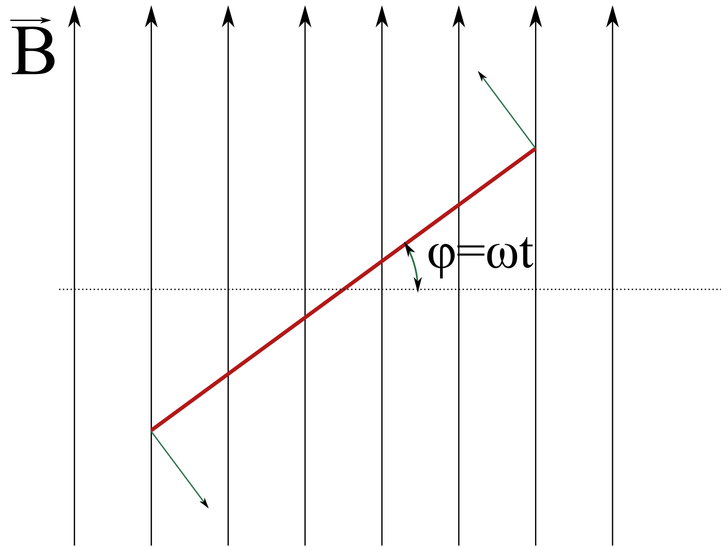


Figure 74: Schematic AC generator. A cable in the shape of a loop rotates inside a magnetic field. Front view.

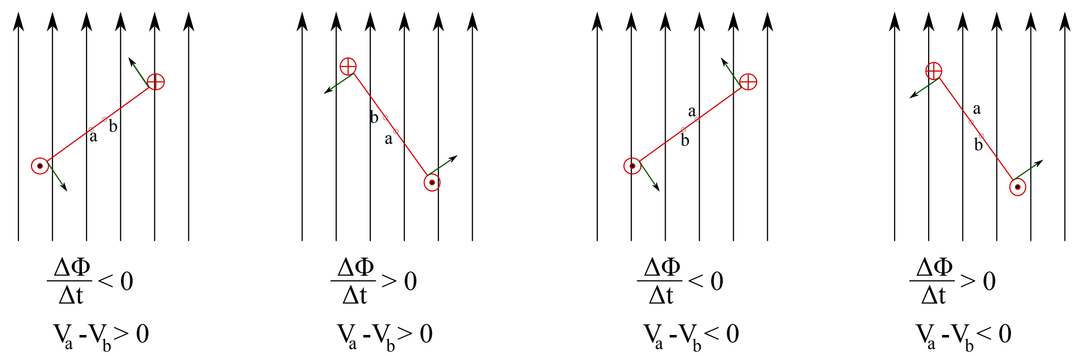


Figure 75: Schematic AC generator. The variation in flux creates a potential between terminals (a) and (b). The polarity alternates as seen in this figure. The symbols at the end of the conductor indicate if the current goes into the page $\oplus$ or out of the page $\odot$.

then its velocity is given by

$$v_x = \frac{\Delta x}{\Delta t} = \frac{\Delta \cos \omega t}{\Delta t} \tag{13.4}$$

We know how to do this, we just need to consider a particle moving in circles as in fig.77. If it rotates with angular velocity
omega, the position is given by

$$x = R \cos \omega t \tag{13.5}$$

$$y = R \sin \omega t \tag{13.6}$$

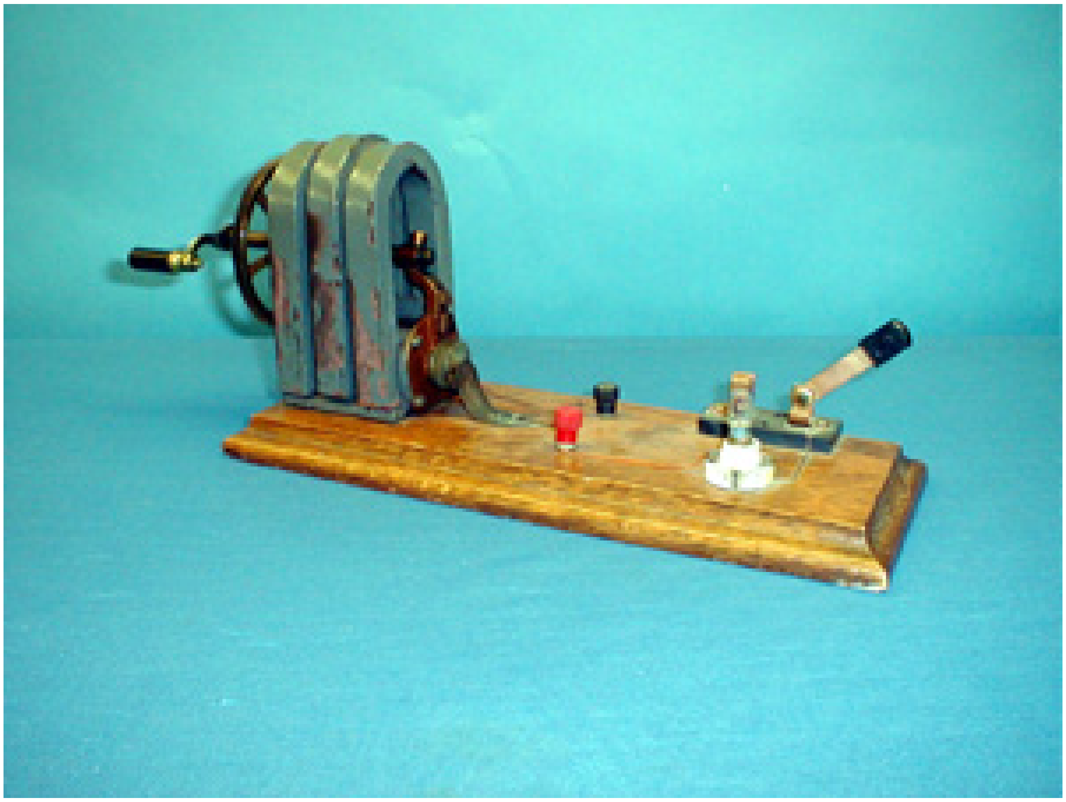


Figure 76: Demo. Actual AC generator. Cranking the handle creates enough power to light up a small light bulb.

The modulus of the velocity is given by the fact that it goes around the circle in time $T = \frac{2\pi}{\omega}$. The length of the circle is $\ell = 2\pi R$ so the velocity is

$$|\vec{v}| = \frac{2\pi R}{T} = \omega R \tag{13.7}$$

The direction of the velocity is as in the figure. Projecting over the x and y axis we get for the velocity:

$$v_x = -\omega R \sin \omega t \tag{13.8}$$

$$v_y = \omega R \cos \omega t \tag{13.9}$$

We then derived the following very important formulas

$$\frac{\Delta \cos \omega t}{\Delta t} = -\omega \sin \omega t \tag{13.10}$$

$$\frac{\Delta \sin \omega t}{\Delta t} = \omega \cos \omega t \tag{13.11}$$

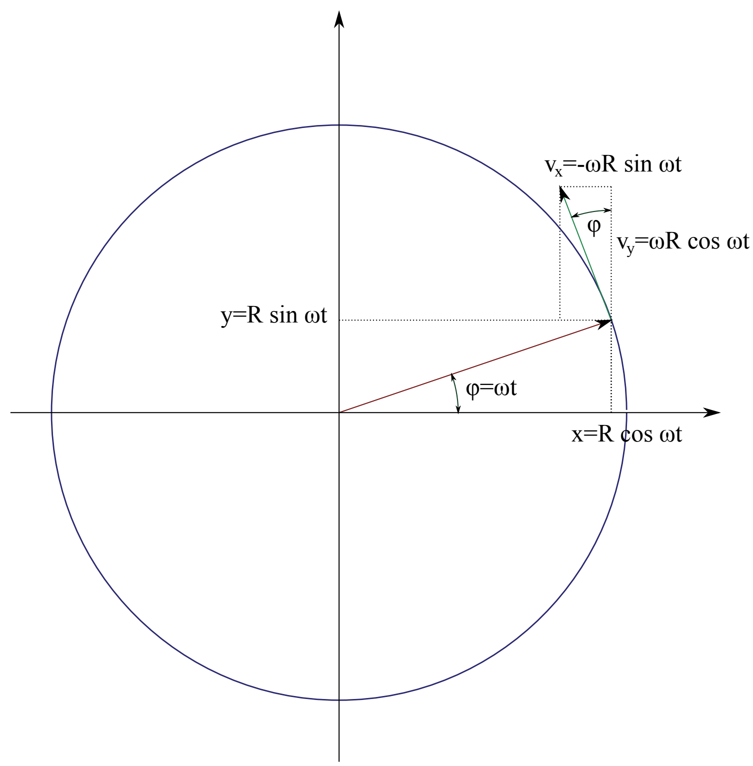


Figure 77: For a particle rotating in circle sit is easy to compute the velocity. This gives us same useful formulas that we can apply in other situations.

The notation is not very precise but what it means is that is something moves as $\cos \omega t$ then its velocity is $-\omega \sin \omega t$ and similarly for the sine. This mechanical analog allows us then to compute the voltage as

$$V_a - V_b = -A|\vec{B}|\omega \sin \omega t \tag{13.12}$$

It is clearly seen form the nature of the motion that the relative polarity of (a) and (b) alternates. For that reason this is called an AC generator. AC stands for alternate current. If connected to a resistor the current circulates in one direction and then the other, that is why it is is called alternate current. The household current alternates 60 times a second. This is called the frequency and is measured in Hertz:

$$1Hz = 1\frac{1}{s} \tag{13.13}$$

The household current therefore has a frequency of 60Hz. It should be noticed that the function

$$V = V_0 \sin \omega t \tag{13.14}$$

repeats itself when we shift $t \to t + \frac{2\pi}{\omega}$ (since sine is a function that repeats itself when the argument is shifted by 2π). therefore the period and frequency of the current are defined as:

$$T = \frac{2\pi}{\omega}, \qquad f = \frac{1}{T} = \frac{\omega}{2\pi} \tag{13.15}$$

13.2 AC resistor circuit

An AC generator connected to a resistor dissipates power. Notice that power is dissipated no matter the direction of the current. According to Ohm's law we have

$$\Delta V = IR \tag{13.16}$$

For the generator we have

$$\Delta V = V_0 \sin \omega t \tag{13.17}$$

therefore

$$I = \frac{1}{R} V_0 \sin \omega t \tag{13.18}$$

the power dissipated is

$$P = I \Delta V = \frac{1}{R} V_0^2 \sin^2 \omega t \tag{13.19}$$

It changes in time but we can compute the average power dissipated. To average over one cycle we notice that, if instead of a sine we had a cosine, the power dissipated would be the same. But we have (denoting average with parenthesis):

$$\langle \sin^2 \omega t + \cos^2 \omega t \rangle = \langle 1 \rangle \qquad \Rightarrow \qquad \langle \sin^2 \omega t \rangle = \frac{1}{2} \tag{13.20}$$

We get then

$$\langle P \rangle = \frac{V_0^2}{2R} = \frac{V \text{rms}}{R} \tag{13.21}$$

where we defined

$$V \text{rms} = \frac{1}{\sqrt{2}} V_0 \tag{13.22}$$

which is known as root mean square voltage (and V_0 is known as peak voltage).

14. Lecture 14

14.1 AC circuits: capacitors and inductors

14.1.1 Capacitors

Consider a circuit as in fig.78 which is called a low-pass filter for reasons we will see shortly. The AC generator determines the potential

$$V_a = V_0 \sin \omega t \tag{14.1}$$

and we want to understand the behavior of the potential V_b at point (b). It is first convenient to see what happens if, instead of a sine we have a square wave as in fig.79. This is equivalent to putting a battery which periodically switches polarities as we see in the same figure. When the battery is of one polarity it charges the capacitor in time $\tau = RC$ as we saw before. After the capacitor charges nothing else happens and V_b remains equal to V_0. When the battery switches polarity the capacitor first discharges and then it charges with the opposite sign so $V_b = -V_0$ and V_b continues to be $-V_0$ until the battery switches back again. We see that, except for a small delay of $\tau = RC$ the potential V_b follows the value at V_a. In that sense the capacitor, for time scales $t \gg RC$ behaves as an open circuit. This last point should be emphasized, we assumed that the period with which we switched the battery was much larger than $\tau = RC$. If we switch the battery very fast then the capacitor has no time to charge and discharge and the potential across it is zero. Namely $V_b = 0$ and the capacitor is like a short-circuit, namely a cable.

To summarize:

$$\text{Capacitor} \begin{cases} T \gg \tau; & \omega \ll \frac{1}{RC}; \quad \text{Capacitor} \to \text{open circuit} \\ T \ll \tau; & \omega \gg \frac{1}{RC}; \quad \text{Capacitor} \to \text{short-circuit (cable)} \end{cases} \tag{14.2}$$

Using this information, if we go back to the circuit in fig.78 we conclude that

$$V_b = \begin{cases} V_a = V_0 \sin \omega t & \text{if} \quad \omega \ll \frac{1}{RC} \\ 0 & \text{if} \quad \omega \gg \frac{1}{RC} \end{cases} \tag{14.3}$$

For that reason it is called a low-pass filter, any high frequency signal does not appear at point (b). On the other hand if we have a circuit as in fig.80 we have, using the same rules:

$$V_b = \begin{cases} 0 & \text{if} \quad \omega \ll \frac{1}{RC} \\ V_a = V_0 \sin \omega t & \text{if} \quad \omega \gg \frac{1}{RC} \end{cases} \tag{14.4}$$

For that reason it is called a high-pass filter. Low frequencies, and in particular DC potentials are blocked by the capacitor. A typical application of the last circuit is in what is called AC coupling. Suppose we have an audio signal mounted on a DC voltage. If we need to input that to an amplifier but we do not want to keep the DC voltage because it might affect the amplifier then we can use a high pass filter as in fig.81. The capacitor and resistors should be chosen so that they do not cut the frequencies we are interested in.

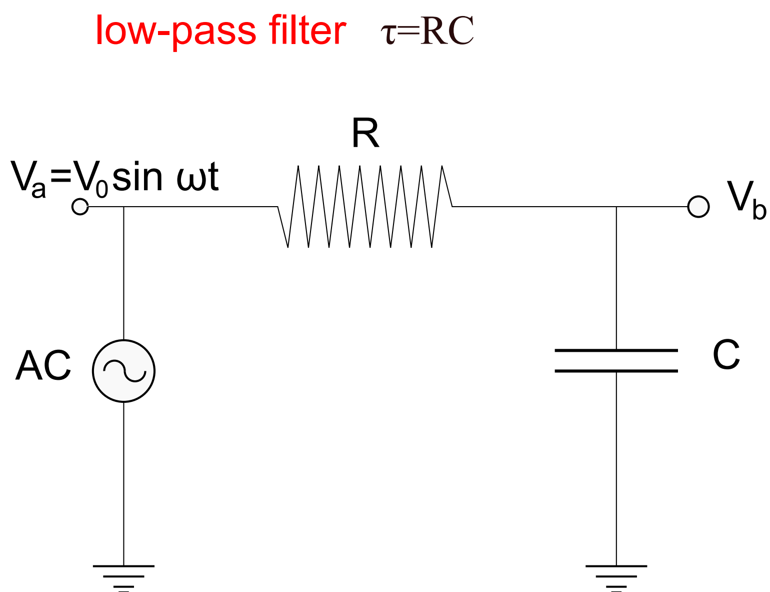


Figure 78: AC generator connected to an RC circuit. Low-pass filter, allows only low frequencies to go through.

14.1.2 Inductors

Inductors behave in the opposite way as capacitors. At low frequencies, including DC current, they behave as a short circuit, just like a cable. There is no voltage across its terminals. At high frequency however, there is no current, because any such current will vary very rapidly and would create a very large voltage across its terminals. What actually happens is that the voltage generated is enough to cancel any voltage applied and very little or no current circulates. So:

$$\text{Inductor}\begin{cases} T \gg \tau = \frac{L}{R}; & \omega \ll \frac{R}{L}; \quad \text{Inductor} \rightarrow \text{short-circuit (cable)} \\ T \ll \tau = \frac{L}{R}; & \omega \gg \frac{R}{L}; \quad \text{Inductor} \rightarrow \text{open circuit} \end{cases} \tag{14.5}$$

Again, looking at the example of fig.82 we have

$$V_b = \begin{cases} 0 & \text{if} \quad \omega \ll \frac{R}{L} \\ V_a = V_0 \sin \omega t & \text{if} \quad \omega \gg \frac{R}{L} \end{cases} \tag{14.6}$$

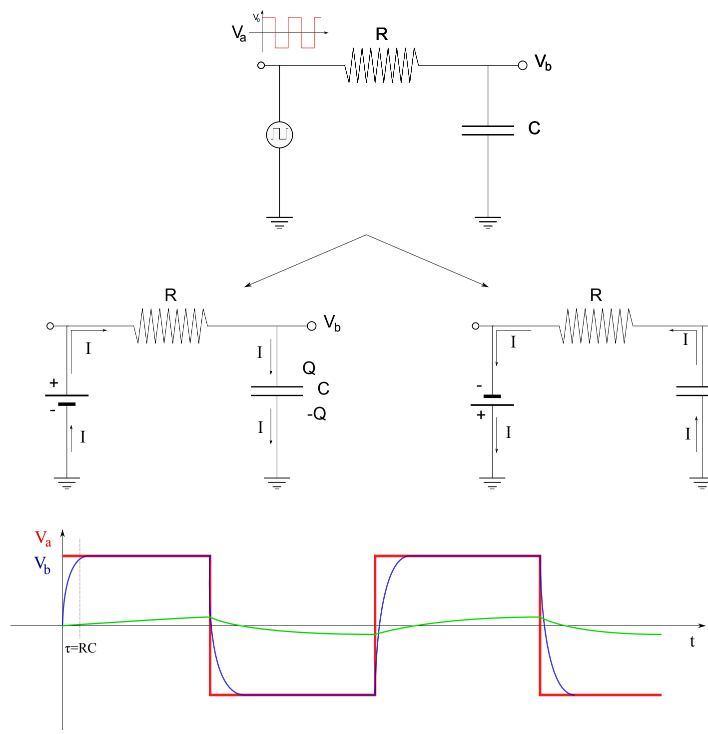


Figure 79: Previous circuit connected to a square wave generator. It is equivalent to a battery that flips polarity. If the flips are at long intervals then the capacitor charges and discharges following the voltage. If they are very frequent then the capacitor has no time to charge and discharge and the voltage across it is very small.

This is called a low-pass filter. In the case of fig.83 we have instead:

$$V_b = \begin{cases} V_a = V_0 \sin \omega t & \text{if} \quad \omega \ll \frac{R}{L} \\ 0 & \text{if} \quad \omega \gg \frac{R}{L} \end{cases} \tag{14.7}$$

This is a high-pass filter.

All these ideas are very nicely illustrated in the demo of fig.84 where changing the frequency of the generator we can see how the behavior of the capacitor and inductor changes according to the rules we saw before. It is interesting to find out the cut-off frequency for the capacitor. From the circuit one can see that $C = 220\mu F$. The light-bulb is a resistor. It is a 23W light-bulb when operated at 12V. From here we find:

$$P = 25W, \quad V = 12V, \quad P = \frac{V^2}{R}, \quad R = 6\Omega \tag{14.8}$$

Therefore

$$\omega = \frac{1}{RC} = 800Hz \tag{14.9}$$

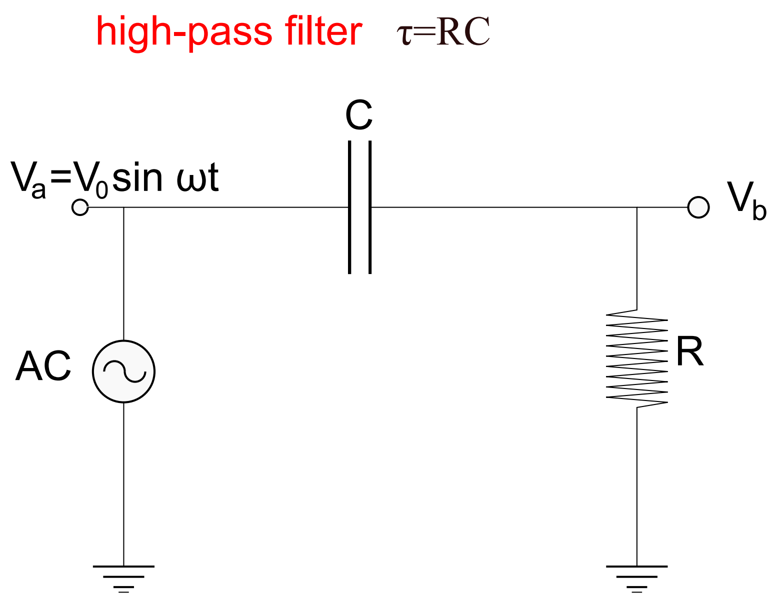


Figure 80: AC generator connected to an RC circuit. High-pass filter, allows only high frequencies to go through.

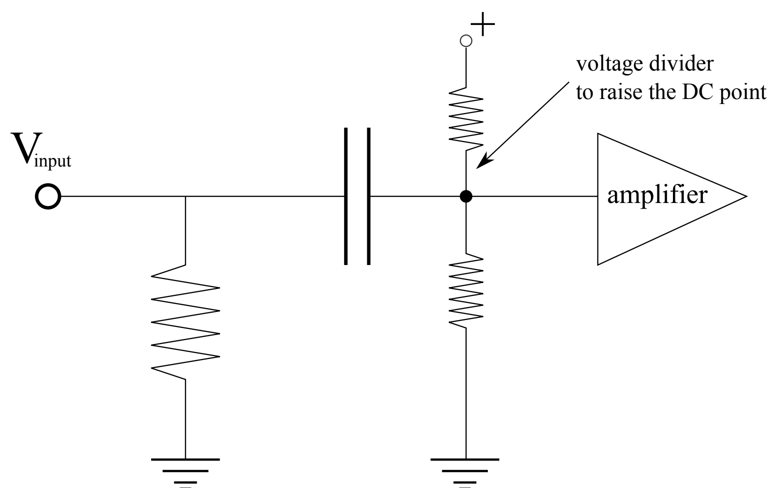


Figure 81: Example of AC coupling. If an AC signal is mounted on top of a DC voltage, only the AC part goes through (for adequate values of the capacitance and resistance).

The frequency, as we saw before is

$$f = \frac{\omega}{2\pi} = \frac{800}{2\pi} Hz \simeq 130 Hz \tag{14.10}$$

which can be easily verified in the demo by changing the frequency of the generator.

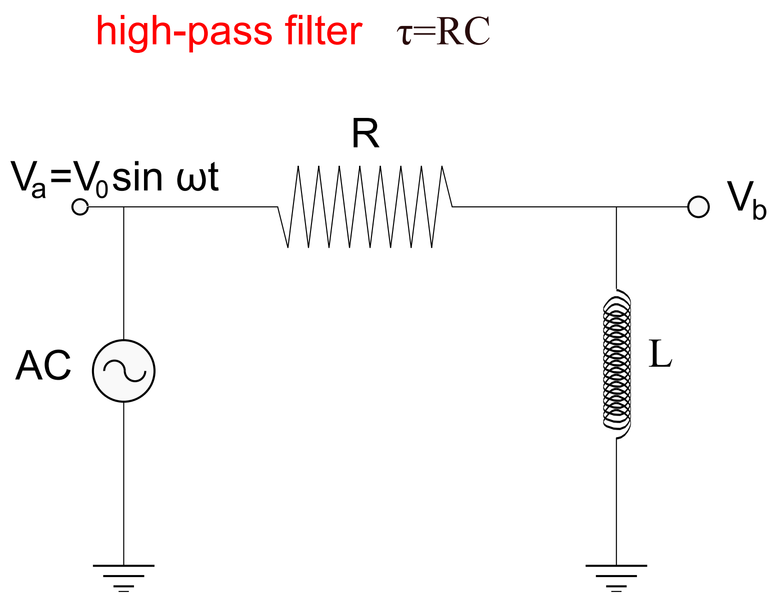


Figure 82: AC generator connected to an RL circuit. High-pass filter, allows only low frequencies to go through.

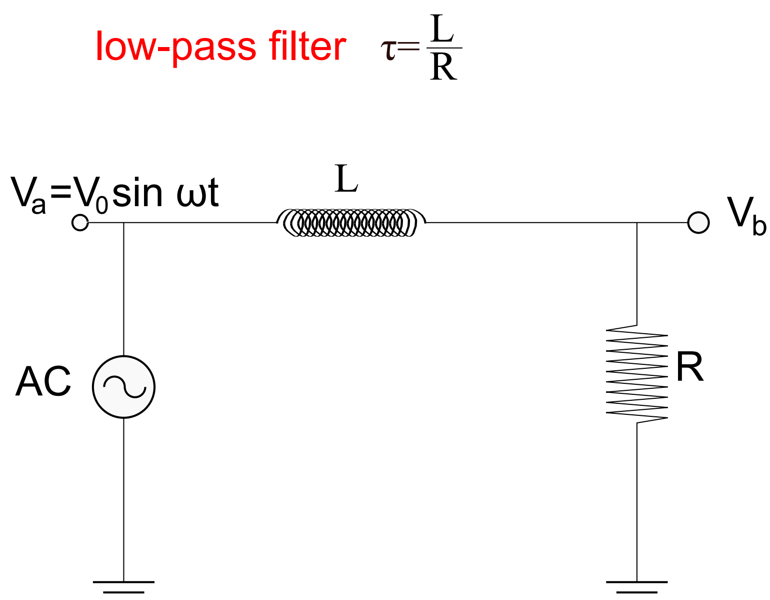


Figure 83: AC generator connected to an RL circuit. Low-pass filter, allows only high frequencies to go through.

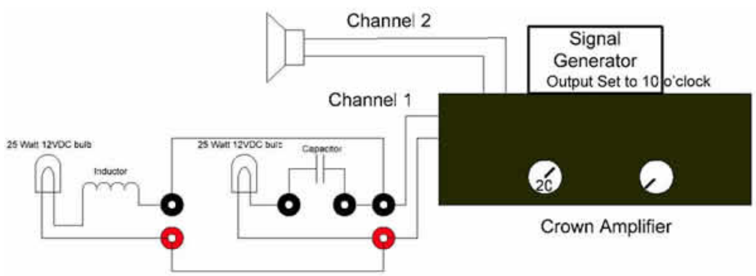


Figure 84: Demo: Variable frequency AC generator connected to an RL and RC circuit to demonstrate the AC properties of capacitors and inductors.

15. Lecture 15

15.1 Demo: Sound transmission with (laser) light

Consider the circuit in figure 85 where and audio source drives the primary of a transformer. The secondary of the transformer is in series with a battery and a laser or LED light. Since the secondary of the transformer acts as a battery whose voltage fluctuates following the original sound, the intensity of the light will also fluctuate accordingly. This fluctuations are small and too fast to be seen by the eye, so the laser pointer seems to be working as usual but in fact contains the information about the sound. This can be detected by using a solar cell. The solar cell generates a voltage that is proportional to the intensity of the light and that therefore fluctuates as the original voltage. If we connect to any device (in this case a children's toy) that has an input for a microphone, it will be the sound manifest either on a loudspeaker (as we do) or by recording it. The idea for this interesting circuit can be found in several websites:
http://www.i-hacked.com/index.php?option=com_content&task=view&id=162&Itemid=44, http://www.wikihow.com/Transmit-Audio-With-a-Laser-Pen, etc. It is always interesting to look around and try to see how the ideas we learn can be used in simple ways to produce interesting results.

15.2 Electromagnetic waves

A varying electric field produces a magnetic field, a (varying) magnetic field. A varying magnetic field creates a (varying) electric field. This gives rise to a periodic wave that propagates in space. In vacuum they propagate at speed $c \simeq 3\times 10^8 \frac{m}{s}$. This is the speed of light which is an example of an electromagnetic wave. The speed of propagation can be written as

$$c = \frac{1}{\sqrt{\mu_0 \epsilon_0}} \tag{15.1}$$

Generically a wave is profile that propagates (see fig.86). The intensity of the field oscillates in time and also changes in space. The distance between to crests is called the wave-length λ. At a given point the intensity oscillates in time and the interval between two maxima is called the period. It is clear that, if the profile propagate at velocity c, the two maxima separated by a distance λ will arrive at a time interval of $T = \frac{\lambda}{c}$. The inverse of the period is called the frequency

$$f = \frac{1}{T} = \frac{c}{\lambda} \tag{15.2}$$

and is measured in Hertz, where $1Hz = 1\frac{1}{s}$ a unit of frequency we already saw for AC current. As a simple example consider a FM radio station transmitting at $f =$

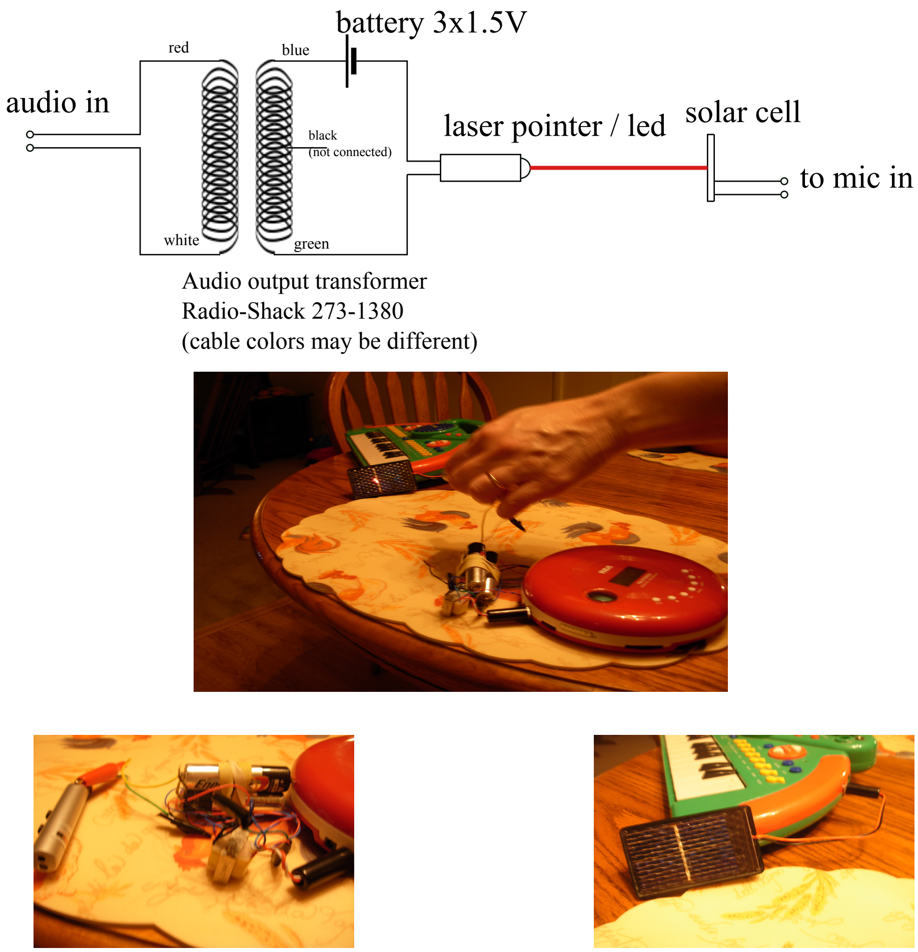


Figure 85: Demo. A transformer is used to modulate, that is to change the intensity, of a laser following an audio input. The variations in the intensity of the light are detected by a solar cell (or photoelectric panel) and converted into sound by an amplifier. The same principle works if we use a LED but the laser beam stay narrow at much longer distances and transmits the signal more effectively.

$100MHz = 10^8 Hz$. The period is $T = \frac{1}{f} = 10^{-8}s$. In that time the wave travels $\lambda = cT = 3m$ which is the wavelength of that particular radio station. In the case of electromagnetic waves the fields that oscillate are the electric and magnetic fields. For that reason we need to know, not only their intensity but also their orientation. Both from theory and experiment it appears that electromagnetic waves are such that the electric and magnetic field are perpendicular to each other and perpendicular to the direction of propagation. The direction of propagation $\hat{k}$ is given by the right hand rule: $\hat{k} \parallel \vec{E} \times \vec{B}$. In fig. 87 we see a depiction of an e.m. wave. The electric field can point in any direction contained in the plane perpendicular to the direction

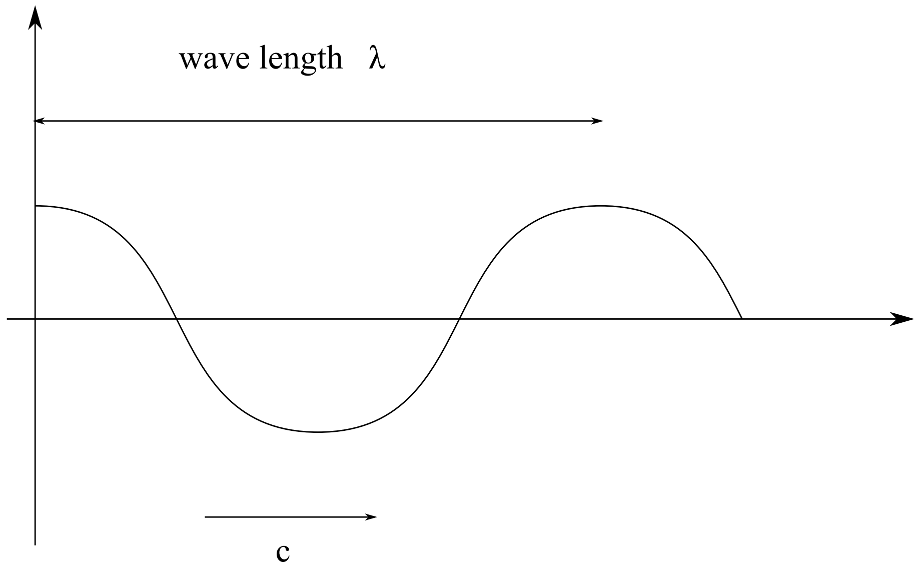


Figure 86: Generic periodic wave. A disturbance propagates with velocity c. The distance between to peaks is called the wave-length λ. At a fixed point the time interval between the passage of two peaks is called the period $T = \frac{\lambda}{c}$.

of propagation of light. If it points always in the same direction then it is said that the wave is linearly polarized and the direction of polarization is the direction of the electric field. In general any wave can be consider as a superposition of two waves with orthogonal polarizations. That is, suppose that the wave propagates in direction $\hat{z}$, then the electric field can point in the direction of $\hat{x}$ or y or can be a superposition of two waves, one polarized along $\hat{x}$ and the other along $\hat{y}$. For example, an interesting case is circular polarization where the electric field rotates in the (xy) plane. This can be considered as a superposition of two oscillations, one in x and the other in y which are out of phase (this is just from projecting $\vec{E}$ in its two components along $\hat{x}$ and $\hat{y}$). A simple demo show how this works (see fig.**??**). The e.m. wave is detected by putting two metallic rods connected by a light-bulb. When a electric field is present it moves the electrons in the metal and generates a current that lights up the electric bulb. The current has a maximum when the rods are parallel to the electric field which allows us to find its direction. Other important facts about e.m. waves are that modulus of electric and magnetic fields are related by:

$$|\vec{E}| = c|\vec{B}| \tag{15.3}$$

and that the intensity of the wave is given by

$$I = \frac{1}{2}\epsilon_0 c|\vec{E}|^2 \tag{15.4}$$

Is is measured in $\frac{W}{m^2} = \frac{J}{m^2\,s}$ and determines how much energy crosses a given area per unit time. For example for a solar cell this gives the power generated by square meter of solar cell (up to the efficiency factor since part of the energy is converted into heat).

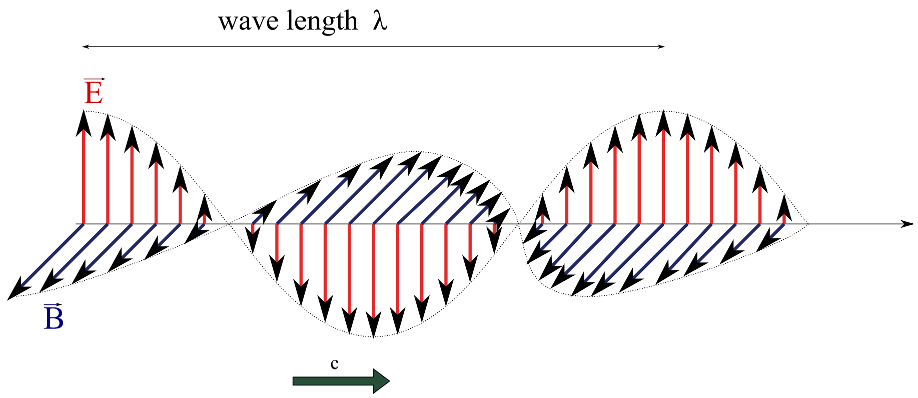


Figure 87: Electromagnetic wave. The electric and magnetic field are perpendicular to each other and to the direction of propagation.

Although electromagnetic waves of different wave-length represent the same physical phenomenon, they interact (namely are generated and absorbed) with matter in a different way since they tend to interact with objects of the size of order the wave-length and/or with processes whose timescale is similar to the period of the wave. For that reason they are known with different names. For example γ-rays interact with the atomic nucleus and therefor have wave-length of order $\lambda \sim 10^{-15} m$ or shorter. X-rays and light interact with atoms and therefore have wave-length in the order of $10^{-10} m$ up to $10^{-6} m$. Radio waves are generated by oscillating circuits and antennas and have wave-length in the meters or even kilometers. Fig.88 gives a more precise classification.

15.3 Light as an electromagnetic wave

In the previous section we saw that light are electromagnetic in a particular narrow region of wave-length. It happens to be the range to which our eyes are sensitive to. The main reason seems to be that water is more transparent in that region allowing us to see under water and also in the atmosphere (which has large amounts of water vapor). In the case of light the polarization can be detected by a polarizer, a medium with a prefer direction which allows only the passage of waves polarized in that direction. If the wave is polarized in a different direction, only the component of the electric field parallel to the preferred direction goes through. That is (see fig.89), if the angle between the electric field and the preferred direction is θ, the magnitude of the electric field after the wave goes through the polarizer if $|\vec{E}_{\text{out}}| = |\vec{E}_{\text{out}}| \cos\theta$. the intensity is

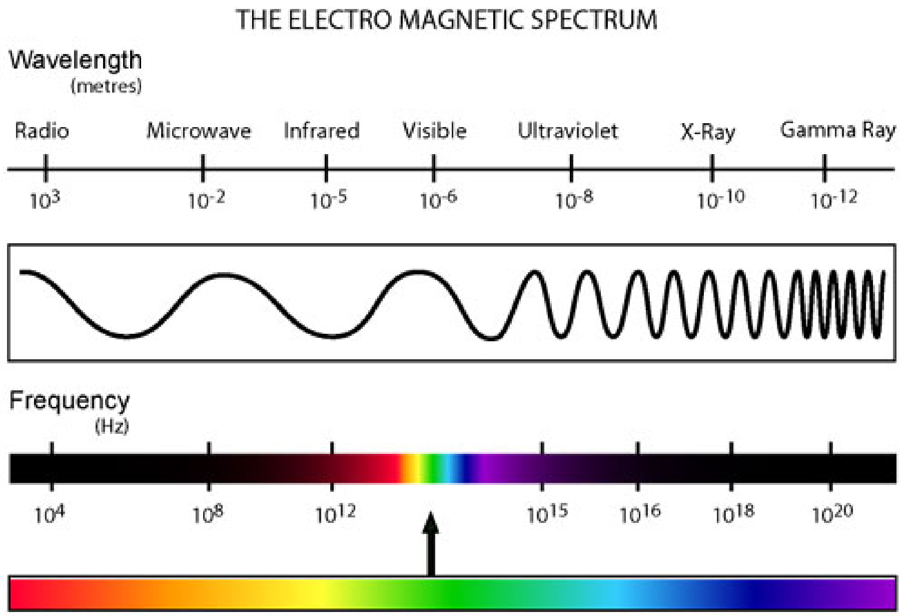


Figure 88: Electromagnetic spectrum from wikipedia. Electromagnetic waves of different wave-length interact with matter differently and also have different applications.

proportional to the square of the electric field (see eq.(15.4)) and therefore:

$$\frac{I_{\text{out}}}{I_{\text{in}}} = \cos^2\theta \tag{15.5}$$

15.3.1 Index of refraction

An important fact about light propagating in a medium is that its speed c_m is slower that c the speed of light in vacuum. The ratio is called the index of refraction:

$$n = \frac{c}{c_m} \tag{15.6}$$

The change in speed gives rise to a phenomenon known as refraction. A wave crossing a medium interfaces is bent. The change in angle is given by Snell's law:

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{15.7}$$

where $n_{1,2}$ are the index of refraction of the two media and the angles are defined in figure 90. This is nicely illustrated in a demo where we can also see that for a given angle the light emerges from the liquid parallel to the interface and for larger angles it does not emerge at all, a phenomenon known as total reflection.

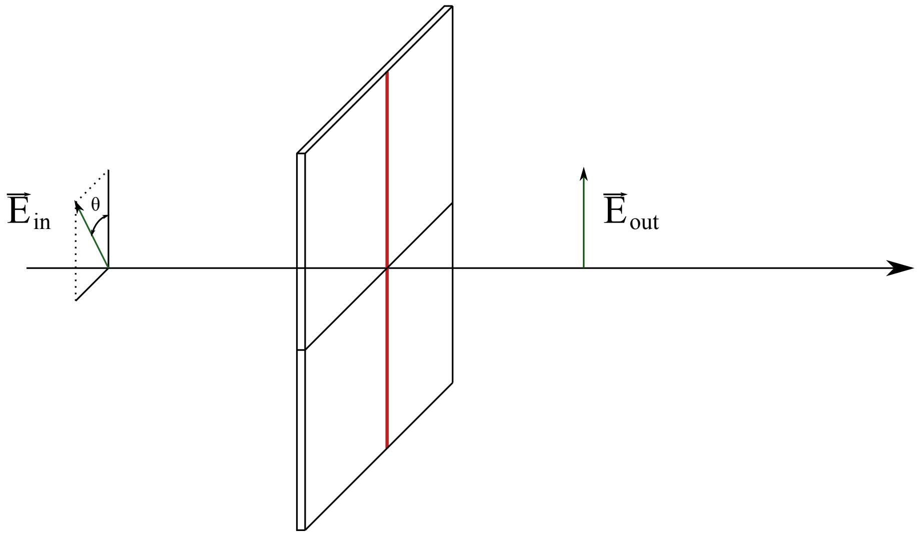


Figure 89: A polarizer lets through only the component of the electric field parallel to a preferred direction (indicated with red).

15.3.2 Fermat's principle

A problem that works in a similar way to the refraction of light is the following. Looking at fig. 92 suppose you need to reach a buoy some distance from the shore in the shortest possible time. If you are in the water you swim straight to it since a straight line is the shortest path. If you are in-land however, you first need to run to shore and since you run faster than swim it is better to take a somewhat longest path on land if it cuts your swimming leg. It turns out that, as pointed out by Fermat already in the sixteen hundreds, the path of least time is such that Snell's law is obeyed. Therefore, if we assume that a ray of light takes always the path of least time then we can derive Snell's law as suggested by Fermat. On small correction is that in certain cases a ray of light would take actually take also the path of largest time. These are generically called extremal paths. A more sever objection is that such principle seems more appropriate to particles rather than waves. We will see later on how to derive Snell's law using that light is a wave. Nevertheless Fermat's principle was a great idea and had a profound impact physics. In fact, most of modern physics is based in similar extremal principles. That is for a certain physical system one finds a quantity called the action such that the system always follows the paths of extremal action.

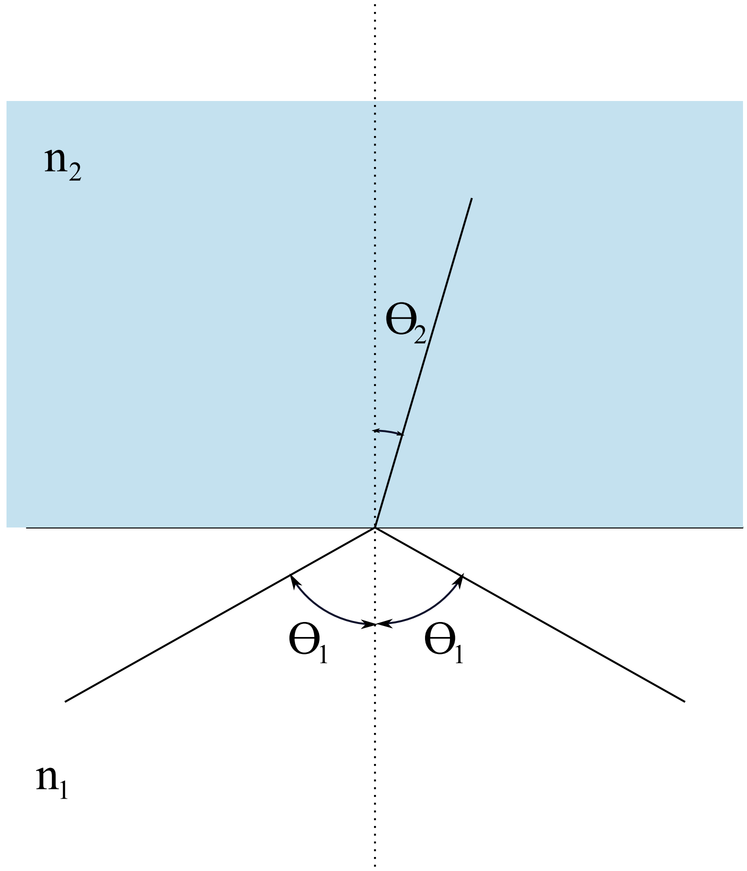


Figure 90: Refraction of light in the interface between two media. Snell's law determines that $n_1 \sin\theta_1 = n_2 \sin\theta_2$. Part of the light is also reflected and the angle of reflection is the same as the angle of incidence θ_1.

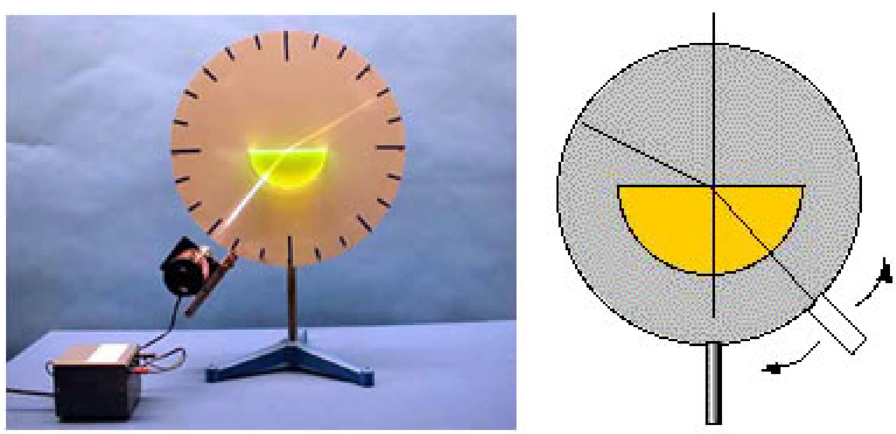


Figure 91: Demo: Refraction of light in the interface between two media. Illustrates Snell's law.

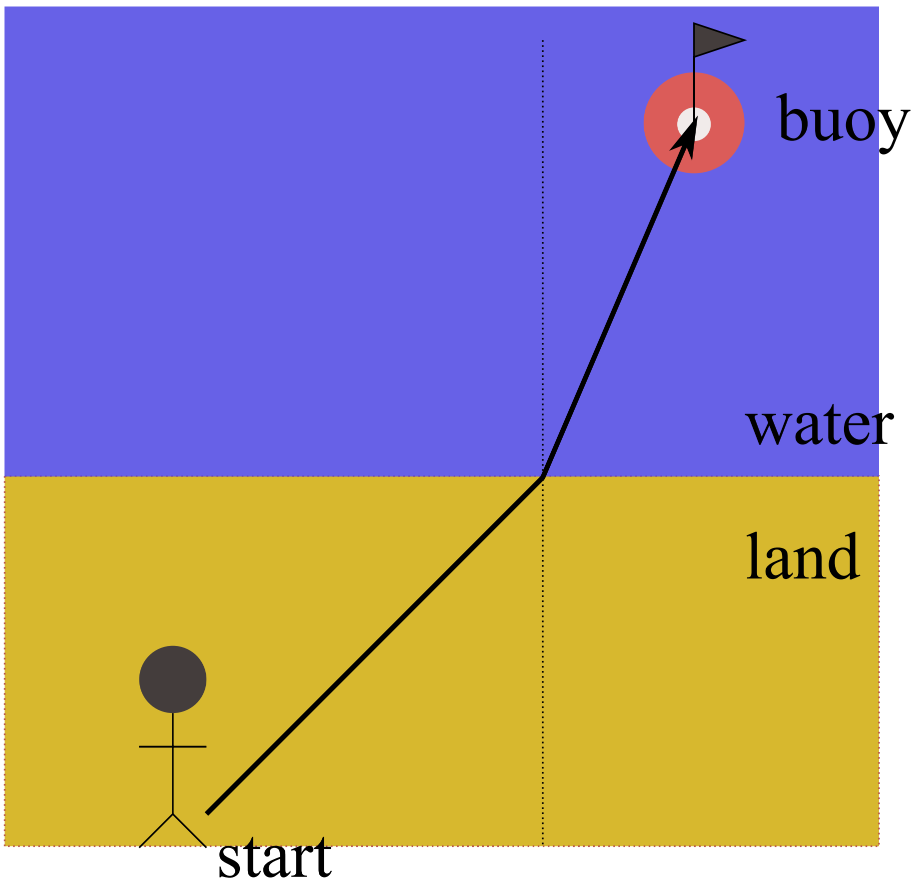


Figure 92: Fermat's principle can be illustrated with a simple consideration. If a person needs to get to a buoy in the water and part of the path is on land where he/she can run a fast and part in the water which slows them down, which is the optimal path?. It turns out that the optimal path is given by Snell's law!.

16. Lecture 16

16.1 Refraction

Let's have a a look at Snell's law

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{16.1}$$

from the point of view of waves. First we can define a wave front as the surface where the waves have maximum amplitude. Such wave front moves with the speed of light. A wave-length behind it there is another wave front where the amplitude is maximum etc., there are in fact many wave fronts. When the wave reaches the interface (see fig.93, each part of the front does so at a different time. Therefore, two arbitrary points A,B separated by a distance L on the interface are not in phase, namely if A is a maximum B is not. In fact B is behind A by a time Δt determined by

$$c_1 \Delta t = L \sin\theta_1 \tag{16.2}$$

To see that we can use some geometry to prove that the angle of incidence θ_1 is the same as the other angle labeled as θ_1 in the figure. On the other hand, the same oscillations of the field are seen from the other side of the interface. This implies that the phase difference between A and B given by Δt has to also be equal to

$$c_2 \Delta t = L \sin\theta_2 \tag{16.3}$$

From here we derive that

$$\frac{L}{c_1} \sin\theta_1 = \frac{L}{c_2} \sin\theta_2 \tag{16.4}$$

and, canceling L and multiplying by c the speed of light in the vacuum we derive that

$$\frac{c}{c_1} \sin\theta_1 = \frac{c}{c_2} \sin\theta_2 \quad \Rightarrow \quad n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{16.5}$$

which is Snell's law. Notice that part of the light is always reflected form the interface. In the case where $n_1 > n_2$ (for example light going from water to air) we see that

$$\sin\theta_2 = \frac{n_1}{n_2} \sin\theta_1 \tag{16.6}$$

Since $\frac{n_2}{n_1} < 1$ there is a critical angle $\theta_1 = \theta_c$ such that

$$\sin\theta_c = \frac{n_2}{n_1} \tag{16.7}$$

In that case we get $\theta_2 = \frac{pi}{2}$ namely the transmitted light emerges parallel to the interface. For angles $\theta_1 > \theta_c$ there is no value of θ_2 that satisfies Snell's law because $\sin\theta_2 \leq 1$ (sine is always smaller than one). In that case we have a phenomenon known as total reflection, all the light is reflected and nothing is transmitted. You can verify that when you go to a swimming pool and look up while being under water. There is am angle beyond which you cannot see outside.

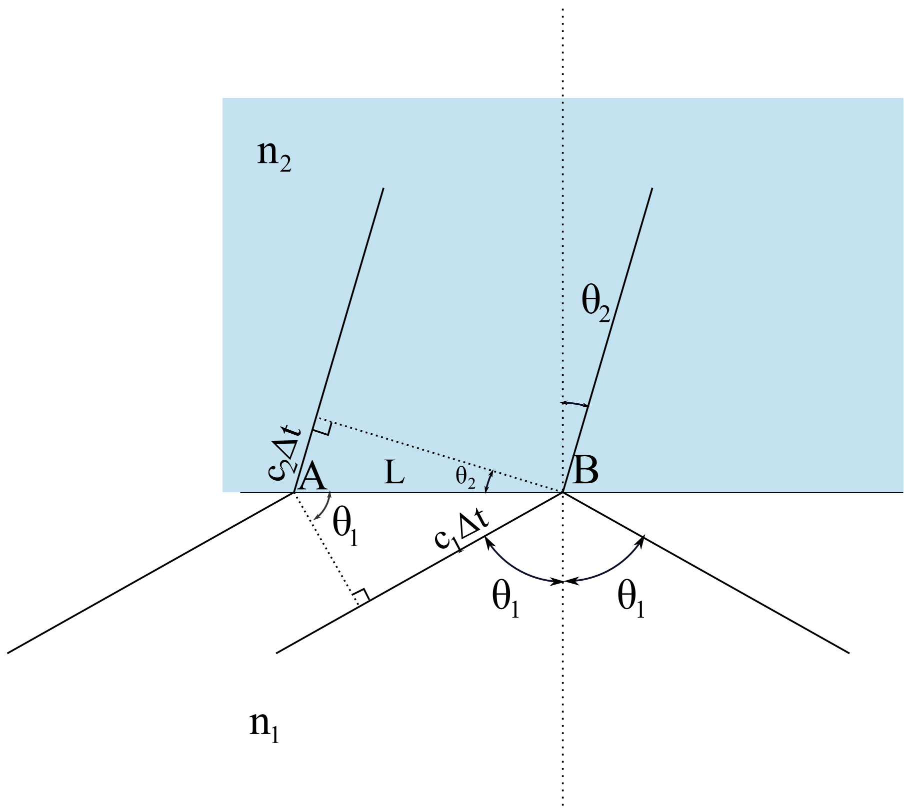


Figure 93: Refraction of light in the interface between two media. Snell's law can be derived from the wave nature of light. The dotted lines represent wave fronts.

16.2 Mirrors

A mirror can be thought as an interface that reflects all the incoming light, that is, nothing is transmitted. There are several types of mirror characterized by their shapes.

16.2.1 Flat mirror

Flat mirrors are the most common ones. As we already discussed, the reflected ray forms the same angle with the normal as the incident ray. This follows form the same principles used to derive refraction. if we look at the front wave now the speed of light is the same so the angle is the same. A consequence of this is that, as we all know, a flat mirror gives an image identical to the original only that is reversed front to back. The image appears to be at the same distance behind the mirror as the original object in front. To see why this is so we can look at figure 94. If you look at the reflected rays originating from a point A, they appear to come from point A'. So you interpret this information as if there were another object in point A' although of course there is none. This method to analyze the formation of images is applied to more complicated cases as we discussed now.

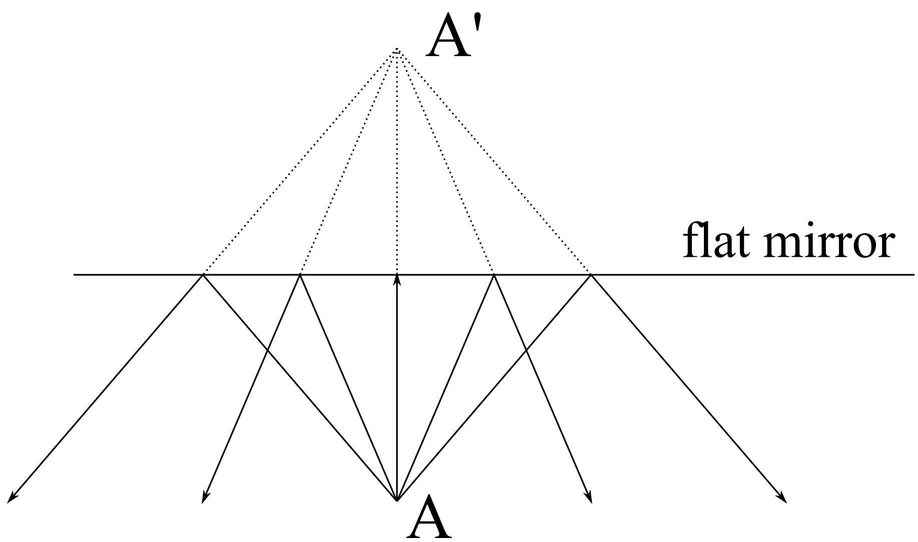


Figure 94: Reflection of light by a flat mirror. The reflected rays appear to emanate from the point A' which is the image of point A.

16.2.2 Concave mirror

A concave mirror has the shape of a section of a spherical surface. They are characterized by R the radius of the corresponding sphere (see fig. 95). The center of the sphere is called the center of curvature of the mirror and R the radius of curvature. If we look at what happens for a ray that originates in point A and moves parallel to the horizontal axis in the same figure, we see that, after being reflected it intersects the horizontal axis at a point which approximately $\frac{R}{2}$ away from the mirror. Such point is called the focal point. The proof of that is that, from the figure we have $\alpha = 2\theta$ and then

$$\sin\theta = \frac{h}{R} \tag{16.8}$$

$$\sin 2\theta = \frac{h}{\ell} \tag{16.9}$$

Now we have to use an important formula valid when an angle is small (in radians):

$$\sin\theta \simeq \theta \quad \text{when} \quad \theta \ll 1 \tag{16.10}$$

Using this we find

$$\left.\begin{array}{l} \theta \ \simeq \frac{h}{R}, \\ 2\theta \simeq \frac{h}{\ell} \end{array}\right\} \Rightarrow \quad \ell = \frac{R}{2} \tag{16.11}$$

From the triangle drawn in the figure we find that the focal point is at the same distance ℓ from the center of curvature. Therefore the focal point is at a distance $f \simeq \frac{R}{2}$ from the mirror. In the approximation of small angles all ray parallel to the horizontal axis are reflected toward the focal point.

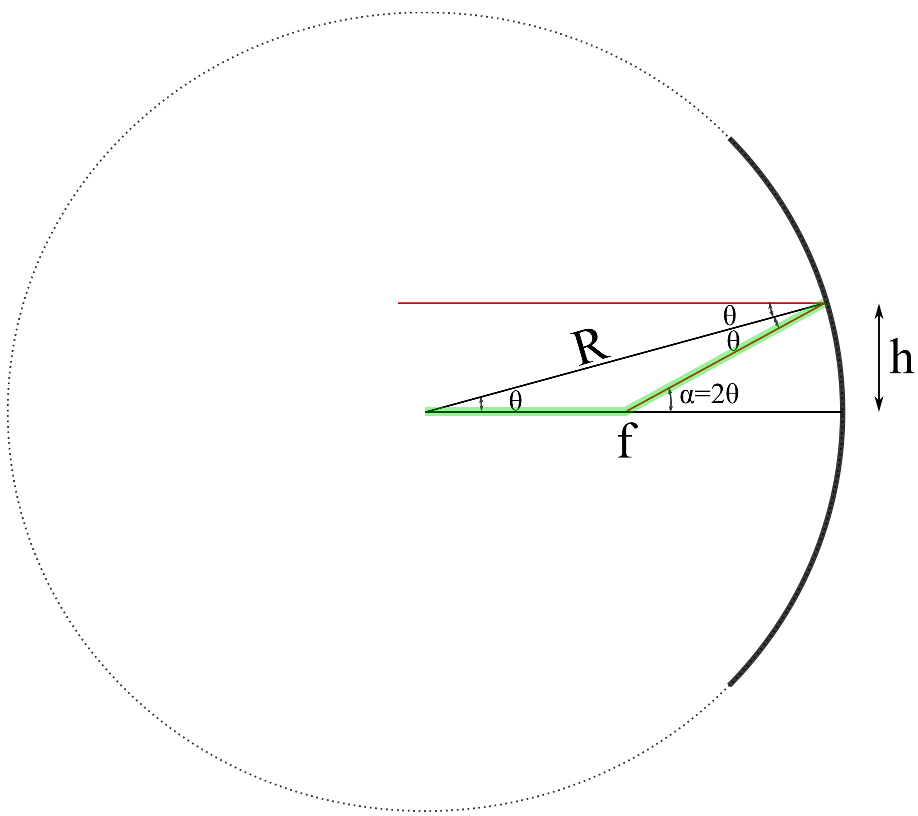


Figure 95: Reflection of light by a concave mirror. A ray parallel to the horizontal axis is reflected toward the focal point at a distance $f = \frac{R}{2}$ of the mirror. The two segments highlighted in green have the same length (denoted as ℓ in the text).

Figure 96: Demo: Reflection of light by a concave mirror.

Figure 97: Demo: an interesting effect of real images. The image of a small object can be easily confused with the real thing.

17. Lecture 17

17.1 Concave mirror

Continuing with our study of the concave mirror we can find now where the image is formed. In the approximation of small angles we are using, all the rays of light emanating from the object that are reflected in the mirror intersect at a point which is the position of the image. To find such position we need to us at least two rays. This is illustrated in fig.98. One is the ray that is parallel to the horizontal axis that is reflected toward the focal point. Another one is the one that goes through the focal point which is reflected parallel to the horizontal axis. This is so because the path of the light can always be backtracked, so if it arrives parallel then goes to the focus implies that if it comes from the focus it reflects parallel. Where these two rays intersect is where the image is located. Another ray that can be used is the one that hits the center of the mirror which reflect with the same angle as it arrives. This is because at that point the normal to the mirror is horizontal so a ray hitting the center behaves as if it were reflecting from a vertical flat mirror. Just to be clear when we say "horizontal" or "vertical" we are assuming that the axis of the mirror is the horizontal direction in the picture. From the figure we see that, in this case, the image is real, namely it is formed in front of the mirror. On the other hand, looking at figure 99 we find that, if the object is closer to the mirror than the focal distance f, then the image is virtual. In the next section we will be more precise and find a mathematical equation that determines the position and height of the image.

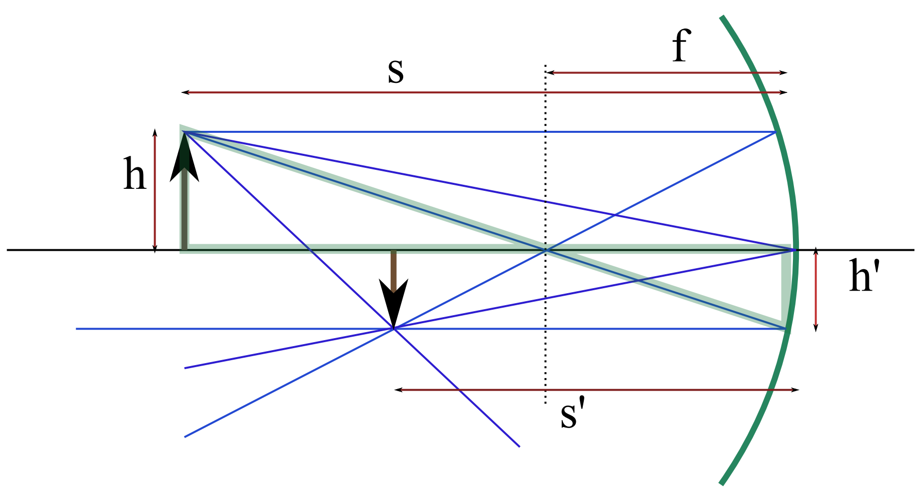


Figure 98: Image of an object by a concave mirror when the object is further than the focal point. Follow the rays and see how the image is constructed. The highlighted triangles are similar to each other and used in deriving the mirror equation.

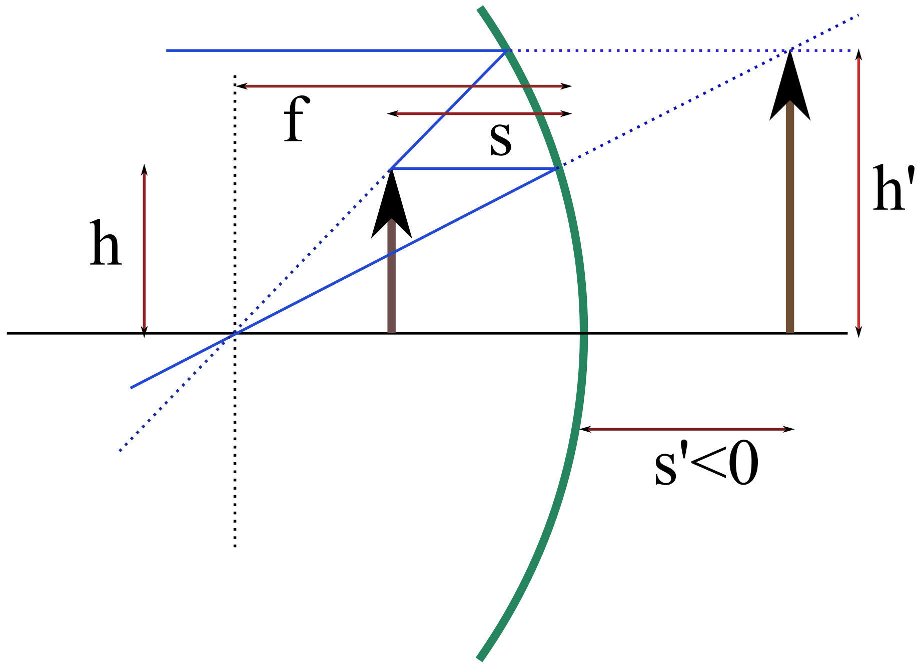


Figure 99: Image of an object by a concave mirror when the object is closer than the focal point.

17.2 Mirror equation

Consider the case of the concave mirror that we have already discussed and illustrated in fig.98. If we put an object of height h at a distance s from the mirror we are interested in finding the height h' and position s' of the image. Before starting let us discuss how we set up the sign conventions. We are going to take $s' > 0$ if the image is formed on the same side as the object (real image) and $s' < 0$ if it is formed behind the mirror (virtual image). Furthermore, if the image is upright we take $h' > 0$ and if it is inverted we take $h' < 0$.

We can now proceed to find the position and height of the image. From fig.98, using that the highlighted triangles are similar (namely they have the same angles and their sides are proportional) we find that:

$$\frac{h}{s-f} = \frac{-h'}{f} \tag{17.1}$$

Here we took into account that $h' < 0$ so the length of the corresponding side of the triangle is actually $-h'$. If we do the same with the image we find another equality:

$$\frac{-h'}{s'-f} = \frac{h}{f} \tag{17.2}$$

which actually is the same as the previous equation after interchanging $s \leftrightarrow s'$, $h \leftrightarrow -h'$. From eq.(17.1) we find

$$h' = -\frac{hf}{s-f} \tag{17.3}$$

replacing this value of h' in the eq.(17.2) we find:

$$\frac{hf}{(s'-f)(s-f)} = \frac{h}{f} \quad \Rightarrow \quad \frac{hf}{s-f} = \frac{h(s'-f)}{f} \quad \Rightarrow \quad s' = \frac{f^2}{s-f} + f \tag{17.4}$$

Taking common denominator in the expression for s' we finally find

$$s' = \frac{sf}{s-f} \tag{17.5}$$

$$h' = -\frac{hf}{s-f} \tag{17.6}$$

where we included eq.(17.3). These two equations completely determine the position and size of the image. Although we derived them for the case where $s > f$, it is straightforward to check they also work when $s < f$. It is conventional to rewrite them in a way that makes more evident the symmetry between the object and the image:

$$s' = \frac{sf}{s-f} \quad \Rightarrow \quad \frac{1}{s'} = \frac{s-f}{sf} = \frac{1}{f} - \frac{1}{s} \quad \Rightarrow \quad \frac{1}{s} + \frac{1}{s'} = \frac{1}{f} \tag{17.7}$$

For the height we have:

$$\frac{h'}{s'} = -\frac{hf}{s-f}\frac{s-f}{sf} = -\frac{h}{s} \tag{17.8}$$

So we get the equivalent equations:

$$\frac{1}{s} + \frac{1}{s'} = \frac{1}{f} \tag{17.9}$$

$$\frac{-h'}{s'} = \frac{h}{s} \tag{17.10}$$

17.3 Convex mirror

A convex mirror is shaped as the exterior of a sphere. Rays which arrive parallel to the axis are divergent after being reflected. They appear to originate from a point behind the mirror called the focal point. On the other hand, a ray that is directed toward the focal point will be reflected parallel. In fig.100 we see that in this case the image is always virtual and upright. With some algebra and using the same ideas as before we can derive that the convex mirror obeys the same mirror equation as the concave one with the only change being that now $f < 0$.

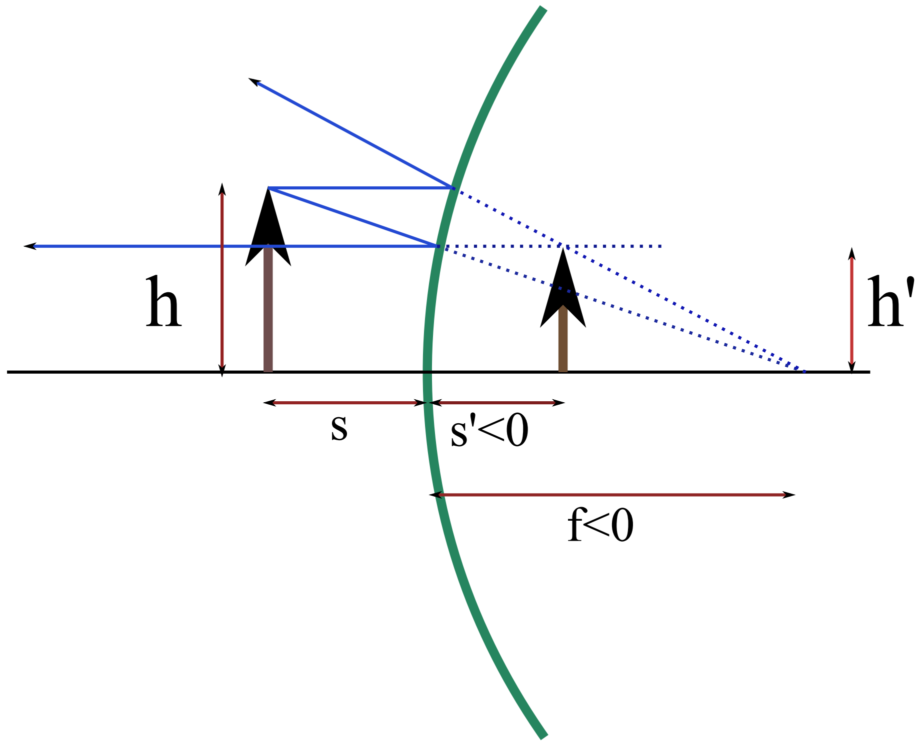


Figure 100: Image of an object by a convex mirror. The image is always virtual and upright. In this case $f < 0$.

17.4 Convergent lens

Another important optical component is the lens. In the ideal situation, rays arriving parallel to the axis of the lens are transmitted and converge into a single point called the focal point. The lens works using the laws of refraction. Typically it is made of glass or a similar material. The rays of light are refracted twice, once in each surface of the lens and that produces the desired effect. As with the mirror, the behavior of the lens is only approximate and valid for rays which are not too far from the axis and/or arrive at small angles. If the rays arrive from the other side of the lens they will converge into another focal point situated at the same distance as the one we already discussed but on the other side of the lens. Furthermore, if a ray goes through a focal point, it will become parallel to the axis after crossing the lens. Finally, a ray going through the center of the lens is transmitted right through. The rules to construct images are similar as with the mirrors. In the figures 101 and 102 we see two cases, one where the object is far from the lens and the other when it is closer that f.

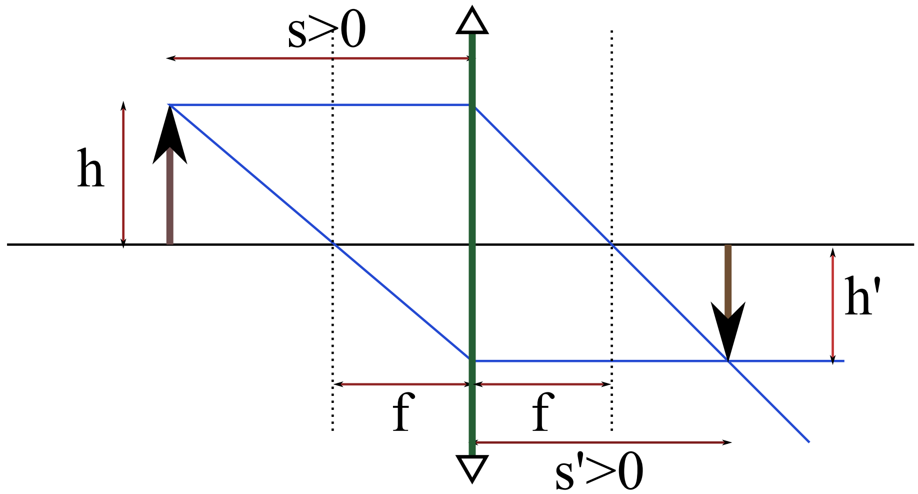


Figure 101: Image of an object by a convergent lens when the object is further than the focal point.

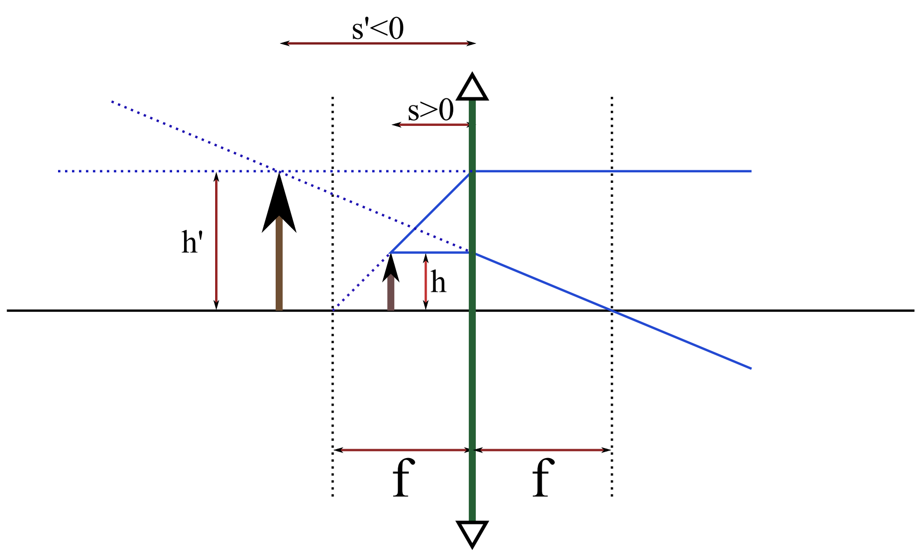


Figure 102: Image of an object by a convergent lens when the object is closer than the focal point.

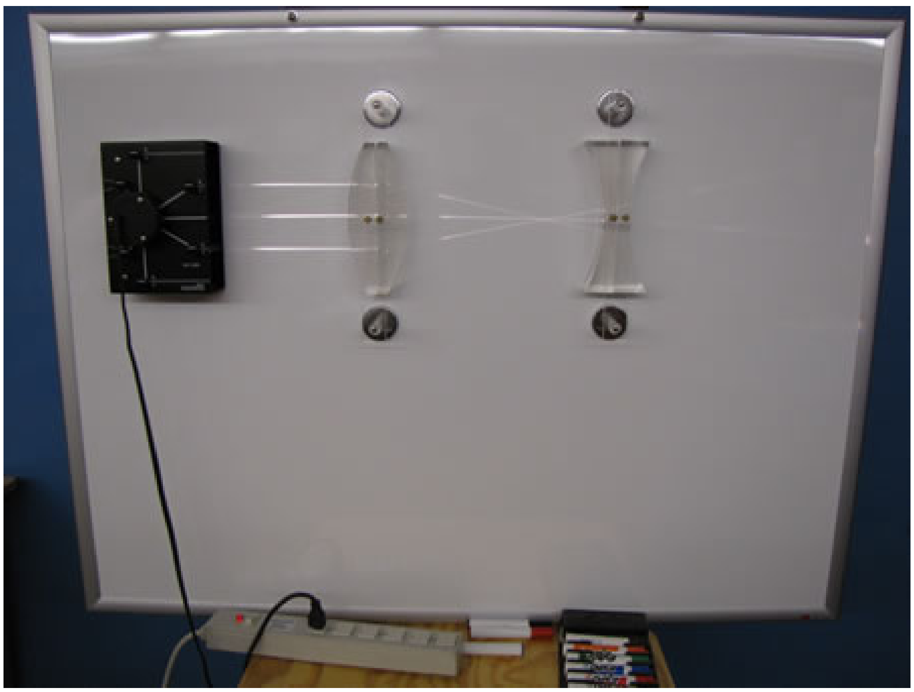


Figure 103: Demo illustrating the properties of lenses

18. Lecture 18

18.1 Lens equation

Going back to the convergent lens, we can once again find an equation that determines the position and size of the image given the position and size of the object. The sign convention for the image's position is now that $s' > 0$ if the image is on opposite sides of the lens as the object and $s' < 0$ if the image is on the same side of the object. This makes sense because we look at the lens from behind whereas we looked at the mirror from the front. That is a real image has $s' > 0$ in both cases. With these conventions and looking at figure 101 we can derive, similarly as for the mirror:

$$\frac{h}{s-f} = \frac{-h'}{f} \tag{18.1}$$

and also

$$\frac{-h'}{s'-f} = \frac{h}{f} \tag{18.2}$$

Since the equations are the same as for the mirror we can immediately derive

$$s' = \frac{sf}{s-f} \tag{18.3}$$

$$h' = -\frac{hf}{s-f} \tag{18.4}$$

or equivalently:

$$\frac{1}{s} + \frac{1}{s'} = \frac{1}{f} \tag{18.5}$$

$$\frac{-h'}{s'} = \frac{h}{s} \tag{18.6}$$

It is easy to see that once again these equations are valid if $s > f$ or $s < f$.

18.2 Divergent Lens

When rays arrive parallel to the axis into a divergent lens they become divergent and seem to originate from a focal point situated on the same side of the lens from where the rays arrived. As seen in fig. 104 the image is always virtual and up-right. The equation for the lens can be applied to this case if we take $f < 0$.

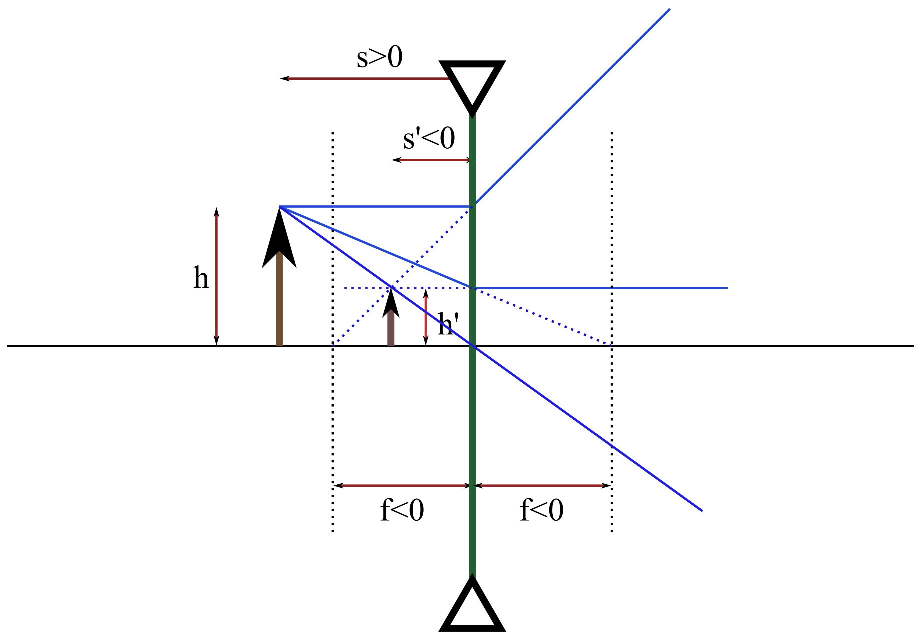


Figure 104: Image of an object by a divergent lens. It is always virtual and upright.

18.3 Camera, microscope, telescope

There are numerous application of lenses and mirrors. The photographic camera is one of them. It actually works on the same principle as our eye. It creates an image on a surface sensitive to light. The image is then read electronically or chemically by producing a reaction on a film. The minimal setup is a convergent lens which creates a real image as in fig.101. The screen is positioned at the point where the image is formed.

Another application is the microscope. The minimal setup in this case is a convergent lens as in fig.102. Actually this is also the usual magnifying glass. Actual microscopes have more lenses that increase the magnification but work on the same principle.

Finally we have the telescope which is used to look at objects far away. When a point is far away, the rays coming from such point, for all practical purposes are parallel to each other. On the other hand if we have an extended object far away, and the object is large enough, the rays originating from different points of the object arrive with different angles. In that way we can talk about the angular size of an object which is the difference in angles between rays arriving from opposite edges of the object. As seen in figure 105, out of parallel rays, a convergent lens forms an image on its focal plane. This is how our eyes form images for objects at infinity. The angular size of the object is determines the size of the image on our retina. In that way we see for example the Moon as an object which has an apparent size which does not seem too big compared to everyday objects even if we know it is much larger than any object we can handle. The reason of course being that the Moon is much further away that any everyday object.

If, as in figure 105 we introduce another lens whose focal point coincides with the one of the first lens, the rays will once again be parallel but now the angular size will be different. We want the angular size to be larger so the object will appear larger and we will be able to see more details. From fig.105 we can derive that

$$\tan\theta = \frac{h}{f}, \quad \tan\theta' = \frac{h}{f'} \tag{18.7}$$

where h is the height of the image created by the first lens and f, f' are the focal distances of both lenses. The angles θ and θ' are the angles at which a given ray arrives and then emerges from the telescope. Since we always consider small angles we have

$$\tan\theta \simeq \theta, \quad \tan\theta' \simeq \theta' \tag{18.8}$$

which implies

$$\text{magnification} = \frac{\theta'}{\theta} = \frac{f}{f'} \tag{18.9}$$

which is defined as the magnification of the telescope. If we have $f' < f$ then we also have $\theta' > \theta$ as we wanted.

18.4 Aberrations

We insist once again that all the equations and image formation we nave studied for mirrors and lenses are approximations for small angles. Sometimes these are called the laws of ideal mirrors and lenses whereas actual ones do not behave exactly the same. The differences between the ideal and actual behavior are called aberrations. Let us discuss two simple ones. First, we said that a spherical mirror forms a sharp image only if the rays arrive close to the axis. namely the mirror is small compared to the radius of curvature. This is illustrated in fig.106 where we see that the rays emanating from the object do not all cross at the same point. For rays reaching the mirror further from the axis this is evident. This is called spherical aberration. If we put a screen or photographic plate at the position of the image the image will be blurred for such a big mirror. On the other hand, if we just take a look from the left with our own eyes this might not be so evident because our eyes are quite small and will only capture a few of the rays. Such rays would intersect at the image point and the image will appear sharp essentially because our eyes are taking advantage only of the central part of the mirror.

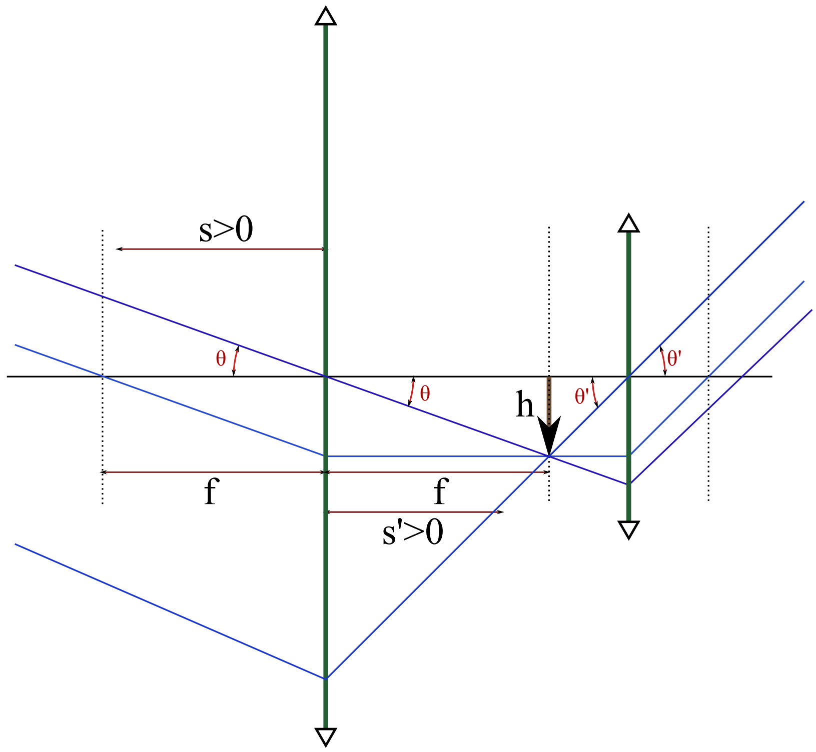


Figure 105: Simple refractive telescope. Two lenses increase the angular size of objects situated at infinity

Another kind of aberration is called color or chromatic aberration. It applies only to lenses. Although we did not discuss this in detail it turns out that the index of refraction is different for different colors as seen in a prism which separate the different colors from white light. A lens works similarly and therefore the focal point is slightly different for different colors. This distorts the images produced by lenses and needs to be corrected in some applications by putting several lenses such that their color aberrations cancel each other.

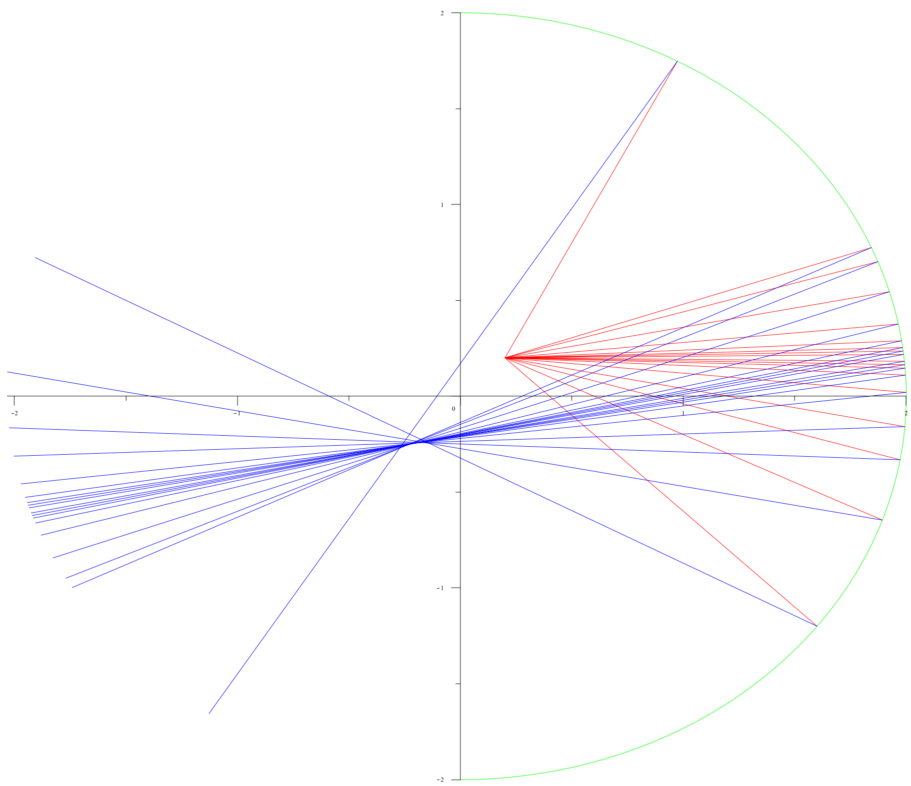


Figure 106: Spherical mirror. Images formed by a big mirror (compared with its curvature radius) are not sharp.

19. Lecture 19

19.1 Interference

The principle of superposition implies than when two electromagnetic waves meet each other, the total electric and magnetic fields are the sum of the electric and magnetic field of each wave. This is quite important since, for example, the electric fields can add up or subtract depending on how they are oriented. When they are opposite and of the same strength they cancel each other and we have the very surprising property that two beams of light can cancel each other and give a dark spot. This is called destructive interference (whereas constructive is when they add up). However it does not seem to be something we see everyday!. In fact we know that two light bulbs always illuminate more than one. So what is going on?. The situation is that in a light bulb the atoms of the filament are emitting light independently, so from each light bulb we have an enormous number of pulses or wave packets that are not in phase with each other. So wave packets from different sources sometimes add up and sometimes cancel each other. In average the intensities just add up. Notice that this is not constructive interference. In constructive interference the electric fields add up. But the intensity is proportional to the square of the electric field so two beams of equal intensity give through constructive interference a beam of four times the intensity. On the other hand random interference gives just twice the intensity.

In fig.107 we see the case of two wave packets interfering constructively, destructively or an intermediate case.

The question is now if we can actually see interference and cancel light with light. The trick is to split each wave packet in two parts and then join them back so essentially each packet interferes with itself. How to do so we know from when we studied refraction, an incident beam is split into a reflected beam an a transmitted beam. This is the principle used in the interferometer.

19.1.1 Interferometer

The interferometer (see fig.108) works by splitting a beam into two using a partially reflecting mirror (drawn in blue). The two beams are reflected from flat mirrors (drawn in green) and joined back together using the same partially reflective mirror. Part of the beam will be sent back to the source and lost. After the two beams are joined they are projected onto a screen. If the difference in the length of the two arms of the interferometer is an integer multiple of the wave-length then the interference in constructive. If it is a half-integer multiple is destructive. Otherwise we have an intermediate situation. The trick is that all wave packets emitted by the source interfere in the same way provided that the source is monochromatic, namely has a single wave

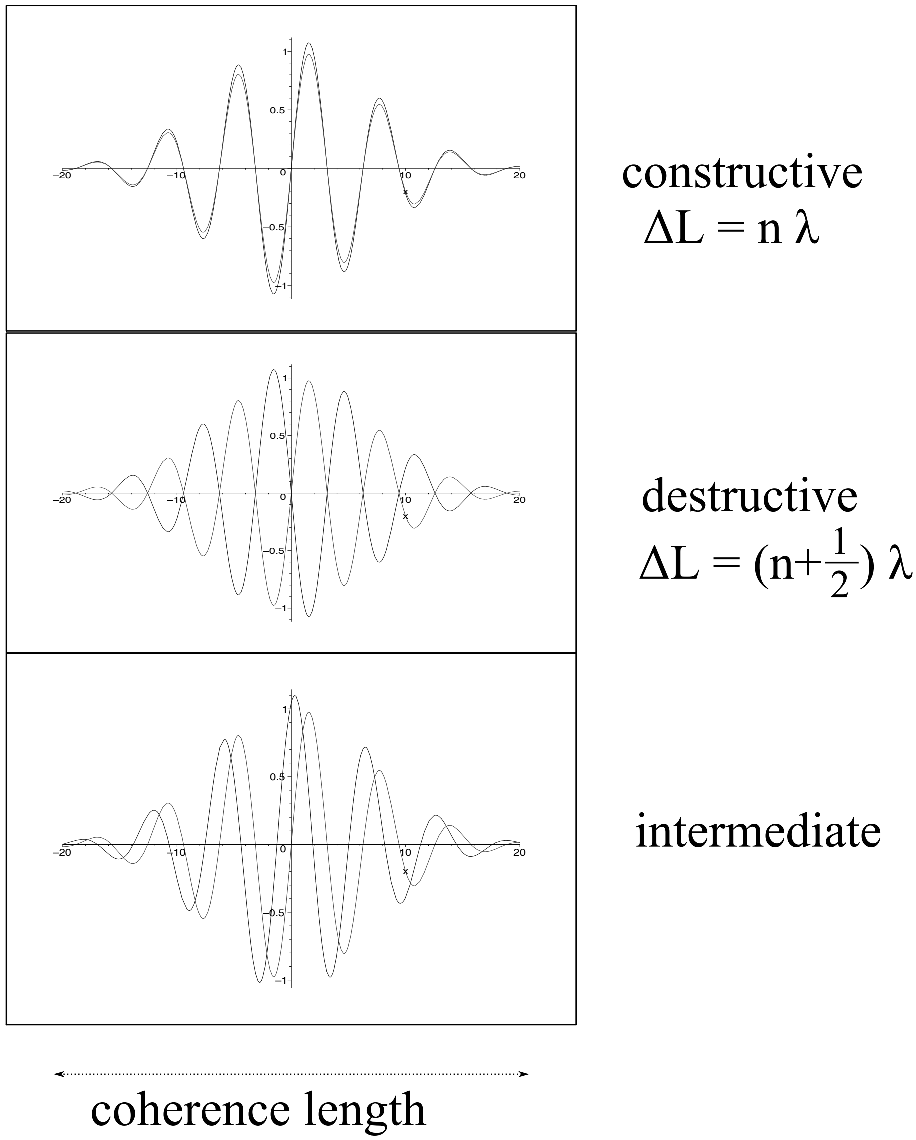


Figure 107: Two wave packets can interfere constructively or destructively. Also there are many intermediate situations as illustrated in the third plot.

length. It is also important that the difference in length between the two arms of the interferometer is smaller than the typical length of the wave-packets (also known as coherence length). Otherwise a packet cannot interfere with itself. To summarize:

$$\Delta L = \begin{cases} m\lambda & \rightarrow \quad \text{Constructive} \\ \left(m+\frac{1}{2}\right)\lambda & \rightarrow \quad \text{Destructive} \end{cases} \tag{19.1}$$

where m is an arbitrary integer. The wave length for visible light is of the order of 500nm so by moving the mirror so slightly one can go from constructive to destructive interference. Not only is the interferometer a clear proof of the wave-like nature of light but it also provides a very sensitive instrument to measure length. Although this is not practical to use in everyday life, for certain physics experiments it is very useful. One such experiment which is under way is the detection of gravitational waves. Such waves would make the mirrors move and such motion can be detected. It requires great care since, for example, even small vibrations that change the position of the mirrors are enough to destroy the interference pattern.

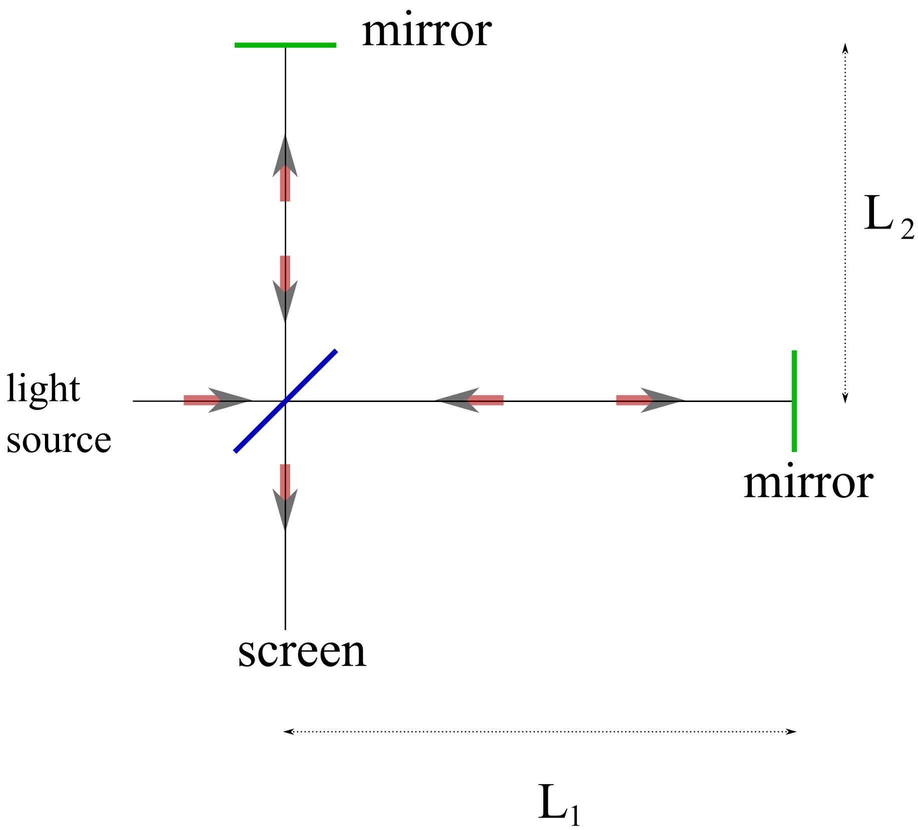


Figure 108: Interferometer.

19.1.2 Thin films

Other situation where interference appears and is actually very easy to see is in thin films. For example soapy water on glass can form a thin film. It is known that when illuminated with white light such films appear colored as a rainbow (e.g. soap bubbles). The reason is that we have interference as shown in fig.109. The two reflected beams will interfere constructively or destructively. The difference in path length is

$$\Delta L = \frac{2d}{\cos\theta'} \tag{19.2}$$

where d is the thickness of the film and θ' is the angle of refraction (given by Snell's law). There is a subtlety which is that, as we are going to see later, when a beam is reflected from an interface there could be a phase shift of 180^o equivalent to a half-wave length shift. The rule is that when coming from medium 1 and reflecting from an interface with medium 2:

$$n_1 < n_2 \;\rightarrow\; 180^o \text{ phase shift} \tag{19.3}$$

$$n_1 > n_2 \;\rightarrow\; \text{no phase shift} \tag{19.4}$$

$$\tag{19.5}$$

where $n_{1,2}$ are the indices of refraction of the two media. The other point to take into account is that the wave-length changes when going to a different medium. The wave length is related to the frequency through

$$\lambda = \frac{c}{f} \tag{19.6}$$

where c is the speed of light in the medium and f is the frequency which is the same in all regions (since the boundary conditions enforce that the electric and magnetic field oscillate in unison in all media). If we use λ_1 the wave-length in medium 1 as a reference, the wave-length λ_2 in medium two is given by

$$\lambda_2 = \frac{c_2}{c_1}\lambda_1 = \frac{n_1}{n_2}\lambda \tag{19.7}$$

since the speed of light is inversely proportional to the refraction index. Of medium 1 is air we take $n_1 = 1$ and obtain

$$\text{If } n_1 < n_2 < n_3, \text{ or } n_1 > n_2 > n_3 \text{ then} \quad \begin{cases} \Delta L = \frac{2d}{\cos\theta'} = m\frac{\lambda}{n_2} & \rightarrow \quad \text{Constructive} \\ \Delta L = \frac{2d}{\cos\theta'} = \left(m + \frac{1}{2}\right)\frac{\lambda}{n_2} & \rightarrow \quad \text{Destructive} \end{cases}$$

$$\text{If } n_1 < n_2 > n_3, \text{ or } n_1 > n_2 < n_3 \text{ then} \quad \begin{cases} \Delta L = \frac{2d}{\cos\theta'} = m\frac{\lambda}{n_2} & \rightarrow \quad \text{Destructive} \\ \Delta L = \frac{2d}{\cos\theta'} = \left(m + \frac{1}{2}\right)\frac{\lambda}{n_2} & \rightarrow \quad \text{Constructive} \end{cases}$$

Since the angle at which there is constructive interference depends on the wavelength, different colors will be seen at different angles giving rise to the rainbow effect. The same happens for example for two flat glass surfaces separated by a thin layer of air. In fact the effect is more evident if the surfaces are not parallel but slightly at angle. Then dark and bright bands appear. An example is when a lens is positioned over a flat surface as seen in fig.111.

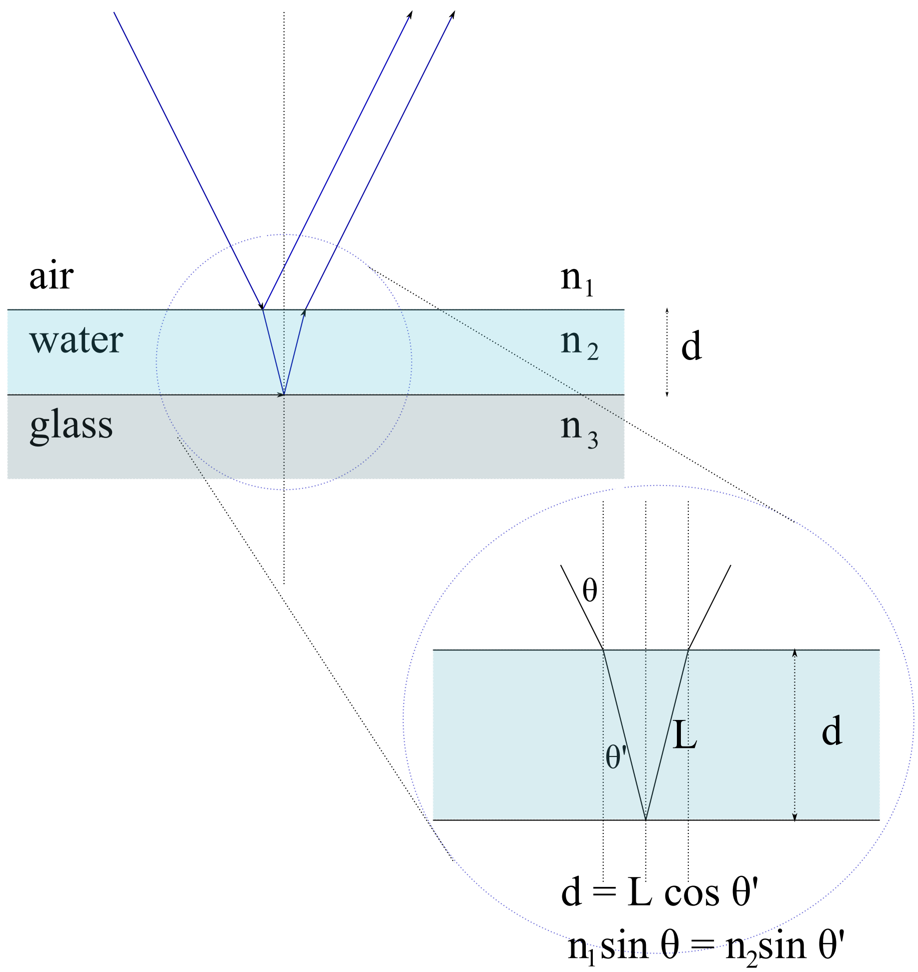


Figure 109: Thin-film interference. Care should be taken to include the extra 180^o phase shift introduce when reflecting form a material of larger index of refraction.

19.1.3 Two slits

Another experiment is the two slit experiment also knows an Young's experiment. Light is shone toward an opaque surface with two very thin slits close to each other. A screen on the other side shows alternative bright and dark bands produces by interference. From fig.112 we see that the bright bands appear whenever[2] :

$$\Delta L = d \sin\theta_M = m\lambda \tag{19.8}$$

where d is the distance between slits and θ_M is the angle at which we find a maximum. Also the integer m is known as the order of the maximum. The bright and dark bands are also called fringes.

[2]Remember that in these formulas it is conventional to use m to denote an arbitrary integer.

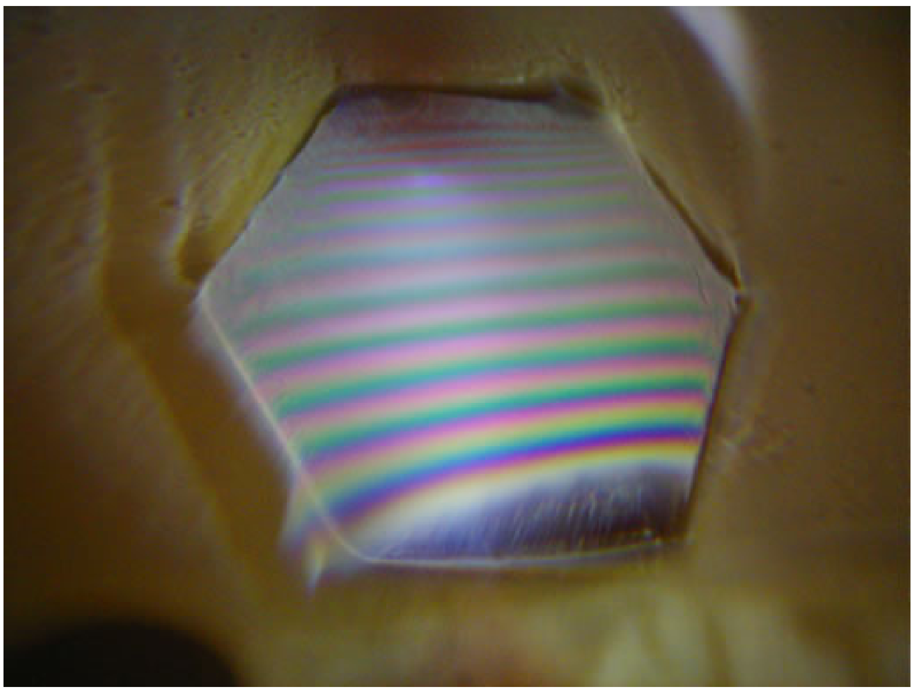


Figure 110: Demo: Thin-film interference. The rainbow colors are produced by interference maxima located at different positions for different wave-lengths.

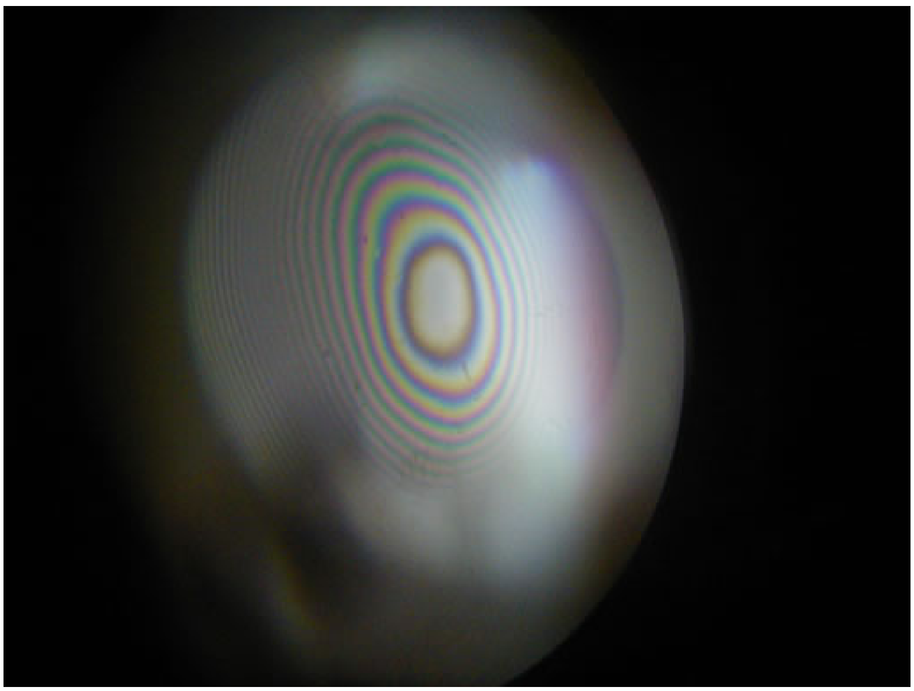


Figure 111: A thin film of air can be obtained between a lens and a flat specular surface. The interference pattern is known as Newton's rings.

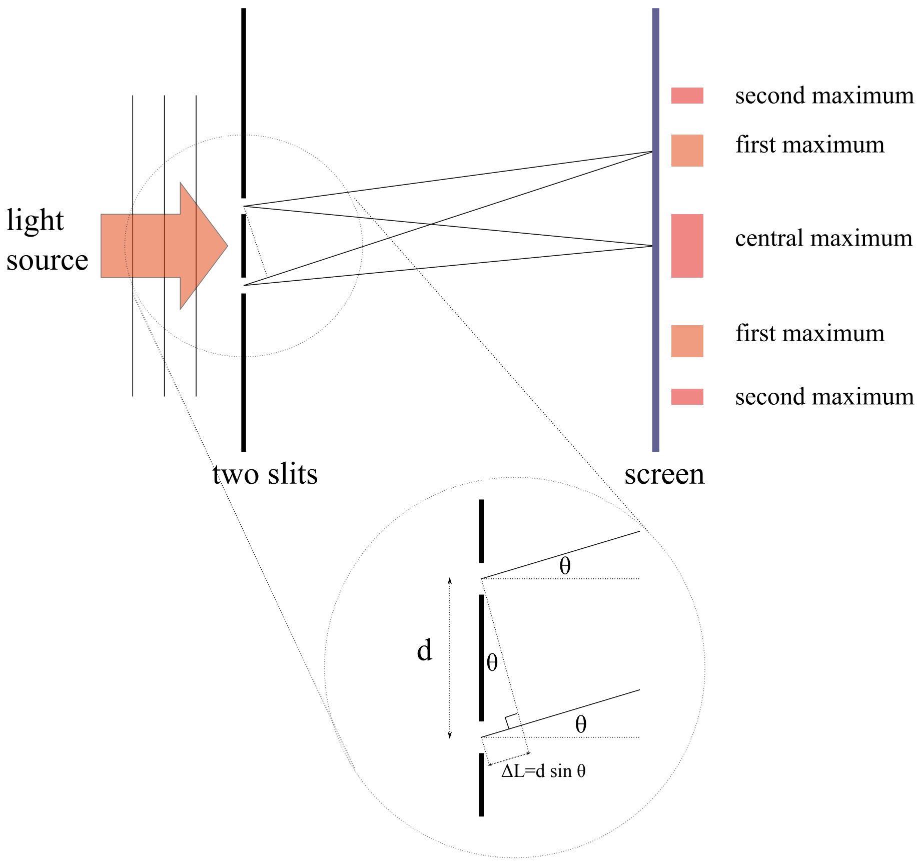


Figure 112: Young's experiment, interference between light coming from two slits illuminated by the same source.

20. Lecture 20

20.0.4 Fresnel Equations

In the previous lecture we discussed that light can experience a 180^0 phase shift when reflected from an interface. Here we are going to show that this follows from the properties of light as an electromagnetic wave. Typical media that we consider are air, water, glass, etc. which are dielectrics. Namely they are made of tiny dipoles that orient themselves with the electric field. In the presence of an electromagnetic wave they oscillate producing a slow down of light and the property of refraction. What is important for us here is that they generate a surface charge and therefore the normal component of the electric field can have a jump at the interface. On the other hand the component of the electric field parallel to the interface is continuous. Besides, the magnetic field is also continuous since the materials we consider are not magnetic and therefore generate no magnetic fields at the interface. To understand the calculation we draw fig.113 were we consider the case when the electric field is parallel to the interface in which case it has to be continuous:

$$E_i + E_r = E_t \tag{20.1}$$

that is, the electric field in medium 1 is the sum of the incident and reflected ones and has to be equal to the electric field in medium two which is the transmitted one. The magnetic field has two components and we have

$$-B_i \cos\theta + B_r \cos\theta = -B_t \cos\theta' \tag{20.2}$$

$$-B_i \sin\theta - B_r \sin\theta = -B_t \sin\theta' \tag{20.3}$$

Notice that the direction of the magnetic field is determined by the fact that $\vec{E} \times \vec{B}$ points in the direction of propagation of the wave. Moreover, the strength of the magnetic field is related to the electric field by the speed of light:

$$B_i = \frac{E_i}{c_1}, \quad B_r = \frac{E_r}{c_1}, \quad B_t = \frac{E_t}{c_2}. \tag{20.4}$$

Using eq.(20.4) to replace the magnetic field in eqns.(20.2) and (20.3) we get

$$-\frac{E_i}{c_1}\cos\theta + \frac{E_r}{c_1}\cos\theta = -\frac{E_t}{c_2}\cos\theta' \tag{20.5}$$

$$-\frac{E_i}{c_1}\sin\theta - \frac{E_r}{c_1}\sin\theta = -\frac{E_t}{c_2}\sin\theta' \tag{20.6}$$

Taking into account that the speed of light $c_{1,2}$ in each media are given in terms of the index of refraction by

$$c_{1,2} = \frac{c}{n_{1,2}} \quad \Rightarrow \quad \frac{c_1}{c_2} = \frac{n_2}{n_1} \tag{20.7}$$

we find that eq.(20.6) reduces to

$$n_1 \sin\theta\,(E_i + E_r) = n_2 \sin\theta' E_t \tag{20.8}$$

Comparing with eq.(20.1) we see that we need to satisfy

$$n_1 \sin\theta = n_2 \sin\theta' \tag{20.9}$$

which is Snell's law!. Therefore we are left with two equations:

$$E_i + E_r = E_t \tag{20.10}$$

$$-E_i \cos\theta + E_r \cos\theta = -\frac{n_2}{n_1} E_t \cos\theta' \tag{20.11}$$

namely eqns.(20.1) and (20.2). We have two equations and we need to compute two quantities, E_r and E_t, it is then just a matter of algebra. Before proceeding we can take the simplest case of normal incidence when $\theta = \theta' = 0$. This gives

$$E_i + E_r = E_t \tag{20.12}$$

$$-E_i + E_r = -\frac{n_2}{n_1} E_t \tag{20.13}$$

from where we derive

$$E_t = \frac{2n_1}{n_1 + n_2} E_i \tag{20.14}$$

$$E_r = \frac{n_1 - n_2}{n_1 + n_2} E_i \tag{20.15}$$

which in particular implies that, if $n_2 > n_1$ the reflected electric field is opposite to the incident one, namely a 180^o phase shift. Equivalent to shifting the wave by a half wave-length. If the incidence is not normal the same calculation gives for the reflected electric field

$$E_r = \frac{n_1 \cos\theta - n_2 \cos\theta'}{n_1 \cos\theta - n_2 \cos\theta'} E_i \tag{20.16}$$

$$= \frac{\sin\theta' \cos\theta - \sin\theta \cos\theta'}{\sin\theta' \cos\theta - \sin\theta \cos\theta'} E_i \tag{20.17}$$

$$= \frac{\sin(\theta' - \theta)}{\sin(\theta' + \theta)} E_i \tag{20.18}$$

where we used Snell's law and the identity $\sin(\alpha + \beta) = \sin\alpha \cos\beta + \cos\alpha \sin\beta$. We again derive that

$$\begin{array}{lllll} n_1 > n_2 & \Rightarrow \ \theta < \theta' & \Rightarrow \ E_r, E_i & \text{same sign} & \Rightarrow \ \text{no phase shift} \\ n_1 < n_2 & \Rightarrow \ \theta > \theta' & \Rightarrow \ E_r, E_i & \text{opposite sign} & \Rightarrow \ 180^o \text{ phase shift} \end{array} \tag{20.19}$$

Besides, from E_r and E_t we can also compute how much of the wave is reflected and how much transmitted. That this calculations agree with the experiment leave very little doubt that light is indeed an electromagnetic wave.

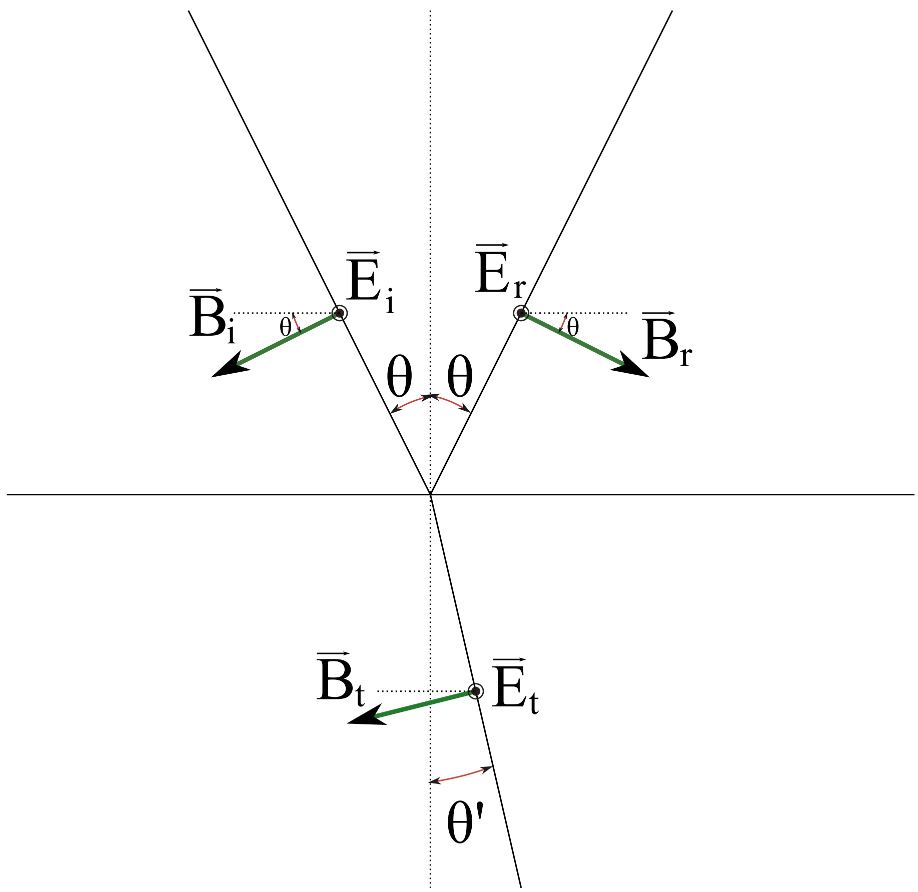


Figure 113: Refraction of an electromagnetic wave.

20.1 Gratings

A grating is a special device composed of a large number of slits parallel to each other. It works similarly to the two slits but it is very effective at separating light in its different constituent colors because the interference maxima are quite narrow. Typical gratings have hundreds of slits per millimeter. In fig.114 we see that the difference in path length between two neighboring rays is given by

$$\Delta L = d \sin\theta \tag{20.20}$$

If it is an integer number m of wave-lengths

$$d \sin\theta_M = m\lambda \tag{20.21}$$

then two consecutive paths interfere constructively. Moreover it is easy to see that any path will interfere constructively with any other one and then we will have a maximum.

It is also clear that if we change the angle slightly then consecutive paths might still interfere constructively but not paths originating from far apart slits because the path difference would be larger. This makes the peaks very narrow. This phenomenon is very important to decompose light in its component wave-lengths. This way we can study light coming from distant stars and understand their chemical composition. The reason is that each element absorbs particular wave-lengths and therefore leave their characteristic imprint in the spectrum. Sort of a "fingerprinting" for chemical elements. Gratings can also be reflective, an everyday example is a CD which has parallel grooves very close together. If you look at light reflected from a CD it is easy to see a rainbow-like pattern. Prisms work similarly using that the refraction index depends on the wave-length but they are not efficient because to separate the spectrum they have to be thick and therefore absorb too much light. Nevertheless the actual rainbow in the sky is produced in this way by refraction inside tiny drops of water.

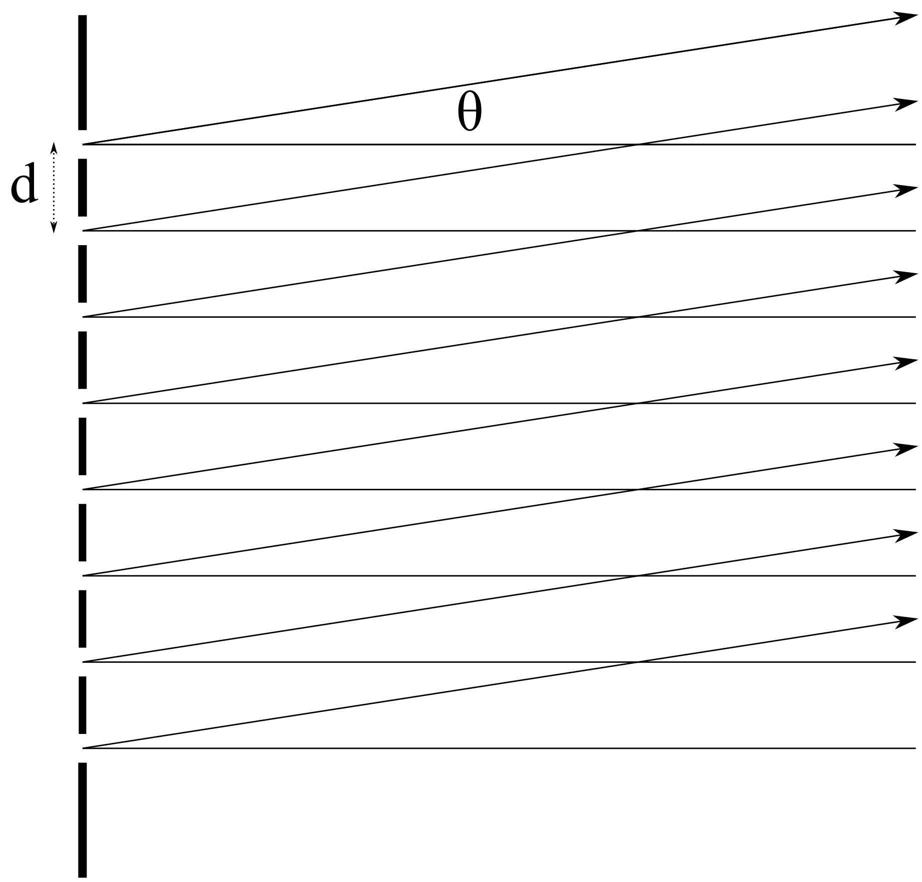


Figure 114: Grating. The angles corresponding to maxima are easily computed as $d \sin \theta_M = n\lambda$.

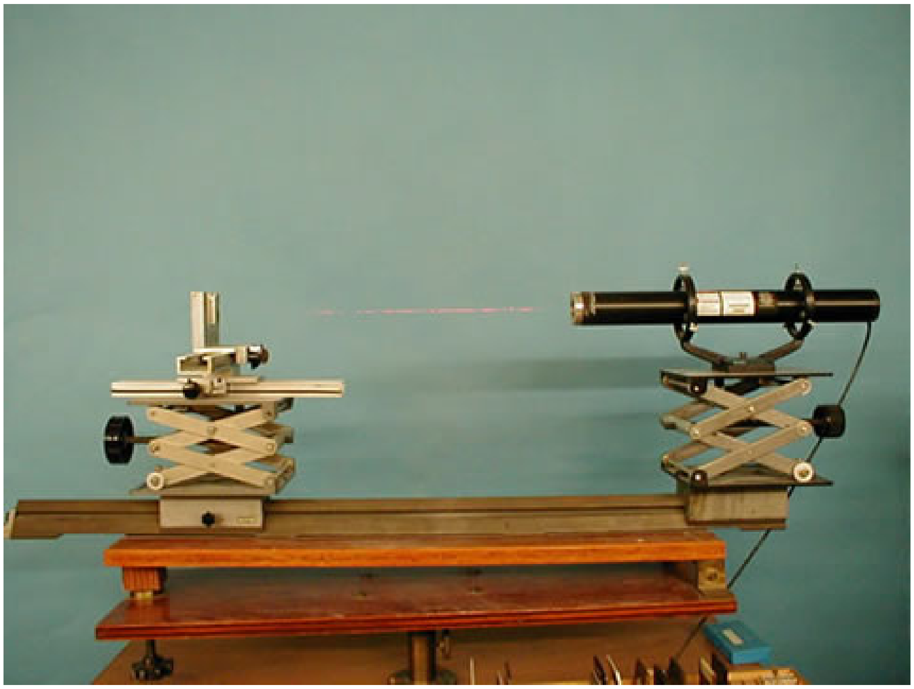


Figure 115: A laser has longer coherence length than a light bulb and makes it easier to look at the interference patterns of gratings as seen in the next figure.

20.2 Diffraction

Interference can also occur when light goes through only one slit in which case is known as diffraction. By shining light on a slit one can see on a screen behind a pattern of light and dark stripes as schematically indicated in fig.117. If we divide the slit into tiny imaginary slits we see that, if a ray coming from the top of the slit interferes destructively with a ray coming from the middle, then for every ray in the top half of the slit there is a ray in the bottom half that cancels it. This gives the condition for a minimum as

$$\frac{a}{2}\sin\theta_m = \left(m + \frac{1}{2}\right)\lambda \tag{20.22}$$

where a is the width of the slit and m is an integer. The first minimum or dark band appears at an angle

$$\sin\theta_m = \frac{\lambda}{a} \tag{20.23}$$

Since the opening is usually significantly larger than the wave-length the angle is small and we can use the approximation $\sin\theta_m \simeq \theta_m$ giving

$$\theta_m \simeq \frac{\lambda}{a} \tag{20.24}$$

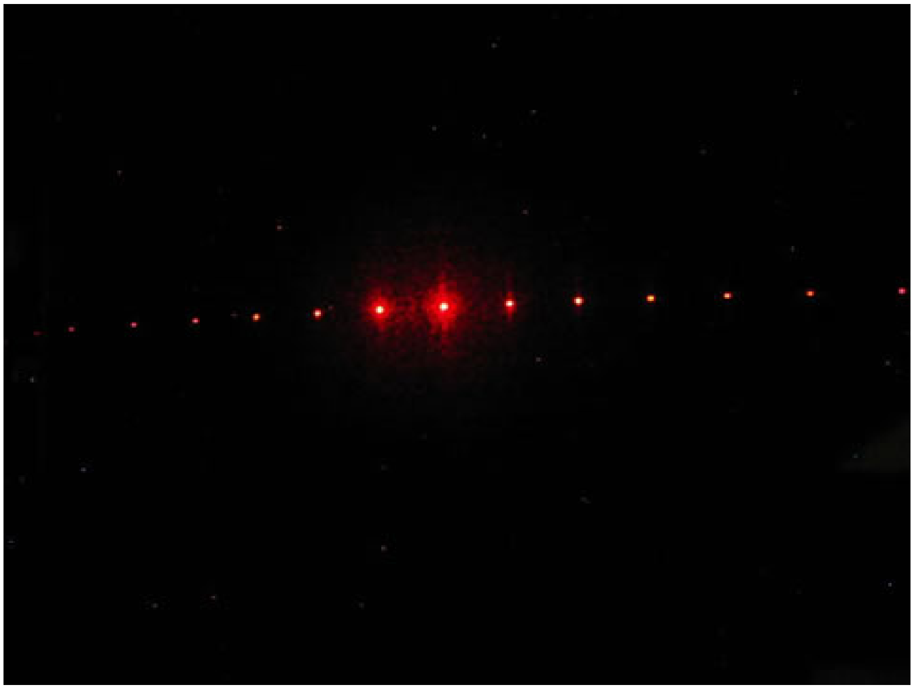


Figure 116: Pattern generated by laser light going through a grating.

If we put a screen at a distance D the width h of the bright band in the middle is

$$h = 2D \tan\theta_m \simeq \frac{2\lambda D}{a} \tag{20.25}$$

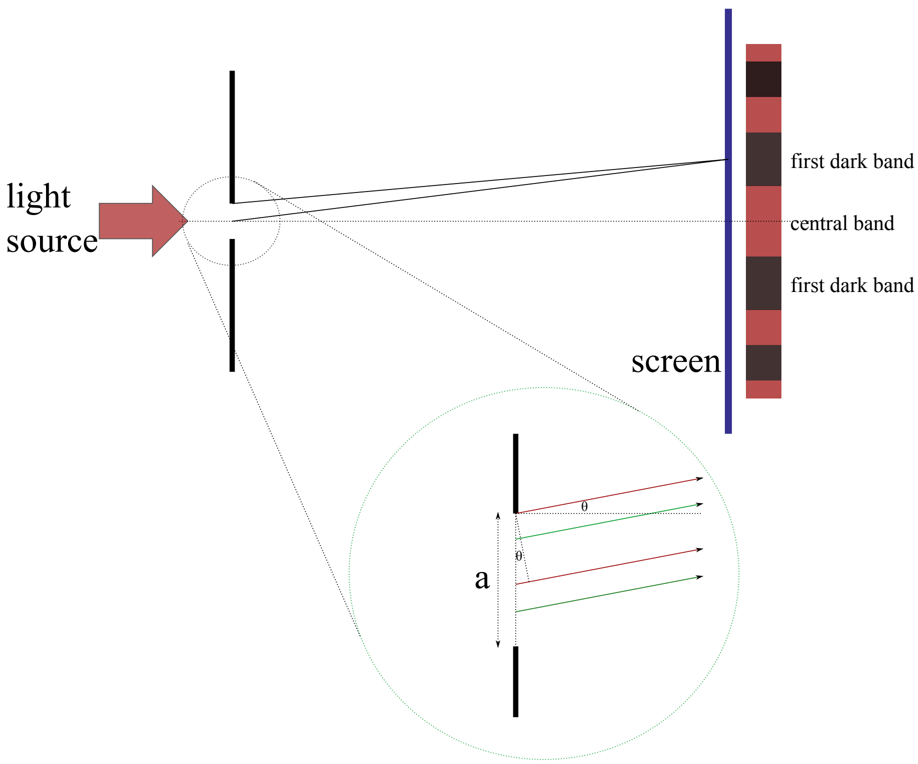


Figure 117: Diffraction through a slit.

21. Lecture 21

21.1 Diffraction and optical instruments

In the last lecture we saw that parallel rays after going through a slit are not parallel anymore but have a range in angles from zero to $\theta_m = \frac{\lambda}{a}$. Either experiment or a more involved calculation show that, if the hole is round instead of a slit, then the minima occurs at an angle

$$\theta_m \simeq 1.22\frac{\lambda}{a}, \quad \text{circular hole} \tag{21.1}$$

Since any optical instrument, for example a telescope has a hole through which light comes in, even perfect parallel rays produce an image blurred as if they were not actually parallel. For example in the simple case of a convergent lens shown in fig.118, horizontal parallel rays produce an image of size r given by

$$r = 1.22\frac{\lambda}{a}f \tag{21.2}$$

where f is the focal distance and a is the diameter of the lens. Namely, although rays arrive parallel to the lens axis, after going through the lens diffraction spreads them over a cone. In fact the image is a solid circle of radius r surrounded by diffraction rings. These rings would make clear that the blurring is from diffraction and not from the other aberrations we have already encountered. Since the wave-length λ is small this blurring does not seem to be very important but we should remember that a telescope further enlarges the image produced by the initial lens or mirror. Diffraction puts a limit to how much this image can be enlarged because once we reach a resolution where we can see the diffraction pattern, further enlargement would be useless. This is sometimes called a theoretical limit, as opposed to other aberrations, nothing can be done to improve it (other than making the telescope wider). Notice also that spherical aberration requires using mostly the region of the lens close to its center. Restricting the opening would accomplish that at the expense of increased diffraction.

One simple but very important device to consider is our own eye. Since the pupil has a diameter of approximately $4mm$ then the image projected on the retina cannot have more angular resolution than

$$\theta_m \simeq 1.22\frac{\lambda}{4mm} \simeq 1.22\frac{400nm}{4mm} \simeq 10^{-4} \tag{21.3}$$

in radians. This is quite small. It also implies that, if Nature was kind to us, we would have enough cells in the retina to reach such resolution. A simple check is to consider, at night, how far we would be able to see the two headlight of an incoming car as separate sources. If the headlights are separated by $1m$ the distance at which we can resolve them is

$$D = \frac{1m}{\theta_m} = 10^4 m = 10Km \simeq 6\text{miles} \tag{21.4}$$

which is the right order of magnitude as we know from our driving experience. This means that the human eye indeed reaches close to its theoretical resolution limit. To improve we would need bigger eyes which would make focusing more complicated.

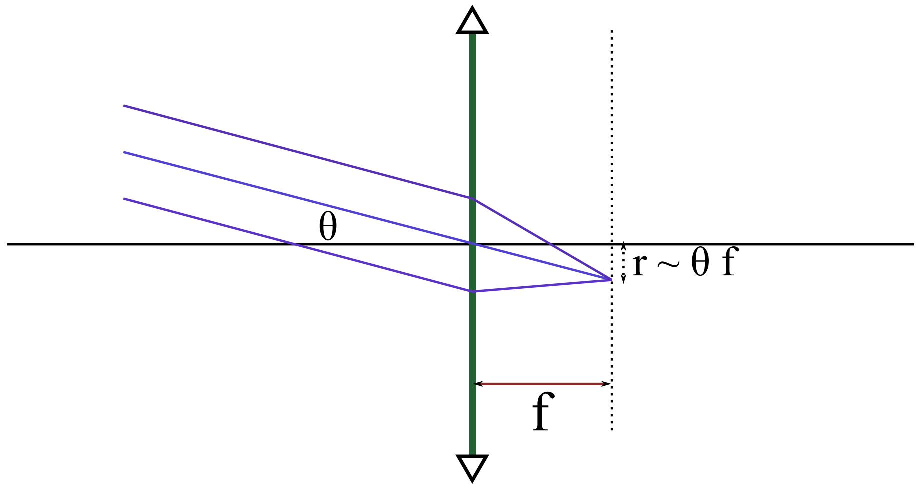


Figure 118: Diffraction through the lens imply that perfectly parallel rays will give rise to a blurred image even for a perfect lens. The effect is very small but detectable when we enlarge the image or look at it with enough resolution. This gives a theoretical limit to the magnification of any optical device (in terms of its width).

21.2 Light-matter interaction: Photoelectric effect

Up to know we have studied the properties of light when it propagates through a medium. However, light interacts with matter in many different ways, for example it is important to study how light is emitted an absorbed by matter. We start this study by considering the photoelectric effect. As we discussed at the beginning of the course, electrons are free to move inside a metal, which explains why they are conductors. When light shines on a metal these electrons can be ripped off the metal. This is called the photoelectric effect. Notice that electrons normally do not leave a metal mainly because if they do so, the metal would be positively charged and would attract them back. Light can kicked them out. In fact certain night-vision systems work in this way by accelerating the electrons and amplifying the resulting current. From a physical point of view we will be mainly concerned with the energy of the electrons which are ripped off the metal. It turns out that the number of electrons coming out is proportional to the intensity of light but their energy is proportional to frequency of the light. This is rather surprising since the intensity of light describes precisely how much energy reaches a certain area of the metal. It make sense that the total energy transmitted to the electrons is proportional to the intensity. In fact it is because the larger the intensity the larger the number of electrons. It is surprising however that the energy of each individual electron depends only on the frequency of light and not the intensity. The experimental result is described in fig.119. No electrons emerge for frequencies smaller than a cut-off f_c. For larger frequencies f the energy of the electrons is given by

$$E_{\text{el.}} = h(f - f_c) \tag{21.5}$$

The explanation of this fact was given by Einstein. He proposed that light is made out of quanta which behave similarly as particles. Each quantum is called a photon and has an energy given by

$$E_{\text{ph.}} = hf = \frac{h}{2\pi}\omega = \hbar\omega \tag{21.6}$$

where h is a universal constant known as Planck's constant. We also used the relation between frequency and angular frequency $\omega = 2\pi f$ and defined

$$\hbar = \frac{h}{2\pi} \tag{21.7}$$

This is an entirely new physical description. Planck had already observed that light is emitted and absorbed in discrete amounts but Einstein took the photons as the real picture of what light is made of. Now the explanation of the photoelectric effect is very simple. To extract an electron of the metal a minimal energy W of needed to overcome the Coulomb attraction. This energy W is called the work function and depends on the substance. If the frequency of the light is smaller that

$$f_c = \frac{W}{h} \tag{21.8}$$

then a photon has not enough energy to kick out an electron. Two or more photons would be needed but the probability of two photons hitting the same electron at the same time turns out to be very small and can be ignored. If the photon has larger energy than W the excess energy is transfered to the electron as kinetic energy:

$$E_{\text{ph.}} = hf = W + E_{\text{el.}} = hf_c + E_{\text{el.}} \tag{21.9}$$

from where eq.21.5 follows.

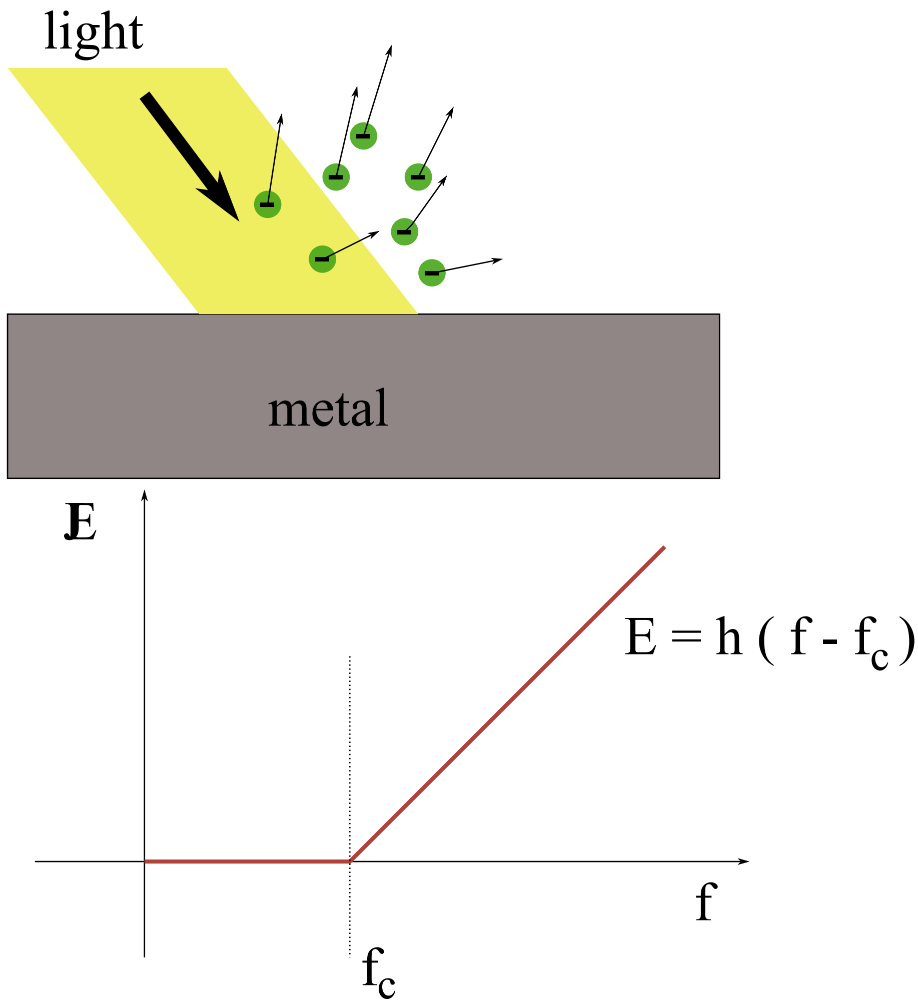


Figure 119: Photoelectric effect. Electrons are ejected from the metal by incident light. For frequencies smaller that f_c no electrons are ejected, for larger frequencies f, the energy of the electrons is given by the formula $E_{\text{el.}} = h(f - f_c)$

22. Lecture 22

22.1 Photons

As already discussed, the correct explanation of the photoelectric effect is that electromagnetic waves are made of small quanta called photons. Each photon has energy and momentum given in terms of the frequency and wave-length by:

$$E_{\text{ph.}} = \hbar\omega = hf, \qquad P_{\text{ph.}} = \frac{E_{\text{ph.}}}{c} = \frac{h}{\lambda} \tag{22.1}$$

That photons have momentum can be seen by shining light on a reflective surface and observing that light exerts a pressure over the surface, namely transfers momentum to it. The value of $\hbar$, the Planck constant is given

$$\hbar c = 197 MeV fm \tag{22.2}$$

where c is the speed of light and as usual $1MeV = 10^6 eV$, $1fm = 10^{-15}m$. Notice that $\hbar$ has units of Energy $\times$ time as expected from the formula $E_{\text{ph.}} = \hbar\omega$. It turns out that usual electromagnetic waves are constituted by huge number of photons and for that reason the particle nature of light is not apparent. However, when interacting with electrons, each photon interacts with an individual electron and the quantum nature of light becomes manifest.

22.2 Hydrogen spectrum and Bohr atomic theory

Now we consider the interaction of light with the simplest atom, namely hydrogen, which is made out of a proton and an electron orbiting around it. When a photon hits the electron it can break it free from the atom in the hydrogen version of the photoelectric effect. However, if the energy is not enough the electron will jump to another orbit and stay bounded to the proton. Similarly, once excited it can decay to a lower orbit and emit a photon. According to classical mechanics the energy of the emitted photon can take any value, the spectrum of energies is continuous. However, very surprisingly at the time, it was observed that when heated up, hydrogen emits photon of very well defined frequencies (or wave-lengths) as shown in fig.120. The same happens with absorption, the same frequencies which are emitted are also the ones that are absorbed by the atom. In fact, experimentally it is seen that the energy of the emitted or absorbed photons fit the very simple formula:

$$E_{\text{ph.}} = 13.6\, eV \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right) \tag{22.3}$$

where $n_{1,2}$ are two positive integers such that $n_2 > n_1$. As Bohr pointed out, the only explanation is that the electron cannot be in any arbitrary orbit but only in orbits with energy:

$$E_{\text{el.}} = -\frac{13.6eV}{n^2}, \quad n = 1, 2, 3, \ldots \tag{22.4}$$

The negative sign means that the electron has less energy than a free electron, namely it is bounded to the proton. When an electron jumps from one orbit to another it can only emit or absorb photons of energies equal to the difference in energy between two of these orbits thus explaining eq.(22.3). Although this explains the hydrogen spectrum it is quite extraordinary. The first surprising point is that the lowest energy is attained for $n = 1$ called the ground state. Classically the electron would radiate all its energy and get stuck to the proton in a state of infinite negative energy. In quantum mechanics this is not what happens, there is a minimum energy that is attained. An explanation for this fact is given in the next subsection in terms of Heisenberg's uncertainty principle. Moreover, above the minimal energy only very specific energies are allowed as we will explain in terms of the de Broglie proposal that particles behave also as waves. To summarize, we are in a realm where Newtonian mechanics does not apply any more and we have to discuss which principles allow us to understand the physical reality at the atomic scale.

22.3 Uncertainty principle

One of the basic principles of quantum mechanics is the uncertainty principle proposed by Heisenberg. In classical mechanics the state of a system is determined by giving the position and momentum (or velocity) of each particle. In a quantum mechanical state, however, the position and momentum of a particle are not simultaneously well defined. If the particle is localized at a point then the momentum is completely undetermined and vice versa if the momentum is well defined, the particle is completely unlocalized. For that reason in quantum mechanics we talk about the probability distribution of position and momentum. If the particle is localized in a region of size Δx and the momentum is in the range $(p,\, p + \Delta p)$ then the uncertainty principle establishes that

$$\Delta x \Delta p \geq \frac{\hbar}{2} \tag{22.5}$$

We can make Δx small at the expense of making Δp large and vice versa.

Consider now the case of the hydrogen atom. The Coulomb potential is attractive and tries to localize the electron as close to the proton as possible. However, if we localize the electron very close to the proton, the momentum distribution is very spread. Since the kinetic energy is given by $K = \frac{p^2}{2m}$ that means that the average value of the

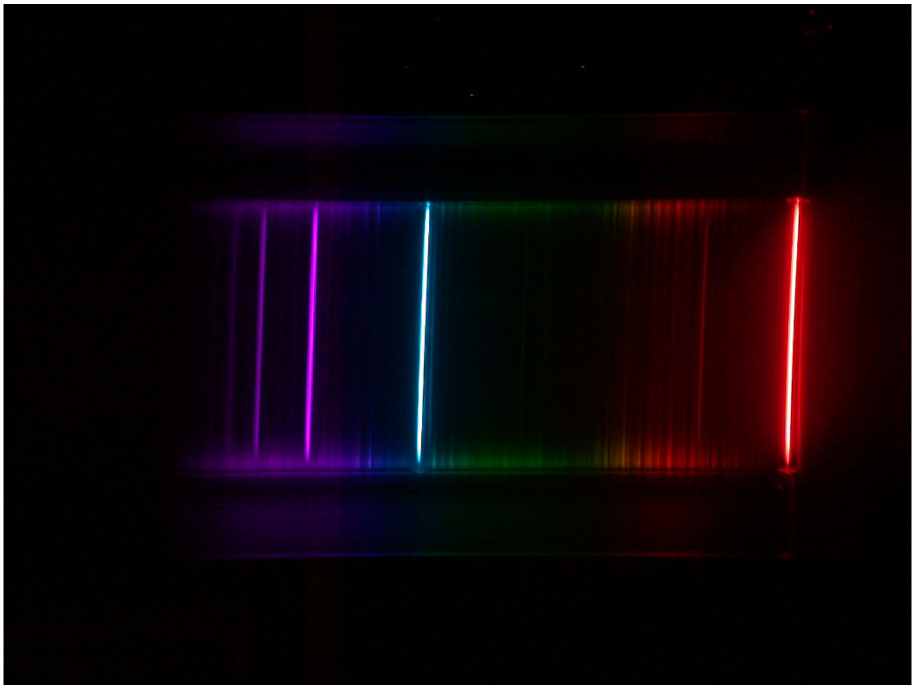


Figure 120: When light is emitted by hydrogen only certain wave-lengths are present as seen in this spectrum where hydrogen light is split using a grating or other similar device.

kinetic energy will be large. This is the way in which the Heisenberg principle operates. The potential energy tends to localize particles at the minimum of the potential. On the other hand the kinetic energy prefers that the particle is spread. For example in metals the conduction electrons are spread all over the metal and conduct electricity whereas there are other electrons which are localized around the atoms and do not contribute to the current. The more we want to localize a particle the stronger the potential we need. For that reason to study the physics at very small scales large energies per particle are required.

Going back to the hydrogen atom we can make a quantitative prediction if we write the total energy as

$$E = \frac{1}{2}mv^2 - \frac{e^2}{4\pi\epsilon_0 r} = \frac{p^2}{2m} - \frac{\bar{e}^2}{r} \tag{22.6}$$

where we used that the momentum is $p = mv$ and also defined

$$\bar{e}^2 = \frac{e^2}{4\pi\epsilon_0} = 1.44\, MeV\, fm \tag{22.7}$$

The last value is computed using the electron charge $e = -1.6 \times 10^{-19} C$. It has the correct units of energy $\times$ length. If we localize the electron in a region of size r, the

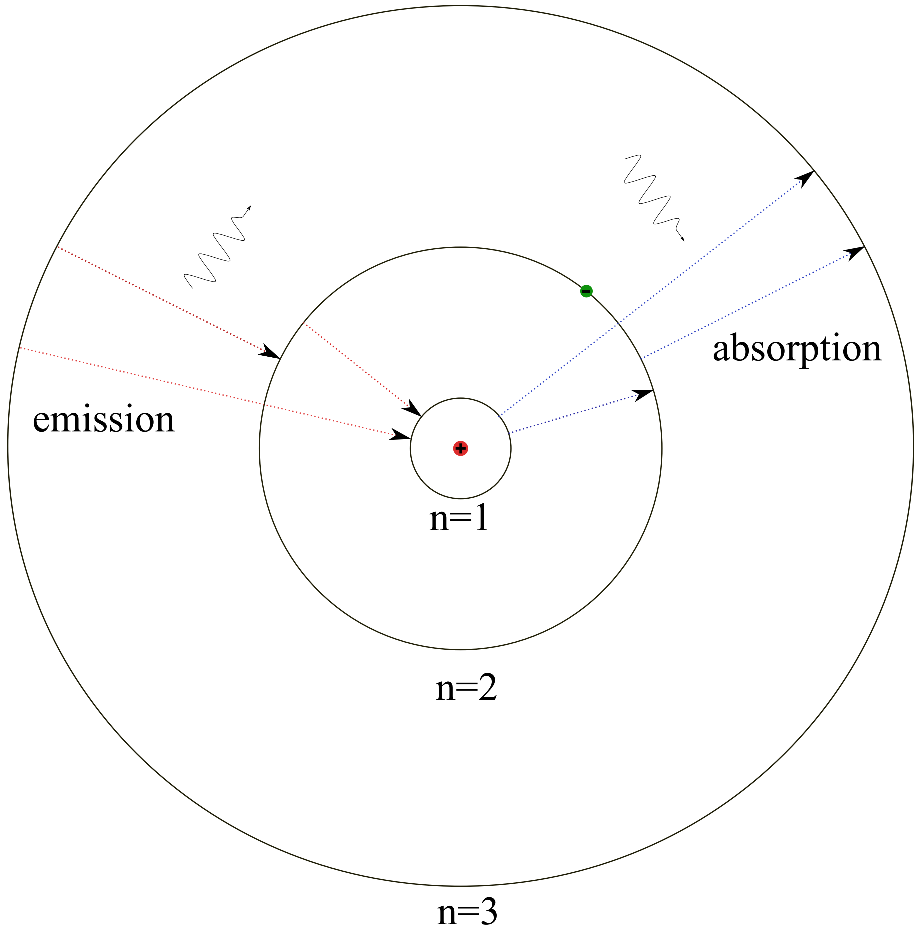


Figure 121: Bohr proposed that only certain discrete orbits are possible based on a quantization principle.

momentum will be spread $\Delta p \sim \frac{\hbar}{r}$ and therefore we estimate:

$$p^2 \sim \frac{\hbar^2}{r^2} \tag{22.8}$$

The total energy therefore is

$$E \simeq \frac{\hbar^2}{2mr^2} - \frac{\bar{e}^2}{r} = \bar{e}^2 \left(\frac{\hbar^2}{2m\bar{e}^2 r^2} - \frac{1}{r} \right) = \frac{\bar{e}^2}{r_0} \left(\frac{r_0^2}{r^2} - \frac{r_0}{r} \right) \tag{22.9}$$

where we defined

$$r_0 = \frac{\hbar^2}{2m\bar{e}^2} \tag{22.10}$$

The radius r_0 has the value

$$r_0 = \frac{\hbar^2}{2m\bar{e}^2} = \frac{(\hbar c)^2}{2mc^2 \bar{e}^2} = \frac{197^2 MeV^2 fm^2}{1MeV 1.44\, MeV fm} = 2.7 \times 10^4 fm = 2.7 \times 10^{-11} m \tag{22.11}$$

where we used that $mc^2 = 0.5MeV$, the energy equivalent of the electron mass. We see now more quantitatively what happens. If r is very small the kinetic energy grows, if r is large then the potential energy grows. We plot the function $f(x) = \frac{1}{x^2} - \frac{1}{x}$ where $x = \frac{r}{r_0}$ in fig.122 and see that it attains a minimum at $x = \frac{r}{r_0} = 2$. If we replace $r \simeq 2r_0$ in the previous computation of the energy we obtain

$$E \simeq \frac{\bar{e}^2}{r_0}\left(\frac{1}{4} - \frac{1}{2}\right) = -\frac{\bar{e}^2}{4r_0} = -13.6eV \tag{22.12}$$

in very good agreement with the experimental result in eq.(22.4) for $n = 1$, the ground state. To be completely honest the full agreement is a coincidence. The uncertainty principle only allows us to get an estimate and this should work well. Namely we expect the energy to be of order of tens of electron volts and it is. That it comes exactly equal is as we said a coincidence and does not always work that way.

In any case, summarizing, the uncertainty principle tells us that we need energy to localize a particle therefore there is a compromise radius where the particle is localized so that the Coulomb energy is low but not too localized that the kinetic energy will grow. Given the mass of the electron and the strength of the interaction (given by the charge) the radius is of order $10^{-10}m$ and the energy of order $10\,eV$.

This principle allows us to understand why there is a minimum energy but we still do not have a principle that tells us that the higher energy states are quantized. We now have to introduce the idea of wave mechanics.

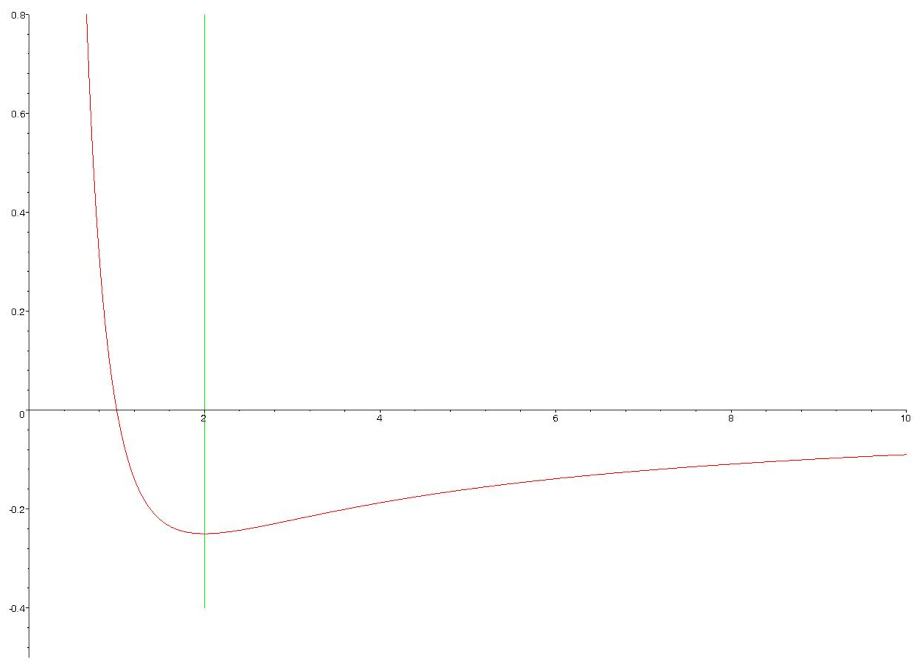


Figure 122: The function $f(x) = \frac{1}{x^2} - \frac{1}{x}$ is plotted and seen to have a minimum at $x = 2$ with $f(2) = -1/4$.

23. Lecture 23

23.1 De Broglie waves

We have seen that light appears to be a wave but, under careful examination, behaves as if composed of particles. Nevertheless all the wave properties of interference and diffraction are still valid. In view of this "particle-wave" duality, de Broglie proposed that particles should also behave as waves. The frequency and wave-length are given by the same relation as for photons:

$$E_{\text{el.}} = \hbar\omega, \quad P_{\text{el.}} = \frac{h}{\lambda} \tag{23.1}$$

In the case of the photon we were more familiar with ω, λ and computed E, P. In the case of the electron we are more familiar with E, P and compute the associated angular frequency ω and wave-length λ. Although these relations are the same as for the photon, it should be emphasized that the relation between energy and momentum is different:

$$E_{\text{el.}} = \frac{P_{\text{el.}}^2}{2m}, \qquad E_{\text{ph.}} = \frac{P_{\text{ph.}}}{c} \tag{23.2}$$

This was a very strange proposal but it received spectacular experimental confirmation when an experiment analogous to the two slit experiment showed an interference pattern for electrons exactly the same as for light. In fact any experiment of diffraction and interference of electrons can be analyzed in the same way as for light. We only need to compute the wave-length using $\lambda = \frac{h}{p}$.

Now we want to see how this helps us in the hydrogen atom. First think of a string in a violin or guitar. It is well known that such a string only vibrates at specific frequencies which is why is used in a musical instrument. In fact the lowest frequency is such that the length of the string is half a wave-length. We say that the vibrations of a string have a discrete spectrum of frequencies. The proposal is that the discrete spectrum of the hydrogen atom is due precisely to the wave-like behavior of the electron. More quantitatively, we require the the size of the orbit is an integer multiple of the wave length of the electron. Namely in each orbit an integer number of wave-lengths fit:

$$2\pi r = n\lambda, \qquad n = 1, 2, 3, 4, \ldots \tag{23.3}$$

Now we use the Newtonian relation

$$m\frac{v^2}{r} = \frac{\bar{e}^2}{r^2} \quad \Rightarrow \quad v = \sqrt{\frac{\bar{e}^2}{mr}} \tag{23.4}$$

to compute the velocity of the electron and thus the momentum and wave-length:

$$p = mv = \sqrt{\frac{m\bar{e}^2}{r}}, \quad \lambda = \frac{2\pi\hbar}{p} = 2\pi\hbar\sqrt{\frac{r}{m\bar{e}^2}} \tag{23.5}$$

Using now eq.(23.3), namely $2\pi r = n\lambda$ we get:

$$2\pi r = 2\pi\hbar\sqrt{\frac{r}{m\bar{e}^2}}n \quad \Rightarrow \quad r = \frac{n^2\hbar^2}{m\bar{e}^2} \tag{23.6}$$

We indeed find a discrete set of orbits!. The energy is computed by replacing the velocity from eq.(23.4) into

$$E = \frac{1}{2}mv^2 - \frac{\bar{e}^2}{r} = -\frac{\bar{e}^2}{r} = -\frac{1}{n^2}\frac{m\bar{e}^4}{\hbar^2} = -13.6eV\,\frac{1}{n^2} \tag{23.7}$$

where we replaced the known values of $\bar{e}$, m and $\hbar$. Now we have the energy of all possible orbits fitting precisely eq.(22.4)!. This is a good check that the discreteness of the hydrogen spectrum is due to the wave-like nature of the electron.

23.2 Other results and applications

We basically have finished what we wanted to discuss about quantum mechanics. For the last hundred years we have been improving our understanding of how quantum mechanics applies to different physical phenomena. For example one can understand the periodic table of the elements, chemistry, material science, particle physics, etc. as fields where quantum mechanics is of paramount importance. However it has only been recently that experimental progress has been enough that one can start thinking of practical application in which controlling the state of a quantum system can be used to our advantage. One such possibility would be to use quantum mechanics to create better computers, a field which is still in its infancy. For illustration we will discuss two devices which are practical applications of quantum mechanics: lasers and atomic clocks.

A laser creates a beam of coherent light which is well collimated, namely does not spread much. In quantum mechanical terms we generate a large number of photons all in the same state. To understand how that is achieved we need another important property: if an atom is in an excited state it can transition to a lower energy state by emitting a photon with the corresponding energy. However, if there is already a number of photons of that energy present, the probability of decay is enhanced for the photon to go to the same state in which those photons are. A laser utilizes this by having a material whose atoms are excited by electric discharges, light, electric fields, etc. In fig. 125 we see an example, the active zone, where the material is excited, is situated

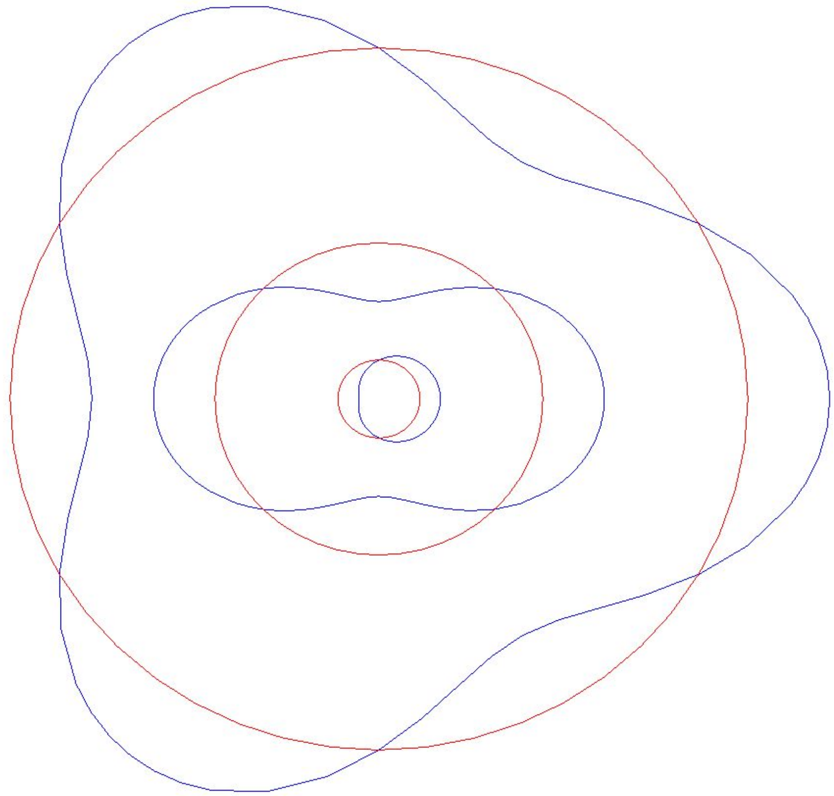


Figure 123: The Bohr atom explained through the de Broglie hypothesis. An integer number of wave-lengths fit into each orbit.

between two mirrors. A decay produces a photon. The other atoms, when they decay prefer to emit the photon in the same state as the one already present. As more and more photons accumulate on a state, more likely is for others to join. In this manner light is amplified which gives its name to the device: Light Amplification by Stimulated Emission of Radiation. The words stimulated emission of radiation makes reference to the idea that a photon already present stimulates the atoms to emit radiation in the same state. The mirrors can be thought as determining standing waves similarly as in a sound waves in a pipe. Alternatively we can think that a beam of light goes back and forth between them being amplified all the time. To extract the energy, one of the mirrors is partially transparent. It is clear that only photons moving along the cylinder, perpendicular to the mirrors are amplified. The others are lost since they are not reflected back into the active zone. Furthermore, all the photons are in the same state, so they are in phase and the light is coherent. This should be contrasted with the usual thermal emission in which the same medium is heated and each atom decays independently of the others. In that case we have a large number of sources all emitting independently. In the laser they all contribute to the same wave generating a highly coherent pulse.

The same idea is used with microwaves. A resonant cavity as it is called contains atoms and an alternating frequency is applied. When the external frequency agrees with the energy of an atomic transition a resonance occurs that can be easily detected. Since the frequency of the atomic transition is very precise, we can create an alternating voltage with a very precise frequency or period. But a periodic signal with very precise period is exactly what we need to build a clock. This type of clocks are called atomic clocks and are the most precise clocks available at the moment. They are used in numerous applications, for example atomic clocks aboard the GPS satellites produce timig signals that allow us to determine our position by getting timing signals from different satellites.

Finally, another phenomenon that we wanted to discuss is Compton scattering. When photoelectric effect occurs but the energy of the photon is much larger than the binding energy of the electron, the electron can be considered as a free particle. In such situation, the photon cannot be absorbed and is deflected. If the deflection angle is θ, see fig.124 then, conservation of energy and momentum give the relation between the wave-length λ of the incoming photon and λ' the wave-length of the outgoing photon:

$$\lambda' - \lambda = \frac{h}{mc}(1 - \cos\theta) \tag{23.8}$$

Here, m is the mass of the electron. In fact to obtain this result you need to use the relativistic relation between momentum an energy $E = \sqrt{m^2c^4 + p^2c^2}$. We leave this as an exercise for anyone who is interested.

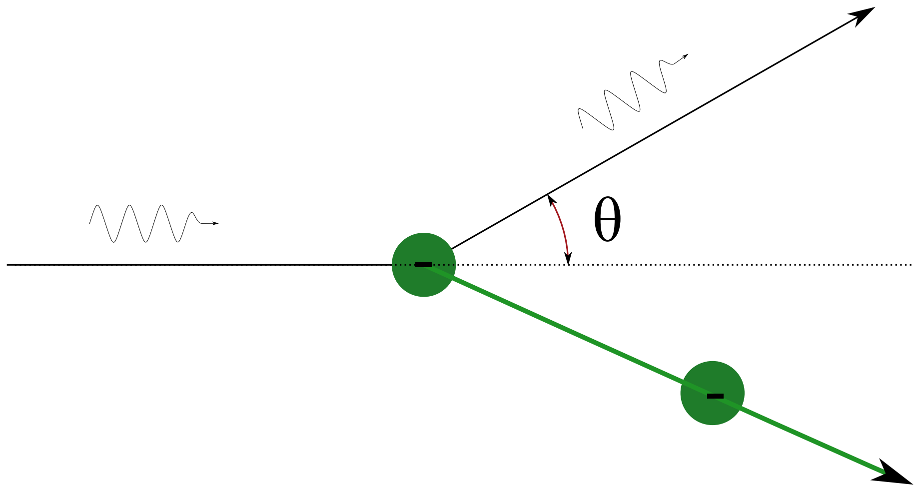


Figure 124: Compton scattering is analogous to the photoelectric effect but the electron either is not bound or the binding energy is small compared to the energy of the photon. The photon deflected by an angle θ and changes its wave-length.

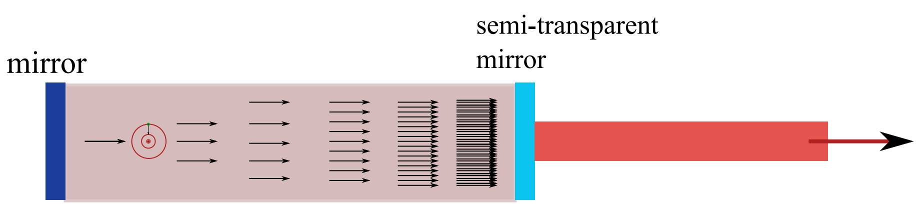


Figure 125: Schematic of a laser. A beam of light goes back and forth in an active medium which amplifies light. One of the mirrors is partially transparent and lets the laser light go out.

24. Lecture 24

24.1 Nuclear Physics

24.1.1 Constituents and binding energy

The atomic nucleus has a typical size of $10^{-15}m$ and therefore is much smaller than the typical atomic size $10^{-10}m$. Using the uncertainty principle we expect then that the energies associated with the nucleus are much larger, of the order of MeV. The nucleus is composed of protons which are positively charged and neutrons which are electrically neutral. The total number of protons and neutrons is called the mass number and denoted by an A. The number of protons is called the atomic number and is denoted by a Z. The number of protons determines the charge and therefore the number of electrons of the corresponding atom and with that its chemical properties. Therefore Z gives the "name" to the nucleus. For example Carbon has six protons as depicted schematically in fig.126. However the number of neutrons in Carbon can vary. In the picture we depicted Carbon fourteen written usually as ${}^{14}_{6}C$. For example there is also Carbon twelve (the most common one). These are called isotopes, they have the same of protons but different number of neutrons.

The mass of the proton and neutron is approximately similar and equal to $m_p \sim m_n \sim 1.6 \times 10^{-27}Kg$. Thus, the mass of the nucleus is given by the mass number A. One important fact is that if one measures the mass of the nucleus M_{nucleus} with precision one finds what is called a mass defect:

$$\Delta M = Zm_p + (A - Z)m_n - M_{\text{nucleus}} \tag{24.1}$$

namely a difference between the mass of a corresponding number of protons and neutrons and the actual mass. To do this computation we should use more precise values for the mass of the proton and neutron:

$$m_pc^2 = 938.272\,MeV \qquad m_nc^2 = 939.566\,MeV \tag{24.2}$$

where we used the more convenient mass equivalent values given by Einstein's formula $E = mc^2$. In fact Einstein formula gives us the clue to understand the mass defect. Since mass and energy are equivalent, when protons and neutrons form the nucleus, the total system has less energy and therefore less mass than the separate components. Therefore $E_b = \Delta Mc^2$ is precisely the binding energy of the nucleus. If we want to split it back in its components this is the energy that must be supplied. This happens for any physical system but in the nucleus the binding energy is large enough that the mass defect can actually be measured. Although they form a bound state not all nuclei are stable. They can decay into other nuclei as discussed in the next section.

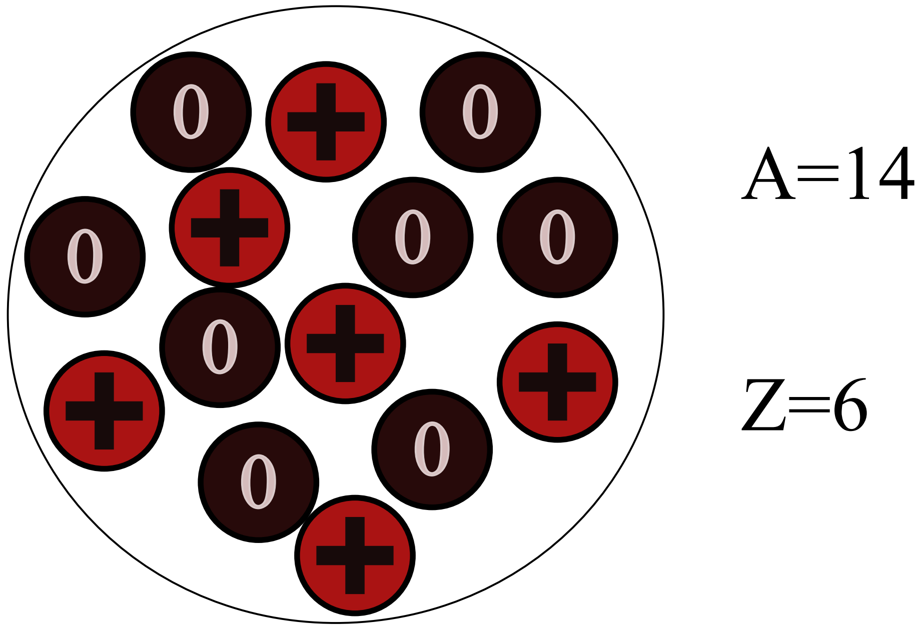


Figure 126: A nucleus of $^{14}_{6}C$ (a radioactive form of carbon) is made out of six protons and eight neutrons

24.1.2 Nuclear decays

Certain nuclei can spontaneously decay into others. Generically the nuclear decays are classified into alpha, beta and gamma decays from the type of particle they produce. Alpha particles are nuclei of Helium four, composed of two protons and two neutrons. It is particularly stable nucleus. Beta particles are simply electrons and gamma particles are photons. The reason for the names alpha, beta, gamma is that they were given before the actual identity of the particles emitted was known.

Alpha decay occurs when a nucleus is unstable. An alpha particle is emitted and the reaction is written:

$$^{226}_{88}Ra \rightarrow ^{222}_{86}Rn + ^{4}_{!}He \tag{24.3}$$

We used the example of Radium (Ra) which a highly radioactive element discovered by Marie Curie. Notice that as a result of alpha decay the nucleus loses two protons and two neutrons. Generically an unstable nucleus X will decay as

$$^{A}_{Z}X \rightarrow ^{A-4}_{Z-2}X + ^{4}_{2}He \tag{24.4}$$

Beta decay occurs due to an interaction that we did not discuss yet. It is called the weak interaction and is responsible for example for the instability of the neutron:

$$n^0 \rightarrow p^+ + e^- + \bar{\nu} \tag{24.5}$$

A free neutron decays into a proton, an electron and an antineutrino ($\bar{\nu}$). The neutrons bound in the nucleus generically do not decay because that would rise the energy of the system. However in certain cases it is possible that a neutron is converted into a proton as in

$$ {}^{14}_{6}C \rightarrow {}^{14}_{7} N + e^- + \bar{\nu} \tag{24.6} $$

Notice that A stays the same but Z is increased by one. The electrons emitted are highly energetic an can be detected as a hallmark of beta decay. It is also possible for a proton in the nucleus to capture an electron an become a neutron. This is called electronic capture. For example

$$ {}^{40}_{19}K + e^- \rightarrow {}^{40}_{18} Ar + \nu \tag{24.7} $$

Finally it can also occur that a proton becomes a neutron plus a positron and a neutrino:

$$ {}^{22}_{11}Na \rightarrow {}^{22}_{10} Ne + e^+ + \nu \tag{24.8} $$

The last type of decay called gamma decay is simply the same electromagnetic decay that we saw for the atom. In this case we can consider that for example a proton is moving in an excited orbit an decays to a lower orbit emitting a photon or gamma ray. Normally this decay occurs as a follow up of one of the other decays. Namely, after alpha or beta decay, the resulting nucleus ends up in an excited state from which it decays through emission of gamma rays.

Both alpha particles and beta particles (electrons) are charged and as such interact electromagnetically with atoms. For that reason they are easily stopped by air for example. However gamma rays, or photons are much more penetrating. They are harmful and potentially lethal at high intensity. They are responsible for the "danger radioactivity" signs. They can be stopped by lead for example.

For any type of decay we consider, quantum mechanics does not allow us to predict the exact time when the decay will occur. However we can compute the probability of a decay occurring and with that what is called the half-life. If we have a sample of N radioactive nuclei, they will start decaying. After a certain time $T_{\frac{1}{2}}$ only half of the nuclei will remain. This $T_{\frac{1}{2}}$ is defined as the half-life. If we want to be more precise, it turns out that the number of radioactive nuclei in the sample decreases exponentially with time:

$$ N(t) = N_0 e^{-\lambda t} \tag{24.9} $$

where N_0 is the initial number of nuclei and $N(t)$ is the number remaining after time t. It is easy to see that $N(t) = \frac{1}{2} N_0$ when

$$ t = T_{\frac{1}{2}} = \frac{\ln 2}{\lambda} \tag{24.10} $$

where $\ln 2$ is the natural logarithm of two, namely $e^{\ln 2} = 2$. One interesting nucleus is ${}^{14}_{6}C$ whose half life is 5730 years. Knowing the initial amount of ${}^{14}_{6}C$ and the present amount one can determine the age of a biological sample. The initial amount is usually known by its ratio to the most common ${}^{12}_{6}C$ which is approximately stable in the atmosphere. A live organism has the same ratio of ${}^{14}_{6}C$ to ${}^{12}_{6}C$ as the atmosphere but after its death, no more interchange with the atmosphere occurs and the amount of ${}^{14}_{6}C$ decreases allowing to determine the time since the death occurred (as long as its not much larger than 5730 years).

The number of decays occurring in 1 second is an important measure of the activity of a radioactive source. In that way the activity is defined and measured in Curies or Becquerels defined as:

$$1Ci = 3.7 \times 10^{10} \frac{\text{decays}}{s}, \quad 1Bq = 1 \frac{\text{decays}}{s} \tag{24.11}$$

24.1.3 Fission and Fusion

Besides the decays we mentioned before there are other two important nuclear phenomena, fission and fusion. In fission a nucleus splits in two approximately equal size nuclei. This can occur spontaneously but more commonly occurs when the nucleus is bombarded with neutrons. Conversely, two light nuclei can join to form a heavier nuclei, a process known as fusion. The most stable nuclei have masses close to iron (Fe) so nuclei lighter than iron can fuse and heavier undergo fission. A notable example for fission is Uranium which can undergo fission for example as:

$${}^{1}_{0}n + {}^{235}_{92}U \rightarrow {}^{139}_{56}Ba + {}^{94}_{36}Kr + 3\,{}^{1}_{0}n \tag{24.12}$$

When ${}^{235}_{92}U$ absorbs a neutron it splits in two releasing also three neutrons. Those neutrons can split more nuclei creating more neutrons and accelerating the reaction. However it should be noted that the most common form of Uranium is ${}^{238}_{92}U$ which does not undergo fission. So in reality most of the neutrons are lost and do not produce more fission. The only way to have a self sustaining reaction is to have large amounts of Uranium so that the neutrons have more distance to travel before leaving the sample or to "enrich" the Uranium by increasing the concentration of ${}^{235}_{92}U$ over ${}^{238}_{92}U$. This is the basis of a nuclear reactor which produce energy out of uranium fission. In fact if the concentration of ${}^{235}_{92}U$ is highly increased the reaction is not only self-sustaining but explosive which is the basis for nuclear weapons. In fact one gram of Uranium completely undergoing fission produces $9 \times 10^{10} J$ of energy to be compared with $1g$ of TNT which produces 4200 J.

Fusion on the other hand occurs between light elements, for example hydrogen with different amounts of neutrons can fuse:

$$^{2}_{1}H +^{2}_{1} H \rightarrow^{3}_{1} H +^{1}_{1} H \tag{24.13}$$

In an exception to the usual rule, hydrogen with one or two extra neutrons have their own names. They are called deuterium and tritium respectively. For example deuterium replaces one atom of hydrogen in what is called heavy water. Since the nuclei undergoing fusion initially repel each other because they have the same charge, a large amount of initial energy is required to produce the fusion. However more energy is liberated afterward. In fact this is the source of energy for the Sun. It actually is also the source of energy for the so called hydrogen bomb which uses a fission bomb as initiator for the fusion. It should be very important to create a controlled fusion reaction which produces more energy that the one put in. This is sometimes consider the source of energy of the future. It has proved more difficult than expected but there is a large international collaboration attempting such feat.