ELECTRICITY, MAGNETISM, and the ELECTROMAGNETIC FIELD

The discovery of various electrical and magnetic phenomena, the realization by Faraday that they were all related, and the invention of a theory by Maxwell that embraced all of these phenomena using the new idea of an electromagnetic (EM) field, is one of the most important developments in the history of science. As discussed in the notes on the field concept, the introduction by Maxwell of this idea has been crucial to physics ever since - we now now that all of Nature seems to be described by fields of one kind or another. The EM field is a model of its kind, and is still the simplest of the fields that have been found. Thus its study is not only of key historical interest, but also of key importance if we want to understand the structure of physical law.

However, before we begin to look what an EM field is, it is important to grasp the amazing diversity of natural phenomena it embraces. The simple fact is that almost everything one is likely to meet in everyday life on earth is explicable in terms of just 2 fields - the gravitational field, and the EM field, and the charges that these 2 fields act on (mass for gravitational fields, and electric charge for EM fields). So we begin by looking at some of the key facts about EM fields, and only then see how they are then incorporated into a theory. To do this we simply go through some of the key experimental results, discovered over a 50-yr period by physicists in Europe during the first half of the 19th century, and, in the first instance, leave them as fascinating and puzzling facts to be digested.

I then give a simple picture of how one can interpret all of these facts in terms of a single field, the EM field, and discuss some of the philosophical consequences of this remarkable result.

A. BASIC EXPERIMENTAL FACTS

In what follows I will present some key experimental facts about electric and magnetic fields. The main aim here is not too teach you all the gory details of electromagnetism (although you will possibly learn a few of these if you want). It is rather to engage in a basic scientific exercise - we look at a set of facts, as a sort of puzzle, of the kind that a detective might be presented with, and then we attempt to make sense of them by looking for what is 'hidden'. We attempt to infer and/or deduce things about the *underlying mechanism*, and the underlying "*hidden entities*" that might be responsible.

A.1: ELECTROSTATIC FIELDS

Let's begin very simply by looking at what happens when we have a single static (ie., motionless) electric charge. It is fairly easy to establish that there are 2 sorts of charge, which we will call positive and negative. Moreover these charges come in well-defined 'lumps'; there is a minimum possible electric charge, and all other charges are just multiples of this. We can also establish fairly easily that like charges repel each other, and unlike charges attract each other. These forces are quite strong, even for very small amounts of charge, and so are very easily measured.

The question then is, *how* do these electrical forces depend on how far apart the charges are from each other? This is an obvious question to ask, since it was found by Newton already nearly 350 years ago that the gravitational force between 2 masses of size M_1 and M_2 varies like the "*inverse square*" of the distance r between them (ie., if the distance is increased by, say, a factor of 2, then the attractive force between them *decreases* by a factor of 4, the square of 2). Moreover the force is proportional to the mass of each object (this seems very reasonable - if we increase a mass M_1, we expect it to create a larger gravitational field, and therefore exert a larger force on another mass, in proportion to its own mass M_1). These 2 statements are summarized in Newton's famous law of gravitational attraction, which says that the force F goes like $F = G(M_1M_2/r^2)$, where G is some constant which tells us how strong gravity is (in fact G is very small - we need very big masses to get an appreciable force).

Remarkably, it is found that the force between 2 charges behaves in the same way - in fact one writes that the strength of the force between 2 charges q_1 and q_2 varies like

$$F = -k\frac{q_1 q_2}{r^2} \tag{0.1}$$

ie., this force also decreases as r increases like the inverse square of r, and it is proportional to each of the two charges. This, along with all the other basic facts about electrostatic forces, were uncovered by Charles Augustin de Coulomb in experiments carried out between 1785-1788. However, he found 2 key differences between electric and gravitational forces. First, the constant k in the Coulomb force is very big compared to G; even very small electric charges can exert very strong forces on each other. Second, the charges q_1 and q_2 can be either positive or negative, so that the force can be either attractive or repulsive (whereas masses can only be positive).

This situation is shown in Fig. 1, where we show (in Fig. 1(b)) these results in the modern language of electric fields. The electric field lines point radially out from a single positive charge Q. These lines indicate the strength and direction of the force created on some infinitesimal test charge q (also positive) by the charge Q. They point outwards, meaning that the test charge q will be repelled away from Q. Now we can think of this in a slightly different way. Imagine that the force lines shown are the result of an electrostatic "potential energy" between the charges Q and q. The test charge will try to lower its energy; and its electrostatic energy will be higher near the the charge Q. There will then be a "potential hill" around the charge Q, and we can then think of the repulsive force as arising from the effect of the electrostatic potential energy - the test charge will "roll down" the potential energy hill around the charge Q, and the field lines simply indicate the direction that the force will be pushing on it.

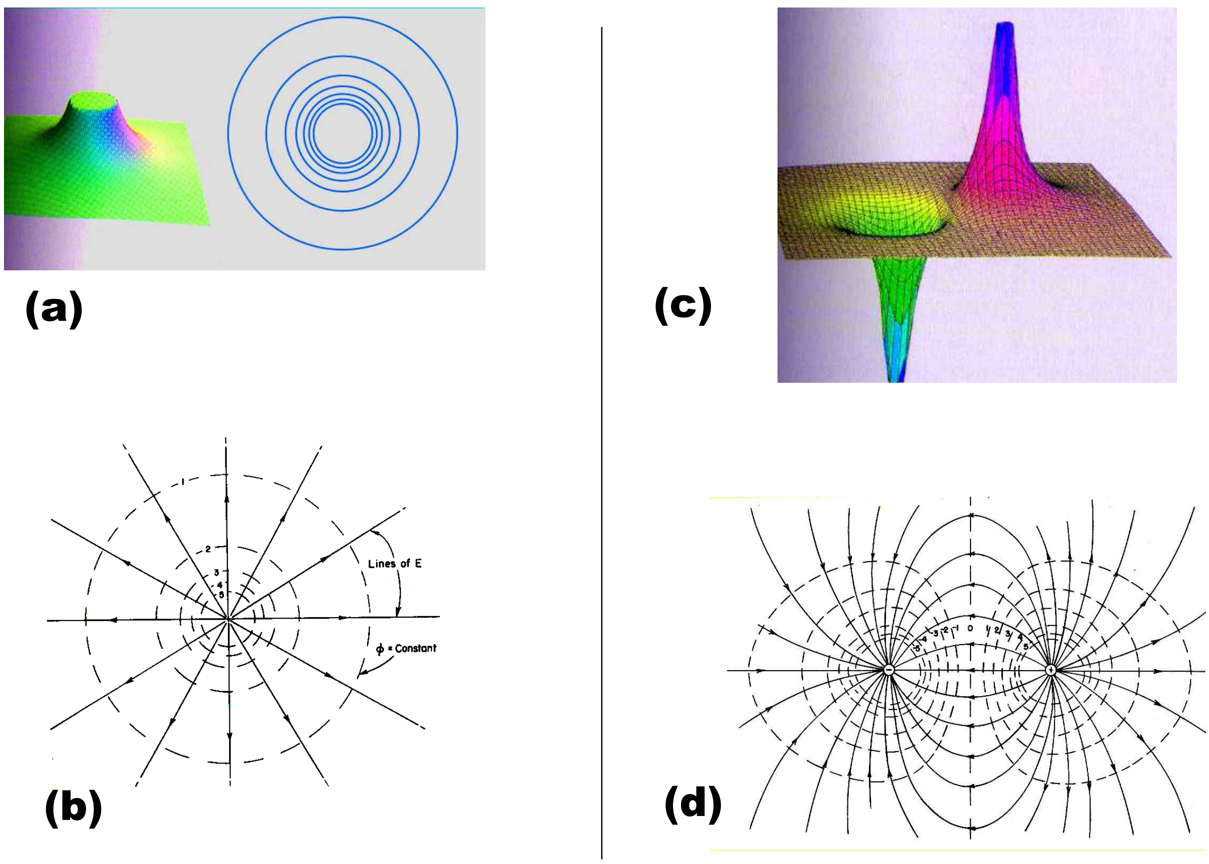

FIG. 1: Static electric fields from static charges. In (a) we see at left, how the electric potential varies near a positive charge, increasing near the charge; at right we show an 'equipotential' contour map of the potential. In (b) the corresponding electric field lines are shown extending radially out from the positive charge - the equipotential lines are shown as dashed lines. In (c) we show the electric potential from a *pair* of charges, one positive and one negative; and in (d) we see the field lines for this pair (with the equipotentials shown as dashed lines.

The potential hill for a positive charge is shown in Fig. 1(a); we show it both as a hill, and also in terms of a contour map, showing equipotential lines (ie., lines on which the electric potential energy is the same). You should think of this as exactly analogous to a topographic contour map, where the contours indicate lines of equal altitude.

We can extend this idea to look at multiple electric charges. In Fig. 1(c) we see the electric potential energy around a pair of charges, in which the one ont he left is negatively charged, and the one on the right is positively charged. Naturally, any test positive charge will be repelled from the 'potential hill' around the positive charge, and will tend to 'fall into' the potential well around the negative charge (ie., it will be attracted towards it). Fig. 1(d) shows the corresponding contour map for equipotentials, along with the electric field lines (the equipotentials are shown as dashed lines).

Note an important conclusion that arises from all of this - since opposite charges are attracted to each other, they will tend to 'fall into' each other, and bind together. This means that opposite charges will tend to pair up into electrically neutral pairs. This is why most large objects are almost exactly electrically neutral - the attractive force between unpaired opposite charges is very strong. Typically the best way to give some object an excess charge is to rub it hard - this will strip some negatively-charged electrons off the surface atoms, leaving the object with a net positive charge. Usually it will not stay this way for long - stray negative charges from other objects (eg., on dust in the air) will quickly be attracted to the surface of the body, to neutralize the excess positive charge.

Let's now summarize this picture. Electric charges generate electric fields, which then exert a force on, or 'interact with', other charges - thereby an interaction between charges is created, mediated by the electric field.

A.2: MAGNETOSTATIC FIELDS

So far so good - we have seen that simple experiments can establish the existence of 2 kinds of electric charge, and the strength of the electric interactions between them. What now of magnetic forces? Both electric and magnetic forces have been known for thousands of years, but at first glance there is no obvious connection between them. In fact it was not until the 19th century that we had any idea how magnetic forces originated. Before this, all that was known was that certain materials (usually ones rich in Iron or Nickel) would tend to orient themselves in a certain direction, caused by an invisible influence (what eventually came to be called a magnetic field); and that the earth generated a magnetic field in some mysterious way. Once humans had begun to voyage over considerable distances (ie., certainly by 1000 AD., and most likely as long ago as 2000 BC - there is good evidence that Javanese sailors were making voyages as far as Madagascar at this time), it was realized that the earth's magnetic field was roughly aligned in a north-south direction.

The breakthrough came when it became possible to generate electrical currents, ie., flows of electric charges through conducting wires. Note that such wires are *electrically neutral*, ie., they have no net electric charge - therefore there are no electrostatic forces coming from them at all. Any forces that do come from the current are therefore not electrostatic; and in fact we will see that that must be magnetic. The first key fact was uncovered by Hans-Christian Oersted in 1820, viz., that a magnetic field was generated by an electric current (ie., a flow of electric charge) in a conducting wire. He could see this because the field created a force which acted on 'test' magnets (eg., small bar magnets, or magnetized iron filings, located in the vicinity of the current-carrying wire. This was followed swiftly, in the same year, by work of Ampère, Biot, and Savart in France, who succeeded in uncovering many of the basic facts about static magnetic fields and the forces they exert on electric currents (and the resulting magnetic forces between current-carrying wires). Finally, in 1831, Faraday made the most important discovery of all, discussed in the next subsection on the *dynamic properties* of fields, that changing magnetic fields in time produced an electric field (and likewise, changing electric fields in time produced a magnetic field).

Let us now go through the key facts about *static* magnetic fields, one by one (it will also be helpful to look at the Course slides at this point). We begin with Fig. 2(a), which shows how a current through a straight wire produces a magnetic field which circulates around the wire. If we now bend the wire around into a ring-shaped "current loop", with current circulating around the ring, the field now continues to circulate around the current, yielding the magnetic field pattern shown in Fig. 2(b). In both of these experiments, the magnetic field distribution in space is mapped out by 'test magnets', in the form of very small iron filings. As Ampere realized, each of these iron filings itself behaved like a miniature current loop, and orients itself along the field. A further increase in the number of loops, as in Fig. 2(c), produces what is called a 'solenoid' coil; the field from each current loop adds to the others. As a matter of fact we can think of a bar magnet (ie., a long bar made from magnetized material, which acts as a 'permanent magnet'), as made up from a huge number of tiny current loops, all oriented in the same direction, with all their fields adding to make a macroscopic field. Each microscopic current loop is in fact created by the electrons circulating around in individual atoms or molecules. In most solids these do not all align together, but are instead pointed in random directions - in this case the fields will on average cancel each other at the macroscopic scale. However it can happen that interactions between the atomic currents on different atoms can cause them to all line up in the same direction, and then we say that the solid is 'magnetized'

Since an electric current creates a magnetic field, it is natural to ask - what is the effect of a magnetic field on an electric current? The general rationale for this kind of question is that one expects that if some physical system or phenomenon A affects another system B, then it is reasonable to suppose that B will also affect A. Two simple experiments to check this out are shown in Fig. 2. First, we consider a pair of current-carrying wires aligned parallel as shown in Fig 2(d). It is easy to check that if we run current through these wires parallel to each other, then the 2 wires will attract each other. On the other hand if we run the currents in the 2 wires in opposite directions to each other, we find that the wires will repel each other.

Note that in each of these cases, the magnetic field generated by one of the wires is 'circulating' around the wire (compare Fig 2(a)), so that when it intersects the other wire, it is intersecting it perpendicularly to that wire -

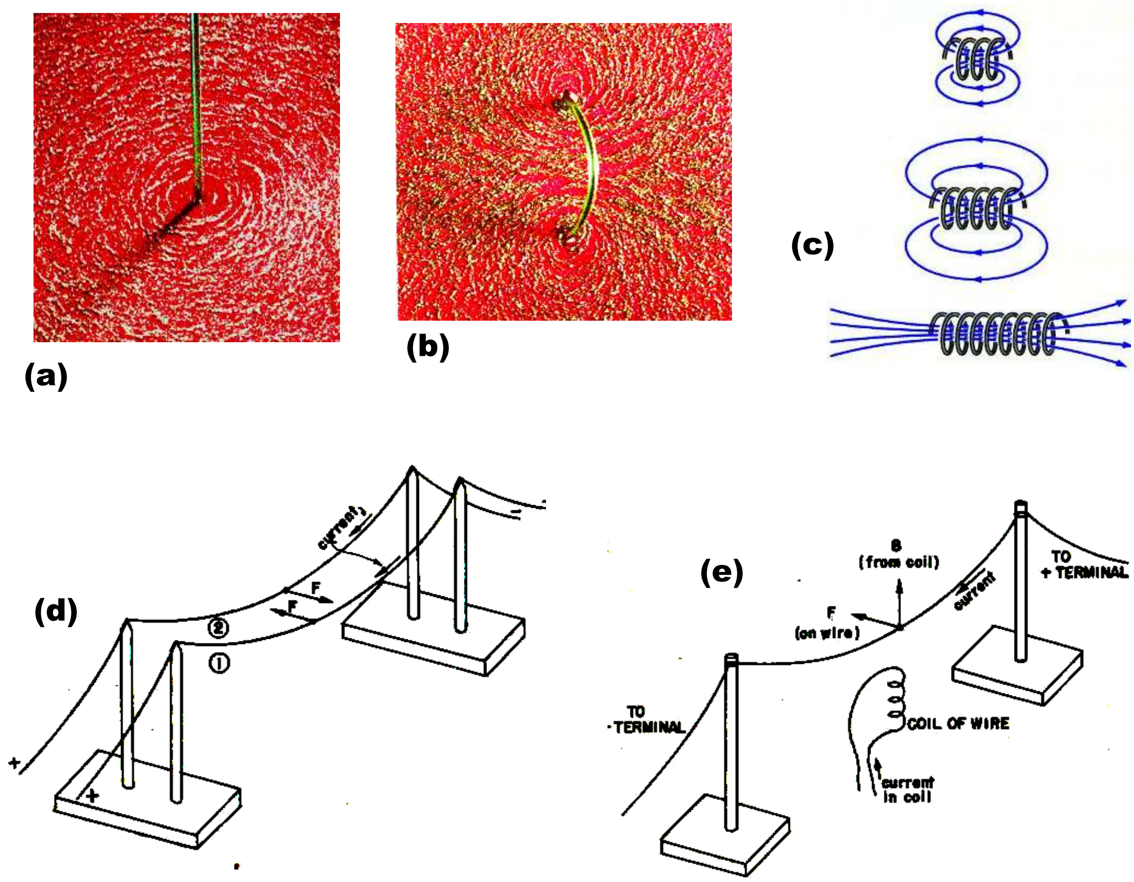

FIG. 2: Static magnetic fields from static current-carrying wires. In (a) we see the magnetic force field lines around a single wire, shown up by iron filings. In (b) we see the field lines near a current-carrying wire ring, and in (c) we show the magnetic field lines from solenoid wire loops of different lengths. In (d) we show at left the force between 2 parallel wires carrying current in the same direction (so they are attracted towards each other); and at right we show how a current through a coil will exert a force on the current-carrying wire above it.

moreover, we see that the field on one wire from the other is actually pointing vertically (up or down). However the force on each wire is not vertical, but horizontal - in a direction perpendicular to both the current and to the field. Thus we see that the force on a current moving in some direction, from a magnetic field which is perpendicular to the current, is perpendicular to both of them. We can explore this further if we take a solenoid, whose magnetic field we can control, and move it around a single current-carrying wire - see Fig. 2(e). By this means we establish the correct relationship between the directions of these 3 vectors (the current vector $\mathbf{J}(\mathbf{r})$ at some position $\mathbf{r}$, the field vector $\mathbf{B}(\mathbf{r})$ at the same position, and the force vector $\mathbf{F}(\mathbf{r})$ acting on the current, also at this position. People studying physics at high school learn this relationship as the 'right-hand rule'; it is not important here. What is important for our purposes is the simple fact that we see that

(i) an electric current causes a magnetic field, circulating around it; and

(ii) that field interacts with any other current around; and

(iii) since the second current is also producing a field, this field acts back on the first current by the same mechanism as in (ii).

Indeed, we can summarize all of these statements by saying that electric currents generate magnetic fields, and these fields act back on currents - so that eventually, 2 currents will interact with each other through the fields they are generating. Or to put it another way, the magnetic field mediates an interaction between 2 currents. And this is precisely what is shown in Fig 2(d). Note again that the static magnetic field has no effect on static electric charges - only the electric field can do this. Thus, the picture we have so far is that electric fields mediate interactions between static charges, and magnetic fields between static currents.

A.3: TIME-VARYING ELECTRIC and MAGNETIC FIELDS

So far we see no obvious connection between electric and magnetic fields. The key contribution of Faraday was to tie all of these observations together by showing what happened when things were no longer static. Faraday realized, first of all, that since currents were nothing but moving electric charges, and since a set of moving charges was necessarily generating electric fields which were changing in time, that the experiment in which electric filed were changing had already been done. Think, for example, of what you would see if a positive electric charge moved past you. At all times, you would see electric field lines radiating out from it - and the lines radiating in your direction would be coming towards you. But as it moves past, this means that both the direction of the field lines and their strength would seem to vary (in the same way that if a light moves past you, the direction of the light coming towards you would vary in strength and direction). Thus, Faraday reasoned, the magnetic field is connected to the fact that the electric field is varying in time.

The obvious question to ask then was - what happens if one varies a *magnetic* field in time?

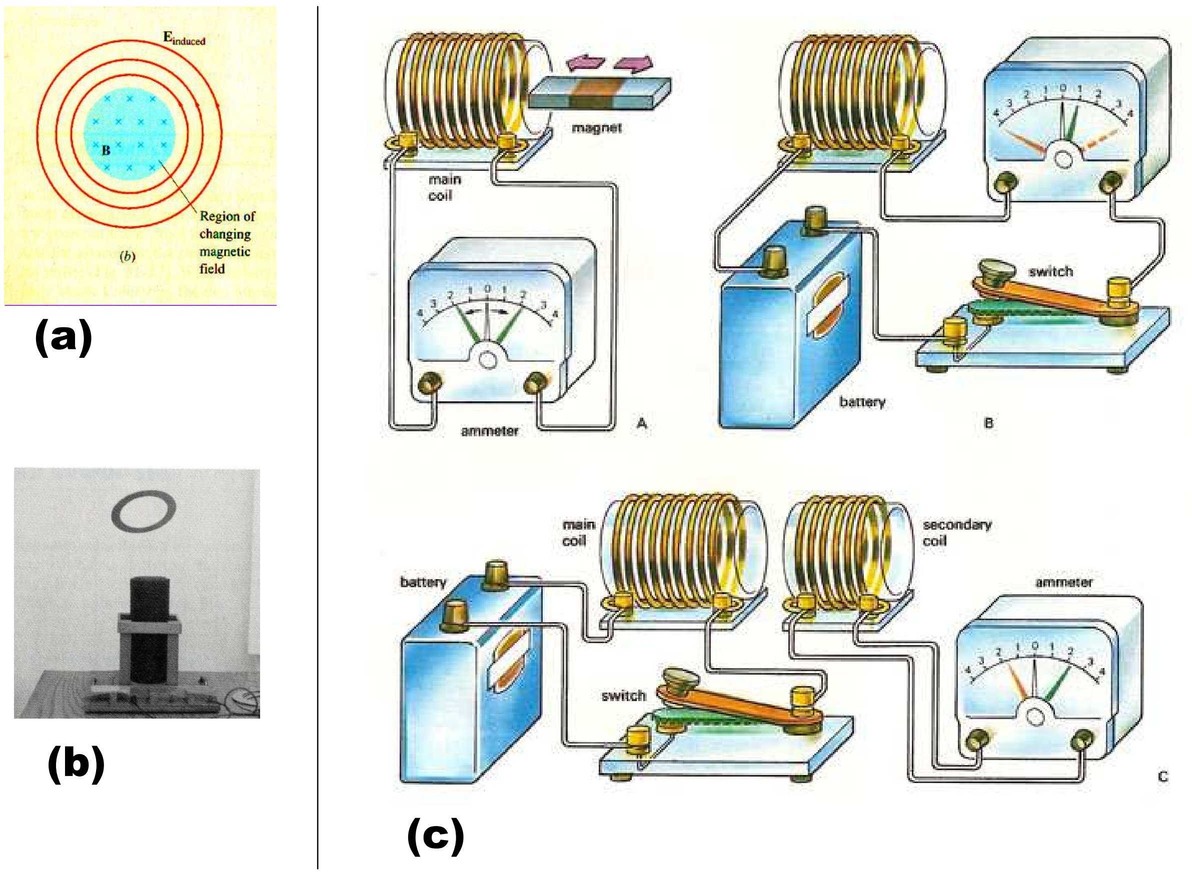

FIG. 3: Experiments showing electromagnetic induction. In (a) we show schematically how changing a magnetic field through a solenoid will create an electric field circulating around the changing magnetic field. In (b) we show an experiment in which a sudden change in the magnetic field through a solenoid will cause a metal ring around it to jump in the air (because the current created in the ring will then interact with the magnetic field causing a repulsive force which force which kicks the ring away from the solenoid). In (c) we see three situations; in A we see how moving a magnet near a solenoid will create current in the solenoid; in B we see how current from a battery in a solenoid creates a magnetic field; and in C we combine these, showing how suddenly switching on the current in the main solenoid coil will induce a current in the secondary one.

In Fig. 3 we see the answer to this question. Fig. 3(a) is a schematic diagram of an experiment that is easy to do. Suppose we take a solenoid through which current is running, so that there is a magnetic field running down the middle of the solenoid, along its axis. Fig. 3(a) shows the magnetic field 'going into the page - but we are also now to imagine that this field is *changing in time*. What happens? It was easy to show in experiments that an *electric field* was created, which circulated around the magnetic field! Note that *there are no charges associated with this electric field*; it is generated solely by the changing magnetic field. We see that this phenomenon - called electromagnetic induction - is the exact mirror of the way in which an electric current (a changing electric field) generates a magnetic field with no magnetic charges (ie., with magnetic field lines that only loop back on themselves, rather than going into or coming out of a charge).

Fig. 3(b) shows a very easy way in which this inductive effect can be seen. To show that there is an electric field being produced by a changing magnetic field, we put a conducting ring around the solenoid. The conducting ring is not connected electrically in any way to the solenoid - no charge can flow between them (it is easy to put a barrier between the two). Now, suppose we suddenly switch on a current in the solenoid, so that the magnetic field through the solenoid quickly jumps from zero up to some finite value. What happens then is dramatic - the metal ring around the solenoid suddenly jumps high into the air. Why has this happened?

The answer is as follows. The rapid change in the magnetic field through the solenoid has caused a large electric field to circulate around the solenoid, as shown in Fig. 3(a). Now this field acts on the electric charge in the metallic ring, and since charge can flow easily through a metal (it is a good conductor), a large electric current immediately begins to flow in the ring, around the ring, pushed by the electric field. So, we now have *two* separate electric currents, and each of them generates its own magnetic field. It then follows that each current will interact with the other current via these fields, and so each current-carrying system (the solenoid and the metal ring) will exert a force on the other. That this force is repulsive is rather obvious from what then happens - the ring is repelled high into the air by the force acting on it (and at the same time it removes itself from the source of the original field). Thus the field from the solenoid source and the 'back-reactive' field from the ring *act against each other*. That it happens this way, and that that the force is repulsive, is a manifestation of what is sometimes called "Lenz's law", viz., that the force generated in electromagnetic induction is always such as to cause a negative back-reaction on the process that is creating it. In fact it turns out that this result follows from energy conservation - if we could use electromagnetic induction to generate currents that back-reacted on the source so as to act with the source, rather than against it, we could continually draw energy from this process, and set up a 'perpetual motion' system, getting energy from nowhere (ie., 'getting something from nothing'). That energy is 'conserved' (ie., that total energy remains the same, so that energy acquired somewhere has to come from somewhere else) is one of the most fundamental principles in all of physics.

Fig. 3(c) picks these processes apart, and also demonstrates their 'conservative' nature (ie., that they are self-contained, with no energy exchange with or connection to the outside world). In A we see what happens if we move a bar magnet in r out of a solenoid coil - the changing magnetic field in the coil induces an electric circulating in the coil, and thence causes current in it, which is measured by the ammeter. In B we show the process also seen in Fig 3(b), where closing a switch allows current to suddenly flow in a coil (again, generating a magnetic field). Finally, in (c), we put the two together - on the left, a switch is closed, causing sufdden current in the left-hand coil, and a magnetic field to be generated. This time-varying field is picked up by the second coil on the right (some of the field from the left coil goes into the right coil), causing a circulating electric field in it, which is then seen in the ammeter. If these 2 coils were allowed to move, they would repel each other.

We see from experiments like these (and others also done by Faraday), that the electric and magnetic fields are inextricably related. Indeed, it seems that there is a remarkable symmetry between them - each is caused by changing the other. The symmetry is not complete, because while we have electric charges in Nature, there seem to be no magnetic charges - despite a very long search which still goes on, no "magnetic monopoles" (another name for magnetic charges) have ever been found. Thus while electric field can be created either by an electric charge, or by changing a magnetic field in time, a magnetic field can only be created by varying an electric field in time.

Faraday's work suggested that some great secret was waiting to be uncovered - a theory which would somehow unify all of these observations. This would have to wait another 34 years - science proceeded at a more leisurely pace in the 19th century.

The ELECTROMAGNETIC FIELD

In 1865, James Clerk Maxwell published "*A Dynamical Theory of the Electromagnetic Field*", the third in a series of papers on electromagnetic phenomena. In fact Maxwell had become interested in the work of Faraday as early as 1855, and by 1862 had already realized that light had to be an electromagnetic (EM) wave. But it was not until the 1865 paper that he achieved a full synthesis in the form of "Maxwell's equations, which have served ever since as the blueprint for field theory in physics. To appreciate the magnitude of his achievement, you have to realize that the language in which I have presented the experimental discoveries listed above is hardly related to that is which they were originally presented. In the work of Coulomb, Oersted, Ampère, Biot, and Savart, the idea of a field nowhere appears. One of Faraday's great innovations was to picture the magnetic and electric effects in space in terms of 'lines of force', but at that time there was no understanding of what might be the cause or seat of such lines of force - indeed, they were regarded as an abstraction.

Maxwell's theory replaced the abstract notion of lines of force with the definite notion of an EM field, which permeated all of space and which had its own dynamics properties. Maxwell's original presentation of this theory was in the form of 20 different differential equations, and this necessarily struck his contemporaries as being extraordinarily complex. Indeed Maxwell himself rewrote them in a quite different form in his great work, "*A Treatise on Electricity and Magnetism*", published in 1873; and writers since then have rewritten them in a variety of different ways (using the modern language of differential forms it is in fact possible to reduce them to a singe equation!).

Rather than wade through the morass of mathematical variants here, I will express the ideas in two much simpler ways. Let us imagine that instead of separate magnetic and electric 'lines of force', or fields, we imagine that there is only one field, which we call the electromagnetic or 'EM' field. Now, our first way of thinking about this is to simply say that we call 'electric' and 'magnetic' fields are now to be thought of as merely 2 different components of one larger entity, the EM field. Following more modern notation, we give these sub-components their own names - we call the electric field the "**E**-field", and the magnetic field the "**B**-field" (and in this language, the EM field is called the F-field; and **E** and **B** are components of F). This is a perfectly good way to talk about the whole thing. We then think of the EM field as having 2 components, each of which is itself a vector, and such that the two are not free to do their own thing - they are part of one entity, and so that behaviour of one is tied to that of the other. Moreover this F-field interacts with electric charges - the **E**-component can in fact produced by them, and acts back on them; and any motion of the charges also engages with the **B**-component.

So far so good. However there is another way of thinking about the EM field which turns out to rather interesting. This is to imagine that there is a more fundamental entity, which we will also call the EM field (the word 'field' unfortunately tends to refer to multiple objects), and the electric and magnetic fields are now simply different kinds of *distortion* of the EM field. To see what I mean by this, consider the following analogy. Imagine we have some completely transparent but deformable medium - a sort of "colourless jelly", which is quite invisible if undisturbed. Imagine now that if we stress the medium it acquires colour, becoming visible. Thus what we see would be the medium under stress - in fact we would see the stress of the medium. However, there are 2 kinds of stress we can apply. One is a compression or expansion - for example, we could squeeze the entire jelly uniformly, or we could imagine that stress was applied locally (eg., by putting some little bubble or grain in it, which would expand it in the vicinity of the grain). Let's suppose this kind of compressional stress would cause the medium to go red in colour. However there is another kind of stress, called a 'torsional stress', in which the medium is twisted. We could twist it locally by inserting a wire to which the medium sticks, and then turning it, forcing the medium to twist around it. Let's suppose this causes the medium to go blue.

Now it turns out that in some ways, this is a rather good analogy to the EM field. We imagine that blue torsional stress is like a magnetic **B**-field, and red compressional stress is like an electric **E**-field; and we further imagine that any change in time of one of these two fields causes the other to appear, just as in electromagnetism. The only further condition we require is that there be 'electric charges', ie., sources and sinks of compressional stress. Positive charges are then equivalent in the analogy to points where the medium is locally expanded; and points where it is locally contracted or compressed, correspond to negative charges. In this analogy, the EM field is like some sort of underlying 'aether', only visible when it is disturbed - and to disturb it, we need electric charges. And notice that this aether only interacts with charges - it does not in any way interact with neutral matter, so that a massive object like a planet (electrically neutral) will move through the EM aether with no resistance whatsoever, neither feeling it nor affecting it in any way.

It was already realized by Maxwell in 1862 that of the many consequences of his theory, one was paramount in importance. This was that the EM-field/EM aether would support wave-like oscillations, which would be able to propagate through it at very high velocity. This was extremely interesting, since these oscillations would necessarily be created by any charge that happened to oscillate in position; and moreover, any wave meeting a charge would then cause it to start oscillating. Thus electric charges could communicate oscillatory motion to each other over long distances.

It did not take Maxwell long to realize that the velocity that he calculated for his waves was precisely that of light - so that light could finally be understood as nothing but an EM wave, which could be emitted by oscillating charges, and absorbed by them. From this idea it was but a small jump to suppose that all of the properties of light could be understood in this way - different colours then corresponded to different waves, of different wavelengths. The phenomenon of colour was simply a manifestation of the different kinds of receptor in our eyes, sensitive to different wavelengths of light (ie., tuned to oscillate at different frequencies, so that some of them absorbed light of one wavelength, and others of other wavelengths. Maxwell was well positioned to understand this - he had spent years, when younger, studying the phenomena of colour and colour perception. Incidentally, we see this sort of coupling between oscillators and waves in a quite different way in our everyday lives, with sound waves and mechanical oscillators. A musical instrument or tuning fork is made so that it preferentially oscillates at certain frequencies, and emits sound waves at those frequencies. Moreover, if these objects receive sound waves at the same frequencies, they will then start to oscillate - they are absorbers/recievers as well as emitters. Our ear receptor system is nothing but a set of mechanical oscillators, tuned to receive various frequencies of sound vibration. Likewise our eyes are electromagnetic receptor systems, tuned to receive light of different frequencies.

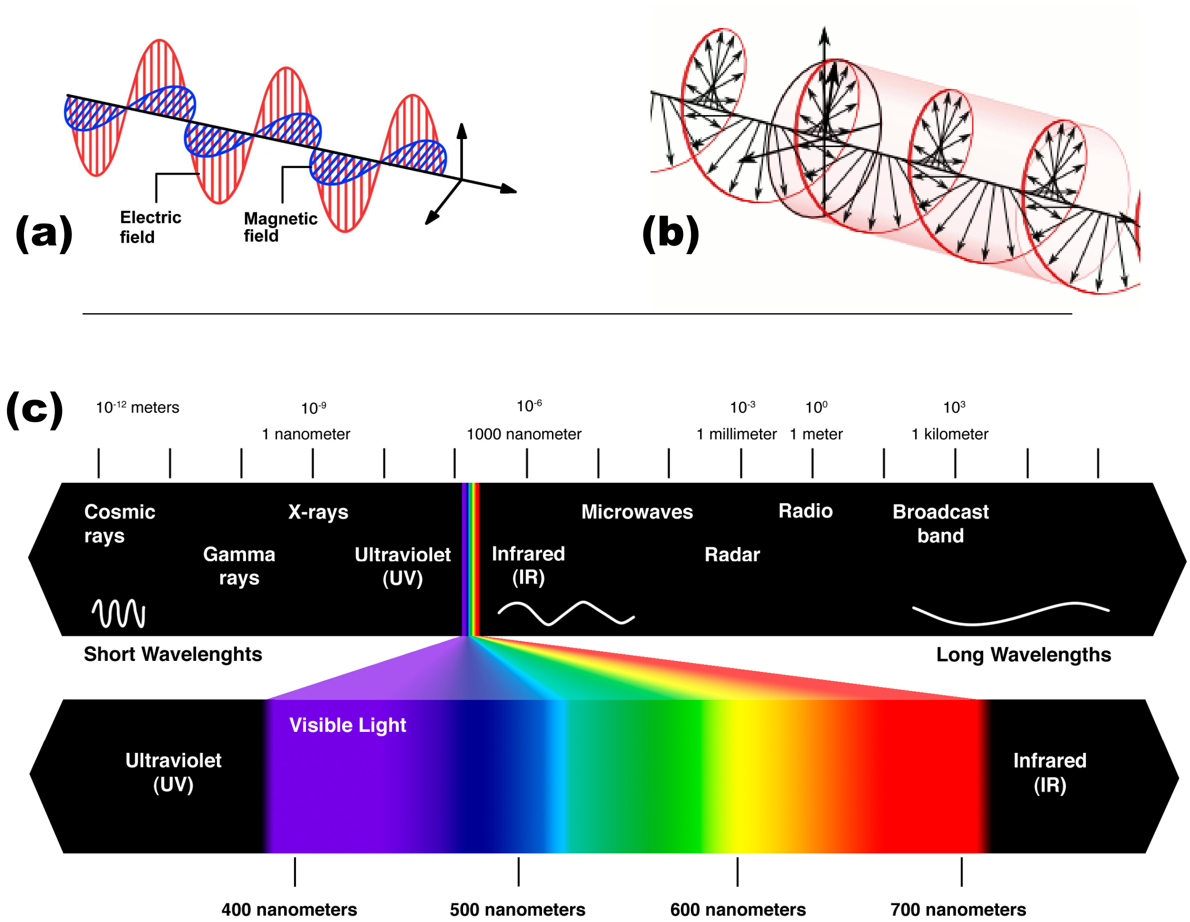

FIG. 4: Electromagnetic waves. In (a) we show how the electric and magnetic fields vary as we move down a linearly-polarized EM wave, and in (b) we see how the electric field varies as we move down a circularly-polarized EM wave (note that these field patterns are moving through space at very high velocity). In (c) we show the spectrum of different EM waves that arise when we vary their wavelength - moving left or right by one step changes the wavelength by a factor of 10 (except in the region of radar and radio waves, which has been compressed). The central part of this spectrum, which refers to visible light, is expanded in the picture below.

Of course the obvious inference one can draw from this is that there must be a whole variety of EM waves of different frequencies and wavelengths, none of which are visible to us. Fig. 4(c) shows what is called the "EM spectrum" of different EM waves. These are shown for different wavelengths, ranging from very short wavelength "gamma-rays" and cosmic rays, having very high frequency, to very long wavelength, low frequency radio waves. In a very narrow range in the middle we have light waves - their range is expanded vastly in the lower image of Fig. 4(c). Notice how narrow is the range of light - extending over less than a factor of 2 in wavelength, compared to the enormous range (a factor of roughly 10^{20}) between very long wavelength radio waves and extremely short wavelength cosmic rays.

And yet all of these waves have the same basic structure, which is shown in Figs. 4(a) and 4(b). In Fig. 4(a) we show what is called a 'plane-polarized' light wave. You should imagine that this wave structure is moving very fast from left to right, at the velocity of light. The wave has both **E**- and **B**-components, and both are oscillating; but they are perpendicular to each other, and out of phase (where there is a maximum amplitude of **E**, there is a minimum of **B**, and vice-versa). This configuration is not too difficult to understand. When **E** is going through a minimum, ie., it is roughly zero, it is also changing fast, so that it causes **B** to be large; and vice-versa. That they are perpendicular is also understandable - we say that changing **E** in one direction causes **B** to appear in a perpendicular direction (and vice-versa). In the slides I show one interesting device that uses the polarization of light, viz, adjustable polaroid glasses. Note, however, that there is no need for the EM wave to be plane-polarized - it can also be 'circularly polarized'. In this case the direction of **E** rotates around as it propagates - this is shown in Fig. 4(b) (note that the **B**-vector will also rotate, but again, out of phase with the **E**-vector).

One other thing is important to understand about these waves. That is that the wavelength and frequency are naturally related for a given wave - this is true of any wave. In fact we can say that if c_o is the velocity of light (equal roughly to 3×10^8 m/sec, or 300,000 km/sec), and λ is the wavelength of the light (the distance between two wavecrests), then the frequency f is given by

$$f = \frac{c_o}{\lambda} \qquad ; \qquad ie., \qquad \lambda = \frac{c_o}{f} \tag{0.2}$$

Let's consider 2 examples. Consider first orange light, which has a wavelength of about 600 nm (ie., $\lambda = 600 \times 10^{-9} m$, or 6×10^{-7}m). Its frequency is then very high - given by $f = c_o/\lambda = 3 \times 10^8/6 \times 10^{-7} = 5 \times 10^{14}$ oscillations per second. This is 500 trillion oscillations per second - this is how many wavecrests pass each second in an orange light wave. It is also how fast the electrons in a rod or cone cell oscillate per second when they absorb an orange light wave.

As a second example consider a radio wave with a frequency of 100 MHz (ie., $10^8 = 100$ million oscillations per second). This is the frequency of a typical radio statio (eg., the name "Rock 101" signifies a frequency of 101 MHz). What then is the wavelength? It is $\lambda = c_o/f = 3 \times 10^8/1 \times 10^8 = 3$m. This ia long wave, 10 ft in length - and yet EM waves travel so fast that 100 million pass each second.

Let's now pause to consider some of the deeper implications of the discovery that magnetism and electricity are connected into one entity, the invisible EM field, which permeates everywhere; and that light is nothing but an EM wave.

The first implication is essentially philosophical. We see that an enormous leap has been accomplished, in that a remarkable abstraction - that of the field - has been given flesh and bones, and connected through the idea of electric charge to a whole host of different phenomena in the real world. Consider the standpoint we started from, that of the Ancient Greeks, who posited on purely philosophical grounds that "Physical Reality" must have some component or components which 'underlie' the visible world, and which are more fundamental than the world of appearances. From this vantage point the idea of the EM field is a colossal advance, perhaps the greatest to occur in the intervening 2000 years. That the defining properties of the EM field and its interaction with charges can be summarized *completely* in a set of pure equations - this would have enormously impressed the Greeks, as it did everyone at the time of Maxwell. That the concept of the EM field, through its theoretical formulation, and the factual connection with observable phenomena, could explain a vast array of experimental phenomena, and be *tested* using these experiments - this would have enormously impressed empiricist philosophers from Bacon to Hume, and would doubtless have changed Kant's ideas. It also had large repercussions at the time, and influenced everyone from philosophical savants, scientists, and politicians, to novelists, and of course charlatans claiming an interest in spiritual media and the paranormal, who used these ideas to take cash off many an unsuspecting believer. A good picture of all this is provided by many British Victorian writers and poets.

Of course, the two thinkers who would have been most impressed of all by the results of Maxwell, would have been Newton and Huyghens. Indeed, we can see that the paradox that confounded both Newton and Huyghens (that their aether caused no resistance to matter moving though it, but was nevertheless able to affect light, which itself interacted with matter) was beautifully resolved. The resolution is simply that the EM "aether" interacts solely with charge, and not with mass - so that a planet can move through a vacuum with no resistance, even though the vacuum and the planet are permeated throughout by the EM field, because, being electrically neutral, the planet does not interact with the EM field (actually this is not quite right - planets do have some excess charge, and some have their own magnetic fields - but the corrections to planetary motion that these cause are extremely small). Maxwell's theory vindicated Huyghens over Newton, in that it made light an EM wave in an "EM aether", and confirmed his predictions that it moved more slowly in dense matter than in a vacuum - and it also explained the polarization properties of light, first studied seriously by Huyghens in the context of his wave theory. But Newton would doubtless have been satisfied that an explanation had been found in terms of forces on objects. Both would have realized how close they had come, with only the vital piece of information - the existence of electric charge, which coupled to electric and magnetic fields - unavailable to them. Newton would have been happy at the key role the 'experimental philosophy' had played - Maxwell had needed the remarkable experiments of his forebears to come to his theory. And Huyghens, with his deep understanding of the key role of prediction in science, would have been impressed at the way in which the remarkable predictions of Maxwell's theory had quickly cemented it as the right theory, in the minds of 19th century scientists.

It seems likely that both Newton and Huyghens would also have quickly realized that the next great question to be answered, again in terms of the field concept, was the nature of gravitation - clearly, in analogy with electromagnetism, the gravitational 'action at a distance' through a vacuum would have to explained in terms of some kind of field. But what was it? The answer to this question would have to wait for Einstein. And it is likely that both of them would have been very surprised at the answer, since it involved the overthrow of the aether idea itself - and the final nature of the gravitational field turned out to be very strange indeed. Perhaps even more extraordinary to both thinkers woud have been the subsequent evolution of the 'particle vs. wave' discussion of the natuer eof the light. The theory of Maxwell seemed to kill of the particle theory of light for good. The great denouement would have to wait this time on new experiments, and the subsequent discovery of quantum mechanics. The ultimate result - that of "wave-particle duality" - would have been quite unimaginable to them, and, one suspects, they would have been just as uneasy with it as everyone else has been since. But all this is to come.

Finally, one suspects that both Newton and Huyghens, along with Maxwell himself (but perhaps not Faraday) would have been quite astonished at the practical consequences of the discovery of magnetic induction, and of the laws governing electromagnetism. If one has to list the three main reasons for the huge technological revolution beginning in the early 19th century, and continuing until now, then they are undoubtedly (i) the understanding of mechanical motion (Newton) (ii) the understanding of heat and thermal phenomena, including chemical reactions (a development beginning with Carnot, and continuing with Helmholtz, Boltzmann, and Gibbs), and finally and most important of all, (iii) the elucidation of electromagnetism. There is not a single device in anyone's home of any sophistication that does not use, indeed rely completely, on our understanding of electromagnetic phenomena - from fridges to coffeemakers, lighting, heating and cooking systems, telephones, TVs, computers, cars, and of course the power to run them all. And very likely not a *single thing* in your house, or indeed anywhere else, that you bought, did not rely for its manufacture, at least at some stage, on electromagnetic devices and electrical power. To really appreciate the importance of this, just imagine what your life would become if all electrical devices were to suddenly disappear - all modern societies would quickly collapse, and necessarily revert to the kind of subsistence society currently prevailing in, eg., a few parts of rural Africa. This gives some indication of the enormous reliance we all place on this technology, and how much human society has been influenced by the work, 150-200 yrs ago, of a small group of researchers. What one thinks of this is of course another story.