1: Recapitulation and re-expression in terms of Fields

So far we have considered the discovery and development of several phenomena and associated ideas that were important for the subsequent growth of understanding of the first set of approximate Laws of Nature concerning electricity and magnetism. Let’s review them before pressing on.

The first we’ve only hinted at in our discussion of the work of Gilbert. The reason is that it’s no longer regarded as more than phenomenological, i.e., rather superficially descriptive. But it had been regarded as important right from Gilbert’s time until the late eighteenth – early nineteenth century. It is the rule (Fig. V. 1):

(1) GILBERT’S RULE: ‘PERMANENT’ MAGNETIC SYSTEMS EACH CONTAIN TWO CENTERS OF MAGNETIC FORCE, TRADITIONALLY CALLED NORTH AND SOUTH MAGNETIC POLES. ANY TWO NORTH POLES REPEL EACH other AS DO ANY TWO SOUTH POLES. ANY NORTH POLE AND SOUTH POLE WILL MUTUALLY ATTRACT ONE ANOTHER. THE STRENGTHS OF THE TWO POLES IN ANY ONE SYSTEM ARE EQUAL. IF A MAGNETIC SYSTEM IS SEPARATED INTO PARTS, EACH CONTAINING ONLY ONE OF THE ORIGINAL POLES, EACH PART WILL THEN BE FOUND TO CONTAIN TWO OPPOSITE POLES, THE ORIGINAL POLES HAVING ‘DISAPPEARED’.

After the discoveries of Oersted and Ampere scientists came to regard the generation of magnetic forces by currents, or moving charges, as more
fundamental than natural magnets and they then tried to understand natural magnets in terms of hypothetical microscopic currents.

On to the other rules:

(2) **COULOMB’S LAW:** EVERY TWO CHARGED PARTICLES EXERT AN ELECTRIC FORCE ON ONE ANOTHER (LIKE CHARGES REPEL AND UNLIKE CHARGES ATTRACT), THE MAGNITUDE OF WHICH IS PROPORTIONAL TO THE PRODUCT OF THE CHARGES AND INVERSELY PROPORTIONAL TO THE SQUARE OF THE DISTANCE BETWEEN THE PARTICLES.

This law is fundamental in the sense that it holds very precisely for charges that aren’t moving too much and the discovery of its’ generalization for moving charges did not occur until the mid nineteenth century and required a two part reformulation in the context of the concept of the **electric field**. We will discuss that reformulation next time.

But we can already express the present form of Coulomb’s Law in terms of the **electric field** as follows: Each charged particle is accompanied, in the space surrounding it, by an **electric field** which can be represented mathematically by a distribution of vectors throughout space that everywhere point in the direction that a second positive charge would be pushed if it were also present (Fig. V. 2). This direction is radially outward if the source charge is positive, and radially inward if the source charge is negative. The magnitude of the distributed vectors is given by,

$$ E = k |q| / d^2 , $$

where |q| is the magnitude of the source charge and d is the distance of the location of the field vector from the source charge. In the presence of many charges, the **total electric field** at any point of space is the resultant of the electric fields from each charge (Fig. V. 3).

For any given point charge the electric field that affects it is the electric field due to all other charges only. If $E$ is the total electric field affecting a point charge, $q$, then the electric force on that charge is given by,

$$ F_E = q E . $$
Fig. V. 2: Directions of electric fields surrounding single point charges

The charges are accompanied by fields. The fields exert forces on the charges.

(3) FRANKLIN’S LAW OF CONSERVATION OF CHARGE:
ELECTRIC CHARGE IS NEVER CREATED OR DESTROYED. IN ANY ELECTRIC PROCESS OR PHENOMENA THE ALGEBRAIC SUM OF ALL THE CHARGES INVOLVED, POSITIVE AND NEGATIVE, REMAINS CONSTANT.

Again, this remains exactly true, so far as anyone knows. The only important modification is to focus on the relation between changing total charge within a finite volume and the flow of current through the surface bounding the volume. We will discuss that “local” version of the conservation of charge shortly.

(4) VOLTA’S RULE: DISSIMILAR CONDUCTORS, CONNECTED BY A SALINE OR ALKALINE SOLUTION WILL BECOME OPPOSITELY CHARGED TO AN ELECTRIC PRESSURE (VOLTAGE) CHARACTERISTIC OF THE PAIR OF CONDUCTORS AND, WHEN PROVIDED WITH A CONDUCTING PATH BETWEEN THEM, WILL REPLACE THE ELECTRIC CHARGE WITH A ‘CONSTANT’ ELECTRIC CURRENT FROM THE ORIGINALLY POSITIVELY CHARGED CONDUCTOR TO THE ORIGINALLY NEGATIVELY...
CHARGED CONDUCTOR. FOR A GIVEN CONDUCTING PATH, THE CURRENT WILL BE PROPORTIONAL TO THE ELECTRIC PRESSURE.

The ‘constancy’ of the electric current is eventually undermined by chemical changes that take place in the solution and on the surfaces of the conductors sitting in the solution. This fact gradually gave rise to the recognition that this phenomena of spontaneous current generation is due to an electro-chemical process at the molecular level. But this rule, itself, while true, is regarded as merely phenomenologically descriptive and not fundamental.

(5) OERSTED’S OBSERVATION: A CURRENT CARRYING CONDUCTOR EXERTS A MAGNETIC TORQUE ON ANY MAGNET BROUGHT NEAR THE CONDUCTOR. THE TORQUE, IF UNOPPOSED, WILL TWIST THE MAGNET SO THAT THE LINE JOINING THE SOUTH POLE TO THE NORTH POLE OF THE MAGNET IS PERPENDICULAR TO THE CONDUCTOR. THE RELATIONSHIP OF THE S→N LINE IN THE MAGNET TO THE DIRECTION OF CURRENT FLOW SATISFIES A RIGHT HAND RULE: WITH THE THUMB OF A RIGHT HAND POINTING IN THE CURRENT FLOW DIRECTION THE CURLED FINGERS OF THE RIGHT HAND WILL POINT IN THE S→N DIRECTION (Fig. V. 4).

This observation contributed to the growth of acceptance of the concept of a magnetic field that exists in the vicinity of a current, as well as around a magnet. Again, the magnetic field is represented by a spatial distribution of vectors, the direction of which, at any point, is the direction into which the S→N line of a small magnet at that point would be twisted.

(6) AMPERE’S RULE: ANY TWO LONG STRAIGHT PARALLEL CONDUCTORS CARRYING CURRENT EXERT MAGNETIC FORCES PER UNIT LENGTH ON ONE ANOTHER (LIKE CURRENTS ATTRACT, UNLIKE CURRENTS REPEL), THE MAGNITUDE OF WHICH IS PROPORTIONAL TO THE PRODUCT OF THE CURRENTS AND INVERSELY PROPORTIONAL TO THE PERPENDICULAR DISTANCE BETWEEN THE CONDUCTORS, i.e.,

\[ f = \kappa \frac{I_1 I_2}{d_\perp}, \]

as we saw before.
Fig. V. 3: Resultant electric field configurations for two point charges; unlike charges at the top, like (positive) charges at the bottom.
Actually Ampere had formulated a much more detailed rule for the magnetic interaction between any two minute portions of current carrying conductors having any directional relationship to one another whatsoever. The simpler rule I’ve presented here is derivable from his more comprehensive rule, which, unfortunately, is mathematically rather complicated. Also, like Coulomb’s Law, it holds precisely only for non-moving conductors.

By analyzing the combination of Oersted’s observation and Ampere’s more detailed rule, Biot and Savart were able to formulate a rule, again rather complicated, for the **total magnetic field** that would accompany an arbitrarily shaped (non-moving) collection of current carrying conductors. One can then recover Ampere’s general rule by stipulating the force a magnetic field exerts on a current.

As with electric fields the **total magnetic field** is the *resultant* of the magnetic fields coming from all the parts of the current carrying conductors *(Fig. V. 5).*
The force exerted by a magnetic field on a current carrying conductor is more complicated than the force of electric fields on charges and again involves a right hand rule. If, at some point, there is a current, $I$, in a conductor and a total magnetic field, $B$, from all the other currents, then the magnitude of the magnetic force per unit length of conductor on $I$ is,

$$f_M = I B_\perp = I_\perp B,$$

where $B_\perp$ is the magnitude of the component of $B$ perpendicular to $I$ and $I_\perp$ is the magnitude of the component of $I$ perpendicular to $B$. The direction of this force is given by the stretched thumb of the right hand if the extended index finger points in the direction of $I$ and the extended second finger points in the direction of $B$ (Fig. V. 6).

The currents are accompanied by fields. The fields exert forces on the currents.

2: Applications and Impediments

We have already considered enough material to be in a position to understand the basic ideas behind an electric motor and an electric generator,
Fig. V. 6: Right hand rule for the magnetic force on current, i.e., moving positive charge or oppositely moving negative charge.

i.e., the conversion between electrical energy and mechanical energy. But this is not the way it happened historically.

Electricity was new and people were not at all clear on the electrical composition of matter, of atoms and molecules. They had to wait for the original formulation of electromagnetic induction by Michael Faraday in the 2nd quarter of the nineteenth century. Very quickly thereafter, however, both electric generators and motors were developed. The very first of each, not very impressive but working nonetheless, were invented by a Professor dal Negro of Padua University shortly after Faraday’s announcement of his discovery. Almost immediately the electromagnetic technological revolution went into high gear.

But we can understand motors and generators, basically, from our present perspective (Figs. V. 7, 8). In the simplest case of a motor (Fig. V. 7) a
conducting wire is laid lengthwise back and forth along opposite sides of a rotating axle in such a way that, as the axle rotates, the two extending ends of the wire come alternately in contact with the two electrodes of a battery.

Fig. V. 7: Basic electric motor design. A is the rectangular wire winding, B is the magnetic field, C is a semi-circular contact, D is a battery electrode and MM is the axle about which the wire winding and the semi-circular contacts rotate.

This, by itself, would force current through the wire in alternating directions as the axle rotated. But what rotates the axle? Well, if the whole assembly is placed so that the axle with the current carrying wire lies between the opposite poles of a permanent magnet, then Ampere’s Law will dictate that the magnetic force on the wire will rotate the axle. The axle needs some rotational inertia so that it keeps rotating during the brief periods when the wire is not in contact with the battery electrodes. But that can be achieved. Many electric motors, from the simplest toy motors in children’s toy stores to some of the most powerful giant motors of industry are based on this arrangement. Electrical energy in the battery is converted to the mechanical energy of the spinning winding, contacts and axle.

An electric generator uses the same arrangement of a conducting wire on an axle between the poles of a magnet (Fig. V. 8). But now the axle is rotated mechanically. The electrons and protons of the conductor are subject to opposite forces as a consequence of moving through the magnetic field. The protons, held rigidly in the conductor, cannot respond to the magnetic force, but the free electrons can. They flow through the conductor in the opposite direction of the current because they are negatively charged. As the axle
goes through a full rotation the current in the axle conductor alternates in direction. Depending on how the free ends of the axle conductor contacts the external circuit, the external circuit receives either alternating or direct current.

![Fig. V. 8: Basic generator design. By mechanically rotating the rectangular wire loop and circular contacts the fixed magnetic field induces an alternating current in the wires which then passes through the ‘brush’ contacts to the external circuit or ‘load’.](image-url)

As soon as Volta had discovered the battery the development of more and more powerful batteries was in demand. For some time they were just bigger and bigger versions of the Voltaic Pile. But they could soon produce substantial sparks and the heating, not to say melting, of conducting wires. This led to the growing awareness of the importance of resistance to current flow in conductors.

Remember that the electric field exerts a force on a charge. If that force is not balanced by some other force, the charge, by Newton’s 2nd Law, will accelerate to higher and higher speeds. While researchers in the late 18th century didn’t know exactly what was moving in current flow, they increasingly understood that higher currents meant more or faster charge flow. Conductors melted when the charge flow through them got too intense. Even when it could be safely maintained, conductors got hot, or at least
warm. In those cases, what other forces in the conductors were balancing the electrical forces trying to accelerate the charges?

In those days the natural assumption was some kind of internal frictional resistance to the flow. It was known that friction could make things hot and the fledgling research into the nature and properties of heat was contemporaneous with early 19th century researches into electricity and magnetism. We, convinced of the existence of atoms and molecules as the building blocks of matter, and of electrons as the moving constituents of currents, recognize the “friction” or resistance as due to collisions of the accelerating electrons with the vibrating molecules of the conductor.

If the electric field inside the conductor is not too intense, it accelerates the electrons at some rate for some short time. They pick up speed and then collide with a molecule and lose some of their speed, maybe all of it. Sometimes they’re even thrown backwards by the collision recoil. The molecules are VERY massive compared to an electron! Then they get accelerated again until the next collision. Quickly we reach the situation where the colliding electrons are losing energy to the molecules, on average, just as fast as they’re gaining energy from the electric field. The molecules absorb the energy as vibrations and the molecules on the surface of the conductor transmit that vibratory energy to molecules in the air. If that can be done fast enough a steady state is reached in which the conductors give off heat energy as fast as the current electrons lose kinetic energy to collisions. The rate at which the conductor can give off heat energy increases with its temperature.

If the electric field is too intense the temperature that the conductor has to reach before the steady state is achieved may be high enough for the conductor to start to glow or even melt.

The quantitative rule that governs the relationship between the electric field in a conductor, the current in the conductor and the heat generated when a steady state is achieved was obtained by Humphrey Davy and George Simon Ohm by 1826.

For a long thin conductor of length, L, with a cross sectional area of A (Fig. V. 9), the current I is proportional to the product, A E , where E is the electric field strength and the proportionality constant depends on the
conducting substance and is called the **conductivity** of the substance. It’s usually denoted by the Greek letter, $\sigma$, and we have,

$$I = \sigma A E.$$ 

![Diagram](diagram.png)

**Fig. V. 9:** Cylindrical conductor of resistance, $R = L / \sigma A$.

On the other hand, the **voltage difference** between the ends of the conductor, which is the work per unit charge done by the electric force, is given by,

$$V = L E.$$ 

Combining these two relations we have,

$$V = L \left( \frac{I}{\sigma A} \right) = (L/A) I.$$ 

Defining the **resistance** as $R := (L/\sigma A)$, we have, finally **Ohm’s Law**,

$$V = I R.$$
Since the current flow doesn’t increase with time (steady state condition) the work done on the current is deposited in the conductor as heat. The rate of heat deposition is

\[(\text{work per unit charge})(\text{charge per unit time flowing through conductor})\]

\[= V I = I^2 R.\]

For fixed \(R\), increasing \(V\), and therefore \(I\), by a factor of 2, 3, 4, --- increases the heat deposition rate by factors of 4, 9, 16, --- .

With a properly resisting conductor in a chemically inert environment raised to a brightly glowing temperature, we have the basic idea of the incandescent light bulb.

3: Electromagnetic Induction

In 1831 Michael Faraday, arguably the greatest experimental physical scientist of the 19\(^{th}\) century, possibly of all time, announced his discovery of what came to be called electromagnetic induction, that a changing magnetic field in the vicinity of a conductor will induce a current in the conductor. The simplest way to see this is just to move a magnet near a wire connected to a galvanometer, i.e. a current meter (Fig. V. 10).

Fig. V. 10: Electromagnetic induction: pulling the magnet through the coil produces a current pulse in the circuit.
Rather quickly it came to be understood that what the changing magnetic field really does is produce an electric field at the same place. If a conductor is present, the induced electric field produces the current.

Expressed in terms of averages, the final precise version of Faraday’s discovery, formulated by the theoretician, James Clerk Maxwell, can be expressed in words. But some preliminaries are needed.

Suppose there are electric and magnetic fields present at all the points of space in some region.

Now consider an open surface, \( S \), of area, \( A \), surrounded by a boundary curve, \( C \), of length, \( L \), both lying within the region. Then imagine a choice being made to call one side of \( S \) the positive side and the other side the negative side and a corresponding choice made to call one direction around \( C \) the positive direction and the opposite direction the negative direction (Fig. V. 11). Having done this, the surface and curve are said to be oriented and will be denoted by \( S^\wedge \) and \( C^\wedge \) respectively. Finally, suppose these two choices are correlated so that if the thumb of the right hand points through the surface, \( S^\wedge \), from the negative side to the positive side, then the curled fingers of the right hand point around the curve, \( C^\wedge \), in the positive direction. When all this is done we say the surface and boundary curve have been jointly oriented in accordance with the right hand rule, or simply, right hand oriented.

![Fig. V. 11: Open surface and boundary curve with right hand orientation.](image-url)
Next consider the average of the algebraic magnitude of the tangential component of the electric field around $C^\wedge$. The algebraic magnitude is equal to the magnitude at any point on $C^\wedge$ where the tangential component of $\mathbf{E}$ points in the positive direction around $C^\wedge$ and is the negative of the magnitude at any point on $C^\wedge$ where the tangential component of $\mathbf{E}$ points in the negative direction around $C^\wedge$. This average will be denoted by $\langle E_T \rangle_{C^\wedge}$. Similarly, consider the average of the algebraic magnitude of the normal or perpendicular component of the magnetic field, $\mathbf{B}$, over the surface, $S^\wedge$. This algebraic magnitude is related to the positive and negative sides of the oriented surface, $S^\wedge$, in the same way as the previous algebraic magnitude was related to the positive and negative directions around the oriented curve, $C^\wedge$. This average will be denoted by $\langle B_N \rangle_{S^\wedge}$. We can now assert:

**FARADAY’S PRINCIPLE OF ELECTROMAGNETIC INDUCTION:**
FOR ANY RIGHT HAND ORIENTED SURFACE, $S^\wedge$, OF AREA, $A$, AND BOUNDARY CURVE, $C^\wedge$, OF LENGTH, $L$, WE HAVE

$$L \langle E_T \rangle_{C^\wedge} = -A \left[ \text{rate of change of } \langle B_N \rangle_{S^\wedge} \right].$$

An immediate application of this principle is the production of induced currents. Suppose we have a time dependent voltage, $V_p$, sending a time dependent current through a circuit, called primary, which includes a cylindrical coil. The magnetic field inside that coil will vary with the current flowing through the circuit and if another coil is placed inside the first coil, the varying magnetic field from the first coil will induce an electric field along the windings of the second coil and, thereby, cause a current to flow in the second coil. If this second coil is part of a circuit, called secondary, with terminals, then the induced current will generate a voltage difference, $V_S$, across the terminals. With care in the design and construction of the circuits and coils, the primary current can be made proportional to the voltage difference, $V_p$, and the magnetic field inside the primary coil, proportional to the current. The induced electric field, induced current and secondary voltage difference can then be proportional to the rate of change of the magnetic field and, thus, $V_S$ be proportional to the rate of change of $V_p$ (Fig. V. 12). With complicated circuit design, a wide variety of subtle variations on the relation between the secondary and primary voltages can be produced.
Roughly \( V_s \propto \text{rate of change of } V_p \), i.e., \( \Delta V_p / \Delta t \)

**Fig. V. 12:** A typical wiring arrangement for electromagnetic induction of current in one circuit (secondary) due to varying current in another circuit (primary).