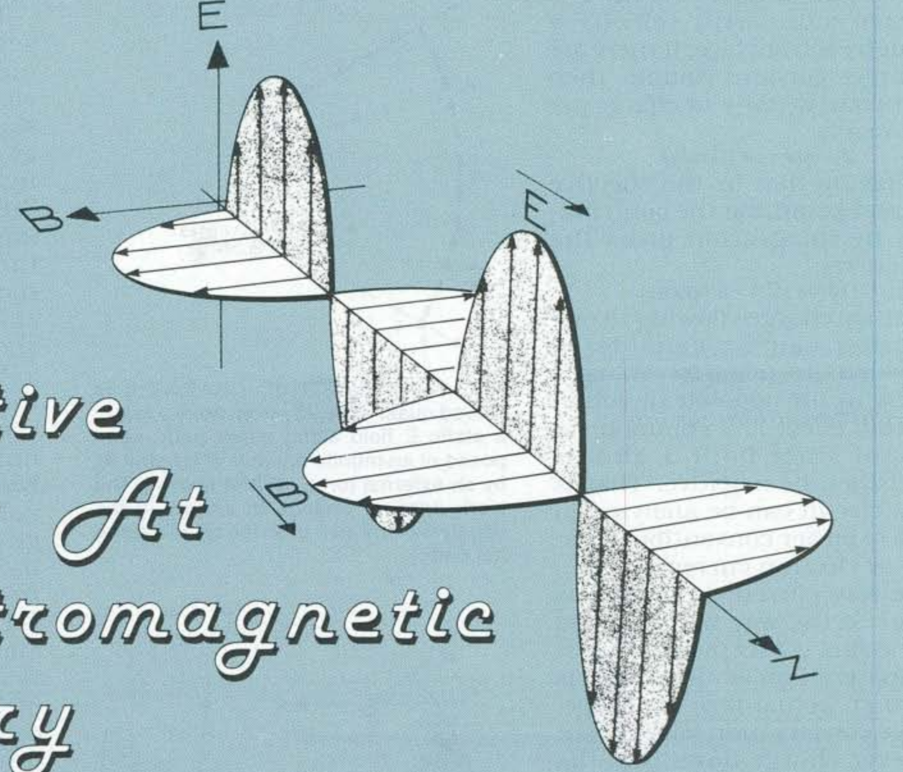
An Intuitive Look At Electromagnetic Theory



WILLIAM P. RICE

LAST TIME WE PRESENTED GENERAL concepts of electric fields and how they are related to static electric charges. We saw that the **E** field in empty space accounts for the forces between such charges. In this article, we'll see how the familiar units of volts and amperes are related to each other. Ohm's law and the concept of an **E** field in materials will be discussed with the help of a simple quantum theory viewpoint.

Potential

To quasi-statically move a charge q from point a to point b in an **E** field, a force that is infinitely close to being equal and opposite to the Coulomb force must be applied to q. That force is $-q\mathbf{E} = -\mathbf{F}_c$, as shown in Fig. 1. As we discussed in our previous article, when moving around a closed path

$$\oint \mathbf{E}\cdot d\mathbf{l} = 0,$$

or

$$\nabla \times \mathbf{E} = 0$$

at all points. So in moving the charge around a closed path

$$-\oint q\mathbf{E}\cdot d\mathbf{l} = 0.$$

The dot product gives the magnitude of force times distance in the direction moved, which is the work done or change in the potential energy ΔU. The energy expended in moving along the path from a to b is just the sum of the contributions along that path, as defined in the calculus notation

$$\Delta U_{ab} = -\int_a^b q\mathbf{E}\cdot d\mathbf{l} \text{ (newton} \times \text{meters = joules).}$$

The energy change is independent of the path taken from point a to b, and the **E** field follows the laws of conservation; whatever energy is expended in moving the charge from point a to b is recovered when the charge moves from b to a. The energy is said to be stored in the **E** field since the field is responsible for the force.

Dividing by the charge gives us the change in energy per unit charge, the potential or voltage at point b with respect to a is

$$V_{ab} = \frac{\Delta U_{ab}}{q} = -\int_a^b \mathbf{E}\cdot d\mathbf{l} \text{ (joules / coulomb = volts).}$$

The use of the name potential is perhaps unfortunate because it's easy to confuse the term with potential energy.

Recall also that since $\nabla \times \mathbf{E} = 0$, **E** must be the gradient of a scaler field, which we now see is the potential V, therefore

$$\mathbf{E} = -\nabla V \text{ (volts/meter = newtons/coulomb).}$$

Along a surface of equal potential, there would be no change in V per length d**l**. Perpendicular to that surface the change in V per length would be a maximum, which is what the gradient tells us.

Since the field is obtainable by linear superposition, the potential difference is simply the sum of the potentials. For example, $V_{ac} = V_{ab} + V_{bc}$. That analysis is the basis of Kirchoff's voltage law, which states that the algebraic sum of the voltage rises and drops around a closed path must equal zero.

Electric current

Imagine a Gaussian surface in space through which a number of q charges are moving, as shown in Fig. 2. (We are not concerned with the type of field influencing the motion, only that there is motion.) The current across that surface is defined as the charge per unit time (in seconds) crossing the surface. In order to calculate that, divide the surface into an infinite number of infinitesimal surfaces, d**s**. The charges move with velocity **v** through each surface. If there are n charges per unit volume, then the current density, or charge per unit area is

$$\mathbf{J} = nq\mathbf{v} = \rho\mathbf{v} \text{ (C/m}^2\text{s)}.$$

Multiplying that by the effective area and summing the contributions by integration gives the total current

$$I = \int \mathbf{J} \cdot d\mathbf{s} \text{ (C/s = amperes)}.$$

Positive charges flowing in one direction can be considered equivalent to negative charges flowing in the opposite direction (the Hall effect is a common exception) since both **J** and d**s** would then be negative. That is why a circuit can be analyzed in terms of either conventional currents or electron currents.

The way current is defined is similar to the way we explained electric flux ω except that flux is an apparent flow while current is due to an actual flow of charge. Charge conservation tells us that whatever charge flows into the surface must also flow out unless the current density inside is changing in time. That is the basis of Kirchhoff's current law, which tells us that the sum of the currents flowing into a junction is equal to the sum of the currents flowing out of that junction. Shrinking the Gaussian surface down to a single point and taking the ratio of the rate of change in current to the rate of change in volume gives the divergence

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t} \text{ (C/m}^3\text{s)}.$$

The partial differential symbol ∂, as in d, means an infinitesimal change in something. It also reminds us that we're only interested in ρ's change with respect to time, t. The negative sign indicates that a decrease in ρ, a negative $\partial\rho/\partial t$, gives a positive divergence. The net charge must therefore flow out through the surface.

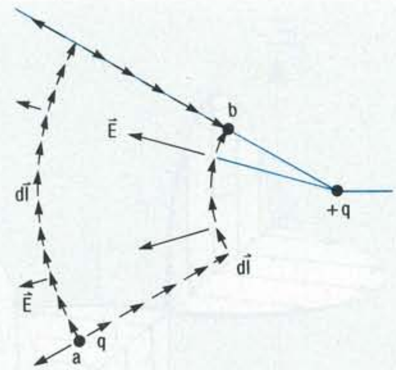


FIG. 1—AN ELECTRIC CHARGE q is moved quasi-statically from point a to b in a static E field along either path, composed of an infinite number of lengths dl, by an external force qE (not shown). The work done or change in energy is the negative of the sum of all the qE·dl's along the path.

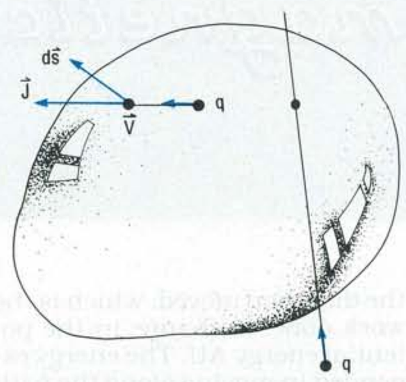


FIG. 2—CURRENT DENSITY J is the number of charges q per unit volume moving with velocity v through an infinitesimal section ds of the Gaussian surface. The total current is found by summing J·ds over the entire surface. Any charge that comes in through one ds must leave through another. Any net outflow must be at the expense of the charge density enclosed by the surface.

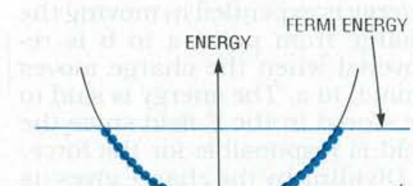


FIG. 3—ENERGY VERSUS MOMENTUM for electrons in a material. Temperature and lattice effects are neglected. Each electron, represented by a dot on the curve, has a unique energy state. Those are the lowest states available. The highest occupied energy is called the Fermi energy.

Conductivity

Up until this point we have been concerned only with charges in empty space. The space of solid materials, however, is far from empty. Atoms are located at positions called lattice points. An external **E** field applied to a solid material causes the electrons with a $-e$ charge to move. Quantum theory must be used to describe the effects of temperature and the lattice upon the motion of charges.

The electrons are in a state described by their energy, momentum, and spin. No two electrons can be in the same state. Electrons can change energy only by moving to a neighboring unoccupied energy state. Figure 3 shows the energy versus momentum states, neglecting the effects of temperature and the lattice. The two possible spin states for each electron are not shown for clarity.

The more electrons there are in the material, the higher the highest occupied energy state, or Fermi level. Only electrons near the Fermi level can respond to external effects such as thermal energy and electric fields. Supplying thermal energy excites some electrons to energies just above the Fermi level, leaving unoccupied states just below. The Fermi level is then taken as the energy with 50% occupancy. Electrons that can change energy, and hence momentum, are called conduction electrons. Thermally excited electrons have random momentum and velocity, and do not produce a net current.

Electrons act as waves and, therefore, experience interference effects due to interaction with the lattice. At certain wavelengths, standing waves result which produce energy gaps, as shown in Fig. 4. If only some of the energy states up to the gap are occupied or the gap is very small, the material will have many conduction electrons since little external energy is required to excite an electron to a higher state. Such materials are good electrical conductors. A good insulator (or dielectric) has occupied states up to a relatively large gap. A large amount of external energy is required to excite electrons to higher energies in a dielectric material. A material with a large gap and many occupied lower states exhibits noticeable electrical resistance.

If a potential difference is maintained across a material, an electric field is established. Conduction electrons will be subjected to a force **F**, which is equal to $-e\mathbf{E}$. Electrons tend to accelerate, and then "collide" and lose energy to the lattice. If τ is the average time between collisions, which is temperature dependent due to thermal motion of the lat-

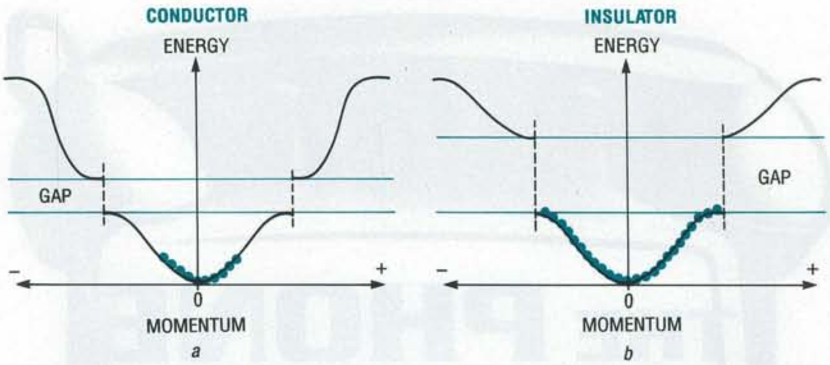


FIG. 4—ENERGY VERSUS MOMENTUM for electrons is a lattice of atoms. The gaps in the curves result from interference effects with the electron waves. In a conductor (*a*) the levels below the gap are partially occupied. External energy excites electrons to the unoccupied energy states. That allows them to participate in an electric current. In an insulator (*b*) the levels below the gap are filled and the energy gaps are large. Electrons cannot participate in a current unless a large amount of external energy is supplied.

tice atoms, then the average electron momentum is

$$\mathbf{F}\tau = -e\mathbf{E}\tau = m + \mathbf{v} \ (N \cdot s = kg \cdot m/s)$$

where m is the electron mass, and **v** is the average velocity. Solving for the velocity and substituting into the equation for current density gives us

$$\mathbf{J} = \frac{ne^2\tau}{m}\mathbf{E}$$

which is the vector form of Ohm's law. Since the number of electrons n and τ are properties of the material, the conductivity

$$\sigma = ne^2\tau/m \ (C^2s/kg = 1/(\Omega \cdot m))$$

is a property of the material. The resistivity is defined as $r = 1/\sigma$. If the material is of uniform cross-sectional area S and of length L, **J** is uniform and normal to d**s**, therefore the current is

$$I = JS = \sigma \frac{V}{L} S$$

or V = IR where R = rL/S is resistance in more familiar units of ohms.

In metals, increasing the thermal energy excites electrons mainly into the unoccupied states of the lower band, but the time between lattice collisions decreases. Increasing the temperature increases the resistance. In some other materials resistance decreases with increasing temperature because the number of conduction electrons exceeds the effect of increased collision time.

Due to the low velocity of electrons in most solids, the magnetic effects can be neglected. Conduction becomes more complicated in gases and liquids since the atoms can also move, and velocities can become greater than in solids.

The electric field in materials

When a material is placed in an external electric field $\mathbf{E}_o$, the wave functions of the atoms are changed. The net effect is that the regions with probability of finding electrons are shifted in the $-\mathbf{E}_o$ direction while the regions with probability of finding the positively charged nuclei are shifted in the direction of $+\mathbf{E}_o$ (Fig. 5). The shifts may not exactly align parallel to $\mathbf{E}_o$, and may not all be uniform except in what we call simple materials. A negative surface charge develops on the material near the source of $\mathbf{E}_o$, and a positive surface charge develops on the opposite side. We say the material has an induced charge, or that it is electrically polarized.

The induced charges produce a field $\mathbf{E}_d$ in the opposite direction to $\mathbf{E}_o$ in the material. In a very good conductor, there are enough free charges so that $\mathbf{E}_d$ equals $\mathbf{E}_o$, and the average field inside is zero. That is why metal is an effective shielding material, at least for static fields. Outside the conductor the $\mathbf{E}_o$ field vectors are changed so that they are normal to the surface.

In dielectrics, the large energy gap means the electrons are elastically attached to the lattice and only slight shifts are experienced. $\mathbf{E}_o$ and $\mathbf{E}_d$ don't cancel each other completely. In a simple dielectric, pairs of internal charges, −q and +q, are separated by a distance **R** taken in the direction of $\mathbf{E}_o$, from −q to +q. Those pairs of negative −q and positive +q charges are called electric dipoles. The vector quantity, q**R**, is called the electric dipole moment. If there are n dipoles per unit volume, then a measure of the polarization can be expressed as

$$\mathbf{P} = n(q\mathbf{R})\zeta \ (C \cdot m/m^3 = C/m^2),$$

which is called the dipole moment per unit volume. ζ is a function of the alignment and ranges from 0 to 1. For simple materials ζ = 1. Since n, q, **R**, and ζ depend on the material,

$$\mathbf{P} = \epsilon_o \varkappa \mathbf{E}$$

where ϰ, the electric susceptibility, is a measure of the ease of polarization of the material. $\mathbf{E}_o$ is present to maintain correct units. The so called depolarization field $\mathbf{E}_d$ is equal to $-\gamma\mathbf{P}/\epsilon_o$, where γ is a number between 0 and 1, and is related to the geometry of the material. $\mathbf{E}_d$ is not, in general, very useful.

The surface charge σ_b is an actual accumulation of charges *continued on page 82*

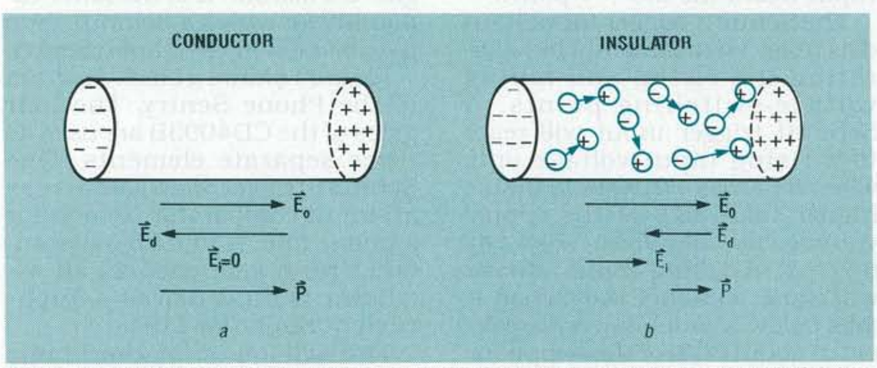


FIG. 5—MATERIALS IN AN EXTERNAL ELECTRIC FIELD E_o exhibit electric polarization. The resulting separation of positive and negative charge regions produce electric dipole moments qR, where q is taken as positive. In a conductor (*a*), enough electrons are free to move to create a depolarization field E_d equal and opposite to E_o. The internal electric field $E_i = E_o - E_d$ is zero. In an insulator or dielectric (*b*), electrons are restricted in movement and E_i is non zero. In both cases, the polarization or dipole moment per unit volume P is related to $-E_d$. The vectors are shown outside the material for clarity.

ELECTROMAGNETIC THEORY

continued from page 59

that are bound directly to the atom and cannot flow. If **N** is of magnitude 1 and is normal to the surface then

$$\sigma_b = \mathbf{P} \cdot \mathbf{N} \text{ (C/m}^2\text{)}.$$

Imagine a Gaussian surface inside the dielectric. With a nonuniform charge distribution some of the bound charges will be displaced across the surface by **P**, leaving a net charge within the surface. In the same manner that we found $\nabla \cdot \mathbf{E} = \rho\epsilon_0$, where ρ is the volume charge density of all the charges contributing to **E**, we can see that the volume charge density in the dielectric is

$$\nabla \cdot \mathbf{P} = -\rho_b \text{ (C/m}^3\text{)}.$$

The negative sign means that the dipole moment per unit volume, **P**, points from negative to positive in the dipoles.

It is customary and convenient to consider a field associated with just the free charge density ρ_f since ρ_b is due to the response of the material. That field must be due to the total charge density less the bound charge density, therefore

$$\rho_f = \rho - \rho_b = \nabla \cdot \epsilon_0 \mathbf{E} + \nabla \cdot \mathbf{P} = \nabla \cdot [\epsilon_0 \mathbf{E} + \mathbf{P}].$$

The term in brackets is called the displacement field vector

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \text{ (C/m}^2\text{)}.$$

In simple dielectrics, **P** and **E** are parallel, and the following relation holds true

$$\mathbf{D} = \epsilon_0(1 + \varkappa)\mathbf{E} = \epsilon\mathbf{E}.$$

ϵ_0 can now be interpreted as the ability of empty space to support an electric field, and is called the permittivity of free space. ϵ is the permittivity of the material. A commonly used quantity is the dielectric constant

$$K = 1 + \varkappa = \epsilon/\epsilon_0.$$

K is greater than 1 for any material, and goes to infinity for a conductor because $\mathbf{E} = 0$ in a conductor. K can be thought of as a measure of the modification of free space by the presence of a material.

From our previous analysis, we have obtained one of Maxwell equations, Gauss' law which reads

$$\nabla \cdot \mathbf{D} = \rho_f.$$

Gauss' law says that the apparent spreading out of the displacement field vector **D** through a Gaussian surface is due to the density of free charges inside. Gauss' law doesn't say, however, that **D** is not producing a swirl. The static **E** contribution can't produce swirling, but the **P** contribution can.

Capacitance

We know that two conductors, separated by a dielectric with dielectric constant k, form a capacitor. If one conductor has charge $+q$ and the other $-q$, the measure of the amount of charge that must be placed on a conductor to change its potential by one volt is called the capacitance, which is in units of coulombs per volt

$$C = q/V \text{ (farads)}.$$

If the free charge q increases, the displacement field vector **D**, which equals the $\epsilon_0 k$ field also increases. That causes a proportionate increase in voltage as **E** rises. Given a particular charge q, the only way to change the capacitance is to change the voltage. That can be done by changing the charge separation distances or by changing the properties of space to give different **E**'s. Simply filling the separation space with a material of greater dielectric constant reduces the **E** field in that space, which reduces the voltage and increases the capacitance.

We can use Gauss' law, without involved calculations, to determine the change in the electric field when any capacitor is filled with a dielectric. In empty space, $\mathbf{P} = 0$ and all the charges are free charges, therefore

$$\nabla \cdot \mathbf{D}/\epsilon_0 = \nabla \cdot \mathbf{E} = \rho_f/\epsilon_0,$$

and

$$\nabla \cdot \mathbf{D}/\epsilon_0 = \nabla \times \mathbf{E} = 0.$$

If the space is filled with a simple dielectric, $\mathbf{D} = \epsilon_0 k\mathbf{E}$, therefore

$$\nabla \cdot \mathbf{D}/\epsilon_0 = \nabla \cdot k\mathbf{E} = \rho_f/\epsilon_0.$$

P is aligned with **E** so there is no apparent rotation and

$$\nabla \times \mathbf{D}/\epsilon_0 = \nabla \times k\mathbf{E} = 0.$$

The divergence and curl of **E** completely characterize the field. By comparison, the **E** for a charged capacitor with empty space as a dielectric is the same as $k\mathbf{E}$ for the same charged capacitor with a dielectric constant k. In a capacitor filled with a dielectric, **E** is reduced by $1/k$. The capacitance $C = q/V$ is increased by k since the voltage potential V is reduced by $1/k$.

In our next edition, we'll look at the effects of electric charges in motion. We'll see that another type of field, the **B** field, is required to describe the magnetic forces associated with them. **R-E**