DOCUMENT RESUME

ED 041 747 SE 009 195

AUTHOR	Phillips, Melba
TITLE	Electricity and Magnetism 1, Electrostatics.
INSTITUTION	Commission on Coll. Physics, College Park, Md.
SPONS AGENCY	National Science Foundation, Washington, D.C.
PUB DATE	66
NOTE	35p.; Monograph written for the Conference on the New Instructional Materials in Physics (University of Washington, Seattle, 1965)
EDRS PRICE	EDRS Price MF-$0.25 HC-$1.85
DESCRIPTORS	*College Science, *Electricity, Energy, Force, *Instructional Materials, Magnets, *Physics, Resource Materials, Undergraduate Study

ABSTRACT

This monograph was written for the Conference on the New Instructional Materials in Physics, held at the University of Washington in summer, 1965. Designed for college students who are non-physics majors, the approach is phenomenological and macroscopic. There are three sections arranged in order of increasing sophistication. The sections are (1) electric forces and fields, (2) electric energy and potential, and (3) electrical properties of matter. A review of the historical development of the concepts of electric forces and fields introduces the first section. The laws of Coulomb and Gauss, and their applications are discussed. The major concepts of electrostatics are presented in section 2. The monograph concludes with a discussion of the electrical properties of material media. Each section has a number of discussion exercises. The author suggests that demonstrations and laboratory work should accompany the presentation of the monograph material. (LC)

Electricity and Magnetism I

ELECTROSTATICS

MELBA PHILLIPS

University of Chicago

COMMISSION ON COLLEGE PHYSICS
DEPARTMENT OF PHYSICS AND ASTRONOMY
UNIVERSITY OF MARYLAND
4321 HARTWICK ROAD
COLLEGE PARK, MD. 20740

GENERAL PREFACE

This monograph was written for the Conference on the New Instructional Materials in Physics, held at the University of Washington in the summer of 1965. The general purpose of the conference was to create effective ways of presenting physics to college students who are not preparing to become professional physicists. Such an audience might include prospective secondary school physics teachers, prospective practitioners of other sciences, and those who wish to learn physics as one component of a liberal education.

At the Conference some 40 physicists and 12 filmmakers and designers worked for periods ranging from four to nine weeks. The central task, certainly the one in which most physicists participated, was the writing of monographs.

Although there was no consensus on a single approach, many writers felt that their presentations ought to put more than the customary emphasis on physical insight and synthesis. Moreover, the treatment was to be "multi-level" --- that is, each monograph would consist of several sections arranged in increasing order of sophistication. Such papers, it was hoped, could be readily introduced into existing courses or provide the basis for new kinds of courses.

Monographs were written in four content areas: Forces and Fields, Quantum Mechanics, Thermal and Statistical Physics, and the Structure and Properties of Matter. Topic selections and general outlines were only loosely coordinated within each area in order to leave authors free to invent new approaches. In point of fact, however, a number of monographs do relate to others in complementary ways, a result of their authors' close, informal interaction.

Because of stringent time limitations, few of the monographs have been completed, and none has been extensively rewritten. Indeed, most writers feel that they are barely more than clean first drafts. Yet, because of the highly experimental nature of the undertaking, it is essential that these manuscripts be made available for careful review by other physicists and for trial use with students. Much effort, therefore, has gone into publishing them in a readable format intended to facilitate serious consideration.

So many people have contributed to the project that complete acknowledgement is not possible. The National Science Foundation supported the Conference. The staff of the Commission on College Physics, led by E. Leonard Jossem, and that of the University of Washington physics department, led by Ronald Geballe and Ernest M. Henley, carried the heavy burden of organization. Walter C. Michels, Lyman G. Parratt, and George M. Volkoff read and criticized manuscripts at a critical stage in the writing. Judith Bregman, Edward Gerjuoy, Ernest M. Henley, and Lawrence Wilets read manuscripts editorially. Martha Ellis and Margery Lang did the technical editing; Ann Widditsch supervised the initial typing and assembled the final drafts. James Grunbaum designed the format and, assisted in Seattle by Roselyn Pape, directed the art preparation. Richard A. Mould has helped in all phases of readying manuscripts for the printer. Finally, and crucially, Jay F. Wilson, of the D. Van Nostrand Company, served as Managing Editor. For the hard work and steadfast support of all these persons and many others, I am deeply grateful.

Edward D. Lambe
Chairman, Panel on the
New Instructional Materials
Commission on College Physics

ELECTROSTATICS

PREFACE

This fragmentary and preliminary material fits into an outline of "multi-level monographs" covering those aspects of electromagnetism which in our view an undergraduate physics major should come to know best. The approach is phenomenological and macroscopic, designed to take advantage of prior experience; we begin magnetostatics with magnets, for example. The material is planned on two levels to lead through the four fundamental empirical laws of electricity and magnetism to electromagnetic radiation as a climax. The propagation of electromagnetic disturbances with velocity <u>c</u>, reached in the "first course" material without use of the calculus and equivalent to the homogeneous wave equation, was written in an elementary way by Oliver Heaviside (<u>Electromagnetic Theory</u>, London, Benn, 1912, Vol. III, p. 3), but only recently has appeared in the regular pedagogical literature. In our treatment we have tried to stress the physical foundations of Maxwell's great synthesis, stating in words the argument corresponding to each mathematical step. This results in a considerably larger proportion of expository writing relative to mathematics than is customarily found in derivations of the wave equation from Maxwell's equations in their usual form. On the other hand, expression of the laws in differential form seems essential for tracing radiation to its sources in a physically meaningful way; the present Chapter 3 of <u>Magnetostatics</u> could be followed almost immediately by Chapter 5 of Monograph III, which would trace radiation fields to retardation effects. We regret having not sufficient time to write such a chapter, as well as the omission of what should have been Chapter 3 of <u>Magnetostatics</u>, an elementary treatment of magnetic materials.

OUTLINE OF MONOGRAPHS ON ELECTRICITY AND MAGNETISM

	I. ELECTROSTATICS	II. MAGNETOSTATICS	III. CIRCULATION LAWS AND THEIR CONSEQUENCES
FIRST COURSE MATERIAL	1. Electric Forces and Fields 2. Electric Energy and Potential 3. Electrical Properties of Matter	1. Magnets and Magnetism 2. Interaction of Steady Currents *Magnetic Properties of Matter	1. Faraday's Law of Induction 2. Ampere's Law Modified 3. Propagation of Electromagnetic Disturbances
UPPER DIVISION COURSE MATERIAL	*4. Electrostatics Reformulated	3. Magnetostatics Reformulated	*Maxwell's Equations and Plane Waves *Radiation Fields

*No textual material was prepared in the summer of 1965 for these chapters.

We have assumed no knowledge of special relativity, but have emphasized the necessity for choosing a frame of reference in which to define electric and magnetic field quantities, thus laying a foundation for the historical development of relativity theory. Unlike mechanics, vacuum electrodynamics needs no modification because of special relativity except in interpretation, so that an excursion into relativity theory could be made before or after study of the present material.

The experiments leading to the four fundamental laws are described at some length, but in use this written material should be accompanied by demonstrations and laboratory work. The basic experiments should come to be a part of genuine experience for students, but a laboratory monograph should be written as an extension of the present outline. Ohm's law and circuitry, for example, do not play an appreciable role in any other projected booklets. We cannot overemphasize the importance of laboratory work, although we were not able to undertake detailed consideration of its content.

We assume that students will have studied mechanics, that they know Newton's laws, the definition of work, the meaning of the Σ symbol, and have a working knowledge of elementary vector algebra before our material is introduced. (We do define the vector cross product as if for the first time.) In the material designed for upper-class work we assume basic calculus. All vector calculus is developed as needed, but we attempt throughout to stress the physics, not the mathematics, and attempt no mathematical rigor.

The first chapters of Monographs I, II, and III should be studied in that order. The few discussion exercises we include can only indicate a type of problem we consider desirable. Numerical problems, which we have made no effort to provide, are also necessary.

M. Phillips
R. T. Mara

C O N T E N T S

1 ELECTRIC FORCES AND FIELDS

In very early times it was observed that certain substances, most strikingly amber after being rubbed with wool, attract chaff and other light objects. What we would call electrostatic effects were particularly troublesome in the spinning of thread; spindles were sometimes made of amber, and attracted chaff and dust. According to Pliny, Syrian women called amber itself harpaga, "the clutcher," and used the same word for spindle. This was probably the first consistent and repeatable observation of an electrical effect.

But Greek science did not include any study of such odd and usually chance effects, and the science of electricity began with William Gilbert, physician to Queen Elizabeth of England. In his book, De Magnete, published in 1600, Gilbert carefully distinguished between the behavior of amber and that of a magnet, and showed that the behavior of amber was shared by a great number of substances. It was Gilbert who gave the name electric (from the Greek word for amber, electron) to the property itself, in order to describe the attraction for light objects shown by glass if rubbed with silk, of wax or resin if rubbed with wool or fur, and so forth for a long list of materials. The title of his book is nevertheless justified: The advances Gilbert was able to make in the knowledge of electricity seem trivial compared with his achievements in disentangling the essential facts of magnetism.

During the great scientific revolution of the seventeenth century, surprisingly little more was learned about electricity. The period was characterized by much writing of a theoretical nature without sufficient recognition of the facts of electrical phenomena, even on the part of such intellectual giants as Descartes and Robert Boyle. The theorizing continued into the eighteenth century, but along with it came the further development of devices for enhancing electrical effects including various "electrical machines" such as the sphere of Fig. 1.1 which can be turned by a handle. The immediate result was that electrical phenomena became an exciting parlor entertainment, but the same devices that shocked (literally!) and delighted ladies and gentlemen in social gatherings also facilitated scientific observations. During the eighteenth century, the essential facts of static electricity became clear.

What are the elementary facts? Electrified objects, such as the amber and glass of Gilbert's observations, are said to possess an electric charge, which is of two kinds: Unlike charges attract each other, whereas like charges repel. In naming the two varieties of charge positive (+) and negative (−), Benjamin Franklin was building into the language an important principle: Unlike charges may cancel each other, but the total amount of charge, with due regard for sign, is never changed. If we begin with a piece of matter such as amber or glass which is electrically neutral (exhibits no

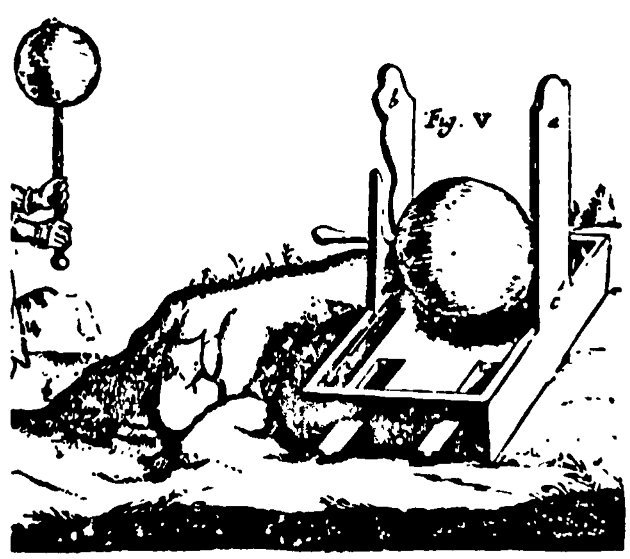

Fig. 1.1

electrical effects such as attracting fluff or dry chaff), we may electrify it by rubbing with cloth, but the cloth then acquires an equal amount of charge of the opposite sign. (In a problem you are asked to devise a test of this statement.) Charge is neither created nor destroyed, although positive and negative charge can neutralize each other, and the two kinds of charge in a neutral body can often be separated from each other. This is the principle of charge conservation, to which no exceptions have ever been found.

Charge is a property of matter which can be described by giving its magnitude and sign, and so is a scalar quantity, like mass, (except that we observe only one kind of mass). Historically, positive charge was defined as that variety which remains on a glass rod if it is rubbed with silk, while negative charge is acquired by amber or sealing wax if rubbed with wool or fur. Franklin found the choice of sign difficult to make, but a choice was necessary to convey the principle of charge conservation. The particular sign convention for charge is actually not important; it is important that some sign convention be established and consistently maintained.

Charged bodies exert forces on each other without actual contact, and

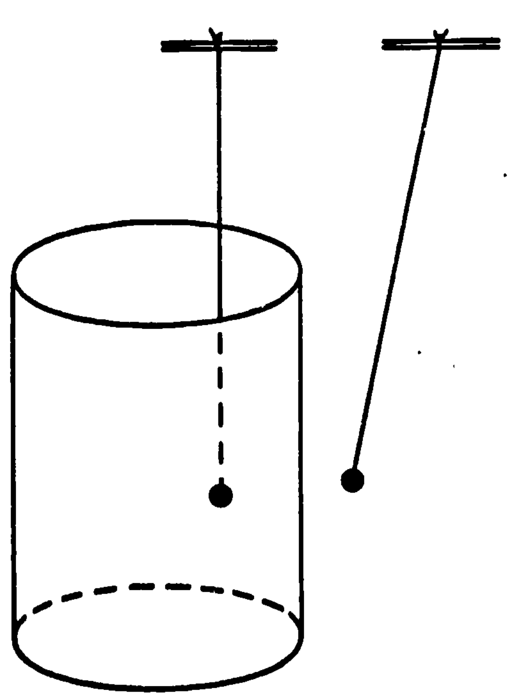

Fig. 1.2

early in the eighteenth century it was recognized qualitatively that the force between two charged objects decreases as the distance between them is increased. But before quantitative aspects of charge could be investigated, it was necessary to distinguish between conductors and insulators. Insulators (once called "electrics," now often called "dielectrics") are materials such as amber or glass which can be held in the hand and electrified by rubbing. Conductors are typically metals, in which charge is free to move and can be conveyed from one part of the material to another. An amber rod is electrified only where it is rubbed, but if a metal sphere on an insulating rod is electrified by stroking it with an electrified amber rod, the charge is distributed over the sphere.

An observation of Franklin's led Joseph Priestley (famous as the discoverer of oxygen) to the first statement of the quantitative relation of electric force to distance. Franklin observed that no electrical effects were to be found inside a charged conductor - no forces on a charged pith ball inside a metal can, as indicated in Fig. 1.2, for example - except very near the rim. Priestley repeated the experiment with a metal sphere that had a small opening for inserting a test charge such as a charged pith ball suspended on a thread. From the absence of any effect on the charge inside, he concluded that the force between charges varies inversely as the square of the distance between them. In reaching this conclusion Priestley reasoned by analogy: It was well known that a uniform spherical shell of matter exerts no net gravitational force on a body inside the shell. This result is a geometrical consequence of the inverse square law in three-dimensional space, and the details are left to a problem. The analogy between electric and gravitational forces should be exact, since on symmetry grounds the charge should be distributed uniformly over the surface of a conducting sphere. Priestley's conclusion, published in 1767, is entirely valid, but it received very little attention at the time.

In 1785, Coulomb measured directly the force between two small charged spheres by means of the torsion balance he had invented (Fig. 1.3), and stated the result in terms of quantity of charge as well as distance. For reasons of symmetry two conducting spheres of the same size should share charge equally if they are touched to each other, and an uncharged sphere should take exactly half the charge of an identical sphere when the two are brought into contact. The forces between charges of known relative magnitude may then be compared at fixed distance of separation, and Coulomb found that both repulsion and attraction are directly proportional to the product of the two charges involved. The effect of distance is then investigated with two charges of constant magnitude, and the force is found to be inversely proportional to the square of the distance between them, both for repulsion and for attraction

If q_1 and q_2 represent the magnitude and sign of two fixed and well-localized charges, the magnitude of the force between them is expressed mathematically as

$$F = kq_1q_2/r^2.$$

Here r is the distance from one charge to the other, both taken as points, and the constant k depends on the choice of units for charge, distance, and force. But force is a vector quantity, and the force exerted by q_2 on q_1 is toward or away from q_2 depending on whether the two charges are opposite in sign or alike. If $\vec{\hat{r}}_{21}$ is a unit vector directed from q_2 to q_1 as in Fig. 1.4, the force on q_1 at distance r from q_2 is

$$\vec{F} = kq_1q_2\vec{\hat{r}}_{21}/r^2. \qquad (1.1)$$

We shall measure q in coulombs, distance in meters, and force in newtons. We shall return to the definition of the coulomb, but may note now that with these units k is found to be 9×10^9, to a very good approximation. It is seen from the equation that the dimensions of k are newton-meter^2/coulomb^2, but it will not often be necessary to write out these dimensions if we are consistent in the use of units. The coulomb is clearly a large charge, since two coulombs at a distance of one meter would exert on each other a force of nearly 10 billion newtons. The stationary charges with which electrostatic experiments are made are small fractions of a coulomb.

If there are more than two charges present, the total force on one of them is found to be the vector sum of the forces exerted by all the others taken

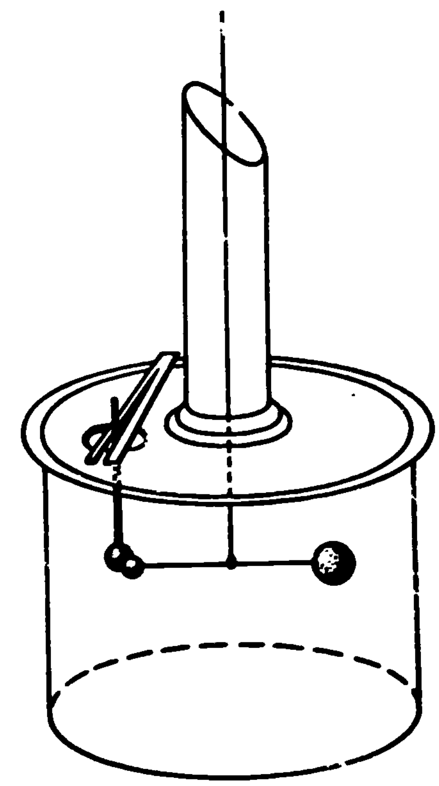

Fig. 1.3

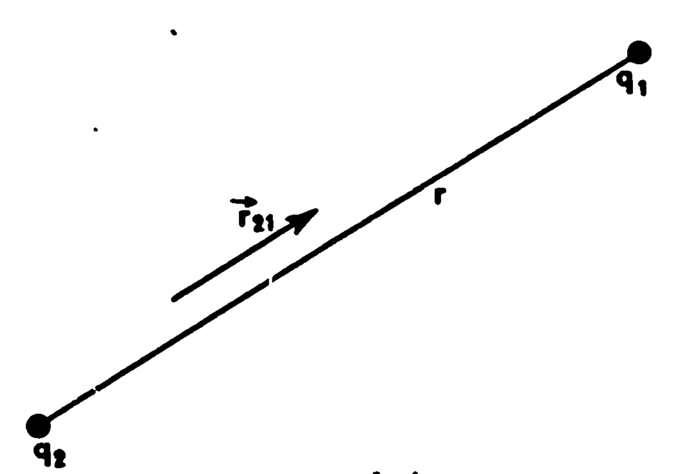

Fig. 1.4

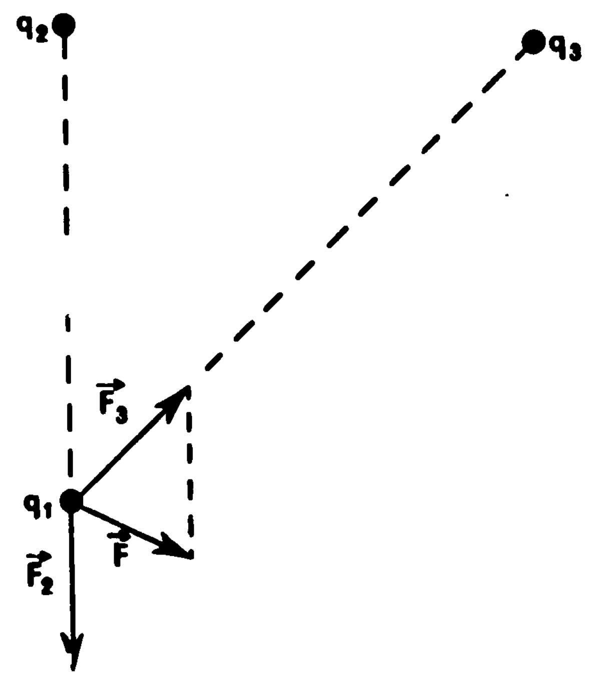

Fig. 1.5

separately, as in Fig. 1.5. This is equivalent to the statement that the force between two charges is not altered by the presence of a third, and is called the principle of superposition. The principle of superposition enables us to find the force on a charge produced by an known fixed distribution of charge, whether well localized or not. The force exerted on q_1 is the vector sum of the effects of all elements Δq of the distribution, each Δq having its own distance r from the position of q_1, and its own unit vector pointed toward that position:

$$\vec{F} = q_1 \sum_{\text{all } \Delta q} \frac{k\Delta q \vec{\hat{r}}}{r^2} \qquad (1.2)$$

This force is proportional to q_1, and depends on the position of q_1; for each different position the unit vectors and the distances r from the fixed charges Δq would be different. Yet the size of q_1 may be considered separately from its position; if $\frac{1}{2}q_1$ were substituted for q_1 at any particular place, the force on it would be reduced by one half, but the direction of the force would be the same as before.

The region around a charge or a configuration of charges where electrical effects could be detected has been called a field, in much the same way that we might speak of the field of influence of a person, or the field (territory) of a traveling salesman. But the electric field may be made quantitative: Coulomb's law enables us to describe the field intensity, as experienced by any small charged body q_1. The field intensity at any point in space is the force per unit positive charge placed at that point. It is a vector quantity, designated by $\vec{E}$, which depends only on the fixed distribution of charges producing the field and the position of the point. The field intensity surrounding an isolated point charge q_2 is

$$\vec{E} = kq_2\vec{\hat{r}}/r^2, \qquad (1.3)$$

where $\vec{\hat{r}}$ is a unit vector radially outward from the position of q_2 to the point at which $\vec{E}$ is considered (Fig. 1.6). If q_2 is a negative charge, the field intensity is directed toward q_2, in the direction of $-\vec{\hat{r}}$.

The field intensity at a particular point P produced by the charges Δq_i of a fixed distribution is

$$E = \sum \frac{k\Delta q_i \vec{\hat{r}}_i}{r_i^2} \qquad (1.4)$$

where every unit vector $\vec{\hat{r}}_i$ points from

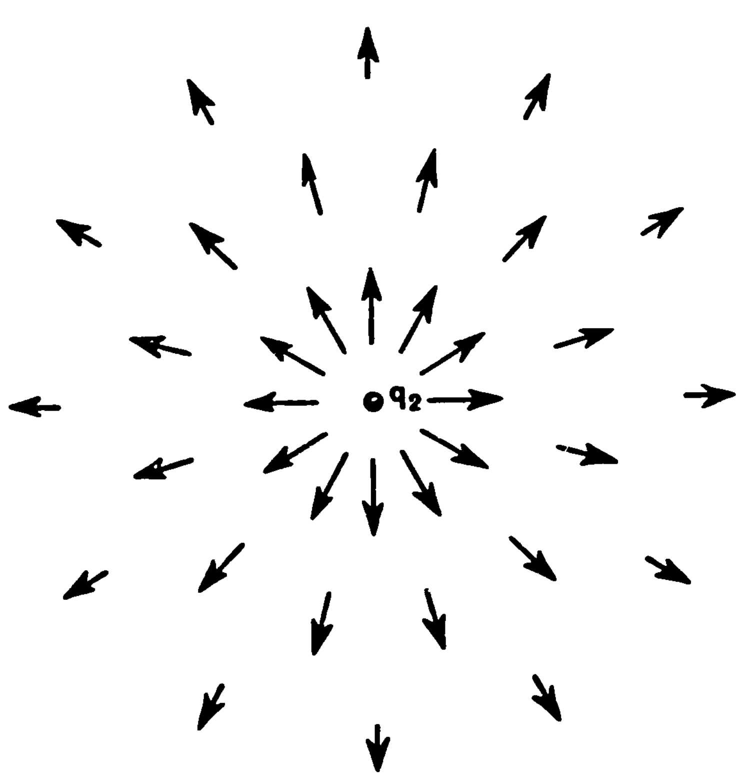

Fig. 1.6

From BERKELEY PHYSICS COURSE, E. M. Purcell, by permission Educational Services, Inc.

the corresponding element of charge of charge Δq_i toward P, and r_i is the distance from Δq_i to P. Positive elements of charge Δq produce contributions to the total field intensity which are directed away from Δq, whereas contributions from negative source charges are directed toward the sources.

Electric fields are often represented by drawing field lines. A field line is a line drawn so that its tangent is in the direction of the field intensity at each point. The field lines of an isolated point charge are simply straight line radii originating at the position of the point charge (Fig. 1.7). The pattern of field lines corresponding to two or more point charges is more interesting (Fig. 1.8). The number of lines drawn does not matter for our purpose, but the lines begin at positive charges and end at negative charges, and are therefore closer together near the charges where the field intensity is stronger. In two dimensions we can show only a cross section of the field, which actually extends through three-dimensional space.

There is no particular advantage in introducing the idea of field intensity for applying Coulomb's law to the interaction of two or even several well-localized charges. All we have done is to divide the problem into two problems, the production of a field by a set of charges considered as sources of a field intensity $\vec{E}$ at each point, and the force $\vec{F} = q\vec{E}$ experienced by a charge q at some particular point. We shall see almost immediately that there is indeed an advantage in the quantitative definition of $\vec{E}$ for determining the force exerted on a point charge by surface and volume distributions of charge, that hard problems often become much easier if they are broken up in this way. But the concept of field intensity becomes almost mandatory when we come to consider changing fields, produced by charges which are not stationary, as we shall see in Monograph III of this series.

Let us again consider the electrostatic field intensity whose source is a point charge. The sign and magnitude of the charge is represented by q, and

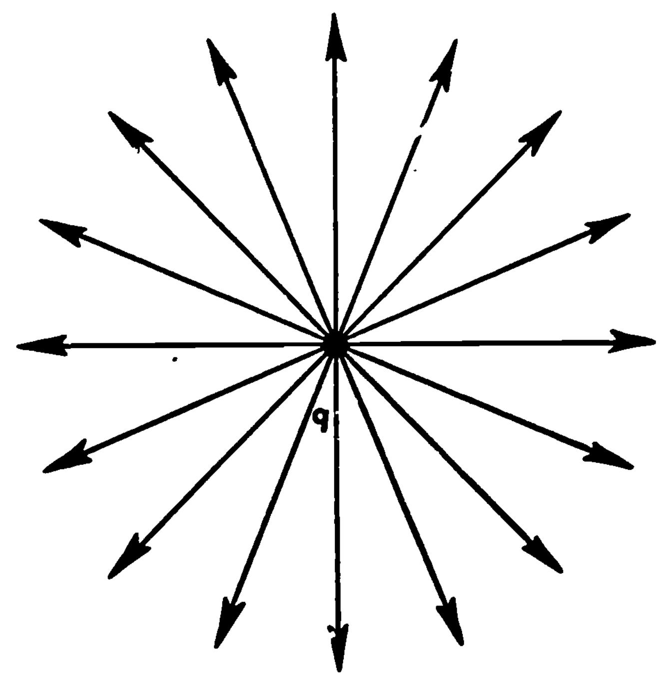

Fig. 1.7

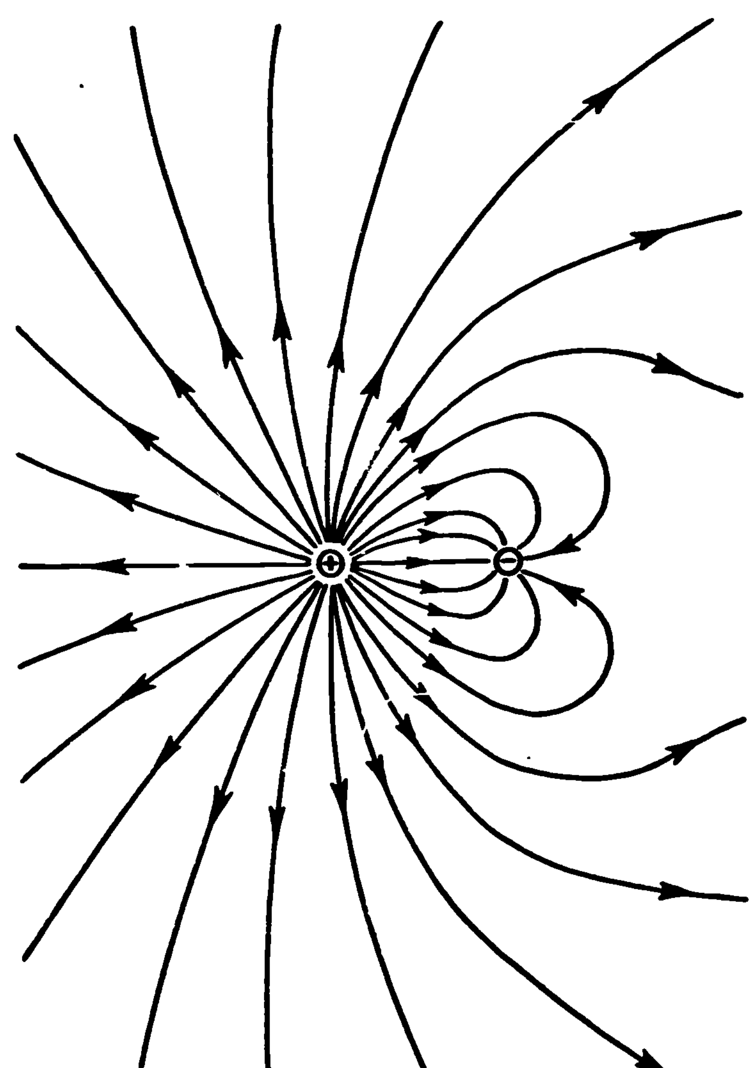
Fig. 1.8

From BERKLEY PHYSICS COURSE, E. M. Purcell, by permission Educational Services, Inc.

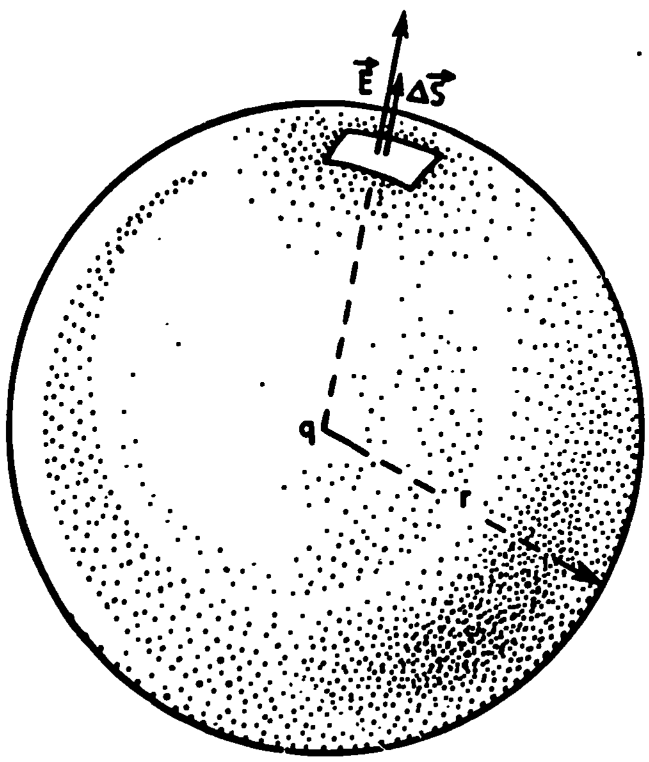

Fig. 1.9

it follows from Coulomb's law that the force per unit positive charge at distance r and in the direction $\hat{r}$ from q is

$$\vec{E} = kq\hat{r}/r^2. \qquad (1.3)$$

A remarkable relation between $\vec{E}$ and its source q can be stated in terms of what is called the flux of $\vec{E}$. The meaning of flux can be demonstrated by considering a sphere whose center is at the position of q. Let the area of the sphere be S, so that a portion of the surface area is ΔS, taken small enough that it is approximately plane. To every such area there corresponds a particular direction, normal to the surface, and thus ΔS may be represented as a vector $\Delta\vec{S}$, whose direction we take positive outward from the sphere. The small areas need not be equal in magnitude, but we can specify each area by an index i; the field intensity at the position of ΔS_i is $\vec{E}_i$. We shall call the scalar product $\vec{E}_i \cdot \Delta\vec{S}_i$ the flux of $\vec{E}$ through this particular increment of area, and we can find the total flux of $\vec{E}$ through the surface of the sphere by summing $\vec{E}_i \cdot \Delta\vec{S}_i$ over the entire surface. If the total flux is called Φ_E, then

$$\Phi_E = \sum \vec{E}_i \cdot \Delta\vec{S}_i. \qquad (1.5)$$

But for our sphere the field intensity $\vec{E}_i$ and $\Delta\vec{S}_i$ are parallel at all points on the surface (Fig. 1.9), so that

$$\sum_{s\ \mathrm{closed}} \vec{E}_i \cdot \Delta\vec{S}_i = E \times [\text{total area}]$$

$$= \frac{kq}{r^2} \cdot 4\pi r^2 = 4\pi kq,$$

since the magnitude of $\vec{E}$ at every point on the sphere of radius r is kq/r^2. The total flux through the surface of a sphere is thus independent of the size of the sphere.

But we can carry the idea further to prove that the total flux of $\vec{E}$ originating from q is the same for any closed surface surrounding q. Since the flux through a portion of spherical surface centered at q is independent of the radius, clearly we could make a complicated surface consisting of spherical segments connected by segments of cones without changing the result. Let us now consider a more general surface as shown in Fig. 1.10. The flux through any ΔS is still $\vec{E} \cdot \Delta\vec{S} = E\Delta S \cos\theta$, just the magnitude of $\vec{E}$ times the component of $\Delta\vec{S}$ parallel to $\vec{E}$, or the projection of $\Delta\vec{S}$ on a sphere whose center is q. Thus the shape of

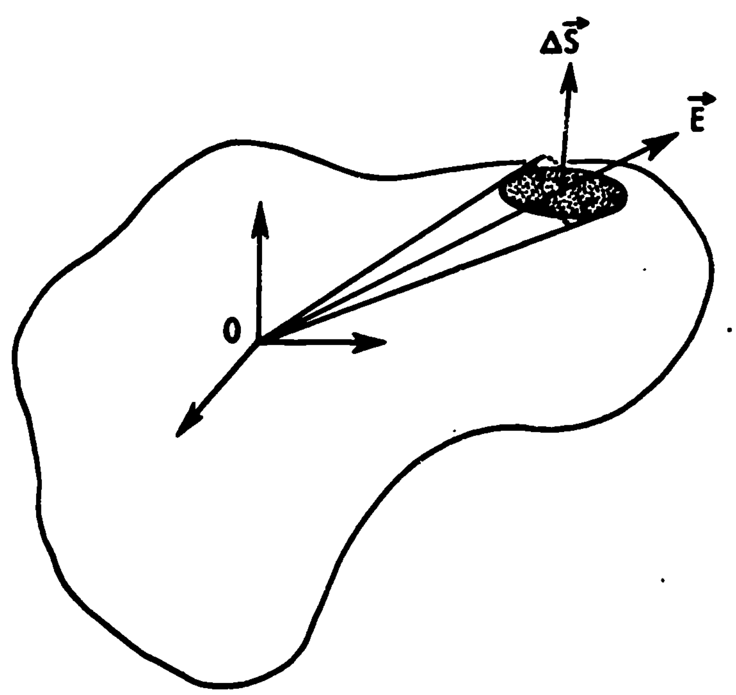

Fig. 1.10
After George Owen, INTRODUCTION TO ELECTROMAGNETIC THEORY. Allyn Bacon, Inc. 1963.

the surface does not matter. What does matter is whether the surface surrounds the charge q. If it does, the outward flux of $\vec{E}$ is $4\pi kq$. If the surface surrounds a volume of space outside the charge, the outward flux is equal to the flux into the volume, so that the net flux is zero.

That the total flux from a closed surface is just $4\pi kq$ does not depend on the localization of the charge to a particular point or small region. In fact, the superposition principle tells us at once that there may be many point charges, or a charge distribution "smeared out" in space, and that the total flux of $\vec{E}$ through a closed surface is $4\pi kQ$, where Q is the total (net) charge enclosed by the surface. Thus

$$\sum_{S\ \text{closed}} \vec{E}_i \cdot \Delta\vec{S}_i = 4\pi kQ, \qquad (1.6)$$

where $Q = \Sigma\,\Delta q$ within the volume enclosed. This remarkable result is known as Gauss's law; we note that the physical content of Gauss's law is the same as Coulomb's law. The original form of the law is essentially a statement of the field intensity in terms of the sources; Gauss's law enables us to locate sources if the field is known. Any closed surface through which there is not net flux contains no net charge, whatever the surface size. But we shall see that Gauss's law also enables us to find very easily the field produced by charge configurations for which the direct sum of the vector field increments is difficult to evaluate.

Before applying Gauss's law to the solution of problems we should note the existence of the geometrical factor 4π. This factor arises from the inverse square law in three-dimensional space: It is simply the area of a sphere divided by r^2. The appearance of 4π is then the consequence of living in a three-dimensional world for any quantity which decreases inversely as the square of the distance from a point source. In this sense the inverse square law is geometrical: The surface density of any quantity which flows from a well-localized (point) source falls off in all directions inversely as the square of the distance from the source, if it is transmitted through space without loss. In the problems you will see that this is true for a stream of particles, and for light.

The geometrical factor 4π in Gauss's theorem will be carried into other equations relating sources and field intensity unless the constant k is defined to suppress its appearance. To save writing 4π again and again in these relations the point charge form of Coulomb's law in mks units is written $k = 1/4\pi\epsilon_0$, so that

$$F = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r^2}, \qquad (1.7)$$

where $\epsilon_0 = 1/4\pi \times 9 \times 10^9 = 8.85 \times 10^{-12}$ coulomb^2/newton-meter^2. We note that the factor 4π is written explicitly in one relation to avert its appearance in others. Its appearance somewhere is unavoidable. So far as we are concerned at this stage the coulomb as a unit of charge is a standard arbitrary unit, as is the length of a meter stick. Once all the units have been decided upon, the constant k must be evaluated by experiment, and we have said that in mks units it is very nearly 9×10^9. In principle the size of the coulomb could be defined by fixing a value for k in advance, but in practice the unit of charge can be much more accurately defined through the interaction of electric currents, as we shall see in Monograph II.

In mks units Gauss's law is simply

$$\Phi_E = \sum_{S\ \text{closed}} \vec{E} \cdot \Delta\vec{S} = Q/\epsilon_0, \qquad (1.8)$$

with Q the total charge enclosed by the surface, as before. Let us apply this theorem to a spherical distribution of charge, for example a charged spherical conductor. Since there is nothing to distinguish one direction of space from another, we can conclude from

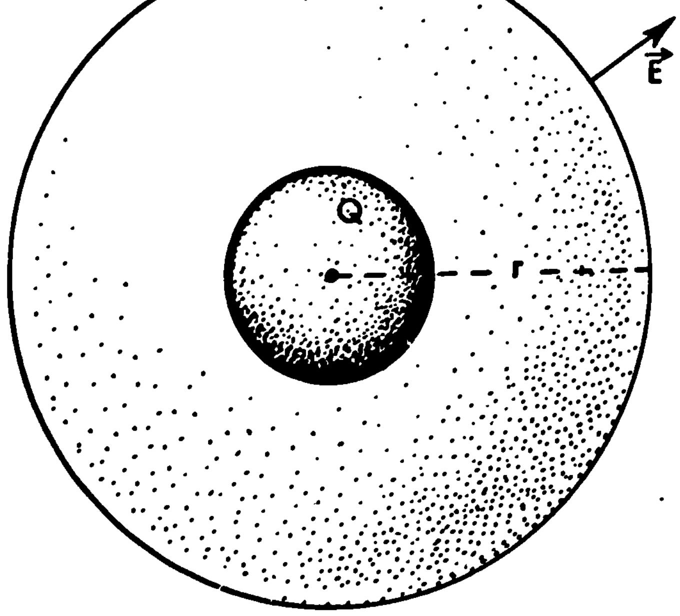

Fig. 1.11

symmetry considerations that the field intensity at the surface of the mathematical sphere (concentric with the sphere of charge) is directed radially outward, but we do not initially know its magnitude. The total flux of $\vec{E}$ through the surface of a sphere of radius r is therefore simply $4\pi r^2 E$, and by Gauss's law this flux is Q/ϵ_0, where Q is the total charge on the sphere. But if

$$4\pi r^2 E = Q/\epsilon_0,$$

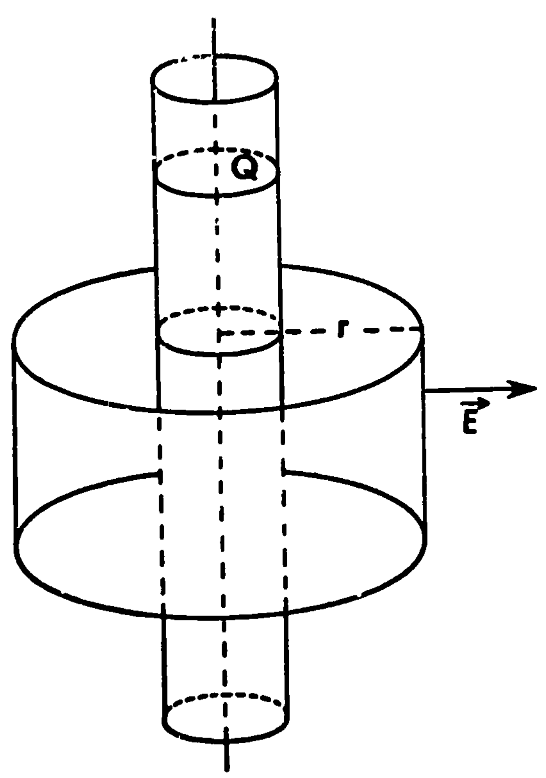

Fig. 1.12

and $\vec{E}$ is directed radially, then

$$E = \frac{1}{4\pi\epsilon_0}\frac{Q\vec{r}}{r^2}, \tag{1.9}$$

where $\vec{r}$ is a unit vector directed radially out from the center of the sphere. This is exactly the same as if Q were a point charge located at the center of the sphere. The same result would be obtained if Q were distributed uniformly, or in any spherically symmetric way, throughout the volume of the sphere. Thus the field of a spherically symmetric distribution of charge is indistinguishable, outside the region of charge, from the field of a point charge (Fig. 1.11). This result is by no means obvious from Eq. (1.4) with the sum of Δq extending over all regions of the sphere.

The application of Gauss's law to a long cylinder of charge to find the field intensity is again much easier than computation of the sum of field increments from the elements of charge in the cylinder. Consider a cylindrical conductor, for example, so long that its end effects are negligible, and let the Gaussian surface also be a cylinder, its axis coincident with that of the conductor (Fig. 1.12). Again we invoke arguments of symmetry: The field intensity $\vec{E}$ is radially symmetric in a plane perpendicular to the axis, and, since the ends of the cylindrical charge are relegated to infinity, there is no axial component of $\vec{E}$. There is then no flux through top and bottom or our Gaussian surface, and the magnitude of $\vec{E}$ is the same at all points on the lateral surface. For a Gaussian cylinder of radius r and length L, the total flux is then E times the lateral area, which is $2\pi rL$:

$$\Phi_E = 2\pi rLE = Q/\epsilon_0,$$

from which

$$E = \frac{1}{2\pi\epsilon_0}\frac{(Q/L)\vec{r}}{r}, \tag{1.10}$$

where $\vec{r}$ is a unit vector from the center of the cylinder at right angles to the axis, and Q/L is the charge per unit length. Again, as with a sphere, we see that the details of the charge distribution do not matter, so long as the distribution is cylindrically symmetric. Any cylinder of charge produces the same field outside the cylinder as would be produced by an axial line of charge with the same charge per unit length, Q/L. The formula for the field intensity accompanying a line of charge is one we have not previously encountered; at the expense of more trouble the same formula can be derived by summing Eq. (1.4).

It may be noted that the fields of both point and line charges become infinite at the source in the limit of geometrical points and lines. Fortunately this need not worry us: Macroscopic charges are always spread out over finite volumes of space, and even elementary charged particles such as the electron are thought to have finite extension in space. And the charge density, defined as the charge per unit volume, must be finite (not infinite) if the total charge of any object is finite. Inside a region of finite charge density the field intensity is also finite, and the determination of the field inside a uniform spherical distribution of charge is left to a problem. The question of the field inside a uniform charge distribution with cylindrical symmetry is equally easy to answer, given Gauss's law.

Let us consider one further application of Gauss's law, a most important one, to find the field just outside the surface of a charged plane conductor. To do so we must examine further the nature of a conductor in electrostatics. A conductor was defined as a substance in which charge is free to move, but in saying <u>statics</u> we demand that the charge <u>not</u> move. Almost by definition, then, a conductor having a static charge can have no field at all <u>inside</u> the conductor, and even at the surface there is no field lying <u>in</u> the surface: Any such fields would produce motions of the charge free to move and we would no longer have a static charge. This tells us two things: The net free charge of a conductor is on its surface, with the interior electrically neutral, and the field intensity just outside the conductor is normal to the conducting surface.

The application of Gauss's law to find the relation of the external field to the surface charge follows immediately. Our Gaussian surface is a short fat cylinder partly inside and partly outside the conductor, its flat external surface parallel to the surface of the conductor (Fig. 1.13). Let the cross section area of this cylinder be S. Since the field intensity is normal to S, and all the flux out of the volume passes through this one surface, the total flux is just ES. Therefore,

$$\vec{E} = (Q/S)\vec{n}/\epsilon_0, \tag{1.11}$$

where $\vec{n}$ is a unit vector normal to the surface of the conductor, and Q/S is the charge per unit area of surface. The surface density of charge is such an important quantity in electrostatics that it is often given a special symbol, $\sigma = Q/S$. The magnitude of the field intensity just outside the surface of a conductor is then simply σ/ϵ_0, in terms of the density of surface charge.

If one is close enough to the surface of any conductor, the surface may

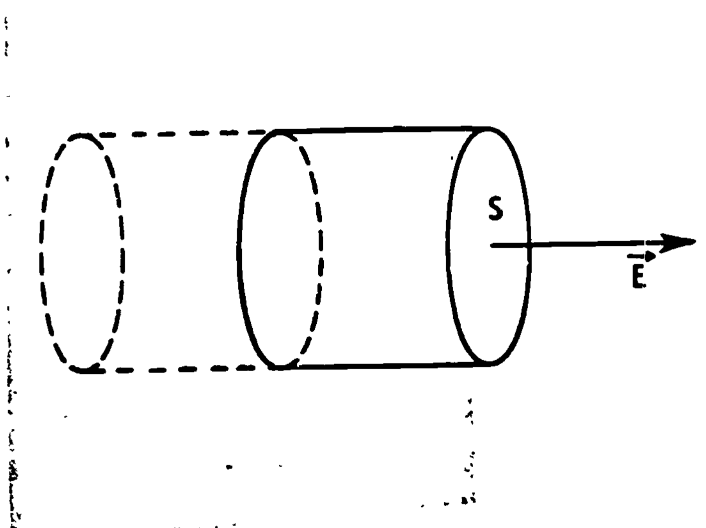

Fig. 1.13

be considered plane, or, to put it in a different way, if one takes a sufficiently small portion of any surface, the portion may be considered plane. The field just outside the surface is given by Eq. (1.11), even for a conducting sphere or cylinder, but then the lines of $\bar{E}$ begin to diverge. For a very large plane, or for a set of two plane conductors with equal and opposite charges, we can obtain a field which is uniform in magnitude and in direction over a considerable volume of space.

PROBLEMS

(The first nine problems were contributed by A. B. Arons.)

1.1 Draw on your own experience to list at least five or six physical situations that you would describe as being associated with "electrification" of various objects. Be sure to identify what was rubbed against what, and indicate what effects provided evidence of the electrification. How can you tell whether or not a particular object is "electrified"? (In addition to cases in which you might have rubbed one material with another, recall circumstances in which you yourself were involved - scuffling over a carpeted floor, handling dacron or nylon clothes, etc.)

Similarly, list five or six effects that you have heard described as "magnetic"; i.e., describe the behavior of magnets (include some of the things you can and cannot do with them). How can you tell whether or not a particular iron bar is a magnet?

What evidence leads us to conclude that we are dealing with different physical effects, justifying the introduction of the two names "electric" and "magnetic"? (Cite some of the <u>differences</u> between the two types of phenomena; recall, for example, the unalterable "two endedness" of magnets; the fact that one can hold a magnet in his hand without having it lose its magnetic property, etc.)

1.2 You have undoubtedly heard the word "charge" used many times in connection with electricity. At this point, what meaning do you associate with this term? Can you see "charge" on an object? How can you tell whether or not an object is "charged"? What experiences with electrified bodies lead us to the notion that we might conceive a body as carrying different quantities or "amounts of charge"?

1.3 How do we arrive at the conclusion that "like charges repel"? (What do we mean by "like" charges? How might we set up a situation in which we can assert with confidence that two objects carry like charges? Describe some possible experiments.)

A piece of amber rubbed with wool and a rubber rod rubbed with fur are observed to repel each other. What is the justification for saying that the amber and rubber carry like charges?

What do we mean by "unlike" charges? How do we know when two bodies carry unlike charges?

1.4 Describe a hypothetical experimental observation that would force you to say, "Here is a body which carries a third kind of electric charge." (Visualize the interactions between this body and suspended rods of rubber and glass carrying respectively the two kinds of charge we have already recognized.) Under these circumstances what would happen to statements and descriptions based on use of the two adjectives "like" and "unlike"? Outline the nature of the accumulated experience that leads us to believe that only two kinds of electrical charge occur in the physical world. Has this assertion been proved?

1.5 During the eighteenth century two other names, "vitreous" and "resinous," competed with Franklin's "positive" and "negative" for acceptance in the description of electrical phenomena. Why do you think Franklin's terminology finally won the competition? Is there anything wrong with the other terminology?

1.6 Between 1729 and 1736 two English friends, Stephen Gray and Jean Desaguliers, reported the results of a series of experiments "showing that the Electric Vertue of a glass tube may be conveyed to any other bodies so as to give them the same property of attracting and repelling light bodies as the tube does when excited by rubbing." They showed, for example, that a cork or other object as much as 800 or 900 feet away could be electrified by connecting it to the glass tube with materials such as metal wires or (moist) hempen string. They found that other materials, such as silk, would not convey the effect. As a matter of fact, they discovered in early painstaking experiments that the distant object would not become electrified if the "transmission line" made contact with the earth and they learned to separate it from the earth by suspending it on silken threads.

Experiences of this kind led investigators to discern that electricity seems to move freely on some materials and not on others. Describe in detail several additional experiments you might perform (or can visualize) with various different objects - experiments that support the findings of Gray and Desaguliers. Why do we introduce the names "conductor" or "nonconductor"? To what experiences do these names refer? Are our own bodies conductors or nonconductors? Cite evidence for your conclusion.

1.7 Suppose we are investigating the force between two small conducting charged spheres A and B of identical size (Fig. 1.14). The force is measured by the twist in the suspension fiber when the center of the spheres are 3.00 cm apart. After measuring the force in a given situation when the spheres repel each other and obtaining a value denoted by F_1, we take an identical, but uncharged, sphere C on an insulating handle and bring it in contact with sphere B. Then we remove C.

(a) What are we inclined to say has happened to the quantity of charge carried by sphere B? On what grounds and with what justification?

(b) In the light of the statements made in the preceding paragraph about Coulomb's investigations, what do we expect will happen to the magnitude of the force between A and B at the previous center to center separation of 3.00 cm. How will the new value of force compare with F_1 numerically?

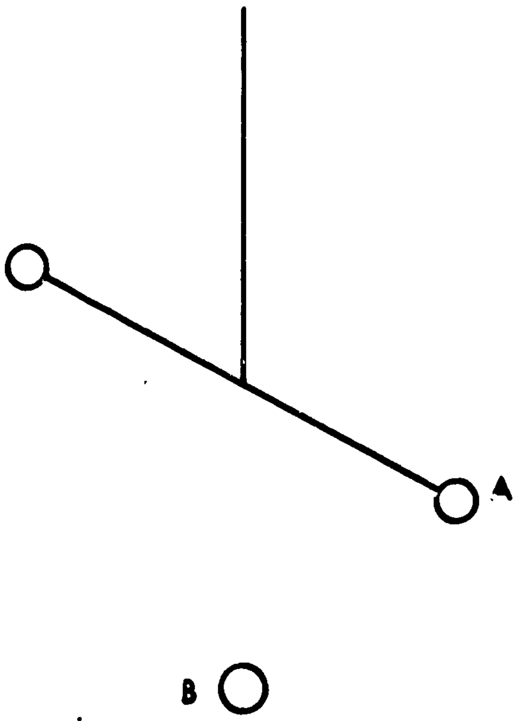

Fig. 1.14

(c) Leaving sphere B as it is after this experiment, we discharge C and bring into contact with A. Then remove C as we did before. We measure the force between A and B at 3.00 cm again. How does it now compare with F_1? Describe the results to be expected after additional steps of this kind.

1.8 Suppose we start with a given charged, conducting sphere, A. Now, as we did in Problem 1.7, we bring this sphere in contact with another, uncharged sphere C, held by an insulating handle. Suppose that C is smaller than A. How do we expect the charge to divide between the two spheres? Suppose C is larger than A? Very much larger than A? (Your responses to these questions are not expected to be numerical. Use words such as "more," "less," "very much less," and explain how you arrive at your inferences.) In the light of this discussion, how would you describe what happens when you touch a charged conducting sphere or establish a conducting path between the sphere and the earth? How would you attempt to describe when transfer of charge from one object to another ceases? What happens when two differently charged conducting spheres are brought in contact? (Note: These questions do not have simple, pat answers; they will eventually have to be reexamined in a broader context. They are being raised at this point to enhance your awareness of some of the problems that lie below the surface of the present discussion.)

1.9 In philosophical discussions of scientific knowledge, it is frequently pointed out that we arrive at the conviction that a particular set of concepts, insights, and descriptions is "correct" not by following one single sequence to a "proof" or an isolated right answer but by finding that the entire network of concepts and experimental observations is internally consistent - that we can criss-cross the network in a variety of different sequences and directions and not develop contradictions. Let's illustrate this notion in connection with our developing conception of electrical phenomena:

Suppose that in a Coulomb torsion balance experiment we charge sphere A on the torsion balance positively. Suppose that B and C are now observed to exert forces of equal <u>magnitude</u> on A at a fixed distance between centers (the forces being, of course, opposite in direction). What would we be led to say about the quantities of unlike charge carried by spheres B and C? If we touch B and C together what would we expect as a final result? It is actually found under such (or analogous) circumstances that the two objects are electrically neutral after contact. In what ways does this reinforce our conceptions of "charge," "quantity of charge," "conservation of charge," neutrality, or unelectrified objects, etc.? The electroscope shown in the diagram (Fig. 1.15) consists of two leaves of flexible gold foil suspended on a conducting rod which passes through the insulating stopper of the protective glass flask to a metal cup at the top. It is observed that if an electrified object, such as amber that has been rubbed with wool, is put into the cup without touching, the two gold leaves diverge as indicated, instead of hanging down.

(a) Account for this behavior.

(b) What will happen if the amber rod is removed, without having touched the metal cup?

(c) Could you use this apparatus to test whether the charge acquired by silk used in electrifying a glass rod is equal and opposite to that acquired by the glass rod? How?

1.11 Show that a uniform spherical shell on matter exerts no net gravitational force on a point mass m inside by considering a double cone of very small aperture, its apex at the point mass, which cuts through the sphere on either side (Fig. 1.16). What is the total gravitational force on the point mass due to the portions of the spherical shell inside the double cone? Complete the argument to include all the mass in the shell.

1.12 (a) Show by applying Gauss's law to the interior of a hollow, conducting, charged sphere that the field inside the sphere is zero.

(b) Assume that you have a uniform volume distribution throughout a sphere. (This is <u>not</u> a conductor!) Show that the field intensity inside the sphere is given by

$$\vec{E} = r\rho\vec{r}/3\epsilon_0,$$

where $\vec{r}$ is a radial unit vector and ρ is the charge density, or charge per unit volume. (Simply apply Gauss's law to a spherical surface of smaller radius than the radius of the charge distribution.)

1.13 (a) Use Gauss's law to show that the field intensity inside a uniform cylindrical shell of charge is zero.

(b) Assume you have a uniform volume distribution of charge throughout a long circular cylinder, and show that $\vec{E}$ inside the cylinder (not near the ends) is

$$\vec{E} = r\rho\vec{r}/2\epsilon_0,$$

where $\vec{r}$ is a unit vector radially directed out from the axis of the cylinder, and ρ is again the charge density.

1.14 If light is emitted constantly and uniformly in all directions by a spherical source, and there

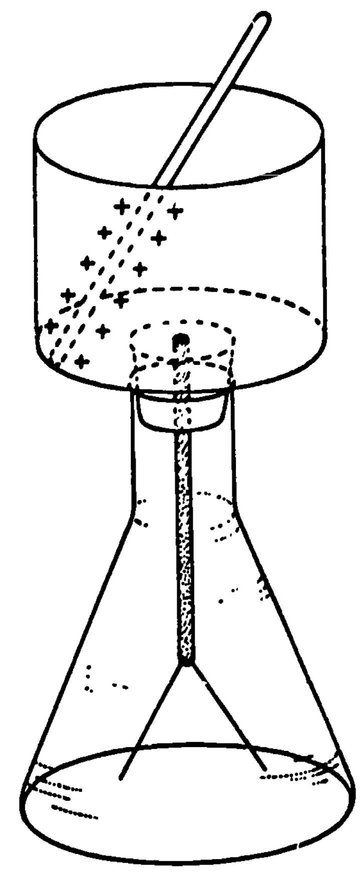

Fig. 1.15

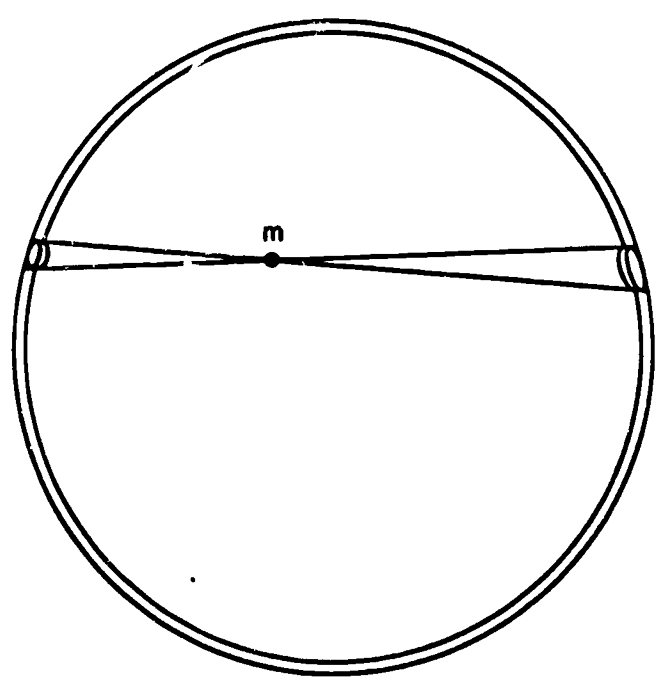

Fig. 1.16

is no intervening material to absorb or reflect it, show that the intensity (amount of light per unit area received by a surface at right angles to its direction) falls off inversely as the square of the distance from the center of the source.

2 ELECTRIC ENERGY AND POTENTIAL

The science of electrostatics began with the study of small and often chance effects such as the attraction of amber for thread and chaff, but in the early days of our planet electrical energy may have played an essential role in the beginning of life itself. According to one theory, incessant flashes of lightning through an atmosphere of nitrogen, carbon dioxide, and water vapor produced the first organic molecules, from which organized life developed. The frequency and intensity of lightning have diminished, but a lightning flash today obviously releases an enormous amount of heat and light. The connection between this impressive natural phenomenon and the behavior of amber was established in the eighteenth century by Benjamin Franklin. The weather is a complicated matter indeed, which scientists are only beginning to understand, but the recognition that lightning is electrical certainly extends electrical phenomena beyond the small theater of parlor entertainment.

In this chapter we are concerned with one aspect of electric energy: work done by, or against, electrostatic forces. That these forces are really very strong is usually masked by the charge neutrality of most objects. Equal amounts of positive and negative charge, however great the amounts separately, can neutralize the effects of each other if they are nearly coincident; this is equivalent to the statement that lines of electric field intensity end, as well as begin, and that ordinarily there is just about as much negative charge as positive charge in any particular region. But if there is the possibility of strong forces there is also the possibility of utilizing these forces to do work. In addition to any possible utility, we shall see that in many instances it is simpler to describe electrical phenomena in terms of work and energy than in terms of forces. For one thing, energy is a scalar quantity while force is a vector, requiring three numbers to specify it instead of only one.

Let us consider the energy associated with the interaction of two point charges. Work is performed against the force of repulsion in bringing two like charges closer to each other, that is, in decreasing their separation r. How much work? We cannot get the answer by simple multiplication of the force given by Coulomb's law and the relative displacement, since the force itself depends on the distance between them and is not the same at the beginning of the displacement as at the end. Rigorous derivation of the formula for computing the energy expended to displace the charge q_1 from r' to r_0 as indicated in Fig. 2.1 requires use of the calculus. Here we shall assume the correct answer and see that it is a reasonable consequence of Coulomb's law. First let us note that we do this work very slowly; no kinetic energy is given to the charged body, and the force we exert is equal to (not greater than) the electrostatic repulsion of the two charges.

The correct answer for the work we must do in moving q_1 from r' to r_0, keeping q_2 fixed, is

$$W = \frac{kq_1q_2}{r_0} - \frac{kq_1q_2}{r'} = \left(\frac{kq_1q_2}{r_0r'}\right)(r' - r_0). \tag{2.1}$$

Here, to save writing in dealing with point charges, we are letting k stand for the constant in Coulomb's law, as

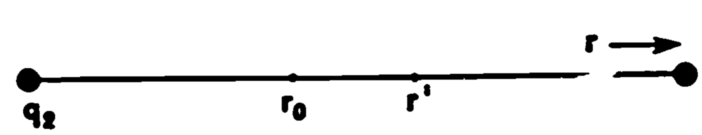

Fig. 2.1

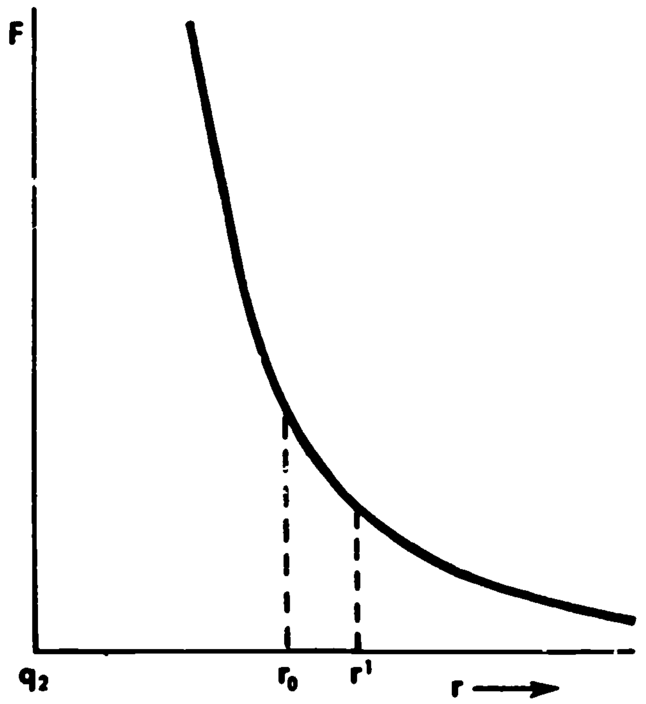

Fig. 2.2

From PSSC PHYSICS, by permission D.C. Heath and Company, 1960

at the beginning of Chapter 1. Equation (2.1) is clearly correct if the displacement $(r' - r_0)$ is very small, so that $1/r_0r'$ is very nearly $1/r_0^2$ or $1/r'^2$. It can be justified for larger displacements by plotting the Coulomb law force against variable r, and finding the area under the curve between any two particular values of r (Fig. 2.2), just as one finds the work done by (or against) a mechanical force which is not constant over the displacement. Actually it is valid for any displacement, large or small, of q_1 along the line on which q_1 and q_2 lie.

If we begin with q_1 so far away from q_2 that the force on it is negligible, the outside work required to bring it up to distance r_0 from q_2 is simply

$$W = kq_1q_2/r_0 \qquad (2.2)$$

Once the work has been done we may say that the pair of charges themselves possess this energy. We have held q_2 fixed while bringing up q_1, so that we could say, alternatively, that q_1 now possesses energy kq_1q_2/r_0 owing to its position with respect to q_2, energy which it did not have when it was far away.

We know from mechanics that energy a body possesses by virtue of its position is called <u>potential energy</u>. If we say that kq_1q_2/r_0 is the potential energy of q_1 at distance r_0 from q_2, we are calling the potential energy of q_1 zero for r = infinity. This is an arbitrary but convenient floor from which to measure electric potential energy of point charges.

The work represented by Eq. (2.1) depends only on the <u>radial</u> displacement. That is because the force between q_1 and q_2 is along the line joining them; the charge q_1 could be moved anywhere on a sphere of radius r_0 about q_2 as a center without costing any work whatever against electric forces (Fig. 2.3). Moreover, in moving the charge q_1 from P' to P_0 in Fig. 2.4, the same amount of work is done for all the paths shown, only the radial portions of the path require any "pushing" against the repulsion of the charges, and in doubling back to larger distances from q_2 one gains energy from the repulsion. The work put into carrying charge q_1 from P' to P_0 is $kq_1q_2(1/r_0 - 1/r')$, regardless of the path. Forces for which the work done in going from one point to another is independent of the path are called <u>conservative</u> forces, and electrostatic forces are conservative.

Let us rewrite the external work required to move q_1 from r' to r_0 in terms of the electric field intensity associated with the charge q_2 as given by Eq. (1.3) of Chapter 1. From the definition of work as the scalar prod-

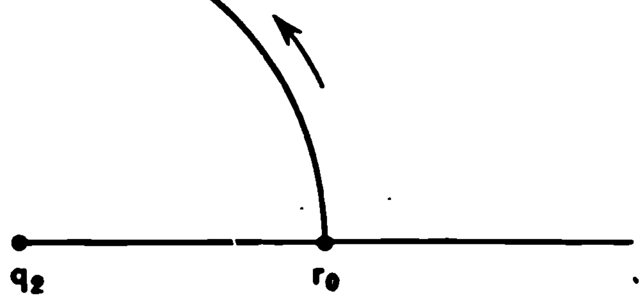

Fig. 2.3

uct of force and displacement, we write the work done against the force $q_1\vec{E}$, for a displacement $\Delta\vec{s}$, so small that $\vec{E}$ is essentially constant over the distance interval, as $q_1\vec{E}\cdot\Delta\vec{s}$. It is exactly the sum of such increments of work that we have considered in obtaining Eq. (2.1).

$$W = \Sigma -q_1\vec{E}\cdot\Delta\vec{s} = q_1(kq_2/r_0 - kq_2/r') = q_1(\phi_0 - \phi'), \quad (2.3)$$

where the sum is over all $\Delta\vec{s}$ on any path from P' to P_0. The last expression for this work makes explicit its dependence on q_1 and the difference of two quantities which depend on the end points of the path in relation to the source of field intensity, q_2. The quantity $(\phi_0 - \phi')$ is called the difference of potential between points P_0 and P'. It is the work per unit positive charge required for the transfer of position from P' to P_0.

The difference of potential between any two points in an electrostatic field may be defined as the external work per unit positive charge required to move the charge from one point to the other, regardless of whether the field is that of a point charge. In going from point P_1 to point P_2,

$$\phi_2 - \phi_1 = -\Sigma\,\vec{E}\cdot\Delta\vec{s}, \quad (2.4)$$

where the sum is to be taken over all line elements of any path beginning at P_1 and ending at P_2. The negative sign is due to the fact that in doing work on the positive charge you are acting against the field if the charge has more potential energy at the end of the path than at the beginning. If $(\phi_2 - \phi_1)$ is positive, we say that P_2, the final point, is at a higher potential than P_1, the initial point.

In order to use the word potential, instead of potential difference, we must establish some conventional floor, just as for potential energy. What floor is established, what position is said to be at zero potential, is a matter of convenience. For a point charge, or a configuration of

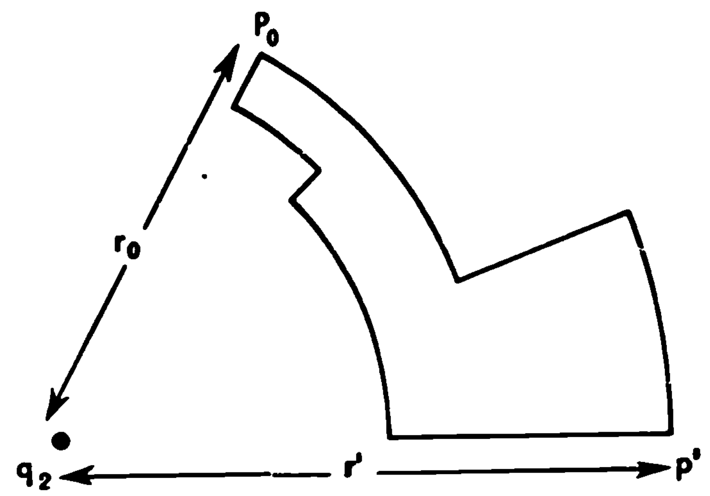

Fig. 2.4

charges localized in a small region of space, it is customary to say that the potential is zero at very large distances from the charge, i.e., at $r = \infty$. Thus the potential at point P which is a distance r from point charge Q is

$$\phi = kQ/r. \quad (2.5)$$

Since, as we have seen, the field intensity outside any distribution of charge having spherical symmetry is indistinguishable from that of a point charge at the center of the sphere, Eq. (2.5) is also the potential at a point r distant from the center of a sphere of charge, so long as the point is outside the sphere.

The law of superposition holds for potential as well as for field intensity, and the superposition is much easier to accomplish for a scalar quantity than for a vector field. The potential at point P in Fig. 2.5 is simply due to the presence of three fixed

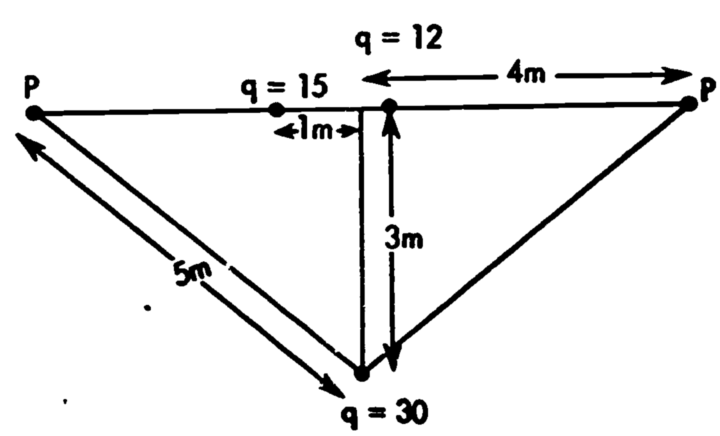

Fig. 2.5

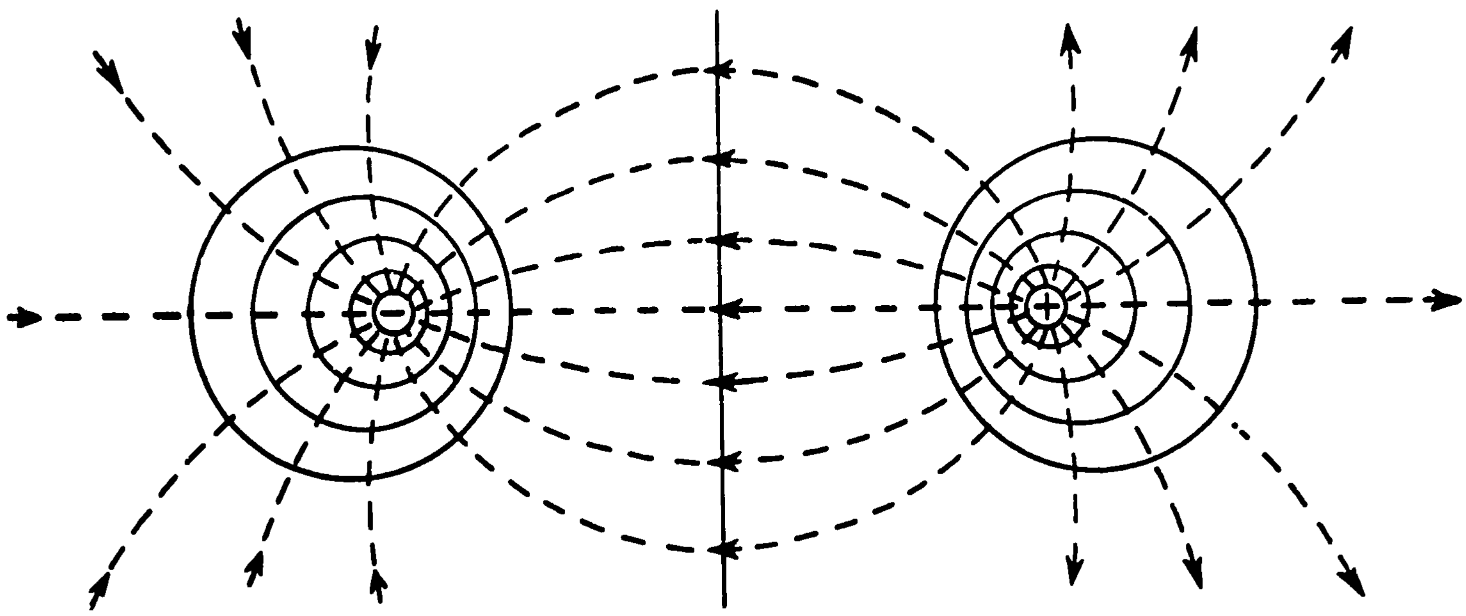

Fig. 2.6

From ELECTRIC AND MAGNETIC FIELDS by D.H. Tombulian, © 1965, by permission Harcourt, Brace & World, Inc.

charges, $k(q_1/r_1 + q_2/r_2 + q_3/r_3)$. It is left to the problems to show that, with the signs and magnitudes of charges indicated, the potential at P is positive, while the potential at P' is zero. This means that work would have to be done by some outside force to bring a positive charge from great distance to P, whereas no network is needed to bring one to P'. Other points in the field are at negative potential. What does that mean?

The applicability of the law of superposition means that a general formula can be written for the potential whose source is a fixed distribution of charge:

$$\phi = \Sigma\, k\Delta q/r \qquad (2.6)$$

is the ordinary numerical sum over all elements Δq of the distribution, each with its own r. This equation can be compared with Eq. (1.2) of the previous chapter; it is much easier to evaluate the potential in many problems, simply because the potential is a scalar.

But of what use is it to know the potential? Presumably we can measure forces on charges with torsion balances or otherwise, and the mapping of electric field lines, as in Chapter 1, tells us a great deal about such forces, but measuring the total work in bringing a test charge from infinity to each point is merely a "thought experiment." Thought experiments are often very useful in understanding physics, but the potential has a more practical value as well. Even in very complicated applications, such as the electrostatic electron microscope, it is often the potential that is mapped, rather than the field itself. The potential is mapped by drawing equipotential lines, or the traces of equipotential surfaces. The device is familiar from contour maps showing the altitude (gravitational potential) of a geographical region, but that is a very simple example, since the gravitational potential varies only with the vertical coordinate of space, while electric potentials vary with three coordinates. The equipotential surfaces corresponding to the two equal and opposite point charges as cut by a plane containing the charges are shown in Fig. 2.6. In this instance the equipotentials are surfaces of revolution about an axis on which the charges lie. The field intensity lines are also sketched; they must meet the equipotentials at right angles, since by definition no work is done against electrical forces in moving from one part of an equipotential surface to another.

The requirement that field lines

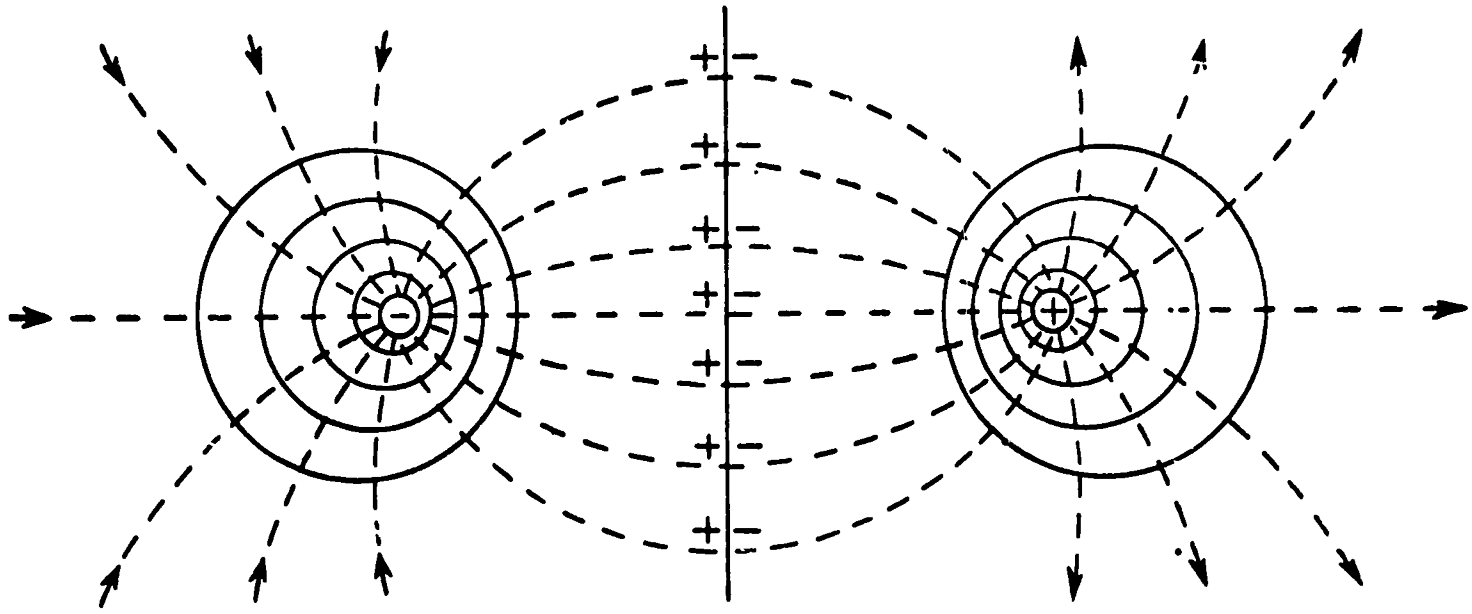

Fig. 2.7

From ELECTRIC AND MAGNETIC FIELDS by D.H. Tombu... © 1965, by permission Harcourt, Brace & World, Inc.

be orthogonal to equipotentials enables us to map the field, knowing the equipotentials. For some charge configurations the potential may be computed (or determined in some other way), equipotentials mapped, and the field lines are ascertained. When conductors are involved, the potential plays a particularly simple and important role. In Chapter 1 we defined a conductor as an object in which charge is free to move, and noted that almost by definition there is no electric field within the substance of a conductor in electrostatics. It follows equally that there is no difference of potential between different points in or on a conductor if the charge is static. The surface of a conductor is an equipotential surface in electrostatics - in fact the entire volume of a conductor is a single potential. If a thin sheet of copper, for example, were substituted for the equipotential plane of Fig. 2.6, as indicated in Fig. 2.7, the field on both sides of the plane would be completely unchanged. The field lines do not penetrate the conductor, to be sure; free charges appear on the conductor in just such a configuration that the plane remains an equipotential. But the conductor is also an electrostatic shield: One of the two point charges could be removed, and the field on the other side would remain just as before. The charge on the conductor is said to be induced by the presence of the point charge Q.

In general an uncharged conductor placed in an electrostatic field distorts the field so its surface becomes an equipotential (Fig. 2.8). This process involves separation of the surface charge in such a way that the field intensity normal to the surface has a magnitude $E = \sigma/\epsilon_0$, in accord with Eq. (1.11). Such charges are said to be induced on the surface of the conductor. The conductor may have a net charge; if so, the charge is redistributed by the presence of another charge nearby, so that the surface is at a single potential. The conductor may be "grounded," connected to the earth, which is a large and reasonably good conductor itself, a reservoir of charge which may be considered infinite (Fig. 2.9). The potential of earth, or any conductor in electrical contact with it, is said to be zero; is this consistent with the zero potential assumed for an isolated point charge?

In configurations of conductors, difference of potential is often more important than any attempt to ascribe potential itself. A particularly simple and instructive set of conductors consists of a pair of parallel plates of equal area, separated by a distance

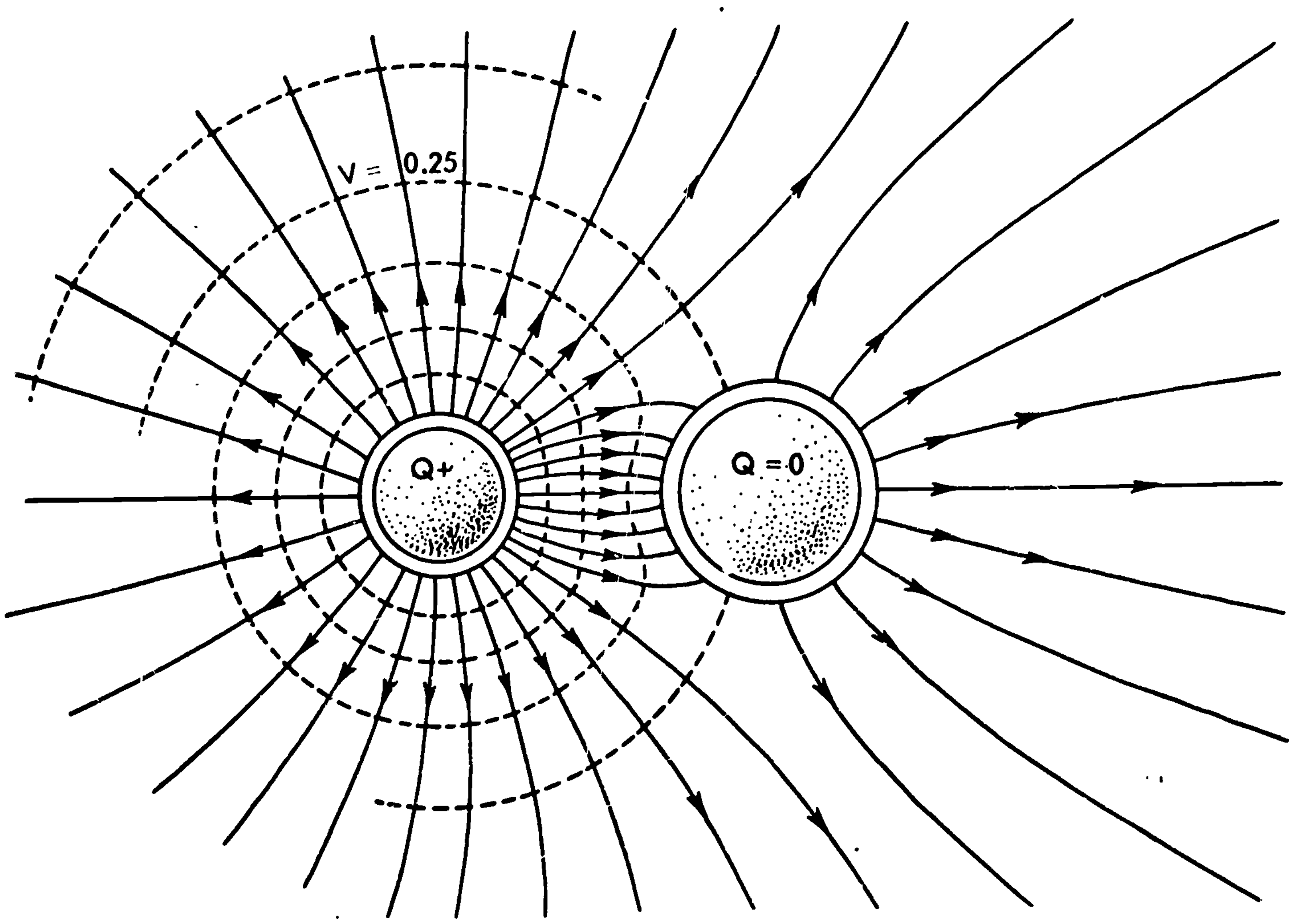

Fig. 2.8

From BERKLEY PHYSICS COURSE, E. M. Purcell, by permission Educational Services, Inc.

d small compared with the dimensions of the plates (Fig. 2.10). Throughout most of the region between the plates the fact that there are edges can be ignored. If the upper plate is given a charge +Q, the field is $\vec{E} = \sigma\vec{n}/\epsilon_0$, where $\sigma = Q/A$, and the equipotential surfaces are planes, as indicated, from the symmetry of the configuration. The equipotential surface which is the lower plate has an induced charge −Q. The field between the plates is uniform and the difference of potential between the plates is $\phi_2 - \phi = \Sigma\vec{E}\cdot\Delta\vec{s}$. In scalar magnitude,

$$\phi_2 - \phi_1 = E\ d = Qd/\epsilon_0 A. \qquad (2.7)$$

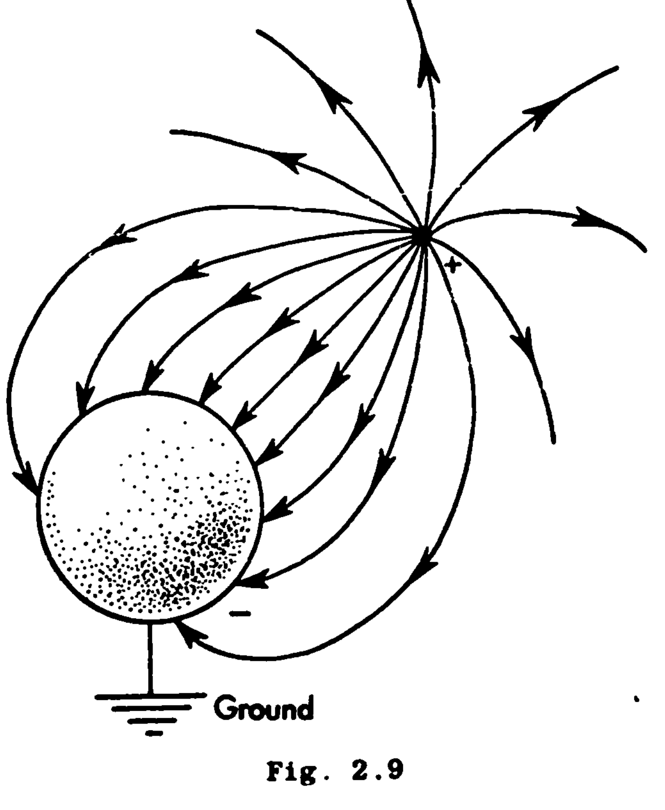

Fig. 2.9

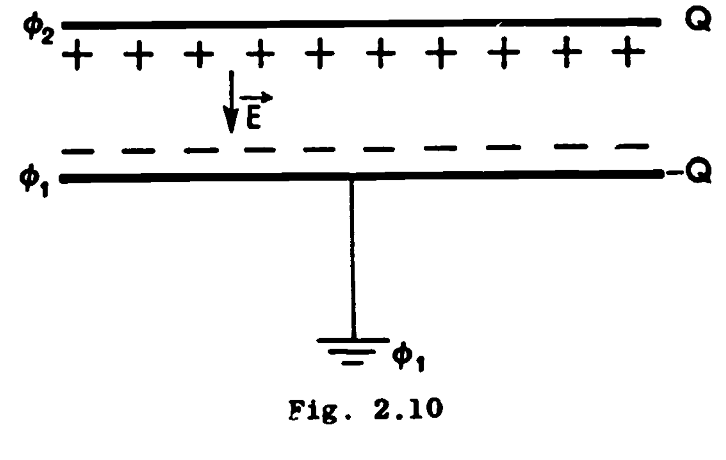

Fig. 2.10

The mks unit of potential difference is the volt, defined as one joule per coulomb: A potential difference of 1 volt means 1 joule of work against electrical forces to move unit charge from one point to the other. From Eq. (2.7) it is easy to see why field intensity is often expressed in volts/meter instead of newtons/coulomb.

Instead of stroking the top plate with a glass rod which has been rubbed with silk and letting a charge be induced on the lower plate, it is more convenient to connect the two plates to some source of difference of potential by means of conducting wires. The most familiar source of potential difference is the chemical cell, or battery. A battery may be defined for our purpose as a device capable of maintaining a constant potential difference between two electrical "terminals," and of supplying equal and opposite amounts of charge when necessary to accomplish this result. The chemical battery was invented by Alessandro Volta of the University of Padua, Italy, in 1800, and greatly facilitated the growth of electrical science.

A source of potential difference is actually a source of energy, a device for transforming energy of some other kind into electric energy. The amount of energy per unit charge it is capable of delivering is called its emf - the letters stand for electromotive force, but the name is an anachronism, since emf is not a force. The potential difference between the terminals of a battery, or between the plates of Fig. 2.11, is equal to the emf of the battery, but an emf cannot be produced electrostatically.

To find the energy required to charge the plates of Fig. 2.10 or Fig. 2.11, we note from Eq. (2.7) that the amount of charge on either plate is directly proportional to the potential difference between the plates. This is usually written

$$Q = C(\phi_2 - \phi_1) = CV, \qquad (2.8)$$

where C is a constant independent of

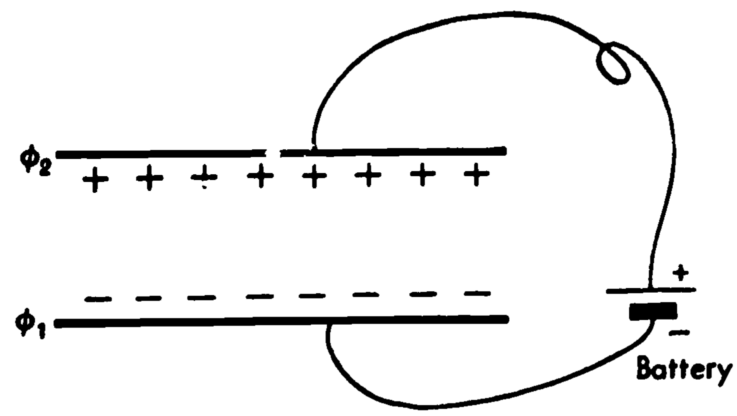

Fig. 2.11

the charge or potential, and $V = (\phi_2 - \phi_1)$ to save writing. Let us begin with uncharged plates, and transfer positive charge Δq from the lower plate to the other, leaving $-\Delta q$ behind. The increment of work Δw required to transport any later Δq is proportional to the potential difference, which builds up as the charge on the plates increases, and is equal to q/C for any net plate charge q. If the potential difference is plotted against q as in Fig. 2.12, and the product $(q/C)\Delta q$ is summed for all Δq from $q = 0$ to $q = Q$, we obtain the total work done in charging the plates; it is equal to the area under the straight line, which may be written in several ways,

$$W = \tfrac{1}{2}QV = \tfrac{1}{2}Q^2/C = \tfrac{1}{2}Q^2 d/A\epsilon_0 = \tfrac{1}{2}QEd$$
$$= \tfrac{1}{2}\epsilon_0 E^2\, Ad \qquad (2.9)$$

by virtue of Eqs. (2.7) and (2.8). This energy is stored in the configuration of charged conductors, but an interesting alternative interpretation is suggested by the last form given in Eq. (2.9). The product of the area A and

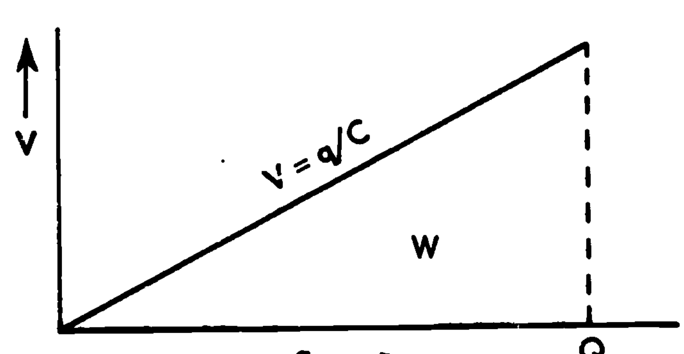

Fig. 2.12

the plate separation d is just the volume of space between the plates where the field intensity is E. Outside this space the field vanishes, if we neglect edge effects. We can, if we like, attribute the energy to the field itself, an energy equivalent to $\frac{1}{2}E^2/\epsilon_0$ joules/cubic meter. The question of whether the energy really resides in the charged conductors or in the space between them has a mechanical analogy: In a loaded air rifle does the bullet have potential energy, or is the potential energy in the compressed air?

Since the plates of Fig. 2.11 are oppositely charged, they attract each other, and mechanical forces are required to hold them in place. From Eq. (2.9) we may determine the amount of force required. Let us consider the work necessary to increase the separation by a small distance Δd, without changing the charge on either plate.

$$W = F\Delta d = \tfrac{1}{2}Q\Delta V = \tfrac{1}{2}QE\Delta d,$$

so that the external mechanical force needed may be identified as the coefficient of the small displacement Δd,

$$F = \tfrac{1}{2}QE = \tfrac{1}{2}\epsilon_0 E^2 A = \tfrac{1}{2}Q^2/A\epsilon_0 \qquad (2.10)$$

which, like the amount of stored energy, may be expressed in a variety of ways.

A configuration of two conductors near each other but not in electrical contact is called a capacitor or a condenser. The proximity of charges of the opposite sign makes it possible to accumulate relatively large charges. The first man-made capacitors were called Leyden jars, and consisted of glass jars with a conducting substance inside and a conductor outside. A jar of water was held in the hand, according to the first record of the device, and a conducting wire connected the water to an "electrical machine" capable of supplying charge. The stored energy became so great that when the other hand was brought near the wire there was a discharge through the air in the form of a spark. To Franklin, who repeated such experiments, the spark was reminiscent of lightning. It was this idea that he tested in his famous experiments with kits in thunderstorms, and so established the connection between lightning and electricity.

Exactly how the clouds become charged is a very difficult and complicated problem, but they do, and they are enormous conductors, near a still larger one, the earth. The discharging spark, which releases the gradually accumulated electric energy, constitutes lightning. Differences of potential between cloud and ground may be as high as one billion (10^9) volts, and a discharge of 20 coulombs is not unusual, so that the energy dissipated in a single stroke of lightning may be 10^{10} joules, equivalent to nearly 3000 kilowatt hours. The discharge itself is of course not an electrostatic phenomenon, but its existence is evidence that electric forces are indeed strong.

3 ELECTRICAL PROPERTIES OF MATERIAL MEDIA

In our consideration of the effects of electric charge we have begun as if there were no intervening matter between charges. The electric field intensity $\vec{E}$ (force per unit charge) is defined and traced to its sources as if the charges existed in vacuum. On the atomic level, this view can be maintained and justified, since we hold that matter consists of "atoms and the void." But in practice we experience gross matter, many of whose properties cannot be traced in any simple way to the behavior of the atoms and subatomic constituents. In fact, many properties of matter have not to date been satisfactorily traced to the behavior of its constituent particles, even if all known complications are taken into account. These problems are the subject of intensive on-going research.

We shall not here restrict ourselves entirely to the gross aspects of phenomena, for a number of them can be understood quite simply, if only qualitatively, in terms of atoms and molecules. Nevertheless we shall begin with the consideration of matter in bulk.

Electrical effects were first observed and studied with what were once called "electrics" and now called insulators. For quantitative observations on static charges, however, it is necessary to use conductors, typically metals, on the surface of which net charge is distributed, and which will maintain their charge if well insulated. Experiments such as Coulomb's can be performed in dry air or in an evacuated vessel with no very appreciable change in the results, and except as convenient supports for charged metal objects, the role of insulators in the science of electricity was minor. (Insulators are, of course, also important for support of current-carrying conductors, but again for their lack of electrical properties, that is, as nonconductors of electric charge.)

But insulators do have electrical effects, even when they possess no net charge. A systematic investigation of these effects was made by Faraday. His experiments were varied, and often complicated, but the essential results may be inferred from consideration of very simple apparatus. Let us consider a pair of fixed parallel metal plates, large in comparison with the separation between them, which can be charged by connecting across a battery, as indicated in Fig. 3.1. The battery imposes a known difference of potential between the plates, which can be measured with a good voltmeter. Now disconnect the battery by opening the switch, leaving the plates charged and well insulated. The space between the plates is vacant (dry air is very nearly equivalent) and the presence of the charge produces a field intensity $E = Q/A\epsilon_0 = \sigma/\epsilon_0$ in that region. Here Q is the total charge per plate of area A, and $\sigma = Q/A$ is the surface density of charge. We are assuming the plates so large that edge effects can be neglected. The difference of potential between the plates is just Ed, the work per unit charge in going from one plate to the other. This last result follows from the definition of potential difference, as force per unit charge times distance, both along the same direction.

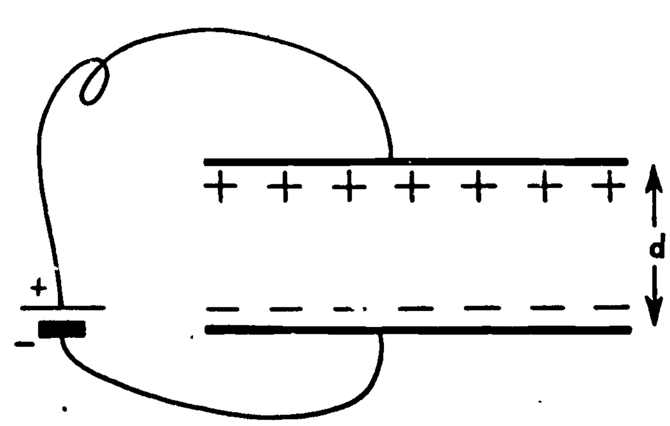

Fig. 3.1

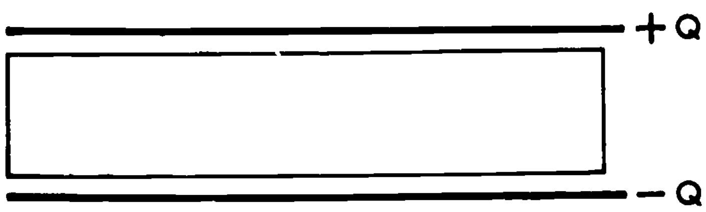

Fig. 3.2

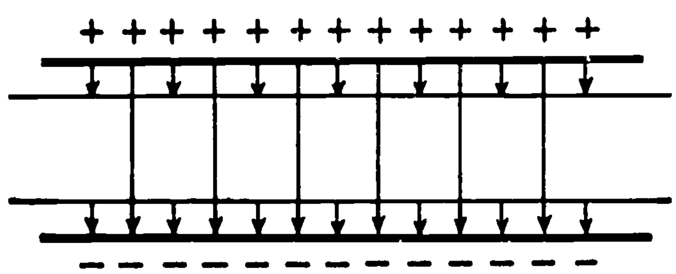
Fig. 3.3

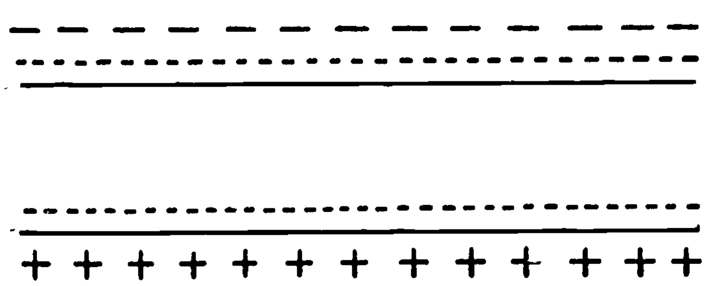
Fig. 3.4

Now insert a slab of uncharged insulating material, such as glass, so as to very nearly fill the space between plates but, as a precaution, without touching them (see Fig. 3.2). Note that the plates are charged just as before. If the difference of potential is now measured again, it is found to have changed. In fact, for glass nearly filling the space the voltage is reduced from its earlier value by more than a factor of two. For all isotropic insulating materials, the effect is to reduce the voltage, and Faraday found that the reduction factor is a constant for any particular material - it does not depend on the particular geometry of the system of conductors.

What must we conclude? The charge has remained unchanged, but if we are to retain the relation between field intensity and potential difference, which follows from the definitions of force and work, we can only conclude that the field intensity E in Ed has changed within the material. Faraday called such materials dielectrics, each with its own dielectric constant; if the voltage is reduced by a factor of 2, the dielectric constant is 2, for example. We should note that most materials, including glass, are found to be entirely unchanged when removed from the apparatus.

This effect of the dielectric can be described in terms of its polarizability. Figure 3.3 shows only lines of $\vec{E}$, and indicates that there are sources (or "sinks") of $\vec{E}$ on the surface of the slab, even though its net charge is zero. This would account for a reduction in the strength of E inside the dielectric, as a factor in V = Ed, the potential difference. What has happened to the dielectric itself? Let us consider the slab alone; the effect of the charged plates remains, but the plates themselves are not shown in Fig. 3.4. The appearance of equal and opposite charge on the two flat surfaces would arise if the normal dielectric consisted of equal and opposite charge densities occupying the same volume and thus cancelling each other, but now one kind of charge is slightly displaced relative to the other. The result is charge neutrality except at the faces perpendicular to the relative displacement, so that each such face has a surface charge. Now a familiar configuration of two equal and opposite point charges is called an electric dipole; its strength, or dipole moment, is the product of the magnitude of one charge and the distance be between them. The dipole moment is actually a vector quantity, whose direction is taken from the negative toward the positive charge as indicated in Fig. 3.5. The slab of dielectric in the previous figure is clearly a dipole, extending throughout a volume instead of being merely a line. If we let the charge density on the surface of the plate be σ_p, so that the total charge is $Q_p = \sigma_p A$, the dipole moment of the slab is $\sigma_p Ad$, where d is the thickness. It is directed from negative σ_p to positive σ_p. But Ad is the volume of the slab, and we may define a dipole moment per unit volume $\vec{P}$, whose magnitude at the surface perpendicular to $\vec{P}$ is σ_p.

We note that the polarization extends throughout the volume, but since it is uniform the only accumulation of charge is at the surface. Even there the charge is not accessible; it is firmly attached to the body of the material, and results only from the slight displacement of positive from negative charge in neutral matter.

The dipole moment per unit volume of a dielectric can be equally well attributed to the atoms or molecules of which it is composed. Suppose that every atom is permanently neutral, but that the positive charge is displaced slightly from the negative charge. Thus every atom is a small dipole, whose dipole moment is $\vec{p}$. If there are N such dipoles per unit volume, all in the same direction $\vec{P} = N\vec{p}$ is an equivalent description per unit volume. We do not explain anything by putting the volume polarization in this form, but we do note that a uniform continuous distribution of charge is not necessary for its definition.

In the case we are considering, with a dielectric such as glass, the polarization exists only in the presence of an electric field $\vec{E}$, and is in fact proportional to $\vec{E}$ over a wide range, but some materials have intrinsic polarizations, and the proportionality of $\vec{P}$ and $\vec{E}$ is not a fundamental fact of electricity. Our description applies both to intrinsic polarization and that produced by the presence of an electric field. Let us note the general relation between the volume polarization of matter and the polarization charge which appears on its surface. The wedge of dielectric in Fig. 3.6 has a cross section which is a right triangle, and is polarized in the direction of one leg. This polarization can be thought of as produced by displacing vertically the whole volume of positive charge from the same volume of negative charge. Charge thus appears on only two surfaces, the base and the slant face, equal and opposite in total amounts. At the base, to which $\vec{P}$ is normal, we can find that σ_p is equal to the magnitude of $\vec{P}$, by the

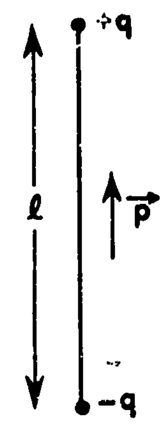

Fig. 3.5

same arguments developed in connection with the parallel slab. But the area of the slant face is greater than that of the base: $A_{slant} \cos\theta = A_{base}$, where θ is the angle between the base and hypotenuse of the cross section. It follows that the surface density of polarization charge on the slant face is smaller than that on the base, i.e., is equal to P cos θ, since the total charge on the two faces is the same in magnitude. The quantity P cos θ is simply the normal component of $\vec{P}$; in general

$$\sigma_p = P_\perp \qquad (3.1)$$

at the plane face of any polarized dielectric.

Since the surface charge on a dielectric is truly inaccessible to direct measurement, as would also be variations in $\vec{P}$ within the dielectric, it is advantageous to write the description of electrical phenomena directly in terms of $\vec{P}$ and accessible

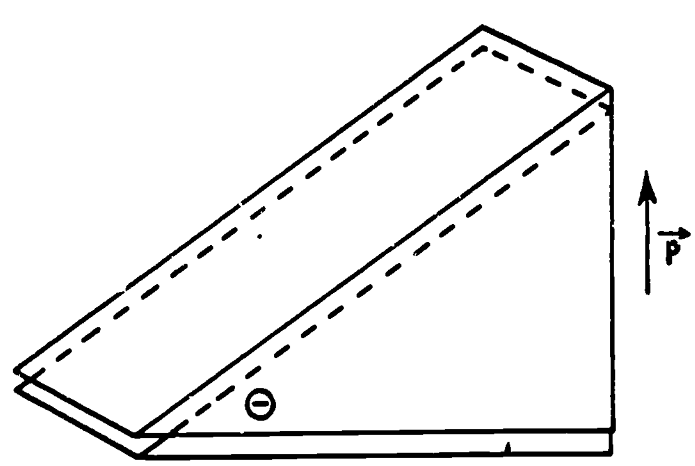

Fig. 3.6

(separable) charge. Let us write

$$\vec{D} = \epsilon_0\vec{E} + \vec{P}, \tag{3.2}$$

where $\vec{P}$ and $\vec{E}$ have now been defined. The new field quantity, $\vec{D}$, has as sources in terms of Gauss's law only accessible charge:

$$\phi = \sum_{\text{closed}} D_\perp \Delta S = Q_{\text{accessible}}, \tag{3.3}$$

the net free charge contained within the volume. Lines of $\vec{D}$ thus begin and end only on separated charges; unlike those of $\vec{E}$, they are continuous through the surface of a polarized dielectric. For historical reasons which go back to the idea of an ether pervading all space and having mechanical properties of its own, $\vec{D}$, is called the displacement field. It is defined here in terms of two quantities which have more immediate physical significance, the electric field intensity, $\vec{E}$, and the electric dipole moment per unit volume, $\vec{P}$, of dielectric material. In empty space $\vec{D} = \epsilon_0\vec{E}$, and within a dielectric we could make a complete description of electrical effects in terms of $\vec{E}$ and $\vec{P}$ alone. But we shall see at once that the combination of $\vec{E}$ and $\vec{P}$ given by Eq. (3.2) is very convenient for expressing electric energy.

The electric energy stored in a pair of charged parallel conducting plates at distance d from each other, given by Eq. (2.9) when there is empty space between the plates, may also be computed for plates separated by a dielectric. The basic equation, $W = \frac{1}{2}QV$, where Q is the magnitude of the charge on each plate and V is the difference of potential between the plates, is still valid, but $\vec{E}$ is no longer so simply related to the charge on the plates. Application of Gauss's law to the surface of a conductor yields

$$\vec{D} = \sigma_{\text{accessible}}\,\vec{n} = (Q/A)\vec{n}, \tag{3.4}$$

where $\vec{n}$ is a unit vector normal to the surface, and $\sigma = Q/A$ for our uniformly charged plates. Equation (3.4) holds whether there is empty space or some dielectric substance outside the surface of the conductors, but $E \neq \sigma/\epsilon_0$ within a dielectric. Since $V = Ed$, however,

$$W = \tfrac{1}{2}QV = \tfrac{1}{2}DEAd, \tag{3.5}$$

for the energy stored in the parallel plate configuration of charges. We note that this expression is equally valid whether the space between the plates is empty or filled with a homogeneous dielectric substance.

The problem of forces between charged bodies embedded in a dielectric must be approached with considerable care. If we knew the volume polarization everywhere, as well as the positions and magnitudes of all free charges, we could in principle apply Coulomb's law directly to find the forces due to all charges, including those which appear on or in the dielectric medium as a result of its polarization. Polarization is usually induced, depending for its very existence on the presence of fields produced by accessible charges. For isotropic homogeneous materials we have already noted that the dipole moment per unit volume, $\vec{P}$, is directly proportional to the field intensity $\vec{E}$. For such materials $\vec{D}$ as given by Eq. (3.2) is then also directly proportional to $\vec{E}$, and the factor κ in the relation

$$\vec{D} = \kappa\epsilon_0\vec{E} \tag{3.6}$$

is called the dielectric constant of the material. The dielectric constant is a pure number, found empirically to be greater than 1 for all substances. If this relation is applicable, a simple expression for the forces between charged conductors may be derived.

Since the mechanical forces which account for the rigidity of solids may be very complicated, let us consider the forces between two charged parallel plates, as in Fig. 3.1, when immersed in a fluid dielectric. A detailed calculation of the forces would include changes in fluid pressure produced by polarization of the fluid, but we may find the total force of attraction between the plates from the expression for electric energy, just as we arrived at Eq. (2.10) in Chapter 2. To increase the separation of the plates by a small distance Δd without changing the charges on the plates we must do an amount of work

$$W = F\Delta d = \tfrac{1}{2}DA\,\Delta d = (Q^2/\kappa A\epsilon_0)\,\Delta d,$$

where the last expression follows from Eqs. (3.4) and (3.6). Therefore

$$F = Q^2/\kappa A\epsilon_0, \qquad (3.7)$$

for the force of attraction between two charged plates of area A immersed in a fluid dielectric whose dielectric constant is κ. The result differs from the force between the plates in empty space by the factor κ in the denominator. Since $\kappa > 1$, the force between the plates is reduced from its vacuum value.

Net forces between charged bodies immersed in a fluid dielectric are always found to be reduced by a factor κ in comparison with the force in empty space. This result is consistent with the fact that $\overline{E}$, the electric field intensity within the dielectric, is also reduced in a homogeneous dielectric from the value it would have in empty space, but the change in the force comes as a result of induced polarization charges in or at the boundaries of the dielectric. Coulomb's law itself is not changed by the presence of the dielectric. The direct electrical interaction between two charges remains the same, but other sources of electric field have been created by the polarization of the dielectric.

Electrostatic forces play an extremely important role in nature. They account for the binding of electrons in atoms, and the binding of atoms into molecules, although the details of such "accounting" are very complicated indeed. Other important aspects of electricity that appear when charges are in motion are the subject of Monographs II and III of this series.

PROBLEM

3.1 The lightweight objects attracted by electrified amber or glass are bits of dielectric polarized by the presence of the electric field.

(a) Describe qualitatively the distribution of charge in such a small object while it is being held by a negatively charged piece of amber.

(b) Is there a net force on a polarized dielectric in a uniform electric field?

(c) Some dielectric substances have an intrinsic electric dipole moment per unit volume; a body possessing this property is called an electret. Would a small electret be attracted to a piece of electrified amber or glass? How could you distinguish between electrets and bits of dielectric which have no permanent polarization?

4 ELECTROSTATICS REFORMULATED

Coulomb's law is all there is to electrostatics, but many aspects and consequences of the law appear clearly only when the subject is formulated somewhat more mathematically. For students wishing to pursue the subject further we recommend Volume II, Chapters 1, 2, 3, and 9 of the Berkeley Physics Course, written by E. M. Purcell. Professor Purcell's approach is "microscopic," based on a qualitative description of the electrical properties of atoms and molecules whose space average yields the field quantities $\vec{E}$ and $\vec{D}$. The theorems which emerge so elegantly from use of the calculus apply equally to continuous distributions of charge or to space averages of atomic charges: They depend essentially on the inverse square law and the superposition principle.