The first relation is just Gauss's law that the total flux of **D** out through the surface is equal to the charge contained inside. The second is the magnetic analog, with no net flux of **B** through a closed surface because of the nonexistence of magnetic charges.

Similarly, let C be a closed contour in space, S' an open surface spanning the contour, $d\mathbf{l}$ a line element on the contour, da an element of area on S', and $\mathbf{n}'$ a unit normal at da pointing in the direction given by the right-hand rule from the sense of integration around the contour. Then applying Stokes's theorem to the middle two equations in (I.1a) gives the integral statements

$$\oint_C \mathbf{H} \cdot d\mathbf{l} = \int_{S'} \left[\mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \right] \cdot \mathbf{n}' \, da \tag{I.15}$$

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = -\int_{S'} \frac{\partial \mathbf{B}}{\partial t} \cdot \mathbf{n}' \, da \tag{I.16}$$

Equation (I.15) is the Ampère–Maxwell law of magnetic fields and (I.16) is Faraday's law of electromagnetic induction.

These familiar integral equivalents of the Maxwell equations can be used directly to deduce the relationship of various normal and tangential components of the fields on either side of a surface between different media, perhaps with a surface charge or current density at the interface. An appropriate geometrical arrangement is shown in Fig. I.4. An infinitesimal Gaussian pillbox straddles the boundary surface between two media with different electromagnetic properties. Similarly, the infinitesimal contour C has its long arms on either side of the boundary and is oriented so that the normal to its spanning surface is tangent to the interface. We first apply the integral statements (I.13) and (I.14) to the volume of the pillbox. In the limit of a very shallow pillbox, the side surface does

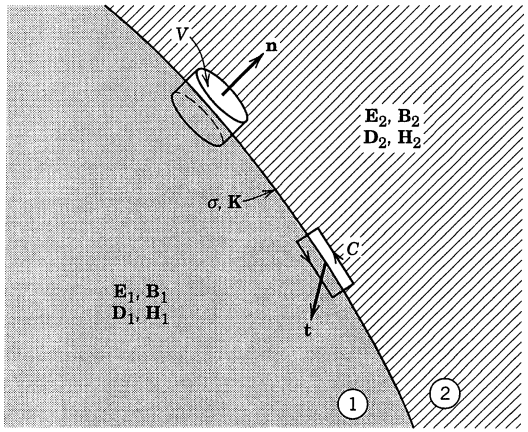

Figure I.4 Schematic diagram of boundary surface (heavy line) between different media. The boundary region is assumed to carry idealized surface charge and current densities σ and **K**. The volume V is a small pillbox, half in one medium and half in the other, with the normal **n** to its top pointing from medium 1 into medium 2. The rectangular contour C is partly in one medium and partly in the other and is oriented with its plane perpendicular to the surface so that its normal **t** is tangent to the surface.

not contribute to the integrals on the left in (I.13) and (I.14). Only the top and bottom contribute. If the top and the bottom are parallel, tangent to the surface, and of area Δa, then the left-hand integral in (I.13) is

$$\oint_S \mathbf{D} \cdot \mathbf{n} \, da = (\mathbf{D}_2 - \mathbf{D}_1) \cdot \mathbf{n} \, \Delta a$$

and similarly for (I.14). If the charge density ρ is singular at the interface so as to produce an idealized surface charge density σ, then the integral on the right in (I.13) is

$$\int_V \rho \, d^3x = \sigma \, \Delta a$$

Thus the *normal components* of **D** and **B** on either side of the boundary surface are related according to

$$(\mathbf{D}_2 - \mathbf{D}_1) \cdot \mathbf{n} = \sigma \tag{I.17}$$

$$(\mathbf{B}_2 - \mathbf{B}_1) \cdot \mathbf{n} = 0 \tag{I.18}$$

In words, we say that the normal component of **B** is continuous and the discontinuity of the normal component of **D** at any point is equal to the surface charge density at that point.

In an analogous manner the infinitesimal Stokesian loop can be used to determine the discontinuities of the tangential components of **E** and **H**. If the short arms of the contour C in Fig. I.4 are of negligible length and each long arm is parallel to the surface and has length Δl, then the left-hand integral of (I.16) is

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = (\mathbf{t} \times \mathbf{n}) \cdot (\mathbf{E}_2 - \mathbf{E}_1) \, \Delta l$$

and similarly for the left-hand side of (I.15). The right-hand side of (I.16) vanishes because $\partial \mathbf{B}/\partial t$ is finite at the surface and the area of the loop is zero as the length of the short sides goes to zero. The right-hand side of (I.15) does not vanish, however, if there is an idealized surface current density **K** flowing exactly on the boundary surface. In such circumstances the integral on the right of (I.15) is

$$\int_{S'} \left[\mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \right] \cdot \mathbf{t} \, da = \mathbf{K} \cdot \mathbf{t} \, \Delta l$$

The second term in the integral vanishes by the same argument that was just given. The *tangential components* of **E** and **H** on either side of the boundary are therefore related by

$$\mathbf{n} \times (\mathbf{E}_2 - \mathbf{E}_1) = 0 \tag{I.19}$$

$$\mathbf{n} \times (\mathbf{H}_2 - \mathbf{H}_1) = \mathbf{K} \tag{I.20}$$

In (I.20) it is understood that the surface current **K** has only components parallel to the surface at every point. The tangential component of **E** across an interface is continuous, while the tangential component of **H** is discontinuous by an amount whose magnitude is equal to the magnitude of the surface current density and whose direction is parallel to $\mathbf{K} \times \mathbf{n}$.

The discontinuity equations (I.17)–(I.20) are useful in solving the Maxwell

equations in different regions and then connecting the solutions to obtain the fields throughout all space.

I.6 Some Remarks on Idealizations in Electromagnetism

In the preceding section we made use of the idea of surface distributions of charge and current. These are obviously mathematical idealizations that do not exist in the physical world. There are other abstractions that occur throughout electromagnetism. In electrostatics, for example, we speak of holding objects at a fixed potential with respect to some zero of potential usually called "ground." The relations of such idealizations to the real world is perhaps worthy of a little discussion, even though to the experienced hand most will seem obvious.

First we consider the question of maintaining some conducting object at a fixed electrostatic potential with respect to some reference value. Implicit is the idea that the means does not significantly disturb the desired configuration of charges and fields. To maintain an object at fixed potential it is necessary, at least from time to time, to have a conducting path or its equivalent from the object to a source of charge far away ("at infinity") so that as other charged or uncharged objects are brought in the vicinity, charge can flow to or from the object, always maintaining its potential at the desired value. Although more sophisticated means are possible, metallic wires are commonly used to make the conducting path. Intuitively we expect small wires to be less perturbing than large ones. The reason is as follows:

> Since the quantity of electricity on any given portion of a wire at a given potential diminishes indefinitely when the diameter of the wire is indefinitely diminished, the distribution of electricity on bodies of considerable dimensions will not be sensibly affected by the introduction of very fine metallic wires into the field, such as are used to form electrical connexions between these bodies and the earth, an electrical machine, or an electrometer.*

The electric field in the immediate neighborhood of the thin wire is very large, of course. However, at distances away of the order of the size of the "bodies of considerable dimensions" the effects can be made small. An important historical illustration of Maxwell's words is given by the work of Henry Cavendish 200 years ago. By experiments done in a converted stable of his father's house, using Leyden jars as his sources of charge, thin wires as conductors, and suspending the objects in the room, Cavendish measured the amounts of charge on cylinders, discs, etc., held at fixed potential and compared them to the charge on a sphere (the same sphere shown in Fig. I.1) at the same potential. His values of capacitance, so measured, are accurate to a few per cent. For example, he found the ratio of the capacitance of a sphere to that of a thin circular disc of the same radius was 1.57. The theoretical value is $\pi/2$.

There is a practical limit to the use of finer and finer wires. The charge per unit length decreases only logarithmically [as the reciprocal of $\ln(d/a)$, where a is the mean radius of the wire and d is a typical distance of the wire from some conducting surface]. To minimize the perturbation of the system below some level, it is necessary to resort to other means to maintain potentials, comparison methods using beams of charged particles intermittently, for example.

*J. C. Maxwell, *A Treatise on Electricity and Magnetism*, Dover, New York, 1954 reprint of the 3rd edition (1891), Vol. 1, p. 96.

When a conducting object is said to be *grounded*, it is assumed to be connected by a very fine conducting filament to a remote reservoir of charge that serves as the common zero of potential. Objects held at fixed potentials are similarly connected to one side of a voltage source, such as a battery, the other side of which is connected to the common "ground." Then, when initially electrified objects are moved relative to one another in such a way that their distributions of electricity are altered, but their potentials remain fixed, the appropriate amounts of charge flow from or to the remote reservoir, assumed to have an inexhaustible supply. The idea of grounding something is a well-defined concept in electrostatics, where time is not a factor, but for oscillating fields the finite speed of propagation blurs the concept. In other words, stray inductive and capacitive effects can enter significantly. Great care is then necessary to ensure a "good ground."

Another idealization in macroscopic electromagnetism is the idea of a surface charge density or a surface current density. The physical reality is that the charge or current is confined to the immediate neighborhood of the surface. If this region has thickness small compared to the length scale of interest, we may approximate the reality by the idealization of a region of infinitesimal thickness and speak of a surface distribution. Two different limits need to be distinguished. One is the limit in which the "surface" distribution is confined to a region near the surface that is *macroscopically small, but microscopically large*. An example is the penetration of time-varying fields into a very good, but not perfect, conductor, described in Section 8.1. It is found that the fields are confined to a thickness δ, called the skin depth, and that for high enough frequencies and good enough conductivities δ can be macroscopically very small. It is then appropriate to integrate the current density $\mathbf{J}$ over the direction perpendicular to the surface to obtain an effective surface current density $\mathbf{K}_{\text{eff}}$.

The other limit is *truly microscopic* and is set by quantum-mechanical effects in the atomic structure of materials. Consider, for instance, the distribution of excess charge of a conducting body in electrostatics. It is well known that this charge lies entirely on the surface of a conductor. We then speak of a *surface* charge density σ. There is no electric field inside the conductor, but there is, in accord with (I.17), a normal component of electric field just outside the surface. At the microscopic level the charge is not exactly at the surface and the field does not change discontinuously. The most elementary considerations would indicate that the transition region is a few atomic diameters in extent. The ions in a metal can be thought of as relatively immobile and localized to 1 angstrom or better; the lighter electrons are less constrained. The results of model calculations* are shown in Fig. I.5. They come from a solution of the quantum-mechanical many-electron problem in which the ions of the conductor are approximated by a continuous constant charge density for $x < 0$. The electron density ($r_s = 5$) is roughly appropriate to copper and the heavier alkali metals.

*N. D. Lang and W. Kohn, *Phys. Rev.* **B1**, 4555 (1970); **B3**, 1215 (1971); V. E. Kenner, R. E. Allen, and W. M. Saslow, *Phys. Lett.* **38A**, 255 (1972).

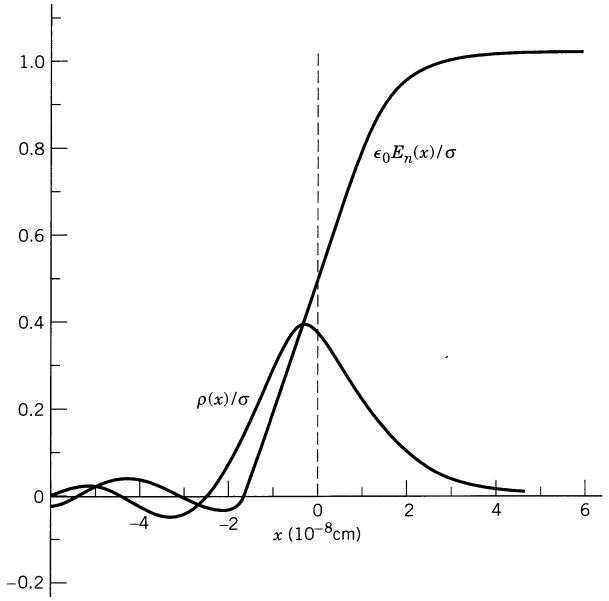

Figure I.5 Distribution of excess charge at the surface of a conductor and of the normal component of the electric field. The ions of the solid are confined to $x < 0$ and are approximated by a constant continuous charge distribution through which the electrons move. The bulk of the excess charge is confined to within ± 2 Å of the "surface."

The excess electronic charge is seen to be confined to a region within ± 2 Å of the "surface" of the ionic distribution. The electric field rises smoothly over this region to its value of σ "outside" the conductor. For macroscopic situations where 10^{-9} m is a negligible distance, we can idealize the charge density and electric field behavior as $\rho(x) = \sigma\delta(x)$ and $E_n(x) = \sigma\theta(x)/\epsilon_0$, corresponding to a truly surface density and a step-function jump of the field.

We see that the theoretical treatment of classical electromagnetism involves several idealizations, some of them technical and some physical. The subject of electrostatics, discussed in the first chapters of the book, developed as an experimental science of *macroscopic* electrical phenomena, as did virtually all other aspects of electromagnetism. The extension of these macroscopic laws, even for charges and currents in vacuum, to the *microscopic* domain was for the most part an unjustified extrapolation. Earlier in this introduction we discussed some of the limits to this extrapolation. The point to be made here is the following. With hindsight we know that many aspects of the laws of classical electromagnetism apply well into the atomic domain provided the sources are treated quantum mechanically, that the averaging of electromagnetic quantities over volumes containing large numbers of molecules so smooths the rapid fluctuations that static applied fields induce static average responses in matter, and that excess charge is *on* the surface of a conductor in a macroscopic sense. Thus Coulomb's and Ampère's macroscopic observations and our mathematical abstractions from them have a wider applicability than might be supposed by a supercautious physicist. The absence for air of significant electric or magnetic susceptibility certainly simplifies matters!

References and Suggested Reading

The history of electricity and magnetism is in large measure the history of science itself. We have already cited Whittaker's two volumes, the first covering the period up to 1900, as well as the shorter account emphasizing optics in

Born and Wolf.

Another readable account, with perceptive discussion of the original experiments, is

N. Feather, *Electricity and Matter*, University Press, Edinburgh (1968).

The experimental tests of the inverse square nature of Coulomb's law or, in modern language, the mass of the photon, are reviewed by

I. Yu. Kobzarev and L. B. Okun', *Usp. Fiz. Nauk* **95**, 131 (1968) [transl., *Sov. Phys. Usp.* **11**, 338 (1968).]

and

A. S. Goldhaber and M. M. Nieto, *Rev. Mod. Phys.* **43**, 277 (1971).

An accessible treatment of the gauge principle in the construction of field theories, building on classical electrodynamics and ordinary quantum mechanics, can be found in

I. J. R. Aitchison and A. J. G. Hey, *Gauge Theories in Particle Physics*, 2nd ed., Adam Hilger, Bristol (1989).

Suggested reading on the topic of the macroscopic Maxwell equations and their derivation from the microscopic equations can be found at the end of Chapter 6. The basic physics of dielectrics, ferroelectrics, and magnetic materials can be found in numerous books on solid-state physics, for example,

Ashcroft and Mermin
Beam
Kittel
Wert and Thomson
Wooten

The second of these is aimed at electrical engineers and stresses practical topics like semiconductors. The last one is mainly on optical properties. The need for spatial nonlocality in treating the surface impedance of metals (the anomalous skin effect) is discussed in several places by

A. B. Pippard, *Advances in Electronics and Electron Physics*, Vol. VI, ed. L. Marton, Academic Press, New York (1954), pp. 1–45; *Reports on Progress in Physics*, Vol. XXIII, pp. 176–266 (1960); *The Dynamics of Conduction Electrons*, Gordon and Breach, New York (1965).

The concept of a wave-vector and frequency-dependent dielectric constant $\epsilon(\mathbf{k}, \omega)$ is developed by

Kittel, *Advanced Topic D.*
D. Pines, *Elementary Excitations in Solids*, W. A. Benjamin, New York (1963), Chapters 3 and 4.
F. Stern, *Solid State Physics*, Vol. 15, eds. F. Seitz and D. Turnbull, Academic Press, New York, pp. 299–408.

The field of nonlinear optics is now nearly 40 years old. Beginnings and introductions can be found in

J. A. Giordmaine, *Phys. Today* **22**(1), 38 (1969).
N. Bloembergen, *Am. J. Phys.* **35**, 989 (1967).

Nonlinear optical phenomena and applications are discussed in

R. L. Sutherland, *Handbook of Nonlinear Optics*, Marcel Dekker, New York (1966).

Some texts and monographs on the subject are

R. W. Boyd, *Nonlinear Optics*, Academic Press, New York (1990).

M. Schubert and B. Wilhelmi, *Nonlinear Optics and Quantum Electronics*, Wiley, New York (1986).

Y. R. Shen, *The Principles of Nonlinear Optics*, Wiley, New York (1984).

CHAPTER 1

Introduction to Electrostatics

We begin our discussion of electrodynamics with the subject of *electrostatics*—phenomena involving time-independent distributions of charge and fields. For most readers this material is in the nature of a review. In this chapter especially we do not elaborate significantly. We introduce concepts and definitions that are important for later discussion and present some essential mathematical apparatus. In subsequent chapters the mathematical techniques are developed and applied.

One point of physics should be mentioned. Historically, electrostatics developed as a science of *macroscopic* phenomena. As indicated at the end of the Introduction, such idealizations as point charges or electric fields at a point must be viewed as mathematical constructs that permit a description of the phenomena at the macroscopic level, but that may fail to have meaning microscopically.

1.1 Coulomb's Law

All of electrostatics stems from the quantitative statement of Coulomb's law concerning the force acting between charged bodies at rest with respect to each other. Coulomb, in an impressive series of experiments, showed experimentally that the force between two small charged bodies separated in air a distance large compared to their dimensions

varies directly as the magnitude of each charge,

varies inversely as the square of the distance between them,

is directed along the line joining the charges, and

is attractive if the bodies are oppositely charged and repulsive if the bodies have the same type of charge.

Furthermore it was shown experimentally that the total force produced on one small charged body by a number of the other small charged bodies placed around it is the *vector* sum of the individual two-body forces of Coulomb. Strictly speaking, Coulomb's conclusions apply to charges in vacuum or in media of negligible susceptibility. We defer consideration of charges in dielectrics to Chapter 4.

1.2 Electric Field

Although the thing that eventually gets measured is a force, it is useful to introduce a concept one step removed from the forces, the concept of an electric field due to some array of charged bodies. At the moment, the electric field can be

defined as the force per unit charge acting at a given point. It is a vector function of position, denoted by **E**. One must be careful in its definition, however. It is not necessarily the force that one would observe by placing one unit of charge on a pith ball and placing it in position. The reason is that one unit of charge may be so large that its presence alters appreciably the field configuration of the array. Consequently one must use a limiting process whereby the ratio of the force on the small test body to the charge on it is measured for smaller and smaller amounts of charge.* Experimentally, this ratio and the direction of the force will become constant as the amount of test charge is made smaller and smaller. These limiting values of magnitude and direction define the magnitude and direction of the electric field **E** at the point in question. In symbols we may write

$$\mathbf{F} = q\mathbf{E} \tag{1.1}$$

where **F** is the force, **E** the electric field, and q the charge. In this equation it is assumed that the charge q is located at a point, and the force and the electric field are evaluated at that point.

Coulomb's law can be written down similarly. If **F** is the force on a point charge q_1, located at $\mathbf{x}_1$, due to another point charge q_2, located at $\mathbf{x}_2$, then Coulomb's law is

$$\mathbf{F} = kq_1q_2 \frac{\mathbf{x}_1 - \mathbf{x}_2}{|\mathbf{x}_1 - \mathbf{x}_2|^3} \tag{1.2}$$

Note that q_1 and q_2 are algebraic quantities, which can be positive or negative. The constant of proportionality k depends on the system of units used.

The electric field at the point **x** due to a point charge q_1 at the point $\mathbf{x}_1$ can be obtained directly:

$$\mathbf{E}(\mathbf{x}) = kq_1 \frac{\mathbf{x} - \mathbf{x}_1}{|\mathbf{x} - \mathbf{x}_1|^3} \tag{1.3}$$

as indicated in Fig. 1.1. The constant k differs in different systems of units.† In electrostatic units (esu), $k = 1$ and unit charge is chosen as that charge that exerts a force of one dyne on an equal point charge located one centimeter away. The esu unit of charge is called the *statcoulomb*, and the electric field is measured in *statvolts per centimeter*. In the SI system, which we employ here, $k = (4\pi\epsilon_0)^{-1} = 10^{-7}c^2$, where $\epsilon_0 \approx 8.854 \times 10^{-12}$ farad per meter (F/m) is called the permittivity of free space. The SI unit of charge is the *coulomb* (C), and the electric field is measured in *volts per meter* (V/m). One coulomb (1 C) produces an electric field

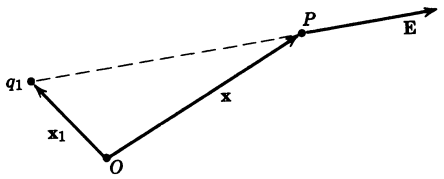

Figure 1.1

*The discreteness of electric charge (see Section I.1) means that this mathematical limit is impossible to realize physically. This is an example of a mathematical idealization in macroscopic electrostatics.

†The question of units is discussed in detail in the Appendix.

of approximately 8.9874×10^9 V/m (8.9874 GV/m) at a distance of 1 meter. One electron ($q \approx 1.602 \times 10^{-19}$ C) produces a field of approximately 1.44×10^{-9} V/m (1.44 nV/m) at 1 meter.

The experimentally observed linear superposition of forces due to many charges means that we may write the electric field at **x** due to a system of point charges q_i, located at $\mathbf{x}_i$, $i = 1, 2, \ldots, n$, as the vector sum:

$$\mathbf{E}(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^{n} q_i \frac{\mathbf{x} - \mathbf{x}_i}{|\mathbf{x} - \mathbf{x}_i|^3} \tag{1.4}$$

If the charges are so small and so numerous that they can be described by a charge density $\rho(\mathbf{x}')$ [if Δq is the charge in a small volume $\Delta x\, \Delta y\, \Delta z$ at the point $\mathbf{x}'$, then $\Delta q = \rho(\mathbf{x}')\, \Delta x\, \Delta y\, \Delta z$], the sum is replaced by an integral:

$$\mathbf{E}(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \int \rho(\mathbf{x}') \frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|^3} d^3x' \tag{1.5}$$

where $d^3x' = dx'\, dy'\, dz'$ is a three-dimensional volume element at $\mathbf{x}'$.

At this point it is worthwhile to introduce the *Dirac delta function*. In one dimension, the delta function, written $\delta(x - a)$, is a mathematically improper function having the properties:

1. $\delta(x - a) = 0$ for $x \neq a$, and
2. $\int \delta(x - a)\, dx = 1$ if the region of integration includes $x = a$, and is zero otherwise.

The delta function can be given an intuitive, but nonrigorous, meaning as the limit of a peaked curve such as a Gaussian that becomes narrower and narrower, but higher and higher, in such a way that the area under the curve is always constant. L. Schwartz's theory of distributions is a comprehensive rigorous mathematical approach to delta functions and their manipulations.*

From the definitions above it is evident that, for an arbitrary function $f(x)$,

3. $\int f(x)\, \delta(x - a)\, dx = f(a)$.

The integral of $f(x)$ times the derivative of a delta function is simply understood if the delta function is thought of as a well-behaved, but sharply peaked, function. Thus the definition is

4. $\int f(x)\, \delta'(x - a)\, dx = -f'(a)$

where a prime denotes differentiation with respect to the argument.

If the delta function has as argument a function $f(x)$ of the independent variable x, it can be transformed according to the rule,

5. $$\delta(f(x)) = \sum_i \frac{1}{\left|\dfrac{df}{dx}(x_i)\right|} \delta(x - x_i)$$

where $f(x)$ is assumed to have only simple zeros, located at $x = x_i$.

In more than one dimension, we merely take products of delta functions in each dimension. In three dimensions, for example, with Cartesian coordinates,

6. $\delta(\mathbf{x} - \mathbf{X}) = \delta(x_1 - X_1)\, \delta(x_2 - X_2)\, \delta(x_3 - X_3)$

*A useful, rigorous account of the Dirac delta function is given by *Lighthill*. See also *Dennery and Krzywicki* (Section III.13). (Full references for items cited in the text or footnotes by italicized author only will be found in the Bibliography.)

is a function that vanishes everywhere except at $\mathbf{x} = \mathbf{X}$, and is such that

7. $\int_{\Delta V} \delta(\mathbf{x} - \mathbf{X})\, d^3x = \begin{cases} 1 & \text{if } \Delta V \text{ contains } \mathbf{x} = \mathbf{X} \\ 0 & \text{if } \Delta V \text{ does not contain } \mathbf{x} = \mathbf{X} \end{cases}$

Note that a delta function has the dimensions of an inverse volume in whatever number of dimensions the space has.

A discrete set of point charges can be described with a charge density by means of delta functions. For example,

$$\rho(\mathbf{x}) = \sum_{i=1}^{n} q_i\, \delta(\mathbf{x} - \mathbf{x}_i) \tag{1.6}$$

represents a distribution of n point charges q_i, located at the points $\mathbf{x}_i$. Substitution of this charge density (1.6) into (1.5) and integration, using the properties of the delta function, yields the discrete sum (1.4).

1.3 Gauss's Law

The integral (1.5) is not always the most suitable form for the evaluation of electric fields. There is another integral result, called *Gauss's law*, which is sometimes more useful and furthermore leads to a differential equation for $\mathbf{E}(\mathbf{x})$. To obtain Gauss's law we first consider a point charge q and a *closed* surface S, as shown in Fig. 1.2. Let r be the distance from the charge to a point on the surface, $\mathbf{n}$ be the outwardly directed unit normal to the surface at that point, da be an

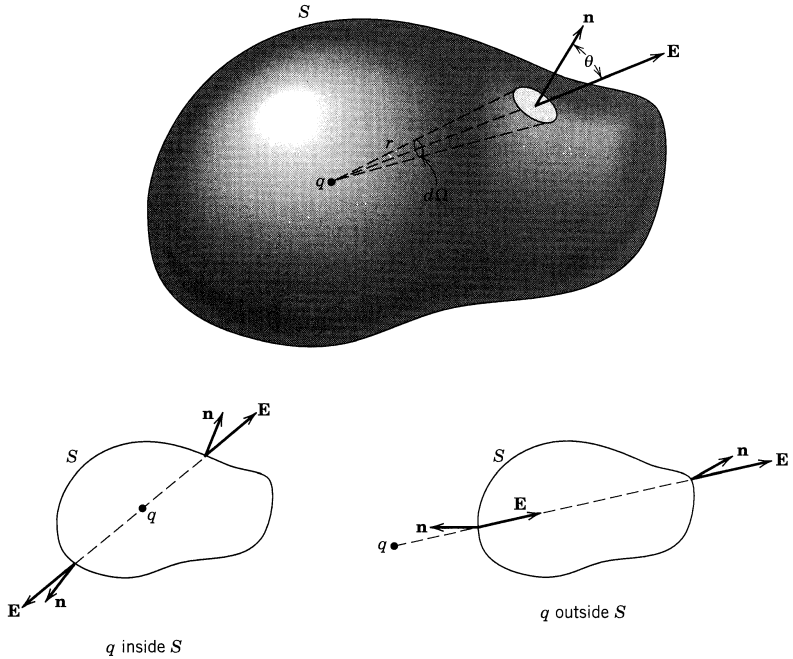

Figure 1.2 Gauss's law. The normal component of electric field is integrated over the closed surface S. If the charge is inside (outside) S, the total solid angle subtended at the charge by the inner side of the surface is 4π (zero).

element of surface area. If the electric field $\mathbf{E}$ at the point on the surface due to the charge q makes an angle θ with the unit normal, then the normal component of $\mathbf{E}$ times the area element is:

$$\mathbf{E} \cdot \mathbf{n}\, da = \frac{q}{4\pi\epsilon_0} \frac{\cos\theta}{r^2}\, da \tag{1.7}$$

Since $\mathbf{E}$ is directed along the line from the surface element to the charge q, $\cos\theta\, da = r^2\, d\Omega$, where $d\Omega$ is the element of solid angle subtended by da at the position of the charge. Therefore

$$\mathbf{E} \cdot \mathbf{n}\, da = \frac{q}{4\pi\epsilon_0}\, d\Omega \tag{1.8}$$

If we now integrate the normal component of $\mathbf{E}$ over the whole surface, it is easy to see that

$$\oint_S \mathbf{E} \cdot \mathbf{n}\, da = \begin{cases} q/\epsilon_0 & \text{if } q \text{ lies } \textit{inside } S \\ 0 & \text{if } q \text{ lies } \textit{outside } S \end{cases} \tag{1.9}$$

This result is Gauss's law for a single point charge. For a discrete set of charges, it is immediately apparent that

$$\oint_S \mathbf{E} \cdot \mathbf{n}\, da = \frac{1}{\epsilon_0} \sum_i q_i \tag{1.10}$$

where the sum is over only those charges *inside* the surface S. For a continuous charge density $\rho(\mathbf{x})$, Gauss's law becomes:

$$\oint_S \mathbf{E} \cdot \mathbf{n}\, da = \frac{1}{\epsilon_0} \int_V \rho(\mathbf{x})\, d^3x \tag{1.11}$$

where V is the volume enclosed by S.

Equation (1.11) is one of the basic equations of electrostatics. Note that it depends upon

the inverse square law for the force between charges,
the central nature of the force, and
the linear superposition of the effects of different charges.

Clearly, then, Gauss's law holds for Newtonian gravitational force fields, with matter density replacing charge density.

It is interesting to note that, even before the experiments of Cavendish and Coulomb, Priestley, taking up an observation of Franklin that charge seemed to reside on the outside, but not the inside, of a metal cup, reasoned by analogy with Newton's law of universal gravitation that the electrostatic force must obey an inverse square law with distance. The present status of the inverse square law is discussed in Section I.2.

1.4 Differential Form of Gauss's Law

Gauss's law can be thought of as being an integral formulation of the law of electrostatics. We can obtain a differential form (i.e., a differential equation) by

using the divergence theorem. The *divergence theorem* states that for any well-behaved vector field $\mathbf{A}(\mathbf{x})$ defined within a volume V surrounded by the closed surface S the relation

$$\oint_S \mathbf{A} \cdot \mathbf{n}\, da = \int_V \nabla \cdot \mathbf{A}\, d^3x$$

holds between the volume integral of the divergence of $\mathbf{A}$ and the surface integral of the outwardly directed normal component of $\mathbf{A}$. The equation in fact can be used as the definition of the divergence (see *Stratton*, p. 4).

To apply the divergence theorem we consider the integral relation expressed in Gauss's theorem:

$$\oint_S \mathbf{E} \cdot \mathbf{n}\, da = \frac{1}{\epsilon_0} \int_V \rho(\mathbf{x})\, d^3x$$

Now the divergence theorem allows us to write this as

$$\int_V (\nabla \cdot \mathbf{E} - \rho/\epsilon_0)\, d^3x = 0 \tag{1.12}$$

for an arbitrary volume V. We can, in the usual way, put the integrand equal to zero to obtain

$$\nabla \cdot \mathbf{E} = \rho/\epsilon_0 \tag{1.13}$$

which is the differential form of Gauss's law of electrostatics. This equation can itself be used to solve problems in electrostatics. However, it is often simpler to deal with scalar rather then vector functions of position, and then to derive the vector quantities at the end if necessary (see below).

1.5 Another Equation of Electrostatics and the Scalar Potential

The single equation (1.13) is not enough to specify completely the three components of the electric field $\mathbf{E}(\mathbf{x})$. Perhaps some readers know that a vector field can be specified almost* completely if its divergence and curl are given everywhere in space. Thus we look for an equation specifying curl $\mathbf{E}$ as a function of position. Such an equation, namely,

$$\nabla \times \mathbf{E} = 0 \tag{1.14}$$

follows directly from our generalized Coulomb's law (1.5):

$$\mathbf{E}(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \int \rho(\mathbf{x}') \frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|^3}\, d^3x'$$

The vector factor in the integrand, viewed as a function of $\mathbf{x}$, is the negative gradient of the scalar $1/|\mathbf{x} - \mathbf{x}'|$:

$$\frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|^3} = -\nabla\left(\frac{1}{|\mathbf{x} - \mathbf{x}'|}\right)$$

*Up to the gradient of a scalar function that satisfies the Laplace equation. See Section 1.9 on uniqueness.

Since the gradient operation involves $\mathbf{x}$, but not the integration variable $\mathbf{x}'$, it can be taken outside the integral sign. Then the field can be written

$$\mathbf{E}(\mathbf{x}) = \frac{-1}{4\pi\epsilon_0} \nabla \int \frac{\rho(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|}\, d^3x' \tag{1.15}$$

Since the curl of the gradient of any well-behaved scalar function of position vanishes ($\nabla \times \nabla\psi = 0$, for all ψ), (1.14) follows immediately from (1.15).

Note that $\nabla \times \mathbf{E} = 0$ depends on the central nature of the force between charges, and on the fact that the force is a function of relative distances only, but does not depend on the inverse square nature.

In (1.15) the electric field (a vector) is derived from a scalar by the gradient operation. Since one function of position is easier to deal with than three, it is worthwhile concentrating on the scalar function and giving it a name. Consequently we define the *scalar potential* $\Phi(\mathbf{x})$ by the equation:

$$\mathbf{E} = -\nabla\Phi \tag{1.16}$$

Then (1.15) shows that the scalar potential is given in terms of the charge density by

$$\Phi(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|}\, d^3x' \tag{1.17}$$

where the integration is over all charges in the universe, and Φ is arbitrary only to the extent that a constant can be added to the right-hand side of (1.17).

The scalar potential has a physical interpretation when we consider the work done on a test charge q in transporting it from one point (A) to another point (B) in the presence of an electric field $\mathbf{E}(\mathbf{x})$, as shown in Fig. 1.3. The force acting on the charge at any point is

$$\mathbf{F} = q\mathbf{E}$$

so that the work done in moving the charge from A to B is

$$W = -\int_A^B \mathbf{F} \cdot d\mathbf{l} = -q \int_A^B \mathbf{E} \cdot d\mathbf{l} \tag{1.18}$$

The minus sign appears because we are calculating the work done *on* the charge against the action of the field. With definition (1.16) the work can be written

$$W = q \int_A^B \nabla\Phi \cdot d\mathbf{l} = q \int_A^B d\Phi = q(\Phi_B - \Phi_A) \tag{1.19}$$

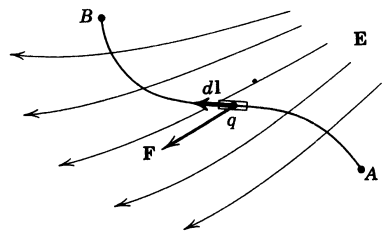

Figure 1.3

which shows that $q\Phi$ can be interpreted as the potential energy of the test charge in the electrostatic field.

From (1.18) and (1.19) it can be seen that the line integral of the electric field between two points is independent of the path and is the negative of the potential difference between the points:

$$\int_A^B \mathbf{E} \cdot d\mathbf{l} = -(\Phi_B - \Phi_A) \tag{1.20}$$

This follows directly, of course, from definition (1.16). If the path is closed, the line integral is zero,

$$\oint \mathbf{E} \cdot d\mathbf{l} = 0 \tag{1.21}$$

a result that can also be obtained directly from Coulomb's law. Then application of *Stokes's theorem* [if $\mathbf{A}(\mathbf{x})$ is a well-behaved vector field, S is an arbitrary open surface, and C is the closed curve bounding S,

$$\oint_C \mathbf{A} \cdot d\mathbf{l} = \int_S (\nabla \times \mathbf{A}) \cdot \mathbf{n}\, da$$

where $d\mathbf{l}$ is a line element of C, $\mathbf{n}$ is the normal to S, and the path C is traversed in a right-hand screw sense relative to $\mathbf{n}$] leads immediately back to $\nabla \times \mathbf{E} = 0$.

1.6 Surface Distributions of Charges and Dipoles and Discontinuities in the Electric Field and Potential

One of the common problems in electrostatics is the determination of electric field or potential due to a given surface distribution of charges. Gauss's law (1.11) allows us to write down a partial result directly. If a surface S, with a unit normal $\mathbf{n}$ directed from side 1 to side 2 of the surface, has a surface-charge density of $\sigma(\mathbf{x})$ (measured in coulombs per square meter) and electric fields $\mathbf{E}_1$ and $\mathbf{E}_2$ on either side of the surface, as shown in Fig. 1.4, then Gauss's law tells us immediately that

$$(\mathbf{E}_2 - \mathbf{E}_1) \cdot \mathbf{n} = \sigma/\epsilon_0 \tag{1.22}$$

This does not determine $\mathbf{E}_1$ and $\mathbf{E}_2$ unless there are no other sources of field and the geometry and form of σ are especially simple. All that (1.22) says is that there

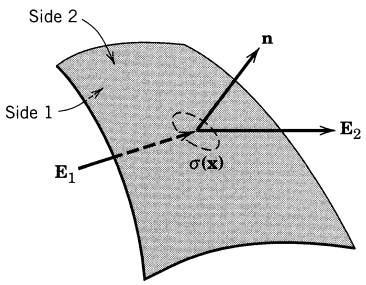

Figure 1.4 Discontinuity in the normal component of electric field across a surface layer of charge.

is a discontinuity of σ/ϵ_0 in the normal component of electric field in crossing a surface with a surface-charge density σ, the crossing being made in the direction of $\mathbf{n}$.

The tangential component of electric field can be shown to be continuous across a boundary surface by using (1.21) for the line integral of $\mathbf{E}$ around a closed path. It is only necessary to take a rectangular path with negligible ends and one side on either side of the boundary.

An expression for the potential (hence the field, by differentiation) at any point in space (not just at the surface) can be obtained from (1.17) by replacing $\rho\, d^3x$ by $\sigma\, da$:

$$\Phi(\mathbf{x}) = \frac{1}{4\pi\epsilon_0} \int_S \frac{\sigma(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|}\, da' \tag{1.23}$$

For volume or surface distributions of charge, the potential is everywhere continuous, even within the charge distribution. This can be shown from (1.23) or from the fact that $\mathbf{E}$ is bounded, even though discontinuous across a surface distribution of charge. With point or line charges, or dipole layers, the potential is no longer continuous, as will be seen immediately.

Another problem of interest is the potential due to a dipole-layer distribution on a surface S. A dipole layer can be imagined as being formed by letting the surface S have a surface-charge density $\sigma(\mathbf{x})$ on it, and another surface S', lying close to S, have an equal and opposite surface-charge density on it at neighboring points, as shown in Fig. 1.5. The dipole-layer distribution of strength $D(\mathbf{x})$ is formed by letting S' approach infinitesimally close to S while the surface-charge density $\sigma(\mathbf{x})$ becomes infinite in such a manner that the product of $\sigma(\mathbf{x})$ and the local separation $d(\mathbf{x})$ of S and S' approaches the limit $D(\mathbf{x})$:

$$\lim_{d(\mathbf{x})\to 0} \sigma(\mathbf{x})\, d(\mathbf{x}) = D(\mathbf{x})$$

The direction of the dipole moment of the layer is normal to the surface S and in the direction going from negative to positive charge.

To find the potential due to a dipole layer we can consider a single dipole and then superpose a surface density of them, or we can obtain the same result by performing mathematically the limiting process described in words above on the surface-density expression (1.23). The first way is perhaps simpler, but the second gives useful practice in vector calculus. Consequently we proceed with

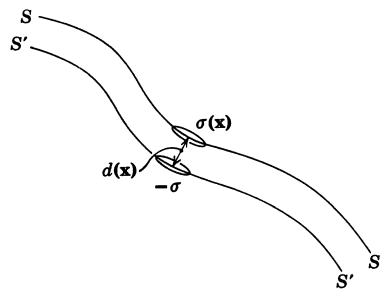

Figure 1.5 Limiting process involved in creating a dipole layer.

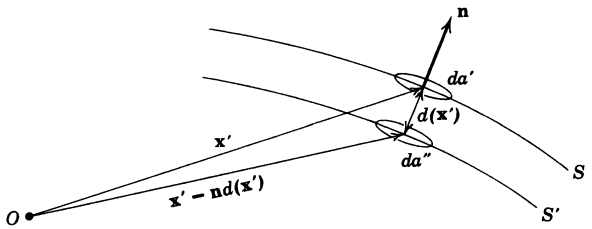

Figure 1.6 Dipole-layer geometry.

the limiting process. With $\mathbf{n}$, the unit normal to the surface S, directed away from S', as shown in Fig. 1.6, the potential due to the two close surfaces is

$$\Phi(\mathbf{x}) = \frac{1}{4\pi\epsilon_0}\int_S \frac{\sigma(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\,da' - \frac{1}{4\pi\epsilon_0}\int_{S'} \frac{\sigma(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'+\mathbf{n}d|}\,da''$$

For small d we can expand $|\mathbf{x}-\mathbf{x}'+\mathbf{n}d|^{-1}$. Consider the general expression $|\mathbf{x}+\mathbf{a}|^{-1}$, where $|\mathbf{a}| \ll |\mathbf{x}|$. We write a Taylor series expansion in three dimensions:

$$\frac{1}{|\mathbf{x}+\mathbf{a}|} = \frac{1}{x} + \mathbf{a}\cdot\nabla\left(\frac{1}{x}\right) + \cdots$$

In this way we find that as $d \to 0$ the potential becomes

$$\Phi(\mathbf{x}) = \frac{1}{4\pi\epsilon_0}\int_S D(\mathbf{x}')\mathbf{n}\cdot\nabla'\left(\frac{1}{|\mathbf{x}-\mathbf{x}'|}\right)da' \tag{1.24}$$

In passing we note that the integrand in (1.24) is the potential of a point dipole with dipole moment $\mathbf{p} = \mathbf{n}\,D\,da'$. The potential at $\mathbf{x}$ caused by a dipole $\mathbf{p}$ at $\mathbf{x}'$ is

$$\Phi(\mathbf{x}) = \frac{1}{4\pi\epsilon_0}\frac{\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|^3} \tag{1.25}$$

Equation (1.24) has a simple geometrical interpretation. We note that

$$\mathbf{n}\cdot\nabla'\left(\frac{1}{|\mathbf{x}-\mathbf{x}'|}\right)da' = -\frac{\cos\theta\,da'}{|\mathbf{x}-\mathbf{x}'|^2} = -d\Omega$$

where $d\Omega$ is the element of solid angle subtended at the observation point by the area element da', as indicated in Fig. 1.7. Note that $d\Omega$ has a positive sign if θ is

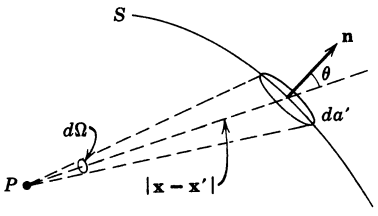

Figure 1.7 The potential at P due to the dipole layer D on the area element da' is just the negative product of D and the solid angle element $d\Omega$ subtended by da' at P.

an acute angle (i.e., when the observation point views the "inner" side of the dipole layer). The potential can be written:

$$\Phi(\mathbf{x}) = -\frac{1}{4\pi\epsilon_0}\int_S D(\mathbf{x}')\,d\Omega \tag{1.26}$$

For a constant surface-dipole-moment density D, the potential is just the product of the moment divided by $4\pi\epsilon_0$ and the solid angle subtended at the observation point by the surface, regardless of its shape.

There is a discontinuity in potential in crossing a double layer. This can be seen by letting the observation point come infinitesimally close to the double layer. The double layer is now imagined to consist of two parts, one being a small disc directly under the observation point. The disc is sufficiently small that it is sensibly flat and has constant surface-dipole-moment density D. Evidently the total potential can be obtained by linear superposition of the potential of the disc and that of the remainder. From (1.26) it is clear that the potential of the disc alone has a discontinuity of D/ϵ_0 in crossing from the inner to the outer side, being $-D/2\epsilon_0$ on the inner side and $+D/2\epsilon_0$ on the outer. The potential of the remainder alone, with its hole where the disc fits in, is continuous across the plane of the hole. Consequently the total potential jump in crossing the surface is:

$$\Phi_2 - \Phi_1 = D/\epsilon_0 \tag{1.27}$$

This result is analogous to (1.22) for the discontinuity of electric field in crossing a surface-charge density. Equation (1.27) can be interpreted "physically" as a potential drop occurring "inside" the dipole layer; it can be calculated as the product of the field between the two layers of surface charge times the separation before the limit is taken.

1.7 Poisson and Laplace Equations

In Sections 1.4 and 1.5 it was shown that the behavior of an electrostatic field can be described by the two differential equations:

$$\nabla\cdot\mathbf{E} = \rho/\epsilon_0 \tag{1.13}$$

and

$$\nabla\times\mathbf{E} = 0 \tag{1.14}$$

the latter equation being equivalent to the statement that $\mathbf{E}$ is the gradient of a scalar function, the scalar potential Φ:

$$\mathbf{E} = -\nabla\Phi \tag{1.16}$$

Equations (1.13) and (1.16) can be combined into one partial differential equation for the single function $\Phi(\mathbf{x})$:

$$\nabla^2\Phi = -\rho/\epsilon_0 \tag{1.28}$$

This equation is called the *Poisson equation*. In regions of space that lack a charge density, the scalar potential satisfies the *Laplace equation*:

$$\nabla^2\Phi = 0 \tag{1.29}$$

We already have a solution for the scalar potential in expression (1.17):

$$\Phi(\mathbf{x}) = \frac{1}{4\pi\epsilon_0}\int \frac{\rho(\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|}\,d^3x' \tag{1.17}$$

To verify directly that this does indeed satisfy the Poisson equation (1.28), we operate with the Laplacian on both sides. Because it turns out that the resulting integrand is singular, we invoke a limiting procedure. Define the "a-potential" $\Phi_a(\mathbf{x})$ by

$$\Phi_a(\mathbf{x}) = \frac{1}{4\pi\epsilon_0}\int \frac{\rho(\mathbf{x}')}{\sqrt{(\mathbf{x}-\mathbf{x}')^2+a^2}}\,d^3x'$$

The actual potential (1.17) is then the limit of the "a-potential" as $a \to 0$. Taking the Laplacian of the "a-potential" gives

$$\begin{aligned}\nabla^2\Phi_a(\mathbf{x}) &= \frac{1}{4\pi\epsilon_0}\int \rho(\mathbf{x}')\nabla^2\left(\frac{1}{\sqrt{r^2+a^2}}\right)d^3x' \\ &= -\frac{1}{4\pi\epsilon_0}\int \rho(\mathbf{x}')\left[\frac{3a^2}{(r^2+a^2)^{5/2}}\right]d^3x'\end{aligned} \tag{1.30}$$

where $r = |\mathbf{x}-\mathbf{x}'|$. The square-bracketed expression is the negative Laplacian of $1/\sqrt{r^2+a^2}$. It is well-behaved everywhere for nonvanishing a, but as a tends to zero it becomes infinite at $r = 0$ and vanishes for $r \neq 0$. It has a volume integral equal to 4π for arbitrary a. For the purposes of integration, divide space into two regions by a sphere of fixed radius R centered on $\mathbf{x}$. Choose R such that $\rho(\mathbf{x}')$ changes little over the interior of the sphere, and imagine a much smaller than R and tending toward zero. If $\rho(\mathbf{x}')$ is such that (1.17) exists, the contribution to the integral (1.30) from the exterior of the sphere will vanish like a^2 as $a \to 0$. We thus need consider only the contribution from inside the sphere. With a Taylor series expansion of the well-behaved $\rho(\mathbf{x}')$ around $\mathbf{x}' = \mathbf{x}$, one finds

$$\nabla^2\Phi_a(\mathbf{x}) = -\frac{1}{\epsilon_0}\int_0^R \frac{3a^2}{(r^2+a^2)^{5/2}}\left[\rho(\mathbf{x}) + \frac{r^2}{6}\nabla^2\rho + \cdots\right]r^2dr + O(a^2)$$

Direct integration yields

$$\nabla^2\Phi_a(\mathbf{x}) = -\frac{1}{\epsilon_0}\rho(\mathbf{x})\left(1 + O(a^2/R^2)\right) + O(a^2, a^2\log a)\,\nabla^2\rho + \cdots$$

In the limit $a \to 0$, we obtain the Poisson equation (1.28).

The singular nature of the Laplacian of $1/r$ can be exhibited formally in terms of a Dirac delta function. Since $\nabla^2(1/r) = 0$ for $r \neq 0$ and its volume integral is -4π, we can write the formal equation, $\nabla^2(1/r) = -4\pi\delta(\mathbf{x})$ or, more generally,

$$\nabla^2\left(\frac{1}{|\mathbf{x}-\mathbf{x}'|}\right) = -4\pi\delta(\mathbf{x}-\mathbf{x}') \tag{1.31}$$

1.8 Green's Theorem

If electrostatic problems always involved localized discrete or continuous distributions of charge with no boundary surfaces, the general solution (1.17) would

be the most convenient and straightforward solution to any problem. There would be no need of the Poisson or Laplace equation. In actual fact, of course, many, if not most, of the problems of electrostatics involve finite regions of space, with or without charge inside, and with prescribed boundary conditions on the bounding surfaces. These boundary conditions may be simulated by an appropriate distribution of charges outside the region of interest (perhaps at infinity), but (1.17) becomes inconvenient as a means of calculating the potential, except in simple cases (e.g., method of images).

To handle the boundary conditions it is necessary to develop some new mathematical tools, namely, the identities or theorems due to George Green (1824). These follow as simple applications of the divergence theorem. The divergence theorem:

$$\int_V \boldsymbol{\nabla}\cdot\mathbf{A}\,d^3x = \oint_S \mathbf{A}\cdot\mathbf{n}\,da$$

applies to any well-behaved vector field $\mathbf{A}$ defined in the volume V bounded by the closed surface S. Let $\mathbf{A} = \phi\,\boldsymbol{\nabla}\psi$, where ϕ and ψ are arbitrary scalar fields. Now

$$\boldsymbol{\nabla}\cdot(\phi\,\boldsymbol{\nabla}\psi) = \phi\,\nabla^2\psi + \boldsymbol{\nabla}\phi\cdot\boldsymbol{\nabla}\psi \tag{1.32}$$

and

$$\phi\,\boldsymbol{\nabla}\psi\cdot\mathbf{n} = \phi\frac{\partial\psi}{\partial n} \tag{1.33}$$

where $\partial/\partial n$ is the normal derivative at the surface S (directed outward from inside the volume V). When (1.32) and (1.33) are substituted into the divergence theorem, there results *Green's first identity*:

$$\int_V (\phi\,\nabla^2\psi + \boldsymbol{\nabla}\phi\cdot\boldsymbol{\nabla}\psi)\,d^3x = \oint_S \phi\frac{\partial\psi}{\partial n}\,da \tag{1.34}$$

If we write down (1.34) again with ϕ and ψ interchanged, and then subtract it from (1.34), the $\boldsymbol{\nabla}\phi\cdot\boldsymbol{\nabla}\psi$ terms cancel, and we obtain Green's second identity or *Green's theorem*:

$$\int_V (\phi\,\nabla^2\psi - \psi\,\nabla^2\phi)\,d^3x = \oint_S \left[\phi\frac{\partial\psi}{\partial n} - \psi\frac{\partial\phi}{\partial n}\right]da \tag{1.35}$$

The Poisson differential equation for the potential can be converted into an integral equation if we choose a particular ψ, namely $1/R \equiv 1/|\mathbf{x}-\mathbf{x}'|$, where $\mathbf{x}$ is the observation point and $\mathbf{x}'$ is the integration variable. Further, we put $\phi = \Phi$, the scalar potential, and make use of $\nabla^2\Phi = -\rho/\epsilon_0$. From (1.31) we know that $\nabla^2(1/R) = -4\pi\delta(\mathbf{x}-\mathbf{x}')$, so that (1.35) becomes

$$\int_V\left[-4\pi\Phi(\mathbf{x}')\,\delta(\mathbf{x}-\mathbf{x}') + \frac{1}{\epsilon_0 R}\rho(\mathbf{x}')\right]d^3x' = \oint_S\left[\Phi\frac{\partial}{\partial n'}\left(\frac{1}{R}\right) - \frac{1}{R}\frac{\partial\Phi}{\partial n'}\right]da'$$

If the point $\mathbf{x}$ lies within the volume V, we obtain:

$$\Phi(\mathbf{x}) = \frac{1}{4\pi\epsilon_0}\int_V\frac{\rho(\mathbf{x}')}{R}\,d^3x' + \frac{1}{4\pi}\oint_S\left[\frac{1}{R}\frac{\partial\Phi}{\partial n'} - \Phi\frac{\partial}{\partial n'}\left(\frac{1}{R}\right)\right]da' \tag{1.36}$$