1 Introduction

1 The field concept

The central concept in the modern theory of electromagnetism is that of the electromagnetic *field*. The forces that electrical charges, currents, and magnets exert on each other were believed by early thinkers to be of the action-at-a-distance type, i.e., the forces acted instantaneously over arbitrarily large distances. Experiments have shown, however, that this is not true. A radio signal, for example, can be sent by moving electrons back and forth in a metallic antenna. This motion will cause electrons in a distant piece of metal to move back and forth in response—this is how the signal is picked up in a radio or cell phone receiver. We know that the electrons in the receiver cannot respond in a time less than that required by light to travel the distance between transmitter and receiver. Indeed, radio waves, or electromagnetic waves more generally, are a form of light.

Facts such as these have led us to abandon the notion of action at a distance. Instead, our present understanding is that electrical charges and currents produce physical entities called *fields*, which permeate the space around them and which in turn act on other charges and currents. When a charge moves, the fields that it creates change, but this change is not instantaneous at every point in space. For a complete description, one must introduce two *vector* fields, $\mathbf{E}(\mathbf{r}, t)$, and $\mathbf{B}(\mathbf{r}, t)$, which we will call the electric and magnetic fields, respectively. In other words, at every time t, and at every point in space $\mathbf{r}$, we picture the existence of two vectors, $\mathbf{E}$ and $\mathbf{B}$. This picture is highly abstract, and early physicists had great trouble in coming to grips with it. Because the fields did not describe particulate matter and could exist in vacuum, they seemed very intangible, and early physicists were reluctant to endow them with physical reality. The modern view is quite different. Not only do these fields allow us to describe the interaction of charges and currents with each other in the mathematically simplest and cleanest way, we now believe them to be absolutely real physical entities, as real as a rhinoceros. Light is believed to be nothing but a jumble of wiggling $\mathbf{E}$ and $\mathbf{B}$ vectors everywhere, which implies that these fields can exist independently of charges and currents. Secondly, these fields carry such concrete physical properties as energy, momentum, and angular momentum. When one gets to a quantum mechanical description, these three attributes become properties of a particle called the *photon*, a quantum of light. At sufficiently high energies, two of these particles can spontaneously change into an electron and a positron, in a process called *pair production*. Thus, there is no longer any reason for regarding the **E** and **B** fields as adjuncts, or aids to understanding, or to picture the interactions of charges through lines of force or flux. Indeed, it is the latter concepts that are now regarded as secondary, and the fields as primary.

The impossibility of action at a distance is codified into the modern theory of relativity. The principle of relativity as enunciated by Galileo states that the laws of physics are identical in all inertial reference frames.[1] One goes from Galilean relativity to the modern theory by recognizing that there is a maximum speed at which physical influences or signals may propagate, and since this is a law of physics, the maximum speed must then be the same in all inertial frames.[2] This speed immediately acquires the status of a fundamental constant of nature and is none other than the speed of light in vacuum. Needless to say, this law, and the many dramatic conclusions that follow from considering it in conjunction with the principle of relativity, are amply verified by experiment.

The application of the principle of relativity also leads us to discover that **E** and **B** are two aspects of the same thing. A static set of charges creates a time-independent electric field, and a steady current creates a time-independent magnetic field. Since a current can be regarded as a charge distribution in motion, it follows that **E** and **B** will, in general, transform into one another when we change reference frames. In fact, the relativistic invariance of the laws of electrodynamics is best expressed in terms of a single tensor field, generally denoted F. The fields **E** and **B** are obtained as different components of F. At low speeds, however, these two different components have so many dissimilar aspects that greater physical understanding is obtained by thinking of them as separate vector fields. This is what we shall do in this book.

2 The equations of electrodynamics

The full range of electromagnetic phenomena is very wide and can be very complicated. It is somewhat remarkable that it can be captured in a small number of equations of

[1] That such frames exist is a matter of physical experience, and actual frames can be made to approximate an ideal inertial reference frame as closely as we wish.

[2] Einstein took the frame invariance of the speed of light as a postulate in addition to the principle of relativity. It was recognized fairly soon after, however, that this postulate was not strictly necessary: the relativity principle alone was enough to show that the most general form of the velocity addition law was that derived by Einstein, with some undetermined but finite limiting speed that any object could attain. That this speed is that of light is, then, a wonderful fact, but not of essential importance to the theory. Some works that explore this issue are W. V. Ignatowsky, *Arch. Math. Phys.* **17**, 1 (1911); **18**, 17 (1911); V. Mitavalsky, *Am. J. Phys.* **34**, 825 (1966); Y. P. Terletskii (1968); A. R. Lee and T. M. Kalotas, *Am. J. Phys.* **43**, 434 (1975); N. D. Mermin, *Am. J. Phys.* **52**, 119 (1984); A. Sen, *Am. J. Phys.* **62**, 157 (1994).

relatively simple form:

Law	**Equation (Gaussian)**	**Equation (SI)**
Gauss's law	$\nabla \cdot \mathbf{E} = 4\pi\rho$	$\nabla \cdot \mathbf{E} = \dfrac{\rho}{\epsilon_0}$
Ampere-Maxwell law	$\nabla \times \mathbf{B} = \dfrac{4\pi}{c}\mathbf{j} + \dfrac{1}{c}\dfrac{\partial \mathbf{E}}{\partial t}$	$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \mu_0\epsilon_0 \dfrac{\partial \mathbf{E}}{\partial t}$
Faraday's law	$\nabla \times \mathbf{E} + \dfrac{1}{c}\dfrac{\partial \mathbf{B}}{\partial t} = 0$	$\nabla \times \mathbf{E} + \dfrac{\partial \mathbf{B}}{\partial t} = 0$
No magnetic monopoles	$\nabla \cdot \mathbf{B} = 0$	$\nabla \cdot \mathbf{B} = 0$
Lorentz force law	$\mathbf{F} = q\left(\mathbf{E} + \dfrac{1}{c}\mathbf{v} \times \mathbf{B}\right)$	$\mathbf{F} = q\,(\mathbf{E} + \mathbf{v} \times \mathbf{B})$

(2.1)

These laws are confirmed by extensive experience and the demands of consistency with general principles of symmetry and relativistic invariance, although their full content can be appreciated only after detailed study. We have written them in the two most widespread systems of units in use today and given them the names commonly used in the Western literature. The first four equations are also collectively known as the *Maxwell equations*, after James Clerk Maxwell, who discovered the last term on the right-hand side of the Ampere-Maxwell law in 1865 and thereby synthesized the, till then, separate subjects of electricity and magnetism into one.[3]

We assume that readers have at least some familiarity with these laws and are aware of some of their more basic consequences. A brief survey is still useful, however. We begin by discussing the symbols. The parameter c is the speed of light, and ϵ_0 and μ_0 are constant scale factors or conversion factors used in the SI system. The quantity ρ is a scalar field $\rho(\mathbf{r}, t)$, denoting the *charge distribution* or density. Likewise, $\mathbf{j}$ is a vector field $\mathbf{j}(\mathbf{r}, t)$, denoting the *current distribution*. This means that the total charge inside any closed region of space is the integral of $\rho(\mathbf{r}, t)$ over that space, and the current flowing across any surface is the integral of the normal component of $\mathbf{j}(\mathbf{r}, t)$ over the surface. This may seem a roundabout way of specifying the position and velocity of all the charges, which we know, after all, to be made of discrete objects such as electrons and protons.[4] But, it is in these terms that the equations for $\mathbf{E}$ and $\mathbf{B}$ are simplest. Further, in most macroscopic

[3] Although modern practice attaches the names of particular scientists to these laws, it should be remembered that they distill the collective efforts of several hundred individuals over the eighteenth and nineteenth centuries, if not more. A survey of the history may be found in E. M. Whittaker (1951). For a more modern history covering a more limited period, see O. Darrigol (2000).

[4] In fact, in dealing with discrete point charges, or idealized current loops of zero thickness, the distributions $\rho(\mathbf{r}, t)$ and $\mathbf{j}(\mathbf{r}, t)$ must be given in terms of the Dirac delta function. A certain amount of mathematical quick-stepping is then necessary, which we shall learn how to do in chapter 2.

situations, one does not know where each charge is and how fast it is moving, so that, at least in such situations, this description is the more natural one anyway.

The four Maxwell equations allow one to find **E** and **B** if ρ and **j** are known. For this reason, the terms involving ρ and **j** are sometimes known as *source terms*, and the **E** and **B** fields are said to be "due to" the charges and currents. However, we began by talking of the forces exerted by charges on one another, and of this there is no mention in the Maxwell equations. This deficiency is filled by the last law in our table—the Lorentz force law—which gives the rule for how the fields acts on charges. According to this law, the force on a particle with charge q at a point **r** and moving with a velocity **v** depends only on the instantaneous value of the fields at the point **r**, which makes it a local law. Along with Newton's second law,

$$\frac{d\mathbf{p}}{dt} = \mathbf{F}, \tag{2.2}$$

equating force to the rate of change of momentum,[5] it allows us to calculate, in principle, the complete motion of the charges.

Let us now discuss some of the more salient features of the equations written above. First, the Maxwell equations are linear in **E** and **B**, and in ρ and **j**. This leads immediately to the *superposition principle*. If one set of charges and currents produces fields $\mathbf{E}_1$ and $\mathbf{B}_1$, and another set produces fields $\mathbf{E}_2$ and $\mathbf{B}_2$, then if both sets of charges and currents are simultaneously present, the fields produced will be given by $\mathbf{E}_1 + \mathbf{E}_2$, and $\mathbf{B}_1 + \mathbf{B}_2$. This fact enables one to simplify the calculation of the fields in many circumstances. In principle, one need only know the fields produced by a single moving charge, and the fields due to any distribution may be obtained by addition. In practice, the problem of addition is often not easy, and one is better off trying to solve the differential equations directly.[6] A large part of electromagnetic theory is devoted to developing the classical mathematical machinery for this purpose. This includes the theorems named after Gauss, Stokes, and Green, and Fourier analysis and expansions in complete sets of orthogonal functions. With modern-day computers, direct numerical solution is the method of choice in many cases, but a sound grasp of the analytic techniques and concepts is essential if one is to make efficient use of computational resources.

The second point is that the equations respect the symmetries of nature. We discuss these in considerably greater detail in chapter 6, and here we only list the symmetries. The first of these is invariance with respect to space and time translations, i.e., the equivalence of two frames with different origins or zeros of time. As in mechanics, this symmetry is connected with the conservation of momentum and energy. The fact that it holds

[5] In the form (2.2) the equation remains relativistically correct. This is not so if we write $\mathbf{F} = m\mathbf{a}$, with m and **a** being the mass and the acceleration, respectively. The reason is that for particles with speeds close to c, $\mathbf{p} \neq m\mathbf{v}$.

[6] Supplemented, one might add, by boundary conditions. Note though, that not all boundary conditions that lead to a well-posed mathematical problem are physically sensible. The physically acceptable boundary conditions are that in static problems, the fields die off at infinity, and in dynamic problems, they represent outgoing solutions, i.e., that there be no flow of energy from infinity into the region of interest, unless such irradiation is specifically known to be present.

for Maxwell's equations automatically leads us to assign energy and momentum to the electromagnetic field itself. The second symmetry is rotational invariance, or the isotropy of space. That this holds can be seen directly from the vector nature of **E** and **B**, and the properties of the divergence and curl. It is connected with the conservation of angular momentum.[7] The third symmetry is spatial inversion, or parity, which in conjunction with rotations is the same as mirror symmetry.[8] We shall find that under inversion, $\mathbf{E} \to -\mathbf{E}$, in the same way that a "normal" vector like the velocity **v** behaves, but $\mathbf{B} \to \mathbf{B}$. One therefore says that **E** is a *polar vector*, or just a vector, while **B** is a *pseudovector* or *axial vector*. The fourth symmetry is time reversal, or what might be better called motion reversal. This is the symmetry that says that if one could make a motion picture of the world and run it backward, one would not be able to tell that it was running backward.[9] The fifth symmetry is the already mentioned equivalence of reference frames, also known as *relativistic invariance* or *Lorentz invariance*.[10] This symmetry is extremely special and, in contrast to the first three, is the essential way in which electromagnetism differs from Newtonian or pre-Einsteinian classical mechanics. We shall devote chapter 23 to its study. Historically, electromagnetism laid the seed for modern (Einsteinian) relativity. The problem was that the Maxwell equations are not Galilean invariant. This fact is mostly clearly seen by noting that light propagation, which is a consequence of the Maxwell equations, is described by a wave equation of the form

$$\nabla^2 f - \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2} = 0. \tag{2.3}$$

Here, f stands for any Cartesian component of **E** or **B**. As is well known, classical wave phenomena are *not* Galilean invariant. Sound, e.g., requires a material medium for its propagation, and the frame in which this medium is at rest is clearly special. The lack of Galilean invariance of Maxwell's equations was well known to physicists around the year 1900, but experimental support for the most commonly proposed cure, namely, that there was a special frame for light as well, and a special medium (the *ether*) filling empty space, through which light traveled, failed to materialize. Finally, in 1905, Einstein saw that Galilean invariance itself had to be given up. Although rooted in electromagnetism, this proposal has far-reaching consequences for all branches of physics. In mechanics, we mention the nonabsolute nature of time, the equivalence of mass and energy, and

[7] The connection between space translation invariance and the conservation of momentum, time translation invariance and the conservation of energy, and rotational invariance and the conservation of angular momentum is a general consequence of Noether's theorem, which states that any continuous symmetry leads to a conservation law and also gives the form of the conserved quantity. Noether's theorem is proved in almost all texts on mechanics. See, e.g., Jose and Saletan (1998), secs. 3.2.2., 9.2.

[8] The weak interactions do not respect this symmetry, but they lie outside the realm of classical physics. The same comment applies to time reversal.

[9] Anyone who has seen the Charlie Chaplin gag where he rises from his bed, stiff as a corpse, while his heels stay glued in one spot, will disagree with this statement. In fact, the laws of physics possess only *microscopic* reversibility. How one goes from this to *macroscopic* irreversibility and the second law of thermodynamics is a profound problem in statistical mechanics, and continues to be a matter of debate.

[10] This term is sometimes expanded to include the previous four symmetries also, and one then speaks of *full* or *general* Lorentz invariance.

the impossibility of the existence of rigid bodies and elementary particles with finite dimensions. Today, relativity is not regarded as a theory of a particular phenomenon but as a framework into which all of physics must fit. Much of particle physics in the twentieth century can be seen as an outcome of this idea in conjunction with quantum mechanics.

Another feature of the Maxwell equations that may be described as a symmetry is that they imply charge conservation. If we add the time derivative of the first equation, Gauss's law, to the divergence of the second, the Ampere-Maxwell law, we obtain the continuity equation for charge,

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \mathbf{j}. \tag{2.4}$$

If we integrate this equation over any closed region of space, and any finite interval of time, the left-hand side gives the net increase in charge inside the region, while, by Gauss's theorem, the right-hand side gives the inflow of charge through the surface bounding the region. Thus, eq. (2.4) states that charge is *locally* conserved. This conservation law is intimately connected with a symmetry known as *gauge invariance*. We shall say more about this in chapter 12.

The last symmetry to be discussed is a certain duality between **E** and **B**. Let us consider the second and third Maxwell equations and temporarily ignore the current source term. The equations would then transform into one another under the replacements $\mathbf{E} \to \mathbf{B}$, $\mathbf{B} \to -\mathbf{E}$. The same is true of the remaining pair of equations if the charge source term is ignored. This makes it natural to ask whether we should not modify the equations for $\nabla \cdot \mathbf{B}$ and $\nabla \times \mathbf{E}$ to include magnetic charge and current densities ρ_m and $\mathbf{j}_m$, in other words, to write (in the Gaussian system),

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} + \frac{4\pi}{c}\mathbf{j}_m, \tag{2.5}$$

$$\nabla \cdot \mathbf{B} = 4\pi\rho_m. \tag{2.6}$$

All the existing experimental evidence to date, however, indicates that free magnetic charges or monopoles do not exist.[11]

In the same connection, we should note that there is another source of magnetic field besides currents caused by charges in motion. All the charged elementary particles, the electron (and the other leptons, the muon and the taon) and the quarks, possess an intrinsic or spin magnetic moment. This moment cannot be understood as arising from a classical spinning charged object, however. The question then arises whether we should

[11] For extremely precise-minded readers, we should note that there is a certain convention implicit in the making of this statement. By adopting a larger set of duality transformations, one could, in fact, modify Maxwell's equations as per eqs. (2.5) and (2.6). Instead of asserting the absence of magnetic monopoles, one would then say that the ratio ρ_m/ρ_e (ρ_e being the electric charge) was the same for all known particles. There is little to be gained from this point of view, however, and it is simpler to pick a fixed representation for **E** and **B** and write Maxwell's equations in the usual form. See Jackson (1999), sec. 6.12, for more on this point.

not add a source term to the equation for $\nabla \cdot \mathbf{B}$ to take account of this magnetic moment. If we are interested only in describing the field classically, however, we can do equally well by thinking of these moments as idealized current loops of zero spatial extent and including this current in the source term proportional to $\mathbf{j}$ in the Ampere-Maxwell law. The integral of the divergence of this current over any finite volume is always zero, so the equation of continuity is unaffected, and we need never think of the charge distribution carried by these loops separately. In fact, the alternative of putting all or some of the source terms into the equation for $\nabla \cdot \mathbf{B}$ is not an option, for it leads to unacceptable properties for the vector potential. We discuss this point further in section 26.

3 A lightspeed survey of electromagnetic phenomena

Having surveyed the essential properties of the equations of electrodynamics, let us now mention some of the most prominent phenomena implied by them. First, let us consider a set of static charges. This is the subject of electrostatics. Then $\mathbf{j} = 0$, and $\rho(\mathbf{r})$ is time independent. The simplest solution is then to take $\mathbf{B} = 0$, and the $\mathbf{E}$-field, which is also time independent, is given by Gauss's law. In particular, we can find $\mathbf{E}(\mathbf{r})$ for a point charge, and then, in combination with the Lorentz force law, we obtain Coulomb's force law—namely, that the force between two charges is proportional to the product of the charges, to the inverse square of their separation, and acts along the line joining the charges. We study electrostatics further in chapter 3.

Similarly, suppose we have a time-independent current density $\mathbf{j}(\mathbf{r})$, and $\rho = 0$. (The current density must be divergenceless to have a well-posed problem, for otherwise the equation of continuity would be violated.) This makes up the subject of magnetostatics. The simplest solution now is $\mathbf{E} = 0$, and a time-independent $\mathbf{B}$, which is given by the Ampere-Maxwell equation (now known as just *Ampere's law*) and the equation $\nabla \cdot \mathbf{B} = 0$. There is now no analog of Coulomb's law, but several simple setups can be considered. One can, for example, calculate the $\mathbf{B}$ field produced by a straight infinite current-carrying wire. A second wire parallel to the first will experience a force which is given by the Lorentz force law. The force per unit length on any wire is proportional to the product of the currents, is inversely proportional to the distance between the wires, and lies in the plane of the wires, perpendicular to the wires themselves. This relationship is the basis of the definition of the unit of current in the SI system, the ampere. We study magnetostatics in detail in chapter 4.

The simplest time-dependent phenomena are described by Faraday's law. This law says that a changing magnetic field, which could be created in several ways—a time-dependent current $\mathbf{j}(\mathbf{r}, t)$, or a moving magnet—produces an electric field. If a metallic wire loop is placed in the region of the electric field, a current and an emf will be induced in the loop. This phenomenon, known as *induction*, is the basis of transformers, generators, and motors, and therefore of the unfathomable technological revolution wrought by these devices. We study this in chapter 5. A related phenomenon is seen when a wire loop, or, more generally, any extended conductor, moves in a static magnetic field. The induced electric field can then drive currents through the conductor. This effect is exploited in dynamos and is believed to lie behind the earth's magnetic field, as we shall see in section 131.

The term $\partial \mathbf{E}/\partial t$ in the Ampere-Maxwell law is needed to make the equations consistent with charge conservation. Its greatest consequence, however, is seen by considering the equations in the absence of any currents or charges. If we take the curl of the Faraday equation, for example, and use the Ampere-Maxwell and Coulomb's laws, we obtain

$$\nabla^2 \mathbf{E} - \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} = 0. \tag{3.1}$$

The same equation is obtained for $\mathbf{B}$ if we take the curl of the equation for $\nabla \times \mathbf{B}$. These two equations have nonzero solutions that are consistent with the first-order equations coupling together $\mathbf{E}$ and $\mathbf{B}$, and with $\nabla \cdot \mathbf{E} = \nabla \cdot \mathbf{B} = 0$. These solutions describe electromagnetic waves or light, and we study them in chapter 7, except for certain observer-dependent properties, such as the Doppler effect, which are covered in chapter 24. We have already commented on the implications of the existence of these solutions for the reality of the electromagnetic field.

Maxwell's equations also describe the production of electromagnetic waves via the phenomenon of radiation. We shall see this in chapters 9 and 10, when we consider the fields produced by moving charges. We shall see that an accelerating charge emits fields that die away only inversely with distance from the charge at large distances and that locally look like plane electromagnetic waves everywhere. These radiated fields carry energy and momentum. This phenomenon underlies radio, TV, cell phones, and all other wireless communications. If the charges are moving at speeds close to that of light, the properties of the radiation change dramatically. This is illustrated by the phenomenon of synchrotron radiation, which we study in chapter 25 after we have discussed special relativity in chapters 23 and 24.

The interaction of radiation or light with matter opens a whole new set of phenomena, which can be divided into large subclasses. First, when the matter is microscopic—individual charges, atoms, and molecules—interest attaches to scattering, i.e., the acceleration of the charges by the incident radiation, and reradiation of an electromagnetic field due to this acceleration. One now obtains the phenomena of Compton scattering, atomic and molecular spectra, etc. A proper treatment of these must be quantum mechanical. Nevertheless, much can be learned even in a classical approach, and we do this in chapter 22 using phenomenological models of atoms and molecules. Second, when the matter is in the form of a bulk medium, the most striking fact is that at certain wavelengths, light can propagate through matter, e.g., visible light goes through window glass. How this happens is examined in chapter 20. We also examine the attendant phenomena of reflection and refraction at interfaces between different media. If the medium is inhomogeneous, then, in addition to propagation, one also gets some scattering of the light. We see this phenomenon every day in the sky, and it also occurs when the medium is denser, such as a liquid. These topics are also discussed in chapter 22. Third, when the matter is in the form of opaque obstacles, large on the scale of the wavelength, application of the superposition principle to light fields leads to distinctive phenomena known as interference and diffraction. We take these up in chapter 8.

Next, let us turn to the behavior of charges in external fields. This is described by the Lorentz force equation. A large variety of motions is obtained, especially in inhomogeneous magnetic fields. We discuss these in chapter 11. Motion of charges in the earth's magnetic field is discussed in appendix G. The motion of magnetic moments in a magnetic field is also discussed in chapter 11.

We have already touched on the phenomena encountered when light interacts with bulk matter, without indicating how these are to be understood. For that, one must first tackle the larger problem of describing electromagnetic fields in matter more generally, not just for radiation fields. This is a very complex problem, as evidenced by the huge variety in the types of matter: conductors, insulators, magnets, and so on. Indeed, matter is itself held together largely by electromagnetic forces, and much of the distinction between the broadly different types of matter we have mentioned above is based on the response of these types to electromagnetic fields. Thus, it would seem that one first needs to develop a theory of matter, so that one may understand how some materials can be, say, conductors, and other materials insulators. Fortunately, one can make substantial progress by relying on intuitive and simplified notions of these terms. The key property that helps us is that matter is neutral on a very short distance scale, essentially a few atomic spacings. Thus, coarse-grained or *macroscopic* electromagnetic fields may be defined by spatially averaging over this length scale. These fields obey equations that resemble those for the fields in vacuum. The resemblance is only skin deep, however. The response of the medium cannot be trivialized. It is modeled through so-called constitutive relations that differ from medium to medium. In conductors, e.g., we have Ohm's law, which says that an internal electric field is accompanied by a proportionally large transport current. In insulators (also known as *dielectrics*), it relates the polarization of the matter to the internal electric field. It is in these constitutive relations that the complexity of the material is buried. Finding them from "first principles" is the province of condensed matter physics and statistical mechanics, which we shall not enter. Instead, we will work with semiempirical and phenomenological constitutive laws. Essentially all phenomena can be understood in this way. The coarse-graining procedure is discussed in chapter 13, and Ohm's law and the related topics of emf and electrical circuits in chapter 17. A simple but widely applicable constitutive model for time-dependent phenomena in many materials is developed in chapter 18.

The simplest kinds of phenomena involving matter are static. Electrostatic fields in the presence of conductors and insulators (or dielectrics) are discussed in chapters 14 and 15. In the first case, the central phenomenon is the expulsion of the electrical field from the interior of the conductors and from any hollow cavity inside a conductor. In the second case, the field is not expelled entirely, but is reduced, and the concern shifts to understanding why and estimating the reduction.

In contrast to the electric case, the response of most materials to static magnetic fields is rather tame. Permanent or ferromagnets are a notable exception. Unfortunately, the most interesting phenomena that they exhibit, such as hysteresis, domain formation, etc., are rather difficult to analyze or even formulate, since the particulars are dominated by material properties and even the shape of the body because of long-range dipole–dipole interactions. Still, several general aspects can be studied, and we do so in chapter 16.

When matter is subjected to time-dependent fields, even more phenomena emerge. In conductors, when the frequency is low, one gets eddy currents, and the expulsion of electric fields that was perfect in the static case is only slightly weakened. We discuss this in chapter 19. When the frequency is high, we get plasma oscillations and waves, as discussed in chapter 21. The near-perfect reflectivity of metals is also discussed in this chapter, along with the waveguides and resonant cavities that this property makes possible.

Needless to say, in attempting to understand such a vast array of phenomena, one must develop and draw upon many general concepts. These include conservation laws, relativistic invariance, thermodynamics and statistical mechanics, causality, stochasticity, the action principle, and the Lagrangian and Hamiltonian formulations of mechanics. The discussion of these concepts is woven into the entire text intimately, for it is in this way that we see the relation between electromagnetism and other branches of physics. The two exceptions are the action formalism, to which we devote an entire chapter, chapter 12, and the formalism of special relativity, which is covered in chapter 23. We shall, by and large, stay away from quantum mechanics, although a knowledge of some elementary quantum mechanical ideas is presumed in a few places.

4 SI versus Gaussian

Two common systems of units and dimensions are used today in electromagnetism. These are the Gaussian and the SI or rationalized MKSA systems. The Gaussian system is designed for use with the cgs (centimeter–gram–second) system of mechanical units, and the SI is designed for use with the MKS (meter–kilogram–second) system. Unfortunately, converting between Gaussian and SI is not as easy as converting between dynes and newtons, or even foot-pounds and newtons, as physical quantities do not even have the same engineering dimensions in the two systems. In the Gaussian system, **E** and **B** have the same dimensions, while in the SI system, **E** has dimensions of velocity times **B**. This means that equations intended for use in the two systems do not have the same form, and one must convert not only amounts but also equations. For example, in the SI system, the factor $1/c$ does not appear in Faraday's law or the Lorentz force law. Additional differences are present in the relations for macroscopic fields **D** and **H** that arise in the discussion of material bodies. Further, the SI system entails two dimensional constants, ϵ_0 and μ_0, known as the *permittivity* and *permeability* of the vacuum, respectively. The net result is that converting between the two systems is almost invariably irritating, but there seems to be little that can be done to bring about a standardization. A prominent twentieth-century magnetician captures the frustration perfectly:

> Devotees of the Giorgi system will not be happy with my units; but I can assure them that the unhappiness that my system inflicts upon them will be no greater than the unhappiness that their system over the last thirty years has inflicted on me.[12]

In this book, we shall give the most important formulas in both systems, but intermediate calculational steps will be given in Gaussian only. From the point of view of physics and conceptual understanding, the Gaussian system is better. From the point of view of practical application, on the other hand, the SI system is better, as it gives currents in amperes, voltages in volts, etc. Thus, a hard insistence on using one system or the other gets one nowhere, and it is necessary for everyone who works with electricity and magnetism to understand *both* systems and to have an efficient and reliable method of going back and forth between them. Conversion tables that achieve this end can be found in almost all textbooks.[13] We too give such tables (see tables 1.1 and 1.2, pages 16 and 17). However, we also show how to derive these conversion factors. For now, we limit ourselves to the basic quantities $\mathbf{E}$, $\mathbf{B}$, charge, etc. Relationships for the macroscopic fields $\mathbf{D}$, $\mathbf{H}$, etc., and related quantities are discussed in chapter 13.

The scheme given here requires knowing (i) that the symbols for all mechanical quantities—mass, length, time, force, energy, power, etc.—are the same in the two systems, and (ii) formulas for three mechanical quantities in both systems. Other choices for this set of three are possible, but the one that we find most easy to remember is tabulated below.

Quantity	**Formula (Gaussian)**	**Formula (SI)**	
Coulomb force	$\dfrac{q^2}{r^2}$	$\dfrac{q^2}{4\pi\epsilon_0 r^2}$	
Energy density	$\dfrac{1}{8\pi}(\mathbf{E}^2+\mathbf{B}^2)$	$\dfrac{1}{2}(\epsilon_0\mathbf{E}^2+\mu_0^{-1}\mathbf{B}^2)$	(4.1)
Lorentz force	$q \quad \mathbf{E}+\dfrac{1}{c}\mathbf{v}\times\mathbf{B}$	$q(\mathbf{E}+\mathbf{v}\times\mathbf{B})$	

The first formula is for the Coulomb force between two equal charges q separated by a distance r. Since the symbols for force and distance are the same, it follows that

$$q_{\text{SI}} = \sqrt{4\pi\epsilon_0}\, q_{\text{Gau}}, \tag{4.2}$$

[12] Brown (1966). By the Giorgi system, Brown means SI. He himself uses a mixed system he calls Gaussian mks, which allows for conversion between SI and Gaussian at the expense of introducing a multiplicative factor in Coulomb's law whose value is different depending on the unit system, and replacement rules for the current and emf.

[13] See, e.g., Jackson (1999), appendix, or Pugh and Pugh (1970), chap. 1.

where the suffix "Gau" is short for Gaussian. The same conversion applies to charge density ρ, current I, and current density $\mathbf{j}$. (Recall that current is the amount of charge flowing through a surface per unit time.)

The second formula is for the energy density in the electromagnetic field. Since we can vary $\mathbf{E}$ and $\mathbf{B}$ independently, this is a "twofer"—it gives us two conversions for the price of one. Since energy and volume are the same in the two systems, we see that

$$\mathbf{E}_{\mathrm{SI}} = \frac{\mathbf{E}_{\mathrm{Gau}}}{\sqrt{4\pi\epsilon_0}}, \tag{4.3}$$

$$\mathbf{B}_{\mathrm{SI}} = \sqrt{\frac{\mu_0}{4\pi}}\mathbf{B}_{\mathrm{Gau}}. \tag{4.4}$$

The Lorentz force formula is the third one. It too is a "twofer." Consider a situation in which there is only an electric field. Since the symbol for force is unchanged in going from one system to the other, so must be the product qE:

$$(q\mathbf{E})_{\mathrm{SI}} = (q\mathbf{E})_{\mathrm{Gau}}. \tag{4.5}$$

But, this is exactly what we get from eqs. (4.2) and (4.3), so we already knew this. Something new is learned when we apply the same reasoning to the magnetic field term. We get

$$(q\mathbf{B})_{\mathrm{SI}} = \frac{1}{c}(q\mathbf{B})_{\mathrm{Gau}}. \tag{4.6}$$

Changing q and B to the Gaussian system using eqs. (4.2) and (4.4), we get

$$\sqrt{\epsilon_0\mu_0} = \frac{1}{c}. \tag{4.7}$$

These relationships are enough to convert any formula in the SI system to the Gaussian, or vice versa. Take, for example, the magnetic field of an infinite current-carrying wire. In the Gaussian system this is given by the formula

$$B = \frac{2I}{cr_\perp}, \tag{4.8}$$

where $r_\perp$ is the distance from the wire to the point where the field is desired. To get the SI formula, we replace B by $(4\pi/\mu_0)^{1/2}B$ and I by $(4\pi\epsilon_0)^{-1/2}I$. This yields

$$\begin{aligned} B &= \frac{1}{4\pi}\sqrt{\frac{\mu_0}{\epsilon_0}}\frac{2I}{cr_\perp} \\ &= \frac{\mu_0 I}{2\pi r_\perp}, \end{aligned} \tag{4.9}$$

where we have used eq. (4.7) to eliminate c.

As another example, let us take the formula for the power radiated by an electric dipole oscillator. In the SI system, this is

$$P = \frac{c^2 Z_0 k^4}{12\pi}|\mathbf{d}|^2; \quad Z_0 = \sqrt{\frac{\mu_0}{\epsilon_0}}. \tag{4.10}$$

Here k is the wave number of the radiation, and $\mathbf{d}$ is the dipole moment. Since the dipole moment for a charge distribution is the volume integral of $\mathbf{r}\rho(\mathbf{r})$, its conversion is the same as that for charge. The quantities P and k are evidently unchanged, so the Gaussian formula is

$$P = c^2\sqrt{\frac{\mu_0}{\epsilon_0}}4\pi\epsilon_0\frac{k^4}{12\pi}|\mathbf{d}|^2 = \frac{ck^4}{3}|\mathbf{d}|^2. \tag{4.11}$$

One check that this is correct is that it is free of ϵ_0 and μ_0.

As the third example, let us change the Ampere-Maxwell law from its SI to the Gaussian form. In the SI system, the law reads

$$\nabla\times\mathbf{B} = \mu_0\mathbf{j} + \mu_0\epsilon_0\frac{\partial\mathbf{E}}{\partial t}. \tag{4.12}$$

Using eqs. (4.2)–(4.4), we see that the Gaussian system form is

$$\sqrt{\frac{\mu_0}{4\pi}}\nabla\times\mathbf{B} = \mu_0\ \ \overline{4\pi\epsilon_0}\mathbf{j} + \frac{\mu_0\epsilon_0}{\sqrt{4\pi\epsilon_0}}\frac{\partial\mathbf{E}}{\partial t}, \tag{4.13}$$

or, dividing by $\sqrt{\mu_0/4\pi}$ and using eq. (4.7),

$$\nabla\times\mathbf{B} = \frac{4\pi}{c}\mathbf{j} + \frac{1}{c}\frac{\partial\mathbf{E}}{\partial t}, \tag{4.14}$$

as given in the table on page 3. The reader should carry out the same exercise for the remaining Maxwell equations.

The rules for converting capacitance, inductance, and conductance, and related quantities such as resistance and impedance will be found later when these quantities are defined.

Finally, let us see how to carry out ordinary or "engineering" dimensional analysis. We will denote the dimensions of a quantity by putting square brackets around it: [$\mathbf{E}$] will denote the dimensions of $\mathbf{E}$, and so on.

In the Gaussian system, all quantities have dimensions that can be expressed in terms of M, L, and T, the dimensions of mass, length, and time. However, these quantities often have to be raised to fractional exponents. Let us see how this happens, starting with charge. From the Coulomb force formula, we have $[q] = [FL^2]^{1/2}$, and since $[F] = MLT^{-2}$, we have

$$[q] = M^{1/2}L^{3/2}T^{-1}. \tag{4.15}$$

The dimensions of $\mathbf{E}$ now follow from a formula such as $E = q/r^2$ for the electric field magnitude due to a point charge. We get

$$[\mathbf{E}] = M^{1/2}L^{-1/2}T^{-1}. \tag{4.16}$$

As a check, we examine the dimensions of the product qE:

$$[qE] = MLT^{-2}, \tag{4.17}$$

which are the same as those of force, as they should be. Similarly, E^2 has dimensions of $ML^{-1}T^{-2}$, which are the same as those of energy density [Energy (ML^2T^{-2})/Volume (L^3)].

In the Gaussian system, the dimensions of **B** and **E** are the same. This can be seen either from the expression for the energy density, or the Lorentz force law. Thus,

$$[\mathbf{B}] = M^{1/2}L^{-1/2}T^{-1}. \tag{4.18}$$

Exercise 4.1 Obtain the dimensions of I, ρ, and $\mathbf{d}$ in the Gaussian system, and verify the dimensional correctness of all formulas given in this chapter in the Gaussian system.

In the SI system, fractional exponents are avoided by including current (I) as a fourth basic unit. The dimensions of all electromagnetic quantities, including the constants ϵ_0 and μ_0, are given in terms of M, L, T, and I.

As the starting point, we again consider the dimensions of charge. This is now very simple. By the definition of current, we have

$$[q] = TI. \tag{4.19}$$

The Lorentz force formula now gives us [**E**] and [**B**]:

$$[\mathbf{E}] = MLT^{-3}I^{-1}, \tag{4.20}$$

$$[\mathbf{B}] = MT^{-2}I^{-1}. \tag{4.21}$$

Note that **E** and **B** do not have the same dimensions in SI; rather **E** has dimensions of velocity times **B**, as already stated.

The dimensions of ϵ_0 and μ_0 can now be obtained from the formula for energy density:

$$[\epsilon_0] = M^{-1}L^{-3}T^4I^2, \tag{4.22}$$

$$[\mu_0] = MLT^{-2}I^{-2}. \tag{4.23}$$

Exercise 4.2 Verify the dimensional correctness of all formulas given in this chapter in the SI system.

Exercise 4.3 A famous text in quantum mechanics states that "in atomic units," the probability per unit time of ionization of a hydrogen atom in its ground state in an external electric field $\mathcal{E}$ (also in atomic units) is given by

$$w = \frac{4}{\mathcal{E}}e^{-(2/3\mathcal{E})}. \tag{4.24}$$

Atomic units are such that $\hbar$, m (electron mass), and a_0 (Bohr radius) all have the value 1. Rewrite the above formula in the Gaussian and SI systems, and find the value of the ionization rate for a field of 10^{10} V/m.

Answer: 10^4 sec^{-1}.

We conclude with a brief history of the two systems of units. Knowing this helps in keeping an open mind about the benefits of one versus the other. In the early 1800s, with the cgs system for mechanical quantities (force, energy, mass, etc.), Coulomb's law provided the natural unit of charge. Likewise, the law for the force between two magnetic poles gave the unit of magnetic pole strength.[14] With Oersted's discovery that currents also produce magnetic fields, and the precise formulation of this discovery via the Biot-Savart law, current could be defined in terms of the magnetic pole strength. All other quantities, such as capacitance, resistance, magnetic flux, etc., could also be connected to the pole strength. This led to the so-called electromagnetic or cgs-emu units. However, current is also the rate of charge flow, so the magnetic field and all other electromagnetic quantities could be related to the unit of charge. This led to the electrostatic or cgs-esu units. It was then noticed by many workers that the ratio of the numerical value of any quantity in cgs-emu units to that in cgs-esu units was very close to 3×10^{10}, or its reciprocal, or the square of one of these numbers,[15] and that this number coincided with the speed of light in cgs units. Gauss saw that by putting a quantity with dimensions of velocity in the denominator of the Biot-Savart law, the cgs-emu and cgs-esu systems could be replaced by a single system; this is how the Gaussian system came to be. Further, this appearance of the speed of light was a key factor behind Maxwell's proposal of the displacement current in 1865, and the idea that light was an electromagnetic wave. All this while, many workers had adopted a "practical" system of units based on the same idea as the cgs-emu, but with units of length and mass equal to 10^9 cm and 10^{-11} g, respectively. In 1901, Giorgi adjusted the constants μ_0 and ϵ_0 to make the practical units compatible with the mks system; this is essentially the SI system in use today. It may surprise some readers that the joule and the newton were not created until the 1930s!

[14] Since magnetic monopoles do not exist, this sentence requires some explanation. The only sources of magnetism known in the eighteenth century were permanent magnets. It was well known that the poles of a magnet could not be separated and that breaking a bar magnet produced new poles at the broken ends. By careful torsion balance experiments with long magnetic needles, however, Coulomb was able to establish in 1785 that they behaved as if there were a force between the poles at the ends of the needles that varied as the inverse square of the separation and the product of the pole strengths.

[15] For example, the ratio for charge was measured by Weber and Kohlrausch to be 3.107×10^{10} in 1856; that for resistance was found to be $(2.842 \times 10^{10})^2$ by Maxwell in 1868, and $(2.808 \times 10^{10})^2$ by W. Thomson in 1869.