1

PHYSICAL & MATHEMATICAL FUNDAMENTALS

Introduction. Since "the world is unitary"—with each part connected (however tenuously) with each other part—it is of some philosophical interest that physics admits of semi-conventional division into semi-autonomous "branches." Most of those branches are concerned with the analysis of fairly *general classes* of physical systems (think, for example, of classical mechanics, or of quantum mechanics, fluid dynamics, thermodynamics), but a few (celestial mechanics, general relativity, ...) are concerned with relatively particularized systems. It is useful to note that electrodynamics is, for all of its incredible richness and variety, a subject of the latter sort: *all that follows will be motivated by a desire to clarify the structure and dynamical properties of a single physical object—the electromagnetic field*.

Our objective, therefore, is to review progress in a field which achieved a kind of maturity (Maxwell, 1873) just about a century ago, a field to which some of the greatest minds (Einstein, Feynman ... and many others) have contributed, a field in which "the last word" has certainly *not* been written. Much of great value can be learned from close study of the (ongoing) *history of electrodynamics* ... but for that I must refer my readers to the relevant literature. A standard source is E. T. Whittaker's *A History of the Theories of Aether & Electricity* (1910). Since this branch of the history of science is currently quite active, it would be well to consult recent issues of (say) *History of Science*. For a good modern account of the "ancient history" of some of the basic notions see Duane & D. H. D. Roller, "The development

of the concept of electric charge: electricity from the Greeks to Coulomb" in J. B. Conant (editor), *Harvard Case Histories in Experimental Science* (Volume II, 1966). You should, of course, take this occasion to become acquainted with the Victorian founding fathers (Faraday, Maxwell) of our subject. I urge you therefore to look into *Michael Faraday: A Biography* by L. P. Williams (1965), *James Clerk Maxwell: Physicist and Natural Philosopher* by C. W. F. Everitt (1975) and/or *Contributions of Faraday & Maxwell to Electrical Science* (1966) ... all of which are informative, yet fun to read. Finally, every student of electrodynamics should peruse the pages of Maxwell's own *A Treatise on Electricity & Magnetism* the (posthumous) 3rd edition (1891) of which was reissued by Dover in 1954. While the history of science is its own reward, the history of electrodynamics (as of classical mechanics, quantum mechanics ...) is also of directly utilitarian value, for it illuminates the processes/circumstances/developments which contribute to the maturation of physics—to the discovery/invention of *new* physics.

That electromagnetic phenomenology (and theoretical understanding of that phenomenology) lies at the base of an elaborate technology—think of electrical power grids, the electric light, motorized devices, electronic communication/computation/mealsurement & control ... none of which were known to the founders of the field—is of course not news. Less well known to the general public are the theoretical contributions of classical electrodynamics, which (directly or indirectly) has stimulated the invention/development of

- special relativity
- quantum mechanics
- the modern theory of gravitation (general relativity)
- elementary particle physics
- many of the methods characteristic of modern applied mathematics

... and much else. One could perfectly well base a course such as this on the *technological applications* of our subject: such an approach would be considered standard in schools of engineering, and is reflected in the design of many existing texts. I prefer, however, to let (my personal view of) the theoretical applications/ramifications of electrodynamics govern the selection, arrangement and presentation of the subject matter. Classical electrodynamics provides a unique "classical window" through which can be glimpsed many of the principles which are now recognized to dominate the structure of the micro-world (also the very-large-scale macro-world ... and much that lies in between). But to gain access to that window we must pay close and critical attention to *structural issues* ... and to that end we must from time to time draw upon mathematical methods which, though of growing importance, have heretofore not been considered standard to the undergraduate education of physicists. The latter material will be developed in appropriate detail as needed.

The "historical approach" (recapitulated psuedo-history) which for a long time dominated instruction in classical and—particularly—quantum mechanics has never been popular in the electrodynamical classroom ... and it is certainly

not my intention to attempt such an experiment. Nor shall I honor the established practice, which is to proceed "crabwise" into the subject, for a pedagogical strategy which places the (allegedly) "easy parts" (electrostatics, potential theory, ...) first necessarily *displaces the fundamentals* ... with the result that Maxwell's equations tend to get lost in the clutter, and relativity to enter (as historically it did) only as an afterthought.

The design of this introductory chapter proceeds therefore from my desire "to put first things first." My goal, more specifically, is to proceed in all feasible haste to a working understanding—however tentative—of *what kind of a thing electrodynamics is, of the physical and computational issues fundamental to the subject*. This will entail review of material to which you have already had some exposure—the

- conceptual innovations and
- physical phenomenology

which historically led James Clerk Maxwell to the *equations of motion* of the electromagnetic field. But we will also begin what will, as we proceed, become a major activity—"looking under rocks:" conceptual rocks, computational rocks, formal rocks. Our intent at this stage is more *to formulate sharp questions* that to formulate sharp answers (the latter can wait). It is interesting to observe that we will be led, even in this introductory survey, to aspects (both deep and numerous) of electrodynamics of which Maxwell died (5 November 1879, at age 48) unaware.

1. Coulomb's law. The phenomenology here goes back to antiquity, and involves the curious behavior of macroscopic samples of certain biogenic substances (amber, fur, silk, paper, pithballs) which are—except for our story—insignificant constituents of the universe. This speculative tradition (to which an allusion survives in the word "electron," from $\eta\lambda\epsilon\kappa\tau\rho o\nu$ = amber) had by ∼1750—owing largely to the work of Benjamin Franklin (1706–1790)—led to the formulation of a recognizable precorsor of the modern concept of *electric charge*. It is today recognized that *electric charge is—like mass—an attribute not merely of bulk matter (pithballs) but of the elementary constituents of such matter*.

Particles announce their charge by exerting forces (forces of a specific yet-to-be-described structural type: "electromagnetic forces") on each other; *i.e.*, by *interacting—electromagnetically* ... and it is from study of how particles respond to such (postulated) forces that we have learned all that we know concerning the existence and properties of the electromagnetic field. The question—the *experimental* question—therefore arises: How are we to make structurally and quantitatively precise the *force law* latent in the preceding remarks?

Prior to ∼1760 (when this question first moved to centerstage) the only "universal force law" known to physics was Newton's

$$F = G\frac{Mm}{r^2}$$

which describes the *instantaneous gravitational interaction-at-a-distance* of mass points M and m. It was widely anticipated that the electrostatic interaction of charged mass points would turn out to be governed by a law of similar form. Experimental evidence in support of this conjecture was published by Daniel Bernoulli in 1760 and by Joseph Priestly in 1767, but the issue was instrumentally delicate, and was definatively resolved only in 1785 by Charles Coulomb (1736–1806), who used sensitive torsion balances and torsion pendula of his own invention (similar to those used years later by Henry Cavendish to measure G). Turning now to the concrete particulars ...

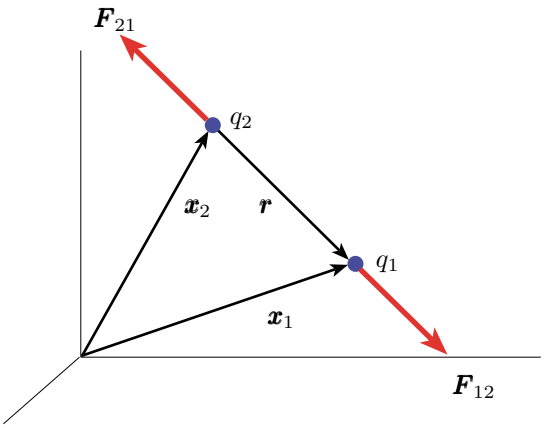


FIGURE 1: *Notation used to describe the relation of one charge to another, and the Coulombic forces which each exerts upon the other.*

Let $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ describe (relative to an inertial Cartesian frame) the positions of a pair of structureless but electrically charged mass points, (m_1, q_1) and (m_2, q_2). For conceptual convenience (*i.e.*, to circumvent the troublesome action-at-a-distance problem) we *assume the point charges to be at rest* ... both now and in the "relevant past." Experimentally

$$\boldsymbol{F}_{12} = k\frac{q_1 q_2}{r^2}\hat{\boldsymbol{r}} = -\boldsymbol{F}_{21} \tag{1}$$

where (see Figure 1) $\boldsymbol{F}_{12}$ is the force exerted on charge #1 by charge #2, and where

$$\begin{aligned}
&\boldsymbol{r} \equiv \boldsymbol{r}_{12} \equiv \boldsymbol{x}_1 - \boldsymbol{x}_2 = -(\boldsymbol{x}_2 - \boldsymbol{x}_1) \equiv -\boldsymbol{r}_{21} \quad : \quad \text{separation vector } 1 \leftarrow 2 \\
&r^2 \equiv \boldsymbol{r}\cdot\boldsymbol{r} \quad : \quad \text{squared separation} \\
&\hat{\boldsymbol{r}} \equiv \boldsymbol{r}/r \quad : \quad \text{unit separation vector } 1 \leftarrow 2
\end{aligned}$$

The *gravitational analog of Coulomb's law* (1) reads

$$\boldsymbol{F}_{12} = -G\frac{m_1 m_2}{r^2}\hat{\boldsymbol{r}} = -\boldsymbol{F}_{21} \tag{2}$$

These familiar results invite several (interconnected) lines of commentary:

1. In the context provided by the 2-body problem we lack the "principle of choice" required to assign $\boldsymbol{F}_{12}$ any direction other than that provided by the "line of centers" (separation vector). The argument would, however, lose its force if

- the point particles were endowed with directionality (spin);
- the forces of interation among n bodies ($n \geqslant 3$) were assembled more complexly than by mere *vectorial superposition*

$$\boldsymbol{F} = \boldsymbol{F}_{12} + \boldsymbol{F}_{13} + \cdots + \boldsymbol{F}_{1n}$$

Happily, complexities of the later type—but not the former!—are *absent from the electrodynamics of point charges* (though they can be expected to arise when one considers the interaction of—say—charged conductors: why?).

2. Comparison of (2) with (1) suggests that we sould construe m_1 and m_2 to be measures not of intertial mass but of "gravitational charge." It is, however, a deep-reaching and surprising fact that to do so would be to make a "distinction without a difference." For measurements of high precision (performed by Baron L. von Eötvös (1848–1919) and refined by R. H. Dicke) have established (to within about one part in 10^{12}) that

$$\frac{\text{gravitational charge}}{\text{inertial mass}} = \text{universal constant}$$

where "universal" here means "the same for all samples and states of matter." This fact was foreshadowed already in Galileo's alleged "Leaning Tower of Pisa experiment" and in the m-independence of the frequency $\omega = \sqrt{g/\ell}$ of a pendulum (why?). By appropriate adjustment of certain conventions (units) we can arrange that

$$= 1 \pm 10^{-12}$$

Such invariable quantitative identity becomes intelligible only if it proceeds from conceptual identity: "inertial mass" and "gravitational charge" must be *different names for the same thing*. This fundamental insight is built into the structure of (2), and entails that (relative to any prescribed system of mechanical units: cgs, MKS, ...) G *becomes an empirical constant of forced/fixed dimensionality*. Actually

$$G = 6.6732 \times 10^{-8}\ \text{dyn} \cdot \text{cm}^2/\text{g}^2$$

For further discussion of the Eötvös–Dicke experiments and their significance see C. W. Misner, K. S. Thorn & J. A. Wheeler, *Gravitation* (1973), pages 1050–1055.

3. Returning in this light to (1) the question arises: Is the value of electric charge/mass invariable, the same for all charged bodies? The answer is an emphatic "no:"

- for macroscopic bodies q/m is highly variable;
- among elementary particles of a given species q/m is constant, but
- the numerical value of q/m varies from species to species.

So in the real (multi-species) world there is *no electrical analog of the so-called "equivalence principle."* This circumstance entails that we must confront the question: What shall be the physical dimensionality $[Q]$ of electric charge? The answer is semi-*conventional* (there is dimensional trade-off between k and q^2), and presents some useful options:

- We might construe k (whatever its *numerical* value) to be *dimensionless*. Such a convention forces

$$[Q^2] = \text{force} \cdot \text{length}^2 = \text{energy} \cdot \text{length} = \text{action} \cdot \text{velocity}$$

  whence

$$[Q] = M^{\frac{1}{2}} L^{\frac{3}{2}} T^{-1} \quad : \quad \text{a "derived quantity"}$$

  If we set $k = 1$ and adopt cgs mechanical units we are led thus to the "electrostatic unit" (esu or "statcoulomb") of charge

$$\text{esu} = \text{g}^{\frac{1}{2}} \text{cm}^{\frac{3}{2}} \text{sec}^{\ -1}$$

  Evidently charges of one esu, if separated by one centimeter, exert upon each other a force on one dyne. The "rationalized" (or Heaviside–Lorentz) esu arises if—to avoid factors of 4π in the field equations—we set $k = 1/4\pi$.

- Since charges/currents/potentials are most conveniently measured by operations/procedures which are alien to mechanics, we might construe charge to be *dimensionally antonomous*: $(M, L, T) \to (M, L, T, Q)$. Such a convention forces

$$[k] = \text{force} \cdot \text{length}^2/\text{charge}^2 = ML^3T^{-2}Q^{-2}$$

  and causes the numerical value of k to assume (like G) the status of an emperical constant. If—following in the footsteps of Faraday—we adopt an electrochemical procedure to define the

$$\text{ampere} \equiv \text{coulomb/second}$$

  then we find by measurement that

$$k = 8.988 \times 10^9 \text{ newton} \cdot \text{meter}^2/\text{coulomb}^2$$

  Circumstances will emerge which make it natural to write

$$= \frac{1}{4\pi\epsilon_0}$$

  and to call ϵ_0 $(= 8.854 \times 10^{-12}\ \text{C}^2/\text{N}\cdot\text{m}^2)$ the "permittivity of free space."

Theoretical physicists tend generally to prefer (rationalized) electrostatic units, and engineers to prefer (rationalized MKS) or "practical" units. Interconversion formulæ follow from

$$\text{coulomb} = 2.997930 \times 10^{9} \text{ esu}$$

From the fact that the

$$\begin{aligned} \text{electronic charge } e &= 4.803250 \times 10^{-10} \text{ esu} \\ &= 1.602189 \times 10^{-19} \text{ coulomb} \end{aligned}$$

we see that the coulomb (also for that matter the esu) is, for most fundamental purposes, an impractably large unit. Often it is preferable to measure charge in multiples of e (as is standardly done in elementary particle physics, nuclear physics, chemistry). For further informatrion concerning the notorious (and—in theoretical contexts—usually irrelevant) "problem of units" see J. D. Jackson, *Classical Electrodynamics* (1962), pages 611–621.[1]

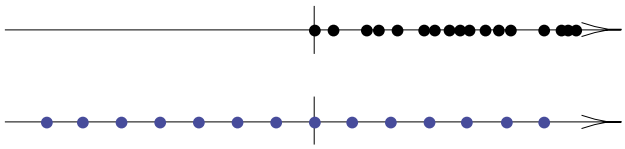


FIGURE 2: *The masses encountered in Nature are shown above, the electric charges below: the former are invariably positive, but are otherwise unconstrained; charges, on the other hand, can occur with either sign, and are always multiples of a fundamental unit.*

4. Gravitational forces are invariably attractive, while charged particles repell or attract each other according as their charges are of the same or opposite sign. These familiar facts trace, *via* the structure of (1) and (2), to the observation that *gravitational charge is invariably positive* while *electric charge can be of either sign*. The situation becomes somewhat more interesting when phrased in the language of elementary particle physics, for in that context the inergial mass concept is somewhat enlarged . . . and an interesting "graininess" reveals itself. One has

$$m \geqslant 0 \quad \text{but} \quad q \gtrless 0$$

as illustrated in Figure 2. Note that $m \geqslant 0$ applies (according to recent experiments) even to *anti*particles. And while "massless particles" exist (photon, graviton, neutrino?), *there are no charged massless particles*: "charge endows mass" (though not *all* mass arises—as was once supposed—by this complex mechanism).

[1] PROBLEMS 1, 2 & 3

5. In Coulomb's law $F = kQq/r^2$ the "2" is, of course, an *experimental* number. *How accurately* can it be said that electrostatic forces (or, for that matter, gravitational forces) "fall off as the square" of distance? If we write

$$F = kQq\frac{1}{r^{2+\epsilon}}$$

then Coulomb himself knew that $0 < \epsilon < 10^{-1}$. Cavendish (in some unpublished work) showed that $\epsilon < 3 \times 10^{-2}$ and Maxwell, by a refinement of Cavendish's technique, showed (1873) that $\epsilon < 5 \times 10^{-5}$. The most recent work known to me (E. R. Williams, 1971) establishes that $\epsilon < 6 \times 10^{-16}$. Interestingly, the quantum mechanical version of our subject (QED) shows that we can expect to have $\epsilon = 0$ *if the photon mass* μ *is precisely zero* ... and enables one to convert the sharpest of the results quoted above into the statement that

$$\mu \leqslant 2 \times 10^{-40}\ \text{g} \approx (\text{electron mass}) \cdot 10^{-20}$$

For a beautiful discussion of this absorbing topic see A. S. Goldhaber & M. M. Nieto, "Terrestrial and extraterrestrial limits on the photon mass," Rev. Mod. Phys. **43**, 277 (1971).[2] Note finally that the (massless) photon, though it "mediates the electromagnetic interaction of electrically charged particles," is itself *un*charged ... and moves always "with the speed of light" only because it is massless. I am, however, ahead of my story.

To describe the force $\boldsymbol{F}(\boldsymbol{x})$ experienced by a charge q if situated at a point $\boldsymbol{x}$ in a region of space occupied (see Figure 3) by a static *population* of charges $\{Q_1, Q_2, \dots\}$ we invoke—but only because it is sanctioned by experience—the *principle of superposition* to write

$$\begin{aligned}\boldsymbol{F}(\boldsymbol{x}) = \sum_i \boldsymbol{F}_i(\boldsymbol{x}) &= \sum_i kq\, Q_i \frac{1}{r_i^2}\hat{\boldsymbol{r}}_i \quad \text{with} \quad \boldsymbol{r}_i \equiv \boldsymbol{x} - \boldsymbol{x}_i \\ &= kq \underbrace{\sum_i Q_i \frac{1}{r_i^3}\boldsymbol{r}_i}_{\text{defines the } \textit{electrostatic field } \boldsymbol{E}(\boldsymbol{x}) \text{ which is established at } \boldsymbol{x} \text{ by the charges } Q_i}\end{aligned}$$

The $\boldsymbol{E}$-field is a *force field*, which in electrostatic units (k dimensionless) has the dimensionality

$$[\boldsymbol{E}] = \text{force/charge}$$

[2] While writing this paragraph I chanced (one midnight at the watercooler) to discuss its substance with Richard Crandall, with consequences that can be read about in R. E. Crandall, "Photon mass experiment," AJP **51**, 698 (1983) and R. E. Crandall & N. A. Wheeler, "Klein-Gordon radio and the problem of photon mass," Nuovo Cimento **84B**, 231 (1984): also the splendid thesis of Richard Leavitt, "A photon mass experiment: an experimental verification of Gauss' law" (1983)—on the basis of which Leavitt became Reed's first Apker Award finalist.

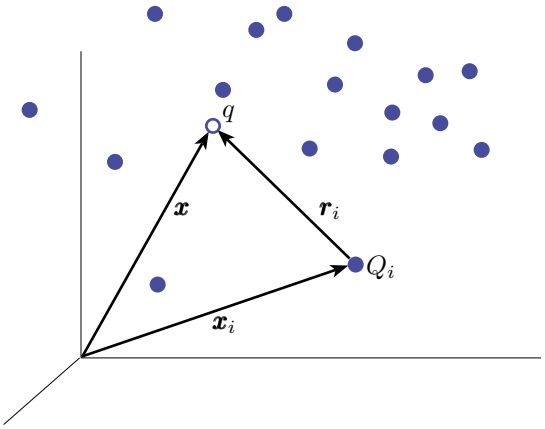


FIGURE 3: *A discrete population of charges acts electrostatically on a test charge*

and which is defined operationally by the dynamical response of the "test charge" (m, q).[3] Mathematically, $\boldsymbol{E}(\boldsymbol{x})$ is a *vector-valued function of position* (which is to say: a "vector field"), given explicitly by

$$\boldsymbol{E}(\boldsymbol{x}) = \sum_i Q_i \, \frac{\boldsymbol{x} - \boldsymbol{x}_i}{|\boldsymbol{x} - \boldsymbol{x}_i|^3} \tag{3.1}$$

Passing to the *continuous limit* we obtain

$$= \iiint \rho(\boldsymbol{\xi}) \, \frac{\boldsymbol{x} - \boldsymbol{\xi}}{|\boldsymbol{x} - \boldsymbol{\xi}|^3} \, d^3\xi \tag{3.2}$$

$\rho(\boldsymbol{\xi})$ is the *charge density* at $\boldsymbol{\xi}$

which gives back (3.1) in the special case

$$\rho(\boldsymbol{\xi}) = \sum_i Q_i \, \delta(\boldsymbol{\xi} - \boldsymbol{x}_i) \tag{4}$$

Though the rich *physics of electrostatic fields* is—in its entirety—latent in (3.2), that equation is susceptible to the criticism that

i) it is, in most contexts, not *computationally* useful

ii) it tells us nothing about the *general structural properties* of $\boldsymbol{E}$-fields.

Thus are we motivated to ask: *What are the differential equations which, in general, constrain/govern/describe the structure of (static) $\boldsymbol{E}$-fields?* That question motivates the following

[3] If the "field sources" Q_i were constrained merely to reside on some prescribed conductors then the presence of the test charge would cause them to rearrange themselves. This effect is minimized by assuming q to become arbitrarily small, though we are in fact constrained by Nature to have $q \geqslant e$ (or at least $q \geqslant \frac{1}{3}e$).

MATHEMATICAL DIGRESSION

For transformation-theoretic reasons which we shall be at pains later to clarify, the differential operators available to us are all latent in the vector-valued "del" operator

$$\boldsymbol{\nabla} \equiv \boldsymbol{i}\frac{\partial}{\partial x} + \boldsymbol{j}\frac{\partial}{\partial y} + \boldsymbol{k}\frac{\partial}{\partial z} \equiv \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix}$$

Readers will (I presume) recall that

1) $\boldsymbol{\nabla}$ applied to a scalar field $\varphi(\boldsymbol{x})$ yields the vector-valued *gradient* of φ: $\boldsymbol{\nabla}\varphi \equiv \operatorname{grad}\varphi$,
2) $\boldsymbol{\nabla}$ dotted into a vector field $\boldsymbol{A}(\boldsymbol{x})$ yields the scalar-valued *divergence* of $\boldsymbol{A}$: $\boldsymbol{\nabla}\cdot\boldsymbol{A} \equiv \operatorname{div}\boldsymbol{A}$, and
3) $\boldsymbol{\nabla}$ crossed into a vector field $\boldsymbol{A}(\boldsymbol{x})$ yields the vector-valued *curl* of $\boldsymbol{A}$: $\boldsymbol{\nabla}\times\boldsymbol{A} \equiv \operatorname{curl}\boldsymbol{A}$.

And they should (now's the time to practice!) be able to produce—"on demand" as it were—identities such as the following:

$$\operatorname{grad}(\varphi\psi) = \varphi\ \operatorname{grad}\psi + \psi\ \operatorname{grad}\varphi \tag{5.1}$$

$$\operatorname{div}(\varphi\boldsymbol{A}) = \varphi\ \operatorname{div}\boldsymbol{A}\ + \boldsymbol{A}\cdot\operatorname{grad}\varphi \tag{5.2}$$
$$\operatorname{curl}(\varphi\boldsymbol{A}) = \varphi\ \operatorname{curl}\boldsymbol{A} - \boldsymbol{A}\times\operatorname{grad}\varphi \tag{5.3}$$

$$\operatorname{div}(\boldsymbol{A}\times\boldsymbol{B}) = -\boldsymbol{A}\cdot\operatorname{curl}\boldsymbol{B} + \boldsymbol{B}\cdot\operatorname{curl}\boldsymbol{A} \tag{5.4}$$
$$\operatorname{curl}(\boldsymbol{A}\times\boldsymbol{B}) = \boldsymbol{A}\ \operatorname{div}\boldsymbol{B}\quad - (\boldsymbol{A}\cdot\boldsymbol{\nabla})\boldsymbol{B} - \boldsymbol{B}\ \operatorname{div}\boldsymbol{A}\quad + (\boldsymbol{B}\cdot\boldsymbol{\nabla})\boldsymbol{A} \tag{5.5}$$
$$\operatorname{grad}(\boldsymbol{A}\cdot\boldsymbol{B}) = \boldsymbol{A}\times\operatorname{curl}\boldsymbol{B} + (\boldsymbol{A}\cdot\boldsymbol{\nabla})\boldsymbol{B} + \boldsymbol{B}\times\operatorname{curl}\boldsymbol{A} + (\boldsymbol{B}\cdot\boldsymbol{\nabla})\boldsymbol{A} \tag{5.6}$$

. . . all of which (though the last three become "easy" only in consequence of some fairly sophisticated technique) are consequences basically of the "product rule:" $\partial(FG) = F\partial G + G\partial F$. Differential expressions of second (and higher) order are obtained from the above by composition. In particular, one has

$$\operatorname{div}\operatorname{grad}\varphi = \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\varphi \equiv \nabla^2\varphi = \underbrace{\left\{\left(\tfrac{\partial}{\partial x}\right)^2 + \left(\tfrac{\partial}{\partial y}\right)^2 + \left(\tfrac{\partial}{\partial z}\right)^2\right\}}_{\text{Laplacian operator}}\varphi$$

And by explicit calculation[4] one establishes that

$$\operatorname{curl}\operatorname{grad}\varphi = \boldsymbol{0} \quad \text{for all scalar fields } \varphi(\boldsymbol{x}) \tag{6.1}$$
$$\operatorname{div}\operatorname{curl}\boldsymbol{A} = 0 \quad \text{for all vector fields } \boldsymbol{A}(\boldsymbol{x}) \tag{6.2}$$

[4] PROBLEM 4

Turning now from broad generalities to some of their more particular consequences, of which we will soon have specific need ... let

$$\varphi(\boldsymbol{x}) = f(r)$$
$$r \equiv r(\boldsymbol{x}) = \sqrt{\boldsymbol{x}\cdot\boldsymbol{x}} = \sqrt{x^2+y^2+z^2}$$

symbolize the assumption that x, y and z enter φ only *via* r; *i.e.*, that φ is *constant-valued on spheres* (*radius* r: $0 \leqslant r < \infty$) *which are concentric about the origin in* $\boldsymbol{x}$*-space* (we assume more briefly that φ is rotationally-symmetric). Immediately (by the chain rule)

$$\boldsymbol{\nabla}\varphi = f'(r)\boldsymbol{\nabla}r$$

where the prime signifies differentiation of $f(\bullet)$ with respect to its sole argument. It is elementary[5] that

$$\boldsymbol{\nabla}r = \hat{\boldsymbol{x}} \equiv \frac{\boldsymbol{x}}{r} \tag{7.1}$$

and also that

$$\operatorname{div}\boldsymbol{x} = 3 \tag{7.2}$$
$$\operatorname{curl}\boldsymbol{x} = \boldsymbol{0} \tag{7.3}$$

so

$$= \frac{1}{r}f'(r)\,\boldsymbol{x} \tag{8}$$

From (5) and (7) it now follows that

$$\nabla^2\varphi = \frac{1}{r}f'\boldsymbol{\nabla}\cdot\boldsymbol{x} + \boldsymbol{x}\cdot\boldsymbol{\nabla}\Big(\frac{1}{r}f'\Big)$$
$$= \frac{3}{r}f' + \underbrace{(\boldsymbol{x}\cdot\boldsymbol{x})\frac{1}{r}\Big(\frac{1}{r}f'\Big)'}$$
$$= r^2\Big\{\frac{1}{r^2}f'' - \frac{1}{r^3}f'\Big\}$$

giving

$$= f'' + 2\frac{1}{r}f' \tag{9}$$

It is a notable consequence of this fact that[6]

$$\text{If } \nabla^2\varphi = 0 \text{ then } f(r) = \frac{a}{r} + b \quad : \quad a \text{ and } b \text{ are constants}$$

and if, moreover, it is required that $f(\infty) = 0$ then necessarily $b = 0$.

---

[5] PROBLEM 5

[6] PROBLEM 6

We are in position now to state and prove the results we will need to carry forward the discussion of (3). They are (some trivial generalizations of) the following:

$$\boldsymbol{\nabla}\,\frac{1}{r} = -\boldsymbol{x}/r^3 \tag{10.1}$$

$$\nabla^2\frac{1}{r} = -4\pi\delta(\boldsymbol{x}) \tag{10.2}$$

The former follows as an immediate corollary of (8). The presence of the *Dirac delta function* $\delta(\boldsymbol{x})$ on the right side of the latter announces that (10.2) has only a *formal* meaning—will be literally meaningful only when encountered in the protective shade of an $\int$-sign—and promises that the proof will be somewhat tricky. To avoid the fact that $1/r$ becomes *singular* at the origin we study the ϵ-parameterized functions

$$g(r\,;\epsilon) \equiv \frac{1}{r+\epsilon}$$

. . . our plan being to "turn ϵ off" at some appropriate moment(s). Immediately

$$\left.\begin{aligned} g^{\,\prime} &= -\frac{1}{(r+\epsilon)^2} \\ g^{\,\prime\prime} &= +\frac{2}{(r+\epsilon)^3} \end{aligned}\right\} \tag{11}$$

so by (9)

$$\begin{aligned} \nabla^2 g &= 2\Big[\frac{1}{(r+\epsilon)^3} - \frac{1}{r(r+\epsilon)^2}\Big] \\ &= -\infty \text{ at the origin (all } \epsilon\text{), but elsewhere vanishes as } \epsilon \downarrow 0 \end{aligned} \tag{12}$$

Next we notice that the result $\iiint \nabla^2 g\, d^3x$ of integrating $\nabla^2 g$ over all space can (by spherical symmetry) be described

$$\begin{aligned} \int_0^\infty \nabla^2 g\, 4\pi r^2\, dr &= \lim_{R\uparrow\infty} 8\pi \int_0^R \left[\frac{r^2}{(r+\epsilon)^3} - \frac{r}{(r+\epsilon)^2}\right] \\ &= \lim_{R\uparrow\infty} 8\pi \left[\frac{\epsilon}{(r+\epsilon)} - \frac{\epsilon^2}{2(r+\epsilon)^2}\right]_0^R \text{ according to } \textit{Mathematica} \\ &= \lim_{R\uparrow\infty} 4\pi \left[1 - \Big(\frac{r}{r+\epsilon}\Big)^2\right]_0^R \\ &= \lim_{R\uparrow\infty} \underbrace{-4\pi\Big(\frac{R}{R+\epsilon}\Big)^2}_{\text{Remarkably, this becomes } \underline{R\text{-independent}} \text{ as } \epsilon \downarrow 0} \quad : \quad \epsilon > 0 \\ &= -4\pi \end{aligned} \tag{13}$$

The function $\nabla^2 g$—see (12)—has, in other words, these seemingly contradictory properties:

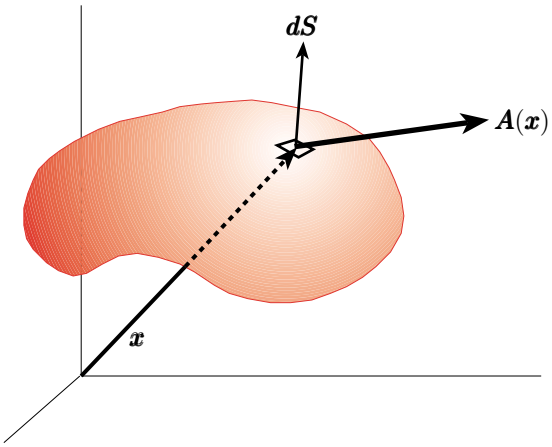


FIGURE 4: *Geometrical context to which* (14) *refers.* $\boldsymbol{x}$ *identifies a point on the boundary* $\partial\mathcal{R}$ *of the "bubble"* $\mathcal{R}$, $\boldsymbol{dS}$ *describes the area and orientation of a surface element, and* $\boldsymbol{A}(\boldsymbol{x})$ *is an arbitrary vector field.*

- it is, for all ϵ (though the fact is masked when $\epsilon = 0$) singular at the origin, but elsewhere
- it vanishes as $\epsilon \downarrow 0$, yet does so in such a way that
- its integral over x-space remains constantly equal to -4π. Finally
- g itself approaches $g(r\,;0) = 1/r$ as $\epsilon \downarrow 0$.

This is precisely the information which the formal equation (10.2) is intended to convey. $\mathcal{QED}$

I should mention that *the preceding line of argument is non-standard*, that the texts argue invariably from the celebrated integral identity

$$\iiint_{\mathcal{R}} \boldsymbol{\nabla}\cdot\boldsymbol{A}\, dV = \iint_{\partial\mathcal{R}} \boldsymbol{A}\cdot\boldsymbol{dS} \tag{14}$$

where (see Figure 4) $\mathcal{R}$ is a "bubble-like" region in 3-dimensional Euclidean space, dV (otherwise denoted d^3x) is an element of volume, $\partial\mathcal{R}$ refers to the (orientable) surface of $\mathcal{R}$, and $\boldsymbol{dS}$ is an outward-directed surface element. That strategy is unavailable to me, since I wish to postpone proof and discussion of *Gauß' theorem* (14) and its relatives. If, however, the reader is content (for the moment) merely to accept (14) then we can

i) take $\mathcal{R}$ to be the sphere of radius R centered at the origin and

ii) take $\boldsymbol{A} = \boldsymbol{\nabla} g$

to obtain

$$\begin{aligned}\iiint_{\mathcal{R}} \nabla^2 g \, d^3x &= \iint_{\partial\mathcal{R}} \boldsymbol{\nabla} g \cdot \boldsymbol{dS} \\ &= -\iint_{\partial\mathcal{R}} \frac{1}{(r+\epsilon)^2} \hat{\boldsymbol{x}} \cdot \boldsymbol{dS} \quad \text{by (8) and (9)}\end{aligned}$$

But $\hat{\boldsymbol{x}} \cdot \boldsymbol{dS} = dS$ since $\hat{\boldsymbol{x}}$ and $\boldsymbol{dS}$ are (for this $\mathcal{R}$) parallel and $\hat{\boldsymbol{x}}$ is a unit vector, so

$$= -\frac{4\pi R^2}{(R+\epsilon)^2}$$

—consistently with a result we obtained *en route* to (13). The surprising fact that this result is (in the limit $\epsilon \downarrow 0$) R-independent is understood as follows: $\nabla^2 g$ is—see again (12)—*singular at the origin* but (in the limit) *vanishes elsewhere*, so $\iiint \nabla^2 g \, d^3x$ *acquires its entire value at/from the singularity* . . . which (again) is the upshot of (10.2). Note finally that by "displacement of the origin" we have

$$\frac{\boldsymbol{x}-\boldsymbol{a}}{|\boldsymbol{x}-\boldsymbol{a}|^3} = -\boldsymbol{\nabla} \frac{1}{|\boldsymbol{x}-\boldsymbol{a}|} \tag{15.1}$$

and

$$\nabla^2 \frac{1}{|\boldsymbol{x}-\boldsymbol{a}|} = -4\pi\,\delta(\boldsymbol{x}-\boldsymbol{a}) \tag{15.2}$$

as trivial generalizations of (10). Equations (15) are fundamental—the results I have been at such pains to derive. END OF DIGRESSION

Returning now with (15.1) to (3.2) we have

$$\boldsymbol{E}(\boldsymbol{x}) = -\iiint \rho(\boldsymbol{\xi}) \boldsymbol{\nabla} \frac{1}{|\boldsymbol{x}-\boldsymbol{\xi}|} d^3\xi \tag{16}$$

Here the operator $\boldsymbol{\nabla}$ looks only to the $\boldsymbol{x}$-dependence of its operand, and since $\boldsymbol{x}$ is not the variable of integration we can take the $\boldsymbol{\nabla}$ outside the $\iiint$, writing

$$= -\boldsymbol{\nabla}\varphi(\boldsymbol{x}) \tag{17}$$

$$\varphi(\boldsymbol{x}) \equiv \iiint \rho(\boldsymbol{\xi}) \frac{1}{|\boldsymbol{x}-\boldsymbol{\xi}|} d^3\xi \tag{18}$$

$$\equiv \underline{\text{electrostatic potential}}, \text{ a } \textit{scalar} \text{ field}$$

Electrostatic $\boldsymbol{E}$-fields are, according to (17), *conservative* (in the sense that they admit of derivation from a scalar "potential," namely the $\varphi(\boldsymbol{x})$ of (18)). The equation

$$\boldsymbol{\nabla} \times \boldsymbol{E} = \boldsymbol{0} \tag{19}$$

—which follows from (17) by (6.1)—provides a compact formulation of the same fundamental fact (and would motivate a hydrodynamicist to remark that such $\boldsymbol{E}$-fields are "irrotational"). Note, however, that (19)—which contains no

reference at all to $\rho(\boldsymbol{x})$—imposes only a weak constraint upon the structure of $\boldsymbol{E}(\boldsymbol{x})$; *i.e.*, that it does not, of itself, enable one to *compute* $\boldsymbol{E}(\boldsymbol{x})$.

Next we take the divergence of (16) to obtain

$$\begin{aligned}\boldsymbol{\nabla}\cdot\boldsymbol{E}(\boldsymbol{x}) &= -\iiint \rho(\boldsymbol{\xi})\,\nabla^2\frac{1}{|\boldsymbol{x}-\boldsymbol{\xi}|}\,d^3\xi \\ &= 4\pi\iiint \rho(\boldsymbol{\xi})\delta(\boldsymbol{x}-\boldsymbol{\xi})\,d^3\xi \quad \text{by (15.2)} \\ &= 4\pi\rho(\boldsymbol{x}) \qquad (20)\end{aligned}$$

And from (20) it follows by (17) that

$$\nabla^2\varphi(\boldsymbol{x}) = -4\pi\rho(\boldsymbol{x}) \qquad (21)$$

Some comments are now in order:

1. Equations (19) and (20)—more specifically

$$\boxed{\begin{array}{cc}\boldsymbol{\nabla}\cdot\boldsymbol{E} = 4\pi\rho \quad , & \frac{\partial}{\partial t}\rho = 0 \\ \boldsymbol{\nabla}\times\boldsymbol{E} = \boldsymbol{0} \quad , & \frac{\partial}{\partial t}\boldsymbol{E} = \boldsymbol{0}\end{array}} \qquad (22)$$

where $\partial\rho/\partial t = 0$ and $\partial\boldsymbol{E}/\partial t = \boldsymbol{0}$ formalize the fact that we are here talking about time-independent physics—provide a *complete <u>local</u> characterization of the laws of electrostatics* ... where "complete" means that *the solution of* (22) *is, for any prescribed boundary conditions, unique*. From (22) one can, in particular, recover the (non-local) statement (3.2) which provided our point of departure.

2. As will be shown later in greater detail, $\boldsymbol{\nabla}\times\boldsymbol{E} = \boldsymbol{0}$ tells us in effect that "there *exists* a (non-unique) φ such that $\boldsymbol{E} = -\boldsymbol{\nabla}\varphi$, while it is the upshot of (21) that <u>in charge-free regions of space</u> φ satisfies *Laplace's equation*

$$\nabla^2\varphi = 0 \qquad (23)$$

In the discrete approximation

$$\begin{aligned}&= \frac{\frac{\varphi(x+\epsilon,y,z)-\varphi(x,y,z)}{\epsilon} - \frac{\varphi(x,y,z)-\varphi(x-\epsilon,y,z)}{\epsilon}}{\epsilon} + y\text{-analog} + z\text{-analog} \\ &= \frac{6}{\epsilon^2}\left\{\frac{\varphi(\text{evaluated at 6 "near neighbors" of }\boldsymbol{x})}{6} - \varphi(\boldsymbol{x})\right\}\end{aligned}$$

so (23) tells us that *in the absence of charge φ "relaxes" until the value assumed by φ at $\boldsymbol{x}$ is the <u>average of the values assumed by φ at the "neighbors" of $\boldsymbol{x}$</u>*. This can be understood to be the "meaning" of Laplace's equation *whatever* the physical/mathematical context in which it is encountered. According to *Poisson's equation*

$$\nabla^2\varphi = -4\pi\rho \qquad (21)$$

the "role" of charge is "to keep φ from relaxing:" φ (locally) exceeds or falls short of the average of neighboring values according as (locally) $\rho \gtrless 0$. Note that if I were to *give* you $\varphi(\boldsymbol{x})$ then you could use (21) to *compute the implied structure of the charge distribution (or "source term")* $\rho(\boldsymbol{x})$.

3. Comparison of (21) with (15.2) shows that we can interpret

$$G(\boldsymbol{x};\boldsymbol{\xi}) \equiv \frac{1}{|\boldsymbol{x}-\boldsymbol{\xi}|}$$

The notation recalls the name of George Green, who (∼1824) was the first to appreciate the power of the general ideas here at issue.

as a description of the *electrostatic potential generated by a unit charge situated at the point* $\boldsymbol{\xi}$ in $\boldsymbol{x}$-space. Now it is fundamental that (see again page 4)

> *Electrodynamics is—like quantum mechanics (but unlike classical mechanics, fluid dynamics, gravitational field physics)—dominated by the principle of superposition.*

This is because the underlying (partial differential) equations are (see (22)) linear: solutions—when

- multiplied by constants and/or
- added to other solutions

—yield solutions. This "build-up principle" pertains, in particular, to (21). Reading the identity

$$\rho(\boldsymbol{x}) = \iiint \rho(\boldsymbol{\xi})\,\delta(\boldsymbol{x}-\boldsymbol{\xi})\,d^3\xi$$

as a formalization of the remark that *arbitrary (even continuous) charge distributions can be synthesized by weighted superposition of point charges*, we infer (by linearity) that $\rho(\boldsymbol{x})$ generates the potential

$$\varphi(\boldsymbol{x}) = \iiint \rho(\boldsymbol{\xi})\,G(\boldsymbol{x};\boldsymbol{\xi})\,d^3\xi \tag{25}$$

Equation (25) is but a notationally disguised re-write of (18), upon which it sheds important new light. If we apply ∇^2 to (25) and insist that the result be—for *all* $\rho(\boldsymbol{x})$—in conformity with (21) then we are forced to the conclusion that

$$\nabla^2 G(\boldsymbol{x};\boldsymbol{\xi}) = -4\pi\delta(\boldsymbol{x}-\boldsymbol{\xi}) \tag{26}$$

which (by (24)) is but a restatement of (15.2), but can be read as the equation that *defines* the *Green's function appropriate to the Poisson equation* (21). Evidently

$$-\boldsymbol{\nabla} G(\boldsymbol{x};\boldsymbol{\xi}) = \text{Coulombic } \boldsymbol{E}\text{-field generated by a unit point charge at } \boldsymbol{\xi}$$

5. To place the preceding remarks in a larger context, and to lend substance to the claim that *the Green's function technique pertains generally to "linear physics"* ... consider the (inhomogeneous linear partial differential) equation

$$(\nabla^2 - \lambda^2)\,\varphi(\boldsymbol{x}) = -4\pi\rho(\boldsymbol{x}) \tag{27}$$

λ is a constant with (evidently) the physical dimensions of $\frac{1}{\text{length}}$

which gives back (21) in the limit $\lambda^2 \downarrow 0$. Equation (25) serves *still* to describe the general solution φ of (27), but the Green's function $G(\boldsymbol{x};\boldsymbol{\xi})$ is constrained now to satisfy not (26) but

$$(\nabla^2 - \lambda^2)\,G(\boldsymbol{x};\boldsymbol{\xi}) = -4\pi\delta(\boldsymbol{x}-\boldsymbol{\xi})$$

which is readily shown[7] to entail

$$G(\boldsymbol{x};\boldsymbol{\xi}) = \frac{1}{|\boldsymbol{x}-\boldsymbol{\xi}|}\,e^{-\lambda|\boldsymbol{x}-\boldsymbol{\xi}|}$$

To reduce the notational clutter and better expose the essence of the matter, write $r \equiv |\boldsymbol{x}-\boldsymbol{\xi}|$ to obtain

$$(\nabla^2 - \lambda^2)G(r) = -4\pi\,\delta(r) \tag{28}$$

$$\Downarrow$$

$$G(r) = \frac{1}{r}\cdot e^{-\lambda r} \tag{29}$$

Equation (29) describes what is sometimes called the "screened Coulomb potential," for reasons that arise from the observation that (28) can be written

$$\left.\begin{aligned} \nabla^2 G(r) &= -4\pi\,\rho(r) \\ \rho(r) &\equiv \delta(r) - \frac{\lambda^2}{4\pi r}\,e^{-\lambda r} \end{aligned}\right\} \tag{30}$$

By quick computation

$$\int_0^\infty \rho(r) 4\pi r^2\,dr = 1 - 1 = 0$$

so (30) can be used to model the *electrostatic environment of a neutral atom* (positively charged point-like nucleus that is "screened" by an exponentially attenuated "electron cloud"—the whole being electrically neutral). A visiting test charge feels an $\boldsymbol{E}$-field given by

$$\begin{aligned} \boldsymbol{E} = -\boldsymbol{\nabla}G &= -\Big(\frac{1}{r}e^{-\lambda r}\Big)'\hat{\boldsymbol{x}} \\ &= \frac{1}{r^2}\underbrace{(1+\lambda r)e^{-\lambda r}}_{\text{attenuation factor}}\,\hat{\boldsymbol{x}} \end{aligned} \tag{31}$$

[7] PROBLEM 7

—the strength of which *falls off faster than* $1/r^2$ (which is to say: "faster than geometrically"), with a "characteristic range" given by λ.

> HISTORICAL NOTE: By 1934 it was known that the so-called "strong force" (the force which overcomes electrostatic repulsion to bind nuclei) is of *short range*. Hideki Yukawa—then 27 years old—saw the opportunity to give the λ-term an important physical job. He recognized that *classical physics*—(e,c)-*physics*—*contains no "natural length"*. Neither does its quantized analog ($(e,c,\hbar)$-physics) . . . but theories of the latter type would acquire a "natural length"—given on dimensional grounds[8] by
>
> $$\text{natural length} \equiv \lambda^{-1} = \frac{\hbar}{\mu c} \tag{32}$$
>
> —if the analog of the photon (Yukawa's hypothetical—but by now very well established—"meson": the particle which mediates the strong interaction) were assigned a non-zero mass μ. Yukawa was led thus to postulate the existence of an elementary particle (it turned out to be a small *population* of particles—the "π-mesons") with mass
>
> $$\mu = \frac{\hbar}{c} \cdot \frac{1}{\text{range of the strong force}} \sim 265 \text{ electron masses}$$
>
> and to suggest that something like the "Yukawa force law" (31) should (in leading approximation) describe the interaction of nucleons. π-mesons were first observed (in nuclear emulsions by Powell & Occhialini) in 1947, and in 1949 Yukawa received the Nobel Prize.

Note finally that

- the "natural length" of (32) *becomes infinite* as $\mu \downarrow 0$;
- the preceding theory becomes "Coulombic" in that limit . . . and could, in particular, be used to construct an alternative to our "non-Gaußian proof" of (10.2);
- we might expect (21) to go over into (27) should it turn out that photons do in fact have a (tiny) mass.

I look finally to the <u>*energetics* of electrostatic fields</u>; *i.e.*, of static chrage configurations. Readers will recall from prior study of elementary mechanics that

1) if $\boldsymbol{F}(\boldsymbol{x})$ describes the forcy environment of a mass point m then the *work that <u>you must perform</u>* to transport m along a prescribed path is given by

$$W[\text{path}] = -\int_{\text{path}} \boldsymbol{F}(\boldsymbol{x})\cdot d\boldsymbol{x}$$

[8] PROBLEMS 8 & 9.

If the path is described parametrically $\boldsymbol{x} = \boldsymbol{x}(\lambda)$: $0 \leqslant \lambda \leqslant 1$ then we can (more specifically) write

$$= -\int_0^1 \boldsymbol{F}\big(\boldsymbol{x}(\lambda)\big) \cdot \frac{\boldsymbol{dx}(\lambda)}{d\lambda} d\lambda$$

2) if the force is *"conservative" in the sense that it admits of description as the gradient of a scalar potential*

$$\boldsymbol{F}(\boldsymbol{x}) = -\boldsymbol{\nabla} U(\boldsymbol{x})$$

then (by $\boldsymbol{\nabla} U \cdot \boldsymbol{dx} = dU$)

$$\begin{aligned} W[\text{path}] &= \int_0^1 \frac{dU(\boldsymbol{x}(\lambda))}{d\lambda}\, d\lambda \\ &= U(\boldsymbol{x}_1) - U(\boldsymbol{x}_0) \end{aligned}$$

Remarkably, the path-dependence of W has dropped away: W has become (not a "path functional" but) a *function of the endpoints* of the path. A simple argument shows, conversely, that path-independence *implies* the existence of U.

We now ask: *What is the work which you must perform to assemble the constellation of charges Q_i first contemplated on page 8*? ... the assumption (mainly of convenience) being that *the Q_i reside initially—far from each other and from us—"at infinity"* (*i.e.*, at the only generally available "standard place").

Evidently we can move the 1st charge Q_1 into position "for free." The 2nd charge Q_2 feels (when at $\boldsymbol{x}$) the Coulombic force

$$\boldsymbol{F}_{12}(\boldsymbol{x}) = kQ_1Q_2 \frac{1}{r_1^3}\boldsymbol{r}_1 = -kQ_2 \boldsymbol{\nabla}\varphi_1(\boldsymbol{x})$$

$$\varphi_1(\boldsymbol{x}) \equiv Q_1 \frac{1}{r_1}$$

exerted by Q_1, and from (33) we infer that to bring Q_2 into position we must do work given by

$$W_2 = kQ_2\big\{\varphi_1(\boldsymbol{x}_2) - \underbrace{\varphi_1(\boldsymbol{\infty})}_{0}\big\} = kQ_2Q_1\frac{1}{r_{21}} \equiv W_{21}$$

$$\boldsymbol{r}_{21} \equiv \boldsymbol{x}_2 - \boldsymbol{x}_1$$

Since electrostatic forces conform to the principle of superposition, the force experienced by Q_3 can be described

$$\boldsymbol{F}_3(\boldsymbol{x}) = \boldsymbol{F}_{31}(\boldsymbol{x}) + \boldsymbol{F}_{32}(\boldsymbol{x}) = -kQ_3\boldsymbol{\nabla}\big\{\varphi_1(\boldsymbol{x}) + \varphi_2(\boldsymbol{x})\big\}$$

. . . and, since $\boldsymbol{F}$ enters *linearly* into the equation $W = \int \boldsymbol{F}\cdot d\boldsymbol{x}$, we infer that to bring Q_3 into position we must do work given by

$$W_3 = kQ_3\big\{\varphi_1(\boldsymbol{x}_3) + \varphi_2(\boldsymbol{x}_3)\big\} = kQ_3Q_1\frac{1}{r_{31}} + kQ_3Q_2\frac{1}{r_{32}} \equiv W_{31} + W_{32}$$

By extension of the same line of argument we obtain

$$W_i = \sum_{j=1}^{i-1} W_{ij}$$

where

$$\left.\begin{aligned} W_{ij} &\equiv kQ_i\varphi_j(\boldsymbol{x}_i) \\ &= kQ_iQ_j\frac{1}{r_{ij}} \quad \text{with} \quad r_{ij} \equiv |\boldsymbol{x}_i - \boldsymbol{x}_j| \\ &= \left\{\begin{array}{l}\text{work done by moving } Q_i \text{ against} \\ \text{the Coulombic force exerted by } Q_j\end{array}\right. \end{aligned}\right\} \tag{34}$$

The <u>energy which we must invest to assemble the entire population</u> is given therefore by

$$\begin{aligned} W = \quad & \bullet && \bullet \\ & W_{21} \quad \bullet && \bullet \\ & +W_{31} + W_{32} \quad \bullet && \bullet \\ & \vdots \\ & +W_{n1} + W_{n2} + \cdots + W_{n,n-1} \quad \bullet \end{aligned}$$

$$= \sum_{i>j} W_{ij} \tag{35.1}$$

but would have been given by

$$= \sum_{i>j} W_{ij} \quad : \quad \text{note the reversed inequality} \tag{35.2}$$

had we reversed the order of assembly. Equations (35) are, of course, equivalent (by $W_{ij} = W_{ji}$: see (34)) . . . and can be combined to give

$$= \tfrac{1}{2}\sum_i \sum_j{}' W_{ij} \tag{36}$$

↑—the prime means that the *self-energy terms* (terms with $i = j$) are to be <u>omitted</u>

which possesses a pleasing formal symmetry. One might be strontly tempted to write

↓—note that the prime is now absent

$$W = \tfrac{1}{2}\sum_i \sum_j W_{ij} - \sum_i W_{ii} \tag{37}$$

were it not for the important fact that *each of the "self-energy terms" W_{ii} is numerically infinite*. Equation (37) possesses, in other words, the latently deceptive structure

$$= (W + \infty) - \infty$$

Computationally/conceptually more useful results can be obtained now by appeal to (34): thus

$$W = \tfrac{1}{2}k \sum_i Q_i \underbrace{\Big\{ {\sum_j}' \varphi_j(\boldsymbol{x}_i) \Big\}}_{= {\sum_j}' Q_j \frac{1}{r_{ij}}} \tag{38.1}$$

$$= \text{potential at } \boldsymbol{x}_i \text{ due to all other charges } Q_j$$

$$= \tfrac{1}{2}k \sum_i {\sum_j}' Q_i Q_j \frac{1}{r_{ij}} \tag{38.2}$$

In the *continuous limit* equations (38) go over semi-plausibly into

$$W = \tfrac{1}{2}k \int \rho(\boldsymbol{x})\varphi(\boldsymbol{x})\, d^3x \tag{39.1}$$

$\int$ means $\iiint$, and extends over any region $\mathcal{R}$ big enough to contain all points where the charge distribution $\rho \neq 0$

$$= \tfrac{1}{2}k \iint \rho(\boldsymbol{x})\rho(\boldsymbol{\xi})\, \frac{1}{|\boldsymbol{x} - \boldsymbol{\xi}|}\, d^3x\, d^3\xi \tag{39.2}$$

which are (by (18)) equivalent. I say "*semi*-plausibly" becuase equations (39) contain no analogs of the primes which decorate (38). Indeed, if we set

$$\rho(\boldsymbol{\xi}) = \sum_i Q_i\, \delta(\boldsymbol{\xi} - \boldsymbol{x}_i) \tag{4}$$

we can perform the $\iint$ and obtain

$$= \tfrac{1}{2} \sum_i \sum_j W_{ij} = W + \text{self-energy terms}$$

We confront therefore this fundamental question (which I must, for the moment, leave dangling): For *continuous* charge distributions $\rho(\boldsymbol{x})$ do "self-energy terms" (ever? sometimes? always?) automatically *vanish*?

We are in position now to review some ideas which are as fundamental as they are pretty. Introducing $\rho = -\frac{1}{4\pi}\nabla^2\varphi$ into (39.1) we obtain

$$W = -\tfrac{1}{8\pi}k \iiint \varphi \nabla^2 \varphi\, d^3x \tag{40}$$

which will strike some readers as reminiscent of the formula

$$\langle E\rangle = \iiint \psi^*\Big\{-\tfrac{\hbar^2}{2m}\nabla^2 + V\Big\}\psi\, d^3x$$

by means of which one computes the *expected value of the average of many energy measurements* if the quantum mechanical system with Hamiltonian $\mathsf{H} = \frac{1}{2m}\mathsf{p}^2 + V(\mathsf{x})$ is known to be in state ψ. Be that as it may ... it follows from (40) (more directly: introduce $\rho = \frac{1}{4\pi}\boldsymbol{\nabla}\cdot\boldsymbol{E}$ into (39.1)) that

$$W = \tfrac{1}{8\pi}k \iiint \varphi\, \boldsymbol{\nabla}\cdot\boldsymbol{E}\, d^3x$$

By (5.2)

$$\varphi\, \boldsymbol{\nabla}\cdot\boldsymbol{E} = -\boldsymbol{E}\cdot\boldsymbol{\nabla}\varphi + \boldsymbol{\nabla}\cdot(\varphi\boldsymbol{E})$$

while by (17)

$$\boldsymbol{\nabla}\varphi = -\boldsymbol{E}$$

So—by what is in effect the 3-dimensional analog of an "integration by parts"—we have

$$= \tfrac{1}{8\pi}k\Big\{ \iiint E^2\, d^3x + \underbrace{\iiint \boldsymbol{\nabla}\cdot(\varphi\boldsymbol{E})\, d^3x}\Big\}$$

$$= \iint \varphi\boldsymbol{E}\cdot d\boldsymbol{S} \quad \text{by Gauß' theorem (14)}$$

We expect $\varphi\boldsymbol{E}$ to fall off asymptotically as $1/r^3$. This is fast enough to cause the later surface integral to *vanish if the surface of integration is "removed to infinity"* ... giving

$$= \tfrac{1}{8\pi}k \iiint E^2\, d^3x \tag{41}$$

where $E^2 \equiv \boldsymbol{E}\cdot\boldsymbol{E}$ and where *the* $\iiint$ *ranges over all space* (or at least over all points where $\boldsymbol{E}(\boldsymbol{x}) \neq \boldsymbol{0}$). Several lines of commentary are now in order:

1. All that has been said concerning W pertains as well to the energetics of *gravitational* (or at least to weak gravitostatic) fields as it does to electrostatic fields. The space-curvature effects associated with very strong fields (whether gravitational or electrostatic) can, of course, be expected to cause our (tacit) Euclidean assumptions to break down ...

2. More familiar to chemists than to physicists—and so general/powerful that it is difficult to formulate except in words—is

> LE CHATELIER'S PRINCIPLE: When an external force is applied to a system in equilibrium the system adjusts so as to minimize the effect of the applied force.

Somewhat similar—in substance and spirit, in its abstract generality, and in its ever-surprising power—is this

> NAMELESSS PRINCIPLE: If the energy E of a system depends upon an adjustable parameter α (of whatever nature) then an "abstract force" $\mathcal{F} = -\partial E/\partial\alpha$ will be associated with variation of that parameter. If α refers to *spatial position* then $\mathcal{F}$ will have literally the nature of a *mechanical* force.

Our electrostatic W is by nature a function of $\boldsymbol{x}_1$, Q_1, $\boldsymbol{x}_2$, Q_2, ..., $\boldsymbol{x}_n$, Q_n (in the discrete case, and a functional of $\rho(\boldsymbol{x})$ in the continuous case). *What is the (literal) force associated with variation of* $\boldsymbol{x}_i$? Bringing (34) to (36) we have[9]

$$-\boldsymbol{\nabla}_i W = -\boldsymbol{\nabla}_i \tfrac{1}{2} k {\sum_{a,b}}' Q_a Q_b \frac{1}{r_{ab}} \tag{42.1}$$

$$= -\tfrac{1}{2} k Q_i \boldsymbol{\nabla}_i \Big\{ \underbrace{{\sum_a}' Q_a \frac{1}{r_{ai}} + {\sum_b}' Q_b \frac{1}{r_{ib}}}_{\text{sums identical by } r_{ab} = r_{ba}} \Big\}$$

$$= -k Q_i \boldsymbol{\nabla}_i {\sum_j}' Q_j \frac{1}{r_{ij}}$$

$$= k Q_i \cdot (\boldsymbol{E}\text{-field at } \boldsymbol{x}_i \text{ due to all other charges})$$

$$= \text{force exerted on } Q_i \text{ by the other charges} \tag{42.2}$$

Note that there is a formal sense in which the prime can be dropped from (42.1): Q_i's self-energy W_{ii}—though infinite—does not change when Q_i is moved ... so $\boldsymbol{\nabla} w_{ii} = 0$:

> A charge Q "carries its self-energy with it," so *does not exert an electrostatic force upon itself.*

Our "nameless principle" can be used to explain why dielectric fluids are lifted into the space between charged capacitor plates, why magnets attract paper clips, where the thermodynamic concepts of "pressure" and "chemical potential" come from ... and much, much else.

3. *Where does W reside*? The structure of (39)—in which the $\int$'s need extend only over that portion of space which contains charge—tempts one to respond "In the charge(s)" ... or perhaps "In the 'Coulombic springs' by which the charges are interconnected." But those "springs" are spooky things, which inhabit empty space. And one is, on the other hand, encouraged by the structure of (41)—where the $\int$ ranges over that portion of space which contains (not charge but) $\boldsymbol{E}$-field—to suppose that W resides "In the $\boldsymbol{E}$-field; *i.e.*, in the empty space which envelops the charge." The question therefore arises: Which viewpoint is correct (= more useful)? The clear answer is "The latter" ... but only on grounds which emerge when one enlarges the conceptual context to contain dynamical (*i.e.*, t-dependent) elements:

> It is most useful to *consider W to reside "in the* $\boldsymbol{E}$*-field."*

[9] I use a and b as summation indices because i is now otherwise engaged.

We are led thus to speak of *field energy* ... and to begin to *think of* $\boldsymbol{E}$*-fields (since they possess energy) as "mechanical objects in their own right."* Such "objects" differ from (say) particles mainly in the fact that they (*i.e.*, *their collective properties*) *are not localized but* <u>*distributed*</u>. Equation (41) can in this light be written

$$\left.\begin{aligned} W &= \iiint \mathcal{E}(\boldsymbol{x})\, d^3x \\ \mathcal{E}(\boldsymbol{x}) &\equiv \tfrac{1}{8\pi} k|\boldsymbol{E}(\boldsymbol{x})|^2 = \text{electrostatic } \textit{energy density} \end{aligned}\right\} \tag{44}$$

4. It is obvious from (44) that *electrostatic energy density* $\mathcal{E}(\boldsymbol{x})$ *is invariably* <u>*non-negative*</u>:

$$W = W_{\text{interaction}} + W_{\text{self}} \geqslant 0$$

The W described by equations (38) is, on the other hand, clearly of *indefinite* sign. This slight paradox is resolved by the realization that (38)—which applies only to discrete charges—pertains only to the *interaction energy*

$$W_{\text{interaction}} \gtrless 0$$

while

$$W_{\text{self}} > 0$$

Were we to use (41 $\equiv$ 44) in problems involving point charges we would (automatically) be taking into explicit account the *energy expended in the* <u>*assembly*</u> *of those point charges* ... which since

- we are in fact physically unable to "assemble" electrons
- a result of the form $W = \infty$ is not very useful

would be poor policy. In discrete problems it is essential that one use (38), not (41/44). One begins to see why, for ∼8o years, physicists have spoken balefully of the "self-energy *problem*" ... which quantum theory transforms, but does not eliminate.[10]

5. According to (41 $\equiv$ 44), W—irrespective of how self-energy terms are handled—is a <u>*non-linear*</u> *number-valued functional of* $\boldsymbol{E}(\boldsymbol{x})$: if $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ give rise to W_1 and W_2, then $\boldsymbol{E} = \boldsymbol{E}_1 + \boldsymbol{E}_2$ gives rise not to $W_1 + W_2$ but to

$$W = W_1 + W_2 + 2\tfrac{1}{8\pi} k \iiint \boldsymbol{E}_1(\boldsymbol{x}) \cdot \boldsymbol{E}_2(\boldsymbol{x})\, d^3x \tag{45}$$

I have on page 6 drawn attention to the conventional status of k, and wish now (actually for practical reasons) to illustrate how that circumstance might be exploited. If we *think of the field equations (22) as fundamental, and of Coulomb's law as arising from a* <u>*particular (spherically symmetric) solution*</u> of those equations, then it becomes natural to suppose that all factors of 4π

[10] PROBLEMS 10 & 11

should attach not to the field equations but to Coulomb's law and its immediate corollaries. Accordingly, we

$$\text{set } k = \tfrac{1}{4\pi} \quad : \quad \text{(dimensionless)} \tag{46.0}$$

so Coulomb's law (1) reads

$$\boldsymbol{F} = \frac{1}{4\pi}\frac{q_1 q_2}{r^2}\,\hat{\boldsymbol{r}} \tag{46.1}$$

... which serves, in effect to *define our ("rationalized electrostatic") unit of charge*. We can further—and quite independently—simplify life by absorbing a k into the definition of $\boldsymbol{E}$, writing

$$\boldsymbol{F}(\boldsymbol{x}) = q\boldsymbol{E}(\boldsymbol{x}) = \text{force on the test charge } q \tag{46.2}$$

Equations (3) become

$$\boldsymbol{E}(\boldsymbol{x}) = \tfrac{1}{4\pi}\iiint \rho(\boldsymbol{\xi})\,\frac{\boldsymbol{x}-\boldsymbol{\xi}}{|\boldsymbol{x}-\boldsymbol{\xi}|^3}\,d^3\xi \tag{46.3}$$

and its discrete analog (which there is no need to write out). If we insist—conventionally—upon retaining the simplicity of

$$= -\boldsymbol{\nabla}\varphi(\boldsymbol{x}) \tag{17}$$

then (arguing as before from (15.1)) we obtain (compare (18))

$$\varphi(\boldsymbol{x}) = \tfrac{1}{4\pi}\iiint \rho(\boldsymbol{\xi})\,\frac{1}{|\boldsymbol{x}-\boldsymbol{\xi}|}\,d^3\xi \tag{46.4}$$

It is evident that in place of (20) and (21) we now have

$$\boldsymbol{\nabla}\cdot\boldsymbol{E} = -\nabla^2\varphi = \rho \tag{46.5}$$

The 4π-factors, which formerly resided in the field equations, are not attached only to expressions which are clearly and directly "Coulombic in character." In place of (39.1) we now have

$$W = \tfrac{1}{2}\int \rho(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,d^3x$$

giving

$$\begin{aligned} &= -\tfrac{1}{2}\int \varphi\,\nabla^2\varphi\,d^3x \\ &= \tfrac{1}{2}\int \varphi\,\boldsymbol{\nabla}\cdot\boldsymbol{E}\,d^3x = \tfrac{1}{2}\int E^2\,d^3x \end{aligned} \tag{46.6}$$

... all of which are conspicuously cleaner that their counterparts on pages 21/22, and which suggest that cleanliness invested in the field equations tends to

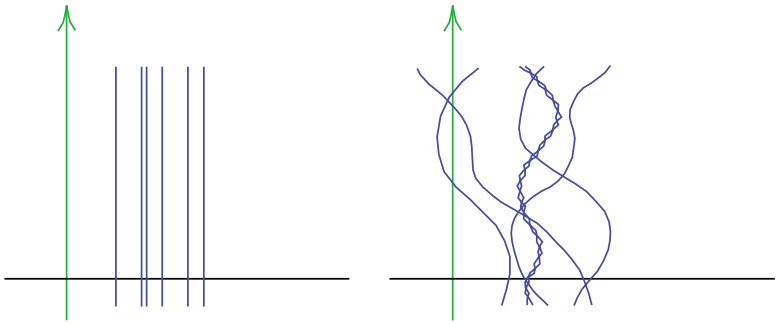


FIGURE 5: *Spacetime diagrams of (on the left) the worldlines traced by a static population of charges, and (on the right) by a dynamical population. The time axis is indicated* ↑.

persist: the best way to clean up a theory is to scour its most fundamental statements. The exercise has involved a good deal of slip and slide: we have

1) assigned to k a conventional value;
2) conventionally adjusted the relationship between $\boldsymbol{F}$ and $\boldsymbol{E}$;
3) conventionally retained the relationship between $\boldsymbol{E}$ and φ.

The whole business admits obviously of many variations . . . and is never much fun.

Our work thus far contains no reference to *time*, no reference to *magnetism*. It turns out—surprisingly, and independently of whether one proceeds in the laboratory or on paper—that to make provision for either is (semi-automatically) to make provison for the other: that

$$\text{electrostatics} \longrightarrow \text{electro}\textit{dynamics}$$

is a program conceptually equivalent to

$$\text{electrostatics} \longrightarrow \text{electro}\textit{magnetism}$$

Now, electrostatics has been seen to proceed from essentially three assumptions:

1) the field sources Q_i don't move; *i.e.*, that they trace worldlines of the form illustrated in the first of the following figures, and that within that specialized context
2) they interact *via* Coulomb's law; moreover
3) electrostatic forces compose by superposition.

Our forward progress requires that we relax the immobility assumption . . . allowing the Q_i to trace worldlines like those shown on the right side of Figure 5.

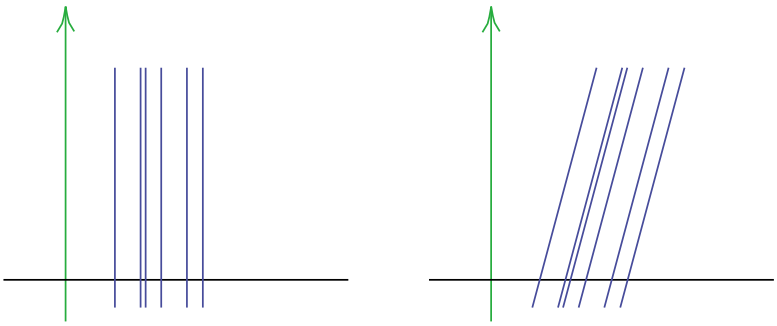


FIGURE 6: *Spacetime diagrams of (on the left) a static population as it appears to O and (on the right) the same population as it appears to us, who see O to be moving by with uniform velocity* $\boldsymbol{v}$.

How to proceed? *How do charged particles interact when they are in relative motion*?

2. Bootstrapping our way to Maxwell's equations. Since my ultimate intent is simply to illuminate the formal/physical ramifications of the *structural properties of the electromagnetic field* (and—as opportunities arise—to illustrate some of the heuristic devices characteristic of modern theoretical physics) . . . I need not apologize for the fact that the discussion which follows is grossly ahistorical. The essential pattern of the argument is due to Julian Schwinger (unpublished notes: ∼1976)[11] . . . but several closely related lines of argument have been around for decades, have been reinvented many times by many people, and have been promoted in the classrooms of Reed College by Dennis Hoffman.

What follows is by nature a "theoretical *bootstrap*" operation, which draws heavily (if interestingly) upon "plausibility arguments" and which leads to results which would remain *merely* plausible in the absence of supporting observational data . . . of which, as it turns out, there is a great deal. The success of the program can itself be read as evidence either of

- the power of hindsight or
- the *extraordinary simplicity of electrodynamics*.

Turning now from anticipatorty generalities to the curious details of our argument . . . let O be an *inertial observer*

1) whom we see to be gliding by with constant velocity $\boldsymbol{v}$;

[11] An edition of those notes was prepared posthumously by several of Schwinger's former associates, and was recently published: see J. Schwinger, L. L. DeRaad, K. A. Milton & W. Tsai, *Classical Electrodynamics* (1998)—especially Chapter 1—and also the review by Jagdish Mehra: AJP **68**, 296 (2000).

2) who possesses the usual "good clock and Cartesian frame" with the aid of which he assigns coordinates $(t, \boldsymbol{x})$ to points in spacetime; *i.e.*, to "events" in his inertial neighborhood;[12]

3) who possesses "complete knowledge of electro*statics*," as developed in §1 and summarized on page 25.

We, on the other hand, possess

1) our own "good clock and Cartesian frame";

2) enough knowledge of physics to know that if O is inertial then so are we ... and (more specifically) enough knowledge of Galilean relativity to "know" that *the coordinates* $(t, \boldsymbol{x})$ *which we assign to an event are related to the coordinates* $(t, \boldsymbol{x})$ *which* O *assigns to that same event* by the equations

$$\left.\begin{aligned} t &= t(t, \boldsymbol{x}) = t \\ \boldsymbol{x} &= \boldsymbol{x}(t, \boldsymbol{x}) = \boldsymbol{x} + \boldsymbol{v}t \end{aligned}\right\} \tag{47}$$

3) no prior knowledge of electrostatics.

Our simple goal—at least at the outset—is to translate O's electrostatic equations in to our variables. The circumstance which makes the enterprise interesting is (see Figure 6) that *while* O*'s charges are at rest with respect both to* O *and to each other ... they are in (uniform) motion with respect to us*.[13] We confront therefore a situation intermediate between those depicted in Figure 5. Though the figures refer (as a matter of graphic convenience) to point charges, we shall find it analytically most convenient to work with continuous charge distributions ρ—a convention which entails no essential loss of generality.

In 2-dimensional spacetime (to which I retreat for merely notational convenience) it would follow from (47) that

$$\begin{aligned} \frac{\partial}{\partial t} &= \frac{\partial t}{\partial t}\frac{\partial}{\partial t} + \frac{\partial x}{\partial t}\frac{\partial}{\partial x} = \frac{\partial}{\partial t} + v\frac{\partial}{\partial x} \\ \frac{\partial}{\partial x} &= \frac{\partial t}{\partial x}\frac{\partial}{\partial t} + \frac{\partial x}{\partial x}\frac{\partial}{\partial x} = \qquad\quad \frac{\partial}{\partial x} \end{aligned}$$

while from (47) itself it follows (similarly) that[14]

$$\left.\begin{aligned} \frac{\partial}{\partial t} &= \frac{\partial}{\partial t} + \boldsymbol{v}\cdot\boldsymbol{\nabla} \\ \boldsymbol{\nabla} &= \qquad\quad \boldsymbol{\nabla} \end{aligned}\right\} \tag{48}$$

[12] The extent of such a neighborhood is set by curvature effects; *i.e.*, by the structure of the gravitational field. We shall eliminate such (typically quite small) effects by *supposing gravitation to have been "switched off"* ($G \downarrow 0$), so all "neighborhoods" become infinite and coextensive: spacetime becomes (not just locally but) *globally flat*.

[13] From O's point of view we are in effect asking: "How does an electrostatic field look to a moving observer (namely: us)?

[14] Here $\boldsymbol{\nabla}$ denotes "del with respect to $\boldsymbol{x}$," while $\boldsymbol{\nabla}$ denotes "del with respect to $\boldsymbol{x}$."

If we assume —plausibly?—that O's charge density can by us be described

$$\rho(\boldsymbol{x},t)=\rho(\boldsymbol{x}-\boldsymbol{v}t)=\rho(\boldsymbol{x}) \tag{49}$$

—*i.e.*, that (relative to the $\boldsymbol{v}$-parameterized Galilean transformations (47)) ρ transforms as a scalar field—then O's equation

$$\tfrac{\partial}{\partial t}\rho=0 \tag{22}$$

can by us (according to (48)) be written

$$\tfrac{\partial}{\partial t}\rho+\boldsymbol{v}\cdot\boldsymbol{\nabla}\rho=0$$

The $\boldsymbol{x}$-independence of $\boldsymbol{v}$ entails $\boldsymbol{\nabla}\cdot\boldsymbol{v}=0$ so (by (5.2))

$$\boldsymbol{v}\cdot\boldsymbol{\nabla}\rho=\boldsymbol{\nabla}\cdot(\rho\boldsymbol{v})$$

and if we define

$$\boldsymbol{j}\equiv\rho\boldsymbol{v}\equiv\text{(electric) }\textit{current density} \tag{50}$$

we have

$$\boxed{\tfrac{\partial}{\partial t}\rho+\boldsymbol{\nabla}\cdot\boldsymbol{j}=0} \tag{51}$$

Postponing (here and below) all physical/formal commentary, I can report that (51) provides a local formulation of the principle of *charge conservation*.

If we assume —plausibly in view of (49) and what we know from mechanics about the Galilean transform properties of force—that O's $\boldsymbol{E}$-field can by us be described

$$\boldsymbol{E}(\boldsymbol{x},t)=\boldsymbol{E}(\boldsymbol{x}-\boldsymbol{v}t)=\boldsymbol{E}(\boldsymbol{x})$$

—*i.e.*, that the *individual components* of $\boldsymbol{E}$ respond to (47) like *scalar* fields—then O's equation

$$\tfrac{\partial\boldsymbol{E}}{\partial t}=0 \tag{22}$$

can by us be written

$$\tfrac{\partial\boldsymbol{E}}{\partial t}+(\boldsymbol{v}\cdot\boldsymbol{\nabla})\boldsymbol{E}=0$$

Expressions of the form $(\boldsymbol{v}\cdot\boldsymbol{\nabla})\boldsymbol{E}$ are—since $\boldsymbol{\nabla}$ is neither dotted nor crossed into $\boldsymbol{E}$—"funny," but they are in fact familiar already from (5). It follows in fact from (5.5) that

$$\begin{aligned}(\boldsymbol{v}\cdot\boldsymbol{\nabla})\boldsymbol{E}=\boldsymbol{v}(\boldsymbol{\nabla}\cdot\boldsymbol{E})&-\boldsymbol{\nabla}\times(\boldsymbol{v}\times\boldsymbol{E})\\&-\underbrace{\boldsymbol{E}(\boldsymbol{\nabla}\cdot\boldsymbol{v})+(\boldsymbol{E}\cdot\boldsymbol{\nabla})\boldsymbol{v}}\\&\qquad 0\ \text{ by }\boldsymbol{x}\text{-independence of }\boldsymbol{v}\end{aligned}$$

O's equation

$$\boldsymbol{\nabla}\cdot\boldsymbol{E}=\rho$$

can by us (and without the assistance of any additional assumptions) be written

$$\boxed{\boldsymbol{\nabla}\cdot\boldsymbol{E}=\rho} \tag{53}$$

It follows therefore by (50) that

$$=\boldsymbol{j}-\boldsymbol{\nabla}\times(\boldsymbol{v}\times\boldsymbol{E})$$

So we have

$$\boxed{\boldsymbol{\nabla}\times\boldsymbol{B} = \tfrac{1}{c}\boldsymbol{j} + \tfrac{1}{c}\tfrac{\partial \boldsymbol{E}}{\partial t}} \tag{54}$$

where now

$$\boldsymbol{B} \equiv \tfrac{1}{c}(\boldsymbol{v}\times\boldsymbol{E}) \tag{55}$$

Here c—a *constant with the dimensions of velocity* (it turns out in a moment to be associated with the speed of light)—has been introduced for no more fundamental purpose than to insure that $\boldsymbol{E}$ and $\boldsymbol{B}$ are dimensionally identical: $[\boldsymbol{E}] = [\boldsymbol{B}]$. $\boldsymbol{B}(\boldsymbol{x}, t)$ itself is a vector field which turns out to be associated with the phenomenology of *magnetism*.[15]

In view of the structure of (54) it becomes natural to inquire after the value of $\boldsymbol{\nabla}\cdot\boldsymbol{B}$. Drawing upon (5.4) we have

$$\begin{aligned}\boldsymbol{\nabla}\cdot\boldsymbol{B} &= \tfrac{1}{c}\boldsymbol{\nabla}\cdot(\boldsymbol{v}\times\boldsymbol{E})\\ &= -\tfrac{1}{c}\boldsymbol{v}\cdot(\boldsymbol{\nabla}\times\boldsymbol{E}) + \underbrace{\tfrac{1}{c}\boldsymbol{E}\cdot(\boldsymbol{\nabla}\times\boldsymbol{v})}_{\substack{0 \text{ by } \boldsymbol{x}\text{-independence of } \boldsymbol{v}}}\end{aligned}$$

But O's equation

$$\boldsymbol{\nabla}\times\boldsymbol{E} = 0 \tag{22}$$

can by us (and again without the assistance of any additional assumptions) be written

$$\boldsymbol{\nabla}\times\boldsymbol{E} = 0 \tag{56}$$

So we have

$$\boxed{\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0} \tag{57}$$

It is a striking fact that the preceding boxed equations contain no reference to $\frac{\partial \boldsymbol{B}}{\partial t}$. But it follows from (55) by differentiation that

$$\tfrac{\partial \boldsymbol{B}}{\partial t} = \tfrac{1}{c}\big(\boldsymbol{v}\times\tfrac{\partial \boldsymbol{E}}{\partial t}\big)$$

which by (54) becomes

$$= \boldsymbol{v}\times(\boldsymbol{\nabla}\times\boldsymbol{B}) - \underbrace{\tfrac{1}{c}(\boldsymbol{v}\times\boldsymbol{j})}_{\substack{0 \text{ because } \boldsymbol{v} \text{ and } \boldsymbol{j} \text{ are, by (50) parallel}}}$$

Reading from (5.6) we have

$$\boldsymbol{v}\times(\boldsymbol{\nabla}\times\boldsymbol{B}) = \boldsymbol{\nabla}(\boldsymbol{v}\cdot\boldsymbol{B}) - (\boldsymbol{v}\cdot\boldsymbol{\nabla})\boldsymbol{B} - \boldsymbol{B}\times(\boldsymbol{\nabla}\times\boldsymbol{v}) - (\boldsymbol{B}\cdot\boldsymbol{\nabla})\boldsymbol{v}$$

[15] PROBLEM 12.

The leading term on the right presents $(\boldsymbol{v}\cdot\bar{\boldsymbol{B}}) = \frac{1}{c}\boldsymbol{v}\cdot(\boldsymbol{v}\times\bar{\boldsymbol{B}})$, which vanishes because $\boldsymbol{v} \perp (\boldsymbol{v}\times\boldsymbol{B})$. And the two final terms vanish because $\boldsymbol{v}$ is $\boldsymbol{x}$-independent. The surviving term can be developed

$$-(\boldsymbol{v}\cdot\boldsymbol{\nabla})\boldsymbol{B} = -\boldsymbol{v}\boldsymbol{\nabla}\cdot\boldsymbol{B} + \boldsymbol{\nabla}\times(\boldsymbol{v}\times\boldsymbol{B}) + \boldsymbol{B}\boldsymbol{\nabla}\cdot\boldsymbol{v} - (\boldsymbol{B}\cdot\boldsymbol{\nabla})\boldsymbol{v}$$

The leading term on the right vanishes by (57): $\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0$. And again: the two final terms vanish because $\boldsymbol{v}$ is $\boldsymbol{x}$-independent. So we have

$$\tfrac{\partial \boldsymbol{B}}{\partial t} = \boldsymbol{\nabla}\times(\boldsymbol{v}\times\boldsymbol{B}) \tag{58}$$

of which, as we saw *en route*,

$$\tfrac{\partial \boldsymbol{B}}{\partial t} + (\boldsymbol{v}\cdot\boldsymbol{\nabla})\bar{\boldsymbol{B}} = \boldsymbol{0} \tag{59}$$

provides an alternative formulation.[16,17]

Reviewing our progress ... we note that equations structurally similar to ours would be written by each of our inertial colleauges (our relation to O being entirely "typical"). Since we are related to O by a Galilean transformation (more specifically: by a *Galilean boost*, as rotation-free transformations of the form (47) are standardly called), and since the compose of two (or more) Galilean transformations is itself Galilean,[18] what we have in effect constructed is a *Galilean-covariant formulation of electrostatics*— a theory which gives back the standard theory (O's theory, as summarized on page 25) when $\boldsymbol{v}$ (whence also $\boldsymbol{j}$ and $\boldsymbol{B}$, by (50) and (55)) vanishes.

But the theory we seek is more ambitious. We seek a theory capable of describing the *electromagnetic interaction of charged mass points Q_i which are free to experience arbitrarily complex relative motions*. Looking in this light to the theory in hand, we note that

1) the $\boldsymbol{v}$ which enters into the definition (50) of $\boldsymbol{j}$ may be interpreted as referring to *our perception of the velocity of an existential thing* (a charged mass point), but that
2) every other reference to $\boldsymbol{v}$ is a source of acute embarrassment, for it is a reference to our perception of the velocity of a non-entity: an observer who sees all charges to be at rest. The inertial observers who perceive any *particular* Q_i to be *momentarily* at rest are easily discovered. But an observer O who sees *all* Q_i to be *constantly* at rest *does (in the general case) not exist*!

How to get along without the assistance of our "preferred observer"? How—with minimal formal damage—to eliminate the embarrassing $\boldsymbol{v}$-terms from our theory?

---

[16] If we apply $\boldsymbol{\nabla}\cdot$ to (58) we obtain $\frac{\partial}{\partial t}\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0$ which, while it does not imply, is certainly consistent with (57): $\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0$.

[17] All that has been said thus far pertains as accurately to our perception of O's *gravitostatics* as to our perception of his electrostatics.

[18] PROBLEM 13.

It is here that we have recourse to the "bootstrap" mentioned on page 27. And here, by the way, that we take leave of Newtonian gravitostatics.

We seem to be forced—both formally (at (55)) and phenomenologically (of which more later)—to retain something like $\boldsymbol{B}$ in our theory. But we eliminate one embarrassing $\boldsymbol{v}$ if

1) we *deny the invariable/general validity of* $\boldsymbol{B} \equiv \frac{1}{C}(\boldsymbol{v}\times\boldsymbol{E})$. This has the effect of *promoting* $\boldsymbol{B}(\boldsymbol{x},t)$ *to the status of an autonomous (if—at the moment—operatonally undefined) field* . . . entitled to all the privileges and respect that we are in the habit of according to $\boldsymbol{E}(\boldsymbol{x},t)$.

The proof of (57) now breaks down, but (note that (57) contains no $\boldsymbol{v}$-term, and must retain at least its electrostatic validity) the situation is saved if

2) we *promote* $\boldsymbol{\nabla}\cdot\bar{\boldsymbol{B}} = 0$ *to the status of a law*.

3) We have every reason to *retain* $\boldsymbol{\nabla}\cdot\boldsymbol{E} = \rho$ *as it stands*. Noting that (53) and (54) conjointly imply charge conservation (which we wish to retain) and that (54) contains as it stands no $\boldsymbol{v}$-term, it seems to make conservative good sense if (tentatively)

4) we *promote* $\boldsymbol{\nabla}\times\boldsymbol{B} = \frac{1}{C}\boldsymbol{j} + \frac{1}{C}\frac{\partial\bar{\boldsymbol{E}}}{\partial t}$ *to the status of a law*.

Our problem assumes at last its full force in this question:

$$\text{What to do with (58): } \tfrac{\partial\boldsymbol{B}}{\partial t} = \boldsymbol{\nabla}\times(\boldsymbol{v}\times\bar{\boldsymbol{B}})?$$

It was at just such a juncture that we were motivated at (55) to define $\bar{\boldsymbol{B}}$. One is therefore tempted to write

$$\begin{aligned}\boldsymbol{C} &\equiv \tfrac{1}{C}(\boldsymbol{v}\times\boldsymbol{B})\\ &= \tfrac{1}{C^2}(\boldsymbol{v}\times(\boldsymbol{v}\times\boldsymbol{E})) = \tfrac{1}{C^2}\big[(\boldsymbol{v}\cdot\boldsymbol{E})\boldsymbol{v} - (\boldsymbol{v}\cdot\boldsymbol{v})\boldsymbol{E}\big]\end{aligned} \tag{60}$$

and then to declare $\boldsymbol{C}(\boldsymbol{x},t)$ "autonomous." But such a program (which would amount to sweeping the dirt under the carpet) must—because of its allusion to $\boldsymbol{v}$—now be dismissed as conceptually unattractive. Besides, it would oblige us to search (by the methods of Galilean electrostatics?) for the field equations satisfied by $\boldsymbol{C}$. Such activity would certainly lead us to the field $\boldsymbol{v}\times\boldsymbol{C}$, and thus oblige us to *keep on* introducing such fields . . . a process which would terminate if and only if it were to turn out that at some stage the resulting "$\boldsymbol{Z}$-field" were a ($\boldsymbol{v}$-independent) linear combination of fields previously introduced, which is unlikely/impossible. How, therefore, to proceed?

When in a theoretical jam, it is never unfair to ask Nature for assistance. In this spirit (following Schwinger) we observe that

> There is abundant observational evidence—none of which was known to Maxwell!—that *light is an electromagnetic phenomenon*, that *charge-motion can give rise to radiation*, that *in charge-free regions of spacetime the electromagnetic field equations must possess wave-like solutions*.

where the *wave equation*, is, we recall, a 2^{nd}-order partial differential equation of the form

$$\underbrace{\left\{\tfrac{1}{c^2}\tfrac{\partial^2}{\partial t^2}-\nabla^2\right\}}_{\substack{\text{the ``wave operator,'' sometimes called}\\ \text{the } d'Alembertian \text{ and denoted } \Box^2.}} f(\boldsymbol{x},t)=0$$

In charge-free regions of spacetime the equations in hand (set $\rho=0$ and $\boldsymbol{j}=\boldsymbol{0}$) read

$$\boldsymbol{\nabla}\cdot\boldsymbol{E}=0 \tag{61.1}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{B}=0 \tag{61.2}$$

$$\tfrac{1}{c}\tfrac{\partial}{\partial t}\boldsymbol{E}-\boldsymbol{\nabla}\times\boldsymbol{B}\quad\ =\boldsymbol{0} \tag{61.3}$$

$$\tfrac{1}{c}\tfrac{\partial}{\partial t}\boldsymbol{B}-\boldsymbol{\nabla}\times\underbrace{\left(\tfrac{\boldsymbol{v}}{c}\times\boldsymbol{B}\right)}_{\text{problematic term}}=\boldsymbol{0} \tag{61.4}$$

Application of $\frac{1}{c}\frac{\partial}{\partial t}$ to (61.3) gives

$$\tfrac{1}{c^2}\tfrac{\partial^2}{\partial t^2}\boldsymbol{E}-\boldsymbol{\nabla}\times\tfrac{1}{c}\tfrac{\partial\boldsymbol{B}}{\partial t}=\boldsymbol{0}$$

which by (61.4) becomes

$$\tfrac{1}{c^2}\tfrac{\partial^2}{\partial t^2}\boldsymbol{E}-\boldsymbol{\nabla}\times\left(\boldsymbol{\nabla}\times\left(\tfrac{\boldsymbol{v}}{c}\times\boldsymbol{B}\right)\right)=\boldsymbol{0}$$

Drawing now upon the general identity[19]

$$\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\boldsymbol{A})=\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{A})-\underbrace{\nabla^2\boldsymbol{A}}_{\substack{\text{means that } \nabla^2 \text{ acts separately}\\ \text{on each of the components of } \boldsymbol{A}}} \tag{62}$$

we obtain

$$\tfrac{1}{c^2}\tfrac{\partial^2}{\partial t^2}\boldsymbol{E}-\boldsymbol{\nabla}\left\{\boldsymbol{\nabla}\cdot\left(\tfrac{\boldsymbol{v}}{c}\times\boldsymbol{B}\right)\right\}+\nabla^2\left(\tfrac{\boldsymbol{v}}{c}\times\boldsymbol{B}\right)=\boldsymbol{0}$$

This would (by (61.1)) go over into the *vectorial wave equation*

$$\left\{\tfrac{1}{c^2}\tfrac{\partial^2}{\partial t^2}-\nabla^2\right\}\boldsymbol{E}=\boldsymbol{0} \tag{63}$$

provided we set

$$\boldsymbol{E}=-\tfrac{1}{c}(\boldsymbol{v}\times\boldsymbol{B}) \tag{64}$$

Equation (64) is, however, *unacceptable*: it contains—as did the rejected equation (55)—an objectionable allusion to $\boldsymbol{v}$ (and would, moreover, imply $\boldsymbol{E}\to\boldsymbol{0}$ as $\boldsymbol{v}\to 0$: we would be *out of business*!). But our objective—(63)—would in fact be realized if we assumed (64) to hold in the specific context afforded by (61.4). Thus are we led—tentatively—

[19] PROBLEM 14.

5) to *write* $\boldsymbol{\nabla}\times\bar{\boldsymbol{E}} = -\frac{1}{C}\frac{\partial}{\partial t}\bar{\boldsymbol{B}}$ *in place of* (61.4)≡ (58).

By this strategy we have, in effect, *short-circuited at first opportunity* the "infinite regress problem" which (in connection with the $\boldsymbol{C}$-field) was discussed earlier. But in so doing we have (as will emerge) also done much else.[20]

The field equations that emerge from the heuristic arguments just outlined are precisely **Maxwell's equations**

$$\boxed{\begin{aligned} \boldsymbol{\nabla}\cdot\boldsymbol{E} &= \rho && (65.1)\\ \boldsymbol{\nabla}\cdot\boldsymbol{B} &= 0 && (65.2)\\ \boldsymbol{\nabla}\times\boldsymbol{B} &= \tfrac{1}{C}\big(\boldsymbol{j} + \tfrac{\partial}{\partial t}\boldsymbol{E}\big) && (65.3)\\ \boldsymbol{\nabla}\times\boldsymbol{E} &= \tfrac{1}{C}\big(\boldsymbol{0} - \tfrac{\partial}{\partial t}\boldsymbol{B}\big) && (65.4) \end{aligned}}$$

Here as always, heuristically generated statements could be dismissed out of hand if it could be shown that they were internally inconsistent. Once that test is passed, we acquire the obligation to *to show that our statements conform to physical experience*. It is those two complementary lines of activity—especially the latter—that will absorb our energy in all the pages which follow.

It is—on methodological grounds, and in view of the preceding allusion to "physical experience"—interesting to notice that the heuristic arguments which led us from

$$\text{electrostatics} \longrightarrow \text{Maxwellian electrodynamics}$$

give rise to *observationally incorrect physics* when applied to Newtonian gravitostatics. Electrodynamics and gravitodynamics "share Coulomb's law in the static approximation" but otherwise differ profoundly. Owing to the equivalence principle, the gravitational field acts as its own source—*i.e.*, is "self-coupled"—and so must be governed by *non-linear* partial differential equations ... but *Maxwell's equations are <u>linear</u> partial differential equations*. The circumstance force to this obvious—but sometimes overlooked—point: *heuristic arguments cannot in general be promoted convincingly to the status of "proof;" by bootstrapping one <u>may infer but cannot expect to demonstrate</u> the "unique and necessary structure" of the enveloping theory*.

Maxwell's equations (+ boundary & initial data) enaable us in principle to compute the (generally *dynamical*) *electromagnetic* $(\boldsymbol{E},\boldsymbol{B})$*-fields which are generated by prescribed source activity* (the latter described by ρ and $\boldsymbol{j}$), but tell us nothing about the converse problem: *How do charged mass points move in response to prescribed ambient electromagnetic fields*? More sharply: *What—given* $\bar{\boldsymbol{E}}(\boldsymbol{x},t)$ *and* $\bar{\boldsymbol{B}}(\boldsymbol{x},t)$ *is the force* $\bar{\boldsymbol{F}}$ *experienced by a test charge* q *situated at* $(\boldsymbol{x},t)$? If we see q to be <u>at rest</u> then—by definition!—

$$\boldsymbol{F} = q\boldsymbol{E} \quad : \quad \text{note the } \boldsymbol{B}\text{-independence} \tag{66}$$

[20] In particular we have denied the universal validity of (56): $\boldsymbol{\nabla}\times\boldsymbol{E} = \boldsymbol{0}$. Note, however, that we recover (56) when $\boldsymbol{B}$ is t-independent: $\partial\boldsymbol{B}/\partial t = \boldsymbol{0}$.

. . . but *What if we see q to be in motion*? An inertial observer O' who sees q to be momentarily at rest would—since (66) is "shared physics"—write $\boldsymbol{F}' = q'\boldsymbol{E}'$. The question therefore arises:

> How do the quantities $(\boldsymbol{E}', \boldsymbol{B}', \ldots)$ seen by O' relate to the quantities $(\boldsymbol{E}, \boldsymbol{B}, \ldots)$ seen by another inertial observer O?

The surprising answer to this question shows that we should in the general case write

$$\boldsymbol{F} = q\big(\boldsymbol{E} + \tfrac{1}{c}\boldsymbol{v}\times\boldsymbol{B}\big) \tag{67}$$

($\boldsymbol{v}$: refers to our perception of q's instantaneous velocity)

The *Lorentz force law* (67) is of importance partly because it removes a problem which has been a source of embarrassment ever since we declared the $\boldsymbol{B}$-field to be "autonomous:" it makes possible an *operational definition* of $\boldsymbol{B}$.

The resolution of the transformation-theoretic question

$$(\boldsymbol{E}', \boldsymbol{B}', \ldots) \xrightarrow{\quad ? \quad} (\boldsymbol{E}, \boldsymbol{B}, \ldots)$$

posed above turns out to be "surprising" in this profoundly consequential sense: The heuristic arguments which *led* us to Maxwell's equations (65) drew strongly upon the Galilean transformation (47). But the Maxwell equations themselves are (as will be shown, and the statements (49) and (52) notwithstanding) *not Galilean covariant; i.e.*, they do not preserve their form under tha action of (47/48). It is attention to this critical point which will lead us—as historically it led Lorentz and Einstein—to the formulation of **special relativity**. One need only glance at the history of 20^{th} Century physics ($t \geqslant 1905$)—indeed: of 20^{th} Century civilization—to get a sense of how incredibly consequential a formal subtlety can be![21]

My objective in the next few sections will be to describe, if only in the most preliminary terms, some of the most *characteristic phenomenological consequences of Maxwell's equations.* Note in this connection that if in (65) we set $\boldsymbol{j} = \boldsymbol{B} = \boldsymbol{0}$ we obtain (as remarked already on page 31) the equations

$$\left.\begin{aligned} \boldsymbol{\nabla}\cdot\boldsymbol{E} &= \rho \\ \tfrac{\partial \boldsymbol{E}}{\partial t} = \boldsymbol{\nabla}\times\boldsymbol{E} &= \boldsymbol{0} \\ \therefore \tfrac{\partial \rho}{\partial t} &= 0 \end{aligned}\right\} \tag{69}$$

which were seen at (22) to be fundamental to *electrostatics*, and of which the phenomenological consequences were discussed in some (by no means exhaustive) detail already in §1.

[21] PROBLEMS 15 & 16

3. Current, and the principle of charge conservation. We begin by discussing a simple *corollary* of Maxwell's equations. If (recall PROBLEM 12) we construct $\frac{\partial}{\partial t}(65.1) + c\boldsymbol{\nabla}\cdot\textbf{(65.3)}$ we obtain

$$\tfrac{\partial}{\partial t}\rho + \boldsymbol{\nabla}\cdot\boldsymbol{j} = 0 \qquad (70 \equiv 51)$$

Equations of this particular structure are (for reasons which will emerge) called "continuity equations:" it is because we attach specific interpretations to ρ and $\boldsymbol{j}$ that (70) becomes the "charge conservation equation."

Important insight into the meaning of (70)—and of continuity equations generally—can be obtained as follows: Let $\mathcal{R}$ be a t-independent "bubble" in $\boldsymbol{x}$-space, and let $\rho(\boldsymbol{x},t)$ and $\boldsymbol{j}(\boldsymbol{x},t)$—fields which we shall assume to be in conformity with (70)—be given. The *total charge* $Q(t)$ *contained within* $\mathcal{R}$ can be described

$$Q(t) = \iiint_{\mathcal{R}} \rho(\boldsymbol{x},t)\, d^3x$$

Looking now to the *rate of temporal variation of* Q we have (see Figure 7)

$$\dot{Q} = \iiint_{\mathcal{R}} \frac{\partial\rho(\boldsymbol{x},t)}{\partial t}\, d^3x$$

NOTE: An additional term—describing the "rate at which $\mathcal{R}$ gobbles up charge"—would be required had we allowed $\mathcal{R}$ to be t-dependent.

$$= -\iiint_{\mathcal{R}} \boldsymbol{\nabla}\cdot\boldsymbol{j}\, d^3x \quad \text{by (70)}$$

$$= -\iint_{\partial\mathcal{R}} \underbrace{\boldsymbol{j}\cdot\boldsymbol{dS}}_{\substack{\text{—}\textit{charge flux}\text{ through the surface element } \boldsymbol{dS}.\\ \text{Since } \boldsymbol{dS} \text{ is "outward directed," we have}}} \quad \text{by Gauß' theorem: (14)}$$

$$= -\{\text{total flux outward through the surface } \partial\mathcal{R} \text{ of } \mathcal{R}\}$$

The implication is that *no "birth" or "death" processes contribute to* $\dot{Q}$ *...* which is what we mean when we say that "charge is conserved." The generality of the argument follows from the observation that it works whenever

- ρ is a *density* and
- $\boldsymbol{j}$ is the corresponding *flux density*.

We see that

$$\tfrac{d}{dt}\iiint_{\mathcal{R}} \rho(\boldsymbol{x},t)\, d^3x + \iint_{\partial\mathcal{R}} \boldsymbol{j}(\boldsymbol{x},t)\cdot\boldsymbol{dS} = 0 \qquad (71)$$

expresses globally the information which (70) expresses locally. From the requirement that (71) hold for *all* t-independent bubbles $\mathcal{R}$ one can in fact *recover* (70).

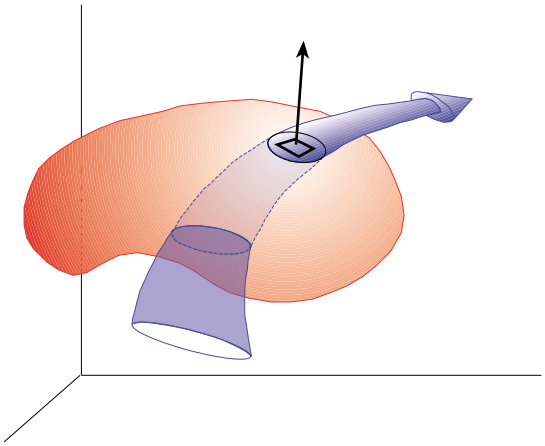


FIGURE 7: *In a spatial region occupied by a drifting charge cloud* $\boldsymbol{j}(\boldsymbol{x},t)$*—represented here by the fat blue arrow—the argument on the preceding page asks us to designate a "bubble"* $\mathcal{R}$*, and to identify the rate of change of enclosed charge with the rate at which charge is transported into* $\mathcal{R}$ *through its surface* $\partial\mathcal{R}$*.*

What, specifically, is the meaning of the statement that $\boldsymbol{j}$ is by nature a measure of (electric) "flux density"? It follows from (70)—whence ultimately from (65)—that

$$\begin{aligned}[\boldsymbol{j}\,] = [\rho\,] \cdot \text{velocity} &= \text{charge} \cdot \text{velocity/volume} \\ &= \text{charge/area} \cdot \text{time}\end{aligned}$$

We infer that

$$\boldsymbol{j}(\boldsymbol{x},t)\cdot\boldsymbol{dS} = \left\{\begin{array}{l}\text{instantaneous rate (at time } t\text{) at which} \\ \text{charge is being transported through a} \\ \text{little "window" } \boldsymbol{dS} \text{ situated at position } \boldsymbol{x}\end{array}\right.$$

Recalling the definition of "$\cdot$" we have

$$= j \cdot dS \cdot \cos\theta$$

$\theta \equiv$ angle between $\hat{\boldsymbol{j}}$ and $\boldsymbol{S}$; *i.e.*, the window's "presentation angle"

It is important to appreciate that the $\boldsymbol{j}$ here under consideration is a more general conception than the $\boldsymbol{j} \equiv \rho\,\boldsymbol{v}$ contemplated at (50) . The latter is literally appropriate only if the charge which flows through the window $\boldsymbol{dS}$ does so coherently—as a unitary entity endowed with a single, well-defined velocity—while the $\boldsymbol{j}$ contemplated in (70) refers only to the *effective mean drift* of the charge at $(\boldsymbol{x},t)$. The distinction is illustrated in Figure 8.

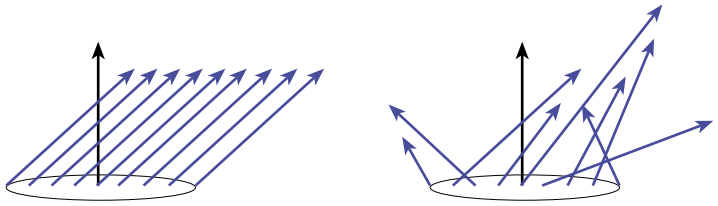


FIGURE 8: *Charges q stream through an inspection window with identical velocities (or "coherently") at left, and with statistically distributed velocities ("incoherently") at right.*

The fields ρ and $\boldsymbol{j}$ pertain most naturally to "continuum physics," and have to strain a bit to accommodate the microscopic physical fact that charge always rides around on localized bits of matter. Let $\boldsymbol{x}(t)$ and $\boldsymbol{v}(t) \equiv \dot{\boldsymbol{x}}(t)$ describe the motion of a charged mass point (m, q). To describe the associated ρ and $\boldsymbol{j}$ we might write

$$\begin{aligned} \rho(\boldsymbol{x}, t) &= q\,\delta(\boldsymbol{x} - \boldsymbol{x}(t)) \\ \boldsymbol{j}(\boldsymbol{x}, t) &= q\,\delta(\boldsymbol{x} - \boldsymbol{x}(t))\,\boldsymbol{v}(\boldsymbol{x}, t) \end{aligned} \tag{72}$$

These singular fields—can you show that they satisfy (70)?—acquire the correct physical dimensionality from the circumstance that

$$[\delta(\boldsymbol{x})] = (\text{volume})^{-1}$$

The restrictive equation

$$\tfrac{\partial}{\partial t}\rho = 0 \tag{73}$$

is familiar from §1, where it was interpreted as referring to "charges that don't move." If, however, we reflect upon the meaning of its mate

$$\tfrac{\partial}{\partial t}\boldsymbol{j} = \boldsymbol{0} \tag{74}$$

we see that (73) admits of a more relaxed interpretation. For (74) *requires that the charges move, but in such a stereotyped manner that they keep replacing each other*. Phrased another way: $\frac{\partial}{\partial t}\rho = 0 \Rightarrow \boldsymbol{\nabla}\cdot\boldsymbol{j} = 0$ (by (70)), but does not, of itself, require $\boldsymbol{j} = \boldsymbol{0}$. Equations (73/74) may be satisfied momentarily, "accidentally," at isolated spacetime points, but tend to be of practical importance only when they hold *globally*. Source fields $\rho(\boldsymbol{x})$ and $\boldsymbol{j}(\boldsymbol{x})$ are—for the reason just mentioned—best described not as "static" but as t-independent or *steady*. In view of the fact that it is so easy to build steady ρ-fields with an isolated point charge, it becomes interesting to note that (except in the trivial sense $\boldsymbol{j} = \boldsymbol{0}$) one *cannot build a steady $\boldsymbol{j}$-field with a single charge*. Interesting to note also that the reason appears to be not logical, not electrodynamical . . . but (see Figure 9) *mechanical*: one runner can't (in the continuous limit) "keep running by with velocity $\boldsymbol{v}$"—even if the racetrack is infinitely short. The

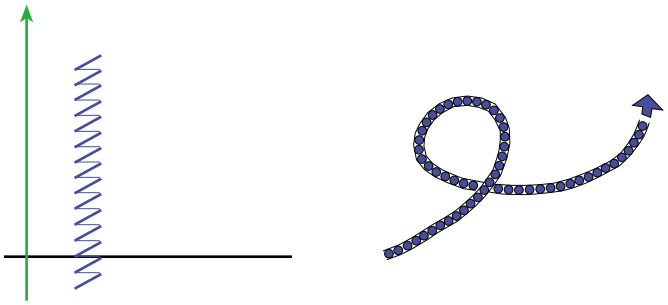


FIGURE 9: *At left, a single charge attempts in vain to "keep running past" an inspection point. At right, entrained charges achieve the intended effect (production of a steady* $\boldsymbol{j}$*) by serially replacing one another. But even with the latter arrangement we cannot produce a steady* $\boldsymbol{j}$ *which vanishes everywhere except at a point.*

problem would disappear if Nature provided not only point charges but true *line charges* (charged strings). Absent those, we are forced to build our steady $\boldsymbol{j}$-fields with the aid of *entrained point charges*: we "glue charges on a string, pull the string ... and pretend not to notice the microscopic granularity." The operation (see again the preceding figure) is most commonly called "sending a current through a wire."

In many practical contexts—particularly those which arise from engineering—it is more common to speak of the *current* I than of the current density $\boldsymbol{j}$. These concepts are related as follows: Let $\mathcal{D}$ be (topologically equivalent—see Figure 10—to) a "disk," and let $\partial\mathcal{D}$ denote its boundary (a closed curve). Given $\boldsymbol{j}(\boldsymbol{x}, t)$, we form

$$I(t;\mathcal{D}) \equiv \iint_D \boldsymbol{j}\cdot\boldsymbol{dS} \tag{75}$$

to obtain a measure of the *instantaneous rate at which charge is (at time t) being transported through* $\mathcal{D}$*, i.e.*, of the total charge flux through $\mathcal{D}$. Evidently

$$[\,I\,] = [\,\boldsymbol{j}\,]\cdot\text{area} = \text{charge/time}$$

Engineers perefer to measure currents I in *Amperes* $\equiv$ *Coulombs/second*. Note that on a disk $\boldsymbol{dS}$ is sign-ambiguous ("outside" being undefined). A disk $\mathcal{D}$ endowed with a sign convention is said to be "oriented." Evidently we are, for the purposes of (75), obliged to require that $\mathcal{D}$ be orient*able*: no Möbius strips allowed! For a given $\boldsymbol{j}$-field one expects to have

$$I(t;\mathcal{D}_1) \neq I(t;\mathcal{D}_2) \quad \text{even when} \quad \partial\mathcal{D}_1 \equiv \partial\mathcal{D}_2$$

It is therefore of some interest that one can show without difficulty[22] that if ρ is steady and if, moreover, $\mathcal{D}_1$ and $\mathcal{D}_2$ share the same boundary ($\partial\mathcal{D}_1 \equiv \partial\mathcal{D}_2$),

[22] PROBLEM 17.

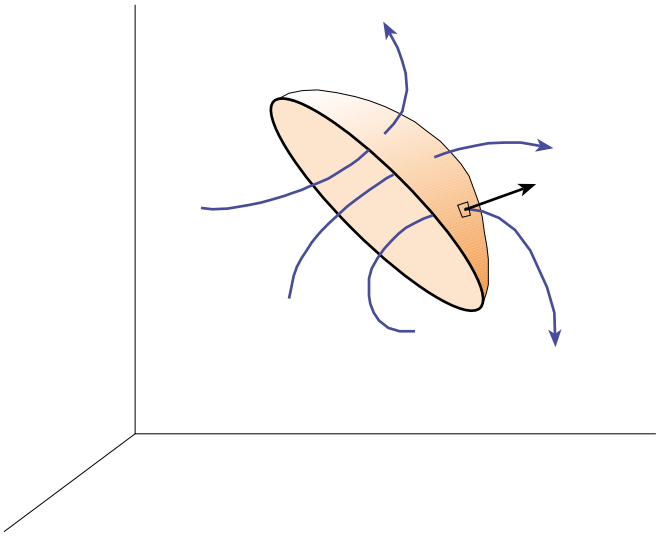


FIGURE 10: *The net current through a hypothetical cap $\mathcal{D}$ is, according to* (75), *found by integrating $\boldsymbol{j}_\perp$—the normal component of current density, the vector field represented here by (blue) directed curves.*

then $I(t;\mathcal{D}_1) = I(t;\mathcal{D}_2)$. In such (frequently encountered) contexts there is no reason to retain any explicit allusion either to t or to $\mathcal{D}$: it becomes more natural to write $I(\partial\mathcal{D})$, and then to drop the $\partial\mathcal{D}$ as "obviouis from the context." Thus does one acquire the privilege of referring simply to "the current I."

To describe the *current I in a wire*[23] we have only to suppose (see Figure 11) that $\partial\mathcal{D}$ circumscribes the wire. Phenomenologically, the current I in a wire of cross-sectiuonal area A can be described

$$I = nevA \quad \begin{cases} n \text{ is the number of charge carriers per unit volume} \\ e \text{ is the charge per carrier (= elecronic charge)} \\ v \text{ is the mean drift velocity of the charge carriers} \end{cases}$$

People are often surprised to discover that (because n is typically quite large) v is typically quite small. If in (76) we assign I, n, e and A the values appropriate to a *1 amp current in a 14-gauge copper wire* (radius $R = 0.0814$ cm) we find that the drift velocity $v = 3.55\times10^{-3}$ cm/sec: evidently *the physics of electrical signal propagation has very little to do with the physics of charge carrier drift*.

It will be appreciated that the currents encountered in Nature, and of fundamental interst to physicists, are for the most part *not confined to wires*

[23] We agree here to overlook the "charge accumulation effects" which may arise at high frequencies: we agree, in other words, to "think DC."

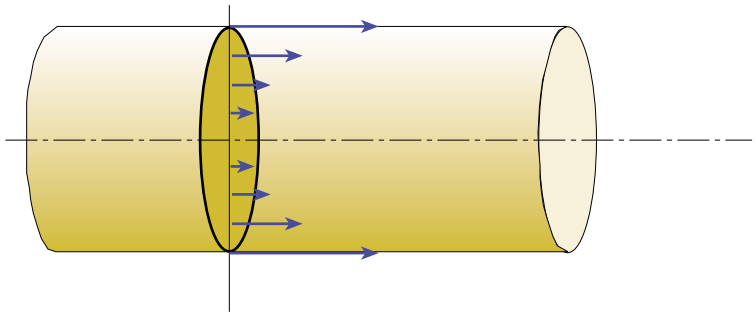


FIGURE 11: *Variant of the preceding figure, adapted to the problem of evaluating the current in a wire. The "cap" has in this instance become simply a cross-section of the wire. The presumption in the figure is that* $\boldsymbol{j}$ *is axially symmetric but non-uniform, being strongest near the "skin" of the wire.*

. . . wires—and nerves—being "rare objects" in the universe. "Unconfined currents" are found in (for example) lightning bolts and throughout the natural world, and in some engineering applications (arcs welders, vacuum tubes, electrochemical process vats, particle accelerators).

Returning now to more theoretical matters . . . the *interests of symmetry* would clearly be served if in place of (65) one had

$$\boldsymbol{\nabla}\cdot\boldsymbol{E} = \rho \quad , \quad \boldsymbol{\nabla}\times\boldsymbol{B} = +\tfrac{1}{c}\big(\boldsymbol{j} \quad + \tfrac{\partial \boldsymbol{E}}{\partial t}\big) \tag{77.1}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{B} = \rho_m \, , \quad \boldsymbol{\nabla}\times\boldsymbol{E} = -\tfrac{1}{c}\big(\boldsymbol{j}_m + \tfrac{\partial \boldsymbol{B}}{\partial t}\big) \tag{77.2}$$

where the subscript $_m$ means "magnetic." Then the argument which when applied to (77.1) gave

$$\tfrac{\partial}{\partial t}\rho \quad + \boldsymbol{\nabla}\cdot\boldsymbol{j} \quad = 0 \tag{70}$$

would when applied to (77.2) give

$$\tfrac{\partial}{\partial t}\rho_m + \boldsymbol{\nabla}\cdot\boldsymbol{j}_m = 0$$

We would, in charge-free regions, still have $\Box^2\boldsymbol{E} = \Box^2\boldsymbol{B} = \boldsymbol{0}$, etc. and all would be well. From this point of view *the actual structure of Maxwell's equations (65) is seen to contain an informative surprise*: (65.2)—$\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0$—states in effect that "point magnetic charges" or (as they are called)

$$\text{"magnetic monopoles" } \ldots \underline{\text{do not exist}} \tag{78}$$

It is, therefore, not at all surprising that the $\boldsymbol{j}_m$ term is absent from (65.4); *i.e.*, that "neither do magnetic currents exist." Several questions—particularly since Maxwell's equations seem in this respect to be so "permissive"—now arise:

1. *Is* (78) *factually true*? The answer must come from the laboratory. Many sophisticated searches have thus far failed to detect a single magnetic monopole. Such activity continues (if as a fairly low-priority item), and reports of the results are always received with interest by the international community of physicists.[24]

2. Is it possible that magnetic *monopoles exist but cannot be observed*? A weak instance of such a situation would arise if monopoles were bound to oppositely charged monopoles by forces so strong that they *cannot be feasibly dislodged and studied in isolation*. The "strong" interpretation—that "monopoles exist but cannot be observed *in principle*—would appear to *strike at the philosophical foundations of science*, to be latently "unscientific" . . . unless it were argued that monopoles (like quarks?) *announce themselves not in isolation but indirectly—by their effects*.

3. Physicists have come to adhere generally—if informally—to the view that

"all which is not forbidden is mandatory"

This heuristic principle suggests that monopoles—if *not* forbidden (by some yet-undiscovered conservation law?)—will eventually (by their direct or indirect effects) be detected, and in the contrary case gives rise to this sharp question: *What (presently unknown) principle effectively "forbids" the existence of magnetic monopoles*?

There is (as will emerge in a subsequent chapter) an interesting—if but little-known—sense in which (78) misrepresents the physical situation: (78) expresses not a fact (?) but a "fact wrapped in a convention." If the (elementary) particles found in Nature carried magnet charge p as well as electric charge q, then to describe the compound charge structure (q, p) of a particle population one might present something like the topmost of the following figures. It is, however, a surprising fact of Nature that (central figure) the observed points lie *on a line; i.e.*, that

$$p/q \equiv \tan\theta$$

has a value shared by all known elementary particles. *This* is the elemental fact which awaits explanation. It is by (seldom remarked) operational *convention*

[24] . . . Particularly since, in the two decades since this paragraph was written, it has become the clear tendency of theoretical developments (in elementary particle physics, cosmology) to *demand* the physical existence of magnetic monopoles!

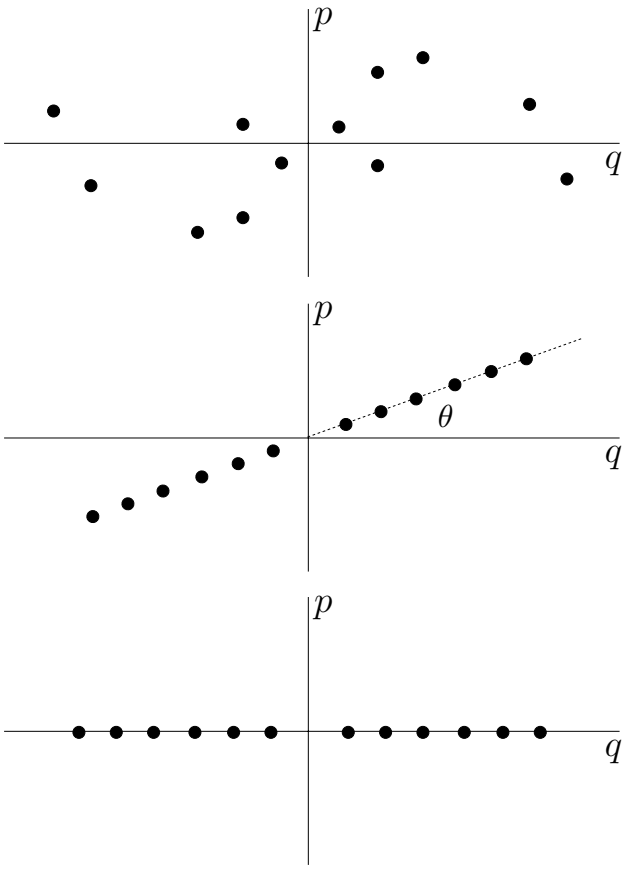


FIGURE 12: *At top: the kind of* (q,p)*-distribution that one naively might expect to encounter in Nature. In the center: the distribution one in fact encounters. At bottom:* θ*-rotational invariance has been used to eliminate the magnetic components from all* (q,p)*-pairs. The surprising fact is that the same rotation works in all cases.*

that we have—essentially by (67): $\boldsymbol{F} = q\big(\boldsymbol{E} + \frac{1}{c}\boldsymbol{v}\times\boldsymbol{B}\big)$—set $\theta = 0$ (bottom figure). It is, in other words, *by convention that we have associated the observed "*(q,p)*-line" with the "electric axis" in "charge space."* Later I will have occasion to discuss the deep formal symmetry (θ-rotational invariance) of Maxwell's equations which permits one to *exercise* such an option.

I turn finally to some historical points. It is roughly—but *only* roughly—correct to state that at some point in the later developmental stages of his electrodynamical work Maxwell realized that

- GAUSS' LAW (of which Coulomb's law is a corollary, and which is not to be confused with Gauß' theorem) can be rendered

$$\boldsymbol{\nabla}\cdot\boldsymbol{E} = \tfrac{1}{\epsilon_0}\rho \tag{79.1}$$

- The NON-EXISTENCE OF MAGNETIC MONOPOLES can be rendered

$$\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0 \tag{79.2}$$

- AMPERE'S LAW[25] can be rendered

$$\boldsymbol{\nabla}\times\boldsymbol{B} = \mu_0\boldsymbol{j} \tag{79.3}$$

- FARADAY'S LAW[26] can be rendered

$$\boldsymbol{\nabla}\times\boldsymbol{E} = -\tfrac{\partial\boldsymbol{B}}{\partial t} \tag{79.4}$$

Here ϵ_0 and μ_0 are empirical constants made necessary by the fact that Maxwell —working close to laboratory experience—used practical units to quantify electromagnetic variables ... while the subscript $_0$ reflects his tendency (born of the same circumstance) to view the vacuum as "degenerate matter." Maxwell realized more particularly that equations (79), while they account for most of the phenomenology known to him, cannot be correct ... for it follows from (79.3) that $\boldsymbol{\nabla}\cdot\boldsymbol{j} = 0$, which conforms to (70), *i.e.*, to the principle of charge conservation, *only in the steady case*.[27] Thus was Maxwell led at length to propose—on no direct observational evidence!—that in place of (79.3) one should write[28]

$$\boldsymbol{\nabla}\times\boldsymbol{B} = \mu_0\big(\boldsymbol{j} + \epsilon_0\tfrac{\partial\boldsymbol{E}}{\partial t}\big) \tag{80}$$

> REMARK: The new term $\epsilon_0\frac{\partial\boldsymbol{E}}{\partial t}$ —because it enters as a kind of companion to $\boldsymbol{j}$—was called by Maxwell the "displacement current." It's introduction represents a somewhat spooky modification of (79.3), for *it involves no charge motion*. Oddly, Maxwell felt no obligation to attach a similar name/interpretation to the $\frac{\partial\boldsymbol{B}}{\partial t}$-term in Faraday's law (79.4).

The first writing of (80) was, in my view, one of the most seminal events in 19th Century physics: indeed, in the entire history of physics. For it gave rise —automatically—to a fully detailed *electromagnetic theory of light* ... and thus by implication to relativity, quantum mechanics and all that follows therefore. How did this come about?

---

[25] Describes the $\boldsymbol{B}$-fields generated by steady currents...of which more later.

[26] Describes the $\boldsymbol{E}$-fields generated by changing $\boldsymbol{B}$-fields...of which again: more later.

[27] This is hardly surprising in view of the fact that *Ampere's law was abstracted from steady-case observations*.

[28] PROBLEM 20.

In charge-free regions of space Maxwell's own equations (79)—as modified —read[29]

$$\boldsymbol{\nabla}\cdot\boldsymbol{E} = 0 \tag{81.1}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0 \tag{81.2}$$

$$\boldsymbol{\nabla}\times\boldsymbol{B} = \mu_0\epsilon_0\frac{\partial\boldsymbol{E}}{\partial t} \tag{81.3}$$

$$\boldsymbol{\nabla}\times\boldsymbol{E} = \quad -\frac{\partial\boldsymbol{B}}{\partial t} \tag{81.4}$$

REMARK: From (81.4) we see that in Maxwell's units

$$[\,\boldsymbol{E}\,] = \text{velocity}\cdot[\,\boldsymbol{B}\,]$$

It follows therefore from (81.3) that

$$[\,\mu_0\epsilon_0\,] = (\text{velocity})^{-2}$$

Equations (81) can be "separated by differentiation,"[30] giving

$$\left(\mu_0\epsilon_0\tfrac{\partial^2}{\partial t^2} - \nabla^2\right)\boldsymbol{E} = \left(\mu_0\epsilon_0\tfrac{\partial^2}{\partial t^2} - \nabla^2\right)\boldsymbol{B} = \boldsymbol{0}$$

It was the observation that the measured values of μ_0 and ϵ_0 entail

$$\tfrac{1}{\sqrt{\mu_0\epsilon_0}} \approx 3\times 10^{10}\ \text{cm/sec}$$

which led Maxwell (1863) to write that "we can scarcely avoid the inference that *light consists of undulations in the medium which is the cause of electric and magnetic phenomena*" [his italics]. This was an idea which had fallen out of the blue into Maxwell's lap, but in which he obviously had great confidence[31] . . . though it remained *merely* an idea at his death, and for eight years thereafter: electromagnetic radiation was first generated/detected by H. Hertz in 1887.

HISTORICAL REMARK: The excitement of discovery experienced by Maxwell would today be impossible . . . because the upshot of his discovery has—by recent international convention—been made a cornerstone of physical metrology:

$$\mu_0 = 4\pi\times 10^{-7} = 12.566370614\ldots\ \text{N A}^{-1}$$

and

$$c = 299792458\ \text{m s}^{-1}$$

are both now held to be *exact*, and

$$\epsilon_0 = (\mu_0 c^2)^{-1}\ \underline{\text{by modern definition}}!$$

---

[29] Note that the right side of (81.3) was introduced by Maxwell to salvage an equation—(70)—which is *not even relevant* in charge-free space!

[30] The procedure was encountered already in PROBLEM 15.

[31] See Maxwell's curiously understated discussion in Chapter XX of his *Treatise on Electricity & Magnetism*. C. W. F. Everitt, in his *James Clerk Maxwell: Physicist & Natural Philosopher* (1975), provides a good brief account of the history of Maxwell's discovery, and on page 101 reports that in 1865 Maxwell "... in a rare moment of unveiled exuberance to wrote to [a friend that] 'I have also a paper afloat, containing an electromagnetic theory of light, which, till I am convinced to the contrary, I hold to be great guns.' "

I draw attention finally to one formal point which was only recently brought to light,[32] but which I find to be of deep interest. If in (81.3) we set $\mu_0\epsilon_0 = 1/c^2$ and then proceed to the limit $c \uparrow \infty$ we find that one and only one thing happens: the $\frac{\partial \boldsymbol{E}}{\partial t}$-term, which Maxwell was at such pains to introduce, is *extinguished*! We recover precisely the charge-free version of (79). What Jean-Marc Lévy-Leblond was evidently the first to notice (1967) is that (as the reader may verify) the equations

$$\begin{aligned} \boldsymbol{\nabla}\cdot\boldsymbol{E} &= 0 \\ \boldsymbol{\nabla}\cdot\boldsymbol{B} &= 0 \\ \boldsymbol{\nabla}\times\boldsymbol{B} &= \boldsymbol{0} \\ \boldsymbol{\nabla}\times\boldsymbol{E} &= -\tfrac{\partial \boldsymbol{B}}{\partial t} \end{aligned}$$

are covariant with respect to[33] the following extension

$$\left.\begin{array}{l} t \longmapsto t' = t \\ \boldsymbol{x} \longmapsto \boldsymbol{x}' = \boldsymbol{x} - \boldsymbol{v}t \\ \frac{\partial}{\partial t} \longmapsto \frac{\partial}{\partial t'} = \frac{\partial}{\partial t} + \boldsymbol{v}\cdot\boldsymbol{\nabla} \\ \boldsymbol{\nabla} \longmapsto \boldsymbol{\nabla}' = \boldsymbol{\nabla} \\ \boldsymbol{E} \longmapsto \boldsymbol{E}' = \boldsymbol{E} + \boldsymbol{v}\times\boldsymbol{B} \\ \boldsymbol{B} \longmapsto \boldsymbol{B}' = \boldsymbol{B} \end{array}\right\} \tag{82}$$

of what at (47/48) we meant by a "Galilean boost." We conclude that—though the point was not appreciated by Maxwell himself—"Maxwell's trick" enforced the *abandonment of Galilean relativity, & the adoption of Einsteinian relativity*.

I hope readers will by now understand why it seems to me not entirely frivolous to suggest that "20$^{\text{th}}$ Century physics is a *grandchild of the principle of charge conservation*" ... or, more precisely, of the *symmetry principle* of which charge conservation is the physical manifestation.

4. Generation of B-fields: Ampere's law. Having reviewed already the "physical upshot" of $\boldsymbol{\nabla}\cdot\boldsymbol{E} = \rho$, $\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0$ and $\partial\rho/\partial t + \boldsymbol{\nabla}\cdot\boldsymbol{j} = 0$ we turn now to a similarly preliminary discussion of the physical significance of the statement

$$\boldsymbol{\nabla}\times\boldsymbol{B} = \tfrac{1}{c}\big(\boldsymbol{j} + \tfrac{\partial \boldsymbol{E}}{\partial t}\big) \qquad (65.3 \equiv 80 \equiv 83)$$

↑ Maxwell's stroke of genius

... which is, unlike the statements studied previously, *vector*-valued.

---

[32] See Section V, §C.1 (page 267) of J-M. Lévy-Leblond's "Group theory and Galilean invariance" in E. M. Loebl (ed.), *Group Theory and Its Applications II* (1971). Also §2 in P. G. Bergmann, "The special theory of relativity" in Volume IV of *Handbüch der Physik* (1962).

[33] Compare (48).

Hans Christian Oersted had expressed his intuitive conviction that "electricity & magnetism must be interrelated" already in 1812 (in his *View of Chemical Laws*), but it was during (!) a lecture in the spring of 1820 that he discovered "electromagnetism;" *i.e.*, that *electric currents give rise to magnetic fields*. Oersted's discovery immediately engaged the excited attention of the leading scientists of the day (J. B. Biot, F. Savant, H. Davy, the young M. Faraday, ...), and when (11 September 1820) Oersted repeated his lecture/ demonstration before members of the Académie des Sciences it came to the attention of André Marie Ampere (1775–1836). I mention these facts partly in order to suggest that it is somewhat inappropriate that we associate with (83) the name of Ampere, rather than that of Oersted. Since Ampere's own work was concerned mainly with the *magnetic interaction of currents* (*i.e.*, of current-carrying wires), it might more appropriately be attached to what we now call the "Biot-Savart law" (see below).

The experimental work to which I have just referred involved steady currents—made possible by Volta's then-recent invention (1800) of the *voltaic cell*.[34] When the sources (whence also their associated fields) are steady the $\frac{\partial \boldsymbol{E}}{\partial t}$-term drops away from (83) and we obtain

$$\boldsymbol{\nabla}\times\boldsymbol{B} = \tfrac{1}{c}\boldsymbol{j} \tag{84}$$

It is with the phenomenological implications (not of (83) but) of (84) that will mainly concern us in the paragraphs which follow. And it is the analytical problem posed by equations of the form (84) that motivates the following

MATHEMATICAL DIGRESSION

A population of elementary theorems of exceptional beauty and power (which could, until rather recently, have been described as "well known to every student of analytical geometry") follows from the idea developed in

> **Problem 21.** Show that the area A of a triangle (012), which is oriented and coordinatized as indicated in the following Figure 13, can be described
>
> $$A(012) = \tfrac{1}{2}\begin{vmatrix} 1 & x_0 & y_0 \\ 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \end{vmatrix} \tag{85}$$
>
> This can be accomplished in many ways: you might, for example, try assembling the triangle from simpler triangles, then drawing upon elementary propeties of determinants. Note that (85) refers

[34] Alessandro Volta' work in this area was stimulated by Luigi Galvani's famous chance observation (1791) that electrical discharges caused the legs of dead frogs, laid out for dissection, to twitch. We may therefore add a frog to the items which already repose in (page 3) our little "museum of biogenic relics."

implicitly to an *orientation convention*, which supplies

$$\begin{aligned} A(012) = \quad A(120) = \quad A(201) \\ = -A(021) = -A(210) = -A(102) \end{aligned}$$

and tells us that $A \gtrless 0$ according as the triangle is right or left handed ($\circlearrowleft$ or $\circlearrowright$).

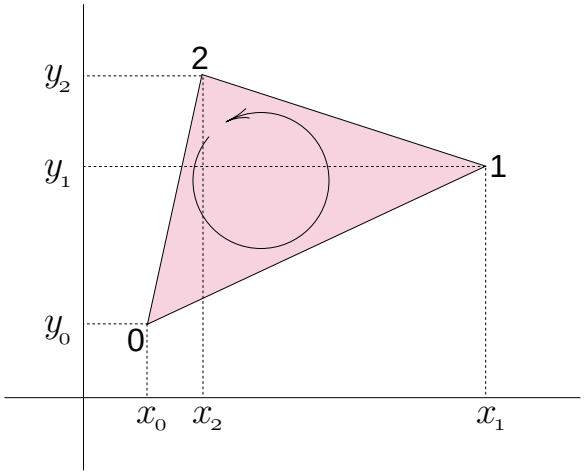


FIGURE 13: *Labeled geometrical construction used in* PROBLEM 21 *to establish the the determinantal description of area.*

Drawing now upon (85) we infer that the *area* $A(\mathcal{D})$ *of an arbitrary plane region* $\mathcal{D}$ (see Figure 14) can be described either

$$A(\mathcal{D}) = \iint_{\mathcal{D}} dxdy \tag{86.1}$$

or

$$= \text{limit of sum of areas } \underbrace{\tfrac{1}{2}\begin{vmatrix} 1 & 0 & 0 \\ 1 & x & y \\ 1 & x+dx & y+dy \end{vmatrix}}_{} \text{ of triangular slivers}$$

But

$$= \tfrac{1}{2}(x\,dy - y\,dx)$$

so

$$= \tfrac{1}{2}\oint_{\partial\mathcal{D}} (x\,dy - y\,dx) \tag{86.2}$$

According to (86.2) one can compute area by operations that are restricted to the boundary $\partial\mathcal{D}$ of the region $\mathcal{D}$ in question. This surprising fact provides the

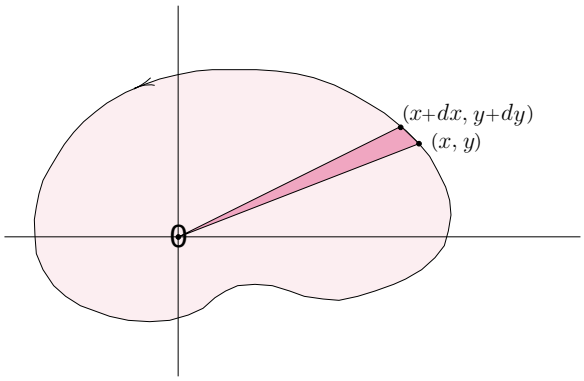


FIGURE 14: *Computing areas by adding triangular slivers—the geometrical basis of* (86.2).

operating principle of the *polar planimeter*—a wonderful device used mainly by architects and engineers.[35]

Let the preceding construction be considered now to be inscribed on the xy-plane in Euclidean 3-space, and let us agree that x, y and z refer henceforth to a *right*-handed frame. Readers will find it very easy to verify that (86)—thus situated—can be formulated

$$\iint_{\mathcal{D}} (\boldsymbol{\nabla}\times\boldsymbol{A})\cdot\boldsymbol{dS} = \oint_{\partial\mathcal{D}} \boldsymbol{A}\cdot\boldsymbol{d\ell} \tag{87}$$

provided we set

$$\boldsymbol{d\ell} = \begin{pmatrix} dx \\ dy \\ 0 \end{pmatrix}$$

$$\boldsymbol{A} = \begin{pmatrix} -y \\ +x \\ 0 \end{pmatrix}$$

$$\boldsymbol{dS} = \begin{pmatrix} 0 \\ 0 \\ dxdy \end{pmatrix}$$

This result provides an instance—and its derivation provides some insight into the proof—of *Stokes' theorem*, according to which (87) holds *generally . . . for <u>all</u> (even non-flat) disks $\mathcal{D}$ in 3-space, and for <u>all</u> vector fields $A(x, y, z)$*.

[35] For a masterful discussion of this pretty topic see Felix Klein's *Elementary Mathematics from an Advanced Standpoint: Geometry* (1925), pages 11–15. The subject is treated also on pages 57–65 of Chapter I in my QUANTUM TOPICS (2000).

It is evident that (see again pages 13 & 19) the following statements

$$\text{GAUSS} \quad : \quad \iiint_{\mathcal{R}} \text{div}\boldsymbol{A}\, d^3x = \iint_{\partial\mathcal{R}} \boldsymbol{A}\cdot\boldsymbol{dS}$$

$$\text{STOKES} \quad : \quad \iint_{\mathcal{D}} \text{curl}\boldsymbol{A}\cdot\boldsymbol{dS} = \oint_{\partial\mathcal{D}} \boldsymbol{A}\cdot\boldsymbol{d\ell}$$

$$\text{NEWTON} \quad : \quad \int_{\mathcal{C}} \text{grad}\,\varphi\cdot\boldsymbol{d\ell} = \varphi\Big|_{\text{endpoints of } \mathcal{C}}$$

are "of a type." They originate in the work of many 19th Century physicist-mathematicians (Gauß, Green, Kelvin, Tait, Maxwell, Cauchy, Stokes, ...), and have come to bear collectively the name of George Gabriel Stokes (1819–1903) for curious reasons that are explained on page *viii* of M. Spivak's *Calculus on Manifolds* (1965: see particularly the cover illustration!). Such identities were first studied *in unified generality* by H. Poincaré (1887), whose work was deepened and given its modern formulation—of which more later—mainly by Élie Cartan (∼1920).[36] "Stokes' theorems" are available even on n-dimensional non-Euclidean manifolds (where there are n such things), and all share the design

$$\int_{\text{region}} \text{differentiated object} = \int_{\text{boundary of region}} \text{undifferentiated object}$$

foreshadowed already in the

$$\text{FUNDAMENTAL THEOREM OF THE CALCULUS} \quad : \quad \int_a^b f'x\,dx = f(b) - f(a)$$

END OF DIGRESSION

Just as

$$\boldsymbol{\nabla}\cdot\boldsymbol{E} = \rho \tag{65.1}$$

and

$$\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0 \tag{65.2}$$

give rise by Gauß' theorem (14) to

$$\iint_{\partial\mathcal{R}} \boldsymbol{E}\cdot\boldsymbol{dS} = \iiint_{\mathcal{R}} \rho\, d^3x \tag{90.1}$$

$$\Updownarrow$$

total "electric flux" through $\partial\mathcal{R}$ = total charge interior to $\mathcal{R}$

and

$$\iint_{\partial\mathcal{R}} \boldsymbol{B}\cdot\boldsymbol{dS} = 0 \tag{90.2}$$

[36] PROBLEMS 22, 23 & 24. Equations (88) and (89)—absent from the text—appear in the first two of those exercises.

so do

$$\boldsymbol{\nabla}\times\boldsymbol{B} = \tfrac{1}{c}\big(\boldsymbol{j} + \tfrac{\partial \boldsymbol{E}}{\partial t}\big) \tag{65.3}$$

and

$$\boldsymbol{\nabla}\times\boldsymbol{E} = \tfrac{1}{c}\big(\boldsymbol{0} - \tfrac{\partial \boldsymbol{B}}{\partial t}\big) \tag{65.4}$$

give rise by Stokes' theorem (87) to

$$\begin{aligned}\int_{\partial\mathcal{D}} \boldsymbol{B}\cdot\boldsymbol{d\ell} &= \tfrac{1}{c}\iint_{\mathcal{D}} (\boldsymbol{j} + \tfrac{\partial \boldsymbol{E}}{\partial t})\cdot\boldsymbol{dS} \\ &= \tfrac{1}{c}\Big\{ I + I_d \Big\}\end{aligned} \tag{90.3}$$

and

$$\int_{\partial\mathcal{D}} \boldsymbol{E}\cdot\boldsymbol{d\ell} = -\tfrac{1}{c}\iint_{\mathcal{D}} \tfrac{\partial \boldsymbol{B}}{\partial t}\cdot\boldsymbol{dS} \tag{90.4}$$

where

$$I \equiv \iint_{\mathcal{D}} \boldsymbol{j}\cdot\boldsymbol{dS}$$

is the *conduction current* through $\mathcal{D}$, and

$$\begin{aligned} I_d &\equiv \iint_{\mathcal{D}} \tfrac{\partial \boldsymbol{E}}{\partial t}\cdot\boldsymbol{dS} \\ &= \tfrac{\partial}{\partial t}\underbrace{\iint_{\mathcal{D}} \boldsymbol{E}\cdot\boldsymbol{dS}}_{\text{“electric flux”}} \quad \text{if } \mathcal{D} \text{ is not itself moving}\end{aligned}$$

is the *displacement current*.[37]

Equations (90) comprise the so-called "integral formulation of Maxwell's equations," and in some conceptual/computational contexts—particularly those which possess a high degree of symmetry—prove more directly informative than their differential counterparts (65).

> **Example.** *What is the $\boldsymbol{E}$-field generated by a static point charge q?* Let the "Gaussian pillbox" $\mathcal{R}$ be spherical, or radius r and centered on q. A familiar symmetry argument implies $\boldsymbol{E} = E(r)\hat{\boldsymbol{r}}$, so the $\iint$

[37] "Displacement current" is standardly given that name but not a symbol; I have borrowed my $_d$ convention from E. M. Purcell, *Electricity & Magnetism: Berkeley Physics Course, Volume 2* (1963), page 261. Its magnetic analog

$$\iint_{\mathcal{D}} \tfrac{\partial \boldsymbol{B}}{\partial t}\cdot\boldsymbol{dS} = \tfrac{\partial}{\partial t}\underbrace{\iint_{\mathcal{D}} \boldsymbol{B}\cdot\boldsymbol{dS}}_{\text{“magnetic flux”}} \quad \text{if } \mathcal{D} \text{ is not itself moving}$$

is standardly given *neither* a name nor a symbol.

on the left side of (90.1) acquires immediately the value $E(r)4\pi r^2$. From (90.1) we are led thus back again to Coulomb's force law

$$\boldsymbol{E} = E(r)\hat{\boldsymbol{r}} \quad \text{with} \quad E(r) = q/4\pi r^2$$

The $\boldsymbol{E}$-field generated by an *arbitrtary* charge distribution ρ could now be assembled by superposition.

Example. *What is the $\boldsymbol{B}$-field generated by a steady current I in an infinitely long straight wire?* Resolve $\boldsymbol{B}$ into parallel, radial and tangential components

$$\boldsymbol{B} = \boldsymbol{B}_{\parallel} + \boldsymbol{B}_r + \boldsymbol{B}_t$$

as indicated in the figure. By symmetry, the magnitude of each can

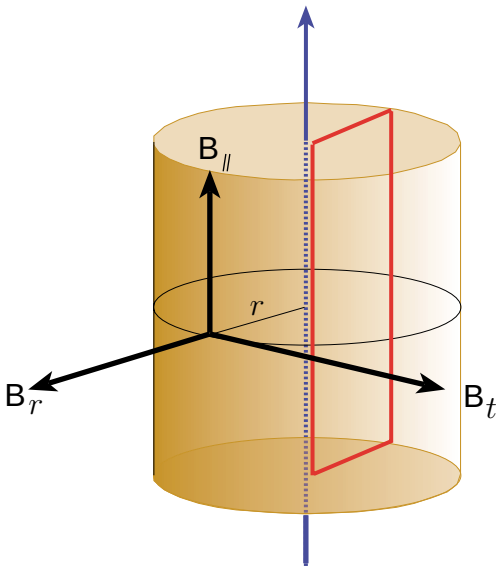


FIGURE 15: *Cylindrical pillbox concentric about a straight wire carrying a steady current. The box has radius r and height h.*

depend only upon r. Equation (90.2) supplies

$$\iint_{\mathcal{R}} \boldsymbol{B}\cdot d\boldsymbol{S} = B_r(r)2\pi rh = 0 \quad \Rightarrow \quad \boldsymbol{B}_r = \boldsymbol{0}$$

while by (90.3) we have

$$\oint_{\text{red rectangle}} \boldsymbol{B}\cdot d\boldsymbol{\ell} = h\big[B_{\parallel}(r_2) - B_{\parallel}(r_1)\big] = 0 \quad \Rightarrow \quad \boldsymbol{B}_{\parallel} = \text{constant}$$

and since we expect to have $\boldsymbol{B}(\infty) = \boldsymbol{0}$ this entails $\boldsymbol{B}_{\parallel} = \boldsymbol{0}$. Finally

$$\oint_{\text{circular cap}} \boldsymbol{B}\cdot d\boldsymbol{\ell} = B_t(r)2\pi r = \tfrac{1}{C}I \tag{91}$$

The implication that the magnetic field "wraps around" the wire, and has a strength that falls off as $1/r$ (*i.e.*, "geometrically," since the system is effectively 2-dimensional). Whether the $\boldsymbol{B}$-field generated by an *arbitrary* steady $\boldsymbol{j}$ could now "be assembled by superposition" (of current-carrying straight wires) remains an interesting open question.[38]

We are in position now to confront the generality of this fundamental question: *What is the $\boldsymbol{B}$-field generated by an arbitrary steady current*? It proves most efficient to proceed not from the integral formulation (90) but from the differential formulation (65) of Maxwell's equations. Just as

the equations		→ the equations	
	$\boldsymbol{\nabla}\times\boldsymbol{E}=\boldsymbol{0}$		$\boldsymbol{\nabla}\cdot\boldsymbol{B}=0$
and	$\boldsymbol{\nabla}\cdot\boldsymbol{E}=\rho$	and	$\boldsymbol{\nabla}\times\boldsymbol{B}=\tfrac{1}{C}\boldsymbol{j}$
give rise to electrostatics, so do			give rise to *magnetostatics*,

the conditions $\partial\rho/\partial t = \boldsymbol{\nabla}\cdot\boldsymbol{j} = 0$ being *shared* by the two subjects in question.

The equation $\boldsymbol{\nabla}\times\boldsymbol{E}=\boldsymbol{0}$ can, by (6.1), be read as stating that there exists a *scalar potential* φ such that

$$\boldsymbol{E} = -\boldsymbol{\nabla}\varphi$$

We note that φ is determined only to within a *gauge transformation*

$$\varphi \to \varphi' = \varphi + \text{constant}$$

and that one can thus arrange that φ vanishes at some given "reference point." Similarly

→ the equation $\boldsymbol{\nabla}\cdot\boldsymbol{B}=0$ can, by (6.2), be read as stating that there exists a *vector potential* $\boldsymbol{A}$ such that

$$\boldsymbol{B} = \boldsymbol{\nabla}\times\boldsymbol{A} \tag{92}$$

We note that $\boldsymbol{A}$ is determined only to within a *gauge transformation*

$$\boldsymbol{A} \to \boldsymbol{A}' = \boldsymbol{A} + \text{grad}\,\chi$$

and that one can thus arrange that $\boldsymbol{A}$ shall in particular satisfy

$$\boldsymbol{\nabla}\cdot\boldsymbol{A} = 0 \tag{93}$$

[38] See page 193 in Purcell's celebrated textbook.[37]

The field equation $\boldsymbol{\nabla}\times\boldsymbol{E}=\boldsymbol{0}$ follows automatically from $\boldsymbol{E}=-\boldsymbol{\nabla}\varphi$, while $\boldsymbol{\nabla}\cdot\boldsymbol{E}=\rho$ becomes

$$\nabla^2\varphi=-\rho$$

This is an inhomogeneous linear equation, the solution of which can, as we have seen (page 16) be described

$$\varphi(\boldsymbol{x})=\iiint G(\boldsymbol{x}-\boldsymbol{\xi})\rho(\boldsymbol{\xi})\,d^3\xi$$

where

$$\nabla^2 G(\boldsymbol{x}-\boldsymbol{\xi})=-\delta(\boldsymbol{x}-\boldsymbol{\xi})$$

entails

$$G(\boldsymbol{x}-\boldsymbol{\xi})=\frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{\xi}|}$$

The $\boldsymbol{E}$-field itself is given therefore by

$$\boldsymbol{E}(\boldsymbol{x})=\iiint -\boldsymbol{\nabla}G(\boldsymbol{x}-\boldsymbol{\xi})\rho(\boldsymbol{\xi})\,d^3\xi$$

where $-\boldsymbol{\nabla}G(\boldsymbol{x}-\boldsymbol{\xi})=\dfrac{\boldsymbol{x}-\boldsymbol{\xi}}{4\pi|\boldsymbol{x}-\boldsymbol{\xi}|^3}$

Similarly

→ The field equation $\boldsymbol{\nabla}\cdot\boldsymbol{B}=0$ follows automatically from $\boldsymbol{B}=\boldsymbol{\nabla}\times\boldsymbol{A}$, while $\boldsymbol{\nabla}\times\boldsymbol{B}=\frac{1}{c}\boldsymbol{j}$ becomes

$$\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\boldsymbol{A})=\tfrac{1}{c}\boldsymbol{j}$$

But $\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\boldsymbol{A})=\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{A})-\nabla^2\boldsymbol{A}$ so if we install the gauge condition $\boldsymbol{\nabla}\cdot\boldsymbol{A}=0$ we obtain

$$\nabla^2\boldsymbol{A}=\tfrac{1}{c}\boldsymbol{j} \qquad (94)$$

This is a trio of inhomogeneous linear equations, the solutions of which can evidently be described

$$\boldsymbol{A}(\boldsymbol{x})=\tfrac{1}{c}\iiint G(\boldsymbol{x}-\boldsymbol{\xi})\boldsymbol{j}(\boldsymbol{\xi})\,d^3\xi \qquad (95)$$

where

$$\nabla^2 G(\boldsymbol{x}-\boldsymbol{\xi})=-\delta(\boldsymbol{x}-\boldsymbol{\xi})$$

entails

$$G(\boldsymbol{x}-\boldsymbol{\xi})=\frac{1}{4\pi|\boldsymbol{x}-\boldsymbol{\xi}|}$$

The $\boldsymbol{B}$-field itself is given therefore by

$$\boldsymbol{B}(\boldsymbol{x})=\tfrac{1}{c}\iiint \begin{pmatrix} 0 & -\partial_3 & \partial_2 \\ \partial_3 & 0 & -\partial_1 \\ -\partial_2 & \partial_1 & 0 \end{pmatrix} G(\boldsymbol{x}-\boldsymbol{\xi})\,\boldsymbol{j}(\boldsymbol{\xi})\,d^3\xi$$

By straightforward calculation[39] we are led thus to

$$=\tfrac{1}{4\pi c}\iiint \frac{\boldsymbol{j}(\boldsymbol{\xi})\times\hat{\boldsymbol{r}}}{r^2}\,d^3\xi \qquad (96)$$

with $\boldsymbol{r}\equiv\boldsymbol{r}(\boldsymbol{x},\boldsymbol{\xi})\equiv\boldsymbol{x}-\boldsymbol{\xi}$.

[39] PROBLEM 25.

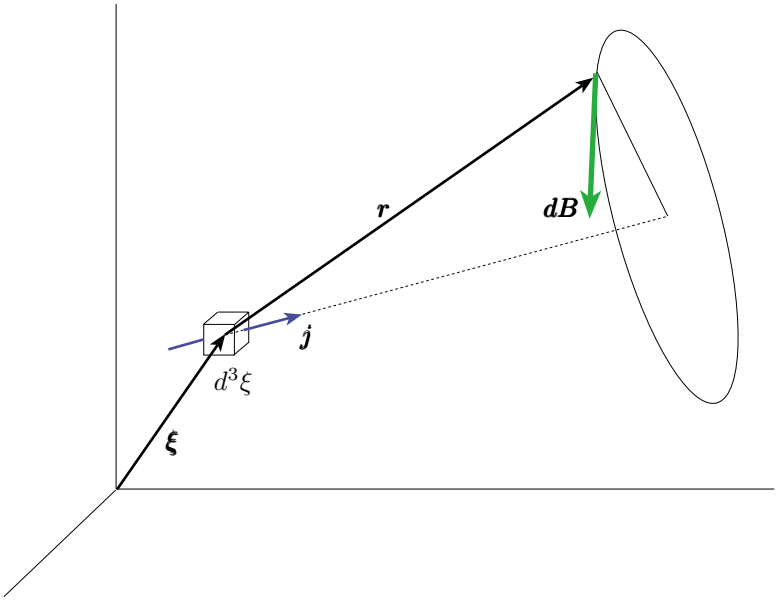


FIGURE 16: *Geometrical meaning of the notations used at* (97) *to describe the differential contribution* $\boldsymbol{dB}$ *to the magnetic field* $\boldsymbol{B}(\boldsymbol{x})$ *at a typical field-point* $\boldsymbol{x}$ *arising from the current differential* $\boldsymbol{j}(\boldsymbol{\xi})\,d^3\xi$ *at a typical source-point* $\boldsymbol{\xi}$.

Equation (96)—though analytically a corollary of Ampere's law (84)—is known standardly (and with more historical justice) as the *Biot–Savart law*. It describes the $\boldsymbol{B}$-field generated by an arbitrary steady current distribution $\boldsymbol{j}$, and invites "interpretation-by-superposition" along lines which emerge if (see the figure) we write

$$\begin{aligned} \boldsymbol{B}(\boldsymbol{x}) &= \int \boldsymbol{dB}(\boldsymbol{x},\boldsymbol{\xi}) \\ \boldsymbol{dB}(\boldsymbol{x},\boldsymbol{\xi}) &= \tfrac{1}{4\pi c}\,\frac{[\boldsymbol{j}(\boldsymbol{\xi})\,d^3\xi]\times\hat{\boldsymbol{r}}}{r^2} \end{aligned} \tag{97}$$

The interpretation of $\boldsymbol{j}(\boldsymbol{\xi})\,d^3\xi$ is, however, a little bit odd. The object in question is perfectly meaningful *in context* (*i.e.*, under the $\int$), but—for the reasons remarked already on page 38—*could not be realized in isolation*.

Later we shall have occasion to study illustrative applications of (96), but for the moment must rest content with a single

Example. *What—according to (96)—is the* $\boldsymbol{B}$*-field generated by a steady current* I *in an infinitely long straight wire*? Taking our notation from the following figure, it is immediate that

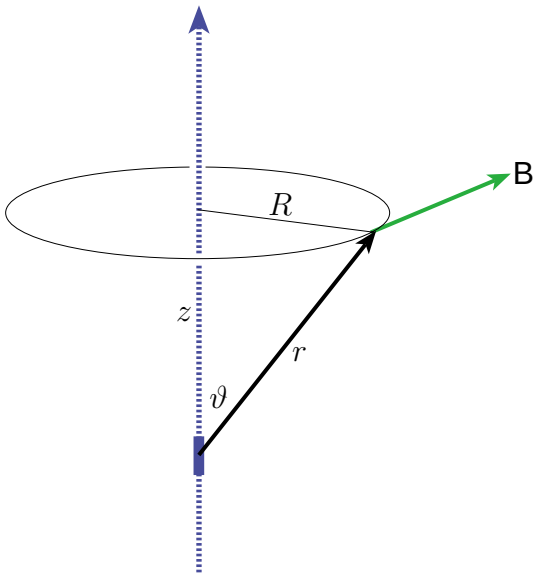


FIGURE 17: *Notation employed in computing the magnetic field* $\boldsymbol{B}$ *generated by current in an infinitely long straight wire.*

$$\begin{aligned} B(R) &= \tfrac{1}{4\pi c} I \int_{-\infty}^{+\infty} \tfrac{1}{r^2} \sin\vartheta \, dz \\ &= \tfrac{1}{4\pi c} I \underbrace{\int_{-\infty}^{+\infty} \frac{R}{(R^2+z^2)^{\frac{3}{2}}} \, dz}_{= \left. \frac{z}{R\sqrt{R^2+z^2}} \right|_{-\infty}^{+\infty} = \frac{2}{R}} \\ &= \frac{I}{2\pi c R} \end{aligned}$$

—which agrees precisely with the result (91) obtained previously by other means.

It should be noticed that if the Biot-Savart law were *postulated* (*i.e.*), abstracted from laboratory experience then the equations $\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0$ and $\boldsymbol{\nabla}\times\boldsymbol{B} = \frac{1}{c}\boldsymbol{j}$—our starting point—could have been recovered as *corollaries* of (96).

I turn now to discussion of the question which was central to Ampere's own work in this area: *What is the force which (steady) currents exert upon one another by virtue of the magnetic fields which they generate*? Suppose, by way of preparation, that

1) impressed fields $\boldsymbol{E}$ and $\boldsymbol{B}$
2) source functions ρ and $\boldsymbol{j}$

are defined on the neighborhood d^3x of a representative point $\boldsymbol{x}$. From the Lorentz force law (67)

$$\boldsymbol{F} = q(\boldsymbol{E} + \tfrac{1}{c}\boldsymbol{v}\times\boldsymbol{B})$$

we infer that the total force experienced by the charges which momentarily occupy d^3x can be described

$$\begin{aligned} \boldsymbol{dF} &= \boldsymbol{\mathfrak{F}}(\boldsymbol{x})d^3x \\ \boldsymbol{\mathfrak{F}} &= \rho\boldsymbol{E} + \tfrac{1}{c}\boldsymbol{j}\times\boldsymbol{B} \equiv \textit{force density} \end{aligned}$$

If we look more specifically to the situation typical of wires (where "charge carriers" drift through a population of oppositely charges) we have $\rho = 0$ whence

$$= \tfrac{1}{c}\boldsymbol{j}\times\boldsymbol{B} \tag{98}$$

> Wires—even wires carrying current—are standardly uncharged, and therefore don't feel ambient $\boldsymbol{E}$-fields.

If (see the first of the following figures) we integrate (98) over a snippet $\boldsymbol{d\ell}$ of wire we obtain

$$\boldsymbol{df} = \int_{\text{snippet}} \boldsymbol{dF} = \tfrac{1}{c} I\boldsymbol{d\ell}\times\boldsymbol{B} \tag{99}$$

Suppose now that $I \equiv I_1$ and $\boldsymbol{d\ell} \equiv \boldsymbol{d\ell}_1$ refer (see the second of the following figures) to a closed loop $\mathcal{L}_1$ of wire, and that $\boldsymbol{B}$ arises from a (steady) current I_2 in a second loop $\mathcal{L}_2$. From (97) and (99) we conclude that the *force* $\boldsymbol{f}_{12}$ *exerted on* $\mathcal{L}_1$ *by* $\mathcal{L}_2$ can be described

$$\boldsymbol{f}_{12} = \tfrac{1}{4\pi c^2} I_1 I_2 \oint_{\mathcal{L}_1}\oint_{\mathcal{L}_2} \frac{\boldsymbol{d\ell}_1 \times (\boldsymbol{d\ell}_2 \times \hat{\boldsymbol{r}}_{12})}{r_{12}^2} \tag{100}$$

It is to this implausible, non-local (*i.e.*, distributed, whence geometry-dependent) result that the name of Ampere is most properly attached. Looking now to some of the implications of (100) ... from $\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c}) = (\boldsymbol{a}\cdot\boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a}\cdot\boldsymbol{b})\boldsymbol{c}$ we have

$$= \tfrac{1}{4\pi c^2} I_1 I_2 \left\{ \oint_{\mathcal{L}_1}\left[\oint_{\mathcal{L}_2} \frac{\boldsymbol{r}_{12}\cdot\boldsymbol{d\ell}_1}{r_{12}^3}\right]\boldsymbol{d}\ell_2 - \oint_{\mathcal{L}_1}\oint_{\mathcal{L}_2} \frac{\boldsymbol{r}_{12}}{r_{12}^3}(\boldsymbol{d\ell}_1\cdot\boldsymbol{d\ell}_2)\right\}$$

But (recall (15.1)) $\boldsymbol{r}_{12}/r_{12}^3 = -\boldsymbol{\nabla}_1(1/r_{12})$ so we have

$$\Big[\text{etc.}\Big] = -\tfrac{1}{r_{12}}\Big|_{\text{starting point}}^{\text{endpoint}} = 0 \quad \text{for a loop}$$

giving

$$\boldsymbol{f}_{12} = -\tfrac{1}{4\pi c^2} I_1 I_2 \oint_{\mathcal{L}_1}\oint_{\mathcal{L}_2} \frac{\boldsymbol{r}_{12}}{r_{12}^3}(\boldsymbol{d\ell}_1\cdot\boldsymbol{d\ell}_2) \tag{101}$$

$$= \left\{ \begin{array}{l} \text{net force on circuit } \mathcal{L}_1 \text{ due to interaction with} \\ \text{magnetic field generated by current in circuit } \mathcal{L}_2 \end{array} \right.$$

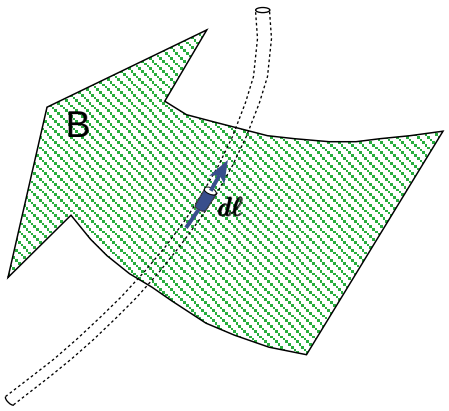


FIGURE 18: *Snippet* $\boldsymbol{d\ell}$ *of wire carrying a current* I *in the presence of an ambient magnetic field* $\boldsymbol{B}$.

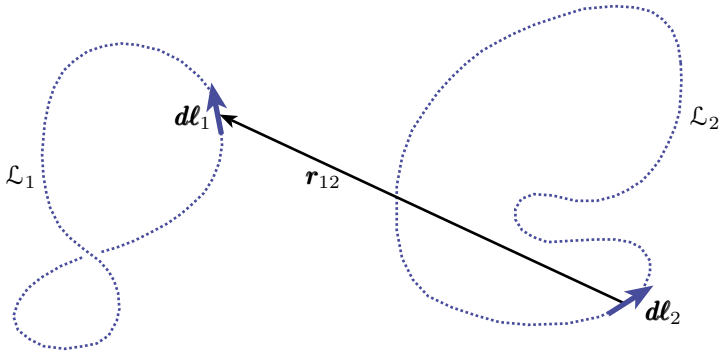


FIGURE 19: *Interaction of current* I_1 *in snippet* $\boldsymbol{d\ell}_1$ *with magnetic generated by current* I_2 *in snippet* $\boldsymbol{d\ell}_2$.

From $\boldsymbol{r}_{12} = -\boldsymbol{r}_{21}$ we conclude that *the forces of interaction between steady current loops conform to Newton's* 3rd *law*:

$$\boldsymbol{f}_{12} = -\boldsymbol{f}_{21} \tag{102}$$

The structure of (100) encourages one to suppose that the equation in question arises by superposition from a statement of the form

$$\text{force on } I_1\boldsymbol{d\ell}_1 \text{ by } I_2\boldsymbol{d\ell}_2 = \tfrac{1}{4\pi c^2}\frac{I_1\boldsymbol{d\ell}_1 \times (I_2\boldsymbol{d\ell}_2 \times \hat{\boldsymbol{r}}_{12})}{r_{12}^3} \tag{103.1}$$

Observing that the vector on the right lies in the plane spanned by $I_2\boldsymbol{d\ell}_2$ and $\boldsymbol{r}_{12}$, we conclude that

$$\neq -\text{force on } I_2\boldsymbol{d\ell}_2 \text{ by } I_1\boldsymbol{d\ell}_1 \tag{103.2}$$

i.e., that *the element-element interaction which purportedly lies at the root of (101) does not conform to Newton's* 3rd *law*.

> People frequently proceed from this fact to the (in my view) profoundly misguided conclusion that Newton's 3rd law is "soft . . . that it holds except when it doesn't." The correct conclusion, it seems to me, is that the isolated current element $I\boldsymbol{d\ell}$ is a hazardous abstraction.

This surprising result is illustrated in the following figure. The figure suggests also that an unknotted current-carrying loop will tend (by magnetic self-interaction) to deform until circular . . . which for a closed loop means "as nearly

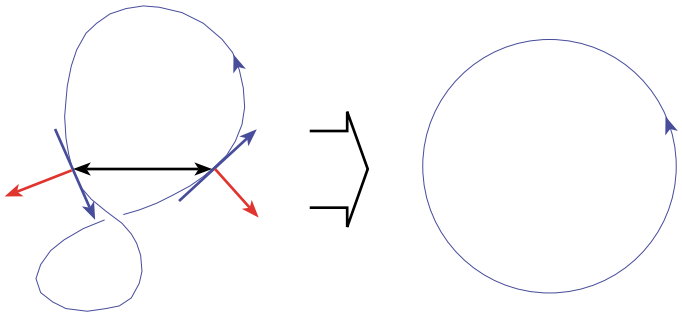


FIGURE 20: *Red arrows in the figure at left refer to the interaction* (103) *of two current elements in a filamentary circuit. If the filament is flexible we expect it to assume the circular form shown at right.*

straight as possible." From (103) we see that current elements in a *straight* wire *do not interact at all*—whence again the inference: "current-carrying wires like, for magnetic reasons, to be as straight as possible." We come away with the impression that electrical devices in which the

1) geometry and/or
2) operative I-values

favor the production of substantial $\boldsymbol{B}$-fields . . . must be strongly constructed, for they will be subjected generally to a tendency to explode![40]

Since wires and electrical devices are "unnatural/artificial" in the sense that they more often the work of engineers than of Nature, it is attractive to suppose that (103) arises as a corollary from

$$\text{force on } \boldsymbol{j}(\boldsymbol{x}_1)d^3x_1 \text{ by } \boldsymbol{j}(\boldsymbol{x}_2)d^3x_2 = \tfrac{1}{4\pi c^2}\frac{\boldsymbol{j}(\boldsymbol{x}_1)\times\big(\boldsymbol{j}(\boldsymbol{x}_2)\times\hat{\boldsymbol{r}}_{12}\big)}{r_{12}^3}d^3x_1d^3x_2 \quad (104)$$

and to view (104) as the *magnetic analog of Coulomb's law*.

[40] PROBLEM 26.

The pattern provided by our prior discussion of electrostatics (see especially pages 19–24) makes it natural to inquire finally into the *energetics* of *magneto*static fields. But we encounter at once some unexpected conceptual difficulties: it is unnatural (taking the argument of pages 19–20 as our model) to attempt to position the current elements $\boldsymbol{j}(\boldsymbol{x})d^3x$ "one at a time" because

1) "isolated point currents" *do not exist*;
2) we would stand in *violation of charge conservation* (*i.e.*, of $\boldsymbol{\nabla}\cdot\boldsymbol{j}=0$) until the assembly is complete;
3) the assembly process entails that we work against forces which *violate Newton's* 3rd *law*.

It is better practice to *build the* $\boldsymbol{j}$*-field by slowly turning it on* ... but this, by (65.4), involves "Faraday emf effects" which we are not presently in position to calculate. My plan, therefore, will be simply to *present* the formula in question (several lines of supporting argument will be reviewed later) and to develop its formal relationship to its electrostatic counterpart:

In electrostatics we obtained

$$W=\int \mathcal{E}(\boldsymbol{x})\,d^3x$$

where $\mathcal{E}\equiv\frac{1}{2}\boldsymbol{E}\cdot\boldsymbol{E}$ defines the electrostatic energy density. Thus

$$W=\tfrac{1}{2}\int \boldsymbol{E}\cdot\boldsymbol{E}\,d^3x$$

which arose (at page 22) from

$$W=-\tfrac{1}{2}\int \underbrace{\boldsymbol{E}\cdot\boldsymbol{\nabla}\varphi}\; d^3x$$

This in turn came—use

$$\equiv -\varphi\boldsymbol{\nabla}\cdot\boldsymbol{E}+\boldsymbol{\nabla}(\varphi\boldsymbol{E})$$

and discard the surface term—from

$$W=\tfrac{1}{2}\int \varphi\boldsymbol{\nabla}\cdot\boldsymbol{E}\,d^3x$$

which we got (by $\boldsymbol{E}=-\boldsymbol{\nabla}\varphi$) from

$$=-\tfrac{1}{2}\int \varphi\,\nabla^2\varphi\,d^3x$$

$$=\tfrac{1}{2}\int \rho\,\varphi\,d^3x$$

$$=\tfrac{1}{8\pi}\iint \rho(\boldsymbol{x})\rho(\boldsymbol{\xi})\tfrac{1}{|\boldsymbol{x}-\boldsymbol{\xi}|}\,d^3x d^3\xi$$

Proceeding similarly (but in reverse), → in magnetostatics we write

$$W=\int \mathcal{B}(\boldsymbol{x})\,d^3x$$

where $\mathcal{B}\equiv\frac{1}{2}\boldsymbol{B}\cdot\boldsymbol{B}$ defines the *magnetostatic energy density*. Thus

$$W=\tfrac{1}{2}\int \boldsymbol{B}\cdot\boldsymbol{B}\,d^3x \qquad (105)$$

giving

$$W=\tfrac{1}{2}\int \underbrace{\boldsymbol{B}\cdot(\boldsymbol{\nabla}\times\boldsymbol{A})}\; d^3x$$

But

$$\equiv \boldsymbol{A}\cdot(\boldsymbol{\nabla}\times\boldsymbol{B})+\boldsymbol{\nabla}\cdot(\boldsymbol{A}\times\boldsymbol{B})$$

so—discarding the surface term—

$$W=\tfrac{1}{2}\int \boldsymbol{A}\cdot(\boldsymbol{\nabla}\times\boldsymbol{B})\,d^3x$$

From $\boldsymbol{B}=\boldsymbol{\nabla}\times\boldsymbol{A}$ and the gauge condition $\boldsymbol{\nabla}\cdot\boldsymbol{A}=0$ it now follows that

$$=-\tfrac{1}{2}\int \boldsymbol{A}\cdot\nabla^2\boldsymbol{A}\,d^3x$$

$$=-\tfrac{1}{2c}\int \boldsymbol{j}\cdot\boldsymbol{A}\,d^3x$$

$$=\tfrac{1}{8\pi c^2}\iint \boldsymbol{j}(\boldsymbol{x})\cdot\boldsymbol{j}(\boldsymbol{\xi})\tfrac{1}{|\boldsymbol{x}-\boldsymbol{\xi}|}\,d^3x d^3\xi \qquad (106)$$

We note that the formal parallel is perfect. Also

1) that the $\int$ in (105) ranges over the field, while the $\int$ in (106) ranges only over its source;
2) that the W of (105) is a *non-negative & non-linear* number-valued functional of $\boldsymbol{B}$;
3) that true line currents give rise to a variant of the familiar *self-energy problem*, and that so (for other reasons) do currents which have their termini at ∞.

The previously-remarked tendency of current-carrying wires to move around can be considered now to follow—by the "nameless principle" of page 23—from the fact that in so doing they may reduce the energy stored in the associated $\boldsymbol{B}$-field.[41]

To conclude: the discussion in recent pages derives mainly from Ampere's law

$$\boldsymbol{\nabla}\times\boldsymbol{B} = \tfrac{1}{c}\boldsymbol{j} \tag{84}$$

—the phenomenological consequences of which have been seen to conform to the diverse physical facts, and to come to this: currents generate and respond to magnetic fields. Maxwell's modification

$$\boldsymbol{\nabla}\times\boldsymbol{B} = \tfrac{1}{c}\big(\boldsymbol{j} + \tfrac{\partial \boldsymbol{E}}{\partial t}\big)$$

implies that $\boldsymbol{B}$-fields are generated also by *temporal variation of* $\boldsymbol{E}$ . . . of which more later.

5. Faraday's law. Coulomb had argued (from evidence) that "electrical and magnetic phenomena are unrelated." It is because Coulomb's view had come to be widely shared that news of Oersted's discovery (of "electromagnetism:" 1820) produced such perplexed excitement among French physicists . . . and generated the developments reported in the preceding seciton of these notes. Across the Channel, Sir Humphry Davy (1778–1829) was receiving regular reports—jumbled and contradictory as they at the time seemed—of the work of his French colleagues. His assistant (Michael Faraday) repeated the basic experiments and entered into correspondence with Ampere, whose work had earned him a position of acknowledged leadership among the French. Faraday confessed openly that he could not understand the mathematical aspects of Ampere's work . . . but took intuitive exception to some of Ampere's interpretive comments/ideas, particularly those concerning the microscopic meaning of "current." Absent the observational data required to settle the issue, Ampere and Faraday "agreed to disagree," and Faraday turnjed to other matters.

By 1830, Faraday (under the influence of Charles Wheatstone: 1802–1875) had developed an interest in physical acoustics . . . and particularly in the Chladni patterns which are set up on one membrane when another (distant)

[41] PROBLEM 27.

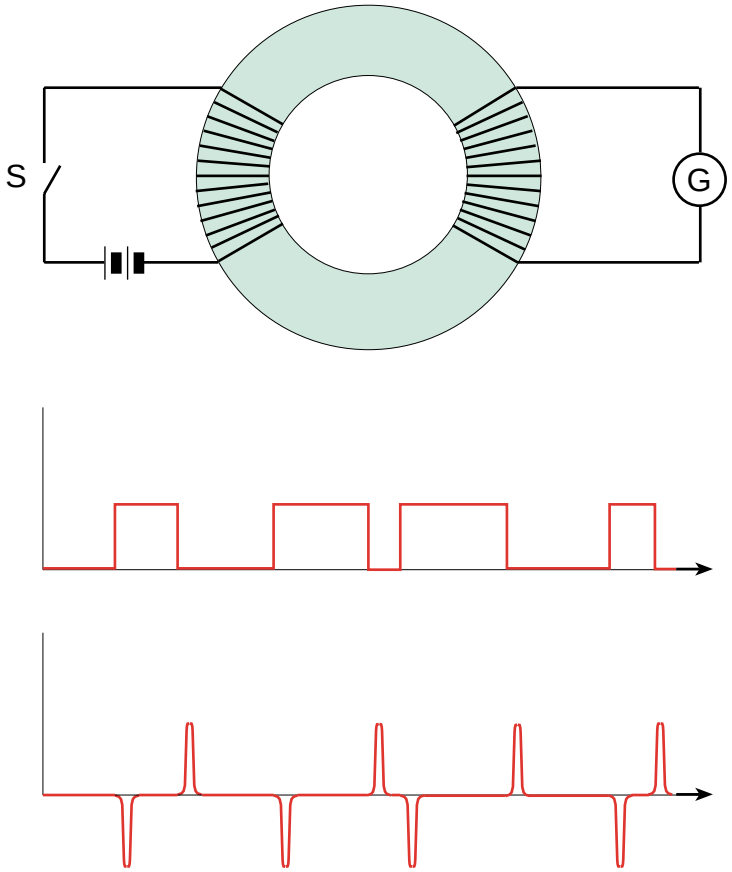


FIGURE 21: *At top, the experimental set-up used by Faraday. The idea of using a soft iron ring to link the primary and secondary coils was borrowed from Joseph Henry; without it the induced current would have been undetectably small. The middle figure shows the current in the primary that results from closing/opening the switch S. The graph of the induced current (lower figure) does not mimic the primary current, but shows spikes synchronized with the switch activity; i.e., with the moments when* $\frac{d}{dt} I_{\text{primary}} \neq 0$. *Jacque d'Arsonval's dates, by the way, are* 1851–1950*; the question therefore arises*: What kind of ballistic galvinometer was available to Faraday in 1831?

membrane is stimulated. Faraday's interest in this topic was reenforced by his reading of an essay by John Herschel ("A preliminary discourse on the study of natural philosophy;" 1830) in which it was argued that the physics of light and the physics of sound must be similar . . . in the sense that both must have root in the vibratory motion of an elastic medium. Faraday speculated that such an analogy might pertain also to electrical and magnetic phenomena. Faraday was aware that such a view—though out of fashion among the French—had been

advanced already in 1806 by Oersted (who had himself cited Chladni patterns as "analogs of electrical action"), and reasoned that electrical/magnetic effects (if such a view were correct)

1) could be understood only in terms of the dynamics of the "elastic medium;" *i.e.*, in the language of a *field theory*
2) entailed *delayed* action-at-a-distance.

Thus did time t become for Faraday a relevant dynamical variable. Faraday's problem was to discover observational evidence which would support or contradict the weight of his intuition.

It was at about this point (1831) that Faraday learned of the *strong electromagnets* which Joseph Henry (of Albany, New York, and later first director of the Smithsonian Institution: 1797–1878) had achieved by *replacing the traditional air core with a soft iron core*. Faraday knew that

1) currents give rise to (and feel) $\boldsymbol{B}$-fields, and anticipated (with an intuition rooted partly in his religious convictions) that
2) $\boldsymbol{B}$-fields should give rise (after some brief delay?) to currents.

More or less thus[42] was Faraday led (September & October, 1831) to the experimental arrangement and discovery outlined in Figure 21. Previous efforts to detect "the currents generated by $\boldsymbol{B}$-fields" had always yielded a negative result. What Faraday had in effect discovered was that currents arise not from $\boldsymbol{B}$ but from $\partial\boldsymbol{B}/\partial t$. The qualitative/quantitative upshot of Faraday's experiments—which were many and diverse, and were in some respects anticipated (1830) by Henry (who, however, was slow to publish his findings)—can be summarized

$$\boldsymbol{\nabla}\times\boldsymbol{E} = -\tfrac{1}{c}\tfrac{\partial\boldsymbol{B}}{\partial t} \tag{107}$$

which was encountered already at (65.4) and (81.4), and is an expression of *Faraday's law of electromagnetic induction*. Several comments are now in order:

1. Comparison of (107) with

$$\boldsymbol{\nabla}\times\boldsymbol{B} = \begin{cases} \tfrac{1}{c}\,\boldsymbol{j} & : \text{ steady case} \\ \tfrac{1}{c}\big(\boldsymbol{j} + \tfrac{\partial\boldsymbol{E}}{\partial t}\big) & : \text{ general case} \end{cases} \tag{83}$$

shows that the geometry of the $\boldsymbol{E}$-field generated by $\partial\boldsymbol{B}/\partial t$ resembles the geometry of the $\boldsymbol{B}$-field generated (in the steady case) by $\boldsymbol{j}$. We see also that the structure of Faraday's law (107) provides formal precedent for Maxwell's $\partial\boldsymbol{E}/\partial t$-term.

2. Faraday's law (107) presents—*Lenz' law*—a "stabilizing minus sign" which is absent from the Ampere-Maxwell law (83) ... of which more later.

---

[42] The ture history of his thought is not known, and I have provided only the grossest outline of what *is* known: see Chapter 4 in L. P. Williams, *Michael Faraday* (1964) or Chapter 3 in R. A. R. Tricker, *The Contributions of Faraday & Maxwell to Electrical Science* (1966) for further details.

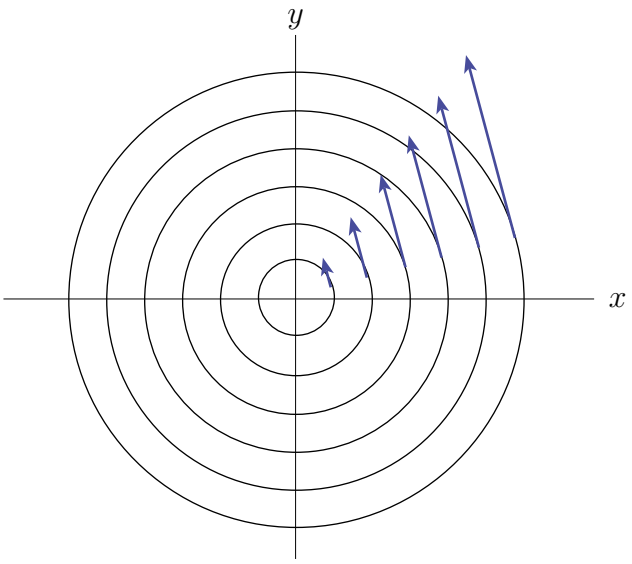


FIGURE 22: *The $\boldsymbol{E}$-field encountered in the following example. The field is divergenceless, but has obvious circulation (or "curl"). Such fields cannot be produced electrostatically, but are typical of the fields produced by time-dependent magnetic fields.*

3. Equation (107)—surprisingly, in view of Faraday's intent (and the nature of his observationss)—contains *no direct reference to current*. It says that $\partial\boldsymbol{B}/\partial t$ generates an $\boldsymbol{E}$-field, which *in the presence of charge* may give rise to charge flow. Suppose, for example, that $\boldsymbol{B}(\boldsymbol{x},t)$ has the (physically implausible) form

$$\boldsymbol{B}=\begin{pmatrix}0\\0\\-c\beta t\end{pmatrix}\quad:\quad\begin{cases}\text{uniformly ramped,}\\ \boldsymbol{x}\text{-independent,}\\ \text{everywhere }\parallel\text{ to the }z\text{-axis}\end{cases}$$

It then follows from (107) that

$$\begin{aligned}\boldsymbol{E}(\boldsymbol{x},t)&=\begin{pmatrix}-\frac{1}{2}\beta y\\ +\frac{1}{2}\beta x\\ 0\end{pmatrix}+\operatorname{grad}\varphi\\ &=\boldsymbol{E}_{\text{faraday}}+\boldsymbol{E}_{\text{electrostatic}}\end{aligned}$$

where—by (6)—$\boldsymbol{E}_{\text{faraday}}$ is <u>divergenceless</u>, but $\boldsymbol{E}_{\text{electrostatic}}$ is <u>curlless</u> (and is fixed not by (107) but by $\boldsymbol{\nabla}\cdot\boldsymbol{E}=\rho$ and the physically appropriate boundary conditions). The structure of the induced field $\boldsymbol{E}_{\text{faraday}}$ (which, it is important to notice, is <u>not conservative</u>: $\boldsymbol{\nabla}\times\boldsymbol{E}_{\text{faraday}}\neq\boldsymbol{0}$) is indicated in Figure 22.

If a charge—let us, for simplicity, say a *solitary* charge—q were released it would move off initially in response to the $\boldsymbol{E}$-field, but after it had gained some velocity it would—by

$$\boldsymbol{F}=q\big(\boldsymbol{E}+\tfrac{1}{c}\,\boldsymbol{v}\times\boldsymbol{B}\big)\tag{67}$$

—also feel the $\boldsymbol{B}$-field. We would, therefore, expect the trajectory of q to be in general quite complicated ... and even more so if (instead of being solitary) q has companions. Faraday was himself protected from the full force of the complications just mentioned by two experimental circumstances:

- his currents were *confined to wires*
- his wires were mechanically *constrained to resist Amperean forces.*

But how—in detail—do the circumstances just noted serve to "simplify" the implications of the Maxwellian field equations (65) and of the Lorentzian force law (67)? To resolve this important question we must digress to acquire some familiarity with

1) a phenomenological law (Ohm's law) and
2) a poorly named but useful concept ("electromotive force").

THEORETICAL STATUS OF OHM'S LAW It is to Georg Simon Ohm (1789–1854) that we owe the experimental discovery (1826) that for samples of the most commonly encountered conductive materials

$$\frac{\text{impressed voltage}}{\text{resulting current}} = \text{constant}$$

i.e.,

$$V/I = R \equiv \text{sample } \textit{resistance}$$

—resistance between those particular contact points (see upper Figure 23). To formulate this result in geometry-independent terms specific to the material in question let the sample be of "standard shape" (*i.e.*, cylindrical, of length L and cross-sectional area A). Using $V = EL$ and $I = JA$ we have

$$E = \rho J$$
$$\rho \equiv RA/L \equiv \textit{resistivity} \text{ of the material}$$

which is more usefully[43] notated

$$\boldsymbol{j} = \sigma \boldsymbol{E} \tag{108}$$
$$\sigma \equiv \textit{conductivity} = \frac{1}{\text{resistivity}}$$

OHM'S LAW (108) provides our first instance of what is called a "constitutive relation." Such relations are denied "fundamental" status not because they are approximate (even Maxwell's field equations[44] are, strictly speaking, only approximate) but because they are subject in (in)appropriately chosen materials

[43] And at less risk of confusing ρ with "charge density"!

[44] What we call "Maxwell's equations" were abstracted from Maxwells' work by Heaviside, Lorentz and others over a period of nearly twenty years. I was surprised to discover that the equations proposed by Maxwell himself *included Ohm's law as a full-fledged partner*; see "Theories of Maxwellian design" (1998).

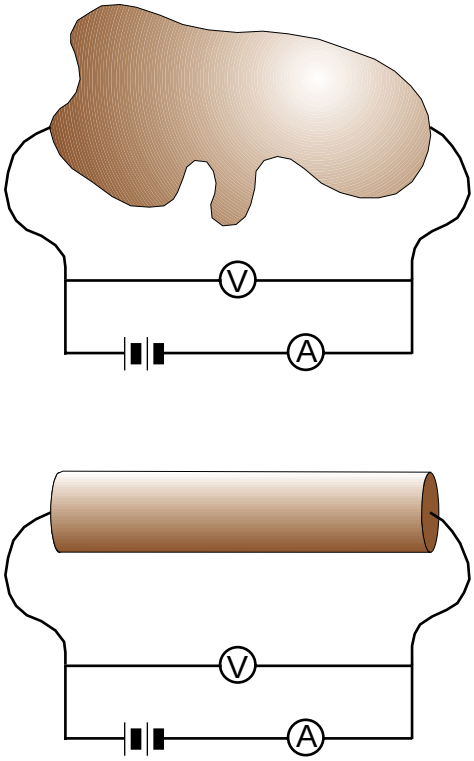


FIGURE 23: *Above: the arrangement used to measure the resistance R between two specified points on the surface of an arbitrary material blob. Below: the standardized sample of homogeneous material used to measure "resistivity" (or "conductivity"), which is an intrinsic property of that material.*

to gross violation. Constitutive relations have always—sooner or later—to be derived from first principles: the task is seldom easy, and entails that *such relations have always the character of macroscopic averages over microscopic complexities*. For *anisotropic* materials (108) assumes the form

$$\boldsymbol{j} = \boldsymbol{\sigma}\boldsymbol{E} \tag{109}$$

$$\boldsymbol{\sigma} \equiv \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \equiv \text{conductivity matrix}$$

Equations (109) look like a truncated version of the still more general relation

$$j_n = \sum_a \sigma_{na} E_a + \underbrace{\tfrac{1}{2} \sum_{a,b} \sigma_{nab} E_a E_b + \tfrac{1}{3!} \sum_{a,b,c} \sigma_{nabc} E_a E_b E_c + \cdots}_{\text{important in specialized materials, or when the } \boldsymbol{E}\text{-field is sufficiently strong}}$$

Evidently Ohm's law refers merely[45] to the "tip of a phenomenological iceberg." The surprising fact, from this point of view, is that (108) contains no reference at all to $\boldsymbol{B}$ ... for reasons which have evidently to do with the fact that the drift velocity v is typically so small that the $(\boldsymbol{v}\times\boldsymbol{B})$-forces experienced by individual charge carriers are negligible ... though the facts are, as will emerge, somewhat more subtle. Bringing (76) to (108) we obtain $nev = \sigma E$ or

$$e\boldsymbol{E} = \frac{ne^2}{\sigma}\,\boldsymbol{v} \tag{110}$$

... according to which the impressed force $e\boldsymbol{E}$ is proportional not to the acceleration but to the (mean) *velocity* of the charge carriers. The situation is (roughly) this: the charge carriers keep *trying* to accelerate, but keep *running into things and getting stopped* (releasing their recently acquired kinetic energy to the obstacle-matrix, which gets hot). The situation can be modeled (Drude, 1900) by writing

$$e\boldsymbol{E} - \underbrace{b\,\boldsymbol{v}}_{\text{damping force: models the effect of collisions}} = m\boldsymbol{a} \tag{111}$$

and supposing that the (mean) acceleration $\boldsymbol{a} = \boldsymbol{0}$. The drift velocity acquires thus the status of a kind of "terminal velocity," and it follows in fact from (110/111) that

$$\sigma = \frac{ne^2}{b}$$

The implication is that *charge carriers keep moving because the ambient $\boldsymbol{E}$-field keeps doing work* on them. How about the $\boldsymbol{B}$-field? It does work at the temporal rate given by

$$P_{\text{magnetic}} = \boldsymbol{v}\cdot\tfrac{e}{c}(\boldsymbol{v}\times\boldsymbol{B}) = 0 \tag{112}$$

Magnetic fields do no work on moving charges and so cannot assist in the transport of charge carriers through a wire ... except perhaps indirectly: one can imagine far-fetched circumstances in which $\boldsymbol{B}$-fields (by deforming the carrier trajectories) might cause charge carriers to hit/miss appropriately deployed obstacles. This would lend $\boldsymbol{B}$-dependence to σ, but would *not* cause an additive $\sigma_{\text{mag}}\boldsymbol{B}$-term to appear on the right side of Ohm's law (108).

ELECTROMOTIVE FORCE Given, therefore, that charge carriers flow through material wires because constantly worked on by $\boldsymbol{E}$-fields, the question arises: *How much* work (per unit charge) do the $\boldsymbol{E}$-fields do? Let the circuit in question be modeled by a closed curve (of loop) $\mathcal{C}$. Immediately

$$q\oint_{\mathcal{C}}\boldsymbol{E}\cdot d\boldsymbol{\ell} = \text{work done in transporting } q \text{ virtually around } \mathcal{C}$$

[45] ... but importantly: the conductivities of common materials range over at least 23 orders of magnitude. Few indeed are the "laws of Nature" that can claim such dynamic range.

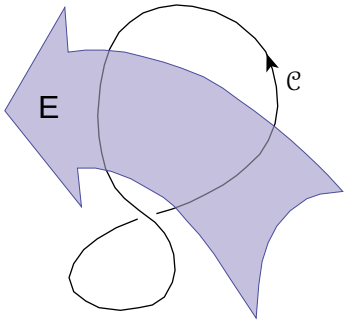


FIGURE 24: *The physical* $\boldsymbol{E}$*-field and "mental loop"* $\mathcal{C}$ *that enter at (113) into the definition of* "emf." *In practical applications it is often natural to identify the "mental loop" with a* <u>metal</u> *loop (wire).*

where the "virtually" means that the pransport takes place "mentally," not physically (*i.e.*, not in real time, with the attendant accelerations, etc.). The $\oint$ defines what is called the "electromotive force" associated with the given circuit and field. It is standardly denoted $\mathcal{E} \equiv \mathcal{E}(\mathcal{C}, \boldsymbol{E})$, and *has actually not the dimensions of "force" but of "work/charge."* I prefer therefore to call

$$\mathcal{E} \equiv \oint_{\mathcal{C}} \boldsymbol{E}\cdot\boldsymbol{d\ell} \tag{113}$$

the "emf" of the circuit/field in question (and to put out of mind the fact that "emf" came into the world as an acronym). What *is* the $\boldsymbol{E}$-field contemplated at (113)? It is the "field experienced by the virtually transported test charge"—a field which (since the interior of matter is a complicated place) is actually *unknown*. Happily, *the complication just noted is—to the (substantial) extent that it is microelectrostatic in origin—irrelevant* ... for this simple reason: electrostatic fields are curlless

$$\boldsymbol{\nabla}\times\boldsymbol{E}_{\text{electrostatic}} = \boldsymbol{0}$$

... from which it follows by Stokes' theorem (87) that (for all circuits $\mathcal{C}$)

$$\mathcal{E}_{\text{electrostatic}} = 0 \tag{114}$$

This means that *the (generally unknown) electrostatic component of the "fields experienced by the transported charge q" can be dropped from all* emf-*calculations*. To make the same point another way: purely electrostatic $\boldsymbol{E}$-fields cannot be used to drive currents in circuits.[46] But while $\boldsymbol{\nabla}\times\boldsymbol{E} = \boldsymbol{0}$ pertains universally to electrostatic fields, it does *not* pertain

- to the $\boldsymbol{E}$-fields generated by chemical action in batteries;

[46] What, in this light, do you make of the physics of *lightening bolts*?

- to the $\boldsymbol{E}$-fields produced by thermal/optical/mechanical action in diverse solid-state devices;

. . . and in particular it does not pertain

- to the Faraday $\boldsymbol{E}$-fields which, according to (107), are induced by time-dependent $\boldsymbol{B}$-fields.

So non-zero values of $\mathcal{E}$ are certainly attainable.[47] Drawing finally upon Ohm's law, we have

$$\begin{aligned}\mathcal{E} \equiv \oint_{\mathcal{C}} \boldsymbol{E}\cdot d\boldsymbol{\ell} &= \oint_{\mathcal{C}} \rho \boldsymbol{j}\cdot d\boldsymbol{\ell} \\ &= \rho \frac{I}{A} L \quad \text{for wires of uniform cross section} \\ &\downarrow \\ \text{giving} \qquad \mathcal{E} &= IR \end{aligned} \tag{115}$$

It should, in view of (115), not be necessary to belabor the claim that $\mathcal{E}$ is—at least for the purposes of practical/applied physics—a "useful[48] concept."

The question posed near the top of page 65 now "answers itself." The integral formulation

$$\int_{\partial \mathcal{D}} \boldsymbol{E}\cdot d\boldsymbol{\ell} = -\tfrac{1}{C} \iint_{\mathcal{D}} \tfrac{\partial \boldsymbol{B}}{\partial t}\cdot d\boldsymbol{S} \tag{90.4}$$

of Faraday's law (107) can now be formulated

$$\mathcal{E}_{\text{faraday}} = -\tfrac{1}{C}\dot{\Phi} \tag{116}$$

$$\Phi \equiv \iint_{\mathcal{D}} \boldsymbol{B}\cdot d\boldsymbol{S} \equiv \textit{magnetic flux} \text{ through } \mathcal{D}$$

> REMARK: Let $\mathcal{D}_1$ and $\mathcal{D}_2$ be distinct caps that share the same boundary $\mathcal{C} = \partial\mathcal{D}_1 = \partial\mathcal{D}_2$. It is (recall the formal upshot of PROBLEM 17) a consequence of $\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0$ that
>
> $$\text{magnetic flux through } \mathcal{D}_1 = \text{magnetic flux through } \mathcal{D}_2$$
>
> and better, therefore, to speak of the "magnetic flux entrapped by $\mathcal{C}$."

If the loop $\mathcal{C}$ is realized physically by a wire of resistanc R then the so-called "flux theorem" (116) states that $\Phi(t)$ and the induced current $I(t)$ stand in the following relationship:

$$IR = -\tfrac{1}{C}\dot{\Phi}(t) \tag{117}$$

This is the physics to which Faraday's induction experiments directly speak. One does not deny the utility of (117)—but does gain a more vivid sense of

---

[47] They arise from *non-conservative* $\boldsymbol{E}$-fields.

[48] For "useful" read "indispensable"!

Maxwell's physics/formal genius—when one stresses that $\boldsymbol{\nabla}\times\boldsymbol{E} = -\frac{1}{c}\dot{\boldsymbol{B}}$ is a much deeper statement than (117): it is independent of the dirty physics which underlies Ohm's law, and it tells us "what is going on" even in the absence of mobile charge. Equation (117) is susceptible also to the criticism that it *does not quite represent the facts* ... for reasons which emerge from the following

EXAMPLE: Let an impressed $\boldsymbol{B}$-field have the spatially uniform and temporally ramped structure

$$\boldsymbol{B} = \begin{pmatrix} 0 \\ 0 \\ -c\beta t \end{pmatrix}$$

encountered already on page 64, and let $\mathcal{C}$ refer to a circular wire ring of radius r and resistance R, oriented as shown in

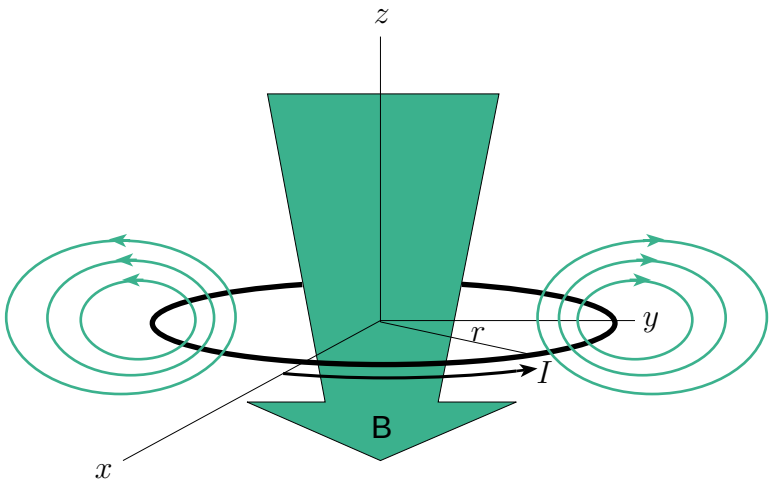


FIGURE 25: *A surging magnetic field stimulates current in the conductive ring, which generates an oppositely oriented time-dependent toroidal field, which ...*

the figure. Time-dependent flux $\Phi(t) = \pi r^2 B(t) = -\pi r^2 c\beta t$ is encircled by the wire, which by (117) induces a current $I = -(cR)^{-1}\dot{\Phi} = \pi r^2\beta/R$. That current itself generates a toroidal magnetic field $\boldsymbol{B}'$ and an associated Φ'. In general, we must take into account the so-called

$$\text{"back emf"}\ \ \mathcal{E}' = -\tfrac{1}{c}\dot{\Phi}'$$

when computing $I(t)$. We are here released from that infinite regress only because in the present (highly artificial) context $\dot{\Phi}' = 0$. Generally, however, we confront this question: *How to describe the quantitative physics of the self-interactive effect just noted*?

It is to get a handle on that issue that we digress now to acquire familiarity with the concept of

INDUCTANCE Let it be arranged/assumed that steady (!) currents $I_1, I_2, \ldots, I_N$ circulate in material loops $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_N$:

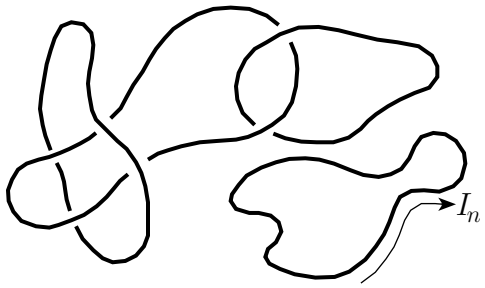


FIGURE 26: *Current-carrying loops interact magnetically. We are not concerned at the moment with the mechanism (batteries?) that in reality would be required to* maintain *the steady currents* I_n.

the figure. The $\boldsymbol{B}$-field at any point $\boldsymbol{x}$ (if we dismiss as irrelevant any $\boldsymbol{B}$-field of extrinsic origin) can be described

$$\boldsymbol{B} = \boldsymbol{B}_1 + \boldsymbol{B}_2 + \cdots + \boldsymbol{B}_n + \cdots + \boldsymbol{B}_N$$

$$\boldsymbol{B}_n \equiv \boldsymbol{B}\text{-field generated by current } I_n$$

From the Biot-Savart law (96) it follows in parlticular that

$$= I_n \cdot \underbrace{\tfrac{1}{4\pi c} \oint_{\mathcal{C}_n} \frac{d\boldsymbol{\ell} \times \hat{\boldsymbol{r}}}{r^2}}_{\substack{\text{vector-valued factor which relates} \\ \boldsymbol{x} \text{ to the geometry of } \mathcal{C}_n}} \tag{118}$$

Let

$$\Phi_{mn} \equiv \text{magnetic flux through } \mathcal{C}_m \text{ due to field generated by } I_n$$

$$= \iint_{\mathcal{D}_m} \boldsymbol{B}_n \cdot d\boldsymbol{S}_m \quad : \quad \mathcal{D}_m \text{ is any cap with } \partial \mathcal{D}_m = \mathcal{C} \tag{119}$$

Introducing (118) into (119) we conclude that Φ_{mn} is proportional to I_n through a factor which depends mutually and exclusively upon the *geometries* of the loops $\mathcal{C}_m$ and $\mathcal{C}_n$:

$$\Phi_{mn} = M_{mn} I_n \tag{120}$$

The analytical evaluation of M_{mn} is—even in simple cases—typically quite difficult[49] ... but some formal progress is possible. Appealing to (92) we have

$$\boldsymbol{B}_n = \boldsymbol{\nabla} \times \boldsymbol{A}_n$$

[49] See Richard Verbeck, "S & M Induction Formulæ" (Reed College 1981).

so

$$\Phi_{mn} = \iint_{\mathcal{D}_m} (\boldsymbol{\nabla}\times\boldsymbol{A}_n)\cdot d\boldsymbol{S}_m = \oint_{\mathcal{C}_m} \boldsymbol{A}_n\cdot d\boldsymbol{\ell}_m$$

But it follows from (95)[50] that

$$\boldsymbol{A}_n = \tfrac{1}{4\pi C}\oint_{\mathcal{C}_n} \tfrac{1}{r} d\boldsymbol{\ell}_n \cdot I_n \tag{121}$$

Thus do we obtain

$$M_{mn} = \tfrac{1}{4\pi C}\oint_{\mathcal{C}_m}\oint_{\mathcal{C}_n} \tfrac{1}{r}\, d\boldsymbol{\ell}_m\cdot d\boldsymbol{\ell}_n \tag{122}$$
$$r \equiv \text{distance between } d\boldsymbol{\ell}_m \text{ and } d\boldsymbol{\ell}_n$$

This pretty result (subject, however, to an alternative interpretation) was first achieved ($\sim$1845) by Franz Neumann[51] (1798–1895). It is known as "Neumann's formula," and carries with it the important implication that

$$M_{mn} = M_{nm} \tag{123}$$

The real numbers M_{mn}—which, though electrodynamically important, refer exclusively to the geometry and relative placement of the loops $\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_N$—are called *coefficients of mutual inductance* when $m \neq n$, and *coefficients of self-inductance* when $m = n$. In the latter case it is standard to adjust the notation:

$$M_{mm} \longmapsto L_m \equiv \text{self-inductance of the } m^{\text{th}} \text{ loop}$$

From the fact (see again page 61) that $\boldsymbol{B} \to \boldsymbol{0}$ near a "filamentary current" (current in a wire of zero radius) we conclude—the associated "self-fluxes" being anavoidably infinite—that

$$\begin{array}{l}\text{The self-inductance of a } \textit{filamentary} \text{ loop is}\\ \text{—irrespective of the loop's geometry—infinite.}\end{array} \tag{124}$$

. . . which I take to be Nature's way of reminding us that "filamentary currents" are a (physical unrealizable and) latently *dangerous abstraction*.

---

[50] The $\iiint$ ranges only over the volume of the wire, since the integrand vanishes elsewhere. Integration over cross-sections converts current density to current. The surviving integral is a $\oint$ along the length of the wire.

[51] Neumann was the *inventor of the vector potential* $\boldsymbol{A}$ (and of much else), but how he obtained (122) at such an early date—and without knowledge of Stokes' theorem—is beyond my understanding! Notice that in (122) all reference to $\boldsymbol{A}$ has dropped away.

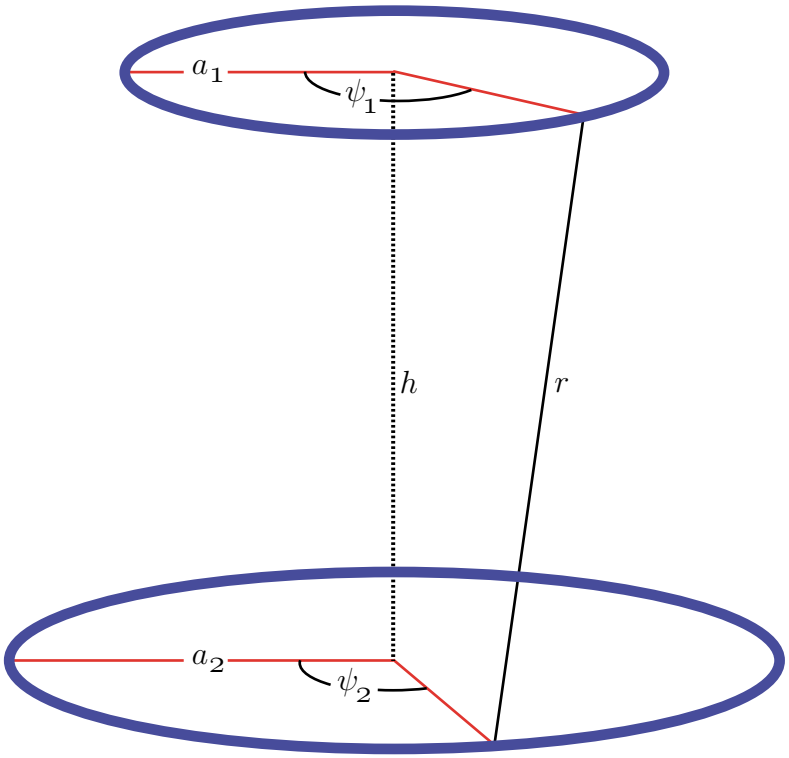


FIGURE 27: *Coaxial filamentary rings. In the text we compute the mutual inductance M—an arduous task made (barely) feasible by the high symmetry of the system.*

I turn now to review of a line of argument which leads to a description of the mutual inductance $M \equiv M_{12} = M_{21}$ of a pair of coaxial filamentary rings.[52] First we establish by geometrical argument that the distance between point ψ_1 on $\mathcal{C}_1$ and point ψ_2 on $\mathcal{C}_2$ can be described

$$r = \sqrt{h^2 + a_1^2 + a_2^2 - 2a_1 a_2 \cos(\psi_1 - \psi_2)}$$

so (122) supplies

$$\begin{aligned} M &= \tfrac{1}{4\pi c}\int_0^{2\pi}\int_0^{2\pi} \frac{a_1 a_2 \cos(\psi_2 - \psi_1)}{\sqrt{h^2 + a_1^2 + a_2^2 - 2a_1 a_2 \cos(\psi_2 - \psi_1)}}\, d\psi_1 d\psi_2 \\ &= \tfrac{1}{2c}\int_0^{2\pi} \frac{a_1 a_2 \cos\theta}{\sqrt{h^2 + a_1^2 + a_2^2 - 2a_1 a_2 \cos\theta}}\, d\theta \\ &= -\tfrac{k}{c}\sqrt{a_1 a_2}\int_0^{\frac{1}{2}\pi} \frac{\cos 2\phi}{\sqrt{1 - k^2 \sin^2\phi}}\, d\phi \end{aligned}$$

[52] See PROBLEM 28, where you are asked to work out the details of the individual steps.

where $\phi \equiv \frac{1}{2}(\theta - \pi)$ and

$$k^2 \equiv \frac{4a_1a_2}{h^2 + (a_1 + a_2)^2}$$

The integral is tabulated, and supplies

$$M = \tfrac{1}{C}\sqrt{a_1a_2}\Big\{\big(\tfrac{2}{k} - k\big)\mathbf{K}(k) - \tfrac{2}{k}\mathbf{E}(k)\Big\} \tag{125.1}$$

where

$$\mathbf{K}(k) \equiv \int_0^{\frac{1}{2}\pi} \frac{1}{\sqrt{1 - k^2\sin^2\phi}}\, d\phi$$
$$\mathbf{E}(k) \equiv \int_0^{\frac{1}{2}\pi} \sqrt{1 - k^2\sin^2\phi}\, d\phi$$

define the "complete elliptic integrals of 1st and 2nd kinds.[53] Writing

$$k^2 = \frac{4a_1a_2}{h^2}\cdot\left[1 + \frac{(a_1 + a_2)^2}{h^2}\right]^{-1} = \frac{4a_1a_2}{h^2}\left\{1 - \frac{(a_1 + a_2)^2}{h^2} + \cdots\right\}$$

we see that $k^2 \sim 0$ corresponds physically to the case in which the rings are *widely separated* ($h \gg \sqrt{4a_1a_2}$). But for k^2 small the handbooks supply[54]

$$\begin{aligned}\mathbf{K} &= \tfrac{2}{\pi}\Big[1 + 2\tfrac{k^2}{8} + 9\big(\tfrac{k^2}{8}\big)^2 + \cdots\Big]\\ \mathbf{E} &= \tfrac{2}{\pi}\Big[1 - 2\tfrac{k^2}{8} - 3\big(\tfrac{k^2}{8}\big)^2 + \cdots\Big]\end{aligned} \quad : \quad k^2 \text{ just greater than } 0$$

From these facts it follows that <u>for loosely coupled coaxial rings</u>

$$M \approx \tfrac{\pi}{16C}\sqrt{a_1a_2}\,k^3 \quad \text{with} \quad k = \sqrt{4a_1a_2}/h \tag{125.2}$$

On the other hand ... we observe that

$$1 - k^2 = \frac{h^2 + (a_1 - a_2)^2}{h^2 + (a_1 + a_2)^2}$$

which shows that $k^2 \sim 1$ corresponds physically to the case in which the rings are *very close together* ($h \sim 0$ and $a_1 \sim a_2$). The handbooks now supply[54]

$$\begin{aligned}\mathbf{K} &= \Lambda + \tfrac{1}{4}(\Lambda - 1)\kappa^2 + \tfrac{9}{64}(\Lambda - \tfrac{7}{6})\kappa^4 + \cdots\\ \mathbf{E} &= 1 + \tfrac{1}{2}(\Lambda - \tfrac{1}{2})\kappa^2 + \tfrac{3}{16}(\Lambda - \tfrac{13}{12})\kappa^4 + \cdots\end{aligned} \quad : \quad k^2 \text{ just less than } 1$$

with $\kappa \equiv \sqrt{1 - k^2}$ and $\Lambda \equiv \log(4/\kappa)$. From these (more intricate) facts it follows

[53] For discussion of the properties of these famous functions see, for example, E. Jahnke & F. Emde, *Tables of Functions* (1945), pages 73–85 or J. Spanier & K. B. Oldham, *An Atlas of Functions* (1987), Chapter 61.

[54] See Jahnke & Emde, page 73.

that for tightly coupled coaxial rings

$$M \approx \tfrac{1}{C}\sqrt{a_1 a_2}\big(\Lambda - 2\big)$$
$$\Lambda = \log \frac{4}{\sqrt{1-k^2}}$$

and that this (by $0 \sim h \ll a_1 \sim a_2 \sim a$) can be formulated

$$\approx \tfrac{1}{C} a\Big(\log \tfrac{8a}{b} - 2\Big) \tag{125.3}$$

where $b = r_{\min} = \sqrt{h^2 + (a_1 - a_2)^2}$ is the shortest distance between the two rings.

Inductance calculations can sometimes (*i.e.*, in a few favorable cases) be accomplished by more elementary means. Suppose, for example, that in the "coaxial 2-ring problem" one ring is very much smaller than the other: $a_1 \ll a_2$. The small ring lies then in the "axial region," where the magnetic field $\boldsymbol{B}_2$ generated by current I_2 in the large ring is easy to calculate:[55] one finds that $\boldsymbol{B}_2$ runs parallel to the axis (in the sense given by the $\circlearrowleft$-rule) and is of magnitude

$$B_2(h) = \frac{I_2}{2c}\frac{a_2^2}{(h^2+a_2^2)^{3/2}} I_2$$
$$= \begin{cases} \frac{I_2}{2ca_2}\Big[1 - \frac{3}{2}\big(\frac{h}{a_2}\big)^2 + \frac{15}{8}\big(\frac{h}{a_2}\big)^4 + \cdots\Big] & : \quad h \ll a_2 \qquad (126.1) \\ \frac{I_2}{2ch}\big(\frac{a_2}{h}\big)^2\Big[1 - \frac{3}{2}\big(\frac{a_2}{h}\big)^2 + \frac{15}{8}\big(\frac{a_2}{h}\big)^4 + \cdots\Big] & : \quad h \gg a_2 \qquad (126.2) \end{cases}$$

If we conceptualize the present 2-ring problem as indicated in Figure 28 then it follows immediately from (126.2) that in leading approximation

$$\Phi_{12} = \pi a_1^2 \cdot \frac{I_2 a_2^2}{2ch^3}$$

whence

$$M_{12} = \pi a_1^2 a_2^2 / 2ch^3$$

which agrees precisely with (125.2). It is instructive to note that the problem would *not* have seemed easy had we on the other hand tried to evaluate M_{21}, for while

- the field $\boldsymbol{B}_2$ intercepted by the small ring $\mathcal{C}_1$ is nearly uniform
- the same cannot be said of the $\boldsymbol{B}_1$ intercepted by the large ring $\mathcal{C}_2$.

Nevertheless—and from this point of view somewhat surprisingly—we know on general grounds that $M_{12} = M_{21}$. I turn now from the calculation of mutual inductances to the *calculation of self-inductances—a problem which (because bedeviled by ∞'s) tends to be rather more difficult.*

[55] PROBLEM 29.

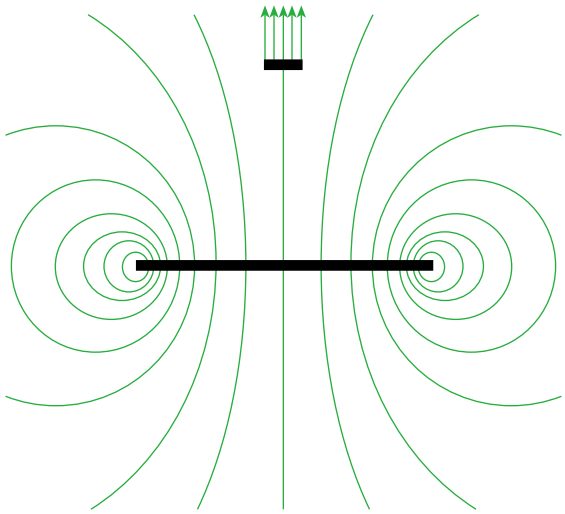


FIGURE 28: *Coaxial ring system, made exceptionally tractable by the circumstance that the small ring intercepts an essentially* uniform *sample of the* $\boldsymbol{B}$*-field generated by the large ring.*

INTERLUDE: It is useful—pedagogically, and because it will gain me a result of which I will have need—to rise for a moment above the details of battle to ask: *Why* are we studying inductance? The answer" *So that we can deal effectively with the magnetic interactions of currents confined to wires*—with one another, and with themselves.[56] And it is the *self*-interaction problem which has just moved to center stage. How are we doing? We are in position now to state that a t-dependent current I_n in $\mathcal{C}_n$ produces in $\mathcal{C}_m$ an emf given (return with (120) to (116)) by

$$\mathcal{E}_{mn} = -\tfrac{1}{c} M_{mn} \tfrac{d}{dt} I_n \tag{127}$$

Setting $m = n$, we expect the "back emf" to be given by an equation of the form

$$\text{back emf} = -\tfrac{1}{c} L \tfrac{d}{dt} I \tag{128}$$

and it is L which we desire now to compute. We proceed, as before, in terms of particular examples.

[56] We take no embarrassment from the fact that "currents confined to wires" are of more importance to people than to God and other physicists. It is, after all, by the activities of men (Faraday) that we *learn* the ways of God: our real intent at the moment is *to understand Faraday's laboratory experience*.

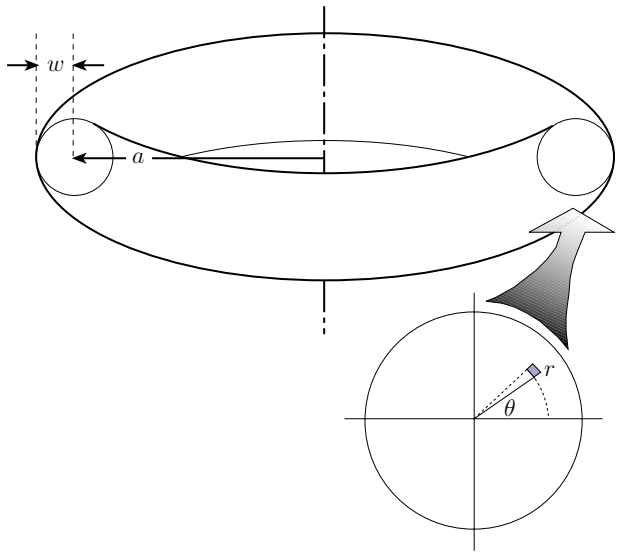


FIGURE 29: *Variables used to describe the gross form and cross-sectional elements of a conductive ring. Our objective is to compute the self-inductance of such a ring.*

Let a ring of radius a be formed from wire of radius w ($w \ll a$). We will assume $w > 0$ (*i.e.*, we exclude the filamentary idealization $w = 0$) in order

1) to avoid the ∞ mentioned at (124), and
2) the better to model engineering reality

but have purchased thus a conceptual problem: *How* to model such a wire? This we do as follows: we agree (tentatively) to ...

> Think of the ring as a "cable" made up of *filamentary sub-rings*, each of cross-sectional area $dA = rdrd\theta$. The current carried by the filament with coordinates (r, θ) can be described
>
> $$dI = j(r, \theta)rdrd\theta \tag{129}$$
>
> and we will assume that the *ratio* of the currents carried by *any pair* of filaments is time-independent; *i.e.*, that they *fluctuate in concert*. This entails
>
> $$j(r, t) = J(r) \cdot K(t) \tag{130}$$
>
> REMARK: Though it makes physical sense, it is really only for analytical convenience that I have assume the current density j to be θ-independent. In that same spirit one could—though I for the moment won't—assume further that $J(r)$ is r-independent.

Now it follows from (130) that

$$I(t) = K(t) \cdot \underbrace{\int_0^{2\pi}\int_0^{w} J(r) r\, dr d\theta}$$

It is convenient to scale $J(r)$ so this factor equals *unity*. This is in effect to write

$$j(r,t) = I(t) \cdot J(r) \tag{131}$$

and to interpret $J(r)rdrd\theta$ as the fraction of the total current I which circulates in the filament with coordinates (r,θ).

The magnetic flux $\Phi(r,\theta)$ through the (r,θ)-filament—produced by the currents circulating in all the other filaments—can be described

$$\Phi(r,\theta) = \int \underbrace{M(r,\theta;r',\theta')}\, dI'$$

This function has (see below) the structure implied by (125.2).

NOTE: In the discrete approximation we would have to write $\sum'$, signaling our intention to omit the infinite self-fields that arise when "filaments of zero cross-section" are imagined to carry finite currents. That problem *does not arise in the present context because our filaments carry currents proportional to their cross-sections.*

$$\begin{aligned} &= \Big\{ \int_0^{2\pi}\int_0^{w} M(r,\theta;r',\theta') J(r')\, r'\, dr' d\theta' \Big\} \cdot I(t) \\ &\equiv W(r,\theta) I(t) \end{aligned} \tag{132}$$

Next—looking to (127/128) for guidance—we note that temporal variation of the current I produces in the (r,θ)-filament an emf

$$\begin{aligned} \mathcal{E}(r,\theta) &= -\tfrac{1}{C}\dot{\Phi}(r,\theta) \\ &= -\tfrac{1}{C} W(r,\theta)\dot{I}(t) \quad \text{by (132)} \end{aligned} \tag{133}$$

which would stimulate a current

$$dI(r,\theta) = \frac{1}{R(r,\theta)}\mathcal{E}(r,\theta)$$

where

$$\begin{aligned} R(r,\theta) &\equiv \text{resistance of the } (r,\theta)\text{-filament} \\ &= \rho\frac{1}{r dr d\theta}2\pi(a + \underbrace{r\cos\theta}_{\text{neglect because } r \leqslant w \ll a}) \end{aligned}$$

can be used to give

$$\begin{aligned} dI(r,\theta) &= \underbrace{\frac{1}{2\pi\rho a}}_{= \frac{1}{R\cdot\pi w^2} \quad \text{where } R \equiv \text{total ring resistance}} \cdot\mathcal{E}(r,\theta)\, r dr d\theta \end{aligned}$$

So we do have

$$\begin{aligned} I = \int dI &= \frac{1}{R\cdot\pi w^2}\iint \mathcal{E}(r,\theta)\, r dr d\theta \\ &= -\frac{1}{R}\Big\{\frac{1}{\pi w^2 c}\iint W(r,\theta)\, r dr d\theta\Big\}\dot{I} \quad \text{by (133)} \end{aligned}$$

but by the effective *definition* (128) of self-inductance expect to have

$$= -\frac{1}{Rc}L\dot{I}$$

Comparison gives

$$\begin{aligned} L &= \tfrac{1}{\pi w^2}\iint W(r,\theta;)\, r dr d\theta \\ &= \tfrac{1}{\pi w^2}\iiiint M(r,\theta;r',\theta')J(r')\, r' r dr' d\theta' dr d\theta \end{aligned} \tag{134.1}$$

where according to (125.3)

$$M(r,\theta;r',\theta') = \tfrac{a}{c}\big(\log\tfrac{8a}{s} - 2\big) \tag{134.2}$$

$$\begin{aligned} s &\equiv \text{distance between } (r,\theta) \text{ and } (r',\theta') \\ &= \sqrt{r^2 + r'^2 - 2rr'\cos(\theta' - \theta)} \end{aligned} \tag{134.3}$$

Equation (134.1) is susceptible to some simplificaiton. From $\iint r dr d\theta = \pi w^2$ and $\iint J(r')\, r' dr' d\theta' = 1$ it follows almost immediately that

$$L = \tfrac{a}{c}\Big\{\log 8a - 2 - \tfrac{1}{w^2}\iiint J(r')\, r' r \log\big[r^2 + r'^2 - 2rr'\cos\theta\big] dr' dr d\theta\Big\}$$

but to obtain a more concrete result let us now assume the current to be *uniformly* distributed: $J(r') = 1/\pi w^2$. Then

$$L = \frac{a}{c}\Big\{ \log 8a - 2 - Q(w)\Big\}$$

$$Q(w) \equiv \tfrac{1}{\pi w^4}\int_0^w\int_0^w\int_0^{2\pi} r'r \log\big[r'^2 + r^2 - 2rr'\cos\theta\big]dr'drd\theta$$

and by interesting calculation[57] we obtain

$$= \log w - \tfrac{1}{4}$$

We are brought thus to the conclusion that the *self-inductance of a homogeneous ring* can—in the approximation

$$\text{wire radius} \ll \text{ring radius}$$

—be described

$$L = \frac{a}{c}\big(\log\tfrac{8a}{w} - \tfrac{7}{4}\big) \tag{135}$$

I apologize for the analytical tedium which has attended the derivation of (135), but in real physics one *frequently* encounters points of principle which can be clarified in no other way than by computational labor.[58] What have we learned?

1. Just as the electrostatic self-energy problem disappears when charged particles are modeled not as "points" but as "pithballs of finite radius," so does the magnetic self-energy problem disappear when currents are imagined to be carried not by idealized "filaments" but by more realistic "wires *of non-zero cross-section*."
2. Such realistic models serve to display self-inductance as integrated mutual inductance.
3. Our progress hinged on our willingness to make certain approximations, of which the physically most interesting was that the r-dependence of $j(r)$ could be neglected. This (in the language of Figure 11) amounts to an assumption that

   $$\text{skin depth} \gg \text{wire radius} \tag{136}$$

   In point of physical fact, skin depth decreases as frequency increases; we should therefore look upon (135) as the low-frequency approximation to a function $L(\omega)$ ... except that at *very* high frequencies—frequencies so high that

   $$\text{period} < \text{optical transit time across the circuit}$$

   we expect the very *concept* of mutual/self-inductance to lose its utility.

[57] PROBLEM 30.

[58] The formal simplicity of (135) suggests the possibility of a "simple derivation"... which—if it exists—is unknown to me.

4. Equation (135) provides a sharpened version of (124). It states that L diverges only *logarithmically* as the wire becomes filamentary ($w \to 0$). This can be understood as reflecting the fact that *logarithmic potentials give* $1/r$ *force laws, which correspond to "geometrical fall-off in a two-dimensional world"—the "world" defined by a linear source in 3-space*.We may expect the logarithmic divergence of L to pertain (not just to rings but) generally. to loops of *every* figure.
5. (Self)-inductance calculations are essentially *geometrical* in nature. They stand prior to electrodynamical calculation just as (say) moment of inertia calculations stand prior to the dynamics of rigid bodies. The question arises: Does the self-inductance of a loop stand in any invariable relationship to any other physically important "shape-sensitive" parameters (for example: the least area and/or fundamental frequency of a spanning membrane, the moments of inertia, etc.)? Can one anticipate on general grounds what happens to $L(\mathcal{C})$ when $\mathcal{C}$ is deformed? Or—see again Figure 20—what $\mathcal{C}$ will minimize $L(\mathcal{C})$?[59]

A surprisingly limited population of *analytical induction formulæ* can be found scattered (sparcely) throughout the literature—particularly the older electrical engineering literature.[49] Experimentally inclined readers may ask: If physically reliable analytical inductance formulæ are so difficult to obtain . . . *why bother*? *Why not must measure the inductance*? I would remind such readers of our primary goal, which is *review the classical basis of the claim that Maxwell's equations do in fact provided a representation of electromagnetic reality*. . . and for that we must be in position to *compare theory with experiment*. Returning now to the physical question which precipated this digression . . .

Figure 30 presents a schematic diagram of Faraday's experimental set-up (see again Figure 21). Working from the diagram, we have

$$\left.\begin{aligned} V(t) - \tfrac{1}{\mathcal{C}} L_1 \dot{I}_1 - \tfrac{1}{\mathcal{C}} M \dot{I}_2 &= R_1 I_1 \\ - \tfrac{1}{\mathcal{C}} M \dot{I}_1 - \tfrac{1}{\mathcal{C}} L_2 \dot{I}_2 &= R_2 I_2 \end{aligned}\right\} \tag{137}$$

and have interest in the currents $I_1(t)$ and $I_2(t)$ that result when the battery is switched on at time $t = 0$:

$$I_1(0) = I_2(0) = 0 \quad \text{and} \quad V(t) = \begin{cases} 0 & : \quad t < 0 \\ V & : \quad t \geqslant 0 \end{cases}$$

We confront at this point the (purely mathematical) problem of solving a coupled system of 1st-order ordinary differential equations, which can be notated

$$\mathbb{L}\, \tfrac{d}{dt} \boldsymbol{I} + \mathbb{R}\, \boldsymbol{I} = \boldsymbol{V}(t)$$

with

$$\boldsymbol{V}(t) \equiv \begin{pmatrix} V(t) \\ 0 \end{pmatrix}, \ \boldsymbol{I}(t) \equiv \begin{pmatrix} I_1(t) \\ I_2(t) \end{pmatrix}, \ \mathbb{L} \equiv \tfrac{1}{\mathcal{C}} \begin{pmatrix} L_1 & M \\ M & L_2 \end{pmatrix}, \ \mathbb{R} \equiv \begin{pmatrix} R_1 & 0 \\ 0 & R_2 \end{pmatrix}$$

[59] Questions of precisely this nature are explored in G. Polya & G. Szegö, *Isoperimetric Inequalities in Mathematical Physics* (1951)—an extraordinary monograph which I recommend very highly to your attention.

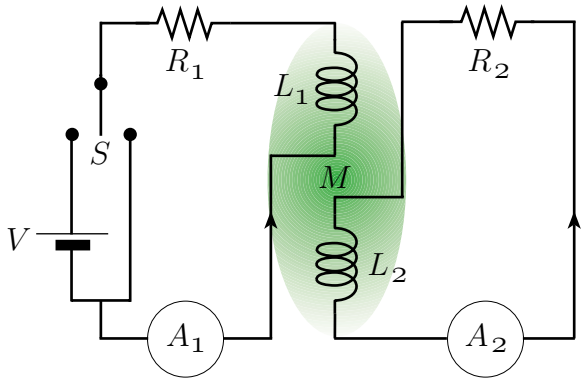


FIGURE 30: *Schematic representaion of Faraday's experimental set-up. The switch S premits one to insert/remove the battery from the primary circuit. The principal effect of the soft iron core was to increase the value of the mutual inductance M, which serves to couple the two circuits. The green shading represents magnetic field. A (usually tacit) presumption—here as always in circuit analysis—arises from the circumstance that "electromagnetic news is propagated with finite speed," that a charge will not be instantly aware of non-local field variations: we assume that* all characteristic circuit times are much greater than the time required for light to transit the circuit. *This, by the way, is the reason ultrahigh frequency devices must be physically small.*

Multiplication by $\mathbb{L}^{-1}$ is possible provided

$$\det \mathbb{L} = \tfrac{1}{c^2}\{L_1 L_2 - M^2\} \neq 0$$

which will later be shown on very general grounds to be invariably the case. We are led thus to

$$\begin{aligned} (\mathbb{D}+\mathbb{W})\boldsymbol{I}(t) &= \boldsymbol{F}(t) \quad : \quad \text{inhomogeneous differential equation} \qquad (138)\\ &\Downarrow \\ \boldsymbol{I}(t) &= (\mathbb{D}+\mathbb{W})^{-1}\boldsymbol{F}(t) \end{aligned}$$

$$(\mathbb{D}+\mathbb{W})\boldsymbol{I}(t) = \boldsymbol{0} \qquad : \quad \text{homogeneous companion of (138)}$$

Here $\boldsymbol{F}(t) \equiv \mathbb{L}^{-1}\boldsymbol{V}(t)$ and

$$\mathbb{D} \equiv \begin{pmatrix} \partial_t & 0 \\ 0 & \partial_t \end{pmatrix}, \ \mathbb{W} \equiv \mathbb{L}^{-1}\mathbb{R} = \underbrace{c(L_1L_2 - M^2)^{-1} \cdot \begin{pmatrix} L_2R_1 & MR_2 \\ MR_1 & L_1R_2 \end{pmatrix}}_{\text{elements have dimensionality of "frequency"}}$$

What meaning are we to assign to $(\mathbb{D}+\mathbb{W})^{-1}$? From the identity

$$\tfrac{d}{dt}e^{Wt}f = e^{Wt}(\tfrac{d}{dt}+W)f \quad : \quad \text{all } f(t)$$

one obtains the "shift rule"

$$\big(\tfrac{d}{dt}+W\big)\bullet = e^{-Wt}\tfrac{d}{dt}e^{Wt}\bullet \quad : \quad \text{to be read as an } \textit{operator identity}$$

of which

$$\big(\tfrac{d}{dt}+W\big)^n\bullet = e^{-Wt}\big(\tfrac{d}{dt}\big)^n e^{Wt}\bullet$$

is a corollary and (in the case $n=-1$)

$$(\mathbb{D}+\mathbb{W})^{-1}\bullet = e^{-\mathbb{W}t}\int^t e^{\mathbb{W}s}\bullet\, ds \tag{139}$$

the matrix analog.[60] The theory of linear differential equations supplies this general proposition:

> general solution of inhomogeneous equation
> = any particular solution of inhomogeneous equation
> + general solution of associated homogeneous equation

Bringing these remarks together, we conclude that the general solution of (138) can be described

$$\boldsymbol{I}(t) = e^{-\mathbb{W}t}\int_0^t e^{\mathbb{W}s}\boldsymbol{F}(s)ds + e^{-\mathbb{W}t}\boldsymbol{I}(0)$$

In the present instance $\boldsymbol{I}(0)=\boldsymbol{0}$ and $\boldsymbol{F}(t)$ is (for $t>0$) a *constant* vector, so we can *perform* the integration, and obtain[61]

$$= \frac{\mathbb{I}-e^{-\mathbb{W}t}}{\mathbb{W}}\boldsymbol{F} = \big(\mathbb{I}-e^{-\mathbb{W}t}\big)\mathbb{R}^{-1}\boldsymbol{V} \tag{140}$$

Our analytical task reduces therefore to the evaluation of $(\mathbb{I}-e^{-\mathbb{W}t})/\mathbb{W}$. This can be accomplished in a great variety of ways, two of which are described on pages 124–129 of the 1980/81 edition of these notes. The details are amusing, and of some methodological interest . . . but distract from the *physical* points at issue: here I will be content to

- assign representative values to the circuit parameters (you are encouraged to try other values) and
- entrust the computational labor to *Mathematica*.

[60] In preceding equations the $\bullet$'s are placeholders for the "operands"—*i.e.*, for unspecified functions of t

[61] PROBLEM 31.

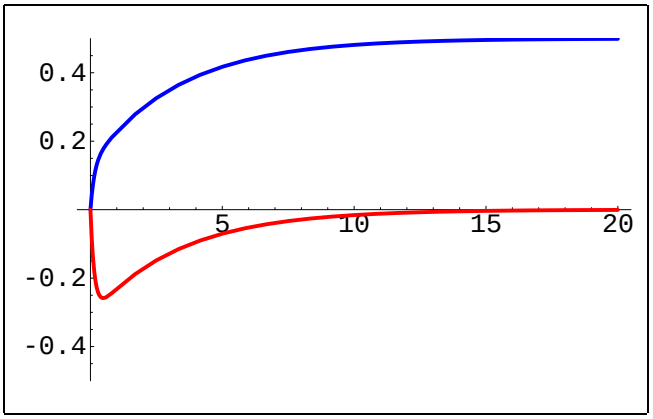


FIGURE 31: *At time $t = 0$ the battery is switched* ON *and the current I_1 in the primary circuit rises (as the sum to two exponentials, one "fast" and the other "slow") to the steady value V/R_1. The current I_2 induced in the secondary circuit is dipping transcient, present only while $\frac{d}{dt}I_1 \neq 0$.*

Specifically, I (semi-randomly) set

$$\boldsymbol{V} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \ , \ \text{unit} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \ , \ \mathbb{L} = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix} \ , \ \mathbb{R} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

which entail

$$\mathbb{W} = \begin{pmatrix} 2 & -2 \\ -4 & 5 \end{pmatrix} \quad : \quad \det \mathbb{W} \neq 0$$

The command `(unit - MatrixExp[-t W]).Inverse[R].V` instantaneously supplies

$$\left.\begin{aligned} I_1(t) &= \tfrac{1}{2}\Big\{1 - \tfrac{\sqrt{41}-3}{2\sqrt{41}}e^{-\omega_1 t} - \tfrac{\sqrt{41}+3}{2\sqrt{41}}e^{-\omega_2 t}\Big\} \\ I_2(t) &= \tfrac{1}{2}\Big\{\quad\ \tfrac{4}{\sqrt{41}}e^{-\omega_1 t} - \quad\ \tfrac{4}{\sqrt{41}}e^{-\omega_2 t}\Big\} \end{aligned}\right\} \tag{141.1}$$

where

$$\begin{aligned} \omega_1 &= \tfrac{7+\sqrt{41}}{2} = 6.70156 \\ \omega_2 &= \tfrac{7-\sqrt{41}}{2} = 0.29844 \end{aligned}$$

are observed to be precisely the *eigenvalues* of $\mathbb{W}$ (of which more later). The "primary" and "secondary" currents $I_1(t)$ and $I_2(t)$ are plotted in Figure 31. Asymtotically the system approaches a steady state, with (as is obvious already from (137))

$$I_1(\infty) = V/R_1 \quad \text{and} \quad I_2(\infty) = 0$$

That state having been achieved, let us stitch the battery *off* (and at the same time restart the clock). We then have $\boldsymbol{I}(t) = e^{-\mathbb{W}t}\boldsymbol{I}_{\text{steady}}$, which in our

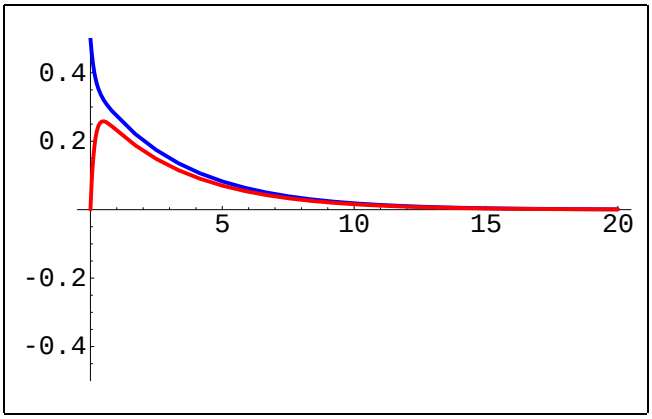


FIGURE 32: *At time $t = 0$ the battery is switched* OFF *and the previously steady current in the primary circuit drops exponentially back to zero. The current I_2 induced in the secondary circuit is now a rising transcient—again present only while $\frac{d}{dt} I_1 \neq 0$.*

numerical example supplies

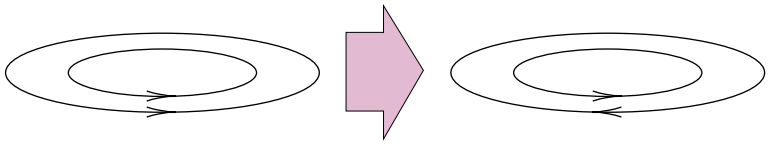
$$\left.\begin{aligned} I_1(t) &= \tfrac{1}{2}\left\{ \tfrac{\sqrt{41}-3}{2\sqrt{41}} e^{-\omega_1 t} + \tfrac{\sqrt{41}+3}{2\sqrt{41}} e^{-\omega_2 t} \right\} \\ I_2(t) &= \tfrac{1}{2}\left\{ -\tfrac{4}{\sqrt{41}} e^{-\omega_1 t} + \tfrac{4}{\sqrt{41}} e^{-\omega_2 t} \right\} \end{aligned}\right\} \tag{141.2}$$

These functions are displayed in Figure 32. It is the *contrary transcience* of the induced current (see again Figure 21) that lies at the heart of Faraday's surprising experimental discovery, and a wonder that Maxwell was able in

$$\boldsymbol{\nabla} \times \boldsymbol{E} = -\tfrac{1}{c} \tfrac{\partial \boldsymbol{B}}{\partial t} \tag{107}$$

to capture its formal essence. Several remarks are now in order:

1. The diagonal elements W_{11} and W_{22} of $\mathbb{W}$ arise from *self*-inductance, and are therefore invariably positive. But the off-diagonal elements $W_{12} = W_{21}$ refer to *mutual*-inductance, and reverse sign when we reverse either of the sign conventions attached to $\mathcal{C}_1$ and $\mathcal{C}_2$:



Mutual inductance is, in other words, *sign-indefinite and convention-dependent.*

2. The temporal aspects of Faraday induction are evidently *under the control of the eigenvalues* of $\mathbb{W}$, which in the general case[62] read

$$\begin{aligned}\omega_\pm &= \frac{R_1L_2 + R_2L_1 \pm \sqrt{(R_1L_2 + R_2L_1)^2 - 4R_1R_2(L_1L_2 - M^2)}}{L_1L_2 - M^2} \\ &= \frac{R_1L_2 + R_2L_1 \pm \sqrt{(R_1L_2 - R_2L_1)^2 + 4R_1R_2M^2}}{L_1L_2 - M^2}\end{aligned}$$

These are clearly real in all cases, but will both be positive—as is required if neither of the factors $e^{-\omega_\pm t}$ is to blow up catastrophically—if and only if

$$L_1L_2 - M^2 > 0 \tag{142}$$

Soon we will be in position to show that the inequality (142)—though not at all an obvious implication of (122)!—follows with elegant simplicity from first principles. Note that "catastrophic blow-up" would result also if the minus sign were dropped from (107): it was with that point in mind that I referred on page 63 to the "stabilizing minus sign." It is a common practice—but, as I will argue, misleading—to point to that minus sign and say "That is Lenz' law." For the minus sign is *always* correct, while Lenz' law (which might better be called "Lenz' rule of thumb") is sometimes violated.

3. Recall the statement of

> LENZ' LAW: The directionality of I_{induced} tends to be such that the resulting magnetic flux Φ_{induced} counteracts the $\dot{\Phi}_{\text{impressed}}$ from which the induced current itself derives.

The word "tends"—though as sharp as it gets in some fields—tells a physicist next to nothing. Where does it come from, and what does it mean? Replace the battery with a signal generator designed to produce some/any *prescribed* $I_1(t)$. Reading from (137) we see that the induced current $I_2(t)$ satisfies $\frac{1}{C}L_2\dot{I}_2 + R_2I_2 = -\frac{1}{C}M\dot{I}_1$ which we may write

$$\left(\tfrac{d}{dt} + \Omega\right)I_2(t) = -f(t), \text{ some prescribed function}$$

with $\Omega \equiv cR_2/L_2$ and $f(t) \equiv (M/L_2)\dot{I}_1(t)$. Arguing as on page 83 we have

$$\begin{aligned}I_2(t) &= -\int_0^t e^{-\Omega(t-s)}f(s)\,ds + e^{-\Omega t}I_2(0) \\ &= -\begin{pmatrix}\text{weighted summary of the} \\ \text{recent history of } f(t)\end{pmatrix} + \begin{pmatrix}\text{start-up transcient} \\ \text{that soon dies}\end{pmatrix}\end{aligned}$$

Evidently Lenz' law speaks to the minus sign, and is made fuzzy by the allusion to "recent history," since the "weighted summary" can be of

[62] Use the *Mathematica* command `Eigenvalues[W]`.

either sign, depending upon *details* of that history. The induced current might, in particular, be found to be flowing momentarily in the "wrong" (anti-Lenzian) direction as a kind of "inertial" effect. Arguing now in somewhat finer detail, we have

$$I_2(t) = -(M/L_2)\int_0^t e^{-\Omega(t-s)}\dot{I}_1(s)\,ds + e^{-\Omega t}I_2(0)$$

which upon integration-by-parts becomes

$$\begin{aligned} &= -(M/L_2)\Big\{I_1(t) - \Omega\int_0^t e^{-\Omega(t-s)}I_1(s)\,ds\Big\} \\ &\downarrow \\ &= -(M/L_2)I_1(t) \quad \text{as the "recall time" } \Omega^{-1} \to 0 \end{aligned}$$

In that limit we have $L_2 I_2 = -MI_1$ or (to say the same thing another way) $\Phi_{\text{induced}} = -\Phi_{\text{impressed}}$. This we might call "Lenz' exaggeration," because "short recall time" means large Ω means small L_2, and by (142) there is a limit to how small L_2 can be: $L_2 > M^2/L_1$.

The preceding discussion—which began on page 61, and has involved digressive looks at several important subtopics

- the physics of Ohmic materials
- emf
- mutual and self-inductance
- techniques for solving coupled circuit equations
- Lenz' rule of thumb

—shows that Maxwell's equations do indeed account for Faraday's experimental results. We have proceeded deductively, but the historic route was (no pun intended) *inductive* (experiment $\longrightarrow \boldsymbol{\nabla}\times\boldsymbol{E} = -\frac{1}{c}\partial\boldsymbol{B}/\partial t$) . . . and clearly required genious of an exceptionally high order. Faraday's work—as experimentalist and as intuitive father of the field concept[63]—was clearly critical to the development of Maxwellian electrodynamics. Remarkably, it opened also some doors which I have not yet mentioned.

6. Some technological & theoretical ramifications of Faraday's law. Faraday was perfectly well aware from the outset that he had discovered a point of fundamental physical principle. He was obliged, however, to leave the theoretical elaboration of his discovery to others (namely to Maxwell, who was his junior by nearly 40 years: Maxwell was, in fact, only three months old when Faraday performed his famous experiements) . . . for while Ampere was celebrated for his mathematical virtuosity, Faraday was, by his own admission, a mathematical ignoramus. The *technological* ramifications of his discovery—that one might

[63] For a good discussion, see Mary B. Hesse, *Forces & Fields: A Study of Action at a Distance in the History of Physics* (1965).

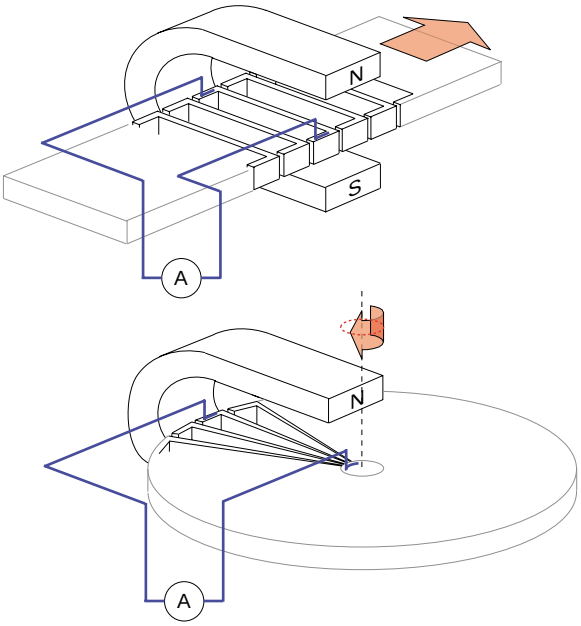


FIGURE 33: *Simple dynamos. At top: an infinite train of "staples" is dragged through the field of a permanent magnet. The magnetic flux enveloped by the circuit (shown in blue) is time-dependent, so an* emf *is developed, which produces a current. The rotational variant of the same device (below) presents no such absurdity as an "infinite train," and could actually be constructed. The placement of the magnet is, in both figures, schematic: in practice one would want to slide the magnet back until the maximal field is positioned to have the maximal effect. Notice that both devices involve sliding contacts—realized in practice by "brushes," which are a source of wear and of electrical noise. For description of a wonderfully ingenious escape from that limitation, see the Reed College thesis "A dynamo without slip rings" by Evan Wedell* (1982).

expect to be able to use not Voltaic cells but changing magnetic fields to generate practical currents—was, on the other hand, instantly apparent to Faraday (who, however, did not immediately foresee that his idea was to have profound ramifications outside the laboratory: he did not imagine rivers strangled by hydroelectric stations, forests dissected by power lines). The invention of the dynamo was essentially *simultaneous* with the discovery of electromagnetic induction, and was followed very swiftly by the first steps toward the "electrification" of the world.

In the figure I show an imagined early chapter in the history of the invention of the dynamo. The second (rotational) design is conceptually so simple that it

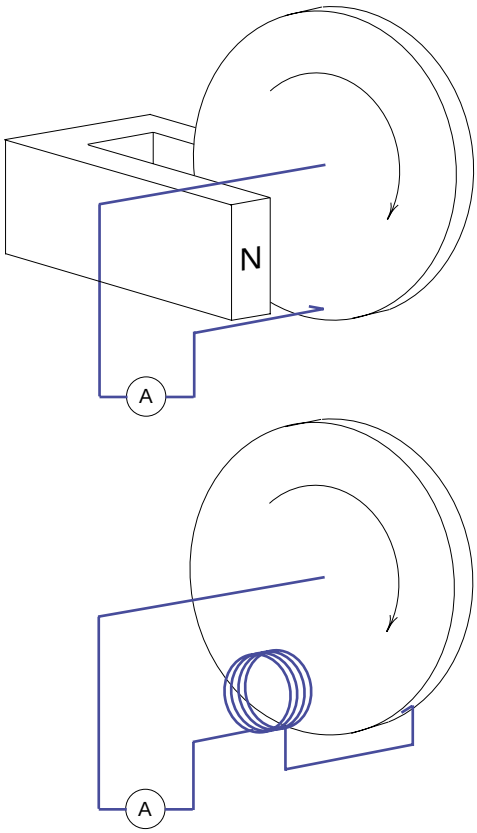


FIGURE 34: *"Homopolar disk dynamos." The design at the top can be looked upon as the "continuous limit" of the design shown at the bottom in Figure 33. Here the lower figure illustrates the operating principle of the* "self-excited *homopolar dynamo:" the permanent magnet has been replaced by an electromagnet that draws its current* from the dynamo itself.

almost "invents itself." But I think most physicists would, on intuitive grounds, be somewhat doubtful that the "homopolar disk dynamos" shown in Figure 34 —evolved from the previous design by proceeding "to the limit of infinitely many fins"—would even work, for they involve *currents which are not confined to moving wires*.[64] They depend, to be more precise, upon the (evidently quite complex) *physics of eddy currents* (interior to the rotating conductive disk). Actually, Faraday was led almost at once to the homopolar design (which works!), and for interesting reasons. In 1824 Dominique F. J. Arago had discovered that *a rotating copper disk exerts torque upon a suspended magnet*, and (see Figure 35) conversely. It seemed clear that some kind of "induced

[64] That circumstance makes it awkward to argue—at least on the geometrical face of the matter—that $\dot{\Phi} \neq 0$.

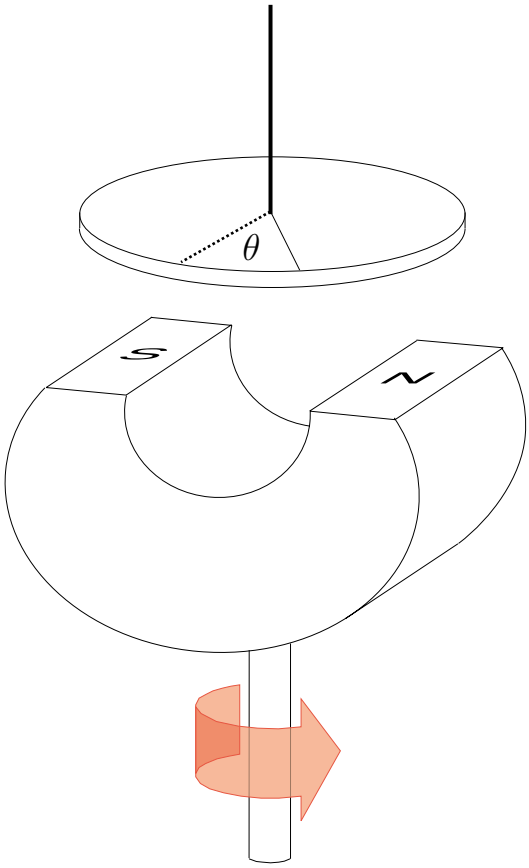


FIGURE 35: *Arago's apparatus. The copper disk is suspended by a torsion fiber. The spinning magnet is found to exert torque on the disk, but the effect is reduced/extinguished when radial slots perforate the disk. The first satisfactory account of the phenomenon was given by Faraday.*

magnetism" was involved, but this mode of explanation could not account for the observation that *when radial slots are cut in the disk the Arago effect is extinguished*. "Arago's extraordinary experiment" was much on the minds of physicists in the late 1820's, and was very well known to Faraday (to whom we owe the correct explanation: the "induced magnetism" arises from induced eddy currents, which Faraday called "whirl currents," and which the slots served to inhibit by "opening the circuits"). So Faraday had "disks on the brain." His homopolar disk dynamo can be understood as a variant of Arago's experimental configuration.

The self-excited homopolar disk-dynamo is a device of such elemental simplicity that it becomes natural to inquire whether it occurs spontaneously in Nature. I do not know enough about "biomotors" (such as twirl flagella) to know whether they provide examples, but a magnetohydrodynamic analog is

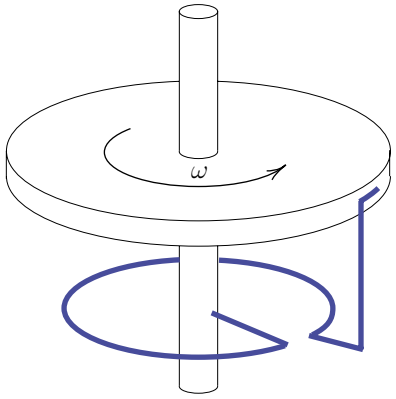


FIGURE 36: *Schematic representation of a self-exciting homopolar disk dynamo, abstracted from the device shown at the bottom of Figure 34.*

believed to be responsible for the geomagnetic field and for the magnetic fields of certain classes of stars.[65] Similar principles may operate at a galactic level. I would like therefore to consider briefly how the physics of such a device might be formalized. Proceeding in reference to Figure 36 ... let an external mechanical agency maintain the constant angular velocity ω of a conductive disk. Evidently

$$\tfrac{1}{c}L\dot{I} + RI = \mathcal{E} \tag{143}$$

where L and R refer to the self-inducatance and resistance of the electrical circuit, and where

$$\mathcal{E} \equiv \text{dynamo emf}$$

We expect[66] $\mathcal{E}$ to be proportional to the rate (set by ω) at which "filamentary constituents of the eddy current cut field lines" of the $\boldsymbol{B}$-field generated by the solenoid. Since $\boldsymbol{B}$ is proportional to I, we expect to have $\mathcal{E} \sim \omega I$. This we will express

$$= \tfrac{1}{c}M\omega I \tag{144}$$

where M is dimensionally some kind of "mutual inductance." We are in no position to *compute* M since

- we don't know *how current is distributed* in the disk (*i.e.*, we lack a "theory of eddy currents") and
- we don't know anything about the $\boldsymbol{B}$-field *interior* to the disk.

[65] See W. M. Elsasser, "Hydromagnetism," AJP **23**, 590 (1955) & **24**, 85 (1956) and "Hydromagnetic dynamo theory," Rev. Mod. Phys. **28**, 135 (1956). For a good and very detailed review of more recent work see H. K. Moffatt, *Magnetic Fields Generation in Electrically Conducting Fluids* (1978).

[66] This is clearest if one argues from the second figure on page 88.

If, however, we assume (144) to be qualitatively correct, then (143) becomes

$$\tfrac{1}{c}L\dot{I} + RI = \tfrac{1}{c}M\omega I$$

—the solution of which

$$I(t) = I_0\, e^{\frac{M\omega - cR}{L}t} \quad : \quad \begin{cases} \text{grows exponentially if } \omega > cR/M \\ \text{dies} \qquad\qquad\qquad\quad \text{if } \omega < cR/M \end{cases}$$

(Physically, exponential growth would proceed only until the rate I^2R of Joule dissipation becomes equal to the power of the external agency which drives the disk.) We conclude that the homopolar dynamo *becomes self-excited only if spun fast enough ... and not at all if spun in the wrong direction*.

There is general agreement among geophysicists that some kind of "dynamo action" (slow convection within the earth's electrically conductive core) must be responsible for the principal component of the geomagnetic field, though details of the mechanism remain inaccessible. Suggestive insight into a characteristic feature of the phenomenon—aperiodic polarity reversal—was obtained by T. Rikitake,[67] who studied the system of *coupled* disk dynamos shown in Figure 37. The two circuits are assumed to have identical resistances R and self-inductances L. The "external agency" is asked not to maintain constant angular velocity but to apply constant and identical torques N to the two disks, which are assumed to have (relative to their spin axes) identical moments of inertia A: $\omega_1(t)$ and $\omega_2(t)$ have joined $I_1(t)$ and $I_1(t)$ as functions to be determined. With these simplifying assumptions one has

$$\left.\begin{aligned} \tfrac{1}{c}L\dot{I}_1 + RI_1 &= \tfrac{1}{c}M\omega_1 I_2 \\ \tfrac{1}{c}L\dot{I}_2 + RI_2 &= \tfrac{1}{c}M\omega_2 I_1 \end{aligned}\right\} \tag{145.1}$$

which describe the *electrical* properties of the system, and

$$\left.\begin{aligned} A\dot{\omega}_1 &= N - \tfrac{1}{c}MI_1I_2 \\ A\dot{\omega}_2 &= N - \tfrac{1}{c}MI_2I_1 \end{aligned}\right\} \tag{145.2}$$

which describe its *mechanical* properties: here $-\frac{1}{c}MI_2I_1$ describes the torque which arises from the Lorentz forces experienced by the eddy current in one disk due to the magnetic field generated by the other ... and *vice versa*. The constant M quantifies the strength of that effect, and acquires its name from the circumstance that dimensionally $[M]$ = "inductance." With Rikitaki, we

[67] "Oscillations of a system of disk dynamos," Proc. Camb. Phil. Soc. **54**, 89 (1958). See also A. E. Cook & P. H. Roberts, "The Rikitake two-disk synamo system," Proc. Camb. Phil. Soc. **68**, 547 (1970) and the final pages of Moffatt's monograph.[65]

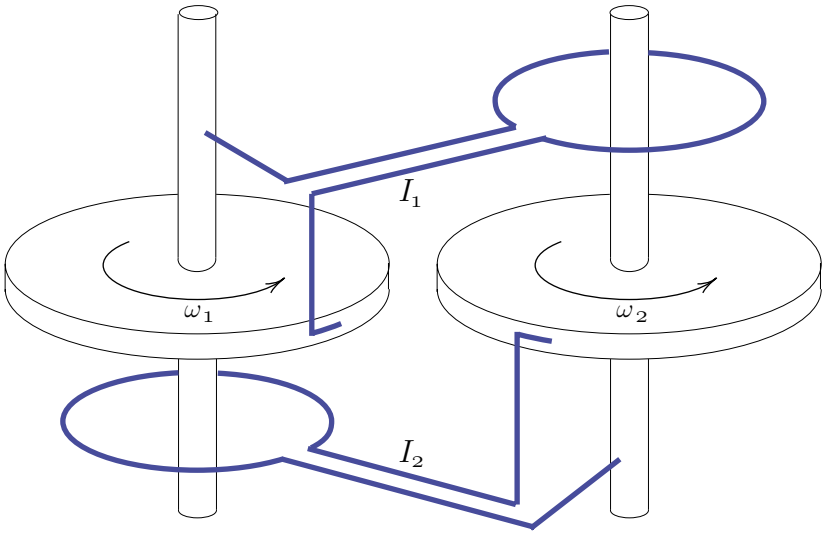


FIGURE 37: *Rikitake's system of cross-coupled disk dynamos, in which the magnetic field experienced by each results from current generated by the other. The external agency, instead of controlling the angular velocities ω_1 and ω_2 of the disks, now applies to each the same constant torque N. Simple though the system is, its behavior is shown in the text to be sometimes* chaotic.

introduce dimensionless variables

$$\tau \equiv \sqrt{NM/AL}\cdot t \quad : \quad \text{dimensionless } \textit{time}$$

$$\left.\begin{aligned} U &\equiv \sqrt{AM/NL}\cdot\omega_1 \\ V &\equiv \sqrt{AM/NL}\cdot\omega_2 \end{aligned}\right\} \quad : \quad \text{dimensionless } \textit{angular velocities}$$

$$\left.\begin{aligned} X &\equiv \sqrt{M/cN}\cdot I_1 \\ Y &\equiv \sqrt{M/cN}\cdot I_2 \end{aligned}\right\} \quad : \quad \text{dimensionless } \textit{currents}$$

and find that equations (145) can be written

$$\left.\begin{aligned} \dot{X} &= -\mu X + UY \\ \dot{Y} &= -\mu Y + VX \\ \dot{U} &= \quad 1 - XY \\ \dot{V} &= \quad 1 - XY \end{aligned}\right\} \tag{146}$$

where $\mu \equiv cR\sqrt{A/LMN}$ is a solitary adjustable parameter, and where the dot now signifies differentiation with respect to τ. Trivially $U - V = \alpha$, where α is

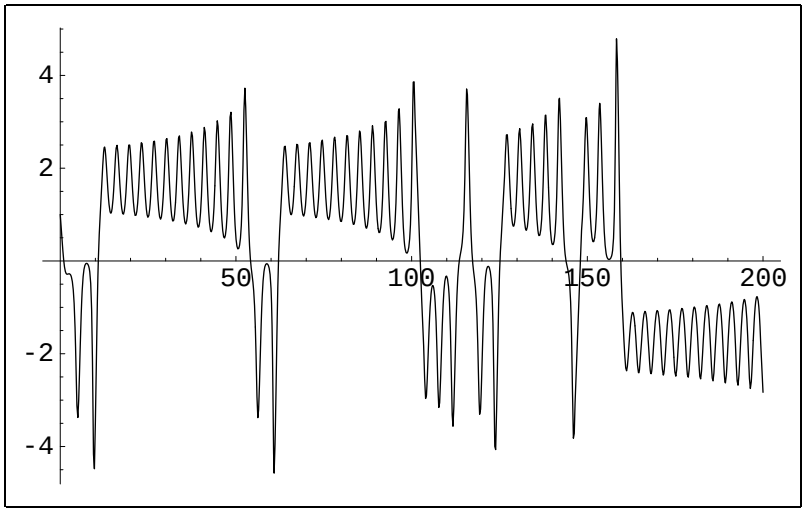


FIGURE 38: *Graph of* $X(\tau)$*, derived from (147) in the case* $\mu = 1.0$*,* $\mu = 2.7$ *with initial conditions* $X(0) = 1.0$*,* $Y(0) = 0$*,* $U(0) = 0.5$*.*

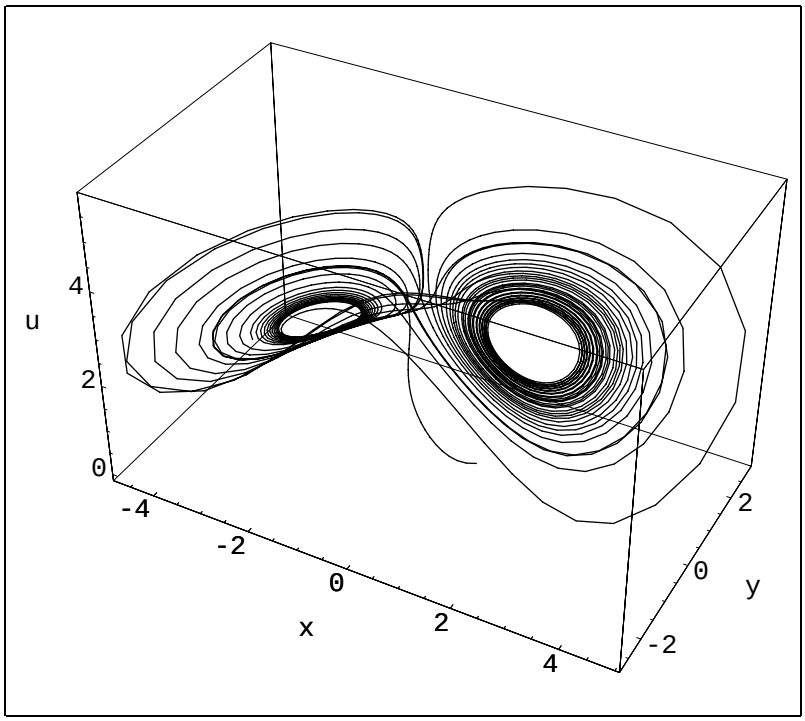


FIGURE 39: *3-dimensional parametric plot of* $\{X(\tau), Y(\tau), U(\tau)\}$ *under those same assumptions.*

a constant which we may without loss of generality assume to be non-negative. Returning with this information to (146) we obtain

$$\left.\begin{aligned} \dot{X} &= \quad -\mu X + UY \\ \dot{Y} &= -\alpha X - \mu Y + UX \\ \dot{U} &= \quad 1 \quad\quad - XY \end{aligned}\right\} \tag{147}$$

which is a triplet of coupled *non-linear* 1st-order differential equations. They defy analytical solution, must be solved numerically ... which in 1958 was a highly non-trivial undertaking, but today lies within the capability of every sophomore. In "Physicist's Introduction to *Mathematica*" (2000) I describe[68] how the resources of *Mathematica* can be brought to bear on the problem, and produce Figures 38 & 39. The point to which Rikitaki drew the attention of his geophysical colleagues was the surprising *aperiodicity of the sign reversals* evident in Figure 38.[69]

I mention finally H. Gruenberg's accidental discovery[70] of a "motor" of astounding simple design. The device can be thought of as a disk dynamo run "backwards—in motor mode."[71]

Returning our "curious devices" to the shelves from which they came, I return now to the theoretical mainline of our subject ... looking specifically to the description of the energy which resides in a magnetostatic field. We saw (pages 18–22) that the analogous electrostatic problem could be formulated as a study of the energetic details of the "source assembly process." But we have seen also (page 60) *the source* $\boldsymbol{j}(\boldsymbol{x})$ *of an magnetostatic field cannot be "assembled"*: *it must be turned on*. This is a process the energetic details of which we are only now—thanks to Faraday—in position to examine. By way of preparation ...

---

[68] See Laboratory 6, Part A

[69] Rikitaki's work did not engage the attention of the broader population of applied mathematicians. But several years later the meterological work of E. N. Lorenz led him (in "Deterministic nonperiodic flow," J. Atmos. Sci. **20**, 130 (1963)) to a triplet of equations

$$\left.\begin{aligned} \dot{x} &= -\sigma x + \sigma y \\ \dot{y} &= rx - y - rxz \quad : \quad \sigma > 0,\ r \text{ and } b > 0 \text{ are parameters} \\ \dot{z} &= xy - bz \end{aligned}\right.$$

which is structurally quite similar to (147), and which yield qualitatively similar solution curves. Lorenz' discovery contributed importantly to the development of the modern *theory of chaotic systems*, but Rikitaki's remains—even today—largely unknown.

[70] "The ball bearing as a motor," AJP **46**, 1213 (1978).

[71] See the Reed College thesis of Peter Miller: "The ball bearing motor: strange torques in spinning conductors" (1981).

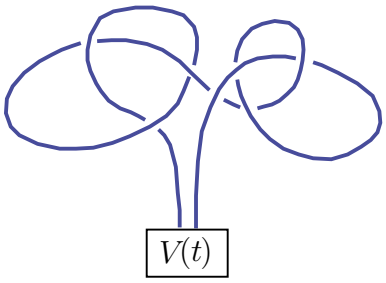


FIGURE 40: *A power supply—drawing is power from some external source (the wall socket, not shown) is used to create and maintain current in a wire loop of resistance R and self-inductance L. $V(t)$ is under the control of the experimenter.*

We have already remarked (page 67) that magnetic fields do no work on moving charges. Time-dependent $\boldsymbol{B}$-fields give rise, however, (by Faraday's law) to $\boldsymbol{E}$-fields, and $\boldsymbol{E}$-fields (of *whatever* origin) do work at a temporal rate given locally by

$$\mathcal{P} = \boldsymbol{E}\cdot\boldsymbol{j} \tag{148}$$

Suppose that $\boldsymbol{j}(\boldsymbol{x})$ refers to the steady current I which circulates in a loop of wire. We then have

$$\begin{aligned} P &\equiv \iiint_{\text{volume of wire}} \mathcal{P}\, d^3x \\ &= \oint_{\text{loop}} \left\{ \iint_{\text{cross section}} \boldsymbol{E}\cdot\boldsymbol{j}\, dA \right\} d\ell \\ &= I \oint \boldsymbol{E}\cdot d\boldsymbol{\ell} \\ &= I\mathcal{E} \end{aligned} \tag{149}$$

as a description of the *temporal rate at which* $\boldsymbol{E}$ *does work* on the charge carriers that comprise the current I. In the *steady* case the power invested by the $\boldsymbol{E}$-field is dissipated (Joule heating) at the familiar rate

$$= I^2 R$$

but in the *non-steady* case some fraction of P may be invested in the $\boldsymbol{E}$ and $\boldsymbol{B}$ fields which are associated with the capacitive and inductive features of the circuit (while another fraction may be dispatched as electromagnetic radiation). Which brings us back to the problem at hand:

Consider (Figure 40) a loop of wire (resistance R, self-inductance L) into which we have introduced an adjustable DC power supply, and let $V(t)$ denote

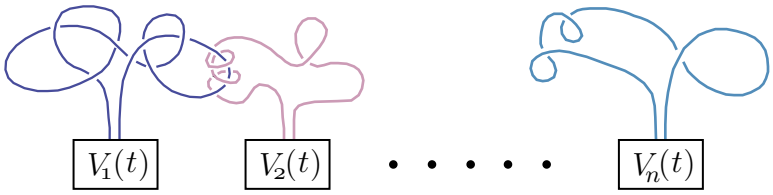


FIGURE 41: *A system of magnetically interactive circuits, each with its own power supply.*

the power supply's *output voltage at time* t (which is under our control). The current $I(t)$ can be computed from

$$V(t) = RI + \tfrac{1}{C}L\dot{I}$$

and the specified value of $I(0)$. The power supply is doing work at a rate given instantaneously by

$$\begin{aligned}\tfrac{d}{dt}W &= VI \\ &= RI^2 + \tfrac{d}{dt}\big(\tfrac{1}{2C}LI^2\big)\end{aligned}$$

so the total energy delivered by the power supply since $t = 0$ becomes

$$\begin{aligned}W \equiv \int_0^t \dot{W}\,dt &= \int_0^t RI^2\,dt + \tfrac{1}{2C}LI^2(t) \\ &= (\text{heat dissipated in } R)+(\text{energy stored in the magnetic field})\end{aligned}$$

Note that the latter term (but not the former) represents an investment which the power supply *recovers when* V *is turned down/off*. Note also that if the wire were replaced by an idealized "filament" then (since for a filament $L = \infty$) the latter term would become infinite (which is to say: the powerless supply would find itself "powerless to drive a current"): this again is the *magnetic version of the self-energy problem*.

Consider now a *system* of wire loops, each with its own adjustable power supply (Figure 41). The currents at time t can be computed from

$$\begin{aligned}V_1 &= R_1 I_1 + \tfrac{1}{C}\big\{\ \ L_1\dot{I}_1 + M_{12}\dot{I}_2 + \cdots + M_{1n}\dot{I}_n\big\} \\ V_2 &= R_2 I_2 + \tfrac{1}{C}\big\{M_{21}\dot{I}_1 + \ \ L_2\dot{I}_2 + \cdots + M_{1n}\dot{I}_n\big\} \\ &\ \ \vdots \\ V_n &= R_n I_n + \tfrac{1}{C}\big\{M_{n1}\dot{I}_1 + M_{n2}\dot{I}_2 + \cdots + \ \ L_n\dot{I}_n\big\}\end{aligned}$$

of which

$$\boldsymbol{V} = \mathbb{R}\,\boldsymbol{I} + \tfrac{1}{C}\mathbb{M}\,\dot{\boldsymbol{I}}$$

is a handy abbreviation. The power supplies are (collectively) doing work at the rate

$$\begin{aligned}\dot W &= \dot W_1 + \dot W_2 + \cdots + \dot W_n \\ &= I_1V_1 + I_2V_2 + \cdots + I_nV_n \equiv \boldsymbol{I}^{\intercal}\boldsymbol{V} \\ &\qquad\qquad = \boldsymbol{I}^{\intercal}\mathbb{R}\,\boldsymbol{I} + \tfrac{1}{C}\boldsymbol{I}^{\intercal}\mathbb{M}\,\dot{\boldsymbol{I}} \\ &\qquad\qquad = \boldsymbol{I}^{\intercal}\mathbb{R}\,\boldsymbol{I} + \tfrac{d}{dt}\big(\tfrac{1}{2C}\boldsymbol{I}^{\intercal}\mathbb{M}\,\boldsymbol{I}\big) \quad \text{by} \quad \mathbb{M}^{\intercal} = \mathbb{M}\end{aligned}$$

so (if we assume that $\boldsymbol{I}(0) = \boldsymbol{0}$)

$$\begin{aligned}W = \int_0^t \dot W\,dt &= \int_0^t \boldsymbol{I}^{\intercal}\mathbb{R}\,\boldsymbol{I}\,dt + \tfrac{1}{2C}\boldsymbol{I}^{\intercal}\mathbb{M}\,\boldsymbol{I} \\ &= (\text{heat}) + (\text{recoverable } \underline{\text{magnetic field energy}})\end{aligned}$$

The Joule dissipation term is (though physically important) for our present purposes uninteresting. Accordingly ...

We restrict our attention henceforth to the term

$$W_{\text{magnetic}} = \tfrac{1}{2C}\boldsymbol{I}^{\intercal}\mathbb{M}\,\boldsymbol{I} \tag{150}$$

This is work which our power supplies would have to perform *even if the wires were non-resistive*. Physically, it records our *effort to overcome the* emf *which results from Faraday induction*. It is (to reemphasize a point already on page 96) the effects not of $\boldsymbol{B}$ (which does no work) but of $\dot{\boldsymbol{B}}$ which lie at the foundation of (150). The question arises:

What—if any—is the relationship between (150) and the formulæ developed (on merely analogical grounds) on page 60? Returning with Neuman's formula (122) to (150) we obtain

$$W_{\text{magnetic}} = \tfrac{1}{8\pi c^2}\sum_m\sum_n \oint\oint \frac{I_m\boldsymbol{d\ell}_m\cdot I_n\boldsymbol{d\ell}_n}{r_{mn}}$$

Evidently

$$= \tfrac{1}{8\pi c^2}\iint \boldsymbol{j}(\boldsymbol{x})\cdot\boldsymbol{j}(\boldsymbol{\xi})\frac{1}{|\boldsymbol{x}-\boldsymbol{\xi}|}\,d^3x\,d^3\xi \tag{151}$$

when the $\boldsymbol{j}$-field is not confined to the interior of wires. But this is precisely (106) and, by the arguments of page 60 (traced in reverse), is known to entail

$$\begin{aligned}&= \int \mathcal{B}(\boldsymbol{x})\,d^3x \\ &\qquad \mathcal{B} \equiv \tfrac{1}{2}\boldsymbol{B}\cdot\boldsymbol{B} \quad : \quad \text{magnetic energy density}\end{aligned}$$

It is on this formal basis that we allow ourselves to state (as we did on the preceding page) that W_{magnetic} describes "energy stored *in the magnetic field*." Several comments are now in order:

1. It was emphasized on page 71 and again on page 81 that the components M_{mn} of $\mathbb{M}$ are "geometrical in nature." They summarize all that is "magnetically relevant" about the current configuration. First encountered in the description

$$\Phi_{mn} = M_{mn} I_n \tag{120}$$

of the magnetic flux which interlinks a population of current loops, their occurrence in

$$W_{\text{magnetic}} = \tfrac{1}{2C} \sum_{m,n} I_m M_{mn} I_n \tag{150}$$

is equally fundamental ... and provides in fact an efficient framework within which to address questions such as those posed on page 81.

2. M_{mn} refers more particularly to the magnetostatics of steady current loops. Since *not every* $\boldsymbol{j}$*-field admits of conceptualization as a "bundle of filamentary loops* ($\boldsymbol{\nabla\cdot j} = 0$ states that "$\boldsymbol{j}$-lines do not have ends," but that does not of itself entail loop-structure), *the concept of inductance would appear to have only limited relevance to the magnetostatics of distributed currents* (such as eddy currents).[72]

3. From results already in hand we have

$$W_{\text{magnetic}} = \tfrac{1}{2C} \boldsymbol{I}^{\intercal} \mathbb{M}\, \boldsymbol{I} = \tfrac{1}{2} \iiint \boldsymbol{B}\cdot\boldsymbol{B}\, d^3x \geqslant 0$$

from which we conclude that

$$\boldsymbol{I}^{\intercal} \mathbb{M}\, \boldsymbol{I} \text{ is a positive definite quadratic form}$$

i.e., that the inductance matrix $\mathbb{M}$ is positive definite. This amounts to a statement that the roots $\lambda_1, \lambda_2, \ldots, \lambda_n$ of the characteristic equation

$$\det(\mathbb{M} - \lambda \mathbb{I}) = 0$$

(which are the "eigenvalues" of $\mathbb{M}$, and which are—by the reality and symmetry of $\mathbb{M}$—necessarily real) are necessarily all positive: $\lambda_i > 0$. An equivalent (and—since they do not require that we *solve* the characteristic equation—more useful) set of "positivity conditions" are[73] the following:

$$L_1 > 0\,, \quad \begin{vmatrix} L_1 & M_{12} \\ M_{12} & L_2 \end{vmatrix} > 0\,, \quad \begin{vmatrix} L_1 & M_{12} & M_{13} \\ M_{13} & L_2 & M_{23} \\ M_{13} & M_{23} & L_3 \end{vmatrix} > 0\,, \ldots, \ \det \mathbb{M} > 0$$

Equivalent inequalities can be obtained by permuting the indices. Thus does

$$L_1 L_2 - M^2 > 0 \tag{142}$$

---

[72] PROBLEM 32.

[73] See G. E. Shilov, *Linear Algebra* (1977), page 209. For a particularly clear discussion of this classic result see (of all people!) Paul Samuelson, *Foundations of Economic Analysis* (1967), pages 365–375.

—seen now to be one of a large *set* of inequalities—follow "with elegant simplicity from first principles," as was asserted on page 86.[74] From this general proposition

$$\text{arithmetic mean} \geqslant \text{geometric mean}$$

it now follows, by the way, that

$$\tfrac{1}{2}(L_1 + L_2) \geqslant \sqrt{L_1 L_2} \geqslant |M|$$

4. Equation (150) can sometimes be used "backwards" to compute the elements of $\mathbb{M}$ without confronting the $\oint\oint$-complexities of (122): one uses $\frac{1}{2}\iiint B^2\, d^3x$ to compute W_{magnetic} and then infers the values of the desired coefficients M_{mn}.[75]

Let currents I_1 and I_2 flow in a pair of loops. The energy which resides in the associated $\boldsymbol{B}$-field can, by (150), be described

$$W = \tfrac{1}{2c}\Big\{L_1 I_1^2 + 2M I_1 I_2 + L_2 I_2^2\Big\} \tag{152}$$

Suppose we work to displace of one loop with respect to the other, bringing about (let us for convenience say) of a rotation-free *translation* of $\mathcal{C}_1$, as illustrated in Figure 42. Our effort has two effects:

- it *modifies the value of* M *(but not of* L_1 *and* L_2*)* and
- it produces Faraday emf's which, if uncompensated, would serve to *modify the values of* I_1 *and* I_2.

We accept as a condition of the problem that $\underline{I_1 \text{ and } I_2 \text{ are to be held constant}}$, and it is to realize that constraint (also to compensate for I^2R-losses) that we have inserted smart power supplies into the circuits. From the conditions just stipulated and (152) it follows that

$$dW = \tfrac{1}{c} I_1 I_2 dM \tag{153}$$

Working from Neumann's formula (122) we have

$$dM = \tfrac{1}{4\pi c}\left\{ \oint\oint \frac{1}{|\boldsymbol{r}_{12}+\boldsymbol{dx}|}\boldsymbol{d\ell}_1\cdot\boldsymbol{d\ell}_2 - \oint\oint \frac{1}{r_{12}}\boldsymbol{d\ell}_1\cdot\boldsymbol{d\ell}_2\right\}$$

and it was established already at (10.1) that by Taylor's theorem

$$\frac{1}{|\boldsymbol{r}_{12}+\boldsymbol{dx}|} = \frac{1}{r_{12}} - \frac{\boldsymbol{r}_{12}}{r_{12}^3}\cdot\boldsymbol{dx} + \cdots$$

so

$$= \left\{ -\tfrac{1}{4\pi c}\oint\oint \frac{\boldsymbol{r}_{12}}{r_{12}^3}\,\boldsymbol{d\ell}_1\cdot\boldsymbol{d\ell}_2\right\}\cdot\boldsymbol{dx} \tag{154}$$

[74] PROBLEM 33.
[75] PROBLEM 34.

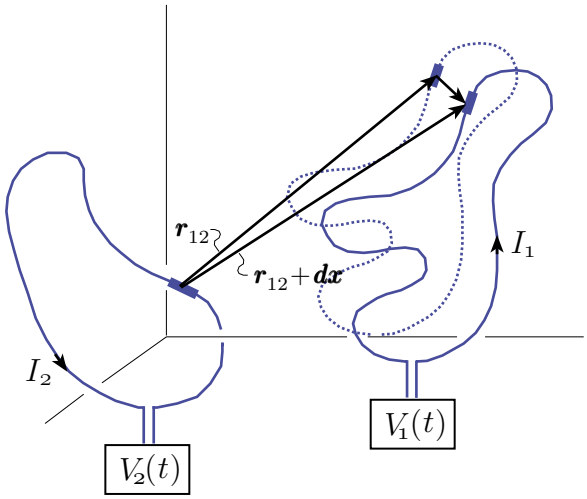


FIGURE 42: *Currents I_1 and I_2 circulate in a pair of loops. We ask*: How is the magnetic field energy altered when one of the loops is displaced with respect to the other? *Resolution of the question leads back—by a tricky argument—to Ampere's description of the* mechanical force *which one loop exerts upon the other.*

But we have encountered {etc.} before—in Ampere's description (101) of the force

$$\boldsymbol{f}_{12} = I_1 I_2 \left\{ -\tfrac{1}{4\pi c^2} \oint\oint \frac{\boldsymbol{r}_{12}}{r_{12}^3}\, \boldsymbol{d\ell}_1 \cdot \boldsymbol{d\ell}_2 \right\}$$

that circuit $\mathcal{C}_2$ exerts on circuit $\mathcal{C}_1$. Returning with this information to (154), we find that (153) can be expressed

$$dW = \boldsymbol{f}_{12} \cdot \boldsymbol{dx} \tag{155}$$

and appear to have encounted a sign problem[76] ... for the work which *we* do, struggling *against* that Amperean force—the energy which we *inject into the magnetic field*—is given by

$$dW_{\text{performed by us}} = -\boldsymbol{f}_{12} \cdot \boldsymbol{dx} \tag{156}$$

Ah! But we are not the sole workers in this story! The power supplies (over and above their obligations to pay the costs of I^2R-losses) have been *working to maintain the constancy of the currents*; *i.e.*, to compensate for the Faraday

[76] To be a physicist is to spend much of a lifetime chasing signs and errant factors of $\frac{1}{2}$, and we have encountered here a classic instance.

inductive effects that arise from the circumstance that the displacement of $\mathcal{C}_1$ takes place (not "virtually" but) in real time. Specifically

$$V_1(t) = I_1 R_1 + I_1 \dot{\Phi}_{12} = I_1 R_1 + \tfrac{1}{C} \dot{M}_{12} I_2$$
$$V_2(t) = I_2 R_2 + I_2 \dot{\Phi}_{21} = I_2 R_2 + \tfrac{1}{C} \dot{M}_{21} I_1$$

The power supplies are delivering energy at instantaneous rates given by

$$P_1(t) = I_1 V_1(t) = I_1^2 R_1 + \tfrac{1}{C} I_1 \dot{M}_{12} I_2$$
$$P_2(t) = I_2 V_2(t) = I_2^2 R_2 + \tfrac{1}{C} I_2 \dot{M}_{21} I_1$$

The I^2R terms will be dismissed as irrelevant to the present discussion: they describe energy dissipated as heat, the unrecoverable "cost of doing business." We are left with

$$\left.\begin{array}{l}\text{rate at which the power supplies are collectively}\\ \text{investing energy in redesign of the magnetic field}\end{array}\right\} = 2 \cdot I_1 I_2 \tfrac{1}{C} \dot{M}$$

where use has been made of $M = M_{12} = M_{21}$. Clearly, the argument that gave (154) gives

$$\dot{M} = \left\{ -\tfrac{1}{4\pi C} \oint\oint \frac{\boldsymbol{r}_{12}}{r_{12}^3}\, \boldsymbol{d\ell}_1 \cdot \boldsymbol{d\ell}_2 \right\} \cdot \frac{\boldsymbol{dx}}{dt}$$

The energy that the power supplies collectively/recoverably invest in time dt is given therefore by

$$\begin{aligned} dW_{\text{performed by power supplies}} &= 2 \cdot I_1 I_2 \tfrac{1}{C} dM \\ &= 2 \cdot \boldsymbol{f}_{12} \cdot \boldsymbol{dx} \\ &= 2\, dW \qquad\qquad \text{by (155)} \end{aligned}$$

In short: the power supplies collectively invest <u>twice</u> the energy dW that shows up in the redesigned magnetic field. But

$$\begin{aligned} dW &= dW_{\text{performed by us}} + dW_{\text{performed by power supplies}} \\ &= dW_{\text{performed by us}} + 2\, dW \end{aligned}$$

from which we immediately recover the desired statement (156).

The preceding argument exposes the sense in Ampere's formula (101) and Neumann's formula (122) make *equivalent statements*. We have used the latter to recover the former. Proceeding similarly, we could study the response dW of the field energy to differential *rotation* of $\mathcal{C}_1$ to obtain a description of the *torque* $\boldsymbol{\tau}_{12}$ which $\mathcal{C}_2$ exerts upon $\mathcal{C}_1$.

The argument shows that we can expect to recover

$$W_{\text{magnetostatic}} = \tfrac{1}{2C} {\sum_{i,j}}' I_i M_{ij} I_j = \tfrac{1}{2} \int \boldsymbol{B} \cdot \boldsymbol{B}\, d^3x$$

(not by "turning on" the currents in the already-positioned loops, but) by assembly of the loop system if we take sufficiently careful account of the *work done against Amperean inter-loop forces* . . . just as (on pages 19 *et seq*) we achieved

$$W_{\text{electrostatic}} = \tfrac{1}{8\pi} {\sum_{i,\,j}}' Q_i \, \tfrac{1}{r_{ij}} \, Q_j = \tfrac{1}{2} \int \boldsymbol{E}\cdot\boldsymbol{E} \, d^3x$$

by taking account of the work done against *Coulombic* forces. The idea is

1) to fabricate the loops "at infinity" and
2) there to invest the self-energy $W_{\text{self}} = \frac{1}{2\mathcal{C}} \sum_i L_i I_i^2$ (infinite, if the loops are "filamentary") required to "switch on" the currents
3) then—quasistatically—to bring the pre-assembled current-carrying loops into their desired local configuration.

It is by "pre-assembly" that we escape the absurdities (both physical and formal: see again pages 58 & 59) that would attend "snippet by snippet" assembly. If the assembly process were "brisk" rather than quasistatic then *radiative effects* would complicate the energetic analysis: a similar restriction pertains to the electrostatic assembly process, since accelerated charges radiate.

7. Recapitulation . . . and a glance ahead. We have—by Schwingerean bootstrap —"derived" Maxwell's equations, and have shown that those equations do account correctly for the experimental discoveries of Coulomb, Oersted, Ampere, Faraday . . . and for some related phenomenology. The foundations of our subject are now in our possession, and many/most of the major formal/ phenomenological ramifications have been hinted at, if *only* hinted at. We have now to examine the details . . . which is quite an assignment, for in terms of

- the subtlety and variety of the relevant points of principle
- its power to inspire mathematical invention
- the diversity and importance of its physical applications

classical electrodynamics stands apart from virtually every other branch of physics.

Here follows—for purposes of orientation—a list of some of the specialized topics into which one might want to inquire. Looking first to formal matters . . .

1. We will want to understand the sense and ramifications of the statement that *electrodynamics is a relativistic classical field theory*. Exploration of this topic leads to certain *reformulations* of the standard theory, which in specialized contexts sometimes prove useful.It leads also to sharpened perception of some fundamental points of principle. And it motivates study of some aspects of *tensor analysis*.

So far as concerns mathematical technique

2. We will want to sharpen our ability actually to *solve* Maxwell's equations. In t-independent contexts (electrostatics, magnetostatics) this objective motivates study of *potential theory* (and of associated mathematics: partial differential equations, higher functions, . . .). In dynamical contexts

the theory of potentials gives rise to the theory of *Green's functions*—a topic of practical but also of deep theoretical importance. The solution of Maxwell's equations has in recent decades acquired also an obvious *numerical* aspect.

The physical ramifications of Maxwellean electrodynamics are so diverse as to require discussion under several headings:

3. We have preferred thus far to work in vacuum, and have alluded to "stuff" only in begrudging recognition of the circumstance that the currents encountered in laboratories tend generally to be confined to wires. That bulk matter is held together by electromagnetic forces (wearing quantum mechanical hats)—and so is inevitably "electromagnetically active"—is, however, a fact of Nature which we cannot forever ignore. We are obliged, therefore, to develop an *electromagnetics of media*. This is a highly model-dependent topic, which fragments into a great variety of subtopics: the solid state physics of dielectrics, of dia/para/ferromagnetic materials, magnetohydrodynamics, ... the list goes on and on, and each subtopic can be approach in various levels of depth.
4. On the other hand ... the electromagnetic field is a highly structured and very "busy" object even in source-free regions of space. We would like to acquire detailed knowledge of the *electrodynamics of light* (physical optics, geometrical optics) ... and of the important "mathematical technology" to which this subject has given rise. We note in particular that it was upon some thermodynamic aspects of this subject that Planck based the theory of *blackbody radiation* which gave rise to quantum mechanics. Also that "optics" must be understood in a sense broad enough to include *radio*. In this connection ...
5. We would like also to study details of the *radiation production/detection process* and of related topics (scattering theory, antenna theory). At issue here is the physics of *fields and sources in dynamic interaction* ... which is electrodynamics in its purest/deepest form. It is at this level that the conceptual limitations of classical electrodynamics come most sharply into focus. The subject exhibits a marked "proto-quantum mechanical" tendency, and inspires some of the imagery fundamental to the physics of elementary particles.
6. If we consider (not the sources but) the *field to be prescribed* then we confront the question: "What is the motion of a charged particle in an impressed field (electron optics, accelerator design)?" And what, more generally, can be said concerning the motion of bulk samples (solid/liquid/gas) of "electromagnetically active" matter?

What I have been describing are some of the principal limbs of a large tree, that dominates its central place in a dense forest. We are not surprised that the limbs, on closer scrutiny, resolve into branches, the branches into twigs ... that intricately intertwine, forming shifting patterns ... which, however, will remain impossible even to begin to describe until we acquire a command of some of the details.