Electrodynamics from Ampère to Einstein

OLIVIER DARRIGOL

Centre National de la Recherche Scientifique, Paris

OXFORD
UNIVERSITY PRESS

This book has been printed digitally in order to ensure its continuing availability

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Singapore Taipei Tokyo Toronto

with an associated company in Berlin

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

First published 2000
Reprinted 2002

A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
Darrigol, Olivier.
Electrodynamics from Ampère to Einstein / Olivier Darrigol.
p. cm.
Includes bibliographical references and index.
1. Electrodynamics-history. I. Title.
QC630.5.D37 1999 537.6'09—dc21 99-049544

ISBN 0-19-850594-9

1

Foundations

1.1 Introduction

In the early nineteenth century electricity was already a wide research field, with diverse methods and multiple disciplinary connections. The oldest and best understood part of the subject was frictional electricity, especially its distribution over conductors and its mechanical effects. In his celebrated memoirs of the 1780s, Charles Coulomb, a military engineer, had founded quantitative electrostatics (later named so by Ampère). He posited two electric fluids, positive and negative, asserted the inverse square law by means of his celebrated torsion balance, and developed its consequences for the equilibrium of conductors in simple configurations. In 1812 Siméon Denis Poisson, one of the first *polytechniciens*, completed the mathematical apparatus of Coulomb's theory. He borrowed from Lagrange's and Laplace's works on gravitation what we now call the potential (V), wrote the corresponding differential equation ($\Delta V + 4\pi\rho = 0$, where ρ is the charge density), solved it in simple cases, and improved the agreement of the theory with Coulomb's experimental results.[1]

Coulomb and Poisson's electrostatics fitted excellently the Laplacian scheme which then dominated French physics. Laplace and his disciples sought to reduce every physical phenomenon to central forces acting between the particles of ponderable and imponderable fluids, in analogy with gravitation theory. In other countries, the number, function, and reality of the electric fluids were controversial issues. The British and the Italians preferred Benjamin Franklin's single-fluid hypothesis, which lent itself equally well to quantitative analysis, as Henry Cavendish had shown before Coulomb. Some of them preferred no fluid at all, or at least avoided direct action at a distance with notions reminiscent of eighteenth century electric 'atmospheres.'[2]

[1] Coulomb 1784–1788; Poisson 1811, 1813. Cf. Whittaker 1951: 57–9, 60–2; Heilbron 1979, 1982: 225–8, 236–40; Blondel 1982: 13–16; Gillmor 1971 (on Coulomb); Blondel and Dörries 1994 (on Coulomb's balance); Grattan-Guinness 1990, Vol. 1: 496–513 (on Poisson).

[2] On Laplacian physics, cf. Crosland 1967; Fox 1974; Heilbron 1993; Grattan-Guinness 1990, Vol. 1: 436–517. On singlism/dualism, cf. Heilbron 1982: 213–18, 228–34 (Cavendish); Blondel 1982: 14–15. On alternative views, cf. Heilbron 1981.

In Germany, the few marginal followers of Friedrich von Schelling's *Naturphilosophie* criticized the general notion of fluids acting at a distance, and sought a deeper unity of nature that would relate apparently disconnected phenomena. They favored a dynamistic, anti-Newtonian view of physical interactions in which matter and force were not to be distinguished: matter was only a balance of two opposite forces, and every action at a distance was to be reduced to a propagating disturbance, or polarity, of this balance. Although these romantic speculations at times bore fruit, they contradicted the basic empiricism of contemporary German physics. For quantitative studies of electricity, the Newtonian fluid theories were the only suitable basis.[3]

The same can be said of magnetism. The chief quantitative theory of this subject was again Coulomb's, based on the assumption of two fluids (austral and boreal) obeying the inverse square law. Most ingeniously, Coulomb explained the impossibility of isolating a magnetic pole by assuming that the magnetic fluids were permanently imprisoned within the molecules of magnetic bodies. His magnetic measurements, however, seemed less reliable than his electric ones, and the arguments in favor of the magnetic fluids were less direct than in the electric case. Hence Coulomb's magnetic theory met more skepticism than his theory of electricity. Yet the analogy between the two theories appealed to Laplace's disciples. Well after Ampère had proposed a contradictory view of magnetism, Poisson applied his mathematical arsenal to Coulomb's view of magnets.[4]

The most popular electric topic was galvanism. It suddenly blossomed in 1800, with Alessandro Volta's discovery of the electric pile. Volta himself regarded the tension and discharge of the pile as an electric phenomenon, therefore belonging to physics. However, other disciplines capitalized on this astonishing device. Its physiological effects and medical applications were intensively pursued, in line with the frog's contribution to Luigi Galvani's discovery. The British discovery of electrolysis attracted the chemists' attention, so that electricity was commonly regarded as a part of chemistry.[5]

In conformity with Volta's original intuition, the electrical, thermal, physiological, and chemical effects of the pile turned out to be the same as those of frictional electricity. It was usually agreed that Volta's device behaved like a battery of Leyden jars that had the mysterious ability to spontaneously recharge itself. When the poles of the pile were connected by a conductor, the discharge unceasingly repeated itself, so that its effects were permanent. In this picture only the state or the pile before discharge seemed amenable to quantitative studies. This may in part explain why quantitative studies of the galvanic current were so scarce before the 1820s.[6]

[3] On *Naturphilosophie*, cf. Caneva 1978; Blondel 1982: 29–30; and Jungnickel and McCormmach 1986, Vol. 1: 27–8 for German rejection.

[4] Coulomb 1785; Poisson 1826. Cf. Whittaker 1951: 59–60, 62–5; Blondel 1982: 16–18; Heilbron 1982: 87–8; Grattan-Guinness 1990, Vol. 2: 948–53 (on Poisson).

[5] Cf. Whittaker 1951: 67–75; Heilbron 1982: 233–6; Blondel 1982: 19–22.

[6] Cf. Brown 1969: 64; Blondel 1982: 21–2; Heilbron 1982: 196.

Beyond the Leyden jar analogy, there were deep disagreements on the cause and nature of the pile's activity. Volta proposed that the electric tension originated in the contact between two different metals. In a series Cu/Zn/mp/Cu/Zn/mp/Cu/Zn . . . (Cu = copper, Zn = zinc, mp = moist paper), the role of the moist paper was simply to avoid the contact Zn/Cu—which would cancel the effect of the previous Cu/Zn contact—without preventing the passage of electricity. Volta verified this assumption by showing that two insulated disks of copper and zinc exhibited opposite electric charges after having been brought in temporary contact. French mathematicians approved Volta's view, in which they saw an opportunity to reduce galvanism to electrostatics. The Swedish chemist Jöns Jacob Berzelius founded his popular doctrine of chemical combination on intramolecular Volta-forces.[7]

The contact theory was less fortunate in England. The leading chemist Humphry Davy found many reasons to assume that chemical changes were responsible for the electric power of the pile. Not only were the pile's effects always accompanied by chemical processes, but the force of the pile appeared to be related to the affinities of the involved chemicals. Davy exploited the latter finding to construct new kinds of pile. He also proposed a mechanism for electrolysis, and suggested, before Berzelius, that chemical forces were of electrical origin.[8]

Altogether, the new science of galvanism offered a striking contrast with electrostatics and magnetism. The latter subjects had reached a state of perfection and were proudly displayed by the French as major achievements of their mathematical physics. On the contrary, galvanism was a rich, disorganized field, growing in multiple directions (physical, chemical, physiological, and medical), but mostly escaping mathematical analysis. In 1820 a radical change occurred: the discovery of electromagnetism suddenly brought galvanism and magnetism in contact, and blurred the methodological and socio-professional borders that separated the two topics. After a summary of Oersted's discovery, the present chapter offers an analysis of Ampère's and Faraday's resulting works that founded electrodynamics.

1.1.1 Electromagnetism

Despite the mathematical analogy of their fundamental laws of equilibrium, electricity and magnetism were generally thought of as completely disconnected phenomena. Their causes and their effects were utterly different: electrification required a violent action and implied violent effects such as sparks and thunder, whereas magnetism seemed a very quiet force. The magnetizing effect of thunder, which had long been known, was regarded as a secondary effect of mechanical or thermal origin. Yet in 1804 an illuminated *Naturphilosopher*, Johann Ritter, believed that he had found an action of the open pile on the magnet, and even

[7] Cf. Whittaker 1951: 71–2; Brown 1969: 76–82 (on the French theory); Blondel 1982: 22–3; Whittaker 1951: 78–9 (on Berzelius).

[8] Cf. Whittaker 1951: 74–6; Blondel 1982: 25–7.

announced the electrolysis of water by magnets. He was soon ridiculed by the French demolition of his claims. Anyone who knew of this episode and assumed distinct fluids for electricity and magnetism was naturally predisposed against similar attempts.[9]

In July 1820, Hans Christian Oersted, a Danish Professor and a friend of Ritter, sent to the leading European physicists a Latin manuscript with the stunning title: *Experimenta circa effectum conflictus electrici in acum magneticam*. Immersed in the depths of German *Naturphilosophie*, he had long expected a connection between electricity and magnetism. He understood the galvanic current as a propagating alternation of decompositions and recompositions of the two electricities, and made this 'electric conflict' the source of heat, light, and possibly magnetism. No more needs to be said of Oersted's philosophy, given that the leading explorers of electromagnetism did not bother to investigate it further.[10]

Most of Oersted's fundamental text was a precise description of a number of experiments performed with a galvanic source, a connecting wire, and a rotating magnetic needle. For the galvanic apparatus, he followed a recipe by Berzelius: 20 copper–zinc cells filled with a sulfo-nitric mixture. He made sure that the wire turned red when connected to the apparatus, as a test of strong electric conflict. He suspended the magnetic needle as is usually done in a compass, let it assume its equilibrium position along the magnetic meridian, approached the wire and connected it to the pile.[11]

In the first of Oersted's experiments, the wire is above the needle and parallel to it. If the Northern extremity of the wire is connected to the negative pole of the pile, the North pole of the needle moves toward the West.

Next, Oersted displaced the wire toward the East or the West, and observed the same action, though a little weaker. He commented: 'The observed effect cannot be attributed to an attraction, because if the deviation of the needle depended on attractions or repulsions, the same pole should move toward the wire whether the latter be on the East side or on the West side.'[12]

Oersted then varied the respective orientations of needle, wire, and magnetic meridian. Two of the resulting experiments deserve special mention, because of their resemblance to later observations by Ampère and Faraday. In the first, the wire is vertical with its lower extremity connected to the positive pole of the pile, and it faces the North pole of the needle. Then this pole moves toward the East. If instead the wire, being still vertical, faces one side of the needle (East or West), between the North pole and the center of the needle, the North pole moves toward the West. In the other interesting experiment, the wire is bent to a vertical U-shape. Then each face of the U attracts or repels the poles of the needle.[13]

From his observations Oersted drew three essential conclusions:

[9] Cf. Blondel 1982: 27–30.
[10] Oersted 1820; 1812, 1813 for the electric conflict. Cf. Meyer 1920; Stauffer 1957; Caneva 1980; Heilbron 1981: 198–9.
[11] Oersted 1820: 215. [12] Oersted 1820: 216. [13] Oersted 1820: 217.

1. The electric conflict acts on magnetic poles.
2. The electric conflict is not confined within the conductor, but also acts in the vicinity of the conductor.
3. 'The electric conflict forms a vortex around the wire.'

To justify the third point, Oersted argued:[14]

> Otherwise one could not understand how the same portion of the wire drives the magnetic pole toward the East when placed above it and drives it toward the West when placed under it. An opposite action at the ends of the same diameter is the distinctive feature of vortices.

Finally, Oersted proposed to complete the picture of the electric conflict in accordance with the vorticity of the magnetic action:

> All the effects we have observed and described on a North pole are easily explained by assuming that the negative electric force or matter follows a *dextrorsum* spiral and acts on the North pole without acting on the South pole. The effects on a South pole are explained in a similar manner by assuming that the positive electric matter moves in the opposite direction and acts on the South pole without acting on the North pole.

The botanic term *dextrorsum* (defining the helicity of climbing plants) did not survive the competition of Ampère's *bonhomme* or Maxwell's cork-screw. But it was the first of the mnemonic devices that physicists proposed for the polarity of the electromagnetic action. From the beginning, Oersted placed the circle–axis duality at the heart of electromagnetism.[15]

In retrospect, Oersted's observations were accurate and his conclusions insightful. He understood the impossibility of reducing electromagnetism to magnetic attractions or repulsions, and yet saw how to mimic such interactions by curving the conjunctive wire. Most important, he perceived that the action of a rectilinear wire on a magnetic pole was a circular one, centered on the wire. Some features of his memoir, however, hindered a full grasp of its contents. He did not provide any figures or diagrams. He operated in conditions for which the electromagnetic effect is comparable to the magnetic action of the Earth, and therefore reached his general conclusions indirectly, by mentally subtracting the effect of the Earth. He formulated these conclusions in terms of a specific picture of galvanic currents, although his description of individual experiments was purely operational. The essential idea of a circular action appeared only in the context of the electric conflict, an alien notion for most of Oersted's readers.

Despite these obscurities, the astonishing claim of an action between a galvanic current and a magnet was easy to confirm. Within a few weeks, the world's best philosophers entered the attractive lands of electromagnetism. Most of them tried to reduce the new phenomenon to a temporary magnetism of the wire. In this way, they could ignore Oersted's dubious speculations on the electric conflict

[14] Oersted 1820: 218.
[15] Oersted 1820: 218. For a philosophical analysis of the role of axis–loop duality, cf Châtelet 1993.

and apply their previous knowledge of magnetic forces. Yet the two men who most influenced the subsequent history of electromagnetism did not follow this natural course.[16]

1.2 Ampère's attractions

The first exception was André-Marie Ampère, a mathematician with an interest in theoretical chemistry and a passion for philosophy. For physics he had done little, save his early unpublished questioning of the principles of electricity and magnetism. The news of Oersted's discovery changed his fate at age 45. In the Summer of 1820 he launched himself into frenetical researches that would make him, according to Clerk Maxwell's judgment, 'the Newton of electricity.'[17]

1.2.1 Undoing the magnet

Ampère first noted the complication of Oersted's experiments due to the magnetic action of the Earth. He conceived what is now called an astatic needle, that is, a magnetic needle whose rotation plane can be made perpendicular to the action of the Earth. In this configuration the orientation of the needle depends only on the action of the wire. Ampère found the needle to be at a right angle to the shortest line joining the center of the needle to the wire. Here was a simple fact of electromagnetism from which Oersted's more complex observations could be derived.[18]

Then Ampère looked for a similar effect produced by the voltaic battery itself. The experiment was by no means superfluous, because of the lack of consensus on the workings of the battery: the existence of a current within the battery was an open question. Ampère thus formed the concept of a 'circuit' in which 'the electric current' was closed. At the same time, he turned the suspended magnetic needle into a universal current detector, which he soon named a 'galvanometer.'[19]

At that stage Ampère reflected:

> Granted that the order in which two facts have been discovered does not make any difference in the available analogies, we could suppose that before we knew about the South–North orientation of a magnetic needle, we already knew the needle's property of taking a perpendicular position to an electric current [...]. Then, for one who tries to explain the South–North orientation, would not it be the simplest idea to assume in the Earth an electric current?

In this view the Earth's magnetic property was reduced to an electric current circulating along the parallels of the Earth. Ampère further imagined that the

[16] For the early reception of Oersted's discovery, cf. Meyer 1920: 101–8; Heilbron 1981: 199–204; Blondel 1982: 44–8.

[17] Maxwell 1873a: #528. On Ampère's biography, and for an accurate bibliography, cf. Hofmann 1995.

[18] Ampère 1820a, 1820b. Cf. Blondel 1982: 69–70; Hofmann 1995: 236–8; Steinle 1998: note 20.

[19] Ampère, 1820a, 1820b. Cf. Blondel 1982: 72–3.

heterogenous composition of the Earth along a parallel made a natural electric pile closed on itself, a device of which he had just proved the magnetic activity.[20]

Ampère then reverted to the analogy between the Earth and a magnet to deduce that every magnet owed its properties to the existence of closed currents in its mass. As a corollary, electric currents had to possess all the properties of a magnet. In particular, an electric current had to attract or repel a magnetic needle. Presumably, a current running in a flat spiral or in a helix would present a North pole and a South pole. Ampère reported these reflections to the French Academy on 18 September, only a few days after Oersted's effect had been demonstrated there, and before he had proven anything but the magnetic action of the current in a battery and the power of an electric current to attract a magnetic needle hung by a thread.[21]

Ampère's new theory of magnetism matched the philosophy of his early unpublished attempts at reforming electricity and magnetism.[22] He believed that a theory based on different kinds of fluids lacked the unity that should be found in God's plans of the universe. There had to be a single fundamental force, preferably one excluding direct action at a distance. The new concept of the magnet was a first step in the right direction, since it eliminated the magnetic fluids. This opinion contradicted Laplacian orthodoxy. Ampère strove, however, to meet other criteria of French mathematical physics. He wished to establish his theory on firm experimental grounds and to cast it in an irreproachable mathematical form.

On 25 September, Ampère showed to the skeptical Academicians that flat helical currents attracted each other and responded to a bar magnet. He had ordered the rather sophisticated apparatus from a competent mechanician. The essential difficulty was to feed the current into the helix without impeding its mobility. Ampère's universal expedient consisted of small mercury cups, in which the extremities of the mobile part of the circuit could rotate and the contact with the battery wires was simultaneously made. With his rotating helices, Ampère believed he had given a 'definitive proof' of the equivalence between magnets and current. Later in the month, he obtained a better imitation of a bar magnet with a helix of current suspended in its middle (Fig. 1.1).[23]

1.2.2 The physical current elements

Ampère's investigations then took a more analytical turn. From the beginning of his researches he expected the interaction between two currents to be analyzable in terms of current elements. Experimentally, this involved the attraction (repulsion) between two portions of parallel (antiparallel), rectilinear currents, demonstrated in October 1820. His device is represented in Fig. 1.2. Except for the mercury cups (R, S, T, U, X, Y) and the surrounding glass box, the construction of the device was entirely dictated by the necessity of isolating the interaction of two current elements, here AB and CD, from the action of the rest of the circuit to which they belong. The

[20] Ampère, 1820b: 238. [21] Ampère 1820a, 1820b. [22] Cf. Ampère [1801].
[23] Ampère 1820a, 1820b. Cf. Blondel 1982: 75–6; Hofmann 1995: 242–4.

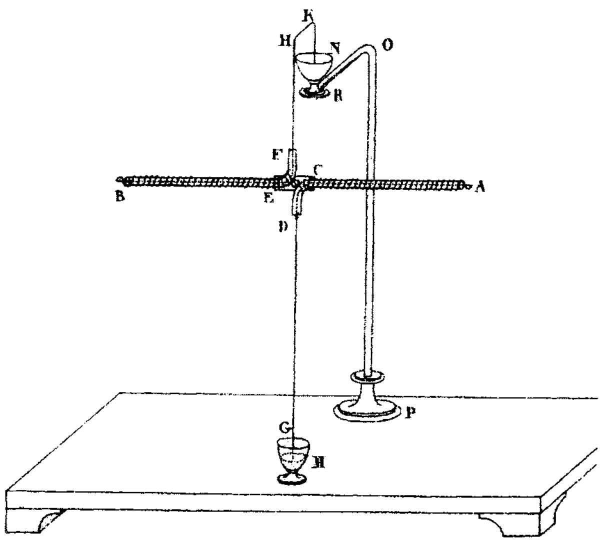

FIG. 1.1. Apparatus for showing the equivalence between a helical current and a bar magnet (Ampère 1820b).

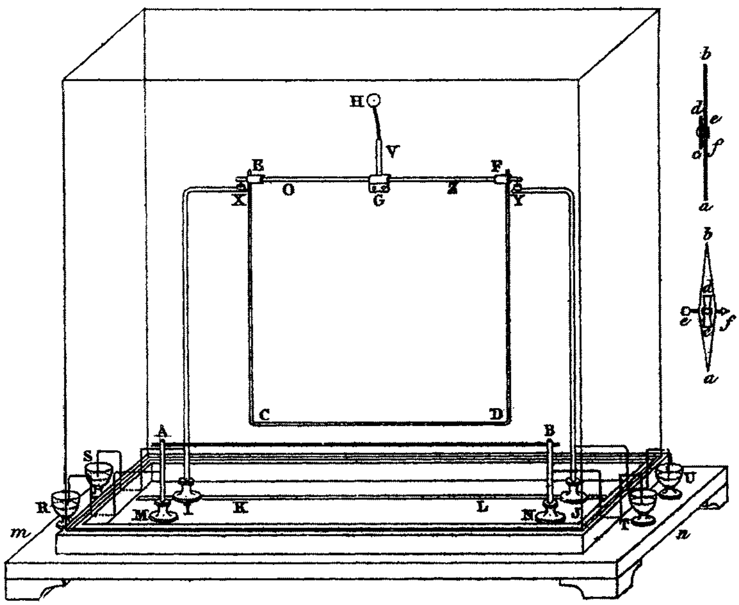

FIG. 1.2. Apparatus for showing the action between two parallel rectilinear currents (Ampère 1820b).

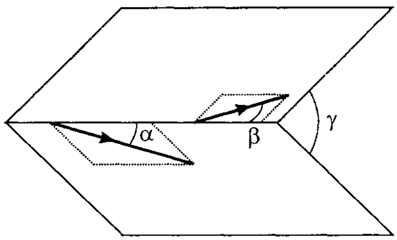

FIG. 1.3. Angles determining the relative orientation of two current elements.

segment AB is longer than CD, and the counterweight H is adjusted so that at the equilibrium position (without currents) CD is very close to AB. Then the action of AB on CD dominates all other electrical actions, and determines the rotation of CDEF around the (non-conducting) axis EF.[24]

In conformity with this concrete possibility of isolating two portions of current, Ampère ascribed a separate physical existence to the force between two current elements. Consequently, he made this force comply with the equality of action and reaction, and he had it lie on the line joining the elements.[25] For information on the angular dependence, he used a device in which the two rectilinear current were free to rotate in planes perpendicular to the line joining their centers. In October he guessed that in the most general configuration the force between two current elements was proportional to

$$\frac{\cos\gamma\sin\alpha\sin\beta}{r^2}, \tag{1.1}$$

the three angles being defined in Fig. 1.3. Analogy with gravitational forces dictated the dependence on the distance r of the two elements, the central character of the forces, and the exclusion of elementary torques. Simplicity, the need to retrieve the properties of magnets, and the two experiments on rectilinear currents suggested the angular dependence.[26]

In the same month, Ampère designed a torsion balance that could measure the force between two current elements in any geometrical configuration, and thus test his conjectured formula. He soon gave up the project, presumably because the variability of his battery and the friction in the mercury cups prevented sufficient precision.[27]

[24] Ampère 1820a (mémoire du 9 octobre), 1820b. The electric forces acting on EC and FG have, to a sufficient approximation, no influence on the motion of the moving part ECDF of the circuit, because the corresponding torque is negligible (assuming with Ampère that the forces between two elements are parallel to the line joining the elements).

[25] Ampère later explained that the forces between two elements had to be central, because if they were not, a perpetual motion could be obtained by rigidly connecting the two elements (Ampère 1826b: 1–2).

[26] Ampère 1820a: 247–8 (mémoire du 9 octobre) has only a brief summary of his analysis. The full version is in Ampère [1820c]. Cf. Blondel 1982: 83–5; Hofmann 1995: 239–41.

[27] Ampère 1820b for the description of two devices of this kind, the first of which was shown to Biot and Arago on 17 October. Cf. Blondel 1982: 84–5; Hofmann 1995: 245–6.

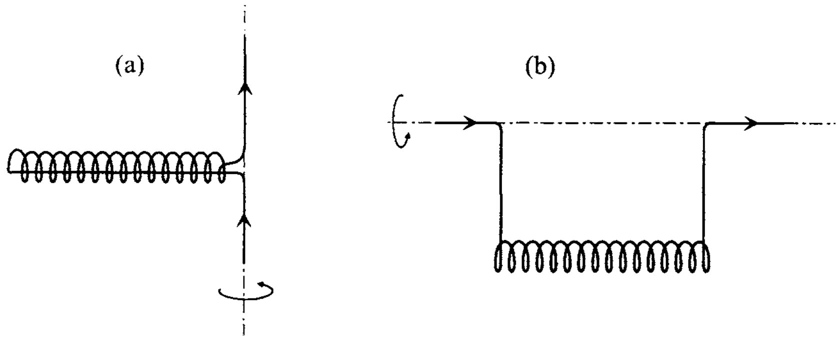

FIG. 1.4. Two kinds of helical current.

Ampère needed another way to justify his fundamental formula. Ironically, he found it in the failure of a less ambitious experiment in which he tested the interactions of two parallel helices. To his surprise, the helices acted like parallel wires instead of imitating parallel bar magnets. He soon recognized the source of the anomaly. In his earlier experiments with a helix, the current was brought to the helix in the manner of Fig. 1.4(a) that permitted rotation around the vertical axis. In the new experiment, it was brought in the manner of Fig. 1.4(b), which permitted rotation around the horizontal axis. The current in a turn of the helix, Ampère surmised, can be regarded as the superposition of a circular current around the axis and a linear current along the axis. Assuming that the action of the composed current is equal to the resultant of the actions of the partial currents, then only the helix of Fig. 1.4(a) can be compared to the parallel circular currents of a magnet; the helix of Fig. 1.4(b) involves the additional action of a linear current, which dominates the former action when the radius of the helix is small.[28]

Ampère detected here a more general principle, according to which any two infinitely short currents with the same extremities were equivalent, no matter how contorted they might be. The principle severely constrained the angular dependence of the force between two current elements. Ampère showed this as follows.[29]

The two elements AG and BH represented in Fig. 1.5 can be decomposed into the elements AM and MG on the one hand, and BP, PQ, and QH on the other. According to the principle and in an obvious notation, the force (AG → BH) is equal to the sum of the forces (AM → BP), (AM → PQ), (AM → QH), (MG → BP), (MG → PQ), and (MG → QH). Call m the force acting between two parallel unit elements of current when they are perpendicular to the line joining their center, and n the similar force when they are on this line. Then, the force (AM, BP) is proportional

[28] Ampère 1820b: 174–6 (mémoire du 6 novembre). Cf. Blondel 1982: 87–8; Hofmann 1995: 246–50.

[29] Ampère [1820d], 1820e, 1820g. In anachronistic terms, we would say that the force is a linear function of each current element regarded as a vector.

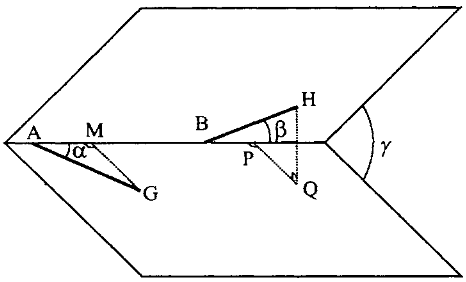

FIG. 1.5. Orthogonal decomposition of two current elements.

to n $\cos\alpha \cos\beta$, and the force (MG, PQ) to m $\sin\alpha \sin\beta \cos\gamma$. All other forces vanish due to symmetry reasons: the geometrical configuration of the corresponding elements is invariant by inversion of one of the currents (neglecting second-order infinitesimals). Consequently, the force acting between two arbitrary current elements has the angular dependence

$$\sin\alpha \sin\beta \cos\gamma + k \cos\alpha \cos\beta, \tag{1.2}$$

where k is equal to n/m. When he announced this result in December 1820, Ampère believed that he could take k equals *zero* 'without inconvenience.'[30]

To consolidate this beautiful reasoning, Ampère conceived experiments that would directly prove the underlying principle. His first idea was to compare the successive actions of a rectilinear and a sinuous current on a magnet. Again, the instability of the galvanic source hampered the project. In the end Ampère had the rectilinear current and the sinuous current act simultaneously on a third mobile current placed at equal distances. Thus was born his famous *méthode de zéro*. As was usual for him, Ampère described the apparatus and the expected results before they were made. On the drawing he provided (Fig. 1.6), SR and PQ are the two currents to be compared, and GH is the test current. Note that Ampère carefully eliminated the effects of the connecting wires. For example, mn and de are placed at equal distance of the test current; the leaders fg and hi to the test current are very close to each other, so that their effects mutually cancel according to a previous experiment. Ampère avoided the magnetic action of the Earth by including the test current in a double loop GFHI, BCDE.[31]

This experiment ended a first phase of Ampère's researches in which Oersted's discovery was the only external stimulus, except for a few remarks by Laplace and by his friends Augustin Fresnel and François Arago. By Christmas 1820,

[30] Ampère 1820e: 229. Cf Blondel 1982: 92–5; Hofmann 1995: 250–2. The symmetry argument for the nullity of the force between perpendicular elements appears for the first time in a note of Ampère 1822a: 209n.

[31] Ampère 1820f (memoire du le 26 décembre), 1822: 162. Cf. Blondel 1982: 96–8; Hofmann 1995: 252–61.

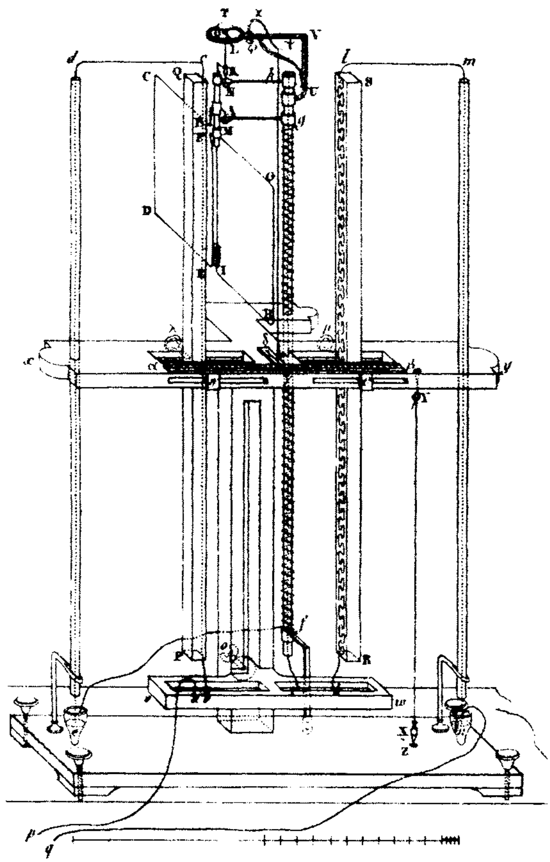

FIG. 1.6. Apparatus for proving the law of sinuous currents (Ampère 1822c).

Ampère had in hand the main elements of his electrodynamics: experimental devices and a mathematical formula for the interaction between two currents, the null method, and the reduction of magnetism to the motion of electricity. Some uncertainty remained on the precise expression of the force between two current elements, and a systematic derivation of the consequences was still missing. However, the main characteritics of Ampère's electric philosophy were already apparent.

1.2.3 Reified theorems

For the most part, Ampère's experiments were planned according to preconceived theoretical ideas. Only the very first experiments had an exploratory value. The more definitive devices were a direct expression of his theoretical beliefs within material constraints such as the compatibility of mobility with current feeding. They served a unique function and could not be transformed to answer new questions. This rigidity was increased by the fact that Ampère, lacking manual skills, always had the apparatus made for him. In general he knew the results of his experiments in advance. Material constraints could, however, lead to instructive surprises, as we saw for the setup with parallel helices.

Ampère's experiments did not yield numbers. In one class of experiments, he showed the qualitative similarity of spirals or helices to magnets. In another class, he examined the more fundamental action between rectilinear currents. There he wished to obtain quantitative results. He failed, however, because the corresponding forces were too small and the voltaic source too unstable. He then switched to the null method, which he believed to provide precision and generality without yielding any number but zero.

To the extent that they reified preconceived ideas, Ampère's experiments played little role in the development of his theory. More instrumental was his critical attitude toward the multiplication of imponderable fluids, which he shared with his friends Fresnel and Arago. He also benefitted from the Newtonian analogy and relevant mathematical techniques, which he learned from the Laplacian circle. At the origin of his intuition that every magnetic action could be reduced to interactions among currents, Ampère saw a virtual history that placed Oersted's discovery before the invention of the compass. His first guesses about the forces between two currents were inspired by his new conception of magnets and by the analogy with gravitational forces.

Yet some of Ampère's experiments contributed to his original intuitions. Most importantly, the failed experiment on parallel helices permitted a basic change of method. From the combination of theoretical conjecture and experimental confirmation, Ampère turned to a more axiomatic method in which the whole theory derived from a few experimentally established principles. The infinitesimal equivalence of rectilinear and sinuous currents was the first of these principles.

Ampère wanted to give his theory a non-speculative outlook. On the nature of the electric current, he followed the French idea of a double flow of negative and positive electic fluids, only adding that the intensity of the flow was the same in all parts of the circuit. He insisted that his deductions did not depend on any particular picture of the electric current, and he kept his speculations on underlying ether processes mostly to himself. He did not regard the new conception of magnets as a speculation: he confused its possibility, demonstrated with helical currents, with a necessity, and he regarded the opposite view as a 'gratuitous supposition.'[32] Lastly, he did not regard the central character of the forces acting between current elements as hypothetical: his rectilinear current apparatus seemed to grant the physical existence of these forces.

[32] Ampère [1820d]: 133.

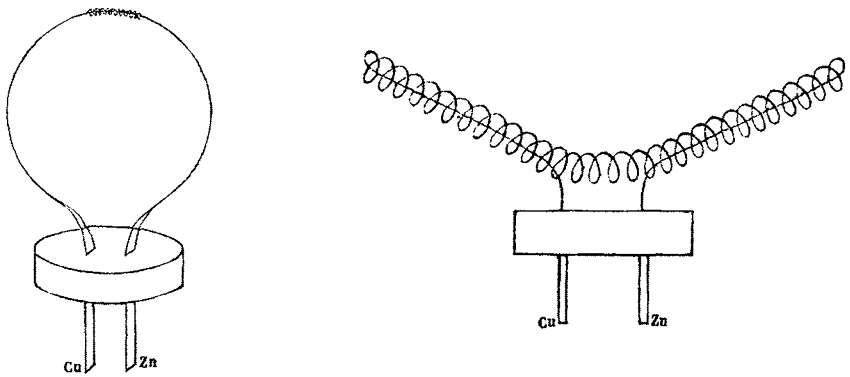

FIG. 1.7. De la Rive's floating devices (G. de la Rive 1821).

1.2.4 Antipathies

In these first months of feverish activity, Ampère's results received more attention than praise. His friends Arago and Fresnel at home, and Gaspard de la Rive in Geneva, seem to have been his only active supporters. Ampère's numerous hurried publications created an impression of confusion, the 'facts' not being clearly distinguished from the theory. The experiments were impressive on paper, but much less so when demonstrated by the inventor. 'Monsieur Ampère is so clumsy,' Laplace said, 'that when his apparatus does not move, he reportedly pushes to shift it.' Oersted was equally unimpressed:

I was at Ampère's by appointment to see his experiments [. . .] He had three considerable galvanic apparatus ready; his instruments for showing the experiments are very complex, but what happened? Hardly any of his experiments succeeded [. . .]. He is dreadfully confused and is equally unskillful as an experimenter and as a debater.[33]

To make it worse, Ampère's devices were much harder to duplicate than Oersted's. His 'ingenious instruments,' de la Rive deplored, 'required skilled workers and fairly high expenses.' As a cheap and easy substitute, de la Rive proposed floating devices made with an acid bath, a cork, a zinc blade, a copper blade, and a piece of wire (Fig. 1.7). With the recipe, he offered the wisdom:[34]

In my opinion we do a favor to Science when we try to diminish the material obstacles that we encounter in our researches and make it possible for a great number of people to study a new experiment: we thus give better chances to new discoveries.

With or without de la Rive's help, physicists quickly accepted Ampère's main 'facts': the attraction between two parallel currents and the analogous behaviors of

[33] Laplace's comment from Colladon 1893: 121; Oersted 1920, Vol. 1: CXIV, both quoted in Blondel 1982: 167n.
[34] G. de la Rive 1821: 201.

a helix and a bar magnet. Yet Ampère's theory met much skepticism or hostility. Foreign physicists were not likely to follow Ampère's mathematical reasoning, and they preferred to see the action between two currents as a consequence of temporary magnetism. The French *savants* were more at ease with Ampère's calculus, but they condemned his departure from Laplacian orthodoxy. Being the heirs of Coulomb's theory of magnetism, they judged Ampère's conception of magnets unclear and questioned the introduction of trigonometric lines in a fundamental force law.[35] They even denied the originality of Ampère's discovery of the interaction between two currents. If, their argument went, a current acted on a magnet and a magnet acted on a current, then a current obviously had to act on an other current. Defending Ampère, Arago objected that two iron keys did not attract each other, although each of them interacted with a magnet.[36]

Ampère's most dangerous critic was the politically and intellectually conservative Jean-Baptiste Biot. More Laplacian than Laplace himself, Biot applied standard French techniques to the determination of the force between a current element and a magnetic pole (that is, the extremity of a long, uniformly magnetized needle). With the help of Felix Savart, he first established by Coulomb's method of the oscillating magnetic needle that the force between a pole and a long rectilinear wire varied as the inverse of their distance. As Laplace told him, this implied a $1/r^2$ dependence for the contribution of a current element to the force. For the angular dependence, he measured the force between a V-shaped current and a pole, varying the aperture of the V. With uncontrolled precision and flawed calculus, he derived the sine we all know. On 18 December 1820, he announced the complete law at the Academy of Sciences.[37]

After one of Ampère's students had pointed to Biot's mathematical mistake,[38] Biot consolidated his proof and argued as follows against Ampère's theory. Ampère's forces between moving electric fluids were 'completely outside the analogies offered by all other laws of attraction.' His interpretation of magnets was a complicated regression to Descartes' vortices. The true course was the Biot–Savart law from which every electromagnetic fact could be deduced without contradicting Coulomb's theory of magnets. Ampère's attractions were nothing but a consequence of the temporary magnetic virtue of the wires carrying the currents.[39]

Biot's criticism could hardly be honest. His own law bore little resemblance to known attraction laws. His explanation of the force between two currents was purely qualitative and depended on a bizarre, if not impossible, distribution of magnetism within the wire, whereas Ampère's description of magnets could be made as precise and quantitative as Coulomb's. Ampère could have used such arguments against Biot. He did not, because he found a more powerful defense in a fact discovered by a British newcomer to the field of electricity.

[35] Cf., e.g., Biot 1824, Vol. 2: 771–2. [36] Cf. Arago 1854, Vol. 2: 58–9.

[37] Biot and Savart 1820, 1821; Biot 1824, Vol. 2: 706–74. Cf. Frankel 1972; Grattan-Guinness 1990, Vol. 2: 923–25.

[38] Savary 1823: 364.

[39] Biot 1824, Vol. 2: 704–74 ('Sur l'aimantation imprimée aux métaux par l'électricité en mouvement'); *Ibid.*: 769–71 (explanation of Ampère's attractions).

1.3 Faraday's rotations

1.3.1 Davy's admirer

When Faraday entered the field of electromagnetism, he was known as the discoverer of a chloride of carbon. His published work was in chemistry, pure and applied. His first interest in science had risen during his apprenticeship in a bookbinder's shop, as he read the books at hand or as he attended popular lectures. The newborn field of galvanism captivated him. With whatever tools and chemicals he could gather, he improvised his own electrochemical experiments.[40]

In his early twenties, Faraday caught Humphry Davy's attention, and became his amanuensis and assistant at the Royal Institution. Founded in 1899, this institute had the official aim of 'diffusing the knowledge, and facilitating the general introduction of useful mechanical inventions and improvements; and teaching, by courses of philosophical lectures and experiments, the application of science to the common purposes of life.' Under Davy's influence, it also became a center for chemical research and popular expositions of science. When Faraday entered the Institution, Davy was a heroic figure both to his peers and to the layman. He was the man who had isolated chlorine and disproved Lavoisier's principle that oxygen was the cause of acidity. Upon Volta's discovery of the pile, he had demonstrated the essential role of chemical reactions in galvanic sources, against Volta's contact theory. He believed in an intimate relation between chemical forces and electricity, and had a critical attitude toward the electric fluids.[41]

Under Davy's lead, Faraday soon became an outstanding chemist in both fundamental and applied matters. He discovered a new compound of carbone and chlorine, a new oil now called benzene, new steels, and new optical glasses.[42] His first excursion beyond pure chemistry occurred in 1820, as a consequence of Oersted's discovery. In the fall of that year he assisted Davy in a series of electromagnetic experiments.

Davy used a much stronger battery than Oersted (100 plates of 4 square inches) and observed that one of the poles of a magnetic needle placed under the wire was 'strongly attracted' by the wire and remained in contact with it. In his opinion, this could be explained only by supposing that the wire itself became magnetic. In order to prove this assumption, he sprayed iron filings on the wire and observed their massive sticking to the wire. He also obtained the magnetization of small pieces of steel. He first interpreted these effects in terms of four magnetic poles in the wire (as Berzelius also did), but soon adopted the idea of 'a sort of revolution of the magnetism around the wire,' which William Wollaston had introduced under the name of 'vertiginous magnetism.' In conformity with this view, small steel needles placed along a circle centered on the wire became magnetized.[43]

[40] Cf. Williams 1965: Ch. 1.

[41] Rumford 1870–1875, Vol. 4: 755. On the Royal Institution, cf. Berman 1978. On Davy, cf. Williams 1965: Ch. 1; Knight 1996.

[42] Cf. Williams 1965: 120–3, 107–8, 109–15, 115–20.

[43] Davy 1821 (read 16 November 1820).

FIG. 1.8. Wollaston's diagram for the attraction of two parallel currents (left) and the repulsion of antiparallel ones (right).

The highly respected Wollaston never published his views in full. In a short published note, he spoke of 'an electromagnetic current' around the axis of the wire, and he provided two drawings explaining the action between two parallel wires (Fig. 1.8). For the same direction of the two currents 'the North and South powers meet' and therefore attract each other. For opposite currents 'similar powers meet' and repulsion results. No one—perhaps not even Wollaston himself—saw clearly what Wollaston had in mind. He seems to have modified Oersted's idea of a helical current to confine it within the conductor. Oersted himself soon interpreted Ampère's forces between two currents in terms of a sympathy, or antipathy, between the corresponding helical motions.[44]

Having assisted Davy in his experiments, Faraday knew the results and Wollaston's speculations. In this period he published an anonymous 'Historical sketch of electro-magnetism,' in which he reviewed the state of experimental and theoretical knowledge in this field. Worth noting are his agnosticism about the electrical current, his fidelity to Davy's conception of electromagnetism in terms of magnetic attractions, and his criticism of Ampère's views. In the first chapter he wrote in unison with Davy:[45]

> There are many arguments in favour of the materiality of electricity, and but few against it; but still it is only a supposition; and it will be well to remember, while focusing on the subject of electro-magnetism, that we have no proof of the materiality of electricity, or of the existence of any current through the wire.

On Oersted's experiment, Faraday repeated Davy's conclusions, insisting on the orientation of the needle across the wire when the effect of the Earth could be neglected, and phrasing everything in terms of *attractions* or *repulsions* of the magnetic needle by one side or the other of the wire. He described Oersted's two spiralling forces with the introduction: 'I have little to say on M. Oersted's theory, for I must confess that I do not quite understand it.' He condemned Berzelius's hasty interpretation of Oersted's experiment in terms of four magnetic poles, and reproached Ampère with lack of clarity. Ampère could not pretend, Faraday went on, to provide an electric explanation of magnetism, for his theory lacked a precise

[44] Wollaston 1821: 363; Oersted 1821: 235–6.

[45] Faraday 1821: 196. Very early, Faraday had read the entry 'Electricity' of the *Encyclopedia Britannica*, in which a certain James Tytler defended a vibrational theory of electricity. Cf. Williams 1965: 14–15.

picture of the electric current. As he confided to de la Rive, Ampère's methods were alien to him:[46]

With regard to [Ampère's] experiments I hope and trust that due weight is allowed to them but these you know are few and theory makes up the great part of what M. Ampère has published and theory in great many points unsupported by experiments when they ought to have been adduced. At the same time M. Ampère's experiments are excellent and his theory ingenious and for myself I had thought very little about it before your letter came simply because being naturally skeptical on philosophical theories I thought there was a great want of experimental evidence.

1.3.2 Rotations and powers

In September 1821 Faraday experimented with a vertical wire and a suspended magnetic needle. Presumably, he did not trust Oersted's results and did not pay much attention to the details of Oersted's account. Also, Davy's repetitions must have been too rough for his taste. According to Davy, the wire attracted one side of the needle and repelled the other. With a large-plate galvanic source, and meticulous variation of the position of the vertical wire, Faraday observed that on a given side of the needle the attraction turned into a repulsion, or vice versa, when the wire passed a certain point located between the center and the extremity of the needle (Fig. 1.9(a)). He concluded that the true poles of the needle were not at its extremities. Most important, the observation excluded his and Davy's view that the motions of the needle resulted from attractions or repulsions between poles and wire. Faraday extrapolated in his mind the motion of a free wire around a fixed magnetic pole (circles of Fig. 1.9(b)), and suspected it to be circular. By trial and error, he soon found a geometrical configuration of a wire and a magnet for which a continuous rotation of the wire occurred (Fig. 1.9(c), (d)).[47]

In Faraday's subsequent elaboration of electromagnetism, this experiment was most basic. All attractions and repulsions observed by Oersted, Davy, and Ampère derived from the simple fact of rotation of a pole around a wire. The action of any wire system on a magnetic pole (concretely, on the extremity of a uniformly magnetized needle) could be traced to the combined circular effects of the different portions of wire. For example, with the drawing of Fig. 1.10 Faraday explained that a system of two parallel wires 'in the same state' (i.e. with the same direction of the current), attracted a pole on one side of their plane, and repelled it on the other. The rotations, real or imagined, connected the various facts of electromagnetism.[48]

The concept of a pole here played a central role. Faraday systematically avoided any reference to the magnetic fluids, and defined the pole as a centre of action. With

[46] Faraday 1822a: 107, 109; Faraday to G. de la Rive, 12 September 1821, *CMF* 1.

[47] *FD* 1: 49–50 (3 Septemper 1821); Faraday 1821, *FER* 2: 127–8. Cf. Gooding 1985.

[48] *FD* 1: 52, #17, #22 (4 September 1821); Faraday 1822, *FER* 2: 133–6, 139–42. Cf. Steinle 1995.

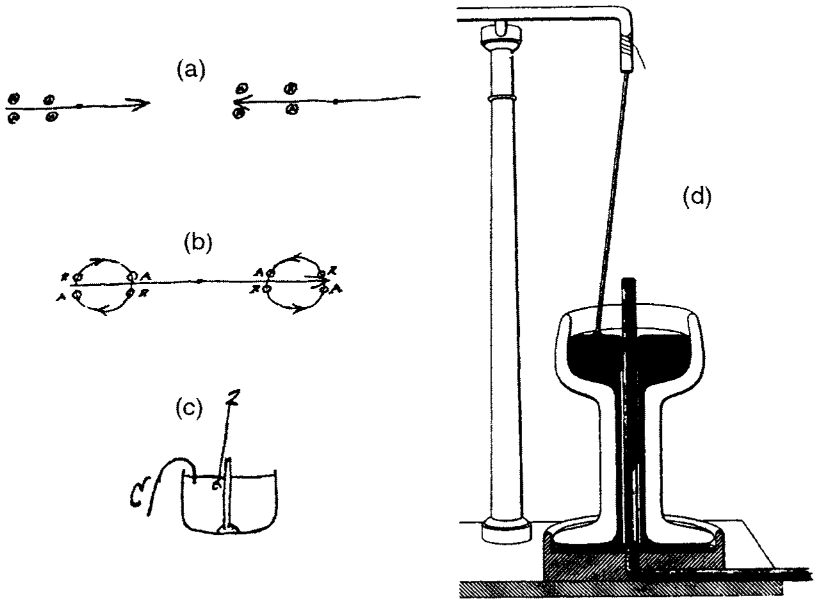

FIG. 1.9. Faraday's steps toward electromagnetic rotations: (a) attractions and repulsions of a wire by a magnetic needle (*FD* 1: 49), (b) imagined rotations (*ibid.*), (c) first rotation device (*FD* 1: 50), (d) classroom version (*FER* 2: plate 4).

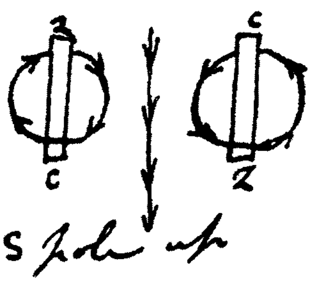

FIG. 1.10. Motion of a magnetic pole between two rectilinear currents (*FD* 1: 51).

this definition, poles were no longer specific to magnets; they could also be produced by electric currents, for example at the extremities of helices. A broader unifying concept was that of 'power,' of which Faraday made abundant use without defining it. Apparently, powers referred to portions of space from which specific actions emanated. Powers could equally belong to a magnet or to the sides of a wire, and they could attract or repel each other:

It has been allowed, I believe, by all who have experimented on these phenomena, that the similar powers repel and the dissimilar powers attract each other; and that, whether they exist in the poles of the magnets or in the opposite sides of conducting wires.[49]

The language of powers was not as universal as Faraday thought it was. But Davy and Wollaston used it, probably to avoid the magnetic fluids, and perhaps out of sympathy for the dynamistic denial of the matter/force dualism. In this language, Faraday explained, the simplest case of magnetic action was that of two centers of concentrated power, that is, the rectilinear attraction or repulsion of two poles. Next came the case of a wire and a pole, which involved three powers: that of the pole, and those of the two sides of the wire.

The pole is at once attracted and repelled by equal powers, and therefore neither recedes nor approaches; but the powers being from opposite sides of the wire, the pole in its double effort to recede from one side and approach the other revolves in a circle.

Then came the case of two parallel wires, which involved four powers, two for each wire. The powers of the facing sides of the wire, Faraday explained, were of the opposite kind when the wires were in the same state and of the same kind for opposite states of the wires, in conformity with the attraction and repulsion observed by Ampère.[50]

The notion of two different powers at the opposite sides of the wire was a bit confusing, as Faraday himself realized:

With regard to the opposite sides of the connecting wire, and the powers emanating from them, I have merely spoken of them as two, to distinguish the one set of effects from the other. The high authority of Dr. Wollaston is attached to the opinion that a single electro-magnetic current passing round the axis of the wire [. . .] is sufficient to explain the phenomena.

However, Faraday needed the notion to explain the actions between currents, and also to express the unity of magnetism and electromagnetism: 'The pole of a magnetic needle presents us with the properties of one side of the wire.'[51]

Faraday thus filled the space around wires and magnets with powers, virtual rotations, and eventually iron filings. In contrast, he left the internal state of wires and magnet undetermined:

I have not intended to adopt any theory of the cause of magnetism, nor to oppose any. It appears very probable that in the regular bar magnet, the steel, or iron is in the same state as the copper wire of the helix magnet; and perhaps, as M. Ampère supports in his theory, by the same means, namely currents of electricity; but still other proofs are wanting of the presence of a power like electricity, than the magnetic effects only.

Note that Amperean currents were not incompatible with Faraday's views. What Faraday rejected was the idea that the attractions and repulsions of two currents were

[49] Faraday 1822, *FER* 2: 128; *Ibid.*: 136. [50] Faraday 1822, *FER* 2: 136–7, 132–3.
[51] Faraday 1822, *FER* 2: 146, 132.

primitive facts. They were a consequence of the distribution of powers demonstrated in electromagnetic rotations.[52]

1.3.3 Opposite styles

Many other differences existed between Faraday's and Ampère's investigations. Whereas Ampère had reached the basic fact of his electrodynamics by speculative reasoning involving virtual history and analogy, Faraday discovered the continuous rotations by patiently exploring the details of the interaction between a magnet and a wire, which others took for granted. In later stages of their research, Ampère and Faraday both used theory, but theory of a very different kind. Ampère exploited the analogy with the theory of gravitation and reasoned in mathematical terms. Faraday knew no mathematics and thought in terms of vaguely defined powers and concretely imaginable actions. For the first, theoretical unity depended on mathematical deduction from a small number of axioms; for the second, theory was always open, and unity derived from the connexity of the various known experimental facts. As Faraday explained to Ampère: 'I am unfortunate in a want of mathematical knowledge and the power of entering with facility any abstract reasoning. I am obliged to feel my way by facts closely placed together.'[53]

Different kinds of theory implied different styles of experiment. Ampère's rigid, professionally designed apparatus was completely at odds with Faraday's improvised, quickly built devices. Whereas Ampère had to give a few days to his mechanician, Faraday managed proper arrangements of wires and needles in a few minutes. He found de la Rive's little floating contrivances 'very simple, easily made, and effectual' and used them abundantly. He wanted to be able to modify and combine the geometrical configurations as easily and quickly as possible, in part to multiply the possibilities of unexpected effects, in part to provide connections between known facts.[54]

With these differences in mind, one easily understands why Ampère could not discover the continuous rotations, although his theory implicitly contained them. Ampère's devices aimed at proving consequences of his theory that he could predict. He did not foresee the continuous rotations, presumably because the analogy with other theories of attractions hid this phenomenon. For a mechanical system moved by gravitational, electrostatic, or magnetostatic forces, it was well known that these forces, being central, could not compensate for the frictional loss of living force (kinetic energy) during a cycle of the system. In other words, the theorem of living forces forbade perpetual motion. Ampère naturally overlooked that his forces, with their angular dependence, contradicted the premisses of the theorem. Having reduced magnets to currents, he could not imagine that electric currents would allow the perpetual motion that no one had ever succeeded in producing with magnets.

[52] Faraday 1822, *FER* 2: 145–146. Faraday used iron filings to show 'the path the pole would follow' (*ibid.*: 140).

[53] Faraday to Ampère, 3 September 1822, *CMF* 1.

[54] Faraday 1821: 288.

1.3.4 How original?

Faraday's discovery of continuous rotations attracted much attention from his British colleagues, though not the kind of attention he expected. The rumor swelled that he had stolen the idea. Wollaston had indeed invoked the possibility of making a wire rotate around its own axis under the influence of a magnet, and had tried—without success—such an experiment with Davy. Faraday defended himself as follows. He was aware of Wollaston's idea and trial, but originally disagreed and interpreted Oersted's phenomenon in terms of attractions and repulsions (as can be verified from the sketch). What led him to the rotations was his closer investigation of the action of a vertical wire on a magnetic needle. And the rotation he produced differed in kind from that anticipated by Wollaston.[55]

There is no reason to doubt Faraday's sincerity. Wollaston himself accepted his explanations. Faraday did not make clear, however, to which extent his discovery was a rediscovery. Were the rotations entirely new? Did previous conceptions of electromagnetism play a role in Faraday's crucial step from the attractions to the rotations?

On the first point, the answer is certainly negative. Oersted's spiralling conflict and Wollaston's electromagnetic current both indicated a circular motion of a magnetic pole around the wire. Moreover, Faraday's observations with the vertical wire added nothing to Oersted's previous observations of the same kind, save the distinction between the poles and the extremities of a magnetic needle. Oersted himself insisted that his observations could not be explained in terms of attractions and repulsions. Faraday innovated in concretizing the rotations, not in imagining them.

If we read Faraday's defense literally, he imagined the rotations only by contemplating the various apparent attractions and repulsions of the needle by the wire. This is possible, but not necessary. Faraday could have unconsciously drawn on Oersted's or Wollaston's idea of a circular action.[56] The essential originality of Faraday lies in the way he could pass from the imagined to the actual rotation. When Wollaston and Oersted understood the circular character of the electromagnetic action, their first impulse was to invent a theoretical cause for it. Wollaston then predicted a case of wire rotation that had little to do with Oersted's original arrangement. In contrast, Faraday avoided speculating on the cause of the circular action. He modified the concrete device on which he conceived the rotation until he could display it in all its splendor.[57]

Imitating de la Rive's advertizing strategy, in September 1821 Faraday mailed to his foreign correspondents a small kit of the rotation device (Fig. 1.11). He provided the cork, the wire, and the glass tube. The happy recipients just had to pour some mercury into the tube and connect it to a galvanic battery. One of them, Jean Hachette, reproduced the wonderful phenomenon in Ampère's company.[58]

[55] Faraday 1823. Cf. Williams 1965: 157–60.

[56] The day he obtained rotations, Faraday noted the absence of rotations of Wollaston's kind: *FD* 1: 50, #7 (3 September 1821).

[57] Cf. Gooding 1985.

[58] Cf. Faraday 1822c: 150–1; Faraday to G. de la Rive, 16 November 1821, *CMF* 1; Hachette to Faraday, 19 Nov 1821, *CMF* 1; Gooding 1985; Blondel 1982: 110.

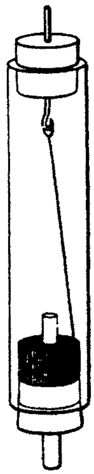

FIG. 1.11. Pocket version of the rotation apparatus (*FER* 2: plate 4).

1.4 *Electro-dynamique*

1.4.1 Awakening

A lung disease and metaphysical broodings had long interrupted Ampère's electromagnetic researches, when he received Faraday's memoir on electromagnetic rotations. Recovery followed magically: 'Metaphysics was filling my head. However, since Faraday's memoir has appeared, all my dreams are about electric currents.' The dreamer soon crafted a number of continuous rotation devices in his own style: with straight wires, mercury cups, and acid baths. The contrivances had less friction than Faraday's, they permitted the substitution of a coil for the magnet, and they were more easily amenable to calculation. On one point Ampère surpassed Faraday: he obtained the rotations of a magnet and a wire around their own axis (Fig. 1.12).[59]

The more theoretical aspects of Faraday's work failed to disturb Ampère: 'This memoir contains very singular electromagnetic facts which perfectly confirm my theory, although the author tries to fight it by opposing one of his invention.' Ampère announced that proper calculations, which he did not provide, explained the rotation in Faraday's original device. More qualitatively, he showed that in his own rotation devices the motion resulted from the forces between the various currents involved. Despite the temporary lack of rigor, Ampère had no doubt: 'These facts comply with the general laws of physics, and one does not have to admit as *a simple primitive fact, a revolutive action* of which nature gives no other example and which we find it difficult to consider as such.' Ampère had other reasons to dislike Faraday's

[59] Ampère to Bredin, 3 December 1821, *CA* 2: 576; Ampère 1821a: 329–33; 1821b, 1822a. Cf. Blondel 1982: 109–16.

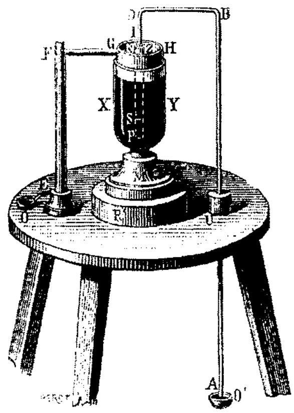

FIG. 1.12. Apparatus for the rotation of a magnet (NS) around its axis (Ampère 1822b). The current enters the magnet through the tip of the vertical wire DI and leaves it through the mercury bath XY. The magnet floats vertically in the bath thanks to the loading SP.

primitive revolutions. They did not provide a sufficient basis for calculation, they involved heterogenous entities (pole and current), and they contradicted the principle of the equality of action and reaction by having a net torque act on the pole–current system. In short, they betrayed every principle of French Newtonian physics.[60]

The lack of understanding was reciprocal. 'I regret that my deficiency in mathematical theory,' Faraday wrote to Ampère, 'makes me dull in comprehending these subjects. I am naturally sceptical in the matter of theories and therefore you must not be angry with me for not admitting the one which you have advanced immediately.' Such statements should not be read as an admission of inferiority. In several occasions Faraday appeared to be proud of his ignorance of mathematics. Upon his later discovery of electromagnetic induction he commented: 'It is quite comfortable to find that experiment needs not quail before mathematics but is quite competent to rival it in discovery.'[61]

Although Ampère misrepresented and rejected Faraday's theoretical ideas, he did not neglect the theoretical consequences of the new fact of continuous rotations. Most strikingly, the rotations offered an apparent exception to the impossibility of perpetual motion. Ampère explained that the continuous supply of living force to the rotating wire came from the electric current. Having thus emphasized the dynamical nature of voltaic electricity, he decided to call 'électro-dynamique' the new science of the interaction of currents. Most important, he used the argument to banish any theory of the temporary magnetism of wires. An arrangement of magnets, no

[60] Ampère to Bredin, 3 December 1821, *CA* 2: 576 (quotation); Ampère 1821b: 370, 374 (quotation); Ampère to A. de la Rive, 14 October 1822, *CA* 2: 605 (against primitive revolutions).

[61] Faraday to Ampère, 2 February 1821, *CMF* 1; Faraday to Phillips, 29 November 1831, *CMF* 1.

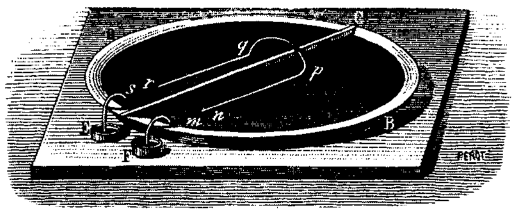

FIG. 1.13. Device for showing the mutual repulsion of the parts of a rectilinear current (Ampère 1822c). The currents in the segments qr and np of the floating wire repel the currents rs and mn in the mercury bath.

matter how complex, could not yield the continuous rotations, since magnetic actions were known to obey the theorem of living forces.[62]

The rotation experiments also played an important role in the determination of the force formula for two current elements. While experimenting on his own rotation devices, Ampère noted that the phenomenon disappeared whenever both wire ends were on the axis of the magnet. Yet calculations with the simple formula (1.1) indicated a positive result in this case. Ampère then returned to the more general formula (1.2), and sought the value of k for which the rotation did not occur. In June 1822 he found $k = -1/2$, which gives

$$d^2 f = ii' \frac{ds ds'}{r^2}\left(\sin\alpha \sin\beta \cos\gamma - \frac{1}{2}\cos\alpha\cos\beta\right) \tag{1.3}$$

for the force $d^2 f$ acting between the elements ds and ds' of the currents i and i' (an attraction being reckoned positively). This formula implies that, contrary to Ampère's early guess, two current elements on the same straight line and with the same orientation repel each other. Ampère soon confirmed this effect in Geneva, with a device which is now familiar to every student of electrodynamics (Fig. 1.13).[63]

The analytical calculations performed in this context are of special interest. Ampère replaced the magnet with a simple circular current, and required that the total torque impressed by an element of this current on any current starting and ending on the axis of the circle should be zero.[64] For Ampère the mathematician, this meant that the torque impressed on any element of current had to be an exact differential with respect to the distance of this element from the axis. With this prop-

[62] Ampère 1822a: 66; Ampère 1822b, 1826b: 97 for 'électro-dynamique'; Ampère 1822a: 65–6; 1826b: 96n.

[63] Ampère 1822c: 235, and 1822d: 418; Ampère [1822e]: 331, and 1826b: 28. Cf. Blondel 1982: 127–8, 132–3; Hofmann 1995: 293–308.

[64] The latter condition does not rigorously result from the similar condition with the whole circular current. Presumably for this reason, Ampère later preferred another equilibrium case. Cf. Blondel 1982: 127–8.

erty in mind, Ampère transformed his trigonometric force formula into another that involved derivatives of the mutual distance of the two elements with respect to the curvilinear abcissae s and s' of the two linear currents to which they belonged (see Appendix 1):

$$\mathrm{d}^2 f = -ii' \frac{\mathrm{d}s\mathrm{d}s'}{r^2}\left(r\frac{\partial^2 r}{\partial s \partial s'} - \frac{1}{2}\frac{\partial r}{\partial s}\frac{\partial r}{\partial s'} \right). \tag{1.4}$$

Although it was born from the consideration of a special class of failed experiments, this formula had a prosperous future.[65]

The new technique proved highly adequate in the important case of closed circuits. Ampère thus proved that the force impressed on a current element by a closed circuit was perpendicular to the element. A few months later, a former student of his harvested other essential results. With a Polytechnician's skills, Félix Savary integrated the force formula over a circular loop of current, and then over a dense pile of such currents—which Ampère later called solenoid, after the Greek $\sigma\omega\lambda\varepsilon\nu$ for canal. When the radius of the circles was much smaller than the length of the canal, the solenoid behaved like two magnetic poles located at its extremities. The force between an extremity and another current obeyed the Biot–Savart law, and the force between two extremities satisfied Coulomb's law (see Appendix 1). Ampère congratulated Savary for having reduced the three basic actions of magnetism under the same law of his, thus proving the validity of his conception of magnets.[66]

1.4.2 The Newton of electricity

By that time, early 1823, Ampère's electrodynamics had reached maturity. With the perfected Ampère law, the Amperean currents, and proper analytical tools, one could calculate every known magnetic or electromagnetic effect. However, a systematic account of the theory was still wanting. This Ampère gave in 1826 with his masterful 'Mémoire sur la théorie mathématique des phénomènes électro-dynamiques, uniquement déduite de l'expérience.'[67]

Imitating the rhetorics of Newton's *Principia* or Fourier's *Théorie Analytique de la Chaleur*, Ampère presented his results as the plain expression of experimental truths: 'I have solely consulted experiment to establish the laws of these phenomena, and I have deduced the only formula that can represent the forces to which they are due.' Later commentators have had no difficulty detecting a few unwarranted hypotheses in Ampère's theory, for example the central character of elementary forces, the absence of elementary torque, and the Amperean currents. There is no reason, however, to doubt Ampère's sincerity. As was mentioned, the concept of *physical* current elements, on which the character of the action between current elements depended, seemed to be materialized in his apparatus. The currents in magnets

[65] Cf. Grattan-Guinness 1990, Vol. 2: 930–33. I use Kirchhoff's notation for the partial differentials.

[66] Ampère 1822d: 419–20; Savary 1823. Cf. Grattan-Guinness 1990, Vol. 2: 934–9. Savary starts with a closed solenoid, motivated by an unpublished experiment of Gay Lussac and Welter.

[67] Ampère 1826b.

were not a hypothesis, as far as they were the only consistent way to unify magnetism and electromagnetism: 'The proofs on which I base [my theory] mostly result from the fact that they reduce to a single principle three sorts of actions which all phenomena prove to depend on a common cause, and which cannot be reduced in a different manner.'[68]

Most important, Ampère's formula for the force between two current elements did not depend on any assumption regarding the nature of the electric current and connected mechanisms: 'Whatever be the physical cause to which we may wish to relate the phenomena produced by this action, the formula obtained will always remain the expression of facts.' As we shall see, this turned out to be largely true, since Ampère's formula (at least its consequences for closed currents) remained an essential basis for the construction of all later theories of electrodynamics. Ampère again compared himself to Fourier, whose equations for heat propagation had survived Fresnel's wave theory of light and heat. Extending the parallel, Ampère did not exclude the search for physical causes. He himself speculated on various mechanisms for the production of electrodynamic forces, as will be seen in a moment. But he required a clean separation between laws and causes.[69]

For the determination of the force between two current elements, Ampère offered a polished version of the null method, which was 'more direct, simpler, and susceptible of great precision.' The first equilibrium case concerned the lack of action of two contiguous opposite currents. The second established the equivalence of rectilinear and sinuous currents, in the manner of 1821. The third replaced the no-rotation devices of 1822 and proved that the force acting from a closed circuit on a current element was perpendicular to the element. The fourth established the scale invariance of the electrodynamic action.[70]

Ampère assumed, as self-evident, that the action between two current elements resulted in equal and opposed forces directed along the line joining the elements and decreasing as the nth power of their distance. Then he used the first case of equilibrium to prove that the force between two orthogonal elements vanished. The second case, as before, determined the angular dependence of the force, up to the constant k. The third and fourth cases gave two relations between k and n, from which $k = -1/2$ and $n = 2$ resulted. The complete expression of the force still involved obvious factors: the lengths of the elements and the intensities of the currents. In Ampère's mind the latter factor constituted a quantitative *definition* of the intensity of a current, including a definite current unit as soon as the unit of force was defined.[71]

The experiments and reasonings of the null method had an air of great systematism. A closer look at them, however, reveals serious flaws. Ampère did not quantify the precision of his apparatus, as if measuring a zero quantity required zero efforts at error analysis. Even worse, his third case of equilibrium was utterly

[68] Ampère 1826b: 2, 83–4. [69] Ampère 1826b: 4. [70] Ampère 1826b: 6, 9–18.
[71] Ampère 1826b: 18–44, and 18–19 for the definition of intensity.

unstable and hardly observable, and the apparatus for the fourth one was never built, on Ampère's own admission.[72] Could it be that Ampère's law rested on paper evidence? Certainly not: Ampère knew that the equivalence between magnets and systems of currents completely determined the values of n and k.[73] For the sake of a reductionist rhetoric, however, he preferred an ideal justification of his formula that would not depend on the complicated physics of magnets.

In the bulk of his memoir, Ampère developed the consequences of his formula for closed currents, Savary's solenoids, and magnets. The diversity of his mathematical techniques must be emphasized. In some reasonings he used the original expression of the force in terms of trigonometric lines, but in most he started with the 'very simple'

$$d^2 f = -ii'dsds'\frac{2}{\sqrt{r}}\frac{\partial^2\sqrt{r}}{\partial s\partial s'}. \tag{1.5}$$

Occasionally, he turned to Cartesian coordinates. For example, he wrote the force (X, Y, Z) acting from a closed circuit on a current element (dx, dy, dz) in the form

$$X = \frac{1}{2}ii'(Cdy - Bdz), \quad etc. \tag{1.6}$$

with

$$A = \int\frac{(y'-y)dz' - (z'-z)dy'}{r^3}, \quad etc. \tag{1.7}$$

where (x', y', z') are the coordinates of the points of the closed circuit (see Appendix 1). This expression exhibits the perpendicularity of the force and the current element. Moreover, it shows that the force is perpendicular to the direction of the vector (A, B, C), which Ampère called the 'directrice' since it depended only on external circumstances. Modern readers should resist the temptation to identify the *directrice* with a magnetic field concept: Ampère considered only the direction, and he did not include the intensity i' in the vector (A, B, C).[74]

As a special case of a closed current, Ampère considered a single infinitesimal loop of current, and showed that it was equivalent to a magnetic dipole. A finite closed current, he went on, could be replaced by a net of infinitesimal current loops

[72] Ampère 1826b, 1st edn: 205 (third case); 2nd edn: 151 (fourth case). In the edition for the Mémoires de l'Académie Royale des Sciences (1827), Ampère omitted the criticism of the third case of equilibrium. Cf. Blondel 1982: 147–8.

[73] Cf., e.g., the remarks in Ampère 1826b: 17, 151, indicating that the properties of the action between current and magnets can be used instead of the fourth case of equilibrium.

[74] Ampère 1826b: 30–1. Cf. Grattan-Guinness 1991; Hofmann 1995: 341–3. Up to a normalization factor, the formulas correspond to the modern $\mathbf{f} = i d\mathbf{l} \times \mathbf{B}$, with $\mathbf{B} = \int i' d\mathbf{l}' \times \mathbf{r}/r^3$.

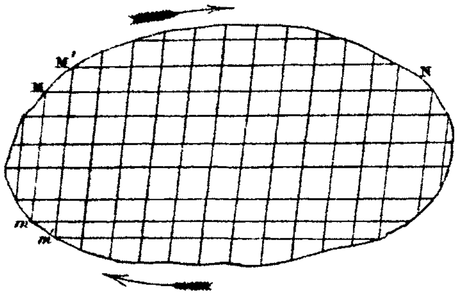

FIG. 1.14. Net of current (Ampère 1826b: plate 1). The same current runs clockwise around each little square, so that the net current in the common side of two square vanishes.

(Fig. 1.14), and was therefore mathematically equivalent to a double sheet of boreal and austral fluid. The ingenious equivalence played little role in Ampère's deductions, save for a proof that the continuous rotations were impossible for closed rigid circuits. Yet it could be very helpful to any one who, unlike Ampère, wished to derive the law of electrodynamics from those of magnetism.[75]

Toward the end of his memoir, Ampère relaxed his severe attitude and indulged in speculations on the cause and nature of electric motions. In his previous researches he had repeatedly tried to understand electrodynamic forces in terms of a propagated action in a medium. In his youth he condemned 'the supposition of an action between bodies that do not touch each other.' In the early 1820s, the success of Fresnel's optical ether revived his desire to reduce all physics to the local motions of a medium. When he discovered the equivalence of rectilinear and sinuous current, he imagined a corresponding superposition of ether motions. Later, the equivalence between a closed circuit and a net of infinitesimal current loops suggested to him a rotary motion in the medium. In each case, the fact preceded the intuition, and Ampère remained very discreet about his ether.[76]

Ampère was more open about his conception of the electric current. In 1821, he gave up Volta's idea of an electric motion of which the substratum of the conductor was the only obstacle. He adopted instead Oersted's idea of a series of compositions and decompositions of the two electricities starting in the battery and propagating along the conductor. In lengthy speculations, he combined this view with the atomistic conception of matter to explain contact tension and electrolysis. More succinctly, he imagined an ether made of the neutral fluid resulting from the combination of negative and positive electricity.[77]

In the memoir of 1826 Ampère expounded his view of the electric current, and mentioned the related conception of the ether. He briefly suggested a propagation of

[75] Ampère 1826b: 41, 101, 145–6; Ampère 1826a. Cf. Blondel 1982: 150–3; Grattan-Guinness 1990, Vol. 2: 956–9. In Chapter 2 it will be shown how Franz Neumann exploited the equivalence.

[76] Ampère [1801]: 175; 1820a: 257; 1826a: 47. Cf. Blondel 1982, 88–9, 152–3; also Caneva 1980.

[77] Ampère 1821c, *MRP* 2: 216 (Oersted's idea); 1822a, *MRP* 2: 249 (ether); [1824a], 1824b (electrochemistry). Cf. Blondel 1982: 155–7 (current), 161–5 (ether).

electromagnetic actions through this ether, but favored a more conservative approach in which Coulomb's electrostatic law remained basic. The idea was to take the average of the Coulomb forces between the separated fluids in the interacting currents. Since the separation was a temporary, spatially directed process, the angular dependence of the net forces could perhaps emerge in this manner.[78]

In sum, Ampère's influential memoir of 1826 was not just the reunion of the equilibrium cases, the Ampère formula, and the Amperean currents in magnets. It also involved a store of mathematical techniques from which successors could borrow, and it prefigured two ways of deepening our understanding of electrodynamic forces: by reducing them to motions in the ether or by summing the direct actions of the electric fluids running in conductors.[79]

The magnificent architecture of the memoir rested on a fictitious three-stage history. In the first stage, fundamental experiments established general properties of electrodynamic forces. In the second, a general force formula was inferred from these properties. In the third, all known phenomena of electrodynamics and magnetism were deduced from the force law and the assumption of Amperean currents. This architecture helped clarify the subject and convince Ampère's readers. At the same time, it obscured the dynamical interplay of experiment, mathematical techniques, and theoretical ideas in the actual genesis of electrodynamics.

Oersted's new effect, Newtonian analogy, and the principle of unity were the sources of Ampère's initial theoretical convictions. Then Ampère conceived, ordered, and used apparatus intended to support these convictions. The infinitesimal analysis of the theory conditioned the structure of the apparatus. Reciprocally, this structure suggested the notion of a physical current element as a separable entity with regard to the principles of mechanics. In general, the experiments confirmed the original intuitions. However, the few failed experiments played a crucial role. They removed previous indeterminations of the theory, they redirected Ampère toward the null method, and they prompted the development of new mathematical techniques. In turn, these techniques permitted a confirmation of the more qualitative components of Ampère's theory, and suggested more fundamental explanations of electrodynamic forces.

This complex history and Ampère's simple reconstruction of electrodynamics share a common trait: the mathematics is rigorous and adaptable, while the experiments lack precision and flexibility. This asymmetry, later regarded as a basic defect of the otherwise impressive French physics, has a natural explanation: the experiments were intended to found the theory at the simplest level of analysis, for which effects are small and geometrical configurations highly constrained. There were two obvious ways of avoiding the difficulty: to deny the control of mathematical theory over experiment, as Faraday did, or to relocate the control at the level of more complex, but still computable systems, as Weber later did.

[78] Ampère 1826b: 87, 97–9.

[79] Faraday's field conception is akin to the first approach, Weber's theory to the second.

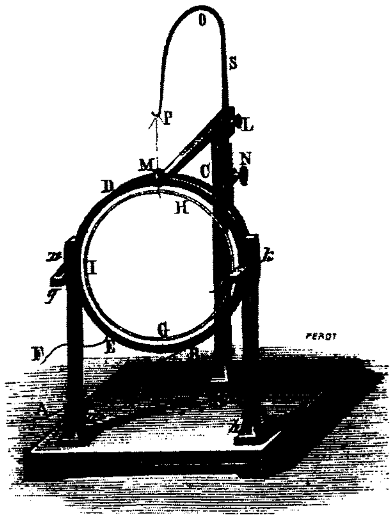

FIG. 1.15. Device for Ampère's induction experiment (Ampère 1821c: 448).

1.5 Electromagnetic induction

1.5.1 First tries

No mention has yet been made of a couple of Ampère's experiments that had nothing to do with the mathematics of current elements. As the hypothesis of molecular currents in magnets played a central role in his theory, Ampère wanted to determine whether these currents preexisted in unmagnetized iron or were created during the magnetization process. For this purpose, in July 1821 he imagined the device of Fig. 1.15, in which the copper ring HIG hangs within the fixed coil BCDE. He placed a magnet on the sticks nq and pk, and fed the coil with a battery. The ring did not move, and therefore did not seem to be the seat of induced currents. Ampère inferred that randomly oriented molecular currents existed in unmagnetized iron, and explained magnetization by an orientation of these currents.[80]

In September 1822 he repeated the experiment in Geneva with a powerful magnet and this time obtained 'alternatively an attraction and a repulsion of the ring.' With this positive result, the experiment could no longer serve to support the existence of molecular currents in iron. Nor was it related to the fundamental law of electrodynamics. Hence Ampère had no theoretical motivation to pursue the subject. He did not even specify whether the effect was permanent or transient, left its direction undetermined, and abandoned the publication to Auguste de la Rive.[81]

[80] Ampère 1821b: 377; 1821c: 448. Cf. Blondel 1982: 118–19; Hofmann 1995: 310–15.

[81] Ampère [1822e]: 333–4; A. de la Rive 1822: 48; G. de la Rive to Faraday, 24 September 1822, *CMF* 1: 291. Cf. Ross 1965; Williams 1986; Hofmann 1987b; Romo and Doncel 1994: 299. For a modern repetition and interpretation of Ampère's experiment, cf. Mendoza 1985.

For Faraday, the induction of currents was not an indifferent matter. If, he might have reasoned, the current-carrying state of a conductor implied magnetic power, the reciprocal effect was likely to exist: magnetic power had to induce electric currents in conductors. In November 1825 he investigated the case of two linear conductors. The first conductor being connected to a battery, its magnetic power could perhaps induce a current in the second conductor. Faraday tried three geometrical configurations: a pair of parallel wires, a straight inducing wire within a helicoidal collecting wire, and a helicoidal inducing wire around a straight collecting wire. In each case his galvanometer gave no deviation.[82]

Perhaps the effect was just too small to be detected, Faraday must have thought. In August 1831, probably impressed by Joseph Henry's and Gerritt Moll's experiments with electromagnets, Faraday imagined a new device that exploited the multiplying effect of coils and the concentrating effect of iron. He had long copper wires coiled around two opposite sides of an iron ring, with proper insulation of the turns. The iron conducted the strong magnetic power of the primary coil to the secondary coil. Three feet of wires connected the latter coil to a primitive galvanometer made of a suspended magnetic needle and a parallel wire. Worth noting is Faraday's distrust of ready-made meters in his search for new effects.[83]

Faraday connected the primary coil to the battery and reported: 'Immediately a sensible effect on needle. It oscillated and settled at last in original position. On *breaking* connection [. . .] with battery again a disturbance of the needle' (Faraday's emphasis). The effect was therefore clear but transient. Faraday spoke of a 'wave of electricity,' meaning a current short and intense as a breaker on the shore.[84]

Faraday did not expect a transient phenomenon. Available analogies, especially that with electrostatic induction, suggested a permanent induced current. Fortunately, he did not need to look for a transient effect in order to see the induced current, for a simple reason: his galvanometric needle had little damping and could perform 'a few oscillations' before it returned to equilibrium. He could therefore observe the perturbation of the needle even if he looked at it well after he had closed the primary circuit. Over previous attempts at detecting induced currents, Faraday's crucial improvement was the amplification of the effect by the coils and the iron core. With the primitive device of November 1825, his galvanometer was too insensitive to show the least disturbance, even transient.[85]

[82] *FD* 1: 279 (28 November 1825). In April 1828 (*FD* 1: 310) Faraday explored a case of induction similar to Ampère's: he placed a magnet within a delicately suspended copper ring, and tried and failed to move the ring by approaching the poles of another, powerful magnet. For his motivations to expect induced currents, cf. *FER* 1: 1–2.

[83] *FD* 1: 367, ##1–5 (29 August 1831). On Henry's and Moll's experiments, cf. Moll to Faraday, 7, 9, 10 June 1831, *CMF* 1.

[84] *FD* 1: 367, #3; *FD* 1: 369, #14 ('wave of electricity'). A careful examination of occurrences of the expression 'wave of electricity' in the diary shows that Faraday did not mean a waving motion (In French one could say he meant *vague*, not *onde*).

[85] *FD* 1: #7. Several commentators have attributed Faraday's success to his supposed anticipation that the effect should be transient, and speculated on various reasons for such an anticipation.

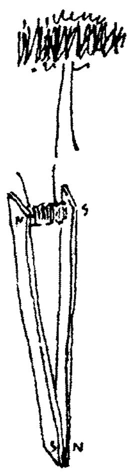

FIG. 1.16. Device for magneto-electric induction (*FD* 1: 372).

1.5.2 From weed to fish

Another key to Faraday's success was his willingness to investigate what looked at first glance like a parasitic phenomenon and thus to transform a crude observation into a full-blown discovery. In the day of the first observation he improved the detection by replacing the straight wire above the compass needle with a flat spiral, and showed that an iron cylinder could be used instead of the iron ring (though less efficiently). He failed, however, to produce two expected effects of the induced current: spark and electrolysis. Three weeks later, on the verge of starting a new series of experiments, he wrote to his friend Richard Phillips: 'I am busy again on electromagnetism and think I have got hold of a good thing but can't say; it may be a weed instead of a fish that after all my labor I may at last pull up.'[86]

The following day Faraday tried induction from coil to coil without iron core, and also induction from moving magnet to spiral. This failed. He then returned to his earlier iron cylinder, made the surrounding wire into a single helix connected to the galvanometric device, and arranged two bar magnets and the iron cylinder in a triangular magnetic circuit (Fig. 1.16). Whenever the magnetic contact was made or broken, the magnetic needle moved. Faraday concluded: 'Distinct conversion of Magnetism into Electricity.'[87]

In October, with a refreshed battery and improved coils, Faraday obtained direct induction from coil to coil, though very weakly. He also managed to produce a spark with the original iron ring and coils. Lastly, he obtained a galvanometric deflection by thusting a bar magnet into a hollow coil and recorded: 'A wave of Electricity was

[86] *FD* 1: #6, #7, #18, #15, #11; Faraday to Phillips, 23 September 1831, *CMF* 1.
[87] *FD* 1: #21, ##25–7, #33 (24 September 1831).

so produced from *mere approximation of a magnet* and not from its formation *in situ*.'[88]

By that time Faraday could realize all cases of electromagnetic induction, except for that given by the relative motion of two circuits, which he obtained at the end of the year.[89] His explorative strategy involved amplifications of the effects with improved coils, batteries, and detectors; and simple mutations of the devices, suggested by the equivalence of currents and magnets with respect to magnetic power. According to this equivalence, the iron in the first induction experiment could not be essential, since it merely channelled and amplified the magnetic power of the primary coil. Conversely, the primary coil could be replaced with a magnet, as Faraday did in the triangular device of Fig. 1.16. Lastly, as the formation and the approximation of a magnet brougth the same change of magnetic power, Faraday conceived induction by a moving magnet.

1.5.3 The electro-tonic state

This series of observations did not depend on a particular view of the induction process. Yet Faraday had one very early on. In his mind, the transient character of the induced current was too surprising to be left unexplained. A transient induced current was conceivable when *closing* the primary circuit: intuitively, the sudden increase of the primary current could be more efficient than a steady current. But the occurence of a transient current of comparable intensity when *breaking* the primary circuit puzzled Faraday. How could a dying current have inductive effects when a steady current had none?

In his first report of the latter effect Faraday underlined the word 'breaking.' In the same day he wrote: 'Recurrence on breaking the connection shews an equilibrium somewhere that must be able of being rendered more distinct.' In this view, which Faraday later exposed at the Royal Society, the conductor assumes a 'tonic' state during the initial transient current and maintains it as long as the primary current exists. When the primary circuit is broken, this state relaxes and an inverse transient current results. In Faraday's own words:

> Whilst the wire is subject to either volta-electric or magneto-electric induction, it appears to be in a peculiar state; for it resists the formation of an electrical current in it, whereas, if in its common condition, such a current would be produced; and when left uninfluenced it has the power of originating a current, a power which the wire does not possess under common circumstances. This electrical condition of matter has not hitherto been recognised, but it probably exerts a very important influence in many if not most of the phenomena produced by currents of electricity. [. . .] I have, after advising with several learned friends, ventured to designate it as the *electro-tonic* state.

In brief, the new state had three essential virtues: it explained the current induced during the suppression of the inducing device, it extended the idea of states induced

[88] *FD* 1: ##36–9, #46, #57. [89] *FD* 1: ##250–1 (26 December 1831).

by states, and it offered a new possibility for developing a picture of the electic current and its effects.[90]

If, Faraday reasoned, the new state of matter truly existed, independent evidence needed to be brought. For example, Faraday sought magnetic actions from conductors in this state, or variations of their conducting power. All attempts were negative. Besides, Faraday soon developed another description of the induction phenomenon, as will be seen in a moment. To his published account of the electrotonic state, he added a footnote mentioning that the notion had become superfluous. He was reluctant to give press to a speculation, but retained his faith in the tonic state for the rest of his life. In 1835 he wrote to Whewell:

> I have given up this electrotonic state for the times as an experimental result (remember, my researches are *experimental*) because I could find no fact to prove it but I cling to it in fancy or hypothesis from general impressions produced by the whole series of researches.

Among Faraday's later tentative proofs of the tonic state we find self-induction in 1834, a quickly discarded difference between the inducing powers of voltaic and magneto-electric currents in 1840, and diamagnetism in 1845. This obstinacy reveals Faraday's commitment to the gradation of cause and effect: he could not admit that an effect would be larger than its direct cause. The interrupted current in a given circuit could not be the cause of a larger current in another circuit without 'a link in the chain of effects, a wheel in the physical mechanism of the action, as yet unrecognized.'[91]

1.5.4 Cut magnetic curves

Nonetheless, the electro-tonic state played little role in Faraday's early experiments on electromagnetic induction. After proving voltaic and magnetic induction, he rather explored the link he suspected between Arago's effect and the new phenomenon. In 1822, while measuring the magnetic force of the Earth near Greenwich, Arago had noticed the damping effect of non-magnetic metals placed in the vicinity of the compass needle. Two years later, he examined and published the reverse effect: the slowing down of a rotating copper disk by a nearby magnet. This new kind of magnetic action attracted much attention, and even triggered a priority quarrel between Arago and David Brewster. Notwithstanding Arago's initial reserve, several assumptions were made about the cause of the new effect, the most popular being a temporary magnetization of the rotating disk.[92]

In the very first day of his induction experiments, Faraday queried: 'May not these transient effects be connected with causes of difference between power of metals in rest and in motion in Arago's expts.?' Faraday had in mind that the force between

[90] *FD* 1: #3, #8; *FER* 1: series 1 (November 1831): #60.

[91] *FER* 1: 16n (footnote); Faraday to Whewell, 19 September 1835, *CMF* 1; *FER* 1, series 9 (December 1834): #1114 (self-induction and quote). On the difference between voltaic and magneto-electric currents, cf. *FD* 4: ##6081–6187 (August 1840). On tonic state and magnetic polarization, cf. *FER* 1: #1729.

[92] Arago 1825. Cf. Arago 1826; Williams 1965: 170–172; Romo and Doncel 1994: 302–303.

the magnet and the rotating plate could be due to currents induced in the plate. Two months later, he proceeded to check this assumption with a copper disk rotating between the jaws of the 'great magnet of the Royal Society.' He placed two collecting blades at two points of the disk, connected them to a galvanometer, and observed a distinct deviation. A new electric machine was born.[93]

In the absence of any precise theory, Faraday assumed that the configuration of the induced currents would imitate the configuration of the inducing current. Accordingly, in his early experiments on Volta-electric induction he judged that a growing current induced currents in the same direction.[94] It took him no less than three months to become aware of this sign mistake. In the rotating-disk experiment, he expected a semi-vortex of currents in the part of the disk situated between the poles of the magnet, in conformity with the configuration of the inducing Amperean currents. Fortunately, this prejudice did not prevent further exploration. Varying the position of the sliding contacts, Faraday soon found that the assumed vortex did not exist. The currents were induced radially, that is, in a direction perpendicular to the motion.[95]

Having in mind a more direct proof of this law, Faraday passed rectangular blades and wires between the jaws of an electromagnet, and by the pole of a cylinder magnet. He concluded: 'The current of electricity which is excited in a metal when moving in the neighbourhood of a magnet depends for its direction altogether upon the relation of the metal to the resultant of magnetic action, or to the magnetic curves.' Not knowing the vector product, Faraday found this direction 'rather difficult to express,' and took three paragraphs to explain it with diagrams, a knife-blade, and a wood-and-threads model. The basic idea was to consider the way the wire cuts the 'magnetic curves,' defined as 'the lines of magnetic forces [. . .] which would be depicted by iron filings or those to which a very small magnetic needle would form a tangent.'[96]

In later experiments regarding the induction under the magnetic action of the Earth and the induction by a rotating cylinder magnet, Faraday proved that the mere cutting of magnetic curves, without change of magnetic intensity, was sufficient to induce a current. He also compared (by opposition) induction in different metals, and found that 'the tendency to generate a current' was the same for all metals. In early 1832 he condensed his results in a single law: 'If a terminated wire moves so as to cut a magnetic curve, a power is called into action which tends to urge an electric current through it.'[97]

[93] *FD* 1: #17; *FD* 1: #85, ##99–109 (28 October 1831).

[94] *FD* 1: #9. Faraday gave the correct direction of the induced currents on 8 December, *FD* 1: 190–208. The manuscript of series 1 read before the Royal Society on 15 December contained the sign mistake. Cf. Romo and Doncel 1994; Doncel 1996.

[95] *FD* 1: #77 (24 October 1831) for the semi-vortex; *FD* 1: ##110–19 (28 October 1831). Cf. Steinle 1994; Romo and Doncel 1994; Doncel 1996.

[96] *FD* 1: ##130–42 (4 November 1831), ##194–213 (8–9 December); *FER* 1, series 1 (November 1831): 33, ##114–16, #114n (magnetic curves).

[97] *FD* 1: ##232–9 (21 December) for terrestrial induction, ##255–7 (26 December) for the rotating magnet; *FD* 1: ##283–87 (26 December) and *FER* 1: series 2 (January 1832): 62 for different metals. Originally, Faraday regarded the equality of the induction in different metals as contradicting the dependency of the Arago effect on the metal. It took him several weeks and Christie's help to understand the

Originally, this statement concerned only the induction produced by the relative motion of a magnet and a conductor. Faraday later included the case of two conductors in relative motion, and finally that of induction by a varying current. In the latter case magnetic curves must be seen as developing during the growth of the primary current and thereby cutting the conductor of the secondary circuit. With this ultimate extension, the law of the cut lines of force became complete and self-sufficient: 'By rendering a perfect reason for the effect produced [the law seems to] take away any for supposing that peculiar condition, which I ventured to call the electro-tonic state.'[98]

Magnetic curves had been widely used in Faraday's circle, including Davy, Sturgeon, and Moll, to represent the magnetic power of magnets, electric currents, and electromagnets. With this geometrical representation of magnetic power, Faraday could bridge his two essential discoveries: electromagnetic rotation and induction. He perceived a basic axis–loop duality that applied to both phenomena: 'The power of inducing electric currents is circumferentially exerted by a magnetic resultant or axis of power, just as circumferential magnetism is dependent upon and is exhibited by an electric current.'[99]

1.5.5 The ambiguities of success

Faraday's extraordinary discovery prompted high excitement among his peers. Without the author's permission, Jean Hachette read to the French Academy a private letter in which Faraday summarized his main findings. Through a French magazine the news reached two distinguished Italian physicists, Leodolfo Nobili and V. Antinori, who immediately experimented on the subject and published their findings. The false rumor of their priority soon circulated, even though they had included the text of Faraday's letter in their paper. Furthermore, an article in *Le lycée* dwelt on French anticipations of Faraday's discovery. Faraday had no difficulty straightening the facts.[100]

However, he hurt Ampère's feelings by attributing to him 'the erroneous result' that induced currents were in the same direction as the inducing currents. In a long, tormented letter to Faraday, Ampère proved that he had never made any pronouncement on the direction of the electric current. Implicitly, he regarded his experiment with the suspended copper ring as an anticipation of Faraday's discovery. Yet this observation had little historical importance, whether or not it had something to do with electromagnetic induction. Where Ampère had done a single,

role of the conducting power of the metal. In terms of Ohm's law, which Faraday did not know, the induced electromotive force does not depend on the metal, but the current does. Cf. Steinle 1996.

[98] *FER* 1, series 2 (January 1832): #232, #238, #231 (quotation). Cf. Steinle 1996.

[99] *FER* 1, series 1 (November 1831): 118. On previous uses of magnetic curves, cf. Simpson 1968: 80–86; Heilbron 1981: 202; Gooding 1985, 1990: Ch. 4.

[100] Faraday to Hachette (lost), extract pub. in *ACP* 48 (1831): 402, and in *Le temps*, 28 December 1831 (read by Nobili); Nobili and Antinori 1831 (dated 31 January 1832, but pub. in Vol. dated November 1831); Rumor in *Literary gazette* (1832): 185; *Le lycée* 36, 1 January 1832; *FER* 1, series 1: 40–41.

doubtful experiment, Faraday offered a long systematic series of researches and gradually constructed the fact of electromagnetic induction.[101]

The key to Faraday's success may be seen in his ability at methodic exploration. He gave to his devices the optimal flexibility, and was attentive to the parts that could be modified according to the opportunities offered by his laboratory. He thus constructed chains or trees of experiments, feeling his way 'by facts closely placed together.' He kept in memory a large stock of previous experiments, to be explained by the new facts (in the case of the Arago effect) or instead to be used in the explanation of the new facts (in the case of the motion of magnetic poles). He settled his views and his experimental activity when a simple, coherent network of actual and virtual experiments was reached.[102]

Faraday avoided two ways of blocking the exploratory function of experiments. First, he did not divert his energies into developing practical applications. He was 'rather desirous of discovering new facts and new relations than of exalting the force of those already obtained.' He was satisfied as soon as the new effects were clear and easily reproduced (eventually in the classroom), and left to others the conception of efficient electric motors and dynamos. Second, Faraday did not let theory invade his researches. Although theoretical prejudices, such as the existence of induced currents, the electro-tonic state, or the vortices in Arago's disk, played a role in orienting his research, they were easily correctible. Faraday was proud and eager of this flexibility, and denounced the sterility of closed mathematical theories:[103]

> I do not remember that Math. have *predicted* much. Perhaps in Ampère's theory one or at most two independent facts. I am doubtful of two. Facts have preceded the math. or where they have not the facts have remained unsuspected though the calculations were ready as in electromagnetic rotation and magneto-electricity generally; and sometimes when the fact was present as in Arago's phenomenon the calculations were insufficient to illustrate its true nature until other facts came into help.

Only at the end of his experimental series on electromagnetic rotation and induction did Faraday offer a synthetic view of the explored field. We may call this view a theory because of its ability at ordering the complex. Yet it was very different from what Ampère (and Faraday) would have called a theory. It was not mathematical and it was not even quantitative.[104] It was an open scheme, in which the nature of the electric current remained undetermined. The central concept, that of magnetic power and its lines of action, had an ambiguous status. Was it a simple convenience of expression, or was it a physical entity? Faraday's operational definition of the magnetic curves suggests the first alternative. But his explanation of the attraction

[101] *FER* 1, series 1 (Nov 1831): #78n; Ampère to Faraday, 13 April 1833, *CMF* 2. Faraday was misled by the description of Ampère's experiment that he found in Demonferrand 1823 (cf. Romo and Doncel 1994: 301). He apologized in *FER* 1, series 3 (Jan 33): 107–9 (note of 29 April 1833).

[102] Faraday to Ampère, 3 September 22, *CMF* 1.

[103] *FER* 1, series 2 (January 1832): ##159; Faraday to Somerville, November 1833, *CMF* 2.

[104] The quantitative concept of the density of lines of force appeared only in 1851: *FER* 3, series 28 (October 1851): #3115, #3122.

between two currents in terms of the corresponding magnetic powers points to the second.[105]

Historians have tried in vain to eliminate the ambiguity. It may instead be seen as an essential characteristic of Faraday's investigations, from the beginning of his interest in electricity and magnetism to his latest works. On the one hand, he wished to keep his researches experimental and to secure a purely instrumental meaning of the lines of force, as a clear and efficient way to formulate the rules of electromagnetism. On the other, he strongly suspected that the magnetic curves and the tonic state were real 'links in the chain of effects' produced by magnets and electric currents. Accordingly, he expected that the magnetic and electric actions were 'progressive and required time,' as the propagation of sound and light already did. This speculation of March 1832 ended up in the safe of the Royal Society. Faraday spent twenty more years of exploration until he made it public.[106]

1.6 Conclusions

Ampère and Faraday co-founded the new science of electrodynamics. In harvesting experimental facts, their contributions were complementary: Ampère demonstrated the forces between two currents and proved the equivalence between magnets and distributions of current (not to be confused with the hypothesis of Amperean currents), while Faraday discovered continuous rotations and electromagnetic induction. However, their experimental and theoretical methods were so different that they could hardly exchange more than uninterpreted facts.

Ampère's physics was dominated by theory. Despite the inductive rhetoric of his chief memoir, he constructed his theory mostly from theoretical resources, including analogy, virtual history, and mathematical unification in neo-Newtonian style. Most of his experiments verified theoretical predictions, or decided between a preconceived range of possibilities. They were highly rigid in their construction, and reflected the mathematical structure of the theory.

In contrast, Faraday knew no mathematics and tried to minimize theoretical prejudice. He regarded the Newtonian notions of electric and magnetic fluids as unproven, and enhanced the exploratory function of experiment by systematically evolving his experimental devices. He did not aim at a closed mathematical theory, but instead maximized the mutual connections of his experimental actions

[105] When he introduced the idea of developing or contracting magnetic curves around a changing current (*FER* 1, series 2: #238), Faraday specified in parentheses that the magnetic curves were 'mere expressions for arranged magnetic forces.' In his experiments with the rotating cylinder magnet, he believed that the magnetic curves did not rotate with the magnet and concluded in a '*singular independence* of the magnetism and the bar in which it resides' (*FER* 1, series 2: #220). Williams (1965: 203–4) takes this to indicate that Faraday's lines of force had already become 'much more real.' However, the conception of the lines as a mere chart of magnetic forces implies their independence of the magnet's rotation. By emphasizing the 'singular' character of this independence, Faraday probably meant a contrast with the behavior of hypothetical magnetic fluids or Amperean currents, which would have had to rotate with the magnet.

[106] Faraday, sealed note of 12 March 1832, Royal Society, quoted in Williams 1965: 181.

and results. His agnosticism about the intimate nature of electricity and magnetism focused him on the *actions* of the various electrified or magnetic bodies rather than on their inner structure. His basic philosophical notion was that of 'power' (or 'force'), comprehending actual and virtual action. In his idiom, different powers induced different states of bodies (states of motion, or internal states). The most essential thing was the distribution of power. Sources (electrified body, magnet, current) were subordinated to the powers they developed.[107]

There is an essential harmony between Faraday's exploratory style of experimentation and his concept of power. A distribution of power may be seen as given by a set of virtual experiments. For instance, magnetic curves represent the motion that would be impressed on a magnetic pole at each point of space. By multiplying and varying experimental configurations, Faraday 'placed facts closely together,' which means that he connected actual effects to virtual experiments that determined the distribution of power. For example, he conceived the electromagnetic rotations by putting together on paper the virtual actions of the sides of a magnetic needle on a wire. Or he explained Ampère's attractions between two currents in terms of virtual rotations revealing the distribution of power around the currents. Hence, exploration could reveal or map powers. Or it could seek new sorts of states induced by a given power. Faraday was doing just that when he looked for electromagnetic induction. As the result did not fit the power-induces-state scheme, he imagined the electrotonic state.

Ampère and Faraday both strove for theoretical unity, but in a different way. The French philosopher imagined an internal structure of the various sources of action that would reduce them to one kind only. By reducing magnets to currents, he unified electrodynamics and magnetism; by reducing currents to flows of electricity, he hoped to unify electrostatics and electrodynamics; by identifying the ether with a neutral compound of the two electricities, he hoped to unify optics and electrodynamics. In contrast, Faraday achieved unity by identifying the powers emanating from various sources. In his view, magnetic, electromagnetic, and electrodynamic effects all derived from the interplay of magnetic powers. Strikingly, he adopted Ampère's statement that the side of a current was like the pole of a magnet. But where Ampère implied that magnets were made of currents, Faraday meant that the same power existed in both cases.

In its mature form Ampère's theory was successful, especially in France and in Germany. It was expressed in clear mathematical language, it relied on largely familiar Newtonian notions and techniques, and it seemed to rest on a sound empirical basis. In brief, anyone who understood it adopted it, save the more controversial Amperean currents. In contrast, Faraday's theory was completely ignored for many years. Not even the rule of the cut lines of force found favor with contemporary physicists. Faraday's first readers had difficulty with his unusual style, and missed

[107] For Faraday, 'force' and 'power' were roughly synonymous. In his later writings, 'force' is more common, and 'power' is used to indicate the kind of force, for example 'magnetic power' versus 'electric power.' Naturally, Faraday also used these two words in their most ordinary sense: for example, the 'force of this argument,' or the 'power of this electromagnet.'

the essential coherence of his material and conceptual practices. The first mathematical theories of electromagnetic induction, which form the subject of the next chapter, did not rest on Faraday's views. Their departure point was Ampère's mathematical theory together with Faraday's facts.

3

British fields

3.1 Introduction

In the first two decades of the nineteenth century, opinions on the nature of electric and magnetic actions still varied, especially in England and in Germany. There even were traces of the eighteenth century 'atmospheres,' invented to avoid direct action at a ditance. After Ampère's electrodynamics and its German extensions, however, the hegemony of Newtonian fluid theories spread from France to other countries. The penetration of the methods of French mathematical physics favored the most readily quantifiable representations of phenomena, and diverted attention from more qualitative notions.[1]

A few British physicists escaped this general evolution, and proposed alternative views of electricity and magnetism. First and foremost was Faraday, who preserved his intellectual independence and his ignorance of mathematical theories. He sought a more immediate connectivity of physical objects and phenomena, and kept exploring the space intervening between electric and magnetic sources. In the 1830s and 1840s, he accumulated major discoveries including electromagnetic induction, electrochemical equivalence, inductive capacity (dielectrics), magneto-optical rotation, and diamagnetism. At the same time he developed the field conception of electric and magnetic actions.

In the 1840s William Thomson revealed analogies between the mathematical laws of electricity and magnetism and the dynamics of continuous media. He thus invented the basic concepts of field mathematics, and pointed to a surprising equivalence between Faraday's reasonings on lines of force and potential theory *à la française*. After 1850, Thomson recurrently speculated on a dynamical ether theory of electricity and magnetism. In most of his early works, however, he avoided commitment on the deeper nature of electricity and magnetism. His concepts and apparatus could be shared by any users of these sciences, be they physicists of various schools, chemists, mathematicians, or engineers. They were immensely successful in their cross-cultural purpose, and provided a lasting foundation for future research on electricity and magnetism. With his analogies,

[1] For pre-1820 field theories, cf. Heilbron 1981.

mechanical models, and energetics Thomson championed a new kind of British physics.

The present chapter is devoted to Faraday's and Thomson's field physics, to their connections and divergences. 'Field' is here used in a loose, meta-historical sense, meaning the introduction of physical or mathematical entities in the space intervening between electric and magnetic sources. With this definition, questions about the exact origin of the field concept or the relative importance of Thomson's and Faraday's contributions become largely meaningless. Field notions in this sense already existed in the eighteenth century. Faraday's lines of force, however, provided the first precise and quantitative concept of a field. Moreover, Faraday advocated a *pure* field theory, in which electric charge and current were derivative concepts. Thomson was first to introduce mathematical field concepts, and to seek their foundation on a dynamical ether theory.[2]

To a large extent, the concepts used today in common applications of electricity are Thomson's invention. This fact eases access to his writings, but it tends to obscure Thomson's merits, which are explained in the two last sections of this chapter. In contrast, Faraday's views are difficult to penetrate, because they differ from modern ones at a very basic level. Yet they provided essential components of the system cultivated by Maxwell and his followers. This is why they are given detailed consideration in the three first sections of this chapter.

3.2 Faraday's electrochemistry

3.2.1 Wet string

In his researches on electromagnetic induction, Faraday repeatedly sought electric effects of the induced currents: spark, action on the tongue, electrolytic decomposion, and heating. He also tried to induce currents by discharging a Leyden jar in the primary of a double coil. Although his attempts met little success, he did not doubt that all forms of electricity were equivalent. Most natural philosophers had shared this opinion since Volta, and Wollaston had supported it with experimental facts. Yet doubts were occasionally expressed. For example, Davy questioned the identity of animal electricity with friction or voltaic electricity. At the end of his series on electromagnetic induction, Faraday judged that a complete, systematic proof of the equivalence of all kinds of electricity was still lacking, and he set himself to fill the gaps.[3]

Trying to improve on Wollaston's proof of electrolytic effects for friction electricity, Faraday hit upon the following device (Fig. 3.1(a)). He moistened a piece of paper with a saline or acid solution and a colour indicator, laid it on a glass plate, and let the ends of two platinum wires rest on it. The first wire was connected to an electrostatic machine through a wet string and a switch made with another glass plate and a tin foil. The second wire led to 'a discharging train,' that is, grounding

[2] On the ambiguities of the field concept and its origins, cf. Nersessian 1985.

[3] *FER* 1: 6; Wollaston 1801; Davy 1829: 17. Cf. *FER* 3, series 3: 76–7; Williams 1965: 211–23.

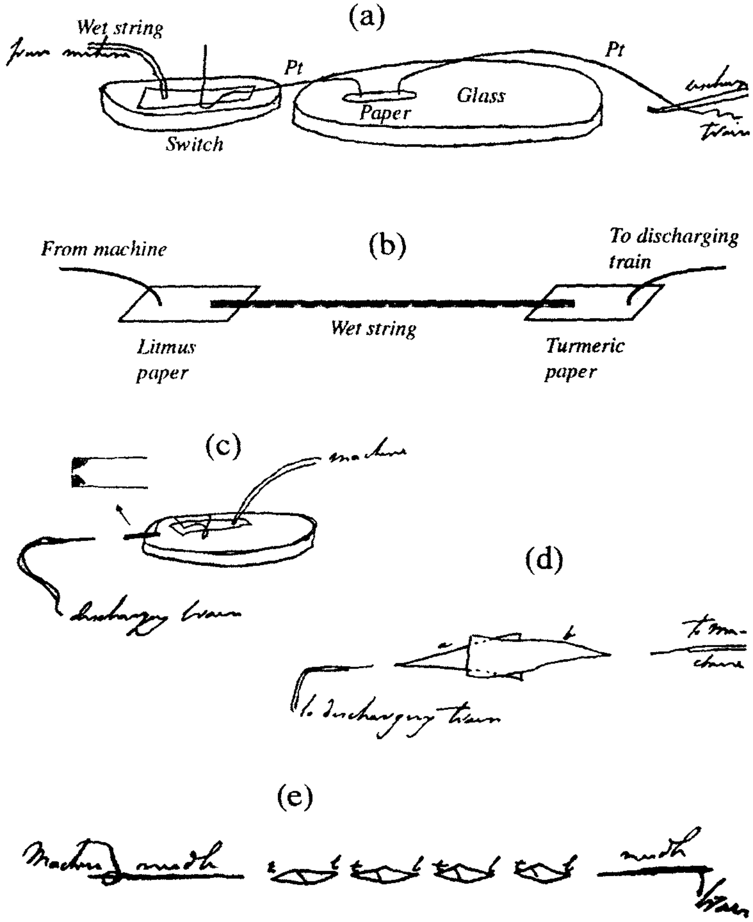

FIG. 3.1. Faraday's experiments on electrolytic decomposition: (a) with two platinum poles on paper (*FD* 2: 9), (b) with remote poles, (c) without pole at the corners of a turmeric paper (*FD* 2: 17), (d) without poles in turmeric and litmus paper, (e) without poles in a chain of turmeric–litmus papers.

by gas or water pipes. A few turns of the electrostatic machine sufficed to produce chemical decomposition at the wire ends. Faraday praised the virtues of paper: 'It makes contact by very minute surface, keeps the decomposed matter on the spot, and by its whiteness well shews the effects of change of color.' This was the starting point of a revealing series of mutations.[4]

The role of the wet string was to 'retard' the electricity by its bad conducting power and thus make it more akin to galvanic electricity. Eventually, Faraday found

[4] *FD* 2: ##46–55 (1 September 1832), quote from #51. Cf. Williams 1965: 220–3.

that the electrolytic effect was just as good without the wet string. He soon exploited another virtue of this cheap resource, that of being a non-metallic conductor. He took two pieces of paper, moistened them with a solution of soda sulphate and colored acid/alkali indicators (litmus/turmeric), and connected them with four feet of the string (Fig. 3.1(b)). The positive platinum pole, resting on the litmus paper, produced acid, and the negative one, resting on the turmeric paper, produced alkali. Faraday found that the effect was about the same as when the two papers where in direct contact. This showed that the decomposition did not depend on the distance between the two metallic poles.[5]

A moment later, Faraday placed the positive platinum pole on a piece of litmus paper, and touched the paper with a wet string connected to the discharge train. Decomposition occurred, despite the lack of a true negative pole. Two days later Faraday further modified his device so as to determine where the alkali produced in the decomposition went. He now used turmeric paper, and replaced the wet string with a metallic point about two inches from the end of the paper (Fig. 3.1(c)), exploiting the conducting power of air near charged metal points. After a few turns of the machine, the corners of the paper turned brown, thus showing the accumulation of alkali. Faraday commented:[6]

> Hence it would seem that it is not a mere repulsion of the alkali and attraction of the acid by the positive pole, etc. etc. etc., but that the current of electricity passes, whether by metallic poles or not, the elementary particles arrange themselves and that the alkali goes as far as it can with the current in one direction and the acid in the other. The metallic poles used appear to be mere terminations of the decomposable substance.

After two more days, Faraday found a better illustration of this view by combining turmeric paper, litmus paper, and two distant metal points (Fig. 3.1(d)). Finally, in April 1833 he provided a visualization of the imagined internal decomposition of electrolytes by means of an alignment of pairs of indicating papers (Fig. 3.1(e)). This case of decomposition, he commented, 'indicates at once an internal action of the parts suffering decomposition, and appears to show that the power which is effectual in separating the elements is exerted there and not at the pole.' This arrangement was similar to the iron files between the poles of magnets: in both cases Faraday took the effects observed in the intervening space to reveal the distribution of power in this space.[7]

3.2.2 The laws of electrolysis

In May 1833 Faraday sought independent confirmation of the distributed decomposing power by passing a beam of polarized light through a conducting solution. This failed. However, a few days earlier Faraday had serendipitously discovered

[5] *FER* 1: #295 (role of wet string); *FD* 2: #57 (1 September); #72 (3 September). Faraday also remarked that in most cases of sparking discharge, the reaction at the two poles was the same.

[6] *FD* 2: #74 (3 September), #81 (4 September); #99, #103 (quote, 6 September). Faraday later confirmed this result with another device in which water played the role of the negative pole: *FD* 2: #577 (30 May 33).

[7] *FD* 2: #108 (8 September); #469 (22 April 33); *FER* 1, series 5: #471 (quote).

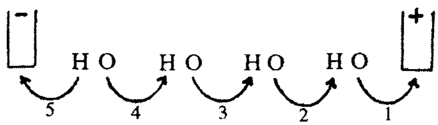

FIG. 3.2. A Grotthus chain in water (H stands for a hydrogen particle, O for an oxygen particle).

another indication that decomposition was essential to electrolytic conduction. He wanted to use ice as a non-metallic conductor in place of the wet string or the air around a metal point. But the ice would not conduct at all. In general, Faraday found that liquid electrolytes lost their conductivity when frozen. He hurried to publish this 'new law of electric conduction,' which 'afforded abundant compensation for [his] momentary disappointment.' And he queried 'whether solidification does not prevent conduction merely by chaining the particles to their places, under the influence of aggregation, and preventing their final separation in the manner necessary for decomposition?'[8]

For Faraday, electrolytic decomposition occurred within the whole substance of the electrolyte. That the products of decomposition were only seen at the ends of the substance only meant that recomposition occurred at the same rate within the substance. This view agreed with the chain process imagined by Grotthus (1806) and Davy (1807), according to which a series of decompositions and recompositions took place on lines joining the ends of the substance (Fig. 3.2). But it contradicted the more recent views of Biot (1824) and Auguste de la Rive (1825), according to which separation of the elements occurred only at the poles and was followed by migration of the charged particles of the elements. Faraday further denied that the attraction or repulsion from the poles caused the decomposition process. In his view the decomposition of a particle of the substance occurred as a consequence of the decomposition of the neighboring particle. This contradicted Grotthus's idea of a direct action of the poles on the particles of the separating elements, and also Davy's more complex idea of an action of the poles communicated by the intervening particles.[9]

From the intimate connection between electrolytic conduction and decomposition, Faraday drew two essential conclusions:[10]

1. 'The *sum of chemical decomposition is constant* for any section taken across a decomposing conductor, uniform in its nature, at whatever distance the poles may be from each other or from the section.'

[8] *FD* 2: ##482–94 (2 May 1833); ##222–49 (23–4 January 1833); *FER* 1, series 4: 110 (quote), 118 (quote).

[9] Grotthus 1806; Davy 1807; Biot 1824, Vol. 1: 636–42; A. de la Rive 1825. According to Davy's view, there should be no decomposing action in the middle of the solution. I have followed Faraday's reading of these texts: *FER* 1: 136–9. See also Ostwald 1896; Whittaker 1951: 75–7; Williams 1965: 227–1.

[10] *FER* 1, series 5 (June 1833): ##504–5. See also the weaker statement of series 7, *FER* 1: #377: '*The chemical power, like the magnetic force is in direct proportion to the quantity of electricity* which passes.' Cf. Williams 1965: 241–57.

2. *'For a constant quantity of electricity, whatever the decomposing conductor may be, whether water, saline solutions, acids, fused bodies, &c., the amount of electro-chemical action is also a constant quantity, i.e. would always be equivalent to a standard chemical effect founded upon ordinary chemical affinity.'*

The latter is what we call Faraday's law. Faraday proved it in May, August, and September 1833 by means of series of electrolytic cells of various kinds. He constructed tables of 'electro-chemical equivalents,' and described 'the only *actual measurer* of voltaic electricity,' the 'Volta-electrometer' (to become 'voltameter') based on the electrochemical decomposition of water.[11]

Faraday was especially interested in the implications for the nature of chemical forces. That the electrochemical equivalents coincided with the chemical ones seemed to confirm Davy's and Berzelius's electric conception of chemical affinity. In Faraday's words:

I think I cannot deceive myself in considering the doctrine of definite electro-chemical action as of the utmost importance. It touches by its facts more directly and closely than any former fact, or set of facts, have done, upon the beautiful idea, that ordinary chemical affinity is a mere consequence of the electrical attractions of the particles of different kinds of matter.

In short: 'ELECTRICITY *determines* the equivalent number, *because* it determines the combining force.'[12]

3.2.3 Redefining the electric current

Why should the electric nature of chemical combinations imply that they involve a definite quantity of electricity? Faraday was aware of a possible atomistic answer: 'If we adopt the atomic theory or phraseology,' he wrote, 'then the atoms of bodies which are equivalent to each other in their ordinary chemical action, have equal quantities of electricity naturally associated with them.' He preferred, however, to avoid this speculation: 'I must confess I am jealous of the term *atom*, for though it is very easy to talk of atoms, it is very difficult to form a clear idea of their nature, especially when compound bodies are under consideration.'[13]

His own explanation of electrochemical equivalence rested on a redefinition of the electric current. As we saw in Chapter 1, he had doubted the existence of electric fluids since his earliest interest in electricity. In his work on the identity of the different forms of electricity, he took an agnostic stance: 'By current, I mean anything progressive, whether it be a fluid of electricity, or two fluids in opposite directions, or merely vibration, or speaking still more generally, progressive forces.' In his series on electrochemical decomposition, he made a more definite choice:[14]

[11] *FER* 1: #739; #1355 (voltameter).

[12] *FER* 1: #248, #256 (Faraday's emphasis). Also *FD* 2: #1917 (5 August 1834): 'The electricities appear to be the forces of attraction by which two particles combine.'

[13] *FER* 1, series 7: #869. [14] *FER* 1, series 3: #283; series 5: #517.

Judging from the facts only, there is not as yet the slightest reasons for considering the influence which is present in what we call the electric current,—whether in metal or fused bodies or humid conductors, or even in air, flame, and rarefied elastic media,—as a compound or complicated influence. It has never been resolved into simpler or elementary influences, and may perhaps best be conceived of as *an axis of power having contrary forces, exactly equal in amount, in contrary directions.*

Faraday meant, like Oersted, that the current consisted of the propagation of a polar state in the conductor. In the case of electrolytic conduction, the polar state corresponds to the brink of decomposition; it propagates by a series of decompositions–recompositions.[15] In brief, the current *is* the series of decompositions and recompositions: 'I have no idea,' Faraday wrote to a friend, 'that in what we call the current in the decomposition of bodies anything but a resolution and recombination of forces occurs between contiguous particles.' Therefore, the strengh of the current determines the amount of decomposition, in conformity with Faraday's law. In other cases of conduction, such as metallic conduction, decomposition could not be involved. Yet Faraday suspected that a similar mechanism occurred. In his study of self-induction, published in December 1834, he proposed that any electric current involved a recurring state of temporary excitation, which could well be the electro-tonic state. Conduction occurred by 'vibrations, or by any other mode in which opposite forces are successively and rapidly excited and neutralized.'[16]

Faraday judged his concept of the electric current to be incompatible with the usual electrochemical terminology. Phrases like 'the positive pole' were impregnated with the fluid concept of electricity, and referred to electrostatic action at a distance. Faraday wanted to replace them with more neutral terms. To the inventor of the word 'physicist,' Reverent William Whewell, he wrote for consultation:[17]

'The ideas of a current especially of *one* current is a very clumsy and hypothetical view of the state of electrical forces under the circumstances. The ideas of *two* currents seems to me still more suspicious and I have little doubt that the present view of electric current and the notions by which we try to conceive of them will soon pass away and I want therefore names [. . .] without involving any theory of the nature of electricity.'

After a friendly exchange, the two men agreed on a new vocabulary for electrochemical decomposition: 'electrode,' 'anode,' and 'cathode' for the terminations of the decomposing matter, 'electrolysis' for the decomposition itself, 'electrolyte' for a substance directly decomposed by the electric current, 'ion,' 'anion,' and 'cation' for the bodies passing to the electrodes. Note that 'anode' and 'cathode' did not merely replace the positive and negative poles of the older terminology; they included cases of decomposition without metallic pole. For example, in the experiment of Fig. 3.1(c) with the turmeric paper placed at a distance of a discharging negative needle, the cathode was the end of the paper facing the needle. Faraday's

[15] *FER* 1, series 5: ##519–20.

[16] Faraday to Lemon, 25 April 1834, *CMF* 2; *FD* 2: #1167 (2 December 1833): 'Consider the transmission of electricity: that there are three modes as in a metal wire, in decomposing fluids, through air, vapour, etc. as spark or brush. Are not these all one?'; *FER* 1, series 9: #1115.

[17] Faraday to Whewell, 24 April 1834, *CMF* 2. Cf. Williams 1965: 257–67.

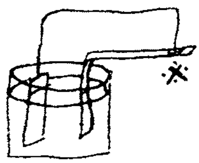

FIG. 3.3. Decomposition of 'hydriodate of potassa' (KI) at X, with no metal–metal contact (*FD* 2: 244). The straight electrode and wire are made of platinum; the bent electrode of amalgamated zinc. Both are immersed in a sulfonitric solution.

definition thus reflected the elimination of attractions and repulsions in favor of decomposition. The ethymology did the same: Whewell forged 'anode' and 'cathode' from the Greek for 'upwards way' and 'downwards way,' alluding to the fact that currents around the Earth's axis would agree with its magnetism if they followed the motion of the Sun. The polarity of the current was thus defined without reference to positive or negative electricity.[18]

3.2.4 *To the trough*

Faraday completed his study of electrolysis in late 1833. In February 1834 he decided 'to go to the trough.' In the electric conception of chemical affinity, a voltaic cell was just an electrolytic cell working backwards, the chemical decomposition now being the cause of the current. Accordingly, Faraday supported Davy's chemical theory of the pile and rejected the more popular contact theory. He showed that the electromotive force of a cell largely depended on the electrolyte, and gave numerous examples of decomposition without the contact of two metals, the first being the decomposition of potassium iodide in the circuit zinc/sulfonitric mixture/platinum/potassium iodide solution/zinc (Fig. 3.3). The argument, however, failed to convince the German and Italian supporters of the contact theory. The latter theory could be generalized to include contact tension between metals and non-metals.[19]

Five years later Faraday accepted the objections of some of his adversaries, especially Stefano Marianini's. He resumed his studies of the voltaic trough, multiplied facts in favor of the chemical theory, and denounced a fundamental defect of the contact theory: it denied 'the great principle in natural philosophy that cause and effect are equal.' If the contact of two metals was the source of the voltaic effect, then one would have a 'production of power without a corresponding exhaustion of

[18] Whewell to Faraday, 25 April, 5 May 1834, *CMF* 2; Faraday to Whewell, 3 May 1834, *ibid.*; *FER* 1, series 7: #661–2.

[19] *FD* 2: #1487 (10 February 1834); #1577 (19 February 1834); *FER* 1, series 8. On the reception, cf. *FER* 1: #1769 and Ostwald 1896, Vol. 1: 476–80 for the leading German supporter of the contact theory, C. Pfaff, and Vol. 2: 693–701 for other Germans.

something to supply it.'[20] In the 1850s, after energy considerations had become central to physics, this argument certainly weakened the positions of the contact theorists. It failed, however, to induce any spectacular conversion to Faraday's views. On the contrary, the British herald of energy physics, William Thomson, restored a form of contact theory.[21]

Most electrochemists admitted voltaic cells without metal–metal contact and electrolysis without metal electrodes. They adopted Faraday's terminology, and they applauded his demonstration of electrochemical equivalents. Yet they did not easily abandon the contact theory, and they completely ignored Faraday's concept of the electric current. Those who wished to explain the law of electrochemical equivalents did it with atoms and electric fluids. No one took seriously Faraday's claim that his views were the mere expression of facts. His exclusion of alternative theories indeed depended on personal judgments: to him electric fluids were unphilosophical, atomistic explanations were too arbitrary, and no acceptable theory could be based on principles whose consequences could only be foreseen mathematically.[22]

3.3 Dielectrics

3.3.1 Redefining electric charge

After founding a new electrochemistry and suggesting a new view of the electric current, Faraday still did not know how to define electricity. In April 1834, he confided to a friend: 'At present my view is very unsettled with regard to the nature of the electric agent. The usual notions attached to Positive and Negative and to the term current I suspect altogether wrong but I have not a *clear view* of what ought to be put in their places.' In the case of electrolysis, he could not distinctly see how the 'resolution and recombination of forces between contiguous particles' came about. A query of November 1835 indicates how he soon hoped to solve this problem:[23]

> Have been thinking much lately of the relation of common and voltaic electricity: of [electrostatic] induction by the former and decomposition by the latter, and am quite convinced that there must be the closest connexion. Will be first needful to make out the true character of ordinary electrical phenomena.

[20] Marianini 1837; *FER* 2, series 17 (January 1840): #2069, #2071. Also *FD* 3: #5112 (26 August 1839): 'By the great argument that no power can be ever be evolved without the consumption of an equal amount of the same or some other power, there is *no creation of power*; but contact would be such a creation.' Cf. Williams 1965: 364–72. Unknown to Faraday, in 1829 Peter Roget had already noted that the contact theory implied the possibility of perpetual motion: cf. *FER* 2: 103n.

[21] Thomson 1862. Thomson had a number of followers, while Maxwell, Lodge, and Heaviside supported the chemical theory. For a penetrating study of the resulting controversy, cf. Hong 1994a. On the persistence of German contact theory, cf. Ostwald 1896, Vol. 2: 731–40.

[22] The modern electrolytic theory does not meet Faraday's criteria, for it requires electric atoms (the electrons) and explains the role of the electrodes by complex electro-kinetic calculations.

[23] Faraday to Lemon, 25 April 1834, *CMF* 2; *FD* 2: #2468 (3 November 1835), anticipated in *FD* 2: ##1846–7 (22 February 1834). Cf. Williams 1965: 287.

Faraday's intuition was that an insulator submitted to an electric source (a battery or an electrostatic machine) was in the same state of tension as electrolytes were before decomposition: 'This is the state of an electrolyte in the circuit before that traversing of the particles has taken place by which the electric force is transferred and the body conducts. It ought to be the state of the electrolyte when it is solid.' There was only one difference: in the insulator the state of tension could last as long as the source was in action, whereas in the electrolyte the tension was continually resolved by decomposition. This state of tension, Faraday speculated, could be the essence of electricity, as continual decomposition was the essence of the electrolytic current. Hence came the next query: 'Does common electricity reside upon the surface of a conductor or upon the surface of the electric [insulator] in contact with it? I think upon the electric, and must work out the results on that view,' with the comment: 'Would be a reason why all upon the surface of conductors.'[24]

According to Faraday's new intuition, the insulator under electrostatic induction was polarized: any part of it, separated in imagination, was positive on one side and negative on the other, just like the Northernness and the Southernness of the parts of a magnet. A surface *within* the polarized insulator had no net charge since the charges of its two sides mutually cancelled, but the surface of contact between the conductor and its insulator could be charged, since the conductor, by definition, could not sustain polarization. Then electric charge was just the termination of polarization; it belonged to the insulating medium, not to the conductor. The insulator now being the *locus* of electricity, Faraday called it 'the electric,' to become the 'dielectric' under Whewell's suggestion. As he later wrote, 'the great point of distinction and power (if it have any) in the theory is, the making the dielectric of essential and specific importance.'[25]

Faraday imagined three ways of testing his new view. First, experiments with hollow conductors would prove the impossibility of absolute charge: if charge derived from the beginning or ending of polarization, every charge was always related to an opposite charge. Second, any effect of the composition of the 'electric' would prove 'that the electricity is related to the electric, not to the conductor.' Third: 'Must try again [as for an electrolyte under decomposition] a very thin plate under induction and look for optical effects, i.e. detect its polarized state.'[26]

3.3.2 No absolute charge

Faraday never succeeded on the third point (the later 'Kerr effect'). But he did on the first two. On 26 November, he verified that an electrified quart pot contained no

[24] *FD* 2: #2511, #2469 (3 November 1835). Cf. also Faraday's historical remarks, *FER* 1: 362: 'As the whole effect in the electrolyte appeared to be an action of the particles thrown into a peculiar state, I was led to suspect that common induction itself was in all cases an *action of contiguous particles*, and that electrical action at a distance (i.e. ordinary inductive action) never occurred except through the influence of the intervening matter.'

[25] *FD* 2: #2507–8 (3 November 1835); Whewell to Faraday, 29 December 1836, *CMF* 2; *FER* 1: 364: The dielectric, 'that substance through or across which the electric forces are acting'; *FER* 1, series 13: #1666.

[26] *FD* 2: ##2474–91; #2497; #2512. Cf. Gooding 1978.

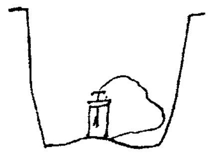

FIG. 3.4. One of Faraday's experiments with an electrified copper boiler (*FD* 2: 408). The gold-leaf electrometer at the bottom of the boiler shows no sign of electricity.

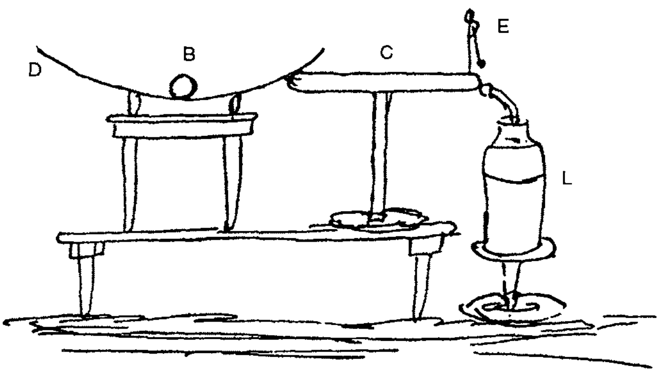

FIG. 3.5. Faraday's contrivance showing induction with no charged body in view (*FD* 2: 417). L denotes a Leyden jar, E an electrometer, C a conductor, D a metal dish, and B a carrying ball to test charge inside the dish.

electricity. In early December he performed more systematic experiments with an electrified copper boiler, electrometers, and carrying balls (Fig. 3.4). He found that the inside bottom of the boiler carried no charge, no matter how electrified the boiler globally was. In Faraday's view, this meant that a part of the surface of a conductor could be charged only if it was inductively related to another conductor. In other words, polarization starting on one conductor had to finish on another, and any charge was inductively related to an opposite charge.[27]

Yet Faraday admitted for a while that the rule could be invalidated when very large lengths of insulator were involved. On 10 December, he experimented with a spherical copper mirror (Fig. 3.5) facing the starry sky on a dry cold night and connected to a charged Leyden jar through a conductor. The bottom of the mirror had no conductor in view, and yet it carried electric charge. Faraday conjectured that the polar tension from the inner surface decayed over long distances and therefore did not need to be terminated by another charged conductor. He was then 'pretty sure' that the inside walls of a very large hollow conductor could be charged. But he soon wondered: 'Can induction through air take place in curves or round a corner? can

[27] *FD* 2: #2634 (26 November 1835); ##2664–736 (5–8 December 1835).

probably be proved experimentally. If so, is not a radiating action, and reasoning as to sky action requires modification.'[28]

In January 1836 Faraday built a twelve foot conducting cube with wood, copper wire, paper, and tin foil. He connected it to an electrostatic machine, and 'lived in it' to check the internal electric state with carrying balls and electrometers. No electricity could be found that would not be explained by the imperfect conducting power of tin foil or by induction at the entrance of the cube. Faraday concluded that induction was 'illimitable,' that there could be no global loss of power along a polarized dielectric. For example, the charge of a conductor placed in a hollow conducting sphere would always be equal to the charge of the inside walls of the sphere, no matter how large this sphere was. Consequently, absolute charge did not exist, all charge was sustained by induction and was related to another, opposite charge.[29]

3.3.3 Specific inductive capacity

Toward the end of 1836, Faraday turned to the other proof of his theory of induction: the dependency of induced charges on the nature of the dielectric. For this purpose, he built a Coulomb torsion balance according to Coulomb's own directions, with a few improvements: for instance he placed a grounded, double conducting belt on the glass cylinder surrounding the balance, 'so that the inductive action within the electrometer might be uniform in all positions.' Then he built two exemplars of a new sort of Leyden jar, in which the dielectric could be changed at pleasure. This apparatus, represented on Fig. 3.6, is made of two concentric brass spheres *aa* and *hn*. The internal sphere *hn* is connected to the conducting ball B through the brass rod *i*. The shell-lac (an excellent insulator) stem *ll* sustains the latter system and seals the collar of the external sphere. A solid or liquid dielectric can be placed in the space *oo* by opening the external sphere at the joint *b*. A gaseous one may be introduced through the stopcock *d* after proper exhaustion.[30]

The principle of Faraday's measurements was simple. He first charged the first exemplar of the apparatus with a Leyden jar and measured its 'tension' or 'degree of charge' (which I call V_0) by making a carrying ball touch the fixed ball B and measuring its resulting charge with the Coulomb electrometer. Then he brought the two exemplars into contact, and measured their resulting tensions (V_1 and V_2), which ought to be equal. Figure 3.7 gives the modern schematics of this procedure. If the two exemplars are filled with different dielectics they have different 'capacities for electric induction' C_1 and C_2. The original charge of the first apparatus, C_1V_0, is divided into C_1V_1 and C_2V_2 by contact with the second apparatus. Consequently, the

[28] *FD* 2: ##2741–6 (10 December 1835); #2766 (12 December 1835). Cf. Gooding 1978: 137.

[29] *FD* 2: ##2808–74 (15–16 January 1836); *FER* 1, series 11: #1174; *FD* 2: #2826 (15 January), 2864 (16 January); *FER* 1: #1178, #1295. Cf. Gooding 1978: 139–42.

[30] *FD* 3: ##3622–4015 (23 December 1836–7 October 1837) and ##3597–3621 (22 November to 21 December 1836) for preliminary experiments; *FER* 1, series 11: 311–409, quote from 368. Cf. Williams 1965: 291–4.

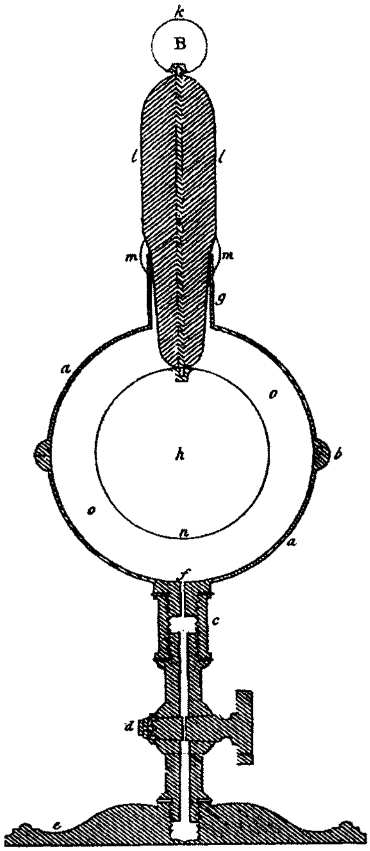

FIG. 3.6. Faraday's apparatus for studying specific inductive capacity (*FER* 1: plate 7).

ratio of the fall of tension $V_0 - V_1$ of the first apparatus to the gain of tension V_2 of the second yields the ratio C_2/C_1 of their capacities. This is how Faraday reasoned, with specific numbers instead of letters.[31]

The experiments were extremely difficult and lasted more than a year. Faraday took great precautions to avoid parasitic induction and charges. This involved long trials and ingenious tricks, for example breathing on the shell-lac and wiping it with a finger wrapped in a silk handkerchief. He had to operate very quickly in order to avoid electric leaking. Furthermore, he found that most dielectrics were not perfect insulators and that they could consequently 'absorb' electric charge. Special care had to be taken to circumvent this phenomenon. Faraday's description of his complete procedures took no less than 20 pages.[32]

[31] *FER* 1: ##1257–9. [32] *FER* 1: #1203, ##1233–1.

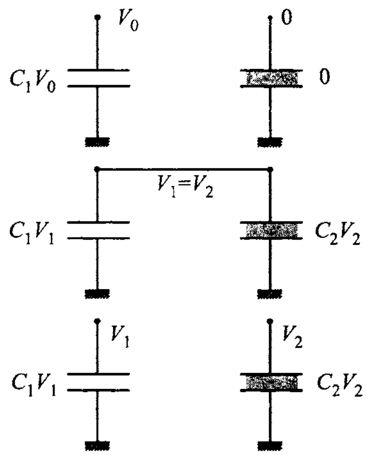

FIG. 3.7. Modern schematics of Faraday's experiments on specific inductive capacity.

For solid and liquid dielectrics, the results were beyond Faraday's expectations: the increase of inductive capacity could exceed 50 per cent. However, gases and rarefied air were a disappointment: Faraday could not find any appreciable difference in their specific inductive capacities. He nevertheless regarded induction by contiguous particles as generally proved, and took the uniform inductive capacity of gases as a further example of the simplicity of their physical properties.[33]

3.3.4 Induction in curved lines

Faraday's extreme attention to details led him to side-discoveries. Electric absorption (later called dielectric after-effect) is one of them. Another was induction in curved lines. Faraday found that parasitic charges on the shell-lac stem could charge the carrying ball when the latter was brought to touch the top of the *grounded* fixed ball B. The carrying ball being small, it was not in sight of the parasitic charges. Since B and all other nearby conductors were grounded, induction could only come from the stem. Therefore, induction had to proceed in curved lines around the ball B.[34]

In the ultimate version of this experiment, Faraday avoided direct contact between the carrying ball and the fixed ball. He grounded the latter ball, brought the carrying ball near the top, uninsulated it for a short while by means of a wire to the ground, insulated it again, and found it to be charged. This version of the experiment excluded the transfer of charge from the stem to the carrying ball through the fixed ball that the defenders of the fluid theory of electricity could have evoked. In Faraday's opinion, this experiment was the best proof that induction was an

[33] *FER* 1: #1260: 'This extraordinay difference was so unexpected in its amount...'; *ibid.*: #1292.
[34] *FD* 3: #4016.

action between contiguous particles, because a direct action would have been rectilinear.[35]

3.3.5 General views

For Faraday, induction in dielectrics was 'the essential principle for the development of electricity.' Electric charge was a spatial discontinuity of induction, and the electric current was a variation in time or a propagation of induction. For metallic conduction or sparking discharge, a temporary polarization of the particles of the medium occurred, followed by a mutual discharge of contiguous particles. In an electric wind (slow discharge of a metallic point in air), electrified particles traveled with their surrounding induction. In an electrolyte, Faraday imagined 'first a polarisation of the molecules of the substance, and then a lowering of the forces by the separation, advance in opposite directions, and recombination of the elements of the molecules.' In every case, an invariable amount of induction was transfered along the conducting channel. As Faraday put it, the electric current was 'constant and indivisible.'[36]

Therefore the magnetic power of a current was 'the same, whether [the current was] passing in an electrolyte, or in a conductor, or in spark, or in [electric] wind or even in inductive action.' Consequently, a convection current and even a dielectric under a varying state of induction had magnetic power, or 'such at least seem[ed] to be the case.' Conversely, Faraday expected that the crossing of magnetic curves would polarize a dielectric. More generally, he anticipated a role of the particles of matter in the communication of magnetic force. However, all the experiments he performed in 1838 on this subject failed.[37]

Induction being central to his views on electricity, Faraday sought to visualize it. From February 1836 to September 1837 he experimented on discharges in gases at various pressures and for various shapes of the conductors. He regarded the beautiful sparks, brushes, and glows as evidence of the previous state of induction of the gas. In the middle of these researches he introduced the 'lines of electric induction' or tension, a concept similar to the magnetic curves or to the lines of current:

> The description of the current as an axis of power, which I have formerly given, suggests some similar general expression for the forces of quiescent electricity. *Lines of electric tension might do*; and through I shall use the terms Pos. and Neg., by them I merely mean the *termini* of such *lines*.

[35] *FD* 3: ##4016–1 (7 October 1837), ##4092–104 (14 October 1837). *FER* 1: 380–3. The grounding of the screen, though essential, was overlooked by Riess, A. de la Rive, and Melloni: cf. Riess to Faraday, 10 December 1855, *SCMF* 2. Faraday had obtained earlier indication of induction in curved lines while experimenting with the cage: *FD* 2: ##2866–7 (16 January 1836).

[36] *FD* 3: #3425 (3 August 1836); *FER* 1, series 12: #1338 (metals), ##1405–1424 (sparks), ##1562–610 (convection); #1347 (electrolysis, also #1622–1624); *FER* 1, series 13 (February 1838): #1627 (indivisible current). Faraday thus anticipated Maxwell's doctrine that all currents are closed.

[37] *FD* 3: #3471 (3 August 1836); also *FER* 1, series 14 (June 1838): #1644, #1654; ##1709–30 (failed experiments). The contribution of the variation of dielectric polarization to the current corresponds to Maxwell's later displacement current. When seeking a role of matter in the communication of magnetic actions, Faraday also had in mind the electro-tonic state (*FER* 1: #1729).

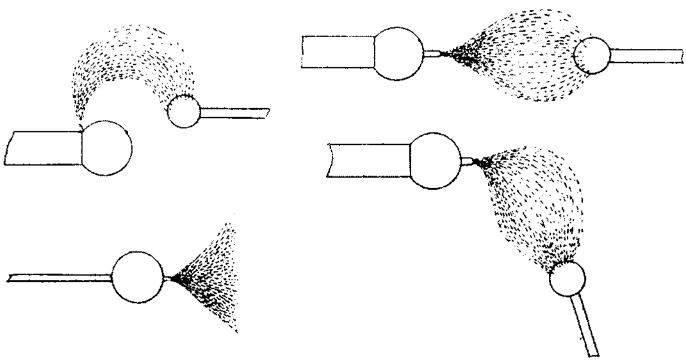

FIG. 3.8. Electric brushes (*FER* 1: plate 8).

In the case of discharge through gases, the observed luminous patterns illustrated the lines of induction (Fig. 3.8), as iron filings did for the magnetic lines of force.[38]

Faraday insisted that the lines of induction were 'imaginary,' as the lines of magnetic force already were. They were 'a temporary conventional mode of expressing the direction of the power in cases of induction.' However, Faraday imagined a mechanical tension along, and a mutual repulsion between the lines. The repulsion explained induction in curved lines and the concentrating power of conducting edges. The tension accounted for the motive forces between charged bodies. Accordingly, two charged bodies could only *attract* each other to the extent of their connection by lines of induction. Electric repulsions were only apparent, they were a consequence of the attraction by other bodies. For example, the leaves of a gold-leaf electrometer did not really repel each other: they were attracted by surrounding bodies charged by induction.[39]

As for the cause of induction, Faraday refused to take a stand. He focused on the 'manner in which the electric forces are arranged in the various phenomena generally [. . .] without committing [himself] to any opinion as to the cause of electricity.' However, he imagined a molecular process for the propagation of induction. In his view all matter was made of polarizable particles scattered in empty space. A polarized molecule polarized its nearest neighbors and only them, so that the action between distant molecules could only be indirect, through chains of intermediate polarized molecules. This transfer of polarization occurred without global loss of intensity, which explained the 'illimitable' character of induction, in conformity with the representation by lines of force. The attraction of successive polar molecules along the same line explained the tension along this line. The repulsion between molecules placed side by side explained the lateral repulsion of the lines of induction. That is to say, the polarized molecules behaved like tiny bar magnets.

[38] *FD* 2: 443–67; *FD* 3: 14–96; *FER* 1, series 12: 447–72, series 13: 473–502; *FD* 3: #3423 (quote, 3 August 1836); *FER* 1, series 11: #1224, #1231. Cf. Gooding 1978: 142–3.

[39] *FER* 1, series 11: #1304 (imaginary), #1231 (temporary), series 12: ##1371–4 (tension and repulsion); *FD* 2: #2642: 'I begin to doubt electric repulsion altogether,' also #2653 (26 November 1835).

Note that Faraday did not completely eliminate action at a distance. The action between nearest neighbors was direct action at a distance. For the time being, he only wished to eliminate the large-scale action at a distance implied in the usual fluid theories of electricity.[40]

Faraday's speculations stopped there. He did not explain the polarity of molecules, he just defined it as 'a disposition of force by which the same molecule acquires opposite powers on different parts.' Whether this disposition resulted from some kind of stress or from the displacement of a fluid, he did not want to decide. This is how one should understand his surprising statement: 'The theory of induction which I am stating does not pretend to decide whether electricity be a fluid or fluids, or a mere power or condition of recognized matter.'[41]

3.3.6 Incommunicability

Faraday's demonstration of specific inductive capacity was immediately hailed as a major electric discovery. His experiments on hollow conductors and on induction in curved lines became popular. His theoretical interpretation, however, fell in deaf ears. Well aware of the iconoclastic character of his views, Faraday concluded his eleventh series with the words:

> I beg to say that I put forth my particular view with doubt and fear, lest it should not bear the test of general examination, for unless true it will only embarrass the progress of electrical science. It has long been on my mind, but I hesitated to publish it until the increasing persuasion of its accordance with known facts, and the manner in which it linked effects apparently very different in kind, urged me to write the present paper.

Faraday's fear was justified. His theoretical statements met more misunderstanding and suspicion than ever.[42]

Experts in mathematical electrostatics had at least one good reason to doubt the soundness of Faraday's reasonings. His statement that induction in curved lines, specific inductive capacity, and impossibility of absolute charge were 'not consistent with the theory of action at a distance' could easily be refuted. A few theorems of French electrostatics covered the first and third facts. Faraday could not see this, for he judged himself 'unfit to form a judgement of [Poisson's] admirable papers' of 1811. With respect to specific inductive capacity, Faraday overlooked that the dependence of induction on the dielectric substance did not necessarily imply that the substance was *entirely* responsible for the capacity. Ottaviano Mossotti and William Thomson soon devised mathematical theories of dielectric polarization based on

[40] *FD* 3: #3423 (3 August 1836); #3512 (6 September 1836): 'Induction [. . .] an action of contiguous particles affecting each other in turn, and not action at a distance'; *FER* 1, series 11: 409–11 and 362n: 'The word *contiguous* is not the best that might have been used [. . .]; For as particles do not touch each other it is not strictly correct [. . .]. By contiguous particles I mean those which are next'; *FER* 1: #1231. Cf. Gooding 1978: 122–7.

[41] *FER* 1: #1304; *ibid.*: 409n. In the diary (*FD* 3: ##4567–70, 1 April 1838), Faraday assumed that the molecules of insulators were conductors insulated from each other (as in Poisson's theory of magnetic polarization). But he did not specify the relevant conduction mechanism.

[42] *FER* 1, series 11: #1306. Cf. Williams 1965: 372.

standard action at a distance. Anyone who could understand their calculations was likely to share the following opinion of a German expert:[43]

Thus [by extension of Coulomb's and Poisson's electrostatics] the electrostatic problems are changed into problems of pure mechanics [. . .]. The advantage of this method is very great, it gives the result of each experiment as the sum of single actions which the mind connects without difficulty, and leaves to the mathematier [sic] the pains to sum up the single effects and to find the amount of the sum [. . .]. Therefore, I have long ago defended this theory against its—indeed not very dangerous—antagonists and I could not abstain continuating the defense, as arose an adversary in the man whom I venerate as the greatest natural philosopher of the age.

Although Faraday failed to disprove the mathematical fluid theory, he could have convinced his readers that his views were a possible alternative. Yet this almost never happened. Even favorably disposed readers found obscurity and absurdity in his statements. Faraday ascribed this communication breakdown to the ambiguities of language:

I feel that many of the words in the language of electrical science possess much meaning; and yet their interpretation by different philosophers often varies more or less, so that they do not carry exactly the same idea to the minds of different men: this often renders it difficult, when such words force themselves into use, to express with brevity as much as, and no more than, one really wishes to say.

A first example of misunderstanding concerned the impossibility of absolute charge. Faraday wrote: 'It is impossible to charge a portion of matter with one electric force independently of the other.' Many of his readers understood that isolated charged particles could not be produced. In fact Faraday only meant that the charge would be the starting point of an induction that would end somewhere (possibly very far) as an opposite charge.[44]

Another difficulty had to do with induction in curved lines. Call A the originally charged body, B the influenced body, and C the screen. According to Faraday, the action of A on B was diminished by the presence of C. Reverent Whewell, Auguste de la Rive, Macedonio Melloni, and Peter Riess, among others, did not see how the action between two bodies could depend on a third body. In their interpretation (Fig. 3.9(b)), the conductor C was polarized under the effect of A, and what Faraday observed was the superposition of the actions of A and C on B. Faraday did not deny that C acted on B if the body C was insulated. But in his experiment C and B were grounded, so that no line of force could connect them. The effect of C was to deflect the lines of force from A and thus to diminish the number of lines reaching B (Fig. 3.9(a)). At that point, misunderstanding became total and reciprocal. Riess maintained that in this case too, the presence of C did not at all diminish the action of A on B. What was observed was the additional effect of the negative charge appearing in C by influence (Fig. 3.9(c)). At the root of the misunderstanding was

[43] *FER* 1: ##1166–8, #1305; Riess to Faraday, 10 December 1855, *SCMF* 2. On Thomson and Mossotti, cf. next section. On Riess, cf. Simpson 1968: 122.

[44] Faraday to Hare, 18 April 1840, *FER* 2: 262; *FER* 1: #1177; *FER* 2: 268. The misunderstanding is, e.g., in Hare to Faraday, *FER* 2: 254.

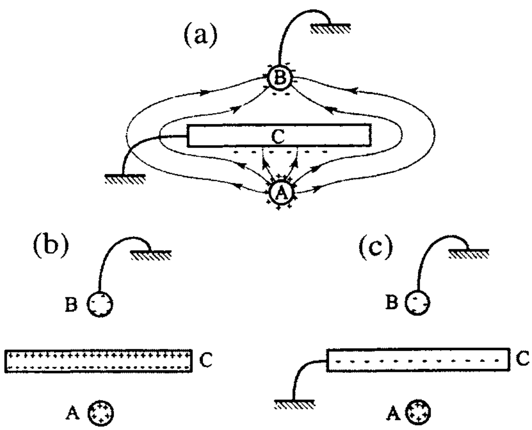

FIG. 3.9. Faraday's 'induction in curved lines': (a) authentic, (b) misinterpreted by Whewell and Riess, (c) reinterpreted by Riess.

a different concept of interaction. For Riess, the interaction of A and B was completely defined by their electric states and their relative position. For Faraday, this interaction was measured by the number of lines of force connecting the two bodies.[45]

Lastly, Faraday's molecular picture of induction arose much suspicion. Robert Hare, a chemistry professor at Penn, reproached Faraday with admitting both contiguous action and direct action at a distance. Indeed Faraday insisted that induction was an action between contiguous particles, but at the same time he admitted direct action through a vacuum, even over one inch. How, asked Hare, could the same law apply to the forces developed in such different circumstances? Should one, in the case of vacuum, admit the existence of an ether whose particles could be polarized as the particles of matter were?[46]

In his public reply Faraday first recalled that he was concerned with the arrangement of electric force, not with the deeper nature of electricity. By action between contiguous particles, he only meant action between successive particles, and not contact action as Hare assumed. A given excited particle could act directly on distant particles if no other particle existed in the intervening space. In this sense contiguous action was perfectly compatible with direct action in a vacuum. Faraday further showed that if the $1/r^2$ law applied to a vacuum it also applied to a dielectric if the latter behaved like a myriad of mutually insulated conducting molecules.[47]

[45] Whewell to Faraday, 22 November 1848, and Faraday to Whewell, 24 November 1848, *SCMF* 1; A. de la Rive 1853, Vol. 1: 143–4; Riess 1854; Faraday to Riess, 19 November 1855, in Faraday 1856 and in *SCMF* 2; Riess to Faraday, 10 December 1855, *SCMF* 2.

[46] Hare to Faraday, July 1840, in *FER* 2: 251–61, esp. 251, 252, 260; *FER* 2: #1616 for the one-inch vacuum.

[47] Faraday to Hare, 18 April 1840, *FER* 2: 262–74, esp. 262, 265–267, 264–5. Cf. Gooding 1978: 119–27.

A few years later, Faraday detected a paradox in his representation of matter: the interstitial vacuum had to be a conductor in conducting bodies and an insulator in insulating bodies. He then proposed a Boscovichian speculation: atoms could be centers of power surrounded by an atmosphere of force. Since the atmospheres never completely vanished, 'matter fill[ed] all space,' and the paradox disappeared. In previous works Faraday had replaced matter-bound imponderable fluids with powers in intervening spaces. Thanks to Boscovichian atoms, he could perhaps subsume all physics under the concept of power. If matter was condensed power, its role in the transmission of electric and other actions became self-evident.[48]

Faraday could hardly hope to disarm his colleagues' criticisms with such speculations. But he could strengthen his immunity against alien approaches. As the acute Riess noted in a letter to Faraday, this was for the better profit of science:[49]

> I have little hope to persuade you, my dear Sir, to modify your views [on electricity] and, I confess, if I could I would scarcely wish it. The great philosopher works best with his own tools, whose imperfections he avoids by dexterous application. But these tools, so efficacious in his hand, are not only useless but very dangerous in the hands of others.

3.4 The magnetic lines of force

3.4.1 Illuminating a magnetic curve

In 1839 overworking, memory loss, and perhaps feelings of intellectual isolation plunged Faraday into a long period of depression. He did not, however, forget his aim of proving the role of matter in the propagation of force. Stimulated by William Thomson, in the fall of 1845 he renewed his earlier attempt to show an effect of electrolytic currents on polarized light. He also tried the similar effect with a polarized, transparent dielectric. Both experiments failed. A week later, he 'worked with lines of magnetic force, passing them across different bodies' (Fig. 3.10). Air and flint glass did not work. However, a special kind of heavy glass that Faraday had earlier fabricated for optical use induced a faint but distinct rotation of the plane of polarization, when the light beam was parallel to the lines of force. Faraday exulted: 'This fact will most likely prove exceedingly fertile and of great value in the investigation of both conditions of natural forces [magnetism and light].'[50]

The rotation of the plane of polarization suggested some kind of rotation within the substance of the glass. 'Is it possible,' Faraday asked, 'that similar electric cur-

[48] Faraday 1844a: 293; also *FER* 3: #2225. Cf. Williams 1965: 375–80. Faraday's reference to Boscovich should be taken in the vague sense of atoms as centers of force. Faraday ignored other essential characteristics of Boscovich's atomism: cf. Spencer 1967; James 1985: 142–3.

[49] Riess 1856: 17. Cf. the defense of incommensurability in Biagioli 1990.

[50] Thomson to Faraday, 6 August 1845, *SCMF* 1; *FD* 4: ##7434–71 (30 August, 1 September 1845), earlier attempt in *FER* 1, series 8: ##951–5; *FD* 4: ##7483–97 (5 September 1845), already planned in *FD* 2: #2512 (3 November 1835); *FD* 4: #7498, #7504 (13 September 1845). Cf. Williams 1965: 384–91; Gooding 1981: 234–6.

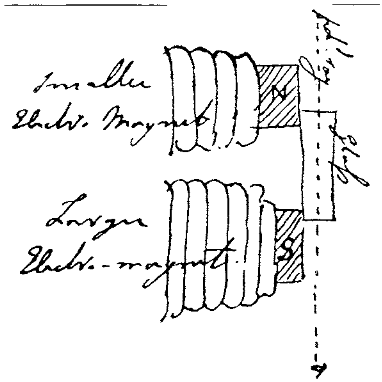

FIG. 3.10. Faraday's first device producing the magnetic rotation of a polarized ray (*FD* 4: 264).

rents are circulating both in the particles of the Iron and the particles of the glass? Or rather, perhaps, may it not be that in the iron there are circular currents, but in the glass only a tension or tendency to circular currents?' Faraday subsequently obtained the effect with other transparent bodies including oil of turpentine, flint glass, rock salt, water, and alcohol. In order to increase the length of action of the magnetic force, instead of a magnet he used a very long solenoid containing a cylinder of the transparent body. In this device, the polarization plane of the light rotated in the same direction as the helical current. The 'beautiful simplicity' of this law strengthened Faraday's idea of an internal rotation: 'Cannot but suppose some relation or similarity of constitution between bodies rotating per se and such as are under the influence of magnetic force. Such bodies are to the latter what ordinary magnets are to Magneto-helices when the current is passing through them.'[51]

If matter was modified by magnetic force, Faraday reasoned, it probably played a role in the communication of magnetic force. He promptly introduced the word 'dimagnetic' (later to become 'diamagnetic' under Whewell's advice) in analogy with 'dielectric,' and speculated that 'magnets act[ed] by intervening particles,' as electrified bodies already did. At the same time, he emphasized the progress he had made in correlating forces of a different kind. As he explained in his publication:

> I have long held an opinion, almost amounting to conviction, in common I believe with many other lovers of natural knowledge, that the various forms under which the forces of matter are made manifest have one common origin; or, in other words, are so directly related and mutually dependent, that they are convertible, as it were, one into another, and possess equivalents of power in their action.

[51] *FD* 4: #7569 (18 September 1845); *FER* 3, series 19: #2200; *FD* 4: #7688 (26 September 1845).

His new discovery confirmed this expectation: 'I have at last succeeded,' Faraday exclaimed, 'in *illuminating a magnetic curve* or *line of force* and in *magnetising a ray of light*.'[52]

3.4.2 Touching magnetic curves

Faraday's more immediate aim was to prove the role of matter in the communication of magnetic action. A few days after his discovery of the effect on light, he tried to show a magnetic condition of the magnetized heavy glass by means of a compass needle. This did not work. In early November 1845, he suspended a bar of heavy glass at the end of a silk thread between the jaws of a new, powerful electromagnet. Upon turning on the current, the bar immediately assumed an equatorial orientation, at a right angle to that taken by an iron bar. Faraday commented: 'Thus touching dimagnetics by magnetic curves and observing a property quite independent of light.'[53]

Faraday then tried other substances. Good conductors like copper displayed complex behavior because of the currents induced during their motion. Most non-magnetic substances, however, behaved like the heavy glass, though not with the same intensity. A poorly conducting metal, bismuth, gave the best effect. Faraday also expected the human body to be diamagnetic: 'If a man could be in the Magnetic field, like Mahomet's coffin, he would turn until across the Magnetic line, provided he was not magnetic.' This is the first known occurrence of the phrase 'magnetic field,' meaning the space between the poles. Faraday may have introduced it here because of the anthropomorphic context: men or prophets explore fields. His growing fondness for the word probably resulted from his awareness that diamagnetic actions were not directly given by the usual magnetic curves.[54]

On 10 November, Faraday 'examined the Magnetic field by the bar of bismuth.' He found that the bismuth tended to move from stronger to weaker points of magnetization. Surprisingly, the direction of the magnetic force played no role here: 'There is no apparently dual character in the force—is an unique phenomenon as to its kind.' This action indicated 'a new set of magnetic curves.' Faraday soon regarded the earlier orientation effect as deriving from the latter action: when the bismuth bar, in the middle of the field, made an angle with the magnetic axis, its ends were

[52] *FD* 4: #7576 (18 September 1845); Whewell to Faraday, 10 December 1845, *SCMF* 2; *FER* 3, series 19: #2146; *FD* 4: #7705 (26 September), #7718 (30 September). Here 'force' is to be understood in the broad, qualitative sense of the cause of physical actions; 'power' has the ordinary meaning of the intensity of the effected transformations. On the correlation of forces, cf. Gooding 1980.

[53] *FD* 4: ##7691–2 (26 September 1845), #7871 (20 October); #7902 (4 November 1845). Cf. Williams 1965: 392–4; Gooding 1981: 236–237. Effects of magnets on bismuth and other non ferromagnetic bodies had been noted long ago, but never been the object of a systematic study: cf. Williams 1965: 393.

[54] *FD* 4: ##7999–8078 (8 November 1845); #8014 (8 November), also #8085 (10 November): 'between the great poles, i.e. in the magnetic field.'

repelled by the axis on which the magnetic intensity was the highest, until the equatorial position was reached.[55]

On 12 November, Faraday imagined two possible explanations of the diamagnetic orientations and repulsions. In the first explanation, which I call Amperean, he invoked the formation of reverse Amperean currents: 'Can there be formation in Bismuth of currents in the *contrary* direction?' In the second explanation, which I call differential, the cause would be a different conductive power of the air and the diamagnetic body for magnetic action:[56]

The Bismuth goes *from strong* to *weak* points of magnetic action. This may be because it is deficient in the inductive force or action, and so is displaced by matter having stronger powers, giving way to the latter. Just as in Electrical induction the best conductors, or bodies best fitted to carry on the action, are drawn into the vicinity of the inducing bodies or into their line of action.

The latter interpretation carried on the analogy between magnetic and electric induction and the view that magnetism was an action between contiguous particles. The Amperean interpretation did not connect as well to Faraday's preconceptions, and seemed to him to have an unwanted consequence: that electromagnetically induced currents in bismuth and copper would have opposite directions. Yet, in print Faraday only mentioned the Amperean explanation:[57]

Theoretically, an explanation of the movements of the diamagnetic bodies [. . .] may be offered in the supposition that magnetic induction causes in them a contrary state to that which it produces in magnetic matter [. . .]. Upon Ampère's theory, this view would be equivalent to the supposition, that as currents are induced in iron and magnetics parallel to those existing in the inducing magnet or battery wire; so in bismuth, heavy glass and diamagnetic bodies, the currents induced are in the contrary direction.

Faraday could not endorse the differential explanation because he failed to verify a clear consequence of it: that the repulsion of a diamagnetic body by a magnet pole should depend on the pressure of the surrounding air, and turn into an attraction in the case of vacuum. One could still imagine, as Edmond Becquerel did, that a magnetic ether played the role of the air. But this option could not satisfy Faraday, who

[55] *FD* 4: #8108 (10 November 1845); #8119 (10 November): 'Its endeavour [the bismuth's] is in fact not to go along or across the curves exclusively—but to get out of the curves going from stronger to weaker points of magnetic action'; #8137 (12 November); #8121 (10 November); *FER* 3, series 20 (December 1845): #2269. Cf. Gooding 1981: 239–43.

[56] *FD* 4: #8138, ##8144–5 (12 November 1845). By 'the best conductors' of the inductive action Faraday probably meant metallic conductors (which do not sustain polarization, but transfer it most efficiently: *FER* 1: #1566, and Faraday to Riess, 19 November 1855, *SCMF* 2): as was well known, an uncharged conducting ball goes to the regions of stronger electric action. A dielectric with high specific inductive capacity would have offered a better analogy, but the effect was too small to be observed.

[57] *FD* 4: #8141 (12 November 1845): 'Would a *Bismuth wire* or rod, carried across the magnetic curves, give a current in the same direction as a wire of copper or a contrary current?' Faraday answered negatively in ##8425–6 (26 November 1845) (however, he saw that microscopic currents could perhaps behave differently from macroscopic ones: *FER* 3: #2431); *FER* 3, series 21: ##2429–30.

generally opposed imponderable fluids and ethers. Instead he reluctantly admitted two sorts of magnetic bodies, with opposite inductive properties: 'I incline, by my view of induction through particles, to think that all bodies are in *one* magnetic list—but the facts as yet rather sustain the view of *two*.'[58]

3.4.3 The 'magnecrystallic' effect

Faraday's new findings immediately attracted the attention of German investigators. In Bonn, Julius Plücker studied the magnetic behavior of birefracting crystals, and detected 'a repulsion of the optical axis' by the poles of a magnet. In August 1848 Faraday meticulously studied this effect with Plücker's collaboration. In September, he explored the magnetic field with crystalline bismuth. Besides the repulsions observed on amorphous bismuth, he found a new orientation effect depending on crystal structure. In a uniform magnetic field, for which the usual repulsions do not exist, a well-defined axis of the crystal, 'the magnecrystallic axis,' positioned itself in a direction parallel to the magnetic curves. In a heterogenous field, a cubic crystal still showed the direction of the magnetic curves, because the cubic shape prevented the equatorial orientation effect. This action, Faraday commented, was 'an important indicator of the direction of the lines of force in a magnetic field,' because the bismuth crystal, unlike a compass needle, did not perturb the lines.[59]

Regarding the nature of the magnecrystallic effect, Faraday first condidered the possibility that a crystal would be less apt for diamagnetic induction along the magnecrystallic axis than along other directions, in the spirit of the Amperean interpretation of diamagnetism. If this were true, the repulsion of the crystal by a magnetic pole would be weaker when the axis pointed toward the pole. With a carefully designed bifilar torsion balance Faraday tested this difference, but could not find it. At that point he could 'not resist throwing forth another view of these phaenomena,' in harmony with his earlier differential interpretation of diamagnetism: the lines of magnetic force could 'pass more freely' in the direction of the magnecrystallic axis, just as light rays traveled faster (or slower) along the optical axis of a crystal. Then the equilibrium position of the crystal would be the position of 'least resistance' to the passage of the lines of force.[60]

[58] *FD* 4: #8257 (15 November 1845): 'What ought a vacuum to do? This is important as regards air, gases and indeed the whole subject'; #8262 (15 November): 'If air be rarefied, ought not different bodies suspended in it to set round in succession into the axial position—Water, Heavy glass, Bismuth, etc.?' [the air becoming successively less magnetic than water, heavy glass, and bismuth]; #8362–78 (22 November): failure to detect any effect of rarefaction; E. Becquerel 1846a, 1846b, 1849, 1850; *FD* 4: #8514 (quote, 6 December 1845). Cf. Tyndall 1870: xiii; Gooding 1981: 249–51. Becquerel independently developed the idea of a differential action, by analogy with Archimedes' push. In his view an ethereal medium was responsible for the magnetic properties of vacuum. Cf. Williams 1965: 420–22.

[59] *FD* 5: ##9378–465 (16 August–1 September 1848); *FER* 3, series 12: #2592–613; *FD* 5: #9467, #9475, #9494 (2 September 1848); *FER* 3: #2479, #2546. Faraday's intense activity in this field was largely motivated by his hope to obtain informations on molecular forces. Cf. Williams 1965: 417.

[60] *FD* 5: ##9920–25 (24 October 1848), rougher try in ##9855–6 (13 October); *FER* 3, series 22: #2551, #2552, #2588, #2591 (quote).

FIG. 3.11. Faraday's replication of Weber's experiment for testing the magnetic polarization of a bismuth block (B) in the field of a horse-shoe magnet (NS), with a compass needle n and a compensatory bar-magnet *S'* (*FD* 5: 153).

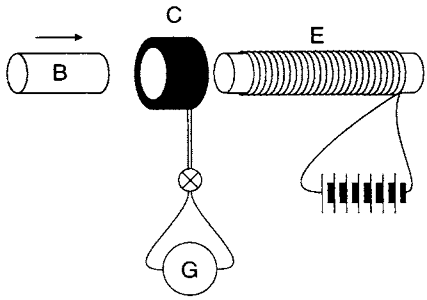

FIG. 3.12. Weber's device for showing electromagnetic induction by a variable diamagnetic polarization.

3.4.4 Weber's diamagnets

Meanwhile, Wilhelm Weber had published a double experimental proof of the interpretation of diamagnetism in terms of reversed Amperean currents. Diamagnetic actions were very small, and Weber deplored that he could not yet study them quantitatively. However, he could use Gaussian techniques to provide direct proofs of diamagnetic polarity. His first experiment aimed at showing the action of diamagnetic polarization on a magnetic needle. In the device represented in Fig. 3.11, the action of the horseshoe magnet on the needle n is exactly balanced by the action of the bar magnet until a block of bismuth B is brought between the jaws of the magnet. With the Gaussian method of the mirror and telescope, Weber detected a motion of the needle in the direction opposite to that which an iron block would give. Weber's second experiment (Fig. 3.12) concerned the electromagnetic induction produced by a variable diamagnetic polarization. The cylinder of bismuth B is periodically thrust into the hollow coil C, placed at the end of a powerful electromagnet E. The induced current is measured by means of a galvanometer connected to the coil by a commutator K compensating the sign changes of the induced currents.[61]

The positive result of these experiments was not the only source of Weber's belief in diamagnetic polarity. He also managed to integrate the phenomenon in his general theory of electricity. Faraday had already suggested that the polarity, if it existed,

[61] Weber 1848b.

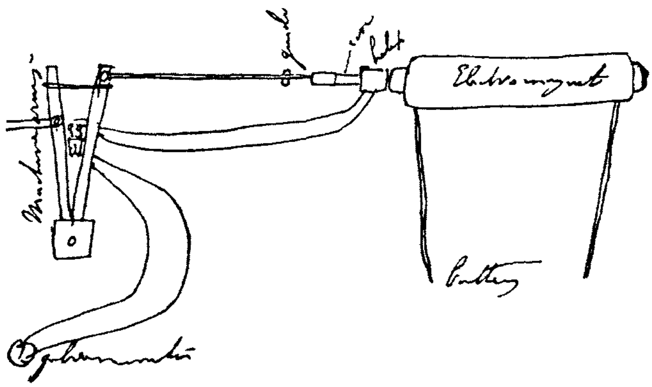

FIG. 3.13. Faraday's device for testing currents induced by the motion of various cores in the field of an electromagnet (*FD* 5: 204). Clockwise, the handwritten words read: galvanometer, machine arms, guide, core, helix, electromagnet, battery.

could be due to the temporary formation of molecular currents running opposite to ordinary Amperean currents. Weber saw that the laws of electromagnetic induction implied the formation of such currents, if only the diamagnetic body contained microscopic circular channels of zero resistance for the electric fluids. In a resistance-less circuit the electromotive force is proportional to the variation of the current (with a coefficient depending on the inertia of the electric fluids and on the self-inductance). Therefore, any variation of the magnetic force at the circuit implies a proportional variation of current. By Lenz's law, this variation must be such that it counteracts the external magnetic force, in agreement with the negative sign of diamagnetic polarization.[62]

Such microsopic reasoning on electric fluids was so alien to Faraday's views that he completely ignored Weber's theory. Faraday did, however, spend several months on Weber's experiments. In March 1849 he repeated the first experiment, but found the bismuth to be '*nil* in its action.' From September to December he worked on Weber's other experiment, with practically the same device, except that he automatized the periodic thrust of the bismuth cylinder (Fig. 3.13). He found it extremely difficult to avoid the communication of vibrations from the mechanism moving the bismuth to the coil and electromagnet. He also struggled to discriminate between currents induced in the bismuth mass and true diamagnetic polarity: 'Astonishing how great the precautions that are needed in these delicate experiments. Patience. Patience,' can be read in his diary. At the term of this painstaking excursion into the territory of Gaussian precision, Faraday concluded that the effect observed by Weber must have been parasitic electromagnetic induction.[63]

[62] Weber 1848b: 267.

[63] *FD* 5: #10050, #10691 (Faraday wrongly attributed this device to Ferdinand Reich); *ibid.*: ##10330–10690, quote from #10462 (16 November 1849). Faraday recycled the oscillating mechanism from earlier experiments on the relation between gravity and electricity.

3.4.5 Conducted lines of force

After this episode, Faraday discarded the possibility of diamagnetic polarity. He found more evidence against the German view of diamagnetism in subsequent experiments on gases. In the fall of 1847 he had learned from Francesco Zantedeschi that flames were repelled by the poles of a strong magnet. Faraday explained this effect by a temperature-dependent (dia)magnetism of gases, and proved that all current gases except oxygen were diamagnetic with respect to the air by observing their motion after being freed in the air between the poles of an electromagnet.[64]

If this motion was due to attractions or repulsions as the polarity interpretation supposed, Faraday reasoned, then a single gas should be more dense (if paramagnetic) or less dense (if diamagnetic) in the more intense parts of the magnetic field. In October 1849 with an optical method and in January 1850 with a closed vessel and a capillary manometer, he proved the absence of such compressions. To which he commented: 'Is then the effect an effect not of attraction or repulsion but a differential effect of another kind between the two bodies which are free to go to the pole?' In April he definitely adopted the idea that 'the conductor [of the magnetic action] which can conduct the most will of necessity be drawn into the place of most intense action.' A great advantage of this view was that it explained why gases were not compressed (or expanded) near magnetic poles, despite their para- or diamagnetic behavior with respect to one another.[65]

In his 26th series, of October 1850, Faraday developed the notion of conducting power 'as a general expression for the capability which bodies may possess of affecting the transmission of magnetic force, implying nothing as to how the process of conduction is carried on.' By definition, a diamagnetic body conducted less and a paramagnetic conducted more than vacuum, and some crystals had different conductivities in different directions. Then all known repulsions, attractions, and orientations resulted from the rule of least resistance to the passage of the lines of force. In sum, Faraday returned to the differential interpretation of diamagnetism of November 1845, but now avoided its main defect, its reliance on an ether, by divorcing conduction from action between contiguous particles.[66]

Faraday multiplied diagrams representing the disturbance of lines of force by dia- and paramagnetic bodies, insisting on the deformations of the lines *outside* the bodies (Fig. 3.14). These deformations, together with the law of least resistance,

[64] Zantedeschi 1847; *FD* 5: ##9066–291 (23 October–18 November 1847); Faraday, 'On the diamagnetic conditions of flame and gases,' *FER* 3: 467–93. Cf. Williams 1965: 396–9. Much later, in July 1850, Faraday was able to compare gases with vacuum and to prove that oxygen was paramagnetic: *FD* 5: ##10896–967.

[65] Plateau to Faraday, 25 March 1849, in *FD* 5: 196–8 (Plateau suggested the optical method); *FD* 5: ##10277–301 (10–5 October 1849): 'No sensible *condensation* or *expansion* of the air or gases in the *intense magnetic field*'; ##10714–43 (7–21 January 1850) and related experiments in February and March; #10744 (21 January); #10793 (4 April 1850). The argument is dubious: if a principle of least resistance to the passage of the lines of force rules the equilibrium of the gas, then a diamagnetic gas should be more expanded where the force is stronger. This effect exists (as today's physicists know) but is too small to be detectable by Faraday's devices.

[66] *FER* 3, series 16: #2797. Cf. Gooding 1981: 268–75.

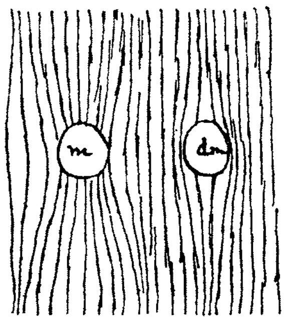

FIG. 3.14. Faraday's drawing of the lines of force around a (para)magnetic (m) body and a diamagnetic (dm) body (*FD* 5: 320).

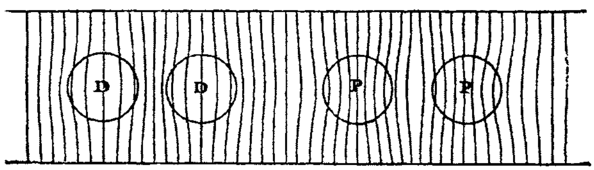

FIG. 3.15. Lines of force deformed by two neighboring diamagnetic (D) or paramagnetic (P) spheres (*FER* 3: 212).

implied new effects, for example the orientation of an oblong piece of bismuth in a homogenous magnetic field, as Thomson had earlier explained to Faraday. Such effects were too small to be observed by means of bismuth in air. Faraday obtained them with bismuth in a strong solution of iron protosulfate, thanks to the higher difference of conductive power.[67]

Effects usually attributed to magnetic polarity and previously denied by Faraday resulted from the rules of conduction. For example, on Fig. 3.15 the two diamagnetic spheres repel each other because the lines of force are compressed between the two spheres, and because diamagnetic bodies, being worse conductors than vacuum, tend to move away from fields of higher intensity. The paramagnetic spheres also repel each other, by a dual mechanism: the lines of force are further apart in the space between the spheres, and paramagnetic bodies, being better conductors than vacuum, tend to move away from places of lower intensity. In this context Faraday still used the world 'polarity,' but in a sense different from Weber's. He only meant the asymmetry of the disturbance of the lines of force when entering and leaving a body. Accordingly, he rejected Weber's idea that a diamagnetic body behaved like a paramagnetic body turned end for end without change of its magnetic state.[68]

[67] *FD* 5: #10832 (8 April 1850); #10921, #10922 (20 July 1850) *FER* 3: #2807, #2810, #2821, #2812; Thomson to Faraday, 19 June 1849, *SCMF* 2.

[68] *FER* 3, series 26: #2815, #2816, #2831, #2820. However, Faraday admitted polarization in Weber's sense for ferromagnetic bodies (#2833).

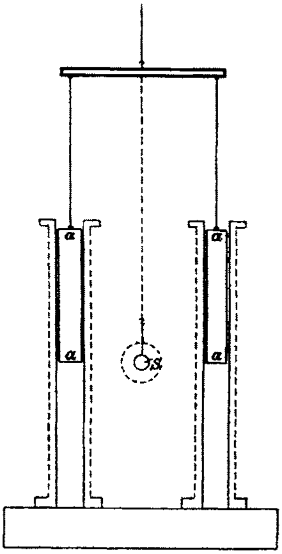

FIG. 3.16. Weber's apparatus for measuring the magnetic moment of 'diamagnets' (Weber 1852: 491).

3.4.6 *Weber's revenge*

In 1852 Weber published a new instalment of his *Maassbestimmungen*, the subject of which was diamagnetism. Presumably hurt by Faraday's criticism of his previous experiments, he claimed that the account he had published was only partial and provisional and that in fact he had taken into account the currents induced in the bismuth mass. He attributed Faraday's failure to duplicate his results to an inferior technique of magnetometric and galvanometric measurement. But he admitted that his previous proofs of diamagnetic polarity were insufficient. 'In order to lead to sure results,' he declared, 'the observation of so weak actions needs *quantitative control*, something that has completely lacked so far.' The main purpose of the new *Maassbestimmungen* was to provide quantitative versions of the two previous experiments on diamagnetic polarity.[69]

In order to measure the magnetic moment of a uniformly polarized bismuth cylinder, he imagined the device of Fig. 3.16. Two identical bismuth cylinders aa can slide within two long, parallel solenoids fed by the same constant current in opposite directions. The south pole S of the suspended magnetic needle of a magnetometer lies at exactly equal distances from the two solenoids. Because of this symmetry, the solenoids have no magnetic action on the needle. Any spurious dissymmetry is compensated by a distant coil (not represented on the figure). Then

[69] Weber 1852: 534–5, 488. Faraday used a ready-made galvanometer by Ruhmkorff (*FER* 3: #2651), whereas Weber used the Gaussian technique.

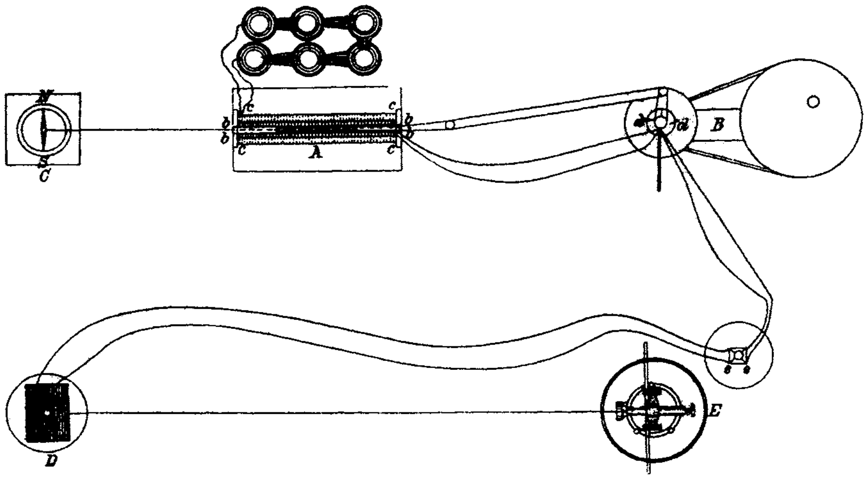

FIG. 3.17. Weber's arrangement for measuring the currents induced by a moving diamagnet (Weber 1852: 508).

only the 'diamagnets' act on the needle. Within the solenoids the magnetic force is uniform and vertical, so that vertical displacements of the bismuth bars cannot induce any currents in their mass. Such displacements occurred in Weber's measurements, because, according to an old Gaussian trick, he multiplied the action of the diamagnets on the needle by swinging the bismuth cylinders synchronically with the oscillations of the magnetometric needle. With four collaborators, including Johann Listing and Bernhard Riemann, Weber determined the magnetic moment of the bismuth bars for a given solenoid current. This moment turned out very small and negative, as was expected.[70]

In the improved version of his second experiment, Weber again exploited the uniformity of the magnetic force within a solenoid. In part A of his drawing (Fig. 3.17), one of the bismuth cylinders of the previous experiment, aa, oscillates within the solenoid cccc fed by the six-cell battery. An induction coil bbbb surrounds the moving cylinder. This coil is made of two oppositely connected halves (not shown on the figure), so that the inductive actions of the two poles of the diamagnet have the same sign. The mechanism in B, borrowed from Faraday, commands the alternating motion of the bismuth bar as well as the commutator dd. The current from the induction coil runs through this commutator and another manual commutator ee, and reaches the galvanometer D, observed with the telescope E. Lastly, the compass SN is used to measure the current in the solenoid cccc. With this sophisticated device, Weber obtained the expected induction from the moving diamagnet, in quantitative agreement with the magnetic moment measured in the first experiment.[71]

[70] Weber 1852: 489–505. [71] Weber 1852: 506–31.

Having thus proved the existence of diamagnetic polarity, Weber turned to its theory. He distinguished four kinds of explanations of magnetism: Poisson's microscopic cells for the separation of the magnetic fluids, elementary pivoting magnets, pivoting Amperean currents, and molecular channels of zero-resistance for the electric fluids. The three first hypotheses led to a magnetic polarization in the direction of the impressed magnetic force. Therefore, they could not explain diamagnetism. The last assumption was the only one left. In his theoretical world of imponderable fluids acting at a distance and organized in microscopic structures, Weber believed that he had explored all possibilities. He therefore asserted the physical existence of undamped microscopic currents. Diamagnetism corresponded to the induction of such currents during the application of a magnetic force, and (ferro)magnetism to the orientation of preexisting currents of this kind. Weber now denied the existence of the magnetic fluids, although he had originally preferred them to Amperean currents.[72]

3.4.7 Indifference

Weber's impressive memoir convinced most experts in the field, even Faraday's friend John Tyndall, who perfected Weber's experiments with Weber's own help. Yet Faraday remained undisturbed. In 1854 he included Weber's work in a long list of 'magnetic hypotheses,' with ambiguous praise:

> Weber stands eminent as a profound mathematician who has confirmed Ampère's investigations as far as they proceeded, and who has made an addition to his hypothetical views [the microscopic induced currents][. . .]It would seem that the great variety of these hypotheses and their rapid succession was rather a proof of weakness in this department of physical knowledge.

Faraday judged that Weber's idea of polar diamagnetism 'involve[d], if not magnetic impossibility, at least great contradiction and much confusion.' If a magnet induced a reverse polarization of the particles of a diamagnetic body, he reasoned, then a reverse induction should also occur from particle to particle and prevent global polarization.[73]

Faraday's antagonism depended on his unwillingness, or incapacity, to conceive secondary sources, and the superposition of their action with primary sources. For him, the disturbance of the field produced by a diamagnetic body was primitive, it was not to be deduced from the formation of a new magnetic state of the body. Faraday's polarity was about the disturbance of the lines of force. Weber's polarity was about secondary sources: 'By *magnetic or diamagnetic polarity* of a body,' Weber wrote, 'I understand a state of this body through which it exerts on other bodies actions that are so constituted that they can be completely explained by *an*

[72] Weber 1852: 538–46, 557–60.

[73] Tyndall 1856; *FER* 3, 'On magnetic hypotheses' (1854): 525–6; *FER* 3, 'On some points of magnetic philosophy' (1855): #3309 (quote), 3310–2. Cf. Tyndall 1870: xvii–xviii. Faraday had two notable supporters: Carlo Matteucci in Italy, and Fabian von Feilitzsch in Germany (cf. Tyndall 1870: 156–8).

ideal distribution of magnetic fluids.' Like Gauss's representation of the Earth's magnetism by a superficial distribution of magnetic fluids, this definition was ontologically neutral, and it gave to the question of diamagnetic polarity a clear-cut empirical meaning. Faraday, being a simple man who 'felt [his] way by facts closely placed together,' could see nothing there but a mathematician's perversion.[74]

The communication breakdown was analogous to what Faraday had already experienced in the chapter of electric induction. Ambiguous concepts were responsible: 'interaction' in one case, 'polarity' in the other. Faraday defined these concepts in terms of the distribution of power in the field, whereas other investigators thought in terms of interacting states of distant objects or sources. Consequently, utterly different interpretations could accompany the same experimental fact. After friendly chats on diamagnetism with Tyndall, Faraday observed: 'I differ from Tyndall a good deal in phrases, but when I talk with him I do not find that we differ in facts. The phrase *polarity* in its present state is a great mystifier.'[75]

3.4.8 Sharpening the lines of force

Since his discovery of 'the Faraday effect,' Faraday had made more and more frequent use of the magnetic lines of force and had become more and more convinced of their physical reality. He could 'illuminate' the lines of force, 'touch' them with diamagnetic bodies, and channel them through 'conductors.' In a study on atmospheric magnetism published in 1850, he opened himself on this speculative matter:[76]

> External to the magnet these concentrations which are named poles may be considered as connected by what are called magnetic curves, or lines of magnetic force, existing in the space around. These phrases have a high meaning, and represent the ideality of magnetism. They imply not merely the directions of force, which are made manifest when a little magnet, or a crystal or other subject of magnetic action is placed amongst them, but these lines of power which connect and sustain the polarities, and exist as much when there is no magnetic needle or crystal than as when there is; having an independent existence analogous to (though very different in nature from) a ray of light or heat, which, though it be present in a given space, and even occupies time in its transmission, is absolutely insensible to us by any means whilst it remains a ray, and is only made known through its effects where it ceases to exist.

Faraday devoted his two next series to the lines of magnetic force. There he refrained from ontological commitment: 'I desire to restrict the meaning of the term *lines of force*, so that it shall imply no more than the condition of the force in any given space, as to strength and direction: and not to include (at present) any idea of the nature or the physical cause of the phenomena.' His aim was to give a more precise definition of the lines and to offer a systematic account of their use in the representation of various phenomena. He gave three definitions of the lines: through the orientation of a compass needle, through the currents induced in a moving wire,

[74] Weber 1852: 486 (his emphasis); Faraday to Ampère, 3 September 1822, *CMF* 1.
[75] Faraday to Matteucci, 2 November 55, *SCMF* 2. [76] *FER* 3: 323.

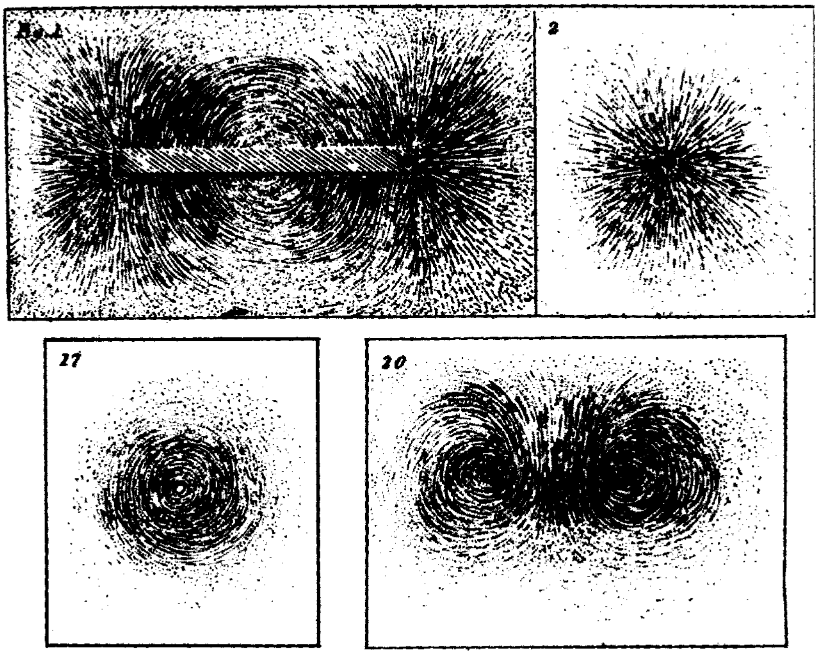

FIG. 3.18. Iron-filing patterns around magnets and conducting wires (*FER* 3: plate 3).

and through the magnecrystallic effect of bismuth. He also provided a plate of illustrations by iron filings (Fig. 3.18). The convergence of the three definitions, and the suggestiveness of the iron filing patterns, must have contributed of his conviction that the lines of force were 'true to nature.'[77]

In previous magnetic studies, Faraday had often made quantitative use of the lines of force: he had implicitly assumed that parallel equally spaced lines represented a homogenous field, and he measured the intensity of the force at a given place by the concentration of the lines. He now made the convention entirely explicit, and called 'unit lines of force' the lines that complied with it. In these terms he could give a quantitative formulation of his electromagnetic induction law: 'The quantity of electricity thrown into a current is directly as the amount of curves intersected.' The unit lines of force also reflected a basic property of the magnetic field, which Maxwell later called the conservation of the magnetic flux.[78]

Among the various ways to define the lines of force, Faraday favored the moving wire, because it could be used to test the conservation of flux inside and outside matter. He experimented with a bar magnet that could rotate around its axis, and

[77] *FER* 3, series 28: #3075 (quotes); series 29: ##3234–40 (filings).

[78] *FER* 3, series 28: #3122, #3115, #3073: 'The sum of the power contained in any one section of a given portion of the lines is exactly equal to the sum of power in any other section of the same lines,' or $\nabla \cdot \mathbf{B} = 0$ according to Maxwell.

with a wire loop connected to a galvanometer, part of which went through the mass of the magnet (with proper insulation). For various forms of the loop, he turned the magnet or the external wire. Taking for granted that the (external) lines of force did not follow the motion of the magnet, he drew the following conclusions: the lines of force traversed the substance of the magnet, they were always closed, and the magnetic flux was conserved upon entering or leaving magnetic matter.[79]

3.4.9 The physical lines of force

In the following year Faraday relaxed his empiricist reserve, and published 'On the physical character of the lines of magnetic force.' Speculations on the deeper nature of forces, he now argued, were 'wonderful aids in the hands of the experimentalist and mathematician.' In his discussion he compared the four known kinds of power: gravitation, light, electricity, and magnetism. All could be represented by lines of force, but these lines could have different meanings in each case. The lines of force for gravitation did not have to be physical, because they were straight, uninfluenced by interposed matter, and instantly acting. In contrast, those for light had a wealth of reasons to be physical: they could be emitted, curved, absorbed, and polarized, and they took time to propagate. In the case of electricity, the facts of decomposition, inductive capacity, and induction in curved lines proved the physical character of the lines: they represented an action between contiguous particles, at least when matter was present.[80]

For magnetism, the situation was unfortunately less clear. Faraday admitted that he had no strict proof of the physical existence of the lines. For sure, they were affected by the presence of matter. But the effects of intervening matter were opposite for diamagnetic and paramagnetic bodies, so that the lines of force could not possibly represent a unique kind of action between contiguous particles of matter.[81] A way out of the difficulty was to introduce, as Edmond Becquerel had already done, a polarizable ether with a polarizability intermediate between that of diamagnetic and paramagnetic bodies. Faraday himself mentioned the possibility 'that all conduction of magnetic force [was] carried on by circular electric currents round the

[79] *FER* 3: ##3090–121.

[80] *FER* 3 (June 1852): #3244; ##3245–51. By gravitational lines of force Faraday could not mean the lines tangent to the net force acting on a point mass (those would be curved in general). What he seems to have meant is the lines representing force emanating *from a given point mass* in the presence of other bodies, that is, the net force minus the forces that the other bodies would exert *if the point mass was not there*. These lines are straight, because the gravitational force exerted by the other bodies does not depend on the existence of the point mass. In the case of electrostatics, the similarly defined lines of force are generally curved, because the force exerted by other bodies (typically neutral polarizable bodies) depends on the existence of the given point charge. This interpretation fits Faraday's statement on #3245: 'One particle gravitating toward another particle has exactly the same amount of force in the same direction, whether it gravitates to that one alone or towards myriads of other like particles, exerting in the latter case upon each one of them a force equal to that it can exert upon the single one when alone: the result of course can combine, but the direction and amount of force between any two given particles remains unchanged.'

[81] Hence Faraday's difficulty to decide between one or two 'magnetic lists': *FD* 4: ##8398–9 (22 November 1845), #8514 (6 December 1845), *FD* 5: ##10806–7 (4 April 1850). Cf. Gooding 1981: 249–53.

line of magnetic force in the whole of its course, and in that case that they must exist in a vacuum itself.' However, such views were too crudely mechanistic to please him. He preferred a direct transference of the magnetic lines of force through pure space, as he already assumed for gravitation and for electric induction across the space between dielectric particles.[82]

As for the curving of magnetic lines of force, Faraday deplored the lack of strong evidence. A proof analogous to the one he had given in the electric case would have required the existence of the magnetic counterpart of electric conductors.[83] Iron filings or compass needles did not necessarily prove the curvature because their presence could alter the distribution of forces. From the beginning of his research Faraday had used entities defined by virtual experiments. Yet he did not mistake counterfactual definitions for proofs of physical existence.[84]

Having exhausted the possibilities for a clear-cut demonstration of the physical existence of the lines of magnetic force, Faraday turned to more speculative and somewhat obscure arguments. In some of them, he insisted on the dual character of magnetic power: if a pole could not be created without the simultaneous creation of an opposite pole, there had to be some kind of physical link between the two poles. Specifically, he related the existence of the lines of force of a magnet to that of inner polarization (proved, for example, by broken magnets): one could not deny an external relation between the poles of a magnet, he suggested, without denying their internal relation.[85]

In another ingenious argument he related the behavior of the magnetic lines of force to the existence of electrodynamic forces. On the one hand, all mechanical actions among magnets could be reduced to a tension along the lines of magnetic force and a mutual repulsion of these lines (suggested by the shape of the lines and by the analogy with magnetic needles placed side by side). On the other hand, electric currents tended to elongate themselves, and parallel currents placed side by side attracted each other. Using the axis–loop relation between lines of current and lines of force, Faraday showed the perfect agreement between the two kinds of forces. In Fig. 3.19(a), the repulsion between the circular lines of force C and C′ tends to elongate the current *i*, while on Fig. 3.19(b), the tension of the line of force L implies an attraction of the current loops *i* and *i*′. The magnetic lines of force thus seemed 'to have a physical existence correspondent to that of their analogue, the electric lines.'[86]

[82] *FD* 5: #10834 (8 April 1850); *FER* 3: #3075; *FER* 1, series 25: #2787, #2788; *FD* 5: #10374: 'Is magnetic action *across space*, through air, water, a vacuum, etc., but between contiguous particles in iron, nickel, etc.?', #10837: 'Is it not probable and most likely that lines of Magnetic force can be transferred across space in the manner of Gravitating and Static Electricity force, and without these circular currents (or their equivalents), which are assumed to exist in iron when it is in the magnetic field?'; *FER* 3: series 21, ##2445–6; *ibid.*, #3258: 'Physical lines of force.'

[83] That is, bodies unable to sustain magnetic induction: nothing to do with the conduction of the magnetic lines of force.

[84] *FER* 3: #3254. [85] *FER* 3: ##3257–64, ##3282–98.

[86] *FER* 3: ##3264–69. This explanation of the attraction of two currents is similar to the one Faraday gave in 1821 in terms of magnetic powers.

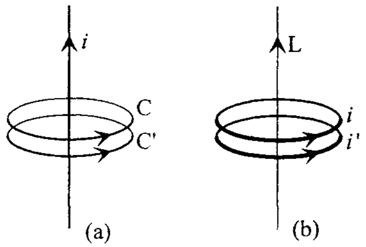

FIG. 3.19. Diagrams for the relation between electrodynamic forces and magnetic field stresses.

Faraday had no definite idea, however, of the physical condition represented by the magnetic lines of force. It could be a dynamic condition, in conformity with the dynamic nature of the electric current and with the hypothesis of Amperean currents. In this case, Faraday argued, there should be a magnetic equivalent to static electricity, and such a thing had never been observed. He therefore preferred to imagine a static condition, 'a state of tension (of the aether?)' that would provide for the long-sought electro-tonic state. As Faraday knew, in 1835 Whewell had proposed a dynamic interpretation of the tonic state as the momentum of a motion connected to the electric current. But Faraday held fast to his older intuition of a state of tension in the magnetic field.[87]

This view raised the question of the thing that was under a state of tension. It could not always be matter, since magnetism acted across a vacuum. It could perhaps be the optical ether. However, Faraday disliked the idea of assimilating empty space with a subtler kind of matter. In most of his works, he maintained a sharp distinction between vacuum and matter. Typical is his 1850 warning against a confusion between the magnetic properties of vacuum and matter: 'To confuse them together would be to confound space with matter, and to trouble all the conceptions by which we endeavour to understand and work out progressively a clearer view of the mode of action and the laws of natural forces.'[88]

Then nothing was under a state of tension. Tension, force, and power existed by themselves. So said Faraday in his occasional dynamistic speculations. In that of 1844 he reduced matter to concentrations of power. In his 'Thoughts on ray vibrations' of 1846, he proposed a natural extension of this view: the subtler kind of matter called ether did not exist; there were only gravitational, electric, and magnetic lines of force crossing empty space. Then light was a transverse vibration of the lines of force. In this view the physical character of the lines of force became a necessity,

[87] *FER* 3: #3269; Whewell to Faraday, 25 April 1834, Faraday to Whewell, 3 May 1834, *CMF* 1. On the latter exchange, cf. Anderson 1994. In the context of the magneto-optical effect, Faraday had imagined circular currents in the magnetized glass (*FD* 4: #7569, 18 September 1845); he did not publish the idea, because he was looking for a state of tension: *FER* 3: #2229.

[88] *FER* 3, series 25: #2787.

since matter and light were both derived from them. Without physical lines of force, nothing in the world would be physical.[89]

3.5 Thomson's potential

3.5.1 British reformers

In the mid-1840s, after many years of general neglect, Faraday's theoretical views attracted the attention of a young mathematical prodigy, William Thomson. This improbable encounter between two very different kinds of mind cannot be understood without first capturing some pecularities of Thomson's education.

Fourier's *Théorie analytique de la chaleur* was Thomson's first intellectual love and inspiration. During his student years at Glasgow and Cambridge, French mathematics and mathematical physics were generally considered best and most promising. Since the 1810s, progressive men like John Herschel, Charles Babbage, William Whewell, and George Airy had denounced the degeneration of the British tradition of mathematical physics. For Herschel, 'the last twenty years of the eighteenth century were not more remarkable for the triumphs of pure and applied matheamtics abroad, than for their decline, and, indeed, all but extinction at home.' The putative cause was a slavish following of Newton's methods. The proposed remedy was a thorough study of French classics.[90]

Initially, the British reformers adopted Laplace's notion of rigor. They abandoned the intuitive, geometrical conception of calculus inherited from Newton, and based mathematical analysis on algebraic definitions and manipulations. They defined derivatives in Lagrange's manner, as the successive coefficients in Taylor series. They solved differential equations by power series. For the mathematization of physics, they adopted Laplace's assumption of point-molecules of ponderable matter and imponderable fluids acting directly at a distance. This model lent itself to perfectly definite mathematical deductions: one just had to integrate the forces acting on a given molecule from all other molecules, to derive a differential equation from the resulting integral equation, and to find the solutions for given boundary solutions. Lastly, these solutions were compared with quantitative experiments.[91]

However, the British interest in Laplacian molecularism and algebraism soon declined. When they became available, Fourier's and Fresnel's memoirs on heat and light captured the attention. Being less speculative and more geometrical than Laplacian works, they pleased the pragmatic and illustrative inclinations of British natural philosophers. Fourier's theory of heat was especially attractive: it refrained from special assumptions on the nature of heat and the constitution of matter; its basic equations had a direct empirical meaning; and it made a central use of the

[89] Faraday 1844a. 1846.

[90] Herschel 1832 [illegible]9–31. Cf. [illegible] 1973; Smith and Wise 1989: 151–5; Crosland and Smith 1978; Grattan-Guinness 1985.

[91] Cf. Smith and Wise 1989: 151–5.

geometrical notion of flux across a surface. The first British works on optics and electricity that went truly beyond previous French theories, George Green's and James MacCullagh's, consciously adopted and perfected Fourier's methodology.[92]

Young William Thomson was in an excellent position to appreciate the virtues of French mathematics. His father James Thomson taught mathematics in Belfast's Academical Institution until 1830, and then at Glasgow University. In his lectures and in his numerous textbooks, he promoted French methods supplemented with geometrical illustrations and practical applications. So did his Glasgow colleagues John Pringle Nichol and William Meikleham, who instructed his son William in natural philosophy. The emphasis on geometry and sensible motions was particularly strong in Scotland. Nichol taught that 'the quality of form is the simplest of all the qualities of matter, and hence geometry, which treats of it, stands at the head of Natural Philosophy.' Like Dr Thomson, he followed another Scottish principle, the unity of art and science, and a latitudinarian value, anti-dogmatism. Both men fought metaphysics and abstraction, as William Thomson would do for the rest of his life.[93]

Among French authors, Fourier best incarnated Scottish values. Nichol's praise of the *Théorie analytique de la chaleur* was so high that William Thomson absorbed the thick volume in two weeks during May 1840, at age 16. Within a few months he was able to correct local misinterpretations of this 'great mathematical poem' and to complete some of its proofs. Most originally, he used an analogy between electrostatics and heat propagation to prove new electrostatic theorems. He sent a highly dense and concise account of this work to the *Cambridge Mathematical Journal* in September 1841, just before going up to Cambridge.[94]

3.5.2 Electrostatics and heat flow

As Thomson knew from Poisson, electrostatic forces in air derived from a function V that satisfied the same partial differential equation as that given by Fourier for the stationary temperature distribution in a homogenous solid: $\Delta V = 0$. However, Fourier treated the sources of heat as surface conditions, whereas Poisson included the sources of electricity in the differential equation. Thomson first extended the analogy by introducing point sources of heat. In this case, Fourier's equation gave a temperature proportional to the inverse distance from the source. Thomson superposed such sources on a surface S with the density σ to reach the following expression of the temperature θ:

$$\theta = \int \frac{\sigma \mathrm{d}S}{r}, \tag{3.1}$$

[92] Cf. Smith and Wise 1989: 155–68; Wise 1981a: 23–32. On Fourier's method, see also Dhombres and Robert 1998: Ch. 8.

[93] Thomson, 'Notebook of Natural Philosophy class, 1839–40,' quoted in Smith and Wise 1989: 210. Cf. *ibid.*: Chs. 1, 2; p. 40 (on Nichol).

[94] Thomson 1841a, 1841b; *TMPP* 3: 296 (poem); Thomson 1842. Cf. Smith and Wise 1989: 167, 203–4.

which is identical to the expression of V corresponding to the electric density σ.[95]

Thomson then imagined heat sources on a closed surface, their distribution being such that the surface was isothermal. He reasoned as follows. The temperature within the surface must be a constant, because if it were not, there would be a flux of heat around a small closed surface surrounding any internal temperature extremum, in contradiction with the absence of internal sources. Analogously, if a solid body has a surface charge such that the corresponding V is a constant on the surface, then V is a constant inside the body and the electric force vanishes there. Consequently, a sufficient condition of equilibrium for a conductor is that the electric force created by the surface charge should be perpendicular to the surface.[96]

Thomson now inverted the analogy. He proved, as Coulomb had done before him, that the electric force immediately outside the closed surface was normal to the surface and equal to 4π times the surface density. Consequently, the density of heat sources that sustain a constant temperature on the surface is equal to the heat flux across the surface divided by 4π. Now, the temperature outside any isothermal surface depends only on the temperature on the surface and on the heat flux at every point of the surface, as long as all sources are within or on the surface. This property is an obvious consequence of Fourier's view of the propagation of heat as an action between contiguous elements of volume. Its electrostatic counterpart is the far less obvious surface-replacement theorem: the electric force due to any distribution of electric charge is the same as the force due to a fictitious distribution of charge on a surface of constant V containing all the real charges, the surface density being equal to the electric force created on the surface by the real charges divided by 4π.[97]

Unknown to Thomson, the new theorems had already been published three times, by Green in 1828, by Gauss and by Chasles in 1839. Green and Gauss, like Thomson, ascribed a central role to the function V. However, their methods were purely analytical, based on partial integration and quadratic forms. Thomson's essential innovation was a new method for finding theorems by formal analogy between two physical theories. In his reasonings, he moved back and forth between two physical theories, transposing notions and theorems from one theory to the other. The starting point of one theory (Coulomb's law), became a new result of the other (the temperature distribution of point sources). An obvious consequence of the physical picture of one theory (Fourier's local transfer of heat) became an essential theorem of the other (the surface-replacement theorem).[98]

Thomson's analogy suggested that electrostatic action could perhaps be a contiguous action in the medium between sources, as heat propagation was. Thomson did not say so much, however, and he did not mention Faraday. For the electric

[95] Thomson 1842: #3, #4. This analogy contains what current textbooks call 'Gauss's theorem': the flux of the electric field across a closed surface must be equal to the total included charge, because the corresponding heat flux is equal to the heat provided by the included sources.

[96] Thomson 1842: #5.

[97] Thomson 1842: ##6–9. Cf. Wise 1981a: 33–9; Smith and Wise 1989: 205–12.

[98] Green 1828; Chasles 1839; Gauss 1839. Cf. Thomson 1845b: 17–18n. On Green, cf. Grattan-Guinness 1995.

density he used the name 'density of electrical matter,' which referred to the fluid conception, but also 'electric intensity' which echoed the view of one of his Glasgow professors that electricity was a *state* of bodies. This liberalness probably indicated a lack of commitment on the nature of electricity. In any case, Thomson avoided a physical exploitation of the analogy between heat and electricity. He did not even give a name to the counterpart of temperature, the function *V*.[99]

3.5.3 Discovering Faraday

Thomson's attitude changed somehow in 1843, as can be seen in his diary:

I have been sitting half asleep before the fire, for a long time thinking whether gravity and electrical attraction might not be the effect of the action of contiguous particles, communicated from one surface of [equal *V*] to another. In Cavendish's experiment, will the attraction of the balls depend at all on the intervening medium?

'Contiguous particles' and 'intervening medium' were Faraday's expressions; surfaces of equal *V* referred to Thomson's thermal analogy. Most likely, Thomson realized the physical implications of the analogy by reading Faraday. Yet he had little respect for Faraday's views. He found himself 'very much disgusted with [Faraday's] way of *speaking* of the phenomena, for this theory can be called nothing else.'[100]

Thomson changed his mind in Paris, when Liouville's account of the late Poisson's worries prompted him to reexamine Faraday's challenge of Coulomb's electrostatics. While developing a mathematical theory of the effect of dielectrics on electrostatic action, he gradually understood the consistency and precision of Faraday's ideas. Most strikingly, he discovered that Faraday's reasonings in terms of electric lines of force were similar to his own reasonings in terms of heat flow:[101]

All the views which Faraday has brought forward, and illustrated or demonstrated by experiment, lead to [my] method of establishing the mathematical theory, and, as far as the analysis is concerned, it would, in most *general* propositions, be even more simple, if possible, than that of Coulomb [. . .]. It is thus that Faraday arrives at some of the most important of the general theorems, which, from their nature, seemed destined never to be perceived except as mathematical truths.

For example, Thomson continued, Faraday knew that tubes of lines everywhere tangent to the electric force and connecting one conductor to another determined equal and opposite charges on the corresponding surface sections of the conductors (Fig. 3.20). For Faraday, the theorem was an immediate consequence of the definition of electric charge in terms of the surging or ending of lines of force. For

[99] Thomson 1842: #5, #7; Thomson, notes on Meikleham's lectures, 1839–40, quoted in Smith and Wise 1989, p. 210: 'Light heat elctricity magnetism, &c are termed imponderable. This is incorrect, as we know them, not as substances, but as states of bodies.' Cf. Smith and Wise 1989: 208–9.

[100] Thomson, Cambridge diary: 24 February, 16–17 March 1843, quoted in Smith and Wise 1989: 203, 213.

[101] Thomson 1845b: 29–30. Cf. Smith and Wise 1989: 213–8.

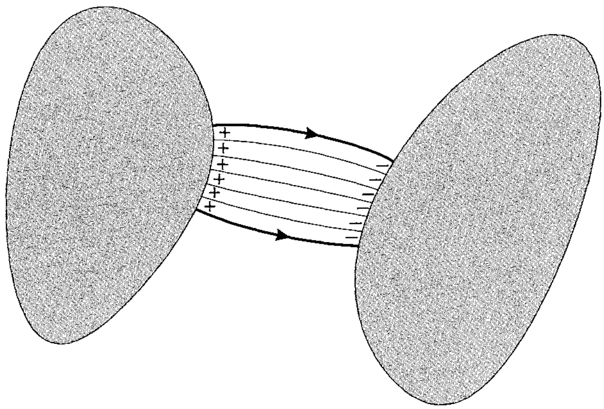

FIG. 3.20. Tube of force and corresponding surface elements of two conductors.

Thomson, it resulted from the conservation of heat flow. Faraday's lines of force had an exact counterpart in the lines of heat flow, which are everywhere perpendicular to the isothermal surfaces.[102]

At that stage, Thomson's heat-flow analogy was no longer confined to the finding of new mathematical theorems. It could be used to bridge two different physical hypotheses on the nature of electricity, Coulomb's and Faraday's. In print, Thomson refused to decide between the two hypotheses, because no known observable mechanical effect could discriminate between them.[103] The main purpose of his study was to show that one of Faraday's supposed proofs of the contiguous action, the existence of specific inductive capacity, could be interpreted in terms of direct action at a distance.

A material dielectric, Thomson showed, could be treated in a manner analogous to Poisson's theory of induced magnetism. Poisson, like Coulomb, imagined microscopic conducting cells (for the magnetic fluids) spread through the mass of the iron. The direct effect of an external magnetic force was to separate the fluids in each cell. Then the resulting dipoles generated secondary magnetic forces, to be superposed to the external one. There was a resulting discontinuity in the net magnetic force at the surface of magnetic bodies: the normal component of the force inside the body was a definite fraction of its value outside the body. The latter result, which no longer depended on microscopic assumptions, was the basis of Thomson's theory of dielectric effects. In the electric case, the surface discontinuity of the force is determined by the specific inductive capacity of the dielectric. This discontinuity can be seen to result from an 'imagined' surface charge, to be included among the sources of electric force. In the case of a Leyden phial filled with a material dielectric of inductive capacity ε, Thomson added the potential created by the imagined surface charges to that of the real charges on the conductors, and showed that for a

[102] Thomson 1845b: 30. [103] Thomson 1845b: 29.

given charge of the phial, its potential was ε times smaller than it would have been in a gaseous dielectric.[104]

This result was all that Thomson needed to explain Faraday's experiments on the 'division of charge' between identical Leyden phials filled with different dielectrics. What Faraday called 'the power or tension' of a phial corresponded to the potential, and the division of charge corresponded to the equalizing of the potentials of the external balls of the phials. Thomson concluded: 'The commonly received ideas of attraction and repulsion exercised at a distance, independently of any intervening medium, are quite consistent with all the phenomena of electrical action which have been here adduced.'[105]

At the same time, Thomson recognized that the heat-flow analogy could be extended to the case of dielectrics. The counterpart of a dielectric would be a solid with a thermal conductivity ε times larger than that of the surrounding medium. This made Poisson's force discontinuity a consequence of the continuity of the flux $-\varepsilon\nabla\theta$ across the surface of the solid. Thomson gave this precision only in a footnote. Unlike Faraday or Maxwell, he did not introduce a specific concept of flux for electrostatics. He limited the speculative use of analogies to a minimum, and tended to avoid conceptual distinctions that had no known empirical counterpart.[106]

3.5.4 The physical potential

While analyzing Faraday's experiments on inductive capacity, Thomson found that the 'power or tension' measured by Faraday by means of a carrying ball and a Coulomb balance was nothing but Green's potential. This was an essential insight, for the potential had previously been an abstract mathematical concept, with no direct operational significance.[107]

Another abstraction of potential theory was the integral $\frac{1}{2}\int\rho V d\tau$. Thomson learned from Gauss that the electricity at the surface of a conductor was in equilibrium if and only if this integral was a minimum. In August 1844 he interpreted this condition in terms of d'Alembert's principle of virtual velocities: for virtual displacements $\delta\mathbf{r}_i$ of the electric particles on the surface, the variation of the discrete version of Gauss's integral is

$$\delta\sum_{ij}\frac{1}{2}\frac{q_iq_j}{|\mathbf{r}_i-\mathbf{r}_j|}=-\frac{1}{2}\sum_{ij}q_iq_j\frac{(\mathbf{r}_i-\mathbf{r}_j)\cdot(\delta\mathbf{r}_i-\delta\mathbf{r}_j)}{|\mathbf{r}_i-\mathbf{r}_j|^3}=\sum_i\mathbf{f}_i\cdot\delta\mathbf{r}_i, \tag{3.2}$$

which must vanish according to d'Alembert.[108]

A few days earlier, Thomson had received a letter from his brother James, the engineer, concerning the efficiency of steam engines and the 'mechanical effect' they

[104] Thomson 1845b: 32–5. Cf. Wise 1981: 44–49; Smith and Wise 1989: 223–6. Ottaviano Mossotti later developed the same analogy, more slavishly and less efficiently (Mossotti 1847, 1850).

[105] *FER* 1: #1258; Thomson 1845b: 37.

[106] Thomson 1845b: 33n. Cf. Wise 1981a: 50–1; Smith and Wise 1989: 228–9.

[107] Thomson 1845b: 35.

[108] Thomson, 'Journal and research notebook': 14 August 1844, quoted in Smith and Wise 1989: 241.

could produce. The expression was synonymous with the 'travail' of French engineers, and referred to the height to which a given weight could be lifted by a machine. In the case of hydraulic engines, the mechanical effect originated from the fall of water. In the case of Carnot's ideal engine, which James and William frequently discussed together, the fall of heat from high to low 'intensity' was the source of the mechanical effect. Along the same lines, William Thomson saw that Gauss's integral was nothing but the mechanical effect needed to produce the distribution ρ.[109]

Around that time Thomson had been calculating the force between two electrified spheres, in connection with electrostatic experiments through which Snow Harris claimed to challenge Coulomb's law.[110] By analogy with an engine producing work, Thomson reasoned that the variation of the mechanical value of two insulated charged spheres during their separation measured the work spent to perform this separation. Consequently, the force between the two spheres could be calculated by taking the derivative of Gauss's integral with respect to their distance. This procedure was much simpler than the direct calculation of the force between the surface distributions of electricity.[111]

In his experiments Snow Harris grounded one of the spheres and connected the other to the ball of a charged Leyden jar. Therefore the force between the spheres depended on their distance, their radius, and the potential difference imposed by the jar, in a manner that Thomson could calculate. Thomson quickly perceived an opportunity for the absolute measurement of potentials.[112] He was familiar with Gauss's memoirs on geomagnetism, which introduced the notion of absolute measurement. Also, he spent part of 1845 working in the laboratory of the French champion of precision measurement, Victor Regnault.[113]

Immersed in a multiple context of steam engines, electrostatic experiments and calculations, geomagnetic measurements, and French steam measurements, Thomson brought physical meaning to the abstract concepts of French electrostatics. His novel insights all appear in a notebook entry of 8 April 1845:[114]

[109] James Thomson to William Thomson, 4 August 1844, discussed in Smith and Wise 1989: 242–3. Thomson, notebook remark of 8 April 1845, quoted *ibid.*: 245.

[110] Thomson 1845b; Harris 1834. Snow Harris operated with two conductors of various shapes, one being suspended on a balance and grounded, the other being fixed and connected to a battery or to a large electrified conductor. The attractive force turned out to be proportional to the square of the charge of the latter conductor. Harris doubted that received theories could explain his law (1834: 245). In response, Whewell and Thomson (1845b: 18–21) argued that his law was a consequence of Coulomb's theory, because the charge induced on the grounded body is proportional to the charge of the inducing body.

[111] Thomson, notebook entry of 8 April 1845, quoted in Smith and Wise 1989: 245; Thomson 1853c: 92–3. Thomson long delayed the publication, probably because he first wanted to develop the method of electrical images (Thomson 1845a, 1848–50): cf. Smith and Wise 1989: 246–7.

[112] For the first two-ball electrometer, cf. Thomson 1853c: 96. In an early draft of his 1845b, dated 12 April 1845, Thomson had already expressed the force of Harris's electrometer in terms of the potential difference of the two conductors: cf. Smith and Wise 1989: 246, 251–2. He noted that Harris's device was unsuited to quantitative measurements because of the lack of screening from inductive effects. Yet he could easily imagine the improvements that would tranform the device into an absolute electrometer.

[113] The head of the British 'magnetic crusade,' Colonel Sabine, had sought Thomson's expertise since the spring of 1844: cf. Smith and Wise 1989: 276–7. On Thomson and Regnault's laboratory, cf. *ibid.*: 106–8.

[114] Quoted in Smith and Wise 1989: 245.

To day, in the laboratory (of Physique at the Coll. de France, M. Regnault, prof.) I got the idea, which gives the mechanical effect necessary to produce any given amount of free electricity, on a conducting or non-conducting body [. . . .] This enables us to find the attraction or repulsion of two influencing spheres, without double integrals. Also the theorem of Gauss that [$\int \rho V d\tau$] is a minimum when V is a constant, shows how the double integral which occurs when we wish to express the action directly, may be transformed into the differential coefficient of a simple integral, taken with reference to the distance between the two spheres [. . .]. This has confirmed my resolution to commence experimental researches, if I ever make any, with an investigation of the absolute force, of statical electricity. As yet each experimenter has only compared intensities [surface charges] by the deviations of their electrometers. They must be measured by pounds on the square inch, or by 'atmospheres' [pressure units[115]].

In this incredibly dense statement, we find the germs of the energetic analysis of electrostatic systems, the energetic definition of force, and a related notion of absolute measurement.

As Gauss had remarked, absolute measurement presumed complete computability of the measuring apparatus in mechanical terms. Thomson met this requirement by means of the engineering notion of balancing mechanical values and effects. In this process, a potential difference became analogous to the difference of water height in hydraulic engines, or the difference of temperature in heat engines. In 1853 Thomson defined the potential directly in terms of a corresponding mechanical effect: 'The potential at any point in the neighbourhood of, or within an electrified body, is the quantity of work that would be required to bring a unit of positive electricity from an infinite distance to that point, if the given distribution of electricity were maintained unaltered.'[116]

3.5.5 *Absolute electrometry*

Thomson's resolution to start experimental researches concretized after he obtained, in 1846, the chair of natural philosophy at Glasgow. One of his first achievements in this area was the design of new electrometers. In his note of April 1845, Thomson meant to measure the 'electrical intensity,' which was, in contemporary terms, the surface charge of conductors. Experimenters commonly measured this quantity by means of Coulomb's '*plan d'épreuve*' or Faraday's 'carrying ball' brought in contact with the conducting surface and then with an electrometer. It seems likely, however, that Thomson also had in mind the absolute measurement of potentials. He knew that Harris's electrometers, if properly improved, would measure potential differences, and he knew that measurements of surface charge could lead to indirect potential measurements: in Faraday's experiments on specific inductive capacity, the surface charge of the connected outer ball was proportional to the potential of the internal conducting sphere. Thomson's later electrometers were all designed to measure potentials.[117]

[115] The square of a surface density has the dimension of pressure.
[116] Thomson 1853c: 87n; 1853a: 522n.
[117] On Thomson's professorship, cf. Smith and Wise 1989: Ch. 5.

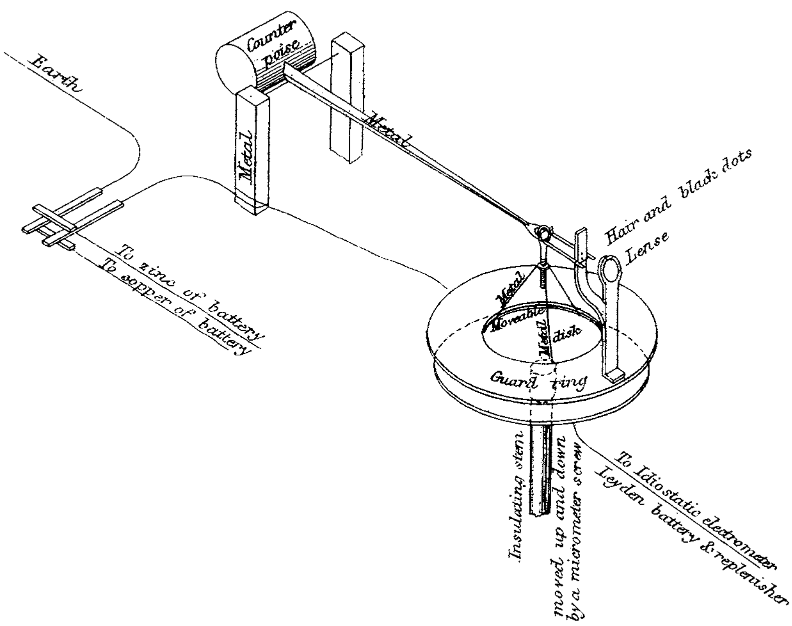

FIG. 3.21. Thomson's absolute electrometer (*TPEM*: 280–1).

Thomson's first project was an absolute electrometer based on Harris's two spheres. He soon preferred two parallel plane disks, which produced stronger forces. In this case, as Thomson had reckoned in 1845, the surface charges are mostly on the facing sides of the disks, and their density is nearly equal to $V/4\pi d$, where V is the potential difference and d the distance between the sides. The resulting attractive force between the two disks is easily found to be $V^2S/8\pi d^2$, if S is the surface of the facing sides. In deploying his device, Thomson demonstrated a Weberian patience and meticulousness. His mathematical power and his practical imagination allowed ingenious improvements. For example, he invented the 'guard-ring' surrounding the moveable disk and maintained at the same potential, so that the field below the moveable disk remained uniform for large values of the distance d (Fig. 3.21).[118]

An essential motivation of Thomson was the determination of the ratio c of the electromagnetic to the electrostatic charge unit. As he explained in 1853, the electromagnetic value of a given electromotive force could be compared with the value measured by an absolute electrometer. In 1860 he performed a first measurement of

[118] Thomson 1853c: 96 (two-ball electrometer built and analyzed); 1845b: 19–20 (force between two plates); 1853b: 553 and 1860a: 238 (description of two-plate electrometer, first exhibited at the Glasgow meeting of the British Association in 1855). For the history and the subsequent improvements, cf. Thomson 1867: 281–92. Cf. also Smith and Wise 1989: 250–2. Thomson also built the first sensitive electrometers, including the divided-ring and the quadrant electrometer. Cf. *ibid.*: 694–7, and Hong 1994a: 284–5.

this kind with a Daniell battery and the two-plate electrometer. The result agreed reasonably well with the value of c that Weber and Kohlrausch had obtained in 1856 by measuring the same electric charge with a Coulomb balance and an electrodynamometer. By that time the knowledge of c had acquired much practical importance, as we will see in a moment.[119]

The theoretical importance of this constant also increased in time. According to Maxwell's theory of 1862, c had to be equal to the velocity of light. Despite his general hostility toward this theory, Thomson found the conjecture worth verifying. In 1867 he set his Glasgow students to an improved determination of c using his absolute electrometer and a standard resistance of known absolute value. The result (2.82×10^8 m/s) agreed no better with Foucault's measurements of the velocity of light (2.98 and 3.08×10^8 m/s) than Weber and Kohlrausch's old value (3.11×10^8 m/s). However, this project nicely illustrated the increasing sophistication of Thomson's absolute electrometry.[120]

3.5.6 *Electromotive force and mechanical effect*

In 1848 Thomson extended his considerations of mechanical effect to electrodynamics. He first considered Neumann's 'very beautiful theorem,' according to which the electromotive force in a conductor moving with respect to a magnet is equal to the time derivative of its electromagnetic potential. 'It has appeared to me,' Thomson announced, 'that a very simple *a priori* demonstration of the theorem may be founded on the axiom that the amount of work expended in producing the relative motion on which the electro-magnetic induction depends must be equivalent to the mechanical effect lost by the currents induced in the wire.' Thomson first determined the mechanical effect lost by the current through the following reasoning. When the strength of the magnet is multiplied by n the induced current is multiplied by n, while the electromagnetic forces acting on this current and their mechanical effect thus provided are multiplied by n^2. The mechanical effect lost by the current i in the time $\mathrm{d}t$ is therefore equal to $ki^2\mathrm{d}t$, where k is a constant depending on the circuit and on the choice of units. Now the work of the electromagnetic forces during this time is $i\mathrm{d}P$, where P denotes the potential of a unit current with respect to the magnet. Balancing this work with the lost mechanical effect yields the expression $(1/k)\mathrm{d}P/\mathrm{d}t$ for the induced current, in conformity with Neumann's potential law.[121]

[119] Thomson 1853b: 553; Thomson 1860a. Maxwell and Thomson's notation for c was v. The constant C of Weber's theory (which Weber denoted c) is equal to $c\sqrt{2}$ (see Appendix 2). In 1855 Thomson had already obtained a rough estimate of c by working back from cable-retardation results: cf. Smith and Wise 1989: 456.

[120] *BAR* (1869): 434. The Glasgow students sent a constant current through a resistance of known absolute value and an electrodynamometer, and measured the potential at the terminals of the resistance with an absolute electrometer. Cf. Maxwell 1873a: #772. On the competition with Maxwell on the same problem, and on the increasing complexity of relevant resources, cf. Schaffer 1995. On later methods and the final convergence between c and the velocity of light, cf. Rosa 1889.

[121] Thomson: 1848b: 91. At that time Thomson was not yet convinced of the kinetic nature of heat, which explains why he does not identify the lost mechanical effect with the Joule heat. His reasoning,

After reading Weber's memoir on absolute resistance measurement (1850), Thomson realized that his analysis of electromagnetic induction provided a fruitful alternative to Weber's definitions. In Weber's absolute units, the electromotive force e in a rectilinear conductor of unit length cutting the lines of force of a uniform magnetic field of unit intensity at right angles is equal to the velocity v of the conductor. If this conductor is the only moving part of a closed circuit, a current is induced in proportion to this electromotive force. The electromagnetic force then acting on the moving conductor is numerically equal to the absolute electromagnetic measure i of the current. Therefore the work needed in a unit of time to move the conductor is vi, or ei. According to the 'principle of mechanical effect' this work must be equal to the mechanical effect consumed in the circuit. For other kinds of source, the mechanical effect produced by an electromotive force e acting on a current i is still equal to the product ei, because it should not depend on the nature of the source. Conversely, Thomson proposed to define electromotive forces by the mechanical effect they produced on a unit current.[122]

These considerations, and the earlier reflections on the potential, may be seen as the electric facet of Thomson's progression toward a general formulation of the energy principle. At the same time they commenced the subsumption of physics under this principle, a process intensifying in the 1850s under the lead of Thomson, Helmholtz, and Rankine. Thomson's notion of absolute measurement made mechanical effect—later to become 'energy'—the measure of all physical quantities. It bore the mark of the engineering culture of his brother James, and it prefigured the definition of forces in terms of energy functions that is found in Thomson and Tait's *Treatise on natural philosophy*.[123]

3.5.7 The transatlantic telegraph and BA units

Thomson's brand of practical physics met spectacular successes in submarine telegraphy. Around 1850 the introduction of gutta-percha, an excellent insulator, permitted the installation of the first subterranean and submarine telegraph lines. However, signalling on such lines proved much less efficient than on air lines. Consulted by the Electric Telegraph Company, Faraday experimented on the new cables, and published the following diagnosis in 1854. A single-wire cable acted as a huge Leyden jar, the gutta-percha corresponding to the glass, the surface of the copper wire to the inner coating, and the sea water or the Earth to the outer coating. Consequently, (electrostatic) induction in the gutta-percha competed with the (electrostatic) induction through the wire, and the discharge of electricity took more time. Faraday regarded this phenomenon as a sensational confirmation of his idea that induction,

similar to one Helmholtz had published a year earlier (see Chapter 6), omits self-induction and assumes that the internal energy of the magnet–circuit system does not depend on the relative position of the magnet and the circuit, which happens to be true (the field energy does not vary in this case, because the magnet's and the circuit's fields are orthogonal in Fourier space). Cf. Knudsen 1995.

[122] Thomson 1851b. Thomson also proposed a thermal measurement of absolute resistance.

[123] Cf. Smith and Wise 1989: 250, Ch. 11. On the growth of energy physics, cf. Smith 1998.

or dielectric polarization, was the essence of electricity and always preceded conduction.[124]

Of Faraday's considerations, Thomson retained only the idea that the electrostatic capacity of the cable had to be taken into account. With a sure feeling for legitimate approximations, he provided simple mathematical relations between directly measurable quantities, and ignored the deeper nature of the process. Calling C the capacitance per unit length of the cable, R its resistance per unit length, V the electrostatic potential, and i the current, he set the rate of change of the electric charge on an element dx of the wire, $Cdx\partial V/\partial t$, equal to the decrease of the current in this element, $-(\partial i/\partial x)dx$. Using Ohm's law $i = -(1/R)\partial V/\partial x$, this leads to the diffusion equation:

$$RC\frac{\partial V}{\partial t} = \frac{\partial^2 V}{\partial x^2}. \tag{3.3}$$

According to Thomson's solution for a sudden rise of potential at the origin of the cable, 'the time required to reach a stated fraction of the maximum strength of current at the remote end [of a cable of length l] will be proportional to RCl^2.'[125]

This 'law of squares' allowed Thomson to predict the signalling performance of long cables, knowing that of smaller ones. Having computed the capacitance of a cylindrical condenser through the flow analogy, he could also indicate how to minimize the retardation. Such theoretical knowledge was essential for the projected transatlantic telegraph. Thomson did more. He designed and patented high-performance apparatus for the emission and reception of signals. He tested the various components of the cable in his laboratory. And he helped solve the many difficulties encountered in laying a 2000 mile cable at the bottom of the ocean. His methods surprised contemporary engineers, who were accustomed to rough empirical procedures. Yet he soon became a director of the Electric Telegraph Company. The success of the transatlantic cable of 1866—after a first unsuccessful trial in 1858—owed much to his advice. As a reward for this major contribution to the wealth of the British Empire, Queen Victoria knighted Thomson in November of the same year.[126]

Thomson's involvement in British telegraphy meant an important transition in the relations between fundamental science and engineering. Thanks to his efforts, the practical advantages of theoretical knowledge became evident, and precision measurement replaced the previous 'rules of the thumb' of British engineers. In the following twenty years, physics laboratories were created in several academic institutions in order to teach Thomsonian methods to future engineers and physicists. In

[124] Faraday 1854. Cf. Smith and Wise 1989: 446–7; Hunt 1991c. On early submarine telegraphy, cf. Bright 1898; Coates and Finn 1979; Smith and Wise 1989: Ch. 19.

[125] Thomson to Stokes, 28 October 1854; Thomson 1855b. Cf. Whittaker 1951: 227–30; Smith and Wise 1989: 447–53. Thomson neglected electromagnetic induction and leakage. Capacitance was indeed the dominant cause of retardation for the cables he was studying. The more complete 'equation of telegraphy' first appeared in Heaviside 1876 (without leakage), 1881 (with leakage).

[126] Cf. Smith and Wise 1989: 661–83.

1861 Thomson easily convinced the British Association for the Advancement of Science (BAAS) to create a committee on standards of electrical resistance. The precise measurement of electric quantities had great commercial importance for the telegraph industry. Thomson imposed an absolute system of electric units based on the mechanical units of work, time, and length.[127]

The BAAS committee set the electromagnetic unit of resistance, the 'ohmad,' to 10^7 m/s; measured the resistivity of pure silver in this unit by an improvement of Weber's method; built a silver-wire standard with the required resistance; and sold copies of this standard throughout the British Empire. When in 1881 the first international congress on electrical standards was held in Paris, Thomson acted as vice-president and imposed a good deal of the BAAS system. The international units were named ohm, volt, farad, coulomb, and ampere.[128]

One basic duty of the BAAS committee was to determine the ratio c of the electromagnetic to the electrostatic charge unit. As Faraday and Thomson knew, this ratio was essential to the analysis of telegraph cables, for the retardation of signals depended on the combination of an electrostatic effect (inductive capacitance) with an electrodynamic one (ohmic resistance). In Thomson's equation (3.3), the resistance R and the capacitance C must of course be given in the same system of units. However, resistances were naturally measured in electromagnetic units, and capacitances in electrostatic units. In order to compute the retardation RCl^2, Thomson needed to know the conversion factor c. In 1855 he did the reverse, that is to say, he used Faraday's retardation measurements for a rough (unpublished) estimate of c. Later, in the 1860s, with his Glasgow students he measured this important quantity for the BAAS committee, as was already mentioned.[129]

To sum up, practical concerns strongly shaped Thomson's works on electricity, as they did for the rest of his physics. This is manifest not only in his interest in instruments, but also in his theoretical approach. His concept of the electrostatic potential acted as a bridge between the different cultures to which he belonged. It provided interesting theorems for Liouville's *Journal de Mathématiques*; it justified Faraday's electrostatic manipulations; it integrated the engineering notion of mechanical effect; and it met the requirements of German absolute measurement. The notion was precise, efficient, and concrete, and yet ontologically neutral. Poisson's fluid theory of electrostatics and Faraday's reasonings on lines of electric induction could both be translated into potential language, without loss of efficiency.[130] Potentials could be measured by their mechanical effect, whereas electric fluids or the tensions of lines of force remained beyond empirical reach. In brief, the physical potential was the paragon of Thomson's most essential qualities: mathematical power, versatility,

[127] On physics teaching laboratories, cf. Gooday 1990. On the BAAS committee, cf. Smith and Wise 1989: 684–90; Schaffer 1992; Smith 1998: Ch. 13.

[128] Smith and Wise 1989: 690–5; Hunt 1994. For a closer analysis of the implied social processes, cf. Schaffer 1992. For the conflict between German and British notions of standards, cf. Olesko 1996.

[129] Thomson to Airy, 2 February 1855 (c from retardation), quoted in Smith and Wise 1989: 456; Thomson 1860a; *BAR* 39 (1869): 434–8. Cf. Smith and Wise 1989: 455–8, 694; Schaffer 1995.

[130] Cf., e.g., Thomson 1860b: 254–8.

pragmatism, and latitudinarianism. It gradually became an indispensible tool to any one interested in electricity.[131]

3.6 Thomson's magnetic field

3.6.1 The strained solid

Thomson excelled at developing ontologically neutral concepts that could be very practical to a variety of users. Yet he did not exclude more speculative representations of electricity and magnetism. These could offer valuable analogies and new techniques for solving problems, even if their physical meaning was uncertain. Since his work on dielectrics, Thomson had admired the consistency of Faraday's theoretical views. The discovery of magneto-optical rotations increased his sympathy for the field conception.

Faraday believed that electric and magnetic forces were propagated through stresses in the intervening medium. He did not try, however, to explain these stresses in terms of specific mechanical strains. In his view, mechanics, especially mathematical dynamics, had no precedence over the broader notions of force and power on which his physics was based. The Scottish-trained Thomson thought differently. In his mind a stress could only be understood by analogy with a strained elastic solid. By a happy coincidence, his friend George Gabriel Stokes completed an elegant study of the elasticity of solids at the time of Faraday's discovery of the magneto-optical rotation.

Stokes's study was based on a new approach to the dynamics of continuous media, with which he derived his famous equation of viscous fluids. Poisson and Navier has already obtained similar equations, starting with the Laplacian picture of molecules interacting by central forces. In conformity with the British tendency to deal directly with elements of the continuum, Stokes 'examined the nature of the most general instantaneous motion of an element of fluid.' Decomposing the velocity differential $d\mathbf{v}$ into its symmetrical and antisymmetrical parts (in the now classical manner), he found that the most general motion was obtained by superposing three dilations or contractions around three othogonal axes (the principal axes of the symmetrical part of $d\mathbf{v}$), and a rotation of angle $\frac{1}{2}\nabla \times \mathbf{v}$. Rotations do not strain the element. The dilations and contractions do, and imply three additional pressures (positive or negative) along the principal axes. From these remarks, Stokes derived the system of stresses acting on an arbitrary surface element, and the resulting equation of motion.[132]

This reasoning and its counterpart for an elastic solid involved a kinematics of continuous media that became instrumental in field mathematics. In particular, the expressions $\partial v_y/\partial z - \partial v_z/\partial y$, etc., which had previously been used only to express

[131] Thomson's potential replaced older notions of tension in three influencial texts: Jenkin 1873, Maxwell 1873a, and Wiedemann 1874. Some conservative electricians fought this evolution: cf. Hong 1994a.

[132] Stokes 1845a: 80. Of course, Stokes used Cartesian coordinates, not vectors.

the condition for the existence of a potential, now indicated a local rotation or twist in the medium if **v** meant a velocity or a displacement. Maxwell had this picture in mind when he introduced the 'curl' of a vector in 1870. So did Thomson in October 1846, when he described three simple kinds of elastic strain that were analogous to the fields of a point charge, of a magnetic dipole, and of a current element.[133]

For the displacement **u** of an incompressible elastic solid, Stokes's equilibrium equations imply that $\Delta\mathbf{u}$ should be a gradient (of pressure). The three following deformations satisfy this condition as well as the condition of incompressibility ($\nabla\cdot\mathbf{u}=0$):

$$\mathbf{u}=\frac{\mathbf{r}}{r^3},\quad \mathbf{u}=\frac{\mathbf{m}\times\mathbf{r}}{r^3},\quad \mathbf{u}=\frac{i d\mathbf{l}}{r}-\frac{1}{2}\nabla\frac{i d\mathbf{l}\cdot\mathbf{r}}{r} \tag{3.4}$$

Thomson identified the first deformation with the electric force of a unit charge, and the curls of the second and third deformations with the magnetic force produced respectively by the magnetic moment **m** and by the current element $i d\mathbf{l}$. In the magnetic case he was inspired by the Faraday effect, which suggested a rotational deformation of the medium. Thus was born the vector potential, of which the magnetic force is the curl. British mechanical analogy produced this concept ten years before German mathematical analysis did.[134]

In a letter to Faraday, Thomson set the limits of his investigation:

> I enclose the paper which I mentioned to you as giving an analogy for the electric and magnetic forces by means of the *strain*, propagated through an elastic solid. What I have written is merely a sketch of the mathematical analogy. I did not venture even to hint at the possibility of making it the foundation of a physical theory of the propagation of electric and magnetic forces, which, if established at all, would express as a necessary result the connection between electrical and magnetic forces.

Again, Thomson avoided commitment to a *physical* explanation of electric and magnetic forces. Yet the new analogy went a little further than the heat-flow analogy. The latter was meant to suggest new theorems and to retrieve some geometrical features of Faraday's view, whereas the strain analogy offered 'a mechanical representation of electric and magnetic forces' that integrated Faraday's notion of stresses in the field.[135] It could be a starting point toward truer mechanical analogies and could ultimately lead to a 'physical theory of the propagation of electric and magnetic forces.' It was a first, mild attack of what Thomson later called his 'ether dipsomania.'[136]

[133] Maxwell 1870: 265; Thomson 1847a. On Maxwell's terminology, see Crowe 1967: 117–39.

[134] Thomson 1847a; For the German vector potential see Kirchhoff 1857b, discussed *supra*, p. 72.

[135] However, the strains described by Thomson imply stresses different from Faraday's, and they do not yield the correct expressions for the mechanical forces acting on charges, magnets, or currents.

[136] Thomson to Faraday, 11 June 1847, in Thompson 1910, Vol. 1: 203–4; Thomson 1847a: title; Lord Kelvin to FitzGerald, 9 April 1896, in Thompson 1910, Vol. 2: 1065. Cf. Smith and Wise 1989: 256–60.

3.6.2 Diamagnetic forces

Thomson promptly returned to more sober physics. The object was the calculation of the mechanical force **f** acting on polarizable small spheres in a magnetic field, in relation with Faraday's recent experiments on diamagnetism. Following Poisson, Thomson replaced the polarized sphere with a magnetic dipole **M**. Then the force acting on the sphere is the sum of the forces acting on the two opposite poles. In symbols, we have

$$\mathbf{f} = (\mathbf{M}\cdot\nabla)\mathbf{H} = \frac{1}{2}\nabla(\mathbf{M}\cdot\mathbf{H}) = \frac{1}{2}k\nabla(H^2), \tag{3.5}$$

where the second and third expressions follow from the irrotational character of the external magnetic force **H** and from the assumption of a linear polarizability k of the sphere. Consequently, the sphere tends to move toward regions of higher magnetic force if k is positive, and the reverse is true if k is negative. Thomson concluded that reverse polarization completely justifed Faraday's law according to which 'a portion of [diamagnetic matter], when under magnetic action, tends to move from stronger to weaker places or points of force.'[137]

Thomson did not speculate on the deeper meaning of this result. He did not publicly espouse Faraday's view that the effect confirmed the physical character of the magnetic lines of force. In a later discussion of the same effect published in 1851, he noted that the variations of $\frac{1}{2}kH^2$ represented a mechanical effect, but did not claim that this effect (energy) was stored in the field. He adopted Faraday's field terminology, but introduced it in a purely operational manner.[138]

3.6.3 A private analogy

In 1847 Thomson tried to generalize his surface-replacement theorem to magnetic forces. One of the problems he examined was: is there a distribution of magnetic force outside a given closed surface so that the normal component of the force **H** immediately outside the surface has a given value (with zero integral, as required by the balance of Northern and Southern magnetic matter)? There is an obvious hydrodynamic counterpart to this problem, obtained by identifying **H** with the velocity of an ideal incompressible fluid (and exchanging the inside and the outside of the surface): is there an irrotational motion of a fluid mass confined within a closed surface whose motion is given? In both cases the mathematical condition determining **H** is that it should be the gradient of a harmonic function.[139]

[137] Thomson 1847b: 493, 497, and Thomson 1850a; *FER* 3: #2418. Cf. Smith and Wise 1989: 261–2.

[138] Thomson 1851a: 475: *ibid.*: 467–8: 'The *total magnetic force at any point* is the force which the north pole of a unit bar-magnet would experience from all magnets which exert any sensible action on it, if it produced no inductive action on any magnet or other body [. . .]. Any space at every point of which there is a finite magnetic force is called "a field of magnetic force" [. . .]. A "line of force" is a line drawn through a magnetic field in the direction of the force at each point through which it passes.'

[139] Thomson notebook, 29 March 1847, quoted in Smith and Wise 1989: 263–4; Thomson to Stokes, 20 October 1847, in Wilson 1990.

Thomson used this reformulation of the magnetic problem in order to bring the question to Stokes, the expert on hydrodynamics. Stokes immediately gave a positive answer. Parallely, Thomson found that Gauss's method of quadratic forms could be extended to this case: the solution exists because it is given by the function **H** for which $\int H^2 d\tau$ is a minimum with the given boundary conditions. In the hydrodynamic problem, this makes the kinetic energy of an irrotational motion a minimum. Thomson used the remark to simplify the proof of some hydrodynamic theorems.[140]

More generally, Thomson found a minimum principle that comprehended all useful existence theorems for hydrodynamics, heat theory, electrostatics, and magnetism. These problems admit a potential V, which satisfies the generic equation

$$\nabla\cdot(\alpha^2\nabla V) = -4\pi\rho. \tag{3.6}$$

The variable parameter α^2 corresponds to the conductivity in heat theory, to the dielectric constant in electrostatics, and to the permeability in the case of magnetism; in hydrodynamics, abrupt variations of this parameter can be used to simulate the boundaries of the fluid. This equation always has a solution, Thomson showed, because it corresponds to the minimum of the quadratic form

$$Q = \int\left(\alpha\nabla V - \frac{1}{\alpha}\nabla U\right)^2 \mathrm{d}\tau, \tag{3.7}$$

where U is the solution for $\alpha = 1$ (which is already known).[141]

It would be tempting to think that the hydrodynamic analogy led Thomson to regard $\frac{1}{2}H^2$ as representing the actual energy distribution in the magnetic field. Yet he did not say so. In his eyes, an essential advantage of considerations of mechanical effect was that they did not depend on the internal make up of physical systems. They made any system a black box, an engine with input and output. Moreover, Thomson could not take the fluid analogy so seriously as to make **H** a linear velocity, because this would have contradicted his intuition of the Faraday effect, according to which **H** meant a local twist of the medium. And he did not know yet how to extend the analogy to induced magnetism. For all these reasons, he did not publicize the hydrodynamic analogy. He published his results either as mathematical theorems or as hydrodynamic laws.[142]

Thomson's major memoir on magnetism of 1849 was quite positivist in tone. There he refrained from any assumption on the nature of magnetism, and avoided

[140] Stokes to Thomson, 10 April 1847, in Wilson 1990; Thomson to Stokes, 20 October 1847; Thomson 1849. Cf. Smith and Wise 1989: 263–2.

[141] Thomson 1848c. Cf. Smith and Wise 1989: 271.

[142] Thomson 1848c, 1849. However, Thomson published the hydrodynamic analogy in 1872a: 455–9, and generalized it to induced magnetism in 1872a: 578–87 (the magnetic 'permeability' being so named in analogy with the permeability of the porous medium in which the fluid circulates: cf. Thomson 1872a: 484).

both magnetic fluids and Amperean currents. He based his theory on the phenomenological notion of elementary magnetic moments, and defined the magnetic force operationally. To measure in thought the magnetic force within the substance of a magnet, he carved out a thin crevasse in the direction of the polarization, and inserted a test unit pole. He introduced magnetic charges only as mathematical aids, defined as the convergence of the polarization ($-\nabla\cdot\mathbf{M}$). In private he also used the equivalent currents given by the curl of the magnetization ($\nabla\times\mathbf{M}$), but did not publish that until the 1870s (after he had adopted Ampère's hypothesis).[143]

Thomson's main concern was a geometrical analysis of the different kinds of magnetic polarization and the resulting magnetic fields. In this context he introduced the distinction between solenoidal and lamellar distributions. For the former kind, the magnet can be decomposed into infinitesimal tubes ($\sigma\omega\lambda\epsilon\nu$ in Greek) with longitudinal polarization. For the latter, the magnet can be decomposed into lamellar sheets with transverse polarization. The corresponding mathematical conditions are $\nabla\cdot\mathbf{M}=\mathbf{0}$, and $\nabla\times\mathbf{M}=\mathbf{0}$. Of course, in the first case Thomson had in mind an incompressible fluid and its tubes of flow. But he relegated the analogy to a footnote.[144]

3.6.4 Mediation

As a consequence of Faraday's developing view of magnetism, Thomson ended up releasing important aspects of the flow analogy. In a letter written in June 1849 he explained to Faraday why an elongated diamagnetic body in a uniform magnetic field should orient itself in the direction parallel to the magnetic force, using Faraday's concept of conducting power for the lines of force. At the 1852 meeting of the British Association, he showed pretty diagrams of lines of force (Fig. 3.22) which he had calculated by a method previously developed in the context of heat theory. He called attention to the remarkable resemblance of these diagrams to those Faraday had recently shown at the Royal Institution to explain his views on diamagnetic action, and claimed to have justified by rigorous mathematical analogy expressions such as 'the conducting power for the lines of force.' He could have added that his minimum principle of 1847 (Eqn. 3.7) could be interpreted as a principle of least resistance, in conformity with Faraday's intuition: the flux corresponding to the conductivity α^2 is indeed $\alpha^2\nabla V$, while the flux for a unit conductivity is ∇U.[145]

Thomson's use of analogy was here quite similar to his previous use of the heat analogy to make sense of Faraday's electrostatics. For any magnetic phenomenon, he could translate an explanation in terms of elementary polarizations acting at a distance into another explanation in terms of Faraday's conducting power for the

[143] Thomson 1849–50: 340, 361–2; Thomson 1872a: 424–5. Cf. Smith and Wise 1989: 279–81. Thomson adopted the Amperean currents after his 1856 analysis of the Fadaday effect (see below): cf. Thomson 1872a: 419n.

[144] Thomson 1849–50: 378–92.

[145] Thomson to Faraday, 19 June 1849 (or 1847: cf. Wise 1981: 59n), in Thompson 1910, Vol. 1: 214; Thomson 1848a (equilibrium of diamagnetic bodies); Thomson 1852 (BA), 1847c (magnetic curves computed), 1843 (heat flow); Thomson 1852: 515 (quote).

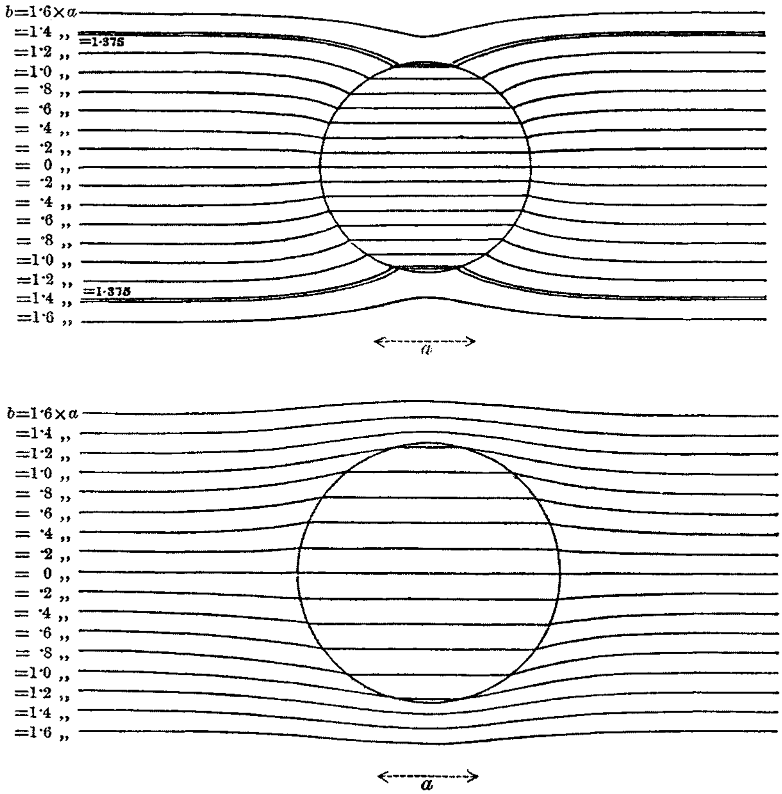

FIG. 3.22. Thomson's computed lines of force around a paramagnetic (above) and a diamagnetic (below) sphere (*TPEM*: 490–1).

lines of force. Where Faraday saw a contradiction, he perceived an exact mathematical equivalence:

> All that Tyndall has done in verifying Weber [. . .] is mere illustration or verification of a conclusion following equally from Faraday's theory, or from the arbitrary assumption [. . .] that a diamagnetic experiences a reverse effect (polarization) throughout its substance, to that experienced by a paramagnetic.

Being perfectly fluent in both languages, Thomson adressed the proponents of the opposite views in their own terms. To Faraday, he explained the orientation of a diamagnetic bar in a uniform magnetic field as an effect of least resistance to the passage of lines of force; to Tyndall he explained the same effect by the mutual interaction of elementary diamagnets.[146]

[146] Thomson, notebook, 6 January 1858, reproduced in Knudsen 1971: 50; Thomson to Faraday, 19 June 1849, in Thompson 1910, Vol. 1: 214; Thomson to Tyndall, 12 March 1855, in Thomson 1872a: 535–8.

FIG. 3.23. Rotation of the plane of polarization of light traveling through a helix for two opposite directions of propagation.

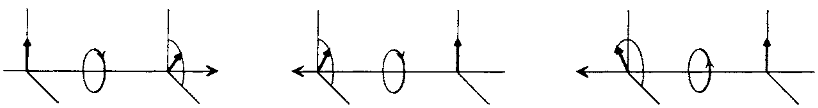

FIG. 3.24. Rotation of the plane of polarization of light traveling through a current loop. The case (b) derives from (a) by inverting the direction of propagation; (c) from (a) by time-reversal.

3.6.5 Molecular vortices

Thomson's neutrality had limits, however. Since his conversion to the kinetic theory of heat around 1850, he was more inclined to speculations on a general theory of ether and matter. In 1856 he had a serious relapse of 'ether dipsomania.' The cause was again the Faraday effect. The rotary power of optically active substances like turpentine could easily be explained by an helicoidal asymmetry of the molecules of the substance: the polarization plane of a light beam should then rotate in the direction defined by the helicity (right- or left-handed) (Fig. 3.23(a)). This explanation could not be extended to the Faraday effect, because it implied that the optical rotation, as seen from the light source, did not depend on the direction of propagation (Fig. 3.23(b)). As Faraday emphasized, the magnetically induced rotation was reversed when the direction of propagation of light was reversed (Fig. 3.24(a),(b)). Consequently, the modification of the medium that was responsible for the rotation had the asymmetry of an oriented circle. Faraday imagined some microscopic rotations in the medium. These rotations could be either static (twist) or dynamic (continuing motion).[147]

In his mechanical representation of magnetic forces of 1847, Thomson flirted with the first possibility. Yet by 1856 he was convinced that the dynamic option was the only one available. Using a general argument and a mechanical model, he explained how a microscopic rotational motion of the medium could affect the polarization of light, and he asserted that there was no other possible explanation of the symmetry properties of the Faraday effect. Unfortunately, he gave no strict proof of this crucial point. A generous reading of his obscure argument leads to the following consideration. By time reversal (Fig. 3.24(c)), the rotation of the polarization of

[147] Thomson 1856; *FER* 3, series 19: ##2231–2. Cf. Knudsen 1976: 244–7, 273–6.

the light beam (with respect to a fixed observer) is reversed. Therefore, the responsible rotation in the medium should also be reversed, which proves its dynamic nature.[148]

In Thomson's new analysis the Faraday effect demonstrated the existence of Amperean currents—which he had previously denied—and it suggested a magnetic exploitation of Rankine's kinetic theory of heat. According to the Scottish engineer, heat was nothing but the rotational motion of space-filling 'molecular vortices.' According to Thomson, magnetization could be an alignment of the vortices, and the common angular momentum would determine the magnetic moment. Strictly speaking, these pictures only applied to matter, heated and magnetized. However, Thomson had reasons to extend them to the ether.[149]

In 1854 he inferred a lower limit for the density of the ether from the mechanical effect of a cubic mile of sunlight, and suggested that the ether was only 'a continuation of our atmosphere.' In a letter to Tyndall of March 1855, he argued that since vacuum had 'perfectly decided mechanical qualities,' it probably had the magnetic property as well. In short, the ether was only a dilute form of matter. Accordingly, Thomson accompagnied his 1856 discussion of the Faraday effect with a vague but bold suggestion:[150]

> The explanation of all phenomena of electro-magnetic attraction or repulsion, and of electro-magnetic induction, is to be looked for simply in the inertia and pressure of the matter [ponderable or not] of which the motions constitute heat. Whether this matter is or is not electricity, whether it is a continuous fluid interpermeating the spaces between molecular nuclei, or is itself molecularly grouped; or whether all matter is continuous, and molecular heterogeneousness consists in finite vortical or other relative motions of contiguous parts of a body; it is impossible to decide, and perhaps in vain to speculate, in the present state of science.

Be it in vain or not, in a notebook entry of 1858 Thomson did speculate on a general picture of ether and matter. He imagined a universal fluid with myriads of rotating motes which could perhaps be further reduced to permanent eddies. The gyrostatic rigidity of the motes or eddies would permit transverse vibrations of the medium, to be identified with light. Heat would be the rotation of the eddies. Electricity would correspond to the less disturbed parts of the fluid between the motes. Then an electric current would alter the rotation of the eddies, as a string pulled between two adhering wheels. This connection would account for the Joule effect, and for magnetism, understood as an alignment of the eddies' axes. Magnetic attractions would result from the centrifugal force of the eddies combined with the pressure of the fluid. Electromagnetic induction would correspond to the storage of momentum in the oriented vortical motions. Lastly, the Faraday effect would result from the influence of these motions on transverse vibrations of the medium. Thomson ended with a prophetical remark: 'A complete dynamical

148 For a closer analysis of Thomson's argument, cf. Knudsen 1976.

149 Thomson 1856: 571.

150 Thomson 1854; Thomson to Tyndall, 12 March 1855, *TPEM*: 535; Thomson 1856: 571. Cf. Smith and Wise 1989: 397–402 (ether), 407–8 (Faraday effect).

illustration of magnetism and electromagnetism seems not at all difficult or far off.'[151]

This was Thomson's first attempt to understand all physics in terms of fluid vortices. Soon he found support in Helmholtz's theorems on vortical motion in an ideal incompressible fluid. For many years he tried to model ether and matter as arrays of vortices. He did not, however, pursue the idea of a vortex-based illustration of electromagnetism. This was left to his most gifted admirer, James Clerk Maxwell.[152]

3.7 Conclusions

Faraday and Thomson invented field theory: they introduced theoretical entities in the space between electric and magnetic sources, and they elaborated powerful techniques for investigating the properties of these entities. They perceived a convergence of their projects and developed a mutual admiration. However, their interests, methods, and concepts were extremely different.

In Chapter 1, we saw how Faraday's first electrodynamic studies depended on the systematic experimental exploration of the 'power' emanating from magnetic sources. He refused to speculate on the internal structure of sources, and focused instead on the intervening space, in which observed actions were regarded as a manifestation of 'magnetic power.' His explorations connected actual and virtual actions, and thus generated mappings of the power, new power-induced states of matter, and rules for the development of these states. None of this required Thomson's advanced mathematics: ordinary language and intuitive geometry sufficed. In fact, Faraday's exploring frenzy resulted in part from his distrust in established mathematical theories. The efficiency of his explorations largely depended on his qualitative concept of power.

In his later works on electrolysis, electrostatics, and diamagnetism, Faraday extended the approach of his earlier researches. In each case, he gleaned new facts and shaped original views in non-mathematical language. His exceptional attention to processes in the intervening space or matter between sources led to his discovery of dielectric and diamagnetic effects. It also instructed his redefinition of charge and current in terms of the cessation of dielectric polarization, and his notion of conducting power for the magnetic lines of force. With these field concepts he was able to predict and explain effects which, in received theories, could not be foreseen without sophisticated mathematics.

Faraday's commitment to the physical existence of field entities evolved in time. At the beginning of his experimental researches he sometimes used wordings and concepts that suggested the reality of the lines of force. The earliest example is his

[151] Thomson notebook, 6 January 1858, reproduced and commented in Knudsen 1971. Cf. also Smith and Wise 1989: 410. The interpretation of induction is not in the notebook, but may be inferred from the remark in Thomson 1856, quoted above.

[152] Cf. Smith and Wise 1989: Ch. 12.

concept of the repulsion between parallel magnetic curves, foreshadowed in the rotations paper of 1821.[153] However, he long resisted the temptation to close the issue. In public he maintained an operational definition of the lines, and did not defend their reality until he had accumulated many favorable arguments and reached the age of unconditional respectability. His arguments of the mid-1830s only concerned the effect of matter on the transmission of force: he believed he had proved that the electric force from a given particle of matter could only reach the nearest particles, more remote actions being indirect, through chains of contiguous particles. To the extent that they referred to polarization in chains, the electric lines of force were real. However, they could also travel across the intermolecular vacuum, in which case Faraday had no proof of their physical character.

The problem of transmission across a vacuum became more acute after the discovery of diamagnetism. With respect to the transfer of magnetic action, vacuum was intermediate between dia- and paramagnetic matter and thus seemed to be on the same footing as matter. Also, the proofs of action between contiguous particles did not transfer smoothly from the electrostatic to the magnetic case. Yet Faraday's confidence in the physical reality of the magnetic lines of force increased as he multiplied the varieties of their uses. He could 'touch' them and 'illuminate' them, at least metaphorically. He could modify their course by means of better or worse 'conductors.' Thus he meant them to exist independently of any ponderable or imponderable medium. In his most daring dynamicist speculations, there was nothing but force, distributed in space with variable qualities and intensities.

According to a widespread misinterpretation, from the beginning Faraday's researches were motivated by the elimination of direct action at a distance. In reality, he regarded an interaction via lines of force as direct action at a distance, whenever no matter contributed to the transmission of the force. He expected the interaction to take time in such cases, but not because a subtle medium or ether was involved. The reason of the retardation was the physical nature of the lines of force. In short, his notion of force transcended the usual dichotomy between direct action and action through a medium.

According to another misinterpretation, Faraday's researches were aimed, from the beginning, at confirming Boscovich's atomist dynamism.[154] Admittedly, Faraday's 'power' and 'force' were dynamicist notions that denied the distinction between action and agent (Newtonian force and imponderable fluid) and eventually the distinction between force and matter. But there is no evidence that Faraday supported a specific dynamicist philosophy, not even Boscovich's. As far as his experimental and conceptual practices were concerned, the focus on power and force only meant the endeavor to express phenomena in terms of virtual actions in intervening spaces. Faraday freely explored the fields of action and the correlations of different

[153] More exactly, in 1921 Faraday spoke of the repulsion between similar magnetic powers in the space between two antiparallel currents. See *supra*, p. 20.

[154] Cf. Williams 1965 for the thesis and Spencer 1967 for the refutation. Levere 1968 denies Davy's and Faraday's interest in speculative metaphysics, but documents their religious inclination toward center-of-force atoms.

powers, and gradually formed his theoretical views in this process. He suspended his opinion on the physical character of the lines of force until exploration ceased to learn him more.

Faraday had no mathematical or mechanical preconceptions, and his theory mostly reflected patient experimental explorations. In contrast, Thomson was originally a mathematician with a strong background in analytical mechanics. His practical bent did not result from familiarity with the laboratory, but from Scottish Common Sense philosophy and interactions with engineers.

Thomson's analogy between electrostatics and heat flow was originally meant to transfer theorems. Implicitly, it also provided new mathematical structures in the space between conductors. Thomson did not wish, however, to commit himself to any specific physical interpretation of these structures. He only showed how his analogy could connect two possible interpretations, Coulomb's and Faraday's. He generally avoided metaphorical concepts that had no direct empirical counterpart. To a large extent, the same remarks apply to his later analogies between magnetism and hydrodynamics.

Aware of his role as a mediator in the cultural complex of mathematics, experimental philosophy, engineering, and geophysics, Thomson forged multi-purpose concepts that transcended cultural barriers and individual theoretical preferences. The most important and successful of these concepts, the physical potential, belonged equally in mathematical theory, electrostatic experiment, engineering considerations of mechanical effect, Gaussian absolute measurability, and the later industrial design of voltmeters. Physicists conversant with French electrostatics could easily express the potential in terms of electric fluid densities. The followers of Faraday's views, if any, could draw the lines perpendicular to the equipotential surfaces and call them lines of force. Energeticists could adopt Thomson's definition of the potential in terms of mechanical effect.

After his major contributions to thermodynamics around 1850, Thomson grew more interested in speculations on the ultimate nature of heat, electricity, and magnetism. He also started to support some aspects of Faraday's new field physics. However, his approach still differed widely from Faraday's. He believed in a mechanical ether, namely a dilute form of matter to which the mechanics of continuous media applied, and of which Faraday's lines of force represented local strains or motions. Inspired by the Faraday effect, Rankine's kinetic theory of gases, and Stokes's hydrodynamics, he figured the ether as an ideal incompressible fluid in which arrays of molecular vortices would represent magnetic fields. All of this was tentative and illustrative. Yet Thomson's hope to reduce all physics to motions in an ultimate medium was well anchored. In particular, he believed that a dynamical illustration of electromagnetism was close at hand.

4

Maxwell

4.1 Introduction

War es ein Gott, der diese Zeichen schrieb?[1] So asks Boltzmann, quoting from Goethe, in an epigraph to his lectures on Maxwell's theory. In the nineteenth century section of the physicists' pantheon, Maxwell's rank remains the highest. The tribute is well deserved. Maxwell wrote the field equations which still form the basis of our understanding of electromagnetism. He subsumed optics under electromagnetism. He founded statistical physics. He created a new style of theoretical physics. As the first director of the Cavendish Laboratory, he contributed to the increasing sophistication of British experimental physics.

Glorification, however, tends to obscure the true nature of Maxwell's achievements. It was not a god who wrote these signs, but a man who had gone through two of the best British universities and had carefully studied Faraday and Thomson for himself. His electromagnetism and his style of physics, innovative though they were, owed much to Thomson, who had already transformed British physics in an even more significant manner and had defined basic concepts and new perspectives of electromagnetism. The heroic account also deforms Maxwell's results. His electrodynamics differed from today's 'Maxwell's theory' in several respects, as basic as the distinction between source and field. It was not a closed system, and it included suggestions for future electromagnetic research. In the present chapter, we will approach this more authentic Maxwell.

4.1.1 Scot and wrangler

James Clerk Maxwell was, like Thomson, a Cambridge graduate first trained in a Scottish university. Despite a seven-year difference in age, the two men's approaches to physics had deep similarities. They both lent a central role to geometry in the expression of mathematical and physical ideas. Following their Scottish professors (John Nichol for Thomson, and James Forbes for Maxwell), they held a broad view of physics, including the full range of experimental subjects and technical

[1] 'Was it a god who wrote these signs?' Boltzmann 1891–1893, Vol. 1: 96, from the introductory monologue of Goethe's *Faust*.

engineering problems. At the same time, they shared the mathematical virtuosity cultivated in the Cambridge Tripos, and had an eye for deeper theory as promoted by John Herschel and William Whewell. By drawing formal analogies between various branches of physics, they combined Baconian diversity and Newtonian unity.[2]

There were, however, perceptible nuances between Thomson's and Maxwell's research styles. Maxwell's involvement in technical, practical matters was less than Thomson's, while his interest in geometry was more sustained and diverse than Thomson's. Following the Clerk family's artistic bent, Maxwell was fascinated by the beauty of geometrical figures. After William Hamilton and Immanuel Kant, he regarded space and time as necessary forms of our intuition of phenomena. His interests and skills in philosophy and literature were exceptionally high for a British scientist. Unlike Thomson, he accompanied his use of dynamical analogies with sophisticated philosophical comment. He wrote good poetry, and brilliantly discussed moral philosophy for the Cambridge Apostles. Lastly, there was an essential psychological difference between Maxwell and Thomson. As an enthusiastic prodigy, Thomson launched essential ideas in numerous concise papers, but rarely found time for their full exploitation or for global syntheses. Maxwell was slower and more dependent on other physicists' innovations, but he could persevere several years on the same subject and erect lofty monuments.[3]

Maxwell first learned electricity and magnetism from James Forbes at Edinburgh University. Forbes adopted an empirical approach, and ignored French or German mathematical fluid theories. Maxwell was still free of theoretical prejudice when in February 1854 he asked his pen-friend William Thomson: 'Suppose a man to have a popular knowledge of electrical show experiments and a little antipathy to Murphy's Electricity [the British rendering of Poisson's electrostatics], how ought he to proceed in reading & working so as to get a little insight into the subject which may be of use in further reading?' Thomson's reply is lost. We know, however, that Maxwell read Faraday and Thomson first, then Ampère and Kirchhoff, and lastly Neumann and Weber. Thus, the young Maxwell assimilated Faraday's field conceptions and developed a distaste for continental theories. He later explained to Faraday: 'It is because I put off reading about electricity till I could do without prejudice, that I think I have been able to get hold of some of your ideas, such as the electro-tonic state, action of contiguous parts &c.'[4]

[2] On Maxwell's biography, cf. Campbell and Garnett 1882; Everitt 1975. For the relative effects of Maxwell's Scottish and Cambridge backgrounds, cf. Wilson 1985, Siegel 1991, and Harman 1998. On Cambridge's Mathematical Tripos, cf. Wilson 1982; Warwick [1999].

[3] On Maxwell and geometry, cf. Harman 1990: 2–3; Harman 1995a: 20–2, 28–9; Harman 1998: 13–15. On Maxwell, Scottish common sense, and Kant, cf. Harman 1985b, 1998: 27–36, and Hendry 1986. For the psychological comparison between Maxwell and Thomson, cf. Everitt 1975: 59–60.

[4] Maxwell to Thomson, 20 February 1854, *MSLP* 1: 237; Maxwell to Faraday, 19 October 1861, *MSLP* 1: 688. When he wrote to Thomson on 13 November 1854 (*MSLP* 1: 262), Maxwell had read Ampère and Kirchhoff, but not Neumann and Weber. He had read Thomson 1849–1850 (mathematical theory of magnetism) before his letter of February 1854 (cf. Harman 1998: 72–3). His interest in electricity was unusual for a Cambridge student, for this subject had been excluded from the Tripos curriculum some years before.

4.2 On Faraday's lines of force

4.2.1 Gridding the field

Before the end of 1854, Maxwell reported substantial progress to Thomson. Following Faraday, he defined the lines of force as the lines everywhere tangent to the force acting on a pole or point charge. Following Gauss and Thomson, he also introduced the surfaces normal to these lines, that is, the equipotentials. His first innovation was to consider simultaneously the lines and the surfaces and to regulate their spacing, in order to allow quantitative geometrical reasoning (Fig. 4.1). He had used similar space-gridding a few months earlier in a discussion of surface folding, and all his previous works involved the geometry of lines or surfaces.[5] In the electric or magnetic context, he required that the potential difference between two successive equipotentials should be a constant. On a given equipotential surface he drew two systems of curves defining cells with a size inversely proportional to the intensity of the electric or magnetic force, and then traced the tubes of force passing through these cells. The tubes played the same role as Faraday's unit lines of force.[6]

Maxwell expressed Faraday's law of electromagnetic induction in terms of the tubes. In the case of a closed circuit, the induced electromotive force depends on the decrease of the number of tubes passing through it. In mathematically precise terms, *the induced electromotive force around a circuit is equal to the decrease of the surface integral of magnetic force across any surface bounded by the circuit.* Maxwell immediately applied the law to a simple analytical case, the induction of currents in a conducting sphere rotating in the magnetic field of the Earth. Twenty years after the discovery of electromagnetic induction, he was the first theorist to take Faraday so seriously as to give a mathematical expression of his induction law.[7]

Maxwell's geometrical representation also helped him reformulate the relation between an electric current and the resulting magnetic field. In his vision, a current in a closed circuit determined a series of equipotentials bounded by the circuit (Fig. 4.2). The number of these equipotentials was a natural geometrical characteristic

[5] Maxwell to Thomson, 13 November 1854, *MSLP* 1: 258; Maxwell [1854a]: 252. Maxwell was aware of Thomson's theory of magnetization (1849–50), which introduced lamellar and tubular analysis of the distributions of magnetism in magnets. In my reconstruction, I assume that Maxwell had the line–surface gridding before he considered the relation between current and magnetic force. The essential point, however, is that he *simultaneously* considered the potential theory of magnetism and Faraday's lines of force.

[6] Originally, Maxwell spoke of lines of polarization instead of tubes of force. On the genesis and meaning of 'On Faraday's lines of force' I have found much inspiration in Norton Wise's insightful paper on 'the mutual embrace' (Wise 1979). Particularly important are his comments on Maxwell's field-geometrical method and on the role of the intensity/quantity distinction.

[7] Maxwell to Thomson, 13 November 1854, *MSLP* 1: 260; *ibid.*: 260–1, and Maxwell 1862: 226–9 for the rotating sphere. At that stage Maxwell did not have yet the distinction between force and flux (intensity and quantity). He used the term 'polarization' (which I have replaced with 'magnetic force') 'to express the fact that at a point of space the south pole of a small magnet is attracted in a certain direction with a certain force' (*MSLP* 1: 256).

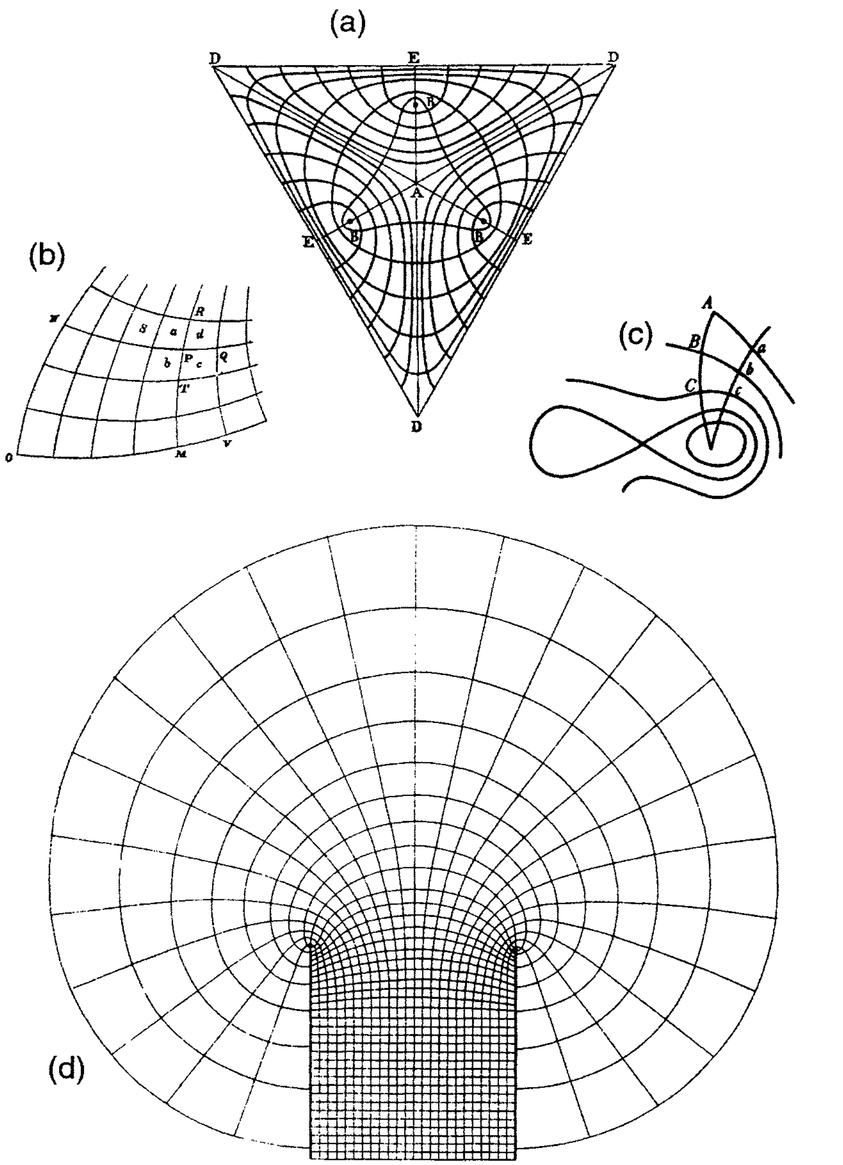

FIG. 4.1. Some of Maxwell's geometrical grids: (a) compression and dilation lines of a glass triangle (Maxwell 1850: 68), (b) lines of surface bending (Maxwell 1854b: 99), (c) electric lines of force and equipotentials (Maxwell [1854]: 252, used by permission of Cambridge University Press), (d) *idem* for a two-plate condensor (Maxwell 1873a: plate 12).

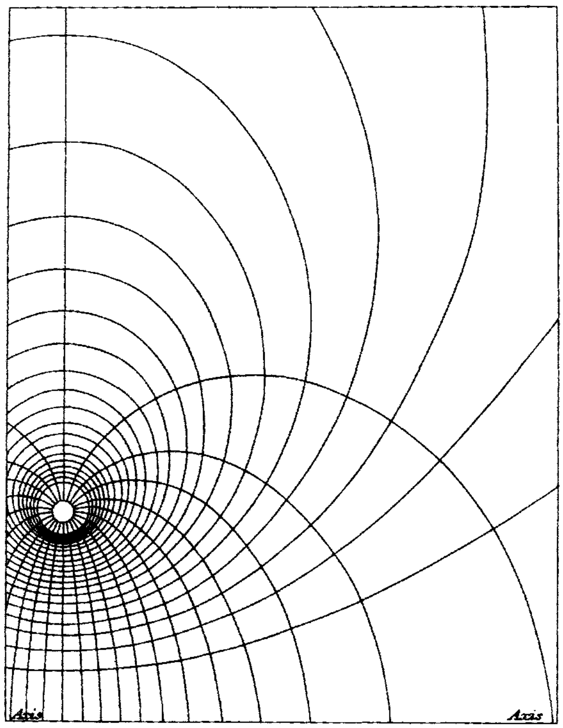

FIG. 4.2. Magnetic lines of force and equipotential surfaces of a circular current, in a half-plane delimited by the axis of the circle (Maxwell 1873a: plate 18).

that obviously depended on the intensity of the current.[8] It also had an energetic meaning, as the work performed by a unit magnetic pole on a curve γ embracing the circuit. In order to determine this number, Maxwell resorted to Ampère's equivalence between a circuit C and a net of contiguous loops (Fig. 4.3) and reasoned as follows.

If the small shaded loop embracing the curve γ were removed, the remaining loops would be equivalent to a double magnetic sheet with a hole at the place of the shaded loop. The corresponding potential would be single-valued, and its total variation on γ would be zero. Consequently, the line integral of the magnetic force, or the number of equipotentials, depends only on the current circulating in the shaded loop, which is equal to the current in C.[9] With a Gaussian eye for topological relations, Maxwell insisted that the integration curve and the current curve had to embrace each other.

[8] This number is well-defined for a proper choice of the potential unit.

[9] The numerial coefficient is determined by considering a particular case, for instance a circular current and its axis regarded as a curve closed at infinity.

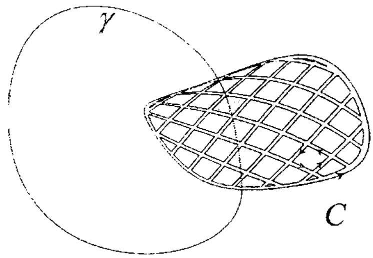

FIG. 4.3. Amperean net and mutually embracing curves for Maxwell's first proof of the magnetic circuital law.

In general, *the line integral of the magnetic force on any closed curve is measured by the sum of the intensities of the embraced currents.*[10]

Maxwell was first to enunciate this result, which is improperly called the Ampère law (or theorem).[11] Together with the induction law, it formed the basis of his own field theory of magnetism. William Thomson was no doubt aware of these two laws.[12] However, they did not appear as explicit, central statements in his papers. Having started from action-at-a-distance theory and energetic considerations, he gave central importance to the potential concept. In contrast, Maxwell started with Faraday's lines of force and expressed fundamental laws directly in terms of the field of force. He regarded the equipotentials as derivative constructs, defined as the surfaces orthogonal to the lines of force, even though they played a role in his derivation of the Ampère law and in his discussion of field energy.

4.2.2 *The resisted-flow analogy*

In the same letter to Thomson, Maxwell applied his line–surface geometry to conduction currents: here the lines refer to electric motion, and the surfaces to equal tension. He also suggested an analogous treatment of induced magnetism, based on Faraday's notion of conductive power for the magnetic lines of force. In the general

[10] Maxwell to Thomson, 13 November 1854, *MSLP* 1: 256–7. Maxwell also stated another theorem: the integral of the magnetic force across a surface bounded by the circuit only depends on the intensity of the current (and on the shape of the circuit, Maxwell should have added), not on the shape of the surface (*MSLP* 1: 257). This results from the equivalence of the circuit with a double magnetic sheet and from the fact that the integral of the magnetic force produced by magnetic masses is zero over any closed surface that does not contain masses, as Maxwell noted a little earlier in his letter. Cf. Wise 1979 and Hendry 1986: 126–30.

[11] I avoid the expression 'Ampère's law,' which is even more misleading.

[12] Thomson was certainly aware of the Ampère law, as appears from his discussion of the potential of a closed current (Thomson 1850b: 426n). However, he did not state it formally, presumably because a true field formulation of electromagnetism was not on his agenda. Regarding electromagnetic induction, Thomson had used Faraday's law and had given its mathematical expression in a particular case (Thomson 1851c: 484). Maxwell knew these papers very well, so he asked Thomson whether he had not 'the whole draught of the thing [Maxwell's "On Faraday's lines of force"] lying in loose papers' (Maxwell to Thomson, 13 September 1855, *SLMP* 1: 322).

case of variable conductivity, however, Maxwell did not know how to prove the existence of the potential. In 1848 Thomson had published a strict but enigmatic proof, based on minimizing a certain positive integral (see p. 129). Maxwell wondered whether his correspondent had a general theory based on the theorem. The answer is lost. In any case, by the spring of 1855 Maxwell was elaborating on the flow analogy that Thomson had so successfully applied to existence theorems. 'Have you patented that notion with all its applications?, for I intend to borrow it for a season,' he wrote Thomson.[13]

Maxwell's resulting analogy, published in the first part of 'On Faraday's line of force,' departed from Thomson's original heat analogy in several respects. Maxwell replaced heat with an 'imaginary incompressible fluid,' arguing that it would provide a more concrete analogy, since heat was no longer regarded as a substance. He treated the most general case of heterogenous and anisotropic conduction, whereas Thomson had mostly confined himself to the homogenous case. Most important, Maxwell integrated his tubes-and-cells geometry in the analogy and thus increased its intuitive appeal and demonstrative power. His aim was to produce a method that 'required attention and imagination but no calculation.'[14]

Maxwell first described the uniform motion of an incompressible and imponderable fluid through a resisting medium with sources and sinks. He parted the fluid into unit tubes, in which one unit of volume passes in a unit of time. The configuration of the tubes completely defines the flow, since their direction gives that of the fluid motion, and their inverse section determines the velocity. Maxwell further assumed the resistance of the medium (a porous body) to be proportional to the fluid velocity. Since the motion is uniform and the fluid has no mass, this implies that the velocity is proportional to the gradient of pressure, as Fourier's heat flux is proportional to the gradient of temperature.[15]

With this illustration, Maxwell proved essentially the same theorems as Thomson had done with the heat-flow analogy. He did not quite meet his aim to prove the existence of the potential—or pressure—in the case of a heterogenous medium.[16] But his reasonings, being based on the geometry of the tubes of flow, were more direct and vivid than Thomson's. For example, he obtained the surface-replacement theorem by the following simple consideration: the flow outside an imaginary closed surface is unchanged if we substitute for the fluid inside the surface a system of sources and sinks on the surface that maintain the flow in each intersecting tube.[17]

[13] Maxwell to Thomson, 13 November 1854, *MSLP* 1: 259–61; Maxwell to Thomson, 15 May 1855, *MSLP* 1: 307. Maxwell also thought of relating the work of electrodynamic forces to the number of cells in the field (*MSLP* 1: 259).

[14] Maxwell [1855]: 306; Maxwell to Stokes, 22 February 1856, *SLMP* 1: 403; Maxwell 1856b: Part I. Cf. Rosenfeld 1956: 1652–5; Heimann 1970; Everitt 1975: 87–93; Moyer 1978; Wise 1979; Hendry 1986: 133–8; Harman 1990: 12–15; Siegel 1991: 30–3.

[15] Maxwell 1856b: 160–4.

[16] He only proved that if the potential flow exists in a heterogenous medium, then it may be regarded as created by an imaginary system of sources spread in a homogenous medium (Maxwell 1856b: 168–71).

[17] Maxwell 1856b: 168 (#20).

Maxwell also introduced 'surfaces of equal pressure' such that a unit pressure difference exists between two consecutive surfaces. He used the cells determined by the intersection of these surface with the tubes of flow to express the energy spent by the fluid to overcome the resistance of the porous medium. In a given cell, a unit mass of fluid experiences a pressure decrease of one unit. Therefore, one unit of energy is spent in each cell, and the total amount of dissipated energy is equal to the total number of cells. This amount must be equal to the work produced or received by the sources and sinks, which is the sum of the products of their rate of flow times the pressure under which they are working. In this picturesque manner, Maxwell justified the interchange of a field integral with a sum over sources, which Gauss and Thomson had obtained by purely analytical means.[18]

Next, Maxwell explained the analogy of the imaginary flow with various domains of electricity and magnetism. For electrostatics, the tubes of flow correspond to Faraday's lines of electric induction, the pressure to the potential, and the resistance of the medium to the inductive capacity of the dieletric. For magnetism, the tubes of flow correspond to Faraday's magnetic lines of force,[19] the pressure gradient to 'the resultant force of magnetism,' and the resistance of the medium to the inverse of Faraday's 'conducting power' for the lines of force. For electrokinetics, the tubes of flow correspond to the lines of current, the pressure to the electrostatic potential or tension, and the resistance of the medium to the electric resistance.[20]

The total number of cells also has a counterpart in each of the three analogies. Clearly, it is equal to the electrostatic energy in the electrostatic case and to the Joule heat in the electrokinetic case. Maxwell only discussed the case of para- and diamagnetism, for it justified Faraday's rule of least resistance to the passage of the lines of force: the total number of cells, or resistance overcome by the flow, is then equal to the total magnetic potential from which mechanical forces are derived. Note, however, that the analogy could not help Maxwell locate the magnetic energy in the field: the number of fluid cells (corresponding to the later $\int \mathbf{B}\cdot\mathbf{H}\,d\tau$) did not measure an energy stored in space, but the energy dissipated by the flow.[21]

4.2.3 *Intensity/quantity*

A more relevant aspect of the analogy was the distinction between force and flux implied in the idea of a resisted flow. Maxwell knew that for electric conduction and electrostatic induction Faraday distinguished between electric intensity and quantity. Intensity meant tension causing the current or the electroscopic effect. Quantity

[18] Maxwell 1856b: 161–2, 173–5. In electrostatic symbols, the interchange reads $\int \rho V d\tau = \int \varepsilon E^2 d\tau$.

[19] This is not quite true, because Faraday's magnetic lines of force have no source, whereas Maxwell's tubes of flow have sources corresponding to the magnetic masses.

[20] Maxwell 1856b: 175–83. In the draft of December 1855 (*MSLP* 1: 364), Maxwell did not introduce the electrostatic potential as a counterpart of the pressure. He did in the final version, as a consequence of his reading Kirchhoff 1849b.

[21] Maxwell 1856b: 178–80. Maxwell also combined this analogy with the equivalence between closed currents and double magnetic sheets, to derive the rule that circuits tend to move in such a way as to maximize the magnetic quantity (flux) across them (*ibid.*: 185).

referred to the strength of the electric current, or to the integral current that a charged condenser could produce. Faraday forcefully defended this usage, even though it departed from Ampère's and Thomson's. Shortly before Maxwell's elaboration of his lines of force, he wrote: 'The idea of intensity or the power of overcoming resistance [to induction or to conduction], is as necessary to that of electricity, either static or current, as the idea of pressure is to steam in a boiler, or to air passing through apertures or tubes; and we must have language competent to express these conditions and these ideas.' Maxwell used his flow analogy to systematize the distinction.[22]

In Maxwell's frame of tubes and surfaces, quantity referred to the number of tubes crossing a surface and intensity to the number of surfaces crossed by a given tube. In terms reminiscent of Faraday's, Maxwell wrote: 'The amount of fluid passing through any area in a unit of time measures the *quantity* of action over this area; and the moving force which acts on any element in order to overcome the resistance, represents the total *intensity* of action within the element.' This distinction immediately became central to Maxwell's field theory. An essential virtue of formal analogies accorging to Maxwell was to provide a classification of physico-mathematical quantities that guided theory construction.[23]

Maxwell's first use of the quantity/intensity distinction was unfortunate. To a given intensity, he surmised, there should correspond one and only one quantity. Therefore, the quantity corresponding to the electric potential differences should be the same in electrostatics and in electrokinetics, and a dielectric should be nothing but a very bad conductor in which the electric quantity or current were too small to be detected. Maxwell believed that he could find support for this idea in Faraday's assertion that 'insulation and ordinary conduction cannot be properly separated when we are examining into their nature,' whereas Faraday only meant that conduction always involved the build up and breakdown of electrostatic induction.[24]

Maxwell made a more felicitous use of quantities and intensities in further reflections on electromagnetic induction. He had already been able to express Faraday's law in mathematical terms, and he had learned from Helmholtz how to derive it by an energetic argument. Yet he was no more satisfied with the form of Faraday's law than Faraday himself was:

> This law, though it is sufficiently simple and general to render intelligible all the phaenomena of induction in closed circuits, contains the somewhat artificial conception of the number of lines *passing through* the circuit, exerting a physical influence on it. It would be better if we could avoid, in the enunciation of the law, making the electromotive force in a conductor depend upon lines of force external to the conductor.

Maxwell wanted to express the electromotive force as the variation of some 'intensity' representing the electrotonic state of the conductor. The quantity/intensity

[22] Faraday 1854: 519

[23] Maxwell 1856a: 371; Maxwell 1856b: 182, 189–92. *Ibid.* on 182 Maxwell referred to the distinction of Faraday 1854: 519. Cf. Wise 1979; Everitt 1975: 89–90; Moyer 1978; Hendry 1986: 136–42.

[24] Maxwell 1856b: 181, including a reference to Faraday 1854: 513n.

distinction, a derived symbolism, and some theorems by Thomson and Stokes provided the answer.[25]

In symbols, the fluid quantity across a surface element $\mathbf{dS}$ is $\mathbf{a}\cdot\mathbf{dS}$, where $\mathbf{a}$ denotes the fluid current. The intensity (pressure difference) along the length element $\mathbf{dl}$ is $\boldsymbol{\alpha}\cdot\mathbf{dl}$, where $\boldsymbol{\alpha}$ denotes the moving force. The incompressibility of the fluid gives $\nabla\cdot\mathbf{a} = 0$ (in the absence of sources). The resistance k of the medium implies $\boldsymbol{\alpha} = k\mathbf{a}$. To specify the magnetic and electric cases, Maxwell inserted the suffixes 1 and 2. Then the Ampère law applied to an infinitesimal closed curve yields

$$\nabla\times\boldsymbol{\alpha}_1 = \mathbf{a}_2 \quad (\nabla\times\mathbf{H} = \mathbf{j}). \tag{4.1}$$

Conversely, this relation implies the Ampère law, because, as Maxwell had learned from Stokes,

$$\int\boldsymbol{\alpha}\cdot\mathbf{dl} = \int\int(\nabla\times\boldsymbol{\alpha})\cdot\mathbf{dS}, \tag{4.2}$$

if the first integration is performed over a curve bounding the surface of the second.[26]

In the same notation, Faraday's law reads:

$$\int\boldsymbol{\alpha}_2\cdot\mathbf{dl} = -\frac{\mathrm{d}}{\mathrm{d}t}\int\int\mathbf{a}_1\cdot\mathbf{dS} \quad \left(\int\mathbf{E}\cdot\mathbf{dl} = -\frac{\mathrm{d}}{\mathrm{d}t}\int\int\mathbf{B}\cdot\mathbf{dS}\right). \tag{4.3}$$

Maxwell wanted to reformulate this law in terms of a state of the circuit itself. From Thomson he knew that any divergenceless vector could be regarded as the curl of another vector. He therefore introduced the intensity $\boldsymbol{\alpha}_0$ such that[27]

$$\mathbf{a}_1 = \nabla\times\boldsymbol{\alpha}_0 \quad (\mathbf{B} = \nabla\times\mathbf{A}). \tag{4.4}$$

According to theorem (4.2), the line integral of this intensity is equal to the magnetic quantity passing through the curve. Consequently, the induced electromotive force is simply given by

$$\boldsymbol{\alpha}_2 = -\frac{\partial\boldsymbol{\alpha}_0}{\partial t} \quad \left(\mathbf{E} = -\frac{\partial\mathbf{A}}{\partial t}\right). \tag{4.5}$$

[25] Maxwell 1856a: 373. On Maxwell's reference to Faraday's electrotonic state, cf. Doncel and Lorenzo 1996.

[26] Maxwell 1856b: 203–5; *ibid.*: 206, with proof of Stokes' theorem based on the equivalence between the curves and a net of infinitesimal loops. This theorem was first stated by Thomson in a letter to Stokes of 2 July 1850 (Wilson 1990: 96–7; Stokes 1880–1905, Vol. 5: 320–1), and published by Stokes in the Smith prize examination for 1854, which Maxwell took.

[27] When there are magnetic masses (magnets), $\mathbf{a}_1$ is not divergenceless; Maxwell extracted the divergence-free part in a manner found in Stokes memoir on diffraction (Stokes 1849: 254–7): Maxwell 1856b: 200–1, 203–4.

Maxwell called α_0 the 'electro-tonic intensity,' for he believed that he had found the mathematical expression of Faraday's long-sought electro-tonic state.[28]

More generally, Maxwell professed to have reached the 'mathematical foundation of the modes of thought indicated in the *Experimental Researches*.' His success depended on a geometric deployment of the resisted-flow analogy, followed by a more symbolic approach in which the quantity/intensity distinction played a crucial guiding role. Unlike Thomson, Maxwell accompanied his use of analogy with philosophical comments. He explained that 'physical analogies' offered 'a method of investigation which allows the mind at every step to lay hold of a clear physical conception, without being committed to any theory founded on the physical science from which that conception is borrowed, so that it is never drawn aside from the subject in pursuit of analytical subtleties, nor carried beyond the truth by a favorite hypothesis.' Weber's theory, elegant though it was, depended on a questionable physical hypothesis. In contrast, Maxwell's own theory did not contain 'even the shadow of a true physical theory; in fact,' Maxwell went on, 'its chief merit as a temporary instrument of research is that it does not, even in appearance, *account for* anything.' The fluid analogy applied indifferently to separate compartments of electric science; it did not account for mechanical forces among charged bodies, currents, or magnets; and it ignored the relation between electricity and magnetism. The incompressible fluid was purely imaginary, the electro-tonic intensity purely symbolic. Nevertheless, 'by a careful study of the laws of elastic solids and of the motions of viscous fluids' Maxwell hoped 'to discover a method of forming a mechanical conception of the electro-tonic state adapted to general reasoning.'[29]

4.3 On physical lines of force

4.3.1 Molecular vortices

In May 1857, after reading Thomson's 'new lights' on the Faraday effect and molecular vortices, Maxwell wrote to his friend Cecil Monro: 'This was a wet day & I have been grinding at many things and lately during this letter at a Vortical theory of magnetism and electricity which is very crude but has some merits, so I spin & spin.' In a letter to Thomson written a few months later he described a gyromagnetic device that would confirm the existence of vortices in magnetized iron, if only the rotating fluid had enough inertia. Three years later, in the first part of 'On

[28] Maxwell 1856a: 374. In the final paper (1856b), instead of assuming Faraday's law, Maxwell used a flawed energetic reasoning inspired by Helmholtz's 'derivation' of electromagnetic induction: cf. Knudsen 1995.

[29] Maxwell 1856b: 207, 156, 207, 188. For Maxwell's reaction to Weber's theory, see also Maxwell to Thomson, 15 May 1855, *MSLP* 1: 305–6. On the analogies of 'On Faraday's lines of force,' cf. Moyer 1978; Wise 1979, 1981a; Hendry 1986: 143–55; Siegel 1991: 30–3, 38–9. On Maxwell's use of analogy in general, cf. Turner 1955; Hesse 1961, 1966, 1973; Kargon 1969; Chalmers 1973a; Hendry 1986; Siegel 1991; Cat 1995.

physical lines of force' he proposed a theory of magnetism based on molecular vortices.[30]

In his 1856 paper on the Faraday effect, Thomson had written: 'The explanation of all phenomena of electro-magnetic attraction or repulsion, and of electro-magnetic induction, is to be looked simply in the inertia and pressure of the matter of which the motions constitute heat.' He then assumed heat to consist of Rankine's molecular vortices and magnetism in the alignment of these vortices. In 1860 Maxwell supported Clausius's kinetic theory, and therefore could not follow the whole of Thomson's suggestion. He did not doubt, however, that magnetism involved vortical motion, as a consequence of Thomson's analysis of the Faraday effect. And he could precisely see why Thomson believed that the pressure and inertia of the revolving matter determined magnetic forces and electromagnetic induction.[31]

If there exist fluid vortices along the lines of force, he reasoned, then the centrifugal force of the vortices implies a larger pressure in the directions perpendicular to the lines of force than along the lines of force. This is equivalent to an isotropic pressure combined with a tension along the lines of force. Maxwell thus retrieved Faraday's intuition of a mutual repulsion of the lines of force and a tension along them. He only had to verify that this stress system implied the known magnetic attractions and repulsions.[32]

Calling p the isotropic pressure, μ the density of the medium, and $\mathbf{H}$ a vector giving the direction of the vortices and the average linear velocity of the fluid, the stress system is

$$\sigma_{ij} = -p\delta_{ij} + \mu H_i H_j, \tag{4.6}$$

in anachronistic tensor notation. From the net effect of these stresses on the sides of an infinitesimal cube, Maxwell derived the force

$$f_i = \partial_j \sigma_{ij} = -\partial_i p + H_i \partial_j (\mu H_j) + \mu H_j (\partial_j H_i - \partial_i H_j) + \mu \partial_i \left(\frac{1}{2} H^2\right), \tag{4.7}$$

or

$$\mathbf{f} = (\nabla \cdot \mu\mathbf{H})\mathbf{H} + (\nabla \times \mathbf{H}) \times \mu\mathbf{H} + \mu\nabla\left(\frac{1}{2} H^2\right) - \nabla p. \tag{4.8}$$

[30] Maxwell to Faraday, 9 November 1857, *MSLP* 1: 552: 'But there are questions relating to the connexion between magneto-electricity and a possible confirmation of the physical nature of magnetic lines of force. Professor W. Thomson seems to have some new lights on this subject'; Maxwell to Monro, 20 May 1857, *MSLP* 1: 507; Maxwell to Thomson, 30 January 1858, *MSLP* 1: 579–80. Cf. Siegel 1991: 33–7; Harman 1990: 30–1; Everitt 1975: 93–5; Everitt 1983: 132–4.

[31] Thomson 1856: 571. Cf. Chapter 3, pp. 133.

[32] Maxwell 1861: 452–5. Cf. Siegel 1991: 56–65.

Maxwell was now on the grounds of his 'On Faraday's lines of force.' Identifying $\mathbf{H}$ and $\mu\mathbf{H}$ with the magnetic intensity and quantity defined there, in the successive terms of eqn. (4.8) he recognized the force acting on the imaginary magnetic masses $\nabla \cdot \mu\mathbf{H}$, the force acting on the current $\nabla \times \mathbf{H}$, and the force responsible for the tendency of paramagnetic (diamagnetic) bodies to move toward places of stronger (weaker) magnetic intensity. Hence Thomson's molecular vortices and the resulting stresses accounted for all known magnetic and electromagnetic forces, with striking mathematical exactitude.[33]

4.3.2 The idle wheels

Maxwell next wondered why a distribution of vortices for which $\nabla \times \mathbf{H}$ did not vanish indicated an electric current. His answer came with the resolution of the following puzzle:

> I have found great difficulty in conceiving of the existence of vortices in a medium, side by side, revolving in the same direction about parallel axes. The contiguous portions of consecutive vortices must be moving in opposite directions; and it is difficult to understand how the motion of one part of the medium can coexist with, and even produce, an opposite motion of a part in contact with it.

Being aware of electromagnetic induction, Maxwell expected the system of vortices to act as a connected mechanism, able to transfer electric motion from one conductor to another. Like his father and his Scottish professors, he was highly interested in practical mechanics. He had read several treatises on this subject, and taught his students the rudiments of kinematics with toothed wheels and cranks. He was surely familiar with the use of 'idle wheels' for transmitting rotation between two toothed wheels without change in the sense of rotation. Accordingly, he somewhat rigidified his fluid vortices and introduced between them a layer of small, round particles that rolled without sliding (Fig. 4.4(a)).[34]

Whenever two contiguous vortices do not rotate at the same speed, the particles between them must shift laterally (Fig. 4.4(b)). For example, if the vortices are parallel to the axis Oz, and if the rotation velocity H_z grows in the direction Ox, the shift occurs in the direction Oy at the rate $-\partial_x H_z$. In general, the shift is given by $\nabla \times \mathbf{H}$, which is equal to the electric current. Maxwell therefore identified the stream of particles with the electric current.[35]

After this purely kinematical analysis, Maxwell examined the dynamics of the new model. As a result of the tangential action $\mathbf{T}$ of the particles on the cells, there

[33] Maxwell 1861: 456–64. Note that the 'quantity' $\mu\mathbf{H}$ differs from the $\mathbf{B}$ of Maxwell's *Treatise* when there are magnets.

[34] Maxwell 1861: 468. Maxwell attended Robert Willis's lectures on mechanism: cf. Maxwell to John Clerk Maxwell, 12 November 1855, *MSLP* 1: 333; and he read a few books on this topic, including Goodeve's *Elements of mechanism* and Rankine's *Applied mechanics* to which he referred in Maxwell 1861: 469n, 458n. On Maxwell's teaching of kinematics, cf. Maxwell to William Thomson, 30 January 1858, *MSLP* 1: 580. On the kinematics of the vortex model, cf. Siegel 1991: 65–9.

[35] Maxwell 1861: 469–71.

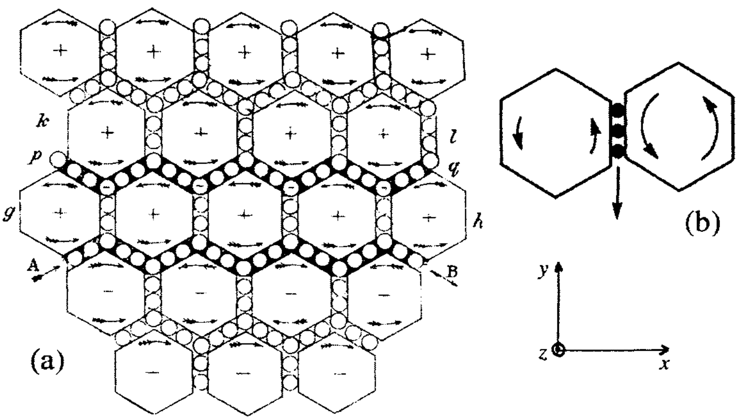

FIG. 4.4. Maxwell's cells and idle wheels (Maxwell 1861: 488 for (a) with mistakes in the arrows from the *MCP* reprint; Siegel 1991: 69 for (b), used by permission of Cambridge University Press).

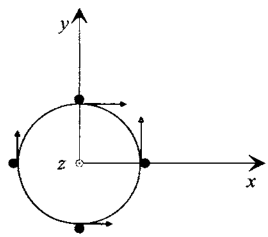

FIG. 4.5. Tangential actions of four idle wheels on a cell.

is a torque acting on each cell. For example, the torque around Oz is proportional to $\partial_x T_y - \partial_y T_x$ (see Fig. 4.5). According to a well-known theorem of dynamics, this torque must be equal to the time derivative of the angular momentum of the cell, which is proportional to $\mu\mathbf{H}$. According to the equality of action and reaction, the force **T** must be equal and opposite to the tangential action of the cell on the particles. Maxwell interpreted the latter action as the electromotive force **E** of magnetic origin acting on the current. In sum, the curl of **E** is found to be proportional to the time derivative of $\mu\mathbf{H}$. The condition that the work of the force **E** on the particles should be globally equal to the decrease of the kinetic energy of the cells determines the coefficient. The final equation of motion is

$$\nabla\times\mathbf{E} = -\frac{\partial\mu\mathbf{H}}{\partial t}, \tag{4.9}$$

in conformity with Maxwell's earlier expression of Faraday's induction law.[36]

Maxwell accompanied his derivation of the fundamental field equations with an intuitive explanation of electromagnetic induction. Consider two conducting circuits separated by an insulator, and let a current be started in one of the circuits. The corresponding flow of particles induces a rotation of the cells immediately outside the conductor. Since in the insulator the particles cannot circulate, they transmit the rotation to the next layer of cells, and so forth until the surface of the second conducting circuit is reached. At this surface the particles are again able to circulate. If there were no electric resistance, they would circulate for ever, and the cells within the conductor would remain at rest. In actual conductors, a frictional force gradually checks the circulation of the particles, and the cells of the conductor are set into rotation. Hence the induced current is only temporary, and the magnetic field in the second conductor is soon the same as it would be in an insulator.[37]

With his wonderful model Maxwell demonstrated the possibility of reducing electromagnetic actions to contiguous mechanical actions. He published his reasoning in the spring of 1861, with a few comments on the awkwardness of the model and on his ignorance of the true nature of electricity. At that time he did not seem to forecast any extension of the model. The obvious limitation to closed currents could not worry him much, since the electrodynamic properties of open currents were experimentally inaccessible.[38]

4.3.3 Electrostatics and light!

A few months elapsed before Maxwell realized that the elasticity of the vortices, which was necessary to their mechanical linking, offered an opportunity to connect electrodynamics with optics and electrostatics. Perhaps a transverse vibration of the substance of the cells could represent light. Perhaps an elastic yielding of the cells under the pressure of the particles could represent dielectric polarization. Specifically, Maxwell imagined that the tangential action of the particles on the cells, which is opposed to the electromotive force by Newton's third law, induced an elastic deformation of the kind represented in Fig. 4.6. Owing to this deformation, the particles in contact with the cells are displaced in a direction opposite to the electromotive force. Calling δ the average displacement, we have

$$\boldsymbol{\delta} = -\varepsilon \mathbf{E}, \tag{4.10}$$

where ε is a constant depending on the elastic constants and on the shape of the cells. The kinematic relation between the flux of particles and the rotation of the cells becomes:

[36] Maxwell 1861: 472–6. Instead of using the theorem of angular momentum, Maxwell used an imperfect energetic reasoning (cf. Darrigol 1993b: note 47). On pp. 479–82 he treated the case of a moving conductor (cf. Darrigol 1993b: 277–9).

[37] Maxwell 1861: 477–8. Cf. Everitt 1975: 96–7.

[38] Cf. Siegel 1991: 75–7; Bromberg 1967: 227; Harman 1970: 191; Everitt 1975: 98–9.

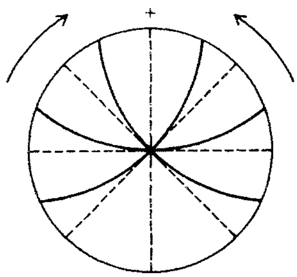

FIG. 4.6. Tangential distortion of a spherical cell (Maxwell to Faraday, 19 October 1861, *MSLP* 1: 684).

$$\mathbf{j} = \nabla \times \mathbf{H} + \frac{\partial \boldsymbol{\delta}}{\partial t}. \quad (4.11)$$

Consequently, the divergence of the current is

$$\nabla \cdot \mathbf{j} = \frac{\partial}{\partial t} \nabla \cdot \boldsymbol{\delta}. \quad (4.12)$$

This equation agrees with the conservation of electricity if the charge density is given by

$$\rho = -\nabla \cdot \boldsymbol{\delta}. \quad (4.13)$$

Although Maxwell does not explictly say so, we may note that ρ also represents an excess of particles, typically occurring at the limit between a conductor and a non-conductor.[39]

Next, Maxwell proceeded to derive the usual electrostatic forces. To this end he considered the elastic energy of the medium,

$$U = \frac{1}{2}\int (-\mathbf{E}) \cdot \boldsymbol{\delta} \mathrm{d}\tau = \frac{1}{2}\int \varepsilon E^2 \mathrm{d}\tau, \quad (4.14)$$

computed it for two point charges q and q', and derived this quantity with respect to the distance d between the charges. The result, $qq'/4\pi\varepsilon d^2$, agreed with Coulomb's

[39] Maxwell 1862: 489–96. For a detailed analysis of the workings of the model, cf. Boltzmann 1898, and Siegel 1986, 1991: 77–119. Most other commentators have misunderstood the mechanics of the model and treated Maxwell's negative sign in the relation between displacement and electromotive force as a mistake. Siegel clarifies this point, and shows how the model accounts for basic electrostatic effects. Some of Maxwell's phrases suggest that he wanted to interpret $\boldsymbol{\delta}$ as a polarization in the Poisson–Mossotti sense. However, in an insulator the displacement of the particles due to the distorsion of the cells must be exactly compensated by a differential rotation of these cells so that the net current $\mathbf{j}$ is zero. As Boltzmann and Siegel argue, the fixity of the particles is essential to the transmission of strain from cell to cell.

law and gave the value of the absolute electrostatic unit of electric charge as $(4\pi\varepsilon_0)^{1/2}$ (the index 0 referring to a vacuum). Consequently, the ratio c of the electromagnetic to the electrostatic charge unit had to be $(\varepsilon_0\mu_0)^{-1/2}$.[40]

At that stage Maxwell had a consistent mechanical model that unified electrostatics and electrodynamics, and he could write the corresponding system of field equations, now called 'the Maxwell equations.' This is not all. He considered transverse waves in the elastic medium. Their velocity is $(k/m)^{1/2}$ if k denotes the transverse elasticity and m the density of the medium. The constant k is inversely proportional to ε, and m is proportional to μ. In order to determine the proportionality coefficients, Maxwell assumed that the cells were spherical and that their elasticity was due to forces between pairs of molecules. He found $k = 1/4\pi^2\varepsilon$, and $m = \mu/4\pi^2$. Then the velocity of transverse waves in a vacuum had to be identical to the ratio c.[41]

Comparing Fizeau's value for the velocity of light and the value of c from Weber and Kohlrausch, Maxwell found agreement within 1% and concluded: 'We can scarcely avoid the inference that *light consists in the transverse undulations of the same medium which is the cause of electric and magnetic phenomena.*' In the same stroke, Maxwell explained the strange proximity of the electromagnetic constant c with the velocity of light and realized Faraday's dream of unifying optics and electromagnetism. Yet the quality of the numerical agreement was accidental. As Duhem pointed out many years later, Maxwell had overlooked a factor of 2 in the transverse elasticity of the cells' substance. In any case, the cells could not be spherical. Moreover, Weber and Kohlrausch's and Fizeau's measurements later proved to be both wrong by 3%. What Maxwell truly had was a rough magneto-mechanical theory of light, based on the elasticity of the substance whose rotation represented the magnetic field.[42]

In the last part of his memoir Maxwell returned to the very phenomenon that had inspired his vortex model, the Faraday effect. The rotation of the cells implied a rotation of the polarization of light in the same direction, by an amount proportional to the radius of the cells. Faraday's observations could be explained if the cells were much smaller in a vacuum than in transparent matter, and if their size depended on the kind of matter. However, Maxwell's model implied that the optical rotation should always be in the direction defined by the magnetic field, whereas Emile Verdet had recently observed an opposite rotation for solutions of iron salts. Maxwell briefly suggested that a proper combination of his cellular model with Weber's molecular currents would explain the anomaly.[43]

[40] Maxwell 1862: 497–9. The stress of the cells, which is linear in **E**, cannot be directly responsible for the electrostatic forces, which require a quadratic stress (see Appendix 6). Maxwell never found a mechanical representation of Faraday's electric stresses (cf. Siegel 1991: 83). Maxwell's notation for c was v. This constant is related to that of Weber's theory by $c = C/\sqrt{2}$.

[41] Maxwell 1862: 499.

[42] Maxwell 1862: 500 (Maxwell's emphasis); Duhem 1902: 208–9, 211–12. Cf. Siegel 1991: 136–41. Bromberg 1967 called Maxwell's theory of light of 1862 'electro-mechanical.' I prefer 'magneto-mechanical' because magnetic vortices were the starting point.

[43] Maxwell 1862: 502–13; Verdet 1854–1863. Cf. Knudsen 1976: 255–8.

4.3.4 An orrery

Maxwell had more to say on the status of his mechanical assumptions. His previous analogies with resisted flow, he recalled, were intended to provide a clear geometrical conception of the lines of force. They did not involve any hypothesis on the deeper nature of electric and magnetic actions. In contrast, his new approach assumed the existence of stresses from which observed mechanical actions derived. The lines of force now referred to these stresses and were therefore as physical as Faraday wanted them to be. Maxwell further adopted Thomson's assumption that the stresses in the magnetic field were due to molecular vortices. These physical hypotheses permitted a unified, dynamical understanding of magnetism and electromagnetism; and they were anchored on the rock of Thomson's argument on the Faraday effect. They remained in the core of Maxwell's theory until his death.[44]

However, Maxwell did not believe in the literal truth of his more specific assumptions regarding the constitution and interconnection of the molecular vortices:

> The conception of a particle having its motion connected with that of a vortex by perfect rolling contact may appear somewhat awkward. I do not bring it forward as a mode of connexion existing in nature, or even as that which I would willingly assent to as an electrical hypothesis. It is however, a mode of connexion wich is mechanically conceivable, and easily investigated, and it serves to bring out the actual mechanical connexions between the known electro-magnetic phenomena; so that I venture to say that any one who understands the provisional and temporary character of this hypothesis, will find himself rather helped than hindered by it in his search after the true interpretation of the phenomena.

Maxwell did not doubt the truth of the relations he had obtained between the electric and magnetic fields, and he believed that these relations derived from the laws of mechanics. But a peculiar combination of vortices and idle wheels could not meet his idea of the simplicity of nature. As he explained to Tait: 'The nature of this mechanism is to the true mechanism what an orrery is to the solar system.'[45]

4.4 The dynamical field

After the publication of 'On physical lines of force,' Maxwell's agenda included the experimental verification of three predictions of his theory. He planned to renew his attempts at detecting gyromagnetic effects. He envisioned precise measurements of the inductive capacity ε of various transparent substances in order to verify the theoretical relation with the optical index ($\varepsilon = n^2$). Most importantly, he intended to verify the identity of the velocity of light with the ratio of absolute electromagnetic and electrostatic charge units by improving on Weber and Kohlrausch's measurement. His enrollment in the British project for electric standards eased this task. In 1864 he imagined an arrangement based on the direct comparison between an elec-

[44] Maxwell 1862: 451–3. Cf. Knudsen 1976: 248–55; Siegel 1991: 39–55.
[45] Maxwell 1861: 486; Maxwell to Tait, 23 December 1867, *MSLP* 2: 337.

trodynamic and an electrostatic force. Four year laters he published the results of a more sophisticated experiment based on the same principle.[46]

Considering that the electromagnetic derivation of the velocity of light was his most important result, Maxwell tried to 'clear the electromagnetic theory of light of all unwarranted assumptions.' The velocity of light could not possibly depend on the shape of vortices or on their kind of elasticity. In 1864 Maxwell managed to reformulate his theory without any specific mechanism and to describe wave propagation in purely electromagnetic terms. In order to understand how he accomplished this, we must return to the electrotonic state.[47]

4.4.1 The reduced momentum

When Maxwell designed the vortex model, he was still looking for a mechanical interpretation of the electrotonic state. He found one of an unexpected sort. Having rewritten the induction law (4.9) as

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} \quad \text{with} \quad \mu\mathbf{H} = \nabla\times\mathbf{A}, \tag{4.15}$$

he noted that **A** played the role of a 'reduced momentum' for the mechanism driven by the flow of particles. He thereby meant a generalization of Newton's second law, in which the force –**E** served to increase the reduced momentum. More concretely, he compared **A** with 'the *impulse* which would act on the axle of a wheel in a machine if the actual velocity were suddenly given to the driving wheel, the machine being previously at rest.' Impulse and momentum were prominent notions in the treatises on mechanism he had been reading, especially Rankine's. Also, impulsive forces played a central role in Stokes's and Thomson's considerations of irrotational flow.[48]

Two simple examples will illustrate what Maxwell had in mind. In the case of a single linear circuit, the current i sets the surrounding cells into rotary motion, as a rack pulled between toothed wheels (Fig. 4.7). If the mass of the axle is negligible, a finite force is still necessary to set it into motion because of the inertia of the connected wheels. By Maxwell's definition, the reduced momentum is the impulse

[46] Maxwell to Thomson, 10 December 1861, *MSLP* 1: 694–8; Maxwell to Thomson, 15 October 1864, *MSLP* 2: 176; Maxwell 1868a. On the gyromagnetic experiments, cf. Maxwell to Faraday, 19 October 1861, *MSLP* 1: 688–9; Maxwell 1861: 485n–6n; Maxwell 1873a: ##574–5; and Galison 1982 for the later history of such effects. For a classification of the devices to measure the units ratio, cf. Jenkin and Maxwell 1863. In Maxwell's 1864 device, the repulsion of two current-fed coils is balanced by the attraction between two electrified disks; the current feeding the coils passes through a resistance of known absolute value; and the potential difference at the ends of this resistance is applied to the disks. On the ensuing project, cf. Schaffer 1995; d'Agostino 1996: 31–6; Simpson 1997: 347–63; Harman 1998: 65–8.

[47] Maxwell to Hockin, 7 September 1864, *MSLP* 2: 164; Maxwell 1865.

[48] Maxwell 1861: 478. I have changed the sign of **A** for consistency with Maxwell's later papers. Reference to Rankine's *Applied mechanics* is found *ibid.*: 458n (for the definition of stresses). On impulsive forces, cf. Moyer 1977: 257–8.

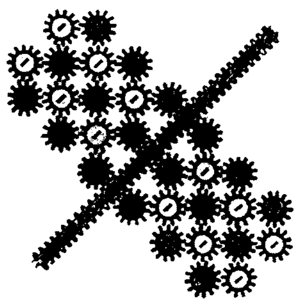

FIG. 4.7. Illustration of self-induction in a linear circuit (from Lodge 1889: 186). The + and – signs indicate the sense of the rotation.

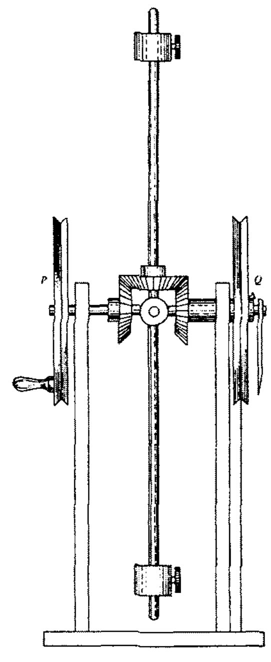

FIG. 4.8. Maxwell's model for mutual induction (Maxwell 1891, Vol. 2: 228).

necessary to obtain a given velocity i. This impulse is proportional to the velocity and to the inertia of the wheels. In electric language, it is equal to Li, where L is the self-inductance of the circuit, and it measures the electro-tonic state.

In the case of two linear cicuits, the rotation of the cells is in a one-to-one correspondence with the two currents i_1 and i_2. This situation is analogous to the mechanism of Fig. 4.8, in which the rotations of the wheels P and Q play the role of the two currents and the rotation of the fly-weights plays the role of the vortex rotation in the magnetic field. The wheels P and Q have a negligible inertia. However, a finite force is in general necessary to set them into motion, because of the inertia of the fly-weights. The reduced momenta at P and Q are the impulses necessary to impart on them the velocities i_1 and i_2. These impulses have the linear forms

$$p_1 = L_1 i_1 + M i_2, \quad p_2 = L_2 i_2 + M i_1. \tag{4.16}$$

They measure the electro-tonic states of the two circuits. Generalizing to a three-dimensional current distribution **J**, the electromotive force necessary to start this current impulsively must be a certain linear function of **J**, to be identified with the electrotonic state **A**.[49]

4.4.2 *Hidden mechanism*

Maxwell reached this mechanical interpretation of the electrotonic state in 1861, on the basis of the vortex model. Three years later, he realized that the interpretation was essentially independent of any specific mechanism and could serve as a more abstract foundation for the dynamics of the magnetic field. He simply admitted that through an unspecified connected mechanism the existence of an electric current implied a motion in the surrounding field. Then, the force necessary to communicate this motion had to be the time derivative of a generalized momentum **A**, which he now called the 'electromagnetic momentum.' In the case of two circuits, this yields the usual equations for inductive coupling (Neumann's)

$$\begin{aligned} e_1' - R_1 i_1 &= \frac{\mathrm{d}}{\mathrm{d}t}(L_1 i_1 + M i_2) \\ e_2' - R_2 i_2 &= \frac{\mathrm{d}}{\mathrm{d}t}(L_2 i_2 + M i_1) \end{aligned} \tag{4.17}$$

where e_1' and e_2' are the impressed electromotive forces and R_1 and R_2 the resistances.[50]

Maxwell next considered the energy brought by the electromotive sources according to Thomson:

$$\begin{aligned} e_1' i_1 + e_1' i_2 = R_1 i_1^2 + R_2 i_2^2 &+ \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}L_1 i_1^2 + \frac{1}{2}L_2 i_2^2 + M i_1 i_2\right) \\ &+ \frac{1}{2}\frac{\mathrm{d}L_1}{\mathrm{d}t} i_1^2 + \frac{1}{2}\frac{\mathrm{d}L_2}{\mathrm{d}t} i_2^2 + \frac{\mathrm{d}M}{\mathrm{d}t} i_1 i_2. \end{aligned} \tag{4.18}$$

[49] Maxwell constructed this model in 1874. Cf. Maxwell 1891: 228, and Everitt 1975: 103–4.
[50] Maxwell 1865: 536–40. Cf. Simpson 1970; Topper 1971; Chalmers 1973; Moyer 1977; Siegel 1981; Buchwald 1985a: 20–3; Hendry 1986: 191–206; Siegel 1991.

The two first terms represent the Joule heat. The third represents the variation of the energy

$$T = \frac{1}{2}(p_1 i_1 + p_2 i_2) \tag{4.19}$$

stored in the hidden mechanism. The three last terms exist only if the geometrical configuration of the circuits varies: they represent the work of electrodynamic forces during this motion. Maxwell thus inverted the prodedure followed by Helmholtz and Thomson; that is, he derived the expression of electrodynamic forces from the laws of induction.[51]

4.4.3 *Lagrangian dynamics*

Maxwell's reasoning appeared in his 'dynamical theory of the electromagnetic field,' published in 1865. With this title he meant to announce a reduction of electrodynamics to hidden motion in the field. In the *Treatise*, published in 1873, he improved his presentation by a recourse to the Lagrange equations. A system of two currents according to Maxwell is a connected system the motion of which is completely defined by two generalized velocities i_1 and i_2. Following Lagrange, the motion of this system is completely determined by the form of its kinetic energy, which is

$$T = \frac{1}{2}L_1 i_1^2 + \frac{1}{2}L_2 i_2^2 + M i_1 i_2, \tag{4.20}$$

for the electrodynamic part. The Lagrange equations with respect to the generalized velocities i_1 and i_2 give

$$e_1' - R_1 i_1 = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial i_1}, \quad e_2' - R_2 i_2 = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial i_2}, \tag{4.21}$$

in conformity with eqns (4.17). If the geometric configuration of the circuits depends on the coordinate ξ, the corresponding Lagrange equation yields the electrodynamic force $\partial T/\partial \xi$ (see Appendix 9 for later three-dimensional generalizations).[52]

Originally, Lagrange designed his analytical method as a way of eliminating the quantities that pertain to the internal connections of a connected mechanical system. Laplacian physicists had little use of the method, since they always started with molecular forces. British physicists were first to appreciate the great power of the method: it gave the equations of motion of a mechanical system by an automatic

[51] Maxwell 1865: 541–2.

[52] Maxwell 1873a: ##578–83. Save for their Lagrangian justification, Maxwell's circuit equations are exactly identical to those of Neumann's theory of induction.

prescription, directly in terms of the controllable elements. For example, in 1837 George Green derived the equations of motion of an elastic solid by expressing its kinetic and potential energy in terms of the local displacements and writing the corresponding Lagrange equations. William Thomson adopted the method, for it shared the virtue of the energy principle of dealing with controllable inputs and outputs. He tried to make it less abstract by combining it with the more physical notions of work and impulse. In his and Tait's *Treatise of Natural Philosophy* (known as TT′), first published in 1867 and proof-read by Maxwell, he defined generalized forces through the work they brought to the system ($\Sigma f_i dq_i$, for a variation dq_i of the generalized coordinates), and the generalized 'momenta' p_i as the impulses necessary to suddenly start the motion of the system from rest. The Lagrange equations, $f_i = dp_i/dt - \partial T/\partial q_i$, thus took a physically transparent form.[53]

Maxwell was very sympathetic to Thomson and Tait's presentation. He developed it in a chapter of his *Treatise*, with the comment: 'We avail ourselves of the labours of the mathematicians [Lagrange and Hamilton], and retranslate their results from the language of the calculus into the language of dynamics, so that our words may call up the mental image, not of some algebraical process, but of some property of moving bodies.' In this process Maxwell was less careful than Thomson and Tait, and erred in a pseudo-derivation of the Lagrange equations based on energy conservation. However, thanks to the new dynamical language he perceived an essential advantage of Lagrange's method: that the motion of the driving points of a connected mechanism could be studied without any knowledge of the internal connections, as some kind of black box. Maxwell used the metaphor of a belfry, the machinery of which is controlled by a number of ropes. The machinery being originally at rest, finite velocities are impressed impulsively on the ropes. If the necessary impulses are measured for every possible value of the positions and final velocities of the ropes, the kinetic energy of the system can be computed as a function of generalized coordinates and velocities (the homogeneity of T implies that $2T = \Sigma p_i dq_i/dt$). Then the motion of the ropes for any applied force is given by the corresponding Lagrange equations.[54]

4.4.4 The electromagnetic momentum

With the momentum interpretation, the vector potential became the central dynamical concept of Maxwell's theory. The induced electromotive force in a circuit was

[53] Green 1838: 246: 'One of the great advantages of this method [of the *Mécanique analytique*], of great importance, is, that we are necessarily led by the mere process of the calculation, and with little care on our part, to all the equations and conditions which are *requisite* and *sufficient* for the complete solution of any problem to which it may be applied'; Thomson to Stokes, 20 October 1847, in Wilson 1990: 32 (for least action applied to impulsively started fluid motion); Thomson and Tait 1867: 217–35. Cf. Siegel 1981: 259–63; Everitt 1975: 105–6; Everitt 1983: 128–9; Buchwald 1985: 60–1; Harman 1987: 287–88; Smith and Wise 1989: 270–3, 390–5 (on TT′).

[54] Maxwell 1873a: #554; Maxwell 1879: 783–84 for the belfry metaphor (for simplicity, I have excluded potential energy). Cf. Moyer 1977; Siegel 1981; Simpson 1970; Topper 1971; Buchwald 1985: 20–3.

just the time derivative of its reduced momentum. Maxwell further assumed that the circuit momentum was the line integral of the 'electromagnetic momentum' **A**. This gives

$$\int \mathbf{E}\cdot d\mathbf{l} = -\frac{d}{dt}\int \mathbf{A}\cdot d\mathbf{l}, \tag{4.22}$$

or

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} + \mathbf{v}\times(\nabla\times\mathbf{A}) - \nabla\phi \tag{4.23}$$

for the electromotive force **E** at a point of the conductor moving with the velocity **v**. Maxwell called ϕ the 'electric potential' and mentioned that it was determined by other conditions of the problem.[55]

In order to relate **A** to the magnetic field, Maxwell followed Faraday's suggestion of defining the magnetic lines of force by the electromotive force induced during their cutting by a linear conductor. Hence the magnetic quantity **B** must be identified with the curl of the electromagnetic momentum **A**. For the determination of **B** in terms of the current **J**, Maxwell used the reasoning of his 'On Faraday lines of force,' which leads to

$$\nabla \times \mathbf{H} = \mathbf{J} \tag{4.24}$$

for the intensity $\mathbf{H} = \mathbf{B}/\mu$.[56]

Maxwell then generalized the expression (4.20) for the kinetic energy of two currents. which gives

$$T = \frac{1}{2}\int \mathbf{J}\cdot\mathbf{A}\,d\tau. \tag{4.25}$$

This expression was most important in the new dynamical theory, for $\frac{1}{2}\mathbf{J}\cdot\mathbf{A}\,d\tau$ meant the energy *controlled* by the current in the volume element $d\tau$. Using $\mathbf{J} = \nabla\times\mathbf{H}$ and a partial integration, it could be transformed back into the expression given by the vortex model,

$$T = \int \frac{1}{2}\mathbf{B}\cdot\mathbf{H}\,d\tau, \tag{4.26}$$

in which $\frac{1}{2}\mathbf{B}\cdot\mathbf{H}\,d\tau$ referred to the energy *stored* in the element $d\tau$.[57]

[55] Maxwell 1865: 555–60.
[56] Maxwell 1865: 550–54, 556–57.
[57] Maxwell 1865: 562–63.

4.4.5 Closing the circuit

The Ampère law (4.24) only applies to divergenceless or closed current. More fundamentally, Maxwell's dynamical reasoning implies the restriction to closed currents, because only in this case is the magnetic field motion completely determined by the currents. If there is any elastic yielding of the field mechanism, as Maxwell assumed in his vortex model, then the motion also depends on the deformation of this mechanism. Maxwell's solution to this difficulty was to change the definition of the electric current. In the vortex model he had defined the current as the flux of particles between the vortices. In his 'dynamical theory,' he tried to follow Faraday's notion that the electric current was a variation or transfer of polarization.[58]

Maxwell first defined the polarization or 'electric displacement' **D** as a displacement of electricity in the molecules of the dielectric, referring here to Mossotti's theory of electrostatic induction. Being elastically resisted, the displacement required an electromotive force $\mathbf{E} = \mathbf{D}/\varepsilon$, and implied a potential energy of the medium

$$U = \frac{1}{2}\int \mathbf{E}\cdot\mathbf{D}\,d\tau. \tag{4.27}$$

A variation of displacement implied an electric current $\partial\mathbf{D}/\partial t$. Electric conduction occurred when electricity was allowed to pass from one molecule to the next at the rate **j**. Hence, in a medium presenting both inductive capacity and conductivity, the total current was

$$\mathbf{J} = \frac{\partial\mathbf{D}}{\partial t} + \mathbf{j}. \tag{4.28}$$

The resulting expression of the Ampère law was the same as that given in 'On physical lines of force':

$$\nabla\times\mathbf{H} = \mathbf{j} + \frac{\partial\varepsilon\mathbf{E}}{\partial t}. \tag{4.29}$$

We must note, however, an important difference of interpretation. In the old theory, what Maxwell called the 'displacement current' was $-\partial\varepsilon\mathbf{E}/\partial t$ and contributed to the conduction current. In the new theory, the displacement current became a contribution to a divergenceless total current.[59]

In conformity with Mossotti's picture of polarization Maxwell took

$$\rho = -\nabla\cdot\mathbf{D} \tag{4.30}$$

[58] Maxwell 1865: 531. [59] Maxwell 1865: 554, 560. Cf. Siegel 1991: 145–52.

to represent the density of 'free electricity.' This brought him into grave difficulties, part of which he solved by reversing the sign in Ohm's law (he took $\mathbf{j} = -\sigma\mathbf{E}$). In fact, his equations were not compatible with the conservation of electricity, as is easily seen by taking the divergence of eqn. (4.28). He was here a victim of his well-known plus-minus dyslexia. He tended to place signs in his equations according to the underlying physical idea, not according to algebraic compatibility. Unfortunately, the physical idea under eqn. (4.30) was incompatible with Faraday's concept of electric charge, as Maxwell later realized.[60]

4.4.6 Electromagnetic light waves

Fortunately, the most important application of the new theory, the derivation of the equation of electromagnetic disturbances in a non-conducting medium, did not depend on the sign of electric charge. Maxwell combined eqns. (4.23), (4.24), and (4.28), and reached, for the magnetic induction,

$$\varepsilon\mu\frac{\partial^2\mathbf{B}}{\partial t^2} = \Delta\mathbf{B}. \tag{4.31}$$

This is a wave equation with the propagation velocity $(\varepsilon\mu)^{-1/2}$. Maxwell also treated the case of a crystalline medium, and determined how conductivity affected transparency. He now had an electromagnetic theory of light *sensu stricto*, since he could describe the waves directly in terms of the electric and magnetic fields. Moreover, his derivation of the velocity of the waves became independent of any assumption on the underlying mechanism.[61]

The electromagnetic momentum **A** being central to his new approach, Maxwell tried to determine how it propagated. Today's physicist knows that **A** is ambiguous: any gradient can be added to it without changing the measurable fields **E** and **H**, provided that a compensating change of the scalar potential is performed. In 1862 Maxwell thought differently. He believed that **A** was unambiguously defined as the impulse necessary to start a given current. Also, he believed that he could maintain the general validity of Poisson's equation ($\Delta\phi + \rho = 0$). On this erroneous assumption he found that the longitudinal part of **A** could not propagate as a wave, in conformity with the transverse character of light waves.[62]

To summarize, by 1865 Maxwell had all the elements of a powerful dynamical theory of the electromagnetic field based on the following principles:

[60] Maxwell 1865: 561. Maxwell reversed the sign in Ohm's law, presumably to mend his theory of electric absorption (*ibid.*: 573–6); but he kept the plus sign in his study of wave absorption by conductors! On the problem of the sign of charge, cf. Siegel 1991: 148–52.

[61] Maxwell 1865: 577–88. Cf. Bork 1966a; Bromberg 1967; Chalmers 1973; Siegel 1991: 152–7.

[62] Maxwell 1865: 580–2. Maxwell was unaware that as the conjugate momentum of the 'velocity' **J**, the potential **A** is ambiguous, because of the constraint $\nabla\cdot\mathbf{J} = 0$ (see Appendix 9). On Maxwell's confusions about the potentials, cf. Bork 1966a: 847–8; Bork 1967; Anderson 1991; Hunt 1991a: 116–17; Cat 1995.

1. Closed currents control a hidden motion in the field.
2. All current are closed.
3. Charge and current derive from polarization, which is an elastic deformation of the medium under electromotive force.

However, the theory was still hampered by confusions regarding the concepts of electromagnetic momentum and dielectric polarization.

4.4.7 Electromagnetic momentum, revised

When in 1868 Maxwell published the results of his new measurement of the ratio c of the electromagnetic to the electrostatic charge unit, he restated the electromagnetic theory of light 'in the simplest form, deducing it from admitted facts, and shewing the connexion between the experiments already described [for the measurement of c] and those which determine the velocity of light.' The 'admitted facts' were Oersted's electromagnetism, Faraday's law of electromagnetic induction, and Faraday's doctrine of polarization. From them Maxwell extracted four simple 'theorems' expressing in words the integrals of the magnetic and electric intensities on closed curves, the relation between electric intensity and displacement, and the displacement current. All reference to the electromagnetic momentum was gone, and the deduction of electromagnetic plane waves became quite elementary.[63]

Maxwell could not, however, renounce the dynamical foundation of his theory. It was an essential part of his later *Treatise*, in the Lagrangian form already described. There he acknowledged the gap in the definition of the electromagnetic momentum $\mathbf{A}$, and introduced the condition $\nabla \cdot \mathbf{A} = 0$ as a convenient way to remove the ambiguity. With this choice the momentum of a given current in a medium of uniform permeability μ became

$$\mathbf{A}(\mathbf{r}) = \frac{\mu}{4\pi} \int \frac{\mathbf{J}(\mathbf{r}')\mathrm{d}\tau'}{|\mathbf{r}-\mathbf{r}'|}. \tag{4.32}$$

The analogy with the scalar potential in a uniform dielectric,

$$\phi(\mathbf{r}) = \frac{1}{4\pi\varepsilon} \int \frac{\rho(\mathbf{r}')\mathrm{d}\tau'}{|\mathbf{r}-\mathbf{r}'|}, \tag{4.33}$$

justified the alternative name 'vector potential' for $\mathbf{A}$.[64]

Unfortunately, in his derivation of the equation for the propagation of electromagnetic disturbances, Maxwell repeated the error of considering the scalar potential formula (4.33) as generally valid, independently of the choice of $\nabla \cdot \mathbf{A}$. In fact

[63] Maxwell 1868: 138. Cf. Everitt 1975: 108–9; Hendry 1986: 220–6; Siegel 1991: 153–4. This simple formulation of Maxwell's theory was largely unnoticed until its reprint by Niven in *MSP* in 1890 (I thank Bruce Hunt for this remark).

[64] Maxwell 1873a: #617.

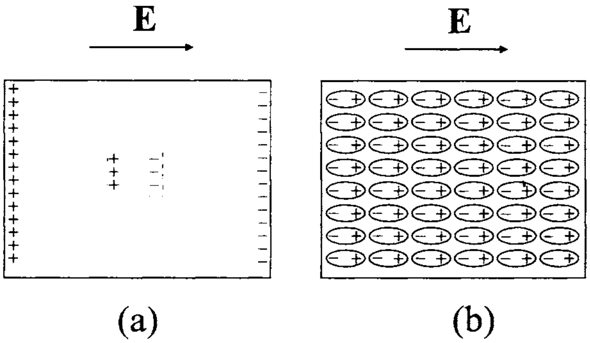

FIG. 4.9. Polarization: (a) according to Maxwell, (b) according to Mossotti.

there is no evidence that he ever combined the equation $\nabla \cdot \varepsilon \mathbf{E} = \rho$ with the electromotive force formula (4.23), except in the electrostatic case. He does not seem to have fully realized that his general assumptions on the electromagnetic field implied a much deeper interconnection of electrostatic and electrodynamic actions than was assumed in continental theories.[65]

4.4.8 Displacement, revised

However, Maxwell managed to clear up his concept of polarization. In the *Treatise* he adopted the positive sign in the relation $\rho = \nabla \cdot \mathbf{D}$, which means that a portion of polarized dielectric is charged positively where the polarization starts and negatively where the polarization ends (Fig. 4.9(a)). This choice agrees with Faraday's definition of positive charge as the starting point of electric lines of force, but contradicts Mossotti's picture of displaced electric charge (Fig. 4.9(b)). Maxwell, like Faraday, avoided the contradiction by considering the concept of polarization as more primitive than the concept of charge. If anything was displaced in the elements of a polarized dielectric, it could not be electric charge. This is an essential point, which must always be kept in mind when reading Maxwell's difficult sections on charge and current.[66]

Maxwell's views then appear to be very similar to Faraday's. Polarization (Faraday's 'induction') is defined as a state of constraint of the dielectric, such that each portion of it acquires equal and opposite properties on two opposite sides. By definition, electric charge is a spatial discontinuity of polarization. Typically, charge

[65] Maxwell 1873a: #783. Several of Maxwell's early readers, including Larmor and J. J. Thomson, inherited Maxwell's confusion on the definition of ϕ. In the same vein, Maxwell gave $-\rho\nabla\phi$ for the force acting on electrified matter (#619), which could be true only in the electrostatic case and contradicted his expression of electric stresses (#108) (see Appendix 6). This mistake is corrected in FitzGerald 1883b, and in the third edition of the *Treatise*.

[66] Maxwell 1873a: ##60–2, #111. See also Maxwell to Thomson, 5 June 1869, *MLSP* 2: 485–6. For a lucid account, cf. Buchwald 1985: 23–34; also Knudsen 1978.

occurs at the limit between a polarized dielectric and a conductor, because by definition a conductor is a body that cannot sustain polarization. As Maxwell explains, 'the electrification at the bounding surface of a conductor and the surrounding dielectric, which on the old theory was called the electrification of the conductor, must be called in the theory of induction the superficial electrification of the surrounding dielectric.'[67]

Conductors cannot sustain polarization. However, they may transfer polarization. This transfer, according to Faraday and Maxwell, results from a competition between polarization build-up and decay in the conductor. A conductor thus appears to be a yielding dielectric: 'If the medium is not a perfect insulator,' Maxwell writes, 'the state of constraint, which we have called electric polarization, is continually giving way. The medium yields to the electromotive force, the electric stress is relaxed, and the potential energy of the state of constraint is converted into heat.' By definition, the electric current is the rate of transfer of polarization. In a dielectric, it is simply measured by the time derivative of the polarization. In a conductor, it also depends on the decay mechanism, the microscopic details of which are unknown. Its expression must therefore be determined empirically (by Ohm's law). Thus defined, the electric current is always closed, for the current in an open conducting circuit is continued through the dielectric.[68]

All of this is quite consistent, and does not involve any of the absurdities later denounced by Maxwell's continental readers. Yet Maxwell's terminology was truly misleading. He called the polarization of a portion of dielectric 'a displacement of electricity.' By this phrase he only meant that a portion of the dielectric, if separated in thought from the rest of the dielectric, would present opposite charges at two opposite extremities. He certainly did not mean that an electrically charged substance was displaced. However, many of his readers understood just that. To make it worse, Maxwell asserted that 'the motions of electricity are like those of an *incompressible* fluid.' Here he only meant that the closed character of the total current made it analogous to the flow of an incompressible fluid. But he was often lent the opinion that electricity *was* an incompressible fluid.[69]

As long as it is used with proper care, the fluid analogy is useful to illustrate the relations between displacement, charge, and conduction. Suppose an incompressible fluid to pervade a space in which a rigid scaffolding has been erected. In 'insulating' parts of this space, the portions of the fluids are elastically linked to the scaffolding. In a 'conducting' part, such links also exist, but when under tension they tend to break down and dissipate their energy into heat; every breaking link is immediately replaced by a fresh, relaxed link. In this illustration, the extension of the links corresponds to Maxwell's displacement (or polarization); the pressure gradient of the fluid to the electromotive force; the flow of the fluid to the electric current; and the discontinuity of the average extension of the links when crossing the limit between conductor and insulator corresponds to electric charge. The analogy

[67] Maxwell 1873a: #60, #111. [68] Maxwell 1873a: #111. Cf. Buchwald 1985a: 28–9.
[69] Maxwell 1873a: #61. See also the fluid-piston illustration of a dielectric, *ibid.* in #334.

properly illustrates the equations $\mathbf{D} = \varepsilon\mathbf{E}$, $\nabla \cdot \mathbf{J} = 0$, $\mathbf{J} = \mathbf{j} + \partial\mathbf{D}/\partial t$, and $\rho = \nabla \cdot \mathbf{D}$. It is, however, misleading when one comes to propagation problems and energy flow, as we will later see.

4.5 *Exegi monumentum*

Around 1867 Maxwell set himself to work on a major treatise on electricity and magnetism. His intention was partly to propel his new theory and Faraday's underlying views. There also was an urgent need for that kind of book. Although the field of electricity and magnetism had grown enormously since Oersted and Ampère, there was as yet no unified presentation of all its experimental, technical, and mathematical aspects. The gap had widened between the practical electricity of telegraphists and the mathematical electricity of learned professors. There was a growing multiplicity of terms, conventions, and theories; and little attempt at uniformization and comparison, despite the high intellectual and economical stakes.[70]

Maxwell was especially sensitive to the neglect of the quantitative aspects of the subject. He believed that the mathematical theories of electricity and magnetism were ripe to be taught in the university, and pressed the Cambridge authorities to introduce them in the Mathematical Tripos. Only the proper reference book was missing, as Maxwell himself judged:[71]

> There are several treatises in which electrical and magnetic phenomena are described in popular way. These, however, are not what is wanted by those who have been brought face to face with quantities to be measured, and whose minds do not rest satisfied with lecture-room experiments.—There is also a considerable mass of mathematical memoirs which are of great importance in electrical science, but they lie concealed in the bulky Transactions of learned societies; they do not form a connected system; they are of very unequal merit, and they are for the most part beyond the comprehension of any but professed mathematicians.—I have therefore thought that a treatise would be useful which should also indicate how each part of the subject is brought within the reach to methods of verification by actual measurement.

Books on electricity were indeed few, and failed to provide a full, systematic introduction to the subject. Auguste de la Rive's *Traité d'électricité* of 1853 was very empirical, had almost no mathematics, and ignored or misrepresented Faraday's theoretical views. Gustav Wiedemann's *Lehre vom Galvanismus* of 1863 gave precise and clear accounts of nearly all works published on the subject, with a fair share of the British views; but its encyclopedic scope and structure made it unsuited to the guidance of students. Maxwell's *Treatise*, published in 1873, filled a major gap in the existing literature.[72]

[70] On the gap between practical and academic electricity, cf. the introduction of Jenkin 1873.

[71] Maxwell 1873a: ix. On the 1867 reform of the Cambridge Mathematical Tripos and on the editorial circumstances of Maxwell's project, cf. Achard 1998.

[72] On the genesis of the *Treatise*, cf. Harman 1995a: 26–33.

4.5.1 Mathematical and empirical foundations

Maxwell's challenge was to expound a new doctrine and at the same time to establish new standards in the treatment of current problems. In order to meet these two conflicting requirements, he carefully separated the basic mathematical and empirical foundations of the subject from more speculative theory. In a preliminary 'on the measurement of quantities' he expounded Fourier's doctrine of dimensions, Hamilton's distinction between scalar and vector, the notions of force and flux corresponding to his older 'intensity' and 'quantity,' various theorems relating the integrals of force and flux, and related topological questions. He regarded the classification of physico-mathematical quantities as a way to short-circuit formal analogies and organize the field of knowledge: 'It is evident that [. . .] if we had a true mathematical classification of quantities,' he had earlier explained, 'we should be able at once to detect the analogy between any system of quantities presented to us and other systems of quantities in known sciences, so that we should lose no time in availing ourselves of the mathematical labours of those who had already solved problems essentially the same.'[73]

Maxwell then defined the basic physical quantities in a neutral manner that could be accepted both by fluid and field theorists. For example, he introduced the quantity of electric charge of a body by means of Faraday's hollow conductors: two charges could be added by bringing their carriers into a hollow conducting vessel and noting the charge of the vessel. He defined the electric potential in Thomson's manner, as the work done on a unit point charge to bring it at a given place. Lastly, he defined the magnetic force **H** and flux **B** in a polarizable substance as the forces acting on a magnetic unit pole (end of uniformly magnetized needle) placed in a small cylindrical cavity, elongated for **H**, and flat for **B**.[74]

With these neutral definitions, Maxwell could conduct much of the mathematical analysis without deciding the nature of electricity and magnetism. This can be seen in his Thomsonian presentation of the potential theories of electrostatics and magnetism. The *Treatise* was in part meant as a source book for computational and experimental techniques for competent electricians, whatever they might think of the essence of electricity. The originators of these techniques were as diverse as their potential users. They could be Laplace on spherical harmonics, Gauss on geomagnetism, Weber on galvanometric measurements, Kirchhoff on circuit theory, Thomson on electrometers, or Maxwell himself on the calculation of inductance.

[73] Maxwell 1873a: ##1–26; Maxwell 1870: 258 (quote). Cf. Harman 1987: 278–87. On dimensions, cf. Jenkin and Maxwell 1863; Everitt 1975: 100–1; d'Agostino 1996: 37–41. On topology, cf. Epple 1998; Harman 1998: 153–6.

[74] Maxwell 1873a: #34 and #63 (for charge), #70 (for potential), ##398–400 (for **B** and **H**). Maxwell and Thomson disagreed on the definition of the electrostatic potential when contact between different metals was involved: cf. Hong 1994a. Maxwell's **B** and **H** corresponded to Thomson 'electromagnetic' and 'polar' definitions of the magnetic field (see Chapter 3, p. 130); however, they referred to two different physical concepts (flux and force), whereas Thomson only meant two different ways of characterizing the same physical entity: cf. Wise 1981a.

These techniques could serve the Mathematical Tripos, German seminars, and telegraphists in the whole industrialized world.[75]

4.5.2 Tolerance

Once equipped with operational definitions and phenomenologico-mathematical theories, the reader of the *Treatise* could enter the realm of higher theory. Maxwell presented the field view, the fluid view, and the relations between the two. Of course, he preferred Faraday's field conception. Compared with the fluid conception, he wrote, it is 'no less fitted to explain the phenomena, and [. . .] though in some parts it may appear less definite, corresponds, as I think, more faithfully with our actual knowledge, both in what it affirms and in what it leaves undecided.' In private, he made fun of the 'learned Germans,' the 'heavy German writers,' or Ampère's 'kind of ostensive demonstration.'[76]

Yet the *Treatise* paid due respect to 'the Newton of electricity' (Ampère) and to the 'eminent' Germans who cultivated action at a distance; and it expounded their theories in sufficient details. This was not only diplomacy: as we will later see, Maxwell integrated some of Ampère's and Weber's atomistics into his own theory. Also, he believed that much could be learn from the comparison between the two kinds of theory:

> In a philosophical point of view, moreover, it is exceedingly important that two methods should be compared, both of which have succeeded in explaining the principal electromagnetic phenomena, and both of which have attempted to explain the propagation of light as an electromagnetic phenomenon [more on this later], and have actually calculated its velocity, while at the same time the fundamental conceptions of what actually takes place, as well as most of the secondary conceptions of the quantities concerned, are radically different.[77]

4.5.3 Field basics

For the essentials of field theory, Maxwell followed Faraday closer than he had ever done. As we have seen, he adopted the field-based definitions of electric charge and current, the concept of conduction as the competition between polarization build up and decay, and the reduction of all electric and magnetic actions to stresses in the field. Even the idea that all currents are closed can be traced back to Faraday's idea of the indivisibility of the electric current (cf. Chapter 3, p. 91). Lastly, Maxwell renounced his earlier theory of magnetism, in which the 'quantity' **B** had sources in

[75] Cf. Maxwell 1873a: Vol. 1, Part 1, Ch. 4 ('General theorems' of potential theory); 2.3.3 (on Thomson's 'Magnetic solenoids and shells'); 1.1.9 ('Spherical harmonics'); 2.3.8 ('Terrestrial magnetism'); 2.4.15 ('Electromagnetic instruments'); 1.2.6 ('Mathematical theory of the distribution of electric currents'); 1.1.13 ('Electrostatic instruments'); 2.4.13 ('Parallel currents').

[76] Maxwell 1873a, Vol. 1: xii; Maxwell to John Clerk Maxwell, 5 May 1855, *MSLP* 1: 294; Thomson to Tait, 1 December 1873, *MSLP* 2: 947; Maxwell to Thomson, 13 November 1854, *MSLP* 1: 255.

[77] Maxwell 1873a: #528; *ibid.*, Vol. 1: xii; Vol. 2, Part 4, Ch. 2 ('Ampère's investigation . . .'); Vol. 2, Part 4, Ch. 13 ('Theories of action at a distance'); Vol. 1: xii.

the magnetic masses of magnets as the electric quantity **D** had sources in charged bodies. In his new theory, **B** was always divergenceless, in conformity with Faraday's notion of magnetic lines of force and with Thomson's flat-cylinder-cavity definition.[78]

For the field equations, Maxwell also depended on Thomson's field mathematics, on the distinction between force and flux, and on the interpretation of Lagrange's equations in terms of energy, force, and momentum. In short, from Faraday's notion of dielectric polarization Maxwell derived the equations

$$\mathbf{D}=\varepsilon\mathbf{E},\quad \nabla\cdot\mathbf{D}=\rho,\quad \mathbf{J}=\mathbf{j}+\frac{\partial\mathbf{D}}{\partial t} \tag{4.34}$$

for the electric force **E**, the electric displacement **D**, the total current **J**, and the conduction current **j**. From his own theory of magnetization and from the equivalence between an infinitesimal current loop and a magnetic dipole, he deduced

$$\mathbf{B}=\mathbf{H}+\mathbf{I},\quad \nabla\cdot\mathbf{B}=0,\quad \nabla\times\mathbf{H}=\mathbf{J} \tag{4.35}$$

for the magnetic force **H**, the magnetic induction **B**, and the intensity of magnetization **I**. From the Lagrangian dynamics of closed currents he obtained

$$\mathbf{E}=\frac{\partial\mathbf{A}}{\partial t}+\mathbf{v}\times(\nabla\times\mathbf{A})-\nabla\phi,\quad \mathbf{f}=\mathbf{J}\times\mathbf{B}, \tag{4.36}$$

where **A** is the electromagnetic momentum, **v** the velocity of the current carrier, and **f** the electrodynamic force acting on the current carrier. The first formula gives Faraday's induction law if **A** is the vector potential such that $\mathbf{B}=\nabla\times\mathbf{A}$. Maxwell further imposed $\nabla\cdot\mathbf{A}=0$ in order to simplify the relation between **A** and the total current. Lastly, in the absence of a specific mechanism for the decay of displacement, he admitted Ohm's law $\mathbf{j}=\sigma\mathbf{E}$. In a separate chapter, he gave the formula

$$w=\frac{1}{2}\mathbf{E}\cdot\mathbf{D}+\frac{1}{2}\mathbf{B}\cdot\mathbf{H} \tag{4.37}$$

for the energy density of the field (in the absence of permanent magnetism) and the expression

$$\sigma_{ij}=D_iE_j-\frac{1}{2}\delta_{ij}\mathbf{E}\cdot\mathbf{D}+B_iH_j-\frac{1}{2}\delta_{ij}\mathbf{H}\cdot\mathbf{B} \tag{4.38}$$

[78] Maxwell 1873a: Part 3, Ch. 2: 'Magnetic force and magnetic induction.'

for the stresses in the field. In principle, all mechanical forces of electric or magnetic origin could be derived from these stresses (see Appendix 6).[79]

4.5.4 Physical ideas and equations

The sheer number of equations (especially in Cartesian notation) was likely to scare Maxwell's reader. So as to please 'the Chief Musician upon Nabla' (his friend Tait), and for the sake of mathematical power and beauty, Maxwell also wrote his equations in quaternion form.[80] This could hardly help the average reader, as Maxwell himself suspected. A more pedagogical step would have been to eliminate the potentials. Maxwell refused to do so in the *Treatise*, arguing that 'to eliminate a quantity which expresses a useful idea would be rather a loss than a gain in this stage of our enquiry.' He wanted to place the electromagnetic momentum at the forefront of his theory.[81]

In general, Maxwell's use of mathematical symbolism differed essentially from continental or modern practice. For him the equations were always subordinated to the physical picture. He sought consistency, completeness, and simplicity in the picture, not necessarily in the equations. The latter were symbolic transcriptions of partial aspects of the picture, and therefore could not be safely used without keeping the underlying picture in mind. This is quite visible in the way Maxwell treated the electrodynamics of moving bodies. His equations included electromagnetic induction in moving bodies, but not other effects of motion that resulted from his pictures of charge and current. For example, he knew that the convection of electrified bodies constituted an electric current because of the corresponding variation of displacement; but there was no convection current in his equations.

4.5.5 Microphysics

Maxwell had another reason not to seek algebraic completeness. He was aware that his theory was essentially incomplete in its treatment of the relation between ether and matter. The general pictures of dielectric and magnetic polarization, and also the idea of currents controlling a hidden motion, implied that ether and matter behaved as a single medium with variable inductive capacity, permeability, and conductivity. Maxwell admitted, however, that some phenomena required a closer look at the interaction between ether and matter. First of all, his picture of electric conduction left the mechanism of polarization decay in the dark. Like Faraday he hoped

[79] Maxwell 1873a: #68, #83, #610 (for eqns. 4.34); #400, #403, #607 (for eqns. 4.35); #598, #603 (for eqns. 4.36); #241 (for Ohm's law); #630, #634 (for eqn. 4.37); #108, #641 (for eqn. 4.38). Maxwell recapitulated the field equations in ##237–38.

[80] Maxwell 1873a: #17, #25, #619. 'Nabla' is an Assyrian harp, of the same shape as Hamilton's ∇; at the BA meeting of September 1871, Maxwell dedicated a poem to Tait, the 'Chief Musician upon Nabla': cf. Campbell and Garnett 1882: 634–6. On the history of quaternions, cf. Crowe 1967. On their use by Maxwell, cf. Harman 1987: 279–82, 1994: 29–30; 1998: 145–53; McDonald 1965; and the related manuscripts and letters in *SMLP* 2.

[81] Maxwell 1873a: #615.

that the study of electric glows, and especially that of electrolysis would shed light on the deeper nature of electricity: 'Of all electrical phenomena,' he declared, 'electrolysis appears the most likely to furnish us with a real insight into the true nature of the electric current, because we find currents of ordinary matter and currents of electricity forming essential parts of the same phenomnenon.'[82]

In his chapters on electrolysis, Maxwell did not follow Faraday's phenomenological approach. As a believer in atomistics, he found it

> extremely natural to suppose that the currents of the ions are convection currents of electricity, and, in particular, that every molecule of the cation is charged with a certain fixed quantity of positive electricity, which is the same for the molecules of all cations, and that every molecule of the anion is charged with an equal quantity of negative electricity.

This assumption accounted for Faraday's law, and could be perfected to explain electrode polarization. But the quantization of charge puzzled Maxwell. It seemed to suggest the existence of 'molecules of electricity,' as if electricity were a discrete fluid. Maxwell bore the contradiction, though not in silence: 'This phrase, gross as it is, and out of harmony with the rest of this treatise, will enable us at least to state clearly what is known about electrolysis, and to appreciate the outstanding difficulties.' He regarded the propounded theory as a provisional mnemonic aid: 'It is extremely improbable that when we come to understand the true nature of electrolysis we shall retain in any form the theory of molecular charges, for then we shall have obtained a secure basis on which to form a true theory of electric currents, and so become independent of these provisional theories.'[83]

Yet Maxwell did not doubt that molecular structure played a role in conduction. He also approved Weber's theory of induced magnetism, which required the existence of permanently magnetized molecules. And he took Ampère's and Weber's molecular currents quite seriously. In his opinion, Verdet's finding that magneto-optical rotation had opposite signs in diamagnetic and ferromagnetic bodies excluded Faraday's doctrine that a diamagnetic was nothing but a lesser conductor of magnetism than vacuum. The *Treatise* had a chapter devoted to the improvement of Weber's theory of ferromagnetism, and another to 'the electric theories of magnetism,' including Weber's induced molecular currents. Maxwell emphasized the simplification of the magnetic field equations when all magnetism was reduced to electromagnetism: molecular currents thus became the only sources, and the fields **B** and **H** became identical and divergenceless.[84]

Maxwell also believed that the molecular structure of matter played a role in the propagation of light. He did not trust his field equations for high-frequency vibrations in material bodies. In dielectrics, these equations did not include optical dispersion and implied a relation between optical index and inductive capacity ($n = \varepsilon^{1/2}$)

[82] Maxwell 1873a: #55 (glow), #255 (electrolysis). [83] Maxwell 1873a: #255, #260.

[84] Maxwell 1873a: Vol. 2, Part 3, Ch. 6 ('Weber's theory of magnetic induction,' with a modification explaining residual magnetization); Maxwell to Tait, 23 December 1867, in *SLPM* 2: 336, and Maxwell 1873a: #809 (Verdet excluding Faraday); Vol. 2, Part 4, Ch. 22 ('Electric theory of magnetism'); #835 (simplicity of Amperean view).

that seemed to hold only very roughly. In conductors, they predicted an absorption of light much larger than that measured on gold leaves. In such cases, Maxwell judged, 'our theories of the structure of bodies must be improved before we can deduce their optical properties from their electrical properties.'[85]

Maxwell's equations did not contain the Faraday effect either: their linearity excluded any action of an external magnetic field on the propagation of light. Remember, however, that Maxwell had earlier given a theory of the Faraday effect, based on his vortex model of the magnetic field. In the *Treatise*, he extracted from this model the basic idea of a magnetic vortex motion perturbing the optical vibrations, and cast it in Lagrangian form. He borrowed the unperturbed part of the Lagrangian from the elastic solid theory of light, and assumed the simple expression $k(\nabla \times \partial\boldsymbol{\xi}/\partial t)\cdot\mathbf{H}$ for the energy density of the magneto-optical interaction; $\boldsymbol{\xi}$ represents the elastic displacement of the medium, $\nabla \times \boldsymbol{\xi}$ twice the corresponding rotation, and $k\mathbf{H}$ the vortical motion implied by the magnetic force $\mathbf{H}$. The latter differs from the impressed magnetic force $\mathbf{H}_0$ by the amount $(\mathbf{H}_0\cdot\nabla)\boldsymbol{\xi}$ if the vorticity depends on the displacement $\boldsymbol{\xi}$ in the manner implied by Helmholtz's theory of vortex motion ($\mathbf{H}\cdot d\mathbf{S}$ invariant). Then the optical Lagrangian involves a new term combining $\boldsymbol{\xi}$ and $\partial\boldsymbol{\xi}/\partial t$, from which the magneto-optical rotation is easily deduced. With this semi-phenomenological reasoning, Maxwell avoided atomistic speculation and absorbed the whole effect of matter into one coupling constant, whose value and sign were to be drawn from Verdet's measurements.[86]

This theory of the Faraday effect, and all of Maxwell's attempts to specify the relation between ether and matter, were meant to be provisional. The macroscopic character of his unification of electrodynamics, electrostatic, and optics, conflicted with the empirical need to introduce the molecular structure of matter. Maxwell did not know to what extent his electromagnetic concepts applied at the molecular scale. He avoided microphysical considerations whenever the macroscopic approach proved sufficient.

4.6 Conclusions

Proceeding from Faraday's and Thomson's writings, Maxwell reached the essentials of his electromagnetic field theory stepwise, in three great memoirs. In 'On Faraday's lines of force' his aim was to obtain a mathematical expression of Faraday's field conception. He found the methods of Thomson's field mathematics particularly useful, but modified them substantially. Thomson gave the electric and magnetic (scalar) potentials a central role, as neutral mediators between the mathematics of action at a distance and Faraday's field reasonings. Instead Maxwell made

[85] Maxwell 1873a: ##788–9 ($\varepsilon \sim n^2$ and dispersion); #800 (gold sheets) and also transparency of electrolytes in #799; #789 (quote). Maxwell gave a molecular theory of anomalous dispersion in a Tripos question of 1868 (*SLMP* 2: 419–21, and Rayleigh 1899), also in an 1873 manuscript (*SLMP* 2: 461–2): see Whittaker 1951: 262; Buchwald 1985a: 236; Harman 1994: 11–12.

[86] Maxwell 1873a: #822–7 (Maxwell also included Cauchy's dispersion terms). Cf. Knudsen 1976: 278–81.

the lines of force the central concept of his theory. He threw a geometrical net of lines of force and orthogonal surfaces over Faraday's field, and caught the mathematical field laws directly in terms of the field quantities. He also used Thomson's flow analogy, and extracted from it an essential structural component of his theory: the distinction between intensity and quantity (force and flux). With these modifications of Thomson's methods, Maxwell invented a powerful field-gridding geometry and obtained two circuital laws $\nabla \times \mathbf{H} = \mathbf{j}$ and $\mathbf{E} = -\partial\mathbf{A}/\partial t$ that captured Faraday's intuition of the relations between electricity and magnetism.

In the first part of 'On physical lines of force' Maxwell exhibited a mechanical model of the magnetic field that closely followed Thomson's insights into the vortical nature of magnetism. Unlike the previous flow analogy, this model accounted for the mechanical forces of magnetic origin and for electromagnetic induction. Maxwell soon modified it to include electrostatics and optics, in a manner totally unforeseen by Thomson. This gave the displacement current, the full set of Maxwell's equations, and an expression of the velocity of light in terms of electromagnetic quantities. Although Maxwell acknowledged the artificiality of his model, he firmly believed in the reality of two features: the mutually connected vortical rotations, and the elastic yielding of the connecting mechanism. The rotations represented the magnetic field, and the elastic yielding the electric field (displacement).

In his 'dynamical theory of the electromagnetic field,' Maxwell replaced his vortex model with a dynamical justification of his field equations. He treated the magnetic field as a hidden mechanism, whose motion was controlled by the electric current. The potential **A** thus acquired a central importance as the reduced momentum of the field mechanism dragged by the electric current. Maxwell combined his field equations to obtain a wave equation, and reached a truly electromagnetic optics in which light became a waving electromagnetic field.

The dynamical approach required that the magnetic motion should be determined by the currents only. Accordingly, Maxwell made the displacement current part of the total current. This move brought him closer to Faraday's concepts of charge and current. In the vortex model, the electric current corresponded to the flow of the particles between the vortices and charge to their accumulation. In the new dynamical theory, and more definitely in the *Treatise*, Maxwell defined the electric current as a transfer of polarization, and charge as a discontinuity of polarization. Here polarization was a primitive concept: any attempt to interpret it as a microscopic displacement of electric charge led to absurdities. Maxwell's theory was a pure field theory, ignoring the modern dichotomy between electricity and field.

In the mature form of the *Treatise*, Maxwell's theory had a central core founding the general theory of the electromagnetic field, and a periphery dealing with less understood phenomena. The core contained the pure field theory of electricity with field-based concepts of charge and current, a dynamical derivation of the equations of motion by the Lagrangian method, and the essentials of the electromagnetic theory of light. The periphery included fragmentary mechanisms for the various kinds of electric conduction, and special theories of magnetization and magneto-optical rotation.

The core was essentially macroscopic, in the sense that the basic concepts of field, charge, and current had a macroscopic meaning. It treated matter and ether as a single continuous medium with variable macroscopic properties (specific inductive capacity, magnetic permeability, and conductivity), and avoided speculation on ether models and matter molecules. At the periphery, Maxwell recognized the need for a more detailed picture of the connection between ether and matter. He tried three different strategies. For magnetization, he modified his theory to integrate molecular assumptions; for electrolysis, he proposed a temporary ionic theory that contradicted his general concept of the electric current; for the Faraday effect, his method was essentially based on a phenomenological modification of the optical Lagrangian, although he invoked a deeper molecular mechanism.

By rejecting direct action at a distance and electric fluids, Maxwell distanced himself from continental physics. Whether he did so in a consistent manner has been a major question for Maxwell's commentators. Recent scholarship has established that Maxwell was far more consistent than has usually been admitted. As Siegel has shown in detail, Maxwell's vortex model holds together very well and accounts for all electrodynamic and electrostatic phenomena known to Maxwell. Most of the inconsistencies perceived by earlier commentators of this model can be traced to their failure to distinguish the relevant concepts of charge and current from those proposed in the *Treatise*.[87] Admittedly, there were genuine inconsistencies in the memoir on the dynamical theory due to the unwarranted mixture of Faraday's and Mossotti's concepts of polarization. In the form given in the *Treatise*, however, Maxwell's concepts of charge and current were quite consistent, as Buchwald has most clearly shown. Here Maxwell's readers were often misled by the metaphor of 'displacement of electricity,' which seems to indicate a shift of electric charge (as occurs in the continental concept of polarization), whereas Maxwell only meant something analogous to the shift of a *neutral* incompressible fluid. Charge is not what is displaced, it is a spatial discontinuity in the strain implied in the 'displacement.' As will be seen in the next chapter, the consistency of Maxwell's views comes out clearly in the more pedagogical presentations offered by Maxwell's followers.

Another logorrhea of Maxwellian scholarship has been about the origin of the displacement current. The excessive focus on this question has resulted in a misrepresentation of Maxwell's overall endeavors and achievements in electric topics. As Wise pointed out, Maxwell's first major innovations were an essentially new geometrization of Faraday's and Thomson's field concepts, and the important distinction between quantity and intensity. The former yielded Maxwell's form of the Ampère law ($\nabla \times \mathbf{H} = \mathbf{j}$), and the latter prepared the ground for the dynamical theory. As for Maxwell's path to the displacement current, it may be summarized as follows.

When Maxwell worked out Thomson's notion of a vortical motion in the mag-

[87] Also, some of them were unable to understand the mechanics of the model.

netic field, he introduced the idle wheels as a direct illustration of the current being the curl of the magnetic force. The original purpose of this mechanism was purely electrodynamic. Maxwell knew, however, that both Faraday's electrostatics and the wave theory of light required an elastic medium. He also knew that the mechanical consistency of his model required an elasticity of the rotating cells. When he took this elasticity into account, he found it to imply a new contribution $-\partial\varepsilon\mathbf{E}/\partial t$ to the current $\mathbf{j}$ of idle wheels. The corresponding modification of the Ampère law allowed for open currents.

In such a dense argument, it would be vain to single out a specific reason for Maxwell's introduction of the displacement current. He sought the most complete and consistent theory that would comply with a number of entangled conditions: expression in terms of Faraday's lines of force and the related intensity/quantity pairs ($\mathbf{E}$, $\mathbf{D}$) and ($\mathbf{H}$, $\mathbf{B}$), existence of vortical motion in the magnetic field, integration of the vortical motion in a mechanical model of the ether, possibility of dielectric polarization, identity of the electromagnetic and optical ethers.[88] To make the story even more complex, in his later dynamical theory and in his *Treatise* Maxwell provided a different justification of the displacement current, based on Faraday's concepts of charge and current. Every current became closed and the Ampère law no longer needed to be modified.

Maxwell's electromagnetic theory exemplified a powerful methodology. Important aspects of this methodology can be traced to other British authors. Maxwell praised Thomson and Tait's 'method of cultivating science, in which each department in turn is regarded, not merely as a collection of facts to be coordinated by means of the formulae laid up in store by the pure mathematicians, but as itself a new mathesis by which new ideas may be developed.' This approach included the dynamical ideas through which the 'two Northern wizzards' conducted their mathematical reasonings. It also provided the illustrations and analogies that Maxwell shared with Thomson. The basics of field mathematics were not born in the brains of pure mathematicians. They required the suggestive imagery of flowing liquids and strained solids.[89]

Maxwell's methodology had more original components. He developed the classification of mathematical quantities as a short-cut through the method of formal analogies. He gave more weight to geometrical reasoning than Thomson did, and filled his papers with beautiful figures of curving lines and surfaces. He had an eye for topological relations, as today's field theorists do. Lastly, he inaugurated a moderate kind of mechanical reductionism, in which the connecting mechanism was no longer exhibited. The mere assumption of the existence of such a mechanism implied the existence of a Lagrangian, from which the evolution of empirically controllable quantities could be deduced. Maxwell still hoped, however, for a more detailed mechanical understanding of field processes. For the time being he made sure that Lagrangian dynamics would not be too abstract. He fleshed it out with metaphors, illustrations, and energetics.[90]

Regarding consistency, economy, and pedagogy, Maxwell's *Treatise* was

[88] Cf. Siegel 1975. [89] Maxwell 1873c: 325.
[90] Cf. the penetrating analysis in Harman 1987.

imperfect, even in its core. For example, Maxwell did not fully realize the ambiguity of his potentials; he refused to eliminate them from the final equations; and he misled many of his readers with his metaphor of displacement. In the periphery, he tolerated the contradiction of quantized electric charge, and he occasionally regressed to the elastic solid theory of light. However, the system of the *Treatise* was sufficiently definite to guide further improvements. Maxwell defined a new kind of theoretical physics in which the classification of mathematical quantities, vector symbolism, and Lagrangian dynamics became major construction tools. He also revealed a tension between field macrophysics and the atomic structure of matter, and inaugurated ways of dealing with this tension. His physics was an unended quest. He provided methods that kept theory open and alive.

5

British Maxwellians

5.1 Introduction

Maxwell's electromagnetic theory remained a private enterprise until the early 1870s. The situation began to change after Maxwell's appointment at the head of the new Cavendish Laboratory in 1871 and the publication of the *Treatise* in 1873. This was a slow process, because the *Treatise* was 'a very hard nut to crack' even to Cambridge wranglers, and because in his new capacity Maxwell could not effectively direct theoretical researches. Yet some English-speaking students of electromagnetism, in Cambridge and elsewhere, were now exposed to the new doctrine. Some of them became Maxwell's disciples and apostles.[1]

That Maxwellian studies did not bloom earlier should not be too surprising. In the forms given in 1862 and 1865 Maxwell's theory was too provisional to effectively challenge well-established conceptions. In addition, the man who would have had the strongest power to publicize Maxwell's ideas, Sir William Thomson, did not do so much. His silence even turned into open hostility after Maxwell's death in 1879. Before studying the later reception of Maxwell's electrodynamics, we will first examine why its main inspirer did not endorse it. This will help define Maxwell's originality and dissolve the myth of the evident superiority of his theory.

5.2 Thomson's antipathy

In his Baltimore lectures, delivered in the fall of 1884, Thomson expressed his 'immense admiration' for Maxwell's vortex model and his interest in linking the velocity of light with electromagnetic measurements. Yet he judged Maxwell's electromagnetic theory of light to be 'a backward step from an absolutely definite mechanical motion' as given by Fresnel and his followers. He insisted upon 'the plain matter of fact dynamics and the true elastic solid as giving what seems to me the only tenable foundation for the wave theory of light in the present state of our knowledge.' He could not accept Maxwell's retreat from mechanical models and

[1] Nanson to Maxwell, 5 December 1873, quoted in Warwick [1999], Section 6.3. For an insightful study of the uses of Maxwell's *Treatise* in Cambridge, cf. Warwick, *ibid.*: Ch. 6.

could not regard the Lagrangian treatment of hidden mechanisms as a sufficient mechanical foundation:[2]

I never satisfy myself until I can make a mechanical model of a thing. If I can make a mechanical model I can understand it. As long as I cannot make a mechanical model all the way through I cannot understand; and that is why I cannot get the electromagnetic theory. I firmly believe in an electromagnetic theory of light, and that when we understand electricity and magnetism and light we shall see them all together as parts of a whole. But I want to understand light as well as I can, without introducing things that we understand even less of. That is why I take plain dynamics; I can take a model in plain dynamics. I cannot in electromagnetics.

It should be noted that Maxwell himself regarded the more abstract dynamical methods as only provisional and deplored his incapacity to 'take the next step, namely, to account by mechanical considerations for these stresses in the dielectrics.' Also, his first works on the electromagnetic theory depended on specific mechanical models. Yet even there Maxwell differed from Thomson. From his analogies Maxwell extracted distinctions and notions that were alien to the primary field of study, and he believed these to transcend specific geometrical or mechanical models. Essential components of his theory were obtained in this manner: the distinction between flux and force, the displacement current, and the expression of stresses. Thomson distrusted such adventurous use of analogy.[3]

Similarly, Thomson must have felt that Faraday's electrostatics overplayed the analogy between a vacuum and a material dielectric. There was no empirical evidence that a vacuum could be polarized, and Thomson's own theory of dielectrics indicated that polarization charges in material dielectrics and standard electrostatics were sufficient to explain all of Faraday's results. Hence for Thomson there was no vacuum- or air-polarization, and no displacement current. He believed that the transmission of electrostatic force did not involve electric currents, that it was much faster than the propagation of light, and that it probably had to do with the missing compression waves in the elastic solid theory of light. He condemned Maxwell's idea of transverse electric waves as pure fantasy.[4]

From an empirical point of view rigidified in Thomson's studies of telegraphic lines, electrostatic and electrodynamic interactions were essentially distinct. The electrostatic potential was in itself a physical entity, whose propagation from varying electrostatic charges could be discussed separately.[5] On the contrary, for Maxwell

[2] Thomson 1884: 132, 6, 270–1. Cf. Smith and Wise 1989: 463–1; Harman 1987: 267–8, 290–91; Knudsen 1985: 177–8; Siegel 1991: 159–60.

[3] Maxwell 1873a: #111. Cf. Wise 1981: 19–21.

[4] On dielectrics, cf. notes of Thomson's Glasgow lectures by William Jack, 1852–53, quoted and discussed in Wise and Smith 1987: 332–3, Smith and Wise 1989: 226–7, 451. On electrostatic retardation, cf. Thomson 1884: 6, 42. Unfortunately, Thomson never explained his dislike of the displacement current. The present interpretation assumes a deep interconnection between his views regarding dielectrics (in the Glasgow lectures), his much later ideas on potential retardation, and his theory of telegraph cables.

[5] Thomson 1884: 5–6, 41–3. Thomson believed that a spherical conductor submitted to a periodic potential would emit spherical longitudinal waves traveling much faster than light (*ibid.*: 41–2, and Thomson 1896). He also considered the case of periodic motion of an electrified conductor, and argued

any variation of the electric potential implied a dielectric current and therefore an electrodynamic coupling with other currents. When in 1888 Thomson faced the discovery of electromagnetic waves and the ensuing excitement of British Maxwellians, his first reaction was defensive. He still believed that the electrostatic potential propagated separately, and tried to prove that Maxwell's 'ingenious [. . .] but not wholly tenable hypothesis' of the displacement current had absurd implications for the telegrapher's closed conduction currents.[6]

Inside a homogenous conductor without changing electrification ($\nabla \cdot \mathbf{j} = 0$), Maxwell's equations lead to the equation $\mu(\sigma + \varepsilon\partial/\partial t)\partial\mathbf{j}/\partial t = \Delta\mathbf{j}$ for the conduction current **j**, which differs from the prediction of previous electrodynamic theories by the ε term. Thomson judged this could not be right 'according to any conceivable hypothesis regarding electric conductivity, whether of metals, or stones, or gums, or resins, or wax, or shellac, or india-rubber, or gutta-percha, or glasses, or solid or liquid electrolytes.' A Maxwellian would have replied that the new term was not a matter of conduction: it was a small correction due to the radiation of electromagnetic energy by the variable current. But Thomson seems to have excluded any consideration that would alter the structure of his telegraph theory.[7]

Of course, Thomson did not deny the need to extend electrodynamics to incomplete circuits. However, this could be done without the displacement current, as Helmholtz had already shown (see Chapter 6). Thomson proposed the generalization that gave the simplest equations for the potentials and thus 'simple and natural solutions, with nothing vague or difficult to understand, or to believe when understood, by their application to practical problems, or to conceivable ideal problems, such as the transmission of ordinary telephonic signals along submarine telegraph conductors, and land lines, electric oscillations in a finite insulated conductor of any form, transference of electricity through an infinite solid, &c. &c.' The practical imperatives of the present dominated his approach to electrical problems. He would not adopt more speculative theories, unless they were supported by a plain dynamical ether, the elastic solid or something better.[8]

For a while Thomson could not completely resist the wave of enthusiasm which followed Hertz's 'verification' of Maxwell's theory. In January 1889 he declared that Maxwell's theory marked 'a stage of enormous importance in electro-magnetic doctrine.' In his 1893 preface to Hertz's *Electric Waves* he praised Maxwell's 'splendidly developed theory.'[9] However, he kept speculating on alternative theories and let FitzGerald know how strong his dislike of the newer Maxwellian symbolism

that the phase retardation of the corresponding potential was in principle measurable (Thomson to Heaviside, 6 November 1888, quoted in *HEP* 2: 490, and discussed in Hunt 1991a: 186–7). Cf. Wise and Smith 1987: 340–1; Smith and Wise 1989: 461–3, who insist on the telegraphic context of Thomson's views.

[6] Thomson 1888: 543. Cf. Knudsen 1985: 172–3; Smith and Wise 1989: 477–8; Hunt 1991a: 162–4.

[7] Thomson 1888: 543.

[8] Thomson 1888: 544. Cf. Smith and Wise 1989: 480. Thomson's equations were identical to the Neumann case ($k = 1$) of Helmholtz's equations.

[9] Thomson 1889: 490; Hertz 1893: xiii. Cf. Hunt 1991a: 167.

(Hertz's and Heaviside's) was: 'It is mere nihilism, having no part of lot in Natural Philosophy, to be contented with two formulas for energy, electromagnetic and electrostatic, and to be happy with a vector and delighted with a page of symmetrical formulas.'[10]

5.3 Picturing Maxwell

5.3.1 Lodge's cord and beads

Even for British physicists Maxwell's notions of charge and current were difficult to grasp. An important task of Maxwell's followers was to explain and clarify these conceptions for a wider audience. The first man to do this was Oliver Lodge, a clay merchant's son who had struggled to escape his father's business and become a physicist. Lodge had no great mathematical skill and no Cambridge education (he got his doctoral degree from University College London). Chiefly an experimenter, he reasoned in terms of sophisticated models and pictures that explained or suggested various phenomena without any calculation.[11]

In 1876 his efforts to understand Maxwell's *Treatise* yielded his first model of Maxwellian charge and current. He imagined and constructed the device of Fig. 5.1, in which an inextensible cord circulates over the pulleys ABCD. The weight W corresponds to an electromotive force, the clamp S to a switch (of infinite resistance), and the eight beads typify atoms of matter. The motion of the cord corresponds to Maxwell's total current. In a dielectric, the beads are firmly attached to the cord, and their elastic links with the rigid supports are stretched when the cord is pulled. This stretching represents electric displacement. The excess of cord at A represents positive charge, and the defect of cord at B negative charge. In a conductor, the beads can slide on the cord. Hence the stretching of the supporting threads is smaller, and vanishes when there is no current. Viscous friction between the beads and the cord represents electric resistance.[12]

Lodge extended his device to explain disruptive discharge, electric absorption, charge by induction, and even electrolytic conduction. The model made clear that Maxwellian charge was a discontinuity in a state of strain, and that electricity in Maxwell's sense could not accumulate anywhere. It helped many physicists and engineers understand Maxwell.[13] However, it reinforced Maxwell's metaphor of the incompressible fluid and suggested that displacement was an actual shift of some substance in the direction of the electromotive force, that something was flowing along electric currents.

[10] Thomson to FitzGerald, 9 April 1896, quoted in Thompson 1910, Vol. 2: 1065. After 1888, Thomson still did not understand that Maxwell's theory did not admit instantaneous propagation of *physical* effects. See his 1896 polemic with FitzGerald as discussed in Wise and Smith 1987: 340–2.

[11] Cf. Hunt 1991a: 25–26; Lodge 1931.

[12] Lodge 1876, and a more elaborate form in Lodge 1889: 32–62. Cf. Hunt 1991a: 88–9.

[13] For example, Henry Rowland benefitted from reading Lodge's article: cf. Buchwald 1985a: 78–9. *Modern Views* (Lodge 1889) was a bestseller.

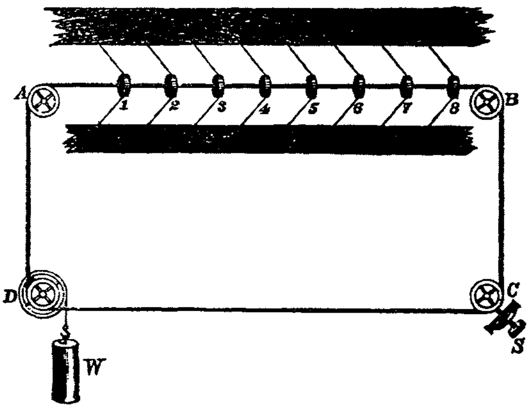

FIG. 5.1. Lodge's cord-and-beads model for a partly dielectric circuit.

5.3.2 Poynting's energy flux

In 1884 John Henry Poynting, third wrangler and Professor of physics at Birmingham, challenged the picture of the electric current as a flow. Unlike Lodge, Poynting was skeptical of any mechanical representation of Maxwell's theory and focused on the more directly observable aspects of the theory: Faraday's lines of force and the energy distribution in the field. With the Cambridge fluency in differential equations and geometrical representations, he had no difficulty answering the question: 'How does the energy about an electric current pass from point to point—that is, by what paths and according to what law does it travel from the part of the circuit where it is first recognisable as electric and magnetic to the parts where it is changed into heat or other forms?'[14]

The question seems obvious to a modern reader. It was not to Maxwell's and Poynting's contemporaries. Energy considerations usually concerned global input and output in a spatially extended system. When Maxwell localized energy in the electromagnetic field, or when the elastic solid theorists expressed the elastic energy, they did not discuss local energy flows. The only exception was the case of light, perhaps as a survivance of older substantial theories. Maxwell himself limited his discussion of energy flow to the case of plane electromagnetic light waves. It was Rayleigh who gave the first considerations of energy flux in a continuum in his *Theory of Sound* of 1877–1878. Poynting was aware of this source when he examined the question of the energy flux in the electromagnetic field.[15]

[14] Poynting 1884: 176. Cf. the obituaries by J. J. Thomson and J. Larmor in Poynting 1920: iv–xxiii, xxiv–xxvi.

[15] Maxwell 1865: 587–8; Poynting 1885a (read on 8 November 1883), where the velocity of sound is derived by consideration of the energy flux following Rayleigh. On the novelty of Poynting's ideas, cf. Buchwald 1985a: 41–3. On the connection with Rayleigh and Cambridge physics, cf. Warwick [1999]: Ch. 6. In his lectures on mechanics published in 1876 (Kirchhoff 1876: 311), Kirchhoff showed that the time derivative of the energy of sound in a given volume was the sum of a surface integral and a volume

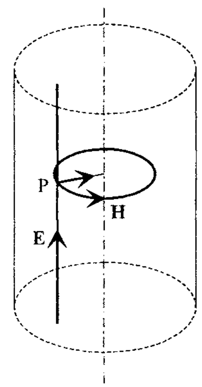

FIG. 5.2. Electric and magnetic line of force passing through a point P of a cylindrical conductor. The arrow at P indicates the direction of the energy flux.

Taking the time derivative of the integral U of the energy density $\frac{1}{2}(\varepsilon E^2 + \mu H^2)$ over a volume V delimited by a surface S, using Maxwell's equations, and integrating by parts, Poynting found

$$\frac{\mathrm{d}U}{\mathrm{d}t} = -\int_V \sigma E^2 \mathrm{d}\tau - \int_S (\mathbf{E}\times\mathbf{H})\cdot \mathrm{d}\mathbf{S} \qquad (5.1)$$

for bodies at rest. Since the first term represents energy lost into Joule heat, the second must be identified with the energy flux across the surface. Without hesitation, Poynting took $(\mathbf{E} \times \mathbf{H})\cdot \mathrm{d}\mathbf{S}$ to represent the energy flux across the surface element $\mathrm{d}\mathbf{S}$. The consequences for the energy paths in usual electric circuits were astonishing.[16]

First consider the case of a long cylindrical conductor (Fig. 5.2) conveying a constant current. The lines of magnetic force are circles centered on the axis of the cylinder, and the lines of electric force within the conductor are parallel to this axis. Therefore, the energy flux within the conductor is inward radial, which implies that the energy comes from the surrounding dielectric and is gradually transformed into heat. Consider now the case of a condenser ALBN slowly discharged through a thin wire LMN (Fig. 5.3). If the wire runs along a line of force and if the current is small, the equipotential surfaces remain those of the disconnected condenser. The energy must flow on these surfaces, since it is perpendicular to the electric force. Specifically, energy travels from the condenser through the dielectric to the wire, which it

integral depending on the included sources (his aim was to give a Gaussian proof of the existence and uniqueness of the velocity potential). In his *Theory of Sound* (Rayleigh 1877–1878: #295), Rayleigh reproduced this derivation, and interpreted the surface term as the energy flux across the surface.

[16] Poynting 1884: 176–81. Cf. Buchwald 1985a: 44. Poynting included the motion of current carriers in his balance.

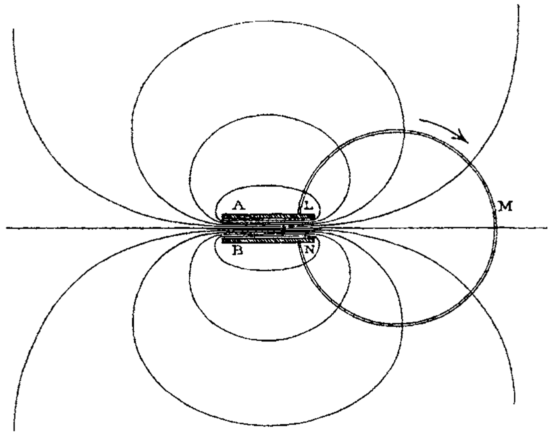

FIG. 5.3. Lines of energy flow for a slowly discharging capacitor (Poynting 1884: 183).

enters perpendicularly. No energy flows along the conductor, against the prevailing intuition.[17]

5.3.3 Moving tubes of force

In his positivistic manner, Poynting regarded Faraday's lines of force and the distribution and flow of energy as the basic properties of the electromagnetic field. In a second paper he tried to connect the energy flow to the behavior of the lines of force. Faraday had already introduced moving magnetic lines of force, so that the electromotive force induced in a linear conductor at rest would be equal to the number of lines cutting the conductor. More generally, Poynting assumed that every circuital electromotive force was due to the motion of tubes of magnetic induction, and also that every 'magnetomotive force' was due to the motion of tubes of electric induction.[18]

Consider in this light the cylindrical conductor of Fig. 5.2. The circular magnetic force around the wire corresponds to a sideways motion of the tubes of electric induction toward the wire. This motion, and a similar motion of the magnetic tubes of induction account for the energy flux into the wire. Poynting then proposed a fitting picture of the electric current: 'The wire is not capable of bearing a continually-increasing induction, and breaks the tubes up, as it were, their energy appearing finally as heat.' He thus maintained Faraday and Maxwell's idea of conduction as a relaxation of polarization, with an essential difference: the polarization now propagated sideways and came from the surrounding dielectric.[19]

[17] Poynting 1884: 181–84. Cf. Buchwald 1985a: 44.
[18] Poynting 1885b. Cf. Buchwald 1985a: 45–9.
[19] Poynting 1885b: 199. J. J. Thomson later regarded Faraday's lateral pressure as the cause of the motion of the tubes: see, e.g., J. J. Thomson 1895a: 277.

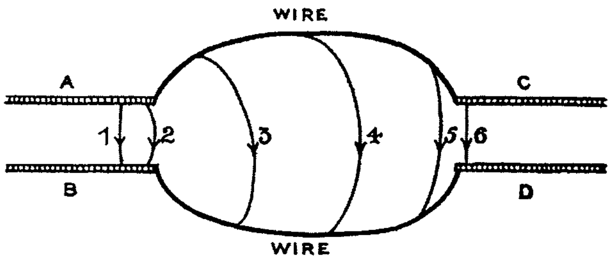

FIG. 5.4. The motion of a tube of electric force during the discharge of a capacitor AB into another CD (Poynting 1885b: 225).

The production of heat is not, however, the most interesting aspect of the electric current. On the simple example of two capacitors (Fig. 5.4), Poynting explained how a pair of wires could transfer energy across space. Originally, the capacitor AB is charged and the capacitor CD uncharged. When the connecting wires are introduced, the induction tubes of the first capacitor move sideways toward the second capacitor, keeping their extremities on the facing sides of the two wires. Opposite electric charges corresponding to these extremities thus travel along the two wires (at their surface), until they are equally distributed between the two capacitors. In this process the tubes are partially dissolved into the wires, which implies a loss of energy into Joule heat. The essential role of the wires, however, is to permit and guide the motion of the induction tubes. Again no energy and no electric charge travels *within* the wires. For an aerial telegraph wire, Poynting explained, the energy travels in the space between the wire and the Earth, with opposite electric charges on their facing surfaces. In a submarine telegraph cable, the energy travels in the insulator between the central copper wire and the surrounding iron sheath.[20]

To cast his ideas into equations, Poynting introduced the vector $\mathbf{A}$ such that $\mathbf{A}\cdot d\mathbf{l}$ gives the total number of tubes of magnetic induction having crossed the length $d\mathbf{l}$ since the origin of time, and the analogous vector $\mathbf{Z}$ in the electric case. The magnetic induction $\mathbf{B}\cdot d\mathbf{S}$ across the surface $d\mathbf{S}$ is equal to the total number of tubes having crossed its border. Therefore, $\mathbf{B} = \nabla\times\mathbf{A}$. The electric induction $\mathbf{D}\cdot d\mathbf{S}$ across the surface $d\mathbf{S}$ is equal to the number $(\nabla\times\mathbf{Z})\cdot d\mathbf{S}$ of electric tubes having crossed its border minus $(\int \mathbf{j}dt)\cdot d\mathbf{S}$, since by definition the electric current is the number of tubes dissolved in a unit of time. Therefore, $\partial(\nabla\times\mathbf{Z})/\partial t = \mathbf{j} + \partial\mathbf{D}/\partial t$. According to Poynting's principles of tube motion, the electric and magnetic forces corresponding to the motion of the magnetic and electric tubes of inductions are simply given by $\mathbf{E} = -\partial\mathbf{A}/\partial t$ and $\mathbf{H} = \partial\mathbf{Z}/\partial t$. The first equation, up to a gradient term, is the same as Maxwell's induction law in bodies at rest, with a new meaning for the vector potential. The curl of the second equation, together with the above relation between $\mathbf{Z}$, $\mathbf{D}$, and $\mathbf{j}$, retrieves Maxwell's form of the Ampère law, including the displacement current.[21]

[20] Poynting 1885c: 225–7; Poynting 1895: 270–1 (telegraph).
[21] This is a simplified rendering of Poynting 1885b: 212–23.

In general the vectors **A** and **Z** depend on the history of the system, so that no general law of motion can be given for the tubes of force. Poynting does not seem to have been aware of this difficulty. In simple cases he could specify the motion of the tubes, and that was sufficient to convince him and his friend J. J. Thomson of the heuristic power of this picture of the electromagnetic field.[22]

Not all of Maxwell's followers adopted Poynting's notion of moving tubes of force. However, the expression of the energy flux quickly became part of the Maxwellian corpus. Also, Poynting imposed his view of the conduction current 'as consisting essentially of a convergence of electric and magnetic energy from the medium upon the conductor and its transformation there into other forms.' Lastly, he contributed to a clarification of Maxwell's displacement. The term 'displacement,' Poynting explained, was ill chosen because it favored the erroneous view that energy was conveyed along the conductor. Even if a true shift of something was responsible for dielectric strain, Maxwell's **D** did not need to be identical with this shift; it only had to be a function of this shift.[23]

5.3.4 FitzGerald's wheels and rubber bands

Among those who promptly endorsed Poynting's views was Maxwell's Irish follower, George Francis FitzGerald. This tall, humorous man also had 'the quickest and most original brain of anybody,' as Heaviside later judged. He had graduated from and won a Professorship at Trinity College Dublin, which harboured as prestigious a mathematical tradition as Cambridge's. His first contributions to Maxwell's theory—which will be discussed later—were amazing blends of mathematical virtuosity and physical insight. Unlike Poynting, FitzGerald was very philosophical. His personal synthesis of Berkeley's idealism and practical materialism led him to expect a reduction of electromagnetism to matter and motion. He was generally sympathetic to the models of his friend Oliver Lodge: rough and provisional as they were, they could indicate true relations of the ultimate mechanical ether.[24]

Upon reading Poynting, FitzGerald sought a new model of the electromagnetic ether that would illustrate the new ideas on energy flux, electric conduction, and displacement. Maxwell's old vortex model and Lodge's more recent models could not do, since they involved a flow of electricity. However, FitzGerald retained two essential components of Maxwell's model: that magnetic force corresponded to local rotation, and that dielectric strain corresponded to an elastic yielding of the mechanism connecting the rotations. In its two-dimensional version, his model consisted in an array of wheels mounted on fixed axes and connected in pairs by elastic rubber bands

[22] On J. J. Thomson, cf. Chapter 7, pp. 295–300; Buchwald 1985a: 49–53.

[23] Poynting 1884: 192; *ibid.*, with a reference to Glazebrook 1881, for whom **D** was the Laplacian of the actual elastic displacement.

[24] Heaviside to Perry [February 1901], quoted in Hunt 1991a: 8. Cf. Lodge's, Larmor's, and Trouton's contributions to the 'Introductory and biographical' of FitzGerald 1902: xix–lxiv; Hunt 1991a: 6–11.

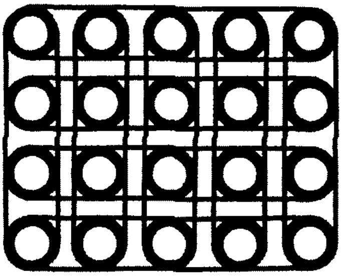

FIG. 5.5. FitzGerald's rubber band model of the electromagnetic field (from Hunt 1991a: 79, used by permission of Cornell University Press).

(Fig. 5.5). The built version was 'rather pretty on a mahogany with bright brass wheels.' FitzGerald used it both for pedagogical and for research purposes.[25]

In this model, the rotation velocity of the wheels represents the magnetic force, and the difference of strain between the two sides of a rubber band represents the electric displacement in the perpendicular direction. In a conductor, the rubber bands can slip on the wheels, thus generating frictional heat. A domain of perfect conduction may be illustrated by removing all the elastic bands in this domain. In order to understand the workings of the model, we may consider the charge and discharge of a condenser. The plates of the condenser and the connecting conductor define an H-shaped region without elastic bands. Charge is obtained by rotating the wheels bordering the channel and bringing back the elastic bands in the channel. The resulting self-locked strain in the region between the two plates corresponds to dielectric polarization. Now suppose that the elastic bands in the channel can slide on the wheels, though with much friction. The strain will be gradually released, its energy being transformed into frictional heat in the channel. In this process the energy travels in the length of the strained elastic bands, that is, in the direction perpendicular to the channel. More generally, the energy flux conveyed by the bands is perpendicular to the displacement, in conformity with Poynting's doctrine.[26]

FitzGerald also discussed oscillatory and sparking discharges, electrostatic induction, and electromagnetic induction. In every case the model faithfully reproduces the predictions of Maxwell's theory. In fact the basic equations of motion are the same. A different rotation of two consecutive wheels implies a different strain of the two sides of the connecting band, or $\nabla \times (\int \mathbf{H} dt) = \mathbf{D}$ in electromagnetic language. If ε measures the elasticity of the bands and μ the angular inertia of the wheels, the bands on a given wheel impress a net torque $\nabla \times (\mathbf{D}/\varepsilon)$, which must be equal to the time variation of the angular momentum $\mu\mathbf{H}$ according to the laws of dynamics. The two circuital laws of Maxwell's theory are thus retrieved.[27]

[25] FitzGerald 1885a, 1885b; FitzGerald to Lodge, 3 March 1894, quoted in Hunt 1991a: 78–9. A three-dimensional extension of the model is sketched in FitzGerald 1885b: 160–1.

[26] FitzGerald 1885a: 143, 145; 1885b: 157–9. For a clear, illustrated explanation, cf. Hunt 1991a: 78–83.

[27] FitzGerald 1885a: 147–8. Instead of Maxwell's **D** and **B**, FitzGerald used their counterparts in MacCullagh's medium, as he had done in 1879 (see below, pp. 190–2).

FitzGerald's model provided an excellent illustration of the central features of Maxwell's theory as Maxwell's followers came to understand it. It showed that displacement was a local change of structure, that nothing circulated along an electric current, that electric charging presumed conduction, and that energy circulated in a direction perpendicular to the electric force. However, as FitzGerald himself emphasized, the model did not represent the connection between ether and matter. Matter was required 'to get a hold on the ether so as to strain it.' It was also necessary to produce electrostatic attractions, because the stress of the rubber bands was linear instead of quadratic. It certainly played a role in magnetized bodies and in the Faraday effect. This raised a difficult question: could there be a simple mechanical representation of Maxwell's system that integrated the effects of matter?[28]

5.3.5 Lodge's cogwheels

Undeterred by this sort of difficulty, Lodge trusted that he could invent a field mechanism for all electromagnetic processes. In 1879 he had already speculated that the ether was made of wheels of positive and negative electricity geared to one another, as in Fig. 5.6. The rotation of the positive wheels (or the opposite rotation of the negative ones) represented the magnetic force, and their elastic yielding corresponded to electric displacement, as in Maxwell's earlier model. Lodge's innovation was the introduction of two electricities instead of one, which he believed to be necessary to explain the double electrolytic motion, the lack of intrinsic momentum of the electric current, and the existence of both positive and negative electric winds, among other things. Ten years later, he published an improved version of this model in his best-selling *Modern Views of Electricity*. Adopting FitzGerald's notion of conduction as a slip in the mechanical connections, he replaced the cogwheels with smooth wheels within conductors (Fig. 5.7). His depictions of basic field processes were similar to FitzGerald's, despite the complication introduced by the two kinds of wheels.[29]

Unfortunately, Lodge multiplied models without clearly stating their limits and interrelations. For the capacitor alone he offered three different models: the cord-and-beads, a hydro-pneumatic device, and the cogwheels. Even though he had the ambition of covering the whole field of electricity and magnetism with a single consistent model, he ended up illustrating various phenomena by a variety of incompatible models. His *Modern Views* prompted Pierre Duhem's famous statement:[30]

[28] FitzGerald 1885a: 142–3: 'I do not intend the model to illustrate at all the connexion between the ether and matter, and indeed think it one of the advantages to be derived from studying this model that it so distinctly emphasizes the distinction between the phenomena, depending on the general properties of the ether itself, and those depending on its connexion with matter.' *Ibid.*: 144, FitzGerald added a clever system of threads to his model in order to represent electrostatic attractions and give a rough idea of the relevant connection between ether and matter.

[29] *BAR* 1879: 258, and Lodge to FitzGerald, 29 February 1880, quoted in Hunt 1991a: 31; Lodge 1889: Chs. 10–11. Cf. Hunt 1991a: 30–1, 89–92.

[30] Duhem 1914: 101. Cf. Hunt 1991a: 87–8.

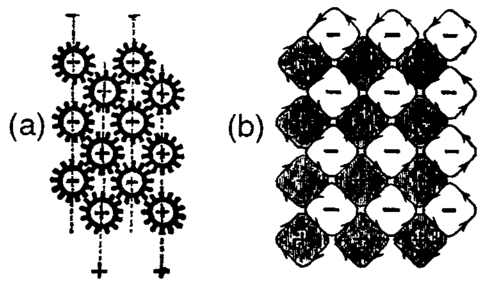

FIG. 5.6. Lodge's cogwheel ether (a), and an improved version (b) (Lodge 1889: 179, 180).

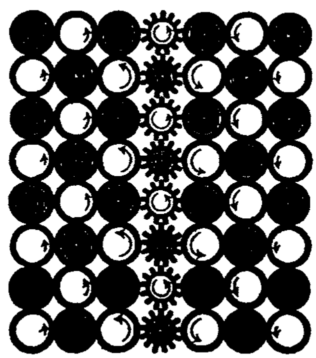

FIG. 5.7. Lodge's illustration of a direct and return current in two conductors (round wheels) separated by a thin insulating layer (cogwheels) (Lodge 1889: 189).

Here is a book intended to expound the modern theories of electricities and to outline a new theory. In it there are nothing but strings running over pulleys, wrapping around drums, going through beads, and carrying weights; and tubes pumping water while others swell and contract; wheels gearing each another and forming pinions for racks. We thought we were entering the tranquil and neatly ordered abode of reason, but we find ourselves in a factory.

Duhem took Lodge's book to be typical of the British inclination for mechanical models. In reality this was an extreme case, even from a British point of view. Poynting, who was generally suspicious of mechanical models, declared that Lodge's explanations were 'merely hypothetical' and were 'solely of value as a scaffolding enabling us to build up a permanent structure of facts, i.e. of phenomena affecting our senses.' He also found that the cogwheel model contradicted an essential feature of Maxwell's theory, the intimate relation between conduction and displacement. In Lodge's cogwheel model, the displacement current involved an actual displacement of the two electricities, whereas the conduction current involved no displacement whatsoever. Poynting kindly offered a modification of Lodge's model in which the conduction current was a double procession of the cogwheels. Nevertheless, he found his moving tubes of force 'much more easy to deal with' and more likely to guide future research: 'I believe that we may symbolise electric and magnetic actions by means of lines of force and their motions in a way which allows us to think clearly of the phenomena, and though the ultimate nature of the lines of force is unknown, we can only say the same of the ether.'[31]

5.3.6 The vortex sponge

Not even Lodge's dear friend FitzGerald liked the cogwheels. By distinguishing positive and negative wheels Lodge had disturbed the central Maxwellian dogma that electric charge was no more than strain discontinuity. FitzGerald exclaimed: 'Oh! I think your model is horrid!' In his philosophy, ether could only be something very simple: a continuous fluid in motion. William Thomson had shown in 1880 that a mass of ideal fluid could exist in a state densely filled with randomly oriented vortex filaments. Because of the gyrostatic inertia of the vortices, the fluid acquired some rigidity and could propagate transverse waves. In 1885, soon after designing the rubber band model, FitzGerald speculated that the ether was such a 'vortex sponge.' The freedom in the arrangement of the vortices offered an *embarras de richesses* for retrieving the known properties of the ether. Until the end of his short life (he died in 1901), FitzGerald and a few sympathizers struggled to construct a convincing ether out of a vortex sponge. They encountered insurmountable mathematical difficulties. The vortex sponge was the string theory of those days: its basis was attractively simple, it could not be refuted, but it could not be developed far enough to be verified.[32]

5.4 Modifying Maxwell's equations

The electromagnetic field equations of the *Treatise* could not be generally valid, whatever the underlying picture was. The clearest hole in their consequences was the Faraday effect. For this special intervention of magnetized matter, Maxwell had to return to the elastic solid theory of the ether. Yet this effect suggested the rotary character of magnetism, an essential feature of Thomson's and Maxwell's conception of the magnetic field. Aware of this strange situation, Maxwell's followers were much interested in magneto-optics. When in 1876 the Glasgow physicist John Kerr announced a new phenomenon of this kind, FitzGerald immediately set himself to work.

[31] Poynting 1893: 264, 267, 267–8. Cf. Hunt 1991a: 94–5.

[32] FitzGerald to Lodge, late September 1889, quoted and dated by Hunt 1991a: 92–3; Thomson 1880; FitzGerald 1885a: 154–6; 1888: 236–40; 1889; 1899. Cf. Whittaker 1951: 295–303; Hunt 1991a: 96–104. FitzGerald's attempt belonged to a strong variety of British mechanical reductionism, in which all energy had to be kinetic: cf. Topper 1971, and Klein 1972a.

5.4.1 FitzGerald's 'very important step'

Magneto-optical rotations, Kerr thought, would be much larger if they could be produced in a strongly magnetic substance like iron. The only obstacle was the opacity of iron, which he circumvented by using reflection instead of transmission for the polarized light. He observed that the polarization of light reflected on a polished magnet pole was altered by the magnetization.[33] FitzGerald promptly explained Kerr's observations by decomposing the incident light into two circularly polarized components, and invoking the different refraction index of the magnetized iron for these two components. A more fundamental theory required an extension of Maxwell's theory of the Faraday effect including the boundary conditions between two different media. This is what FitzGerald obtained in 1879.[34]

The problem of boundary conditions in optics was reputed to be difficult. Nearly all elastic solid theories required an *ad hoc* omission of some of the dynamically necessary conditions.[35] The source of the difficulty was that transverse vibrations did not remain so after crossing the border between two media. If the refracted vibration was artificially required to be transverse, not all boundary conditions could be satisfied. Today's physicists know that the electromagnetic theory of light provides correct boundary conditions in a very simple manner and thus eliminates the outstanding difficulty of elastic solid theories. Maxwell did not. Deterred by the apparent complexity of the issue, and distrusting his equations when applied to quickly variable phenomena, he gave up the derivation of the laws of refraction.[36]

Light came to FitzGerald from a fellow countryman, James MacCullagh. In 1839 this brilliant mathematician had cut the Gordian knot of optical theory by choosing the potential energy of the elastic medium so that the true boundary conditions of this medium would yield Fresnel's formulas for the intensities of reflected and refracted rays. If ξ denotes the local shift of the medium and ε its elasticity constant, MacCullagh's potential is simply given by $(1/2\varepsilon)(\nabla \times \xi)^2$. MacCullagh then used the action principle of another famous Irish mathematician, Rowan Hamilton, to derive the equation of motion

$$\mu \frac{\partial^2 \xi}{\partial t^2} = -\nabla \times (\varepsilon^{-1} \nabla \times \xi) \tag{5.2}$$

(μ is the density of the medium) as well as the two boundary conditions: continuity of ξ, and continuity of the tangential component of $\varepsilon^{-1}\nabla \times \xi$. By itself, this equation of motion excludes longitudinal waves.[37]

[33] Kerr 1876, 1877. If the incident light is linearly polarized, the reflected light becomes elliptically polarized, and the major axis of the ellipse is rotated away from the original plane of polarization: cf. Buchwald 1985a: 102n, 109–10.

[34] FitzGerald 1876. Kerr and Thomson had the same idea: cf. *ibid.*: 14.

[35] This was true for the theories of Poisson, Cauchy, Green, Neumann, and Kirchhoff, but not for Cauchy's labile ether nor for MacCullagh's medium: cf. Whittaker 1851: Ch. 5; Schaffner 1972.

[36] Maxwell to Stokes, 15 October 1864, *MSLP* 2: 186–8; Manuscript notes on the reflection and refraction of light, *MSLP* 2: 182–5. Cf. Harman 1995b: 79–80, 85–6. The solution of the problem was first announced in Helmholtz 1870, and given in Lorentz 1875.

[37] MacCullagh 1839. Cf. Whittaker 1951: 142–4; Schaffner 1972: 59–68, 187–93; Stein 1981: 310–15; Buchwald 1985a: 283–4; Hunt 1991a: 9–10.

MacCullagh was familiar with Green's memoir of 1838, which gave the most general form of the potential of an elastic solid in terms of two moduli for rigidity and compression. In Green's terms MacCullagh's potential corresponds to a negative compressibility, which Green naturally excluded. Aware of this paradox, MacCullagh simply accepted that the ether was very different from any natural solid. Other physicists thought differently, expecially after Stokes had proved, in 1862, that MacCullagh's medium violated the principle of action and reaction: the *absolute* rotation of an element of the medium calls forth a restoring elastic torque. No one took MacCullagh's ingenious theory seriously, until FitzGerald resurrected it in 1879.[38]

MacCullagh's equation for ether motion, FitzGerald discovered, resulted from Maxwell's theory if only Maxwell's displacement was identified with the curl of $\boldsymbol{\xi}$. Indeed, the Ampère law $\nabla \times \mathbf{H} = \partial \mathbf{D}/\partial t$ then implies $\mathbf{H} = \partial \boldsymbol{\xi}/\partial t$, so that MacCullagh's equation (5.2) becomes identical to $\partial \mathbf{B}/\partial t = -\nabla \times \mathbf{E}$, in conformity with the Faraday–Maxwell induction law. Thanks to this equivalence, FitzGerald could write the boundary conditions in the electromagnetic theory of light. He also gave an electromagnetic interpretation of Maxwell's theory of the Faraday effect.[39]

The latter theory was based on a coupling between the optical motion of the ether and the vortical motion produced by the external magnetic field. As Maxwell explained, this was a hybrid approach, because the optical ether motion received no electromagnetic interpretation.[40] By contrast, in FitzGerald's version of the theory all quantities received a double interpretation, an electromagnetic one, and another in terms of MacCullagh's medium. Adding Maxwell's magneto-optical term to MacCullagh's Lagrangian, FitzGerald derived the equations of motion and the Faraday effect in elegant quaternion form: a true Halmitonian feast. Then he used MacCullagh's boundary conditions[41] to derive the effect of the external magnetic field on the reflection of a polarized light wave. He thus retrieved those of Kerr's observations that did not depend on the metallic character of the second medium.[42]

As a referee of FitzGerald's paper Maxwell commented: 'If he has succeeded in explaining Kerr's phenomena as well as Faraday's by the purely electromagnetic hypothesis, the fact that he has done so ought to be clearly made out and stated, for it would be a very important step in science.' FitzGerald followed this advice in the printed conclusion which Maxwell did not live to see:

The investigation is put forward as a confirmation of Professor Maxwell's electromagnetic theory of light [. . .]; if it induced us to emancipate our minds from the thraldom of a

[38] Green 1838; Stokes 1862. Cf. Schaffner 1972: 66, 71–4.

[39] FitzGerald 1879a. MacCullagh's conditions are the continuity of ξ and the continuity of the tangential component of the torque $\varepsilon^{-1}\nabla \times \xi$. In FitzGerald's electromagnetic interpretation, they imply the continuity of the tangential components of **E** and **H**, and the continuity of the normal components of **B** and **D**, as given by the modern electromagnetic reasoning. FitzGerald eliminated the alternative electromagnetic interpretation, in which ξ is identified with **A**: see Appendix 9.

[40] Maxwell to Stokes, 6 February 1879, in Stokes 1907, Vol. 2: 43. See Chapter 4, pp. 172.

[41] However, he dropped the continuity of the normal component of ξ, because it was not consistent with the magneto-optical term. Cf. Larmor's explanation of the difficulty in FitzGerald 1880: 66n.

[42] FitzGerald 1879a. Cf. Hunt 1991a: 15–21.

material ether [it] might possibly lead to most important results in the theoretic explanation of nature.

This statement should not be read as a renunciation of the mechanical ether. FitzGerald only meant to corroborate MacCullagh's intuition that the ether did not resemble any of the substances found in our material world. The ether was no elastic solid, no jelly, no foam, whatever Stokes and Thomson might think; but it could be a mechanical medium of a very different kind, for instance the vortex sponge which later absorbed FitzGerald's hopes.[43]

Not only did FitzGerald's paper redirect the search for the ultimate mechanical medium, but it also inaugurated a powerful strategy for integrating magneto-optical phenomena into Maxwell's theory without entering microphysical speculation. The basic idea was to modify the field equations in conformity with the dynamical foundation of the theory. FitzGerald did that by first writing a Lagrangian giving Maxwell's field equations, and then adding small new terms to this Lagrangian. The first step was in itself an important innovation, for Maxwell had only written the Lagrangian of a system of linear currents, not including the elastic energy stored in electric displacement.[44]

5.4.2 The Hall effect

One weakness of FitzGerald's theory of the Faraday and Kerr effects was that the new term in the Lagrangian was added in a purely *ad hoc* manner, without independent electromagnetic justification. One year later the young American physicist Edwin Hall discovered that a strong magnetic field, when applied perpendicularly to an electric current, implied an electromotive force in the direction perpendicular to both the field and the current. The effect was very small and required refined galvanometry, which explains why it was not discovered earlier. Varying the conditions of the experiment, Hall found that the new electromotive force could be expressed as $h\mathbf{H}_0 \times \mathbf{j}$, where h is a small constant, $\mathbf{H}_0$ the external magnetic force, and $\mathbf{j}$ the electric current.[45]

As Boltzmann pointed out in 1886, the effect is easily justified by assuming that the electromagnetic force acting on the conductor results from a force acting on the current itself. However, Maxwell insisted that electromotive forces and electrodynamic forces were essentially different things, since in his system currents were not the flow of electric charge. As a good Maxwellian, Hall could not consider Boltzmann's kind of explanation. Instead he ascribed the new electromotive force to a fundamental modification of Maxwell's field equations. So did his mentor Henry

[43] Maxwell to Stokes, 6 February 1879, in Stokes 1907: 43; FitzGerald 1879a: 73. Cf. Stein 1981; Hunt 1991a: 20–3.

[44] For FitzGerald, the magnetic induction **B**, not the current **J**, is the generalized velocity (see Appendix 9). This dynamical method does not include the conduction current. Heaviside 1893–1912, Vol.1: ##146–59, showed that the conduction current cannot be obtained by introducing dissipation in MacCullagh's medium: cf. Buchwald 1985a: 68–70. For the general issue of applying Hamilton's principle to the field, cf. Buchwald 1985a: Chs. 6–7.

[45] Hall 1879, 1880a. Cf. Buchwald 1985a: Chs. 9–10.

Rowland, who soon imagined a connection between Hall's effect and the Faraday effect.[46]

Hall's effect, Rowland reasoned, may be seen as the curving of the conduction current by an external magnetic force. Since in Maxwell's theory conduction and displacement currents are on the same dynamical footing, the latter should also be curved by an external magnetic force. Specifically, Rowland added a term $h\mathbf{H}_0 \times \partial\mathbf{D}/\partial t$ to Maxwell's expression of the electromotive force in a magnetic field, and found that the modification implied the magneto-optical rotation calculated by Maxwell. His theory was undoubtedly superior to Maxwell's, for it was entirely electromagnetic, did not require any reference to ether vortices, and related the Faraday effect to a purely electromagnetic phenomenon. He proudly claimed to have given 'a demonstration of the truth of Maxwell's theory of light,' and more generally considered that Maxwell's theory had been 'raised to the realm of fact.'[47]

Rowland did not compare his theory with FitzGerald's. A Cambridge Maxwellian, Richard Glazebrook, soon did that for him. The Maxwell–FitzGerald addition to the field Lagrangian, $(k\nabla \times \partial\xi/\partial t)(\mathbf{H}_0 \cdot \nabla)\xi$, implies a corrective term $-2k(\mathbf{H}_0 \cdot \nabla)\mathbf{H}$ in Maxwell's expression of the electromotive force. This differs from $2k\mathbf{H}_0 \times (\nabla \times \mathbf{H})$ by a gradient which can be absorbed in the scalar potential. Therefore FitzGerald's new term is just what is needed to submit the displacement current to the Hall effect. For Cantabrigian physicists weaned with Lagrangians, Glazebrook's argument marked a major step in the unification of physics, as well as a major victory of Maxwell's theory.[48]

Magneto-optics remained a major Maxwellian topic well into the 1890s both in England and on the continent. In 1884 Hendrik Lorentz and his student W. van Loghem pursued Rowland's connection between the Hall effect and magneto-optical rotation, and were first to take into account metallicity in the Kerr effect. However, their theory did not explain why iron and nickel had similar Kerr effects despite the opposite sign of their Hall effects. For this reason in 1893 J. J. Thomson and Paul Drude abandoned Hall's idea that the Hall effect implied a new kind of electric field both for conduction and for dielectric currents. Even so, their theories turned out to be incompatible with Remmelt Sissingh's excellent data on the Kerr effect, published in 1891. As Lorentz's student Cornelius Wind demonstrated in 1898, the difficulty could only be solved by replacing Maxwell's unanalyzed, macroscopic displacement with Lorentz's ionic polarization. From a major vindication of Maxwell's theory, magneto-optical researches had evolved into one of its major

[46] Boltzmann 1886b (who used the effect to determine the velocity of electricity); Maxwell 1873a: #501; Hall 1880b. A minor Maxwellian, John Hopkinson, preferred to modify Ohm's law (Hopkinson 1880). Maxwell himself had conceived the possibility of a rotary part of the resistivity matrix in magnets (Maxwell 1873a: #297): cf. Buchwald 1985a: 96. The modern explanation is similar to Boltzmann's.

[47] Rowland 1880a, 1880b, 1881: 261. Cf. Buchwald 1985a: 102–6. Rowland 1880b included a strange reformulation of Maxwell's theory that caused a public rebuttal by J. J. Thomson (1881b).

[48] Glazebrook 1881. See also J. J. Thomson 1888: #43, and Larmor 1893a. Cf. Buchwald 1985a: 111–9. The main purpose of Glazebrook's paper was to take Thomson and Maxwell's vortical interpretation of magnetism to the letter: he identified **H** with the vorticity $\nabla \times \partial\xi/\partial t$ in an elastic solid. This attempt had no sequel, probably because it led to a field-energy distribution different from Maxwell's.

insufficiencies. This change was part of a more general historical transition to be explained in Chapter 8.[49]

5.5 A telegrapher's Maxwell

There was, in the flock of Maxwell's followers, one maverick who ignored magneto-optics, questioned attempts at finding the mechanical structure of the ether, judged that 'so-called models' were 'harder to understand than the equations of motion,' and ridiculed the Cambridge fashion for Lagrangians. Yet the changes he brought to the formulation of Maxwell's theory were of the most lasting value. This man was Oliver Heaviside, the son of a wood-engraver, the nephew of the British inventor of the electric telegraph (Charles Wheatstone)—and 'a first rate oddity' even to his closest friends. After a brief attempt at creative writing (including an essay entitled 'Muscular characters'), he devoted himself entirely to electrical science. He spent most of his life as a virtual ermit. He avoided the society of other scientists, and enjoyed denouncing the incompetences of established authorities. His sarcastic wit won him powerful ennemies, and occasionally compromised the diffusion of his works.[50]

5.5.1 Telegraphic circuits

From a scientific point of view, much of Heaviside's originality came from his seven-year experience as a telegraph operator. Unlike other Maxwellians, most of his researches were aimed at solving or easing the solution of practical problems of telegraphy and telephony. In physics and mathematics he had no academic training and acquired his vast knowledge through reading. He was highly impressed by the writings of the knight of the telegraph, Sir William Thomson. Through his expertise in the Atlantic cable project, Thomson had not only placed the art of electric communication on a sound theoretical basis, but he had also enriched electrical science with reliable standards and measuring techniques. Heaviside later praised Thomson's 'invaluable labours in science, inexhaustible fertility, and immense go.'[51]

Heaviside's first papers were devoted to the theory of the electric circuits and apparatus used by telegraphers. Like Thomson and Kirchhoff, he worked directly in

[49] Lorentz 1884; van Loghem 1883 (Lorentz and van Loghem used Helmholtz's version of Maxwell's theory); J. J. Thomson 1893a; Drude 1893; Sissingh 1891; Wind 1898, 1898–1899. Also, Rowland's idea of modifying Maxwell's expression of the electromotive force must be replaced with the idea of modifying the relations involving the constitutive parameters ε and σ. Cf. the thorough study in Buchwald 1985a: 205–9 (Lorentz–van Loghem); 123–9 (J. J. Thomson); 215–7 (Drude); 210–4 (Sissingh); 242–7 (Wind).

[50] Cf. Yavetz 1995: 276 (no magneto-optics); Heaviside to Hertz, 14 August 1889, quoted in Hunt 1991a: 105 (no models); Buchwald 1985c (no Lagrangians); Searle 1950: 96 (oddity); Appleyard 1930 (muscles). For Heaviside's biography, cf. Whittaker 1929; Appleyard 1930: 212–20; Nahin 1988; Hunt 1991a; Yavetz 1995: 5–28.

[51] Heaviside 1885: 418. Cf. Hunt 1991a: 58. On Thomson and the telegraph, cf. *supra* pp. 122–5, and Smith and Wise 1989.

terms of measurable electromotive force, current, resistance, 'capacitance,' and 'inductance,' and avoided speculation on the deeper nature of electricity. His mathematical solutions were extremely thorough, and proceeded elegantly by a constant return to the physical problem. He had the British intolerance for dry mathematical developments and required a physical interpretation for each step of the reasonings. Conversely, the physical meaning of mathematical operations could suggest to him new mathematical methods. For example, he treated the resistance in a circuit and operators such as $L\mathrm{d}/\mathrm{d}t$ (L being the inductance) as part of an operational 'impedance.' This practice led to the 'operational calculus,' a non-rigorous anticipation of modern distribution theory.[52]

With his 'electrical mathematics' Heaviside solved numerous problems of signal propagation. From a practical point of view, his most important result was the derivation of the possibility of distortionless telephony by inductive loading of the lines. Had he not let an American engineer patent this discovery before him, he would have been rich.[53]

5.5.2 The principle of activity

Dynamical considerations played an essential role in Heaviside's physics. 'All the physical sciences,' he declared, 'are bound to become branches of dynamics in the course of time, and anything contradicting the principles of dynamics should be unhesitatingly rejected.' However, he made little use of Thomson's flow analogy for telegraphic cables, and no attempt at specifying the hidden mechanisms. He instead relied on the general dynamical concepts of energy, force, and momentum, as developed in some of Thomson's early papers, Thomson and Tait's *Natural Philosophy*, and Maxwell's *Treatise*. From the early Thomson he borrowed the energetic definition of electromotive forces, from TT′ the more general 'principle of activity,' and from Maxwell the notion of the 'electromagnetic momentum' of an electric current.[54]

By the 'Principle of activity' Heaviside meant Thomson and Tait's interpretation of Newton's scholium to his third law:

> If the Activity [*actio*, so translated in the second edition of TT′ to avoid confusion with Maupertuis's action] of an agent be measured by its amount and its velocity conjointly; and if, similarly, the Counter-activity of the resistance be measured by the velocities of its several parts and their several elements conjointly, whether these arise from friction, cohesion, weight, or acceleration;—Activity and Counter-activity, in all combinations of machines, will be equal and opposite.'

Thomson and Tait interpreted this statement as prefiguring d'Alembert's and Lagrange's formulations of mechanics as well as energy conservation. Heaviside

[52] On Heaviside's circuit theory, cf. Yavetz 1995: Ch. 2 (p. 39 for 'electro-mathematical reasoning'). On the operational calculus, see *ibid.*: 306–20, and Hunt 1991b.

[53] On distortionless transmission, cf. Jordan 1982a; Yavetz 1995: 209–18.

[54] Heaviside 1885: 419; 1878: 95–7 (for a use of the water-pipe analogy); 1885–1887: 451: 'Energy definition of impressed forces due originally, if not explicitly, at least substantially, to Sir William Thomson'; Heaviside 1876: 54, 59, and 1878: 97 for the electromagnetic momentum.

retained the idea of balancing the 'activities,' that is, the rates at which the various forces acting in and on the system perform work. He also adopted Thomson and Tait's generalized concept of force, for which the basic equation 'force × velocity = activity' remains true, even when the 'velocity' no longer refers to the motion of a substance.[55]

5.5.3 Maxwell for the many

Although in his early works on circuit theory Heaviside avoided discussing the nature of electricity, his use of the concept of electric momentum betrayed a sympathy for Maxwell's system. He also shared Maxwell and Thomson's belief that the motion responsible for this momentum was located in the magnetic field. However, he said nothing on the nature of the electric current, although he later remembered that he never accepted the fluid picture:

> It so happened that my first acquaintance with electricity was with the dynamic phenomena, and after I had read with absorbed interest that instructive book, Tyndall's 'Heat as a mode of motion.' This may explain why, when it came later to book learning regarding electricity, I had the greatest possible repugnance to all the explanations, and could not accept the electric current to be the motion of electricity (static) through a wire, but thought it something quite different.

Heaviside then believed that electricity was a mode of motion and the electric current something similar to heat flow. He was prepared to accept another non-substantialist view: Maxwell's.[56]

In 1882 Heaviside started a series in *The Electrician* on 'the relations between magnetic force and electric current' according to Maxwell. He wanted to strip the 'higher conceptions' of 'eminent mathematical scientists of their usual symbolical dress' and to make them 'appeal to the sympathies of the many.' To make Maxwell more accessible to the intelligent telegrapher, he invented the modern vector notation, gave geometrical definitions of the curl and divergence operators, and proved the corresponding integral theorems. His method was largely reminiscent of Maxwell's 'On Faraday's lines of force.' For example, he proved Stokes's theorem by means of a net of infinitesimal loops; he introduced a series of vectors **A**, **B**, **C** (and even a fourth one) deduced from each other by 'curling' and representing the vector potential, the magnetic force, and the current; and he used symmetry arguments to determine particular distributions of magnetic force. In Thomson's and Maxwell's manner, Heaviside imparted life to his symbols by relating them to simple geometrical operations or physical processes. His vector notation, and Maxwell's

[55] Heaviside 1893–1912, Vol. 3: 178; Thomson and Tait 1879–1883, Vol. 1: 247; Heaviside 1883–1884: 291; 1885–1887: 435. Cf. Hunt 1991a: 122–3; Yavetz 1995: 131–6, 269–71. On the reference to Newton in TT', cf. Smith 1998: Ch. 10.

[56] Heaviside 1885–1887: 435, 436. Cf. Yavetz 1995: 143–4.

'curl' and 'convergence' did not only save writing. They provided intuitive guidance in the mathematical thicket of Maxwellian electromagnetism.[57]

In his next series Heaviside argued in favor of Maxwell's way of distributing the magnetic and electric energies in the field. In this context his most decisive insight occurred in 1884 during a study of the currents induced in a conducting core within a solenoid. Having in view applications to electromagnets, transformers, and self-inducting coils, he focused on the energy processes in the core. Combining the Ampère law, Faraday's induction law, and Ohm's law, he computed the time variation of the magnetic energy density, and found

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\mu H^2\right) = -\sigma E^2 - \sigma^{-1}\nabla\cdot(\mathbf{j}\times\mathbf{H}). \tag{5.3}$$

The first term represents the Joule heat and the second an energy flux directed toward the axis of the core. A year later Heaviside generalized this result to include displacement currents, and reached the general expression $\mathbf{E}\times\mathbf{H}$ of the energy flux, independently of Poynting. He described the energy transfers in simple electric systems, and he predicted that quickly varying currents, such as those involved in rapid signaling, would be confined to the surface of the conductors, because the energy coming from the outer dielectric would have no time to penetrate the mass of the conductor before the reversal of the electromotive force. He soon found that the corresponding violation of Ohm's law explained measurements performed by the electrical inventor David Hughes, and claimed priority for the discovery. In reality, the skin effect does not require the Poynting flux nor even Maxwell's theory, and it had been anticipated by several other authors, including Rayleigh, Larmor, and Lamb.[58]

5.5.4 The rough sketch

Heaviside's prediction of the skin effect was part of a major reformulation of Maxwell's theory, modestly entitled 'Rough sketch.' Heaviside started with Ohm's law ($\mathbf{j} = \sigma\mathbf{E}$), Maxwell's electric displacement ($\mathbf{D} = \varepsilon\mathbf{E}$), and magnetic induction ($\mathbf{B} = \mu\mathbf{H}$), from which he built the expressions of the energies dissipated and stored in the volume element $d\tau$ of the medium: $\sigma E^2 d\tau$ and $\frac{1}{2}\mathbf{E}\cdot\mathbf{D}d\tau + \frac{1}{2}\mathbf{E}\cdot\mathbf{H}d\tau$, respectively. He disregarded Maxwell's pictures of charge and current, but retrieved the basic distinction between force and flux, as conjugate factors in energy densities. Then he

[57] Heaviside 1882–1883: 195; *ibid.*: 211–12, and Maxwell 1856b: 206 for Stokes's theorem (in the *Treatise* Maxwell used partial integration, probably for the sake of rigor); Heaviside 1882–1883: 205 for **A**, **B**, **C**, **D**; *ibid.*: 200–1, 224–8, and Maxwell 1878: 140 for combining symmetry arguments and the Ampère theorem. Cf. Yavetz 1995: 66–112. On Heaviside's vector notation, cf. Crowe 1967; Hunt 1991a: 105–7; Yavetz 1995: 85–7.

[58] Heaviside 1883–1884; 1884–1885: 378; 1885–1887: 440–1. On the energy flux, cf. Hunt 1991a: 120–1; On the skin effect, cf. Jordan 1982b; Yavetz 1995: 191–208; and also Chapter 6, pp. 221, 226, for Helmholtz's similar effect.

turned to the activities (rates of producing work) of the forces **E** and **H**. These are obtained by multiplication with the corresponding current. In the electric case, Heaviside simply adopted Maxwell's expression of the current ($\mathbf{J} = \mathbf{j} + \partial\mathbf{D}/\partial t$), which yields:

$$\mathbf{E}\cdot\mathbf{J} = \sigma E^2 + \frac{\partial}{\partial t}\left(\frac{1}{2}\mathbf{E}\cdot\mathbf{D}\right), \tag{5.4}$$

The electric actitivity is therefore equal to the Joule heat plus the electric energy stored in the medium. Heaviside wanted a similar result in the magnetic case. He therefore defined the 'magnetic current' $\mathbf{G} = g\mathbf{H} + \partial\mathbf{B}/\partial t$, the magnetic conductivity g being there only for more symmetry. Then the magnetic activity is

$$\mathbf{H}\cdot\mathbf{G} = gH^2 + \frac{\partial}{\partial t}\left(\frac{1}{2}\mathbf{B}\cdot\mathbf{H}\right), \tag{5.5}$$

in perfect analogy with the electric case.[59]

For the cross-connections between electric and magnetic force, Heaviside wrote Maxwell's form of the Ampère law, and a similar relation between electric force and magnetic current:

$$\begin{aligned} \nabla\times\mathbf{H} &= \mathbf{J}, \\ -\nabla\times\mathbf{E} &= \mathbf{G}. \end{aligned} \tag{5.6}$$

This is the modern form of Maxwell's equations, if we forget the magnetic conduction current. Heaviside was proud to have 'murdered' the potentials, which he held responsible for various physical misconceptions. Of course, he knew Maxwell's interpretation of **A** as the electromagnetic momentum. But he accepted the notion only for complete circuits, and rejected its Lagrangian justification. In his opinion the principle of least action was 'the golden or brazen idol' of the Cambridge over-educated, and it interfered with the better physical insight brought by the principle of activity.[60]

For 'dynamical completeness,' Heaviside further introduced the electromotive sources of chemical and thermoelectrical origin, and also permanent magnetism. He did this by means of the impressed forces **e** and **h**, whose activities $\mathbf{e}\cdot\mathbf{J}$ and $\mathbf{h}\cdot\mathbf{G}$ measure the energy brought to the electromagnetic system in a unit of time. Heaviside regarded this definition as an obvious generalization of Thomson's corresponding definition for linear circuits, and noted that it

[59] Heaviside 1885–1887: 429–34, 441. Cf. Yavetz 1995: 142–62.

[60] Heaviside 1885–1887: 447–448; 1889b: 468 ('Thus Ψ and **A** are murdered'), 483–5 ('On the metaphysical nature of the potentials'); Heaviside 1893–1912, Vol. 3: 175 ('golden or brazen idol'). Cf. Buchwald 1985c; Yavetz 1995: 268–9. The two circuital equations or laws had already played a central role in Maxwell 1861, 1862 and in Maxwell 1868.

had long been well recognized by most writers on electrical subjects, especially since the practical introduction of dynamos, machines, accumulators, etc., which raise the energy transformations concerned in electrical phenomena from being matters of almost purely scientific interest to matters of the extremest commercial importance.

The impressed forces add to the forces determined by the electric–magnetic coupling, so that the full 'duplex equations' read:[61]

$$\nabla\times(\mathbf{H}-\mathbf{h})=\mathbf{J}, \qquad (5.7)$$
$$\nabla\times(\mathbf{e}-\mathbf{E})=\mathbf{G}.$$

Using these equations, the activity of the impressed forces can be re-expressed as

$$\mathbf{e}\cdot\mathbf{J}+\mathbf{h}\cdot\mathbf{G}=\mathbf{E}\cdot\mathbf{J}+\mathbf{H}\cdot\mathbf{G}+\nabla\cdot[(\mathbf{E}-\mathbf{e})\times(\mathbf{H}-\mathbf{h})]. \qquad (5.8)$$

According to eqns. (5.4) and (5.5), the two first terms correspond to the Joule heat and the energy stored in the field. The vector $(\mathbf{E}-\mathbf{e})\times(\mathbf{H}-\mathbf{h})$ must therefore represent the energy flux. With this generalization of Poynting's theorem Heaviside met his own criterion of intelligibility: 'It is a necessity of a rationally intelligible scheme (even if it be only on paper) that the transfer of energy should be explicitly definable.'[62]

5.5.5 Moving bodies

The original form of the duplex equations did not include bodies in motion, for Heaviside was mostly interested in problems of propagation along conducting lines. However, he was well aware of the necessity of terms depending on the velocity of matter. He even was the first Maxwellian physicist to give an exhaustive list of these terms. In 1885 he noted the $\mathbf{v}\times\mathbf{B}$ contribution to the electric force $\mathbf{E}$, as a consequence of Faraday's law applied to moving circuits. Similarly, he introduced a $\mathbf{D}\times\mathbf{v}$ contribution to the magnetic force $\mathbf{H}$: if a displacement current could magnetize a body at rest, then the motion of the body with respect to a constant field of displacement should also have a magnetizing effect. Heaviside then examined the activity of these 'motional forces.' For the electric one, the activity $\mathbf{J}\cdot(\mathbf{v}\times\mathbf{B})$ exactly balances the work $\mathbf{v}\cdot(\mathbf{J}\times\mathbf{B})$ of the electrodynamic force $\mathbf{J}\times\mathbf{B}$. In order to obtain a similar balance in the magnetic case, Heaviside introduced a new 'magnetoelectric force' $\mathbf{D}\times\mathbf{G}$ (with $\mathbf{G}=\partial\mathbf{B}/\partial t$) which was sufficiently small to have eluded observation.[63]

[61] Heaviside 1885–1887: 449, 451. Cf. Yavetz 1995: 154–62. Maxwell had introduced impressed forces only at the level of linear circuits. In Hertz's later formulation of Maxwell's theory, the impressed forces were included in Ohm's law, not in the circuital equations.

[62] Heaviside 1885–1887: 450; 1886–1887: 172.

[63] Heaviside 1885–1887: 448, 446 (with a numerical estimate of the magnetic motional force), 545–6; 1886–1887: 175.

Heaviside also included convection currents in his scheme. Both Faraday and Maxwell admitted the magnetic action of such currents, and Rowland proved it experimentally in 1875 by testing the action of a rapidly rotating charged disk on a compass needle. There was, however, no mathematical treatment of the effects of charge convection in Maxwell's theory, until in 1881 the young J. J. Thomson published an inspired but flawed paper on this subject (see Appendix 10). J. J. Thomson's motivation was to determine the electromagnetic behavior of the charged particles which constituted cathode rays according to Crookes. Adopting Maxwell's displacement, he reasoned that charge convection implied a varying displacement and a corresponding magnetic field. As long as the particle's motion is slow enough, the only change in the electric field is a uniform translation of its lines of force. The intensity of the corresponding magnetic field is proportional to the velocity of the charge, and therefore its energy is proportional to the square of the velocity, which means an increase of the effective mass of the charged particle. In an external magnetic field, there is an interaction energy proportional to the velocity and to the external field, and a corresponding deflecting force (our Lorentz force). These were essential results, to which later electrodynamicists frequently referred.[64]

Unfortunately, the relevant calculations suffered from the slavish following of Maxwell that Heaviside condemned. J. J. Thomson uncritically maintained Maxwell's expression of the displacement current, and sneaked his way around the resulting contradictions. Luckily, he reached the correct form of the final formulas; but the numerical coefficients were wrong. FitzGerald soon showed the necessity of a new contribution $\rho\mathbf{v}$ to Maxwell's dielectric current, where ρ is the charge density and $\mathbf{v}$ the velocity of the electrified matter. It was left to Heaviside to give, in 1885 and 1889, the correct expression of the Lorentz force ($q\mathbf{v}\times\mathbf{B}$), and the correct electromagnetic mass formula for a uniformly charged spherical shell ($q^2/6\pi a$ in rationalized electromagnetic units, with q for the charge, and a for the radius). Heaviside gave the clearest justification of the $\rho\mathbf{v}$ term: it meets Maxwell's requirement that all currents should be closed. Indeed, the divergence of the total current,

$$\nabla\cdot\left(\frac{\partial\mathbf{D}}{\partial t}+\rho\mathbf{v}\right)=\frac{\partial\rho}{\partial t}+\nabla\cdot(\rho\mathbf{v}), \tag{5.9}$$

is identically zero, since charge is conserved during its convection (see Appendix 10).[65]

In their most complete form involving all cases of motion, Heaviside's duplex equations are

[64] Faraday, *FER* 1: #1644, #1654; Maxwell 1873a: ##769–70; Helmholtz 1876 (report on Rowland); Rowland 1878; J. J. Thomson 1881a. Cf. Buchwald 1985a: 74–77 (on Rowland); 269–76 (on theories of convection); Darrigol 1993a: 287–8 (on Maxwell), 303–6 (on J. J. Thomson).

[65] J. J. Thomson 1881a; FitzGerald 1881; Heaviside 1885–1887: 446; 1889. Cf. Buchwald 1985a: 272–3.

$$\nabla\times(\mathbf{H}-\mathbf{D}\times\mathbf{v}-\mathbf{h})=\mathbf{j}+\frac{\partial\mathbf{D}}{\partial t}+\rho\mathbf{v},$$
$$\nabla\times(\mathbf{E}-\mathbf{v}\times\mathbf{B}-\mathbf{e})=-\frac{\partial\mathbf{B}}{\partial t}. \quad (5.10)$$

These equations only determine the electric and magnetic forces. In order to determine the mechanical forces, Heaviside used the principle of activity. The result is in general ambiguous, because there are many different manners to write the energy balance. We will see later how Heaviside found in Hertz a solution to this difficulty.[66]

5.5.6 *The crop*

With his usual computational power and his 'redressed' Maxwell, Heaviside solved numerous problems of wave propagation that later proved very useful, and answered fundamental questions raised by other theorists of electricity. One of his most impressive achievements was the general solution he gave in 1888 to J. J. Thomson's problem of a point charge in uniform rectilinear motion. He found that the electric field was still radial, but compressed toward the meridian plane, to an extent determined by $(1 - v^2/c^2)^{-1/2}$ (see Appendix 10). When the velocity of the particle increases from zero to c, the field evolves from an electrostatic field to an electromagnetic plane wave confined in the meridian plane. Heaviside used this result to show the '*physical* inanity' of the electrostatic potential and reject a suggestion by William Thomson for measuring its propagation.[67]

Originally, no one paid much attention to Heaviside's difficult and lengthy series, except for the insulted telegraph authorities, who managed to have them suspended. In 1887, however, Heaviside convinced William Thomson of the pertinence of his theory of distortionless transmission. The following year, he had the pleasure to read Lodge remarking on 'what a singular insight into the intricacies of the subject, and what a masterly grasp of a most difficult theory, are to be found among the eccentric, and in some respects repellent, writings of Mr. Oliver Heaviside.' Heaviside soon joined the epistolary circle of Lodge, FitzGerald, and Hertz, and convinced them of the superiority of his rendering of Maxwell. FitzGerald was most eloquent in his praise:[68]

Maxwell, like every pioneer who does not live to explore the country he opened out, had not had time to investigate the most direct means of access to the country, or the most systema-

[66] Heaviside 1886–1887: 174–5 (without the convection current); 1888–1889: 497 (with the convection current).

[67] Heaviside 1888–1889: 490–9; 1889c: 510–11. Cf. Hunt 1991a: 186–7; Darrigol 1993b: 313, 316–318. On Thomson's suggestion, see *supra*, p. 178, note 5.

[68] Lodge 1888a: 236; FitzGerald 1893: 299. On Heaviside's publication difficulties and their relation with his quarrel with William Preece, cf. Hunt 1991a: 137–43; Yavetz 1995: 242–56. On his recognition, cf. Hunt 1991a: 143–51; Yavetz 1995: 259–63 (on W. Thomson's support). Lodge's interest in Heaviside's papers was related to his recent experiments on lightning: see below, p. 204.

tic way of exploring it. This has been reserved for Oliver Heaviside to do. Maxwell's treatise is cumbered with the *débris* of his brilliant lines of assault, of his entrenched camps, of his battles. Oliver Heaviside has cleared those away, has opened up a direct route, has made a broad road, and has explored a considerable tract of country.

5.6 Electromagnetic waves

Great though they were, the British Maxwellians missed the discovery of electromagnetic waves, which is now regarded as the most definitive proof of Maxwell's system. Maxwell himself was remarkably silent on the production of electromagnetic waves. He discussed the characteristic spectrum of a substance in terms of a 'disturbance of the luminiferous medium communicated to it by the vibrating molecules,' with no mention of anything electromagnetic. Plausibly, he did not believe that purely electromagnetic processes could generate waves. His doctrine of closed currents indeed obscured the propagation of interactions. By putting conduction and displacement currents on the same footing, he confused sources and their effects. Considered as a function of the total current, his vector potential did not propagate.[69]

5.6.1 The question

Oliver Lodge, the first man to anticipate the electric production of electromagnetic waves, did not reason in terms of the misleading form of Maxwell's equations. His inspiration came from a primitive version of the cogwheel model, which he described at the 1879 meeting of the British Association. There he assumed the ether to be positive and negative electricity bound together (the two kinds of wheels), and interpreted light as a periodic displacement of the two electricities, with an electrostatic restoring force. The view was opposite to Maxwell's, for it made displacement depend on electric forces, and not vice versa. Yet it suggested that light could be excited electrically. Lodge imagined several experimental devices—none of which would have worked, as we can now judge. His best guess was to use the oscillatory discharge of a condenser, although the frequency of light waves could never have been reached in this manner.[70]

Lodge soon abandoned his project, because FitzGerald convinced him that Maxwell's theory forbade the electric production of electromagnetic waves. In a first paper FitzGerald gave two different impossibility proofs, one based on the non-propagation of Maxwell's vector potential as a function of the total current, the other

[69] Maxwell 1875. Chalmers 1873b; Hunt 1991a: 28–30. However, Maxwell referred to Faraday's 'Thoughts on ray vibrations' (Faraday 1846: 450), according to which the sudden motion of an electrified or magnetized body would produce transverse vibrations of the emerging lines of force (Maxwell 1864: 194). On early, uninterpreted observations of electromagnetic radiation, cf. Süsskind 1964.

[70] This is based on Hunt's reconstruction from the following unpublished documents: Lodge to FitzGerald, 26 and 29 February 1880, quoted in Hunt 1991a: 31–2; Lodge to Larmor, 1 January 1902, which contains extracts from Lodge's notebooks for 1879–80. Cf. Hunt 1991a: 30–33.

on the conservative character of a system of closed currents. In a second paper, he confirmed the lack of propagation with a standing-wave solution of the wave equation for the vector potential. Three years later, he found a similar problem treated in Rayleigh's *Theory of Sound*, with progressive solutions that expressed the emission of waves! FitzGerald had to apologize 'for having ventured to investigate these matters when [he] was so ignorant of what had already been done as to make mistakes requiring such serious corrections as are contained in this paper.' He admitted to having erred in his impossibility proofs by including the displacement current in the sources of the field.[71]

To complete the volte-face, FitzGerald suggested that electromagnetic waves could be produced in measurable amount by discharging a condenser through a circuit of small resistance. The following year, he published the retarded vector potential formula, with a discouragingly small estimate of the energy radiated by an oscillating current loop. In his notebooks he calculated the oscillation frequency of simple circuits, and discussed various ways of detecting the emitted waves. But he did not persevere. As Heaviside regretted, 'he saw too many openings. His brain was too fertile and inventive.' Worse, no one followed up the idea, not even his friend Lodge, who had little time for research in this period of his life. Had there been sustained efforts to produce and detect the waves, they would not necessarily have met success. None of the detecting procedures imagined by FitzGerald would have worked; and that later used by Hertz was based on an unexpected property of the electric spark, as we will later see.[72]

5.6.2 Waves on wires

In early 1888 Lodge, whose skill as a scientific speaker was well known, was asked to lecture on lightning protection. In order to simulate thunderbolts, he discharged Leyden jars through a spark gap. For the sake of visibility, he fed the Leyden jars continuously with a powerful electrostatic machine. His arrangement is represented in Fig. 5.8 (without the dotted line L for the moment). The jars stand on the same, low-conducting, wooden table. At the beginning of a cycle, the Voss machine (top) charges the two Leyden jars slowly until the breaking tension of the gap A is reached. The resulting spark short-circuits the gap A, so that a potential difference appears at the gap B. As long as the latter gap is not too wide, the jars discharge through it. All sparking ceases at the end of this process, and a new cycle can begin.[73]

[71] FitzGerald 1879b; 1880; 1882: 101. Cf. Hunt 1991a: 33–2.

[72] FitzGerald 1882: 100; 1883a; Heaviside, in FitzGerald 1902: xxvi. Cf. Hunt 1991a: 46–7. Retarded potentials had already appeared in Lorenz 1867, but in a different context: cf. Chapter 6, pp. 212–13. FitzGerald's imagined oscillator was of the dipolar magnetic kind, which is very inferior to Hertz's electric dipolar oscillator. In 1884 J. J. Thomson discussed the emission of electromagnetic waves by a perfectly conducting spherical shell returning to electric equilibrium (J. J. Thomson 1884c). He judged these waves to be undetectable because they were emitted too suddenly (in a few periods only, with a wavelength comparable to the radius of the sphere). This may explain why experiments of this kind were not attempted at the Cavendish Laboratory.

[73] Lodge 1888a: 234.

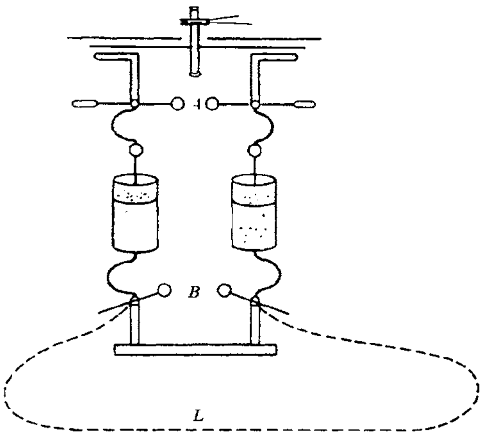

FIG. 5.8. Lodge's experiment of the alternative path (Lodge 1888a: 234).

With this lightning simulator, Lodge proceeded to compare different lightning conductors. He introduced conductors L (see Fig. 5.8) of various shape and constitution, and determined the minimal size of the gap B for which the jars preferred to discharge through the alternative path L. For an audience accustomed to reason exclusively in terms of ohmic resistance, the results were quite counter-intuitive. Even when the resistance of L was a small fraction of an ohm and the gap B was as wide as A, the discharge preferred the gap. Lodge explained this fact by the self-induction of the wire L, which obstructed quickly varying currents. He was more surprised to find out that an iron wire led the discharge better than a similar copper wire. His tentative explanation was twofold. First, the high magnetic permeability of iron, which should have enhanced the self-induction, did not exist for fast-varying currents. Second, iron was better than copper because Heaviside's skin effect, which increases the resistance by confining the current to the surface of the conductor, was more important for the better conductor, copper. Lodge then provided a striking proof of the skin effect by showing that flat conductors conducted the discharge better than round ones.[74]

Lodge knew well that the discharge of a Leyden jar through a small resistance was oscillatory, and he expected the same to be true for the discharge of clouds through lightning. His conclusions largely depended on the high frequency of the oscillations, which enhanced the effect of self-induction. At the Bath meeting of 1888, William Preece, the *bête noire* of the Maxwellians, maintained against Lodge the received wisdom of lightning protection. For this once Preece was right: later studies proved the non-oscillatory character of lightning, and thus ruined a good deal of Lodge's conclusions.[75]

[74] Lodge 1888a: 235–6.
[75] Cf. Hunt 1991a: 146–51; Yavetz 1996 (for the later evolution of the subject).

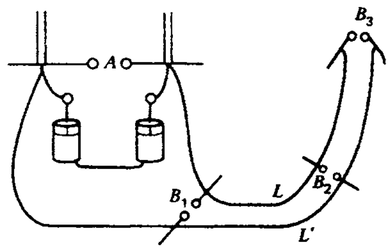

FIG. 5.9. Lodge's 'recoil-kick' experiment (Lodge 1888a: 275).

Aside from the lightning problem, Lodge's focus on oscillatory discharge bore interesting fruit. In a variant of the alternative-path experiment, represented in Fig. 5.9, he had the jars concurrently discharge through the gap A and through a discharger B bridging two long wires attached to the poles of the machine. He expected the sparking at B to cease whenever the gap B was larger than the gap A. Experiment decided differently. Moreover, this sparking proved stronger when the discharger B was farther from the source. Lodge suspected a resonance phenomenon. The high sparking at B, he propounded, corresponded to the 'recoil kick' of reflected waves, and increased when the length of the leads approached the half-wave length. In further experiments showing nodes and anti-nodes on longer wires, Lodge estimated the waves to be 30 yard long. This number agreed with his own estimate of the frequency of the oscillator from its capacity and self-induction.[76]

For continental physicists, Lodge's experiments had little theoretical significance, for they could be interpreted in terms of Kirchhoff's waves of electricity in wires. In contrast, from a Maxwellian point of view Lodge had done no less than produce electromagnetic waves by electrical means. He published his beautiful results in the early summer of 1888, and went on vacation to the Tyrol. During his train ride to the Alps, he read the latest issue of Wiedemann's *Annalen*, and found that a young German physicist, Heinrich Hertz, had obtained 'much better and more striking evidence of these electromagnetic waves.' Lodge swallowed his bitterness, and soon proclaimed his joy over this splendid development.[77]

5.7 Conclusions

Maxwell left his electromagnetic theory in a state full of imperfections and obscurities. The superiority of his views was not self-evident to his contemporaries. The highest British authority on electricity, Sir William Thomson, disliked Maxwell's theory for it ventured far from empirical facts without offering a mechanical repre-

[76] Lodge 1888a: 275; Lodge 1888b, 1888c. Cf. Aitken 1985: 89–95; Hunt 1991a: 148–9. This version of the alternative path was the most obvious to realize, because the commercial Voss machine came with the two Leyden jars already attached to it (for storing the electricity).

[77] Lodge 1888c: appendix written in Tyrol, dated 24 July 1888. Cf. Hunt 1991a: 153.

sentation of basic field processes. According to Thomson, Faraday's concept of charge and Maxwell's displacement current were unwarranted extensions of a partial analogy between vacuum and dielectrics; and there were other ways to deal with open currents. without leaving the conceptual framework of transmission lines. Between the practical concerns of the telegraph and the ideal of a simple elastic solid ether, Thomson tolerated no *via media*. Consequently, he condemned Maxwell's new style of theoretical physics.

Through his indispensable treatise on electricity and magnetism, Maxwell nevertheless managed to transmit his and Faraday's views to a few British physicists. Some of these assiduously perfected and extended his system until it won, in the late 1880s, the preference of most English-speaking electricians. In this process Maxwell's theory was significantly transformed, and acquired several features that are now judged central, for example the four-equation formulation, the Poynting flux, and the electric production of electromagnetic waves.

There were several kinds of Maxwellian works. In the most conservative, Maxwell's equations were blindly applied to computable versions of old problems, for example Arago's disk or rotating conducting spheres. This involved solving systems of differential equations with simple boundary conditions, and therefore incited much Cambridge Tripos activity. To spare the technical details, the ensuing publications have not been discussed in this chapter. However, they contributed to establish a Maxwellian paradigm, and they helped clarify some issues. For example, in 1887 Horace Lamb explained that Maxwell's scalar potential was completely determined by Maxwell's equations alone and generally differed from the electrostatic potential given by Poisson's equation. Maxwell, Larmor, and J. J. Thomson had previously missed this important point.[78]

The most important clarifications of Maxwell's system were obtained either by mechanical pictures or by dynamical, energetic considerations. With his cord-and-beads model, Lodge illustrated Maxwell's concept of charge and current, and various processes in simple electric systems. With wheels and rubber bands, FitzGerald showed that Maxwell's displacement did not have to be the linear displacement of some substance. More likely, this quantity corresponded to local strains of a different kind. Then the electric current no longer resembled the flow of an incompressible fluid. A displacement current meant a variation in the strain of the mechanism transfering rotational motion in the field; a conduction current implied a slipping of this mechanism.

As Lodge and FitzGerald acknowleged, their models were good only to illustrate some aspects of Maxwell's theory. They were, for instance, unable to explain electrostatic attractions. Yet the two friends' ambition was to determine the ultimate constitution of the ether. Somewhat naïvely, Lodge believed that a system of cogwheels, or something similar, was a useful step in this direction. More philosophically, FitzGerald hoped to reduce electromagnetism and optics to the motions of an ideal

[78] Lamb 1887; Maxwell 1873a: #783; Larmor 1884a; J. J. Thomson 1884a. Cf. Darrigol 1993b: 294–7.

fluid. His mathematical power and his rare physical intuition did not suffice, however, to bring the project to fruition.

Not every British physicist shared Lodge's and FitzGerald's trust in mechanical models and pictures. Poynting and Heaviside instead relied on dynamical concepts that had more direct empirical significance. They both determined the energy flux in the electromagnetic field as a necessary consequence of Maxwell's equations and field-energy distribution. They described conductors as sinks and guides for the energy traveling in the surrounding dielectric. For Poynting, the primitive dynamical notion was the motion of tubes of force, which provided an intuitive understanding of basic field processes despite the lack of a mechanical foundation. The role of a conductor, for example, was to guide and partially dissolve the tubes of force moving in the surrounding dielectric.

For Heaviside, the basic dynamical notions were generalized force and velocity, controlled by the 'principle of activity' borrowed from Thomson and Tait. Heaviside required that the activity of the forces **E** and **H** at a given point—that is, their product by the corresponding current—should determine the energy stored and dissipated at this point. He directly expressed the field equations in terms of forces and currents, and completed them to include all cases of bodily motion. Maxwell's potentials were gone, 'as a hip of metaphysics.' In Heaviside's eyes the loss of Maxwell's Lagrangian foundation was largely compensated by a better insight into practical problems. A former telegrapher and a proudly independent thinker, he developed his own efficient methods to study signal propagation. He invented the vector notation, our impedance, inductance etc., and a version of the operational calculus. In his hands Maxwell's theory became more transparent, more complete, and more ready-to-use.

Other Maxwellians pursued the relation between light and magnetism. Old and new effects of that kind offered an opportunity to explore vortical rotation in the magnetic field and eventually to confirm Maxwell's electromagnetic theory of light. FitzGerald started the trend with a remarkable theory of Faraday's and Kerr's magneto-optical effects. His work could be read in a variety of instructive ways: as an indication that MacCullagh's strange rotational medium was the only plausible concept of a mechanical ether, as a proof that the reflection and refraction of light could be treated in a purely electromagnetic manner, and as a general strategy for modifying Maxwell's equations to integrate new, non-linear effects. In the third register, FitzGerald's prescription was to add new terms to the electromagnetic field Lagrangian, so that dynamical principles would be automatically satisfied. The method had powerful adepts in Cambridge, where the abstract dynamics of Thomson, Tait, and Maxwell had a growing influence. It culminated with Glazebrook's noting, in 1881, that FitzGerald magneto-optical Lagrangian term could be justified in a purely electromagnetic manner in relation to the Hall effect. For Maxwellian physicists, this remark meant a major confirmation of Maxwell's theory of light.

Glazebrook's theory did not survive further magneto-optical research in the 1890s. The essentially macroscopic approach with field Lagrangians and effective field equations proved insufficient, and had to be replaced with microphysical considerations. This is not to say that Maxwellian theorists of the 1880s always avoided atomistics. On the contrary, Oliver Lodge and J. J. Thomson devoted much time to the atomistic periphery of Maxwell's system. This Maxwellian microphysics, and its evolution into a different kind of microphysics will be treated in Chapter 7.

By the late 1880s, FitzGerald, Poynting, Lodge, Heaviside, and other British Maxwellians had convinced a large number of English-speaking physicists and electricians that electric fluids and direct action at a distance should be replaced with more philosophical and truly dynamical field notions. The coincidence between the electrostatic/electromagnetic charge units ratio and the velocity of light and the success of derived magneto-optical theories increased the plausibility of Maxwell's theory. Yet direct electromagnetic proofs of its superiority were lacking. Perhaps the Maxwellians were too convinced of the truth of Maxwell's system to pursue such crucial experimenting. The best Lodge did was to show electric waves on wires, which Kirchhoff's theory predicted just as well as Maxwell's.