real fluids in their solution. Even among these writers there was little consistency; they employed very different versions of ether theory, their choice reflecting both wider philosophical issues and also the current state of scientific knowledge. Moreover, those who used ethers for theological purposes extended across almost the whole spectrum of religious opinion; among those discussed here are fundamentalists, Christian mystics, High Anglicans, liberal Anglicans, dissenters, spiritualists, and atheists. Yet this diversity of uses to which ethereal fluids were put suggests that they may be viewed as a scientific resource that could fulfil any of a number of functions within a specific range of philosophical systems. There were, of course, many writers, such as Joseph Priestley, David Brewster, and Thomas Reid, who explicitly rejected ethers from their scientific, philosophical, and theological systems. With this division of opinion, questions concerning ethereal fluids often feature in controversies that spanned scientific and theological issues. The following discussion alludes to some of these controversies, but since the principal concern of this synoptic chapter is to explicate the different theological functions ethers have fulfilled, I will not pay great attention to the scientific, theological, and philosophical commitments of individual protagonists. However, the concluding section suggests how theological attitudes towards ether theories may be related to broader social and intellectual issues.

Ether and natural theology

Throughout the eighteenth century, and less pervasively in the nineteenth, it was widely considered that the universe had been designed by God and as such manifested signs of its wise, intelligent, and good Designer. This natural theological theme appeared frequently in the Boyle Lectures, the Bridgewater Treatises, and a variety of other works, the most famous of which was William Paley's *Natural theology* (1802). In these writings the functions of different parts of the universe were discussed within the total economy, and every part had its pre-ordained role. One traditional argument relating to the role of ether concerned the existence of what appeared to be empty spaces between the stars and planets. Some objected that the Creator did not allow void space, since this would have performed no function. Thus the universe had to be full and the apparent spaces were in fact filled with some form of ethereal fluid, which constituted a plenum.[1]

A more common claim concerning the significance of ether for natural theology was that it was the principal secondary cause by which God governed the normal operations of the universe. For example, for John Cook, an Essex doctor, the discovery of ether was the key referred to in the title of his book *Clavis naturae; or, the mystery of philosophy unvail'd* (1733). In this work

he claimed that God employed ether as the immediate cause of all motion: '*AEther* is the *Rudder* of the Universe, or as the *Rod*, or whatever you will liken it to, in the *Hand* of the *Almighty*, by which he *naturally* rules and governs all *material* created *Beings* . . . Now how beautiful is this *Contrivance* in God'.[2] Similarly, Richard Lovett of the Cathedral Church, Worcester, who equated ether with fire and electricity, argued that

> though God alone is the Author and Preserver of all Things, and which he continually upholds with his immediate Hand; yet, the only instrumental Cause of our Being is this subtil Spirit . . . In a Word, this pure AEther or Fire, contain'd in Air, is the Cause of all Motion, animal and vegetable.[3]

A particularly detailed example of the use of ether within a natural theological framework dates from the early 1830s when William Whewell sought the theological significance of the recently revived wave theory of light. In November 1830, Whewell received a commission from the president of the Royal Society, Davies Gilbert, to write one of the Bridgewater Treatises. Two and a half years later his treatise, *Astronomy and general physics considered with reference to natural theology* (London, 1833), was published. In his section on light, Whewell indicated his commitment to the wave theory and to the existence of a luminiferous ether that produced a wide variety of optical phenomena. Ether, he claimed, was a providentially designed mechanism that enabled man to perceive the physical world through the sense of vision. He suggested furthermore, albeit in a rather speculative manner, that ether was the source of all activity and the cause of chemical, electrical, and vital phenomena. In his view ether constituted one of the three fundamental substances in the universe, the other two being solid matter and aeriform fluids. Whewell therefore argued that owing to the crucial role of ether in the economy of the universe, 'if the world had no ether, all must be dead and inert'. Furthermore, he claimed that if we contemplate the remarkable intricacy of the ethereal mechanism and its importance in the universe, we are forced to acknowledge that it was designed by a 'most wise and good God'.[4]

Whewell also composed a work entitled 'View of the modern theory of light'[5] that may have been a draft for the Bridgewater Treatise but was never published. It was, however, far more than an introductory overview of physical optics, since Whewell not only analysed the significance of the wave theory for natural theology but, more precisely, identified five aspects of the luminiferous ether that were of importance to theology; these were the ether's simplicity, scale, fullness, variety, and harmony.[6] His arguments, which will be discussed in turn, are somewhat naive, however, and tend to overlap.

1. Whewell was impressed by the way in which the luminiferous ether, obeying the simple laws of mechanics, produced a great diversity of optical phenomena. Thus the simplicity of the cause contrasted with the complexity of its effects, and this relationship between the two indicated that the Creator had designed ether in a highly efficient manner.
2. Divine power, Whewell considered, was manifested by the interstellar spaces, which involved distances many orders of magnitude greater than those usually encountered. Similarly, 'the range of density and elasticity which we find in the mechanism of light [i.e., the luminiferous ether] is upon a scale equally extraordinary, or still more so'. Since the ether's elasticity was considerably greater than that of terrestrial bodies and its density very much less than that of steam, these contrasts in scale indicated God's power in being able to create so great a diversity in the universe.
3. The mechanism of the luminiferous ether also indicated to Whewell 'how full the world is of [God's] contrivances' and thus of manifestations of intelligent design. He speculated further about the existence of other micro-mechanisms in space that might produce thermal, magnetic, and electrical phenomena. This argument appears to be a variation on the traditional principle of plenitude.
4. The fourth argument concerned the variety of ethereal fluids in the universe and their design in such a way that each fulfilled a specific function. By considering different examples of action at a distance, such as occur with light, sound, electricity, and heat, Whewell claimed that their means of propagation must be different. This argument for a diversity of ethers (which seems incompatible with argument [1]) indicated specific design and thus a choice of means on the part of the Creator. Furthermore, to Whewell this argument undermined the claims of materialists who considered that phenomena resulted from the necessary properties of matter.
5. Whewell's final point was that from empirical evidence the luminiferous ether appeared to interact with other subtle fluids, such as those responsible for chemical, vital, and electrical phenomena. The connections between these various classes of phenomena indicated an overall providential design and harmony in which the luminiferous ether played an integral part.

Whewell is generally portrayed as a liberal Anglican for whom natural theology, as opposed to revealed religion, was of central importance.[7] He disagreed with the Scottish Evangelical David Brewster on many issues, including theology,[8] Brewster insisting on a literal interpretation of the Bible as opposed to Whewell's latitudinarianism. Morse (1972) has even suggested that Brewster's commitment to an inductivist methodology that omitted unobservable entities (such as ether) from science stems from his Calvinist upbringing, which emphasised the manifest nature of God's works. Brewster's antipathy towards 'broad' natural theology and ether theories can be seen in his published response to Whewell's Bridgewater Treatise.

Brewster received a copy of the book from Whewell soon after it was published. His initial response was civil and complimentary; he wrote to Whewell that it 'has been my companion on two Journies to the Highlands, and I need scarcely say that I never derived more pleasure or instruction from any other book.'[9] However, writing anonymously in the *Edinburgh Review*,[10] he attacked, among other points, Whewell's claim that the luminiferous ether indicated God's design. Brewster did not accept the wave theory of light, and he resented the fact that Whewell had presented that theory as portraying the true nature of light. Moreover, he rejected unobservable fluids like the luminiferous ether that Whewell had identified with the wave theory. On the subject of natural theology, Brewster claimed that its legitimate scope involved only the 'Power, Wisdom, and Goodness of God as [actually] *manifested* in the CREATION'.[11] Thus he considered that it was legitimate to argue from the known structure of the eye to the existence and attributes of its 'Designer'. However, hypothetical entities, like the luminiferous ether, were not appropriate subjects for natural theology, since they could not be considered as true premises of the argument. Moreover, he feared that when the hypothesis was refuted by scientists, as would surely happen in this case, the apologetic argument would readily be exploited by atheists. Thus, Brewster rejected arguments from design employing ethereal fluids.

Ether and matter

Throughout the period under discussion natural philosophers frequently speculated about the structural relationship between ether and gross matter, and the role of God in establishing this relationship. Ether was usually considered a more basic substance than gross matter, and thus it was suggested that ether was the protoplast out of which God had formed matter. The most famous example of this use of ether occurs in a letter of Newton's to Henry Oldenburg published in 1757, although written more than eighty years earlier. Here, Newton speculated that the whole diversity of the physical world had been produced by a variety of 'aethereal spirits, or vapours' in a condensed state. Initially, God brought about this process but subsequently he delegated it to the 'power of nature'.[12] It has been suggested that Newton's concern with the ether as 'protoplast' was related to the alchemical tradition with which he was familiar (Dobbs, 1975).

Two centuries after this letter of Newton's was written, a similar view was articulated with respect to a contemporary scientific theory, the vortex theory of the atom (Silliman, 1963; Siegel, this volume). A number of eminent physicists, such as Lord Kelvin, Peter Guthrie Tait, and James Clerk Maxwell, advocated this theory in which vortical motions in homogenous, incompressible, and frictionless ether constituted the atoms of gross matter. Since these vortex atoms were stable and perfectly elastic they behaved in the same way as the particles postulated by the kinetic theory of gases. Furthermore, the vortex theory appeared to offer a means of explaining all chemical and electromagnetic phenomena within a unified world view in which ether was the fundamental substance. Writing in the ninth edition of the *Encyclopaedia Britannica*, Maxwell expressed the view that the vortex theory was a most important and promising scientific hypothesis.[13]

The vortex theory of the atom, together with other ideas drawn from contemporary science, was adopted by Balfour Stewart and Peter Guthrie Tait in their attempt to defend the coherence of science and religion in a popular book entitled *The unseen universe; or, physical speculations on a future state*, published in 1875. They suggested not only that gross matter was composed of ether, but that the particles of that ether were themselves composed of an even more subtle ether, and so on. Furthermore, the subtler higher-order ethers contained greater energy, so that any particular ether was able to form the next ether lower on the subtlety scale. From this thesis they also argued that the subtler the ether, the more stable and long-lasting, and the higher its energy content. In accordance with the purpose of the book – 'to show that the presumed incompatibility of Science and Religion does not exist'[14] – Stewart and Tait postulated that their scale of ethers rose towards God and tended to become identical with his attributes of omnipresence, omnipotence (having infinite energy), and existence for all time. Thus in their scheme they employed a recessive scale of protoplasmic ethers that bore a specific relationship to God. Thus, God's act of creating our world involved the localisation of energy in the formation of gross matter out of the lowest-order ether.[15]

Stewart and Tait's idea of a scale of ethers between God and matter may have been derived from the popular concept of the Great Chain of Being. A more direct source, however, and one that they cited, was a passage in Thomas Young's *A course of lectures on natural philosophy and the mechanical arts* (Cantor, 1970). Young, who came from a Quaker family but who became an Anglican primarily for careerist reasons, was committed to the

existence not only of matter but also of such 'immaterial' substances as the electric fluid and

> either caloric or a universal ether; higher still perhaps are the causes of gravitation, and the immediate agents of attractions of all kinds . . . It seems therefore natural to believe that the analogy may be continued still further, until it rises into existences absolutely immaterial and spiritual. We know not but that thousands of spiritual worlds may exist unseen for ever by human eyes.[16]

In contrast to Young, other writers, such as Oliver Lodge (see the section of this chapter entitled 'Ether as intermediary'), maintained that only one ether was required, which performed both physical and theological functions.

Ether and dynamics

Mechanical philosophers in the seventeenth century, such as Descartes, employed ethereal fluids filling space as the immediate cause of all motion and activity. Indeed, the idea of motion without a mover in contact with the moving body was deemed impossible and incomprehensible. It has often been contended that this theory of contact action was displaced early in the eighteenth century by Newtonian central forces acting at a distance. However, throughout the eighteenth century and well into the nineteenth there were a number of natural philosophers who rejected action at a distance and instead adopted a contact-action plenal ether. Many of these writers were theologically motivated and claimed the Bible as the source of their natural philosophy. Although they did not form a coherent group, a number of these authors were intellectually indebted to John Hutchinson (1674–1737), who had attacked Newton's natural theology and natural philosophy early in the eighteenth century (Metzger, 1938; Wilde, 1980). We shall take as our example of this genre the book by Samuel Pike entitled *Philosophia sacra; or, the principles of natural philosophy extracted from divine revelation* (1753). Pike (1717?–73) was connected with the dissenting ministry and, towards the end of his life, appears to have been associated with the Sandemanians – a sect that later numbered Faraday among its members.

Unlike Hutchinson, Pike found no conflict between Newton and the physical system revealed in the Bible.[17] He considered that Newton had discovered the law of gravitation but not the underlying cause, which was motion in the pure and subtle ether filling the heavens. In Pike's cosmology there was no void space; all actions resulted from impulsion derived from this ether, which underwent a cyclical, mechanical, and everlasting motion. The inert particles of ether were projected at high speed away from the sun and towards the circumscribing firmament, but as they approached it they slowed down, congealed into larger clusters, and returned to the sun. Here the clusters were broken down and the individual particles reemitted. On this theory gravitation was explained by the pressure produced on bodies by the moving clusters of ether particles. Likewise, the planetary motions and rotations were explained by the forces produced by both the individual particles flowing away from the sun and the clusters moving in the opposite direction. Similarly, Pike explained cohesion, electrical, and optical phenomena on mechanistic principles.

As Pike's title suggests, he attempted to 'extract' the whole of his natural philosophy from the biblical text. Indeed, he conceived his book as contributing to the reconciliation between science and what he understood to be a fully biblical theology. Adopting an ingenious but somewhat crude philological method, Pike reinterpreted several biblical passages and 'derived' from them his theory of a mechanical ether. For example, he translated the Hebrew word *SHa'HaKIM* (skies) as meaning 'the *strugglers* or *aethers in conflict*'.[18] From analysis of this kind he claimed that Scripture teaches us that all natural causes and effects are mechanical.

Central to Pike's theological concerns was the despiritualisation of the physical world. Neither gross matter nor the ether particles were endowed with intrinsic power. Instead, ethereal impulse not only solved the problem of action at a distance but also allowed the whole of the physical world (since the Creation) to remain ontologically distinct from God, in whom alone power resided. Yet, Pike argued, this mechanical universe was not independent of God. At the Creation, God formed both matter and ether and gave motion to the machine. Furthermore, he maintained the mechanism and could at will interfere with its running, for example, by performing miracles.[19]

It is, perhaps, paradoxical that while a fundamentalist writer like Pike employed a contact-action ether to explain phenomena, many natural theologians considered the attempt to explain gravitation by purely material causes tantamount to atheism. The spectre of materialism, traditionally associated with the names of Epicurus, Lucretius, and Hobbes, raised its head again at the turn of the century with the publication of the *Méchanique celeste* (1798–1825) of Laplace. For many this work became associated not only with atheism but also with the social upheaval of the French Revolution. Samuel Vince, professor of astronomy and experimental philosophy at Cambridge, was one of those perturbed by French atheism, in response to which he wrote his *Observations on the hypotheses which have been assumed to account for the cause of gravitation from mechanical principles* (1806). In his preface, Vince made clear his concern:

> Many of the most eminent Philosophers upon the Continent have been endeavouring to account for all the operations of nature upon merely mechanical principles, with a view to exclude the Deity from any concern in the government of the system, and thereby to lay a foundation for the introduction of Atheism.[20]

Vince surveyed various ether hypotheses and proceeded to 'demonstrate' that no hypothesis specifying the elasticity and density of an ethereal medium could account for the inverse square law of gravitation. He then argued that the failure of all ether hypotheses to explain both gravitation and the stability of the celestial system indicated that God did not act by material causes; instead, God must act directly on matter in accordance with the law of gravitation. Finally, Vince claimed that contemplation of this universal God-given law led to appreciation of the power, wisdom, and goodness of the Creator.

A number of other writers also contended that no quiescent or impelling fluid could account for gravitation. Alexander Crombie, for example, adopted the fairly standard distinction between the law of gravitation as a descriptive rule about the behaviour of bodies and the idea of gravity as a causal agent.[21] Like Vince and Thomas Reid (see the next section), Crombie accepted the former but not the latter. There was, however, a dissenting voice raised against Vince's pamphlet. Writing in the *Edinburgh Review* for 1808,[22] John Playfair, professor of mathematics at Edinburgh, pointed out that Vince's argument was mathematically false, since there were an infinite number of ether hypotheses from which the inverse square law could be deduced. Playfair, who was a minister in the Church of Scotland, accepted that it was perfectly legitimate to inquire into the cause of gravitation and that such research should not be constrained by ideological considerations. Moreover, he was 'convinced that the issue of this argument is quite immaterial to the truths of natural religion, which must rest on the same immoveable foundation, whether the physical cause of gravity be discovered, or not'.[23] Playfair, then, rejected the connection that Vince claimed existed between ether and atheism, and he appears to have viewed Vince's pamphlet merely as the reaction of a frightened man to political developments in France.

Consideration of the effect of a ubiquitous stationary ether on the progress of the world is related to this controversy. Unlike Pike's cosmology, which ensured that no physical cause could terminate the solar system, those natural philosophers who adopted a stationary ether were faced with the theological implications of the ether's effect on planetary motions and the ultimate decay of the world. As Schaffer (1977) has shown, Edmund Halley in 1693 recognised that ether's resistance would lead to the secular acceleration of the planets and result in their spiral trajectories into the sun. However, Schaffer claims, Halley did not draw any strong theological conclusions from this argument. Newton, by contrast, abandoned altogether the idea of a Cartesian ether, since he demonstrated its incompatibility with celestial motions. His own concept of the stationary ether – as a rare, tenuous, and highly elastic fluid – was reintroduced in the 1717 edition of the *Opticks,* where he confronted the problem of the effect it would produce on planetary motions. In solving, or rather sidestepping, this problem he showed that if ether were rarer and more elastic than air by factors of 700,000, its presence 'would scarce make any sensible alteration in the Motions of the Planets in ten thousand Years'.[24] Thus, for Newton, ether by itself might produce decay in the world, but the rate of decay would not be observable.

Two very different theological arguments intersect in considering the history of the solar system. Within a natural theological context the stability of the system would indicate God's intelligence in creating a perfect mechanism. Though this view was logically compatible with a creationist theology it might also be employed by the atheist, who would argue that the physical world, being self-sufficient, has existed from eternity. The other factor concerned the physical implications of the creationist doctrine. Although a steady-state solar system did not imply that it was divinely created, a decaying system (without a compensatory process such as that envisaged by Kant) unequivocally implied creation at some past time and the gradual death of the world at a future time. Some theological purchase could thus be obtained from both of these theories, but particularly the theory of decay. Ethers were occasionally employed as a cause of this decay, although there were many other agents that might produce a similar result. However, this role for ether took on new significance in the 1820s and 1830s when the theory of the stationary ether gained support from two empirically based arguments. One was the interpretation of the wave theory of light that required ether to be stationary with respect to the sun and stars in order to account for stellar aberration. The other arose from the secular acceleration of Encke's comet, which was widely interpreted as the effect of resistance by an ethereal fluid. It was the latter that particularly attracted the attention of William Whewell. In his Bridgewater Treatise of 1833 he extrapolated from this explanation of Encke's comet in order to predict the behaviour of the earth and other heavy bodies. He argued that although the resistance to the earth's motion would be slight, nevertheless it would inevitably lead to its destruction, albeit in the distant future. This argument provided further evidence for the 'universal law of decay', which, Whewell believed, dictated changes in both the earth and the cosmos – in accordance with his interpretation of the nebular hypothesis[25] – and also in human societies. For Whewell the decay of the

world had a specific theological implication: that it must also have had a beginning. In turn, this necessarily implied the existence of God, the Creator, 'a First Cause which is not mechanical'.[26]

Ether as intermediary

An important theme in Western philosophy has been the doctrine of dualism in which two distinct substances, mind and matter, are postulated in order to distinguish God from the material universe and our souls (or minds) from our bodies. Although these two substances are considered to be incommensurable they must be able to interact with each other, for example, in the acts of perception and volition. However, this dualistic ontology poses the problem of how these two radically different substances – the one incorporeal, penetrable, and intelligent and the other extended, impenetrable, and passive – are able to affect one another. Dualists, in proposing a variety of solutions to this problem, have sometimes adopted an ethereal fluid as the intermediary between mind and matter. An attraction of this solution is that ethereal fluids, as sometimes defined, partake of some of the qualities of both mind and matter; for example, both activity and extension have simultaneously been attributed to ether.

Probably the most frequently cited and most influential example of the ether as intermediary appeared in queries 23 and 24 to Newton's *Opticks* (1717). Here, Newton proposed the scarcely original theory that an ether fills the nervous system. In his explanation of vision, Newton suggested that rays of light falling on the retina cause this medullary ether to vibrate, and this vibration is then transmitted along the optic nerve 'into the place of sensation'. Likewise, in explaining how we move parts of our bodies, Newton considered that the 'power of the Will' causes the ether in the brain to vibrate. This vibration in the medullary ether is then propagated along the capillaries of the nerves and produces movement through either contraction or dilation of the muscles.[27] Both of these explanations were based on the hypothesis that Newton's ether could interact with both matter and mind, an issue that is discussed in some of his unpublished writings (McGuire, 1968).

Among those who adopted Newton's physiological use of ether was David Hartley, whose *Observations on man* first appeared in 1749. Hartley's ether was, however, unequivocally a form of matter, and hence this theory raised the problem of how a material ether could interact with an immaterial soul. Aware that this was a sensitive issue, Hartley was at pains to assert that his theory neither challenged the doctrine of the soul's immortality nor attributed sensation to brute matter. He explained that he merely posited vibrations in the medullary ether as physiological correlates to our sensations, ideas, and motions.[28] Despite his statements on this issue, other authors attributed to Hartley the view that the mind was merely matter in motion. Among these detractors was the Scottish dualist Thomas Reid, who contrasted the humble way in which Newton proposed a few open-ended conjectures about ether with the atheistic implications of Hartley's fully fledged system. Reid complained that since we have no direct evidence for the existence of ether, let alone its vibrations, Hartley had employed the illegitimate method of hypothesis. In comparison with God's infinite intelligence, man could not comprehend the extent of God's works, and it was therefore presumptuous of Hartley to have framed a hypothesis: Hypotheses were the creation of 'human imagination, [and] not the work of God'. Reid's other theological offensive concerned causality. He claimed that science was not concerned with efficient causes but that the scientist (*pace* Newton) should instead seek lawlike relations between observables. Hartley's ether, which mediated between the sensory organs and the soul, functioned as a spurious efficient cause of sensation, since for Reid the immediate source of sensation 'must be resolved into the will of God'. Any causal explanation of the interaction between mind and matter was theologically unacceptable, because this interaction was known only to God.[29]

Our final example of the use of ether as intermediary takes us to the early part of this century. Sir Oliver Lodge, the eminent physicist, attempted to solve the mind–body problem in a manner similar to Newton's two centuries earlier. Lodge's interest in the problem grew out of a general concern for a reaffirmation of Christianity in the face of modern science and rampant atheism. Moreover, after the death of his son in the World War I, Lodge became progressively involved with the problem of life after death, which he related to his long-standing interest in spiritualism. More precisely, he connected the central Christian doctrine of the immortality of the soul with the idea that we also possess an immortal 'spiritual body'. After death, which is only the death of the physical body, a person's soul remains localised in the 'spiritual body'. The existence of this 'spiritual body' is asserted in 1 Corinthians 15:44: 'There is a natural body, and there is a spiritual body'. Furthermore, Lodge employed this theory in his account of a number of 'spiritual' phenomena, particularly communication with the dead. In elaborating his theory he gave prominence to the ubiquitous ether, which he also utilised in his physics. He considered the ether a physical substance differing from both gross matter and spirit (soul) but able to interact with each. Moreover, he maintained that the 'spiritual body', which is immortal, is composed of ether. Hence every person has an etheric or 'spiritual body' in which the spirit (or soul) is localised: 'Our spiritual and real home is in the ether of space'. In putting forward this theory, Lodge offered a solution not only to the traditional mind–body problem, but also to the problems of psychical research. Thus, for example, the ethereal body carrying the soul of a departed person is able to be present at a seance and to affect material objects by means of localised energy derived from ether.[30]

Lodge's interest in spiritualism reflected the late nineteenth- and early twentieth-century concern with a range of psychic phenomena. A large number of eminent scientists – including Lodge, J. C. Adams, A. R. Wallace, Lord Rayleigh, J. J. Thomson, and W. Crookes – were concerned with or actively investigated these phenomena, and many joined the Society for Psychical Research founded in 1882. Several recent studies have shown that many scientists of the period held deep-seated religious views that they attempted to integrate with contemporary scientific theories. Frequently ethereal fluids performed this mediating role (Wilson, 1971; Heimann, 1972; Kottler, 1974; Turner, 1974; Wynne, 1977, 1979).

Ether and spirit

A common view, particularly during the eighteenth century, was that matter was inert and thus some superimposed active agent was required in order to account for all types of phenomena ranging from gravitation and cohesion to animal growth. Frequently an ethereal fluid was proposed as the source of this activity. However, when ether was considered a purely material substance, the problem arose of explaining its own activity. To overcome this difficulty, Bryan Robinson suggested that 'Spirit' supplied this activity in the form of a repulsive force between the particles of a Newtonian elastic ether.[31] Another writer, Abraham Tucker, employed a 'spiritual substance', presumably identical to the neo-Platonist theory of the mundane soul, to account for the activity of ether.[32] Others, including George Berkeley and Michael Ramsay, claimed that ether was intrinsically active.

A particularly important and influential discussion of the active powers in nature (Ritterbush, 1964) appeared in Bishop Berkeley's *Siris* (1744).[33] Here, Berkeley maintained that an ether, which he equated with fire and the substance of light, was the secondary cause responsible for animating the universe. Yet Berkeley's ether was not conceived in mechanistic terms but was instead an active animate entity. Thus, Berkeley claimed that ether is 'the vegetative soul or vital spirit of the world' and is infused with divinity. In propounding this view he drew in part on early eighteenth-century chemical philosophy and particularly on Boerhaave's discussion of fire as the active agent. More significantly, he cited at length the *prisca sapientia* tradition[34] in which spiritualized ethers had been propounded by the ancient Egyptians, Greeks, and Persians and also in the Hermetic corpus and the Bible.

A slightly earlier writer who drew extensively on the *prisca sapientia* tradition was the Chevalier Ramsay, a Scot who spent most of his life in France and who wrote a popular exposition of the ancient theology in the form of a fictional travelogue. The second book of Ramsay's *The travels of Cyrus* (1727) concerned Cyrus's visit to the Magi of Persia. The Magi could supposedly disengage themselves from matter and thus communicate directly with the spiritual realm. Zoroaster, the Archmagus, initiated Cyrus into the hidden mysteries of the physical universe. Here the French and English texts diverge. In the latter, Ramsay outlined Newton's theory of gravitation (which he considered compatible with his ether theory), whereas in the French text he explicated the role of a pure and invisible ethereal fluid. Here he maintained that ether, which he identified with the *primum mobile* of the Greeks, was infinitely divisible and formed numerous fluids that were responsible for all celestial and terrestrial activity, including the motions of stars and planets, gravitation, light, and the growth of plants. Despite Ramsay's reference to *ce systême Cartesien,* it is clear that his ether bore little resemblance to that of Descartes. He denied that ether acted according to the laws of blind mechanical necessity. By contrast, he claimed that ether 'is the *body* of the Great [God] Oromazes, whose soul is truth' and who is the 'first principle of things. He diffuses himself everywhere'.[35] Ramsay allied this belief in a divine function of ether with Newton's suggestion in the final query of the *Opticks* that 'God, being present everywhere by His will, moves all bodies in His infinite, uniform *sensorium,* and so shapes and reshapes according to His pleasure all parts of the universe'.[36]

Not only did Ramsay distinguish his ether from the subtle fluids of the mechanical philosophers; he also explicitly connected it with the *prisca sapientia* tradition. Walker (1972) suggests the identification of Ramsay's pure ether with the neo-Platonist doctrine of the mundane spirit. Moreover, he conjectures that a connection existed between Ramsay and Newton, and this may help explain the revival of Newton's concern with ether late in the first decade of the century.

Whereas Ramsay and Berkeley drew heavily on the ancient theology, other writers explicitly related ether to their Trinitarian beliefs. One such author was the Dublin clergyman Richard Barton, who discussed the analogies between the natural and moral worlds in a published series of lectures entitled *The analogy of divine wisdom* (1750). One topic under discussion was the analogy between

> the infinite divine Spirit or HOLY GHOST, and the UNIVERSAL AETHER or elemental Fire . . . Because as the mechanic philosophers make the Aether the cause of attraction, muscular motion and other extraordinary phaenomena of matter: So is the HOLY GHOST the cause of all spiritual conduct, which is consonant to the divine Law.[37]

Barton identified fire with ether, and also with electricity, and he cited Berkeley's *Siris* in order to show not only the role of fire in the world economy, but also that fire was the spirit of the world. In explicating the analogy between ether and the Holy Ghost, Barton pointed out that the Bible likened the appearance of the Holy Ghost to fire (Acts 2:3) and that Christ was able to baptise with fire and the Holy Ghost (Matthew 3:11). Another analogical connection was that ether functioned as 'a principle instrument in the conservation of the material world', and similarly the divine spirit had conserved religion and virtue in the moral realm. Furthermore, both ether and the Holy Ghost were universally present, yet both were unequally diffused. Thus while some material objects contained an excess of fire, some men by their virtuous acts had become more endowed with divine spirit. Barton's more general claim was that although our knowledge of ether and its functions was far from complete, science had shown that the whole economy of nature depended on ether. Likewise, he conceived a parallel dependence of the moral world on the divine spirit. From these and other analogies, Barton attempted to show the essential coherence between natural philosophy and divine revelation.[38]

In contrast to Barton, John Hutchinson not only drew analogies between ether and each of the three persons in the Trinity but also found evidence for his Trinitarian beliefs in the functions of the contact-action ether that he claimed was the source of all observed motions. Hutchinson believed that parallels existed between the divine realm and the structure of the physical world, such that the former could be appreciated by an analysis of the latter. Moreover, the key to understanding both was the Bible. In Hutchinson's physical theory the material world was governed by a single ethereal substance that was capable of three modifications: fire, light, and spirit (air). The analogue of this physical situation involved the three persons of the Trinity within the single Godhead (Elohim). In Hutchinson's scheme fire was related to the Father, light to Christ, and spirit (air) to the Holy Ghost. Furthermore, these relationships were sanctioned by their functional similarities and by lexical relations in the original Hebrew.[39]

Although Hutchinson rejected those natural philosophies that attempted to 'spiritualise' matter and instead insisted that the whole physical universe consisted of matter in motion, he was strangely silent about the human soul, which appears to have played no part in his ontology. Paradoxically, his emphasis on man as a mechanism, albeit one designed by God, brought him close to the traditional kind of materialism that has often been associated with atheism. The programme to reduce mind (and the soul) to the action of material particles has been articulated on numerous occasions both within and beyond the time period covered in this chapter. A central problem facing adherents to this programme, and one quickly seized on by their antagonists, has been the explanation of how inert matter could produce those kinds of activity that dualists considered belonged only to mind. However, one response available to materialists that went some way to meet this objection employed one or more ethereal fluids. To illustrate this strategy we will examine the theory proposed in a biography of Lucretius, published in 1805, by the surgeon and classical scholar John Mason Good.

Good, who appears to have been a Unitarian, clearly believed in the existence of God and in the truth of the Scriptures. However, he maintained that Lucretius had been correct in claiming that whereas man's body is composed of gross matter his soul is made up of a more refined, subtle form of matter. Certainly, he admitted, the soul was different from gross matter, but this did not imply either the soul's immateriality (which was not mentioned in the Bible) or the inertness of all matter. There was much evidence to the contrary: Matter manifested gravitation, and magnetic and chemical action; and more highly organised matter in animals and vegetables was subject to irritability. From animal sensation it was but a small step for Good to claim that the soul was material, but composed of a very subtle and volatile ethereal matter. In explicating this theory he pointed to Lucretius's doctrine that

> . . . the mind, in every act, we trace
> Most voluble, from seeds of subtlest size,
> Rotund and light, its mystic make must spring.[40]

Moreover, Good maintained that modern experimentally based physiology expounded a related doctrine involving the medullary ether and that there was no good reason why this theory, which explained the constitution of the nerves, should not be extended to include the mind. Thus he envisaged that the mind could modify the medullary ether.[41]

Even if Good was no atheist, his theory became associated with the materialist programme expounded by atheists.[42] Indeed, the belief that the mind or soul was composed of matter – probably a subtle form of matter – was expounded by such self-confessed atheists as Julien de la Mettrie and Baron d'Holbach (French, 1969).[43] In response to this challenge orthodox dualists recurred to the argument used by Samuel Clarke in his controversy with Henry Dodwell: Particles of matter irrespective of their subtlety or arrangement cannot have the property of consciousness. In Clarke's words: '*Consciousness* therefore cannot at all reside in the Substance of the *Brain*, or *Spirits*, or in any other *material System* as its *Subject;* but must be a Quality of some *Immaterial Substance*'.[44] Indeed, for many 'moderate' clergymen, gross matter and soul were sufficient and there was no need for recourse to ethereal fluids.

Concluding remarks

Having surveyed the theological functions that ethereal fluids fulfilled during the eighteenth and nineteenth centuries, we may now attempt to draw some general but tentative conclusions. One such conclusion concerns the temporal distribution of the writers cited. With the exception of Whewell's Bridgewater Treatise, all the works discussed herein that advocate ethereal fluids date from either before 1810 or after 1875. An explanation is required for this distribution (assuming that it does not result from an inadequate search of nineteenth-century literature). This earlier period corresponds roughly to the widespread interest among speculative natural philosophers in the unifying role of ethereal fluids (see Heimann, this volume). Moreover, the majority of our examples date from the mid-eighteenth century, the period when the ether theories of Newton and Boerhaave attracted most attention. After about 1830 a major scientific concern was the construction of mathematical and mechanical models of the luminiferous ether. However, the ether theories developed in this context were unlikely candidates for fulfilling four out of the five functions discussed in the section 'Ether and natural theology', although they still could be employed in the argument from design, which was their principal function in Whewell's unpublished manuscript. The late Victorian resurgence of interest is perhaps explained by the 'crisis of faith'[45] that affected many scientists, who then frequently attempted to reconcile science and religion through the somewhat simplistic strategies offered by spiritualism.

The following brief and very tentative comments, which relate to the earlier period, attempt to locate specific groups that tended to employ ethers for theological purposes. Initially, however, it may be helpful to reiterate the point that several different versions of ether theory have been discussed during the course of this chapter. Within this range two distinct types of ether may be identified; one class comprises 'animate' ethers, which were closely allied to spirit, and at the opposite end of the spectrum are mechanistic ethers, which were intended to 'despiritualise' nature. Those writers, such as neo-Platonists, mystics, and certain High Anglicans, who emphasised the immanence of God in an active universe and who rejected the notion of passive matter found the first class of ether theory acceptable but positively rejected the second. Among those discussed here, Berkeley, Ramsay, and Barton fall into this category, and it was also principally this group of writers who stressed the ancient roots of their theory. By contrast, contact-action ethers implied that God did not act directly on each particle of matter. This type of theory was most acceptable to those like Hutchinson and Pike for whom the Bible was the principal source of all spiritual and physical knowledge. This kind of ether was also acceptable to atheists, as was the related theory attributing all activity to the smallness or organisation of ether particles and denying the existence of immaterial souls. This last type of theory, together with others that emphasised ether as the 'active principle' in nature, appears to have appealed particularly to Low Churchmen such as Young and Good, and it may not be inappropriate to include Hartley and even Newton in this group. (Priestley, a Unitarian, likewise attempted a monistic synthesis, although not founded on ether theory.)

Missing from this discussion and underrepresented in the examples cited in this chapter is one significant cluster of theological writers – what might be called the 'moderate' or 'broad' church faction. The members of this group tended to be dualists who firmly distinguished mind from matter. This ontology denied the existence of active ethereal matter. Moreover, dualists usually considered God the immediate source of all activity in nature, so that no role was allowed for unobservable subtle media. It is no coincidence that most of those cited herein as objecting to ethereal fluids on theological grounds were dualists of the 'moderate' persuasion.

The theological functions of ether have too rarely been discussed by historians, who have usually concentrated on scientific theorising. Yet not only does this aspect of ether fail to exhaust its historical role, it also does not take us far in understanding why some writers were enthusiastic champions of ether while others vehemently refused to countenance its deployment or existence. The notion of an ether diffused through space involved far-reaching ontological and epistemological commitments (Laudan, this volume). Moreover, for those scientists deeply concerned with natural theology or the theology of nature, ether has helped solve such problems as the relation between God and the physical universe or the possibility of life after death. Whether this resource has been utilised, in preference to other strategies, has depended on individual theological and philosophical commitments and the position of ether theory in contemporary science.

The problematic relation between science and theology gives rise to two opposing theses. One posits the primacy of theology and conceives ether theorising as simply an extension or elaboration of an already held theology. The other suggests that ether theorising arises principally from the scientific arena and that scientists subsequently – often in old age – explore their extrascientific functions. Examples could be found to confirm each of these theses (Barton would be an example of the first and Lodge of the second), with the latter receiving more support from the nineteenth- than the eighteenth-century cases discussed in this chapter. At the same time many – both scientists and others – held strong predispositions for or against ether on theological grounds. In contrast to these two extreme positions, a more helpful general model would involve a dynamic interaction between scientific theorising and theology, with ether potentially providing an important element in this interaction.

One particular connection has recently received attention from Shapin (1980), who suggests that matter theory should be interpreted as reflecting the perceived social order: Thus the dualist symbolises in his matter theory – through the total separation of mind and matter – the distinction between classes in society, whereas Priestley is interpreted as having collapsed both mind and matter into a monistic theory analogous to his political egalitarianism. In his interesting but speculative scheme, Shapin concentrates on just one of ether's many functions, that of mediating between mind and matter. The social significance of this type of ether theory is not explained in detail by Shapin but perhaps reflects a stratified, socially mobile society. To what extent this thesis is helpful in understanding Lodge or indeed Newton – the socially mobile lad from Lincolnshire – is open to further research. However, this type of scheme, with significant modifications, has been utilised indirectly by Christie (this volume) in discussing the secular context of the work of Adam Smith, David Hume, and William Cullen. At the same time the evidence offered in the present chapter precludes the conclusion that ether was always associated with secularism, or indeed atheism.

While it may be difficult to identify *the* theological function of ether, ether theory may perhaps best be conceived as a resource that could help solve some recurrent problems in natural theology and the theology of nature. The present chapter has attempted to delineate five such functions employed by eighteenth- and nineteenth-century writers, and in this concluding section several general theses, founded on a limited sample of cases, are offered, albeit tentatively.

Acknowledgments

The author would like to thank John Hedley Brooke and Jonathan Hodge for their helpful comments.

Notes

1 E. Chambers, 'Aether', *Cyclopedia; or, an universal dictionary of arts and sciences*, 4th ed., 2 vols. (London, 1741), 1:n.p.
2 J. Cook, *Clavis naturae; or, the mystery of philosophy unvail'd* (London, 1733), 284–6.
3 R. Lovett, *The subtil medium prov'd; or, that wonderful power of nature so long ago conjectur'd by the most ancient and remarkable philosophers, which they call'd aether but oftener elementary fire, verify'd* (London, 1756), 64–5.
4 W. Whewell, *Astronomy and general physics considered with reference to natural theology* (London, 1833), 141.
5 W. Whewell, 'View of the modern theory of light', Whewell Papers, Trinity College, Cambridge: R. 18. 17[3]. This appears to have been written in the early 1830s.
6 Ibid., 195–215.
7 W. B. Cannon, 'Scientists and broad churchmen: an early Victorian intellectual network', *Journal of British Studies 4* (1964), 65–88; J. H. Brooke, 'Natural theology and the plurality of worlds: observations on the Brewster–Whewell debate', *Annals of Science 34* (1977), 221–86; Cantor (1975); Morse (1972).
8 Brooke, 'Natural theology'; Cantor (1975); Morse (1972).
9 D. Brewster to W. Whewell, 10 June 1833, Whewell Papers: Add. MS a. 201[92]. I am grateful to the Librarian of Trinity College, Cambridge, for permitting me to quote from this letter and also from Whewell, 'View'.
10 [D. Brewster], review of Whewell, *Astronomy*, *Edinburgh Review 58* (1834), 422–57.
11 Ibid., 428.
12 T. Birch, *The history of the Royal Society of London*, 4 vols. (London, 1756–7), 3:250. See also Heimann (1973 and this volume).
13 J. C. M[axwell], 'Atom', *Encyclopaedia Britannica*, 9th ed., vol. 3 (Edinburgh, 1875), 36–49.
14 [B. Stewart and P. G. Tait], *The unseen universe; or, physical speculations on a future state* (London, 1875), vii.
15 Ibid., 154–72.
16 T. Young, *A course of lectures on natural philosophy and the mechanical arts*, 2 vols. (London, 1807), 1:610–1.
17 S. Pike, *Philosophia sacra; or, the principles of natural philosophy extracted from Divine Revelation* (London, 1753), viii, 1–14.
18 Ibid., 15–33; G. N. Cantor, 'Revelation and the cyclical cosmos of John Hutchinson', in *Images of the earth: essays in the history of the environmental sciences*, ed. L. Jordanova and R. Porter (Chalfont St. Giles, 1979), 3–22.
19 Ibid., 77.
20 S. Vince, *Observations on the hypotheses which have been assumed to account for the cause of gravitation from mechanical principles* (Cambridge, 1806), 4.
21 A. Crombie, *Natural theology; or, essays on the existence of a Deity and of providence, on the immateriality of the soul and a future state*, 2 vols. (London, 1829), 1:106–7.
22 [J. Playfair], review of Vince, *Observations*, *Edinburgh Review 13* (1808), 101–16.
23 Ibid., 103. See also C. Maclaurin, *An account of Sir Isaac Newton's philosophical discoveries* (London, 1748), 389.
24 I. Newton, *Opticks*, 2nd ed. (London, 1717), 327.
25 Brooke, 'Natural theology'.
26 Whewell, *Astronomy*, 191–209. Cf. B. Powell, *The connexion of natural and Divine truth; or, the study of the inductive philosophy considered as subservient to theology* (London, 1838), 164–6.

27 Birch, *History*, 252–4; Newton, *Opticks*, 328.
28 D. Hartley, *Observations on man, his frame, his duty and his expectations*, 5th ed., 2 vols. (Bath, 1810), 1:34, 525–6.
29 T. Reid, *The works of Thomas Reid, D.D., with notes and supplementary dissertations by Sir William Hamilton, Bart.*, 8th ed., 2 vols. (Edinburgh, 1895), 1:248–53. Related aspects of Reid's thought are discussed by Laudan (1970).
30 See, particularly, O. Lodge, *My philosophy representing my views on the many functions of the ether of space* (London, 1933).
31 B. Robinson, *A dissertation on the aether of Sir Isaac Newton* (Dublin, 1743), 122.
32 A. Tucker [E. Search, pseud.], *The light of nature pursued*, 2 vols. (London, 1768), 2:pt. 2, 81.
33 G. Berkeley, *Siris: a chain of philosophical reflections and inquiries concerning the virtues of tar-water, and divers other subjects connected together and arising from one another* (Dublin, 1744), 69–105.
34 This theme recurs in many of the works cited in this chapter and also in L. Dutens, *An inquiry into the origin of the discoveries attributed to the moderns: wherein it is demonstrated, that our most celebrated philosophers have, for the most part, taken what they advance from the works of the ancients; and that many of the important truths in religion were known to the pagan sages* (London, 1769), 181–92. See also Walker (1972); J. E. McGuire and P. O. Rattansi, 'Newton and the "Pipes of Pan"', *Notes and Records of the Royal Society 21* (1966), 108–43.
35 A. M. Ramsay, *The new Cyropaedia; or, the travels of Cyrus: with a discourse on the theology and mythology of the ancients* (London, 1760), 73–85.
36 Walker (1972), 255–6. Cf. Newton, *Opticks*, 379.
37 R. Barton, *The analogy of divine wisdom, in the material, sensitive, moral, civil, and spiritual system of things, in eight parts* (Dublin, 1750), 62.
38 Ibid., 55–66.
39 J. Hutchinson, *Glory or gravity essential and mechanical*, in *The philosophical and theological works of the late truly learned John Hutchinson, esq.*, 12 vols. (London, 1748–9), 6:21–33.
40 Lucretius, *The nature of things: a didactic poem: translated from the Latin of Titus Lucretius Carus, accompanied with the original text, and illustrated with notes philological and explanatory* [*by J. M. Good*], 2 vols. (London, 1805), 1:405–7; J. M. Good, *The book of nature*, 3 vols. (London, 1826), 3:1–32.
41 Lucretius, *The nature of things*, 1:lxxxv–xciii.
42 J. Buchanan, *Faith in God and modern atheism compared, in their essential nature, theoretical grounds, and practical influence*, 2 vols. (Edinburgh, 1864), 2:70–134.
43 J. O. de la Mettrie, *Man a machine*, 2nd ed. (London, 1750); P. T. d'Holbach, *The system of nature: or, laws of the moral and physical world*, 2 vols. (Boston, 1889), 1:50, 62–3.
44 S. Clarke, 'A second defence of an argument made use of in a letter to Mr. Dodwell, to prove the immateriality and natural immortality of the soul', in S. Clarke, *The works of Samuel Clarke, D.D.*, 4 vols. (London, 1738), 3:799.
45 O. Chadwick, *The Victorian church*, 2 vols. (London, 1966–70), 2:112–50; A. Gauld, *The founders of psychical research* (London, 1968).

5

The medium and its message: a study of some philosophical controversies about ether

LARRY LAUDAN
Center for Philosophy of Science, University of Pittsburgh, Pittsburgh, Pennsylvania 15260 USA

It is by now a commonplace that the emergence of scientific theories has sometimes occasioned extensive philosophical discussion of the conceptual well-foundedness of the ideas on which such theories depend.[1] The eighteenth-century controversies about action at a distance and the nineteenth-century atomic debates are obvious cases in point. It has not, however, been widely appreciated that ether theories during their heyday produced a philosophical discussion at least as profound as, and probably more far-reaching than, those associated with atomism and noncontact action. My aim in this chapter is to explore briefly some of the philosophical aspects of 'the ether debates', with a view to documenting their impact, both on the fortunes of subtle fluid physics and on the nineteenth-century revision of the philosophy of empiricism.

Taking the long view, the ether debates erupted intermittently from Aristotle to Lorentz and Fitzgerald, but the special features of those debates that I shall discuss appear between 1745 and 1850. It is chiefly this period that exhibits a very striking interaction between ethereal physics and empiricist epistemology, an interaction that is the focus of this study. My claim is that during this period the character of this interaction shifted profoundly in ways that were to modify both science and philosophy.

The central theses of this chapter are these:

1. The epistemology prevalent in the second half of the eighteenth century was altogether incompatible with the various ether theories that emerged in the natural philosophy of that period.
2. Some of the early proponents of ethereal explanations chose to abandon or modify that prevailing epistemology so as to provide a philosophical justification for theorising about ether.
3. The modifications so introduced were unconvincing and inadequate,

leaving the scientific status of ether theories very unclear by the beginning of the nineteenth century.

4. The emergence of the optical ether in the early nineteenth century prompted a more radical critique of classical epistemology, a critique that produced some highly innovative and historically influential methodological ideas.

The first phase, 1740–1810

Our story should begin, as any account of Enlightenment epistemology must, by recalling the triumph of Newtonian mechanics and the trenchant inductivism associated with Newton's achievement. As numerous authors have shown, the half century following publication of the *Principia* was marked by a growing antipathy to hypotheses and speculations.[2] Induction and analogical reasoning were all the rage and Newton's doctrine of *verae causae* – adumbrated in dozens of eighteenth-century glosses on his first *regula philosophandi* – was thought to exclude any entity or process not strictly observable.[3] Whether we look to Berkeley and Hume in Britain, to 'sGravesande and Musschenbroek in the Netherlands, or to Condillac and D'Alembert in France, the refrain was similar: Speculative systems and hypotheses were otiose; scientific theories had to deal exclusively with entities that could be observed or measured. For half a century, many natural philosophers sought to develop theories satisfying those demanding strictures; 'moral philosophers' (e.g., Berkeley, Condillac, and Hume), for their part, explored the logical and epistemological ramifications of this new view of the nature of *scientia*.

However, long before epistemologists of science were able to digest these new challenges to the traditional demonstrative ideal of science, scientific developments themselves conspired to produce a significant shift. For, especially during the period from 1745 to 1770, many emerging theories within the sciences moved well beyond the inductive, observational bounds imposed by erstwhile Newtonians. Nowhere is this clearer than with respect to the development of mediumistic or ethereal explanations. In the 1740s alone, there were at least half a dozen major efforts to explain the behaviour of observable bodies by postulating a variety of invisible (and otherwise imperceptible) elastic fluids. In 1745, Bryan Robinson published his *Sir Isaac Newton's account of the aether*. A year later, Benjamin Wilson's *Essay towards an explication of the phaenomena of electricity, deduced from the aether of Sir Isaac Newton* appeared. Of greater moment, Benjamin Franklin developed his account of electricity as a subtle fluid; the Swiss physicist George LeSage articulated an ethereal explanation of gravity and chemical combination; and the highly controversial David Hartley embarked on a programme, culminating in his *Observations on man* (1749), to give a mechanistic theory of mind and perception, whose crucial ingredient was the transmission of vibrations in a subtle fluid or ether through the central nervous system. By the 1760s, the scientific literature abounded with ethereal explanations of heat, light, magnetism, and virtually every other physical process.

Two general points about these developments are especially relevant for our purposes. In the first place, by the 1770s, ethereal or subtle fluid explanations were very widespread among natural philosophers (with the exception, soon to be explained, of many Scottish scientists). Second, such explanations invariably violated the prevailing epistemological and methodological strictures of the age, strictures that, as already noted, would not countenance the use of theoretical or 'inferred' entities to explain natural processes. (After all, an entity that is regarded as in principle unobservable is scarcely consistent with an empiricist epistemology that restricts legitimate knowledge to what can be directly observed.)

Indeed, there was scarcely any domain of scientific theorising in the eighteenth century that left as much scope for speculative hypotheses about unseen agents as did ether theories. As Joseph Priestley remarked in his *History and present state of electricity:*

> Indeed, no other part of the whole compass of philosophy affords so fine a scene for ingenious speculation. Here the imagination may have full play, in conceiving of the manner in which an invisible agent produces an almost infinite variety of visible effects. As the agent is invisible, every philosopher is at liberty to make it whatever he pleases, and ascribe to it such properties and powers as are most convenient for his purpose.[4]

Not a tolerant epistemology at the best of times, classical empiricism (by which I mean the empiricism of Berkeley, Condillac, Hume, and Reid) left no scope for entities like ether. A few natural philosophers of the period failed to perceive the tension between the received epistemology and ethereal theorising. Leonhard Euler, for instance, could simultaneously maintain that the transmission of light depended upon vibrations in an imperceptible medium and insist, in his *Letters to a German princess*, that science should proceed by enumerative induction, eschewing all nonobservable entities. But most scientists and philosophers of the time saw the strain between the emergence of subtle fluid theorising on the one hand and the subscription to naive inductive empiricism on the other. Among this latter group some – such as Thomas Reid – were persuaded that epistemological doctrines took priority over physical theories and thus should be allowed to legislate fluid theories out of the scientific arena. Others, like Hartley and LeSage, saw that option as self-defeating and preferred instead to seek to develop new and more liberal versions of empiricist epistemology that could sanction subtle fluid theories. I want to examine both of these reactions.

The partisans of ether

Among the most persistent, not to say the most notorious, proponents of subtle fluid theories in the last half of the eighteenth century were David Hartley and George LeSage. Hartley foresaw many explanatory roles for the subtle, elastic fluid that he called ether: among them, explaining the transmission of heat, the production of gravity, electricity, and magnetism. Hartley's central concern, however, was to utilise ether, or rather vibrations within ether, to explain a large range of problems about perception, memory, habit, and other activities of the mind. On Hartley's view (which was a detailed elaboration of one of the more speculative conjectures of Newton's *Opticks*),[5] the brain and nervous system are filled with a highly subtle fluid that transmits vibrations from one point in the perceptual system to another. The vibrations in this ether, which are initiated by some external stimuli, subsequently cause the medullary matter composing the nerves and brain to vibrate in ways that are characteristic of the stimuli in question. In his *Observations on man* (1749), Hartley utilised the vibrations in the 'nervous' ether to explain a remarkably divergent range of phenomena, including 'sensible pleasure and pain' (pp. 34–44), sleep (45–55), the generation of simple and complex ideas (56–84), voluntary and involuntary muscular motions (85–114), the sensation of heat (118–25), ulcers (127), paralysis (132–4), taste (151–79), smell (180–90), sight (191–222), hearing (223–38), sexual desire (239–42), memory (374–82), and the passions (368–73). Indeed, most of the five-hundred-odd pages of part I of the *Observations on man* constitute a litany of phenomena that are explicable by Hartley's hypothesis of a vibratory ether.

As Hartley was perfectly aware, the whole structure of his argument was radically out of step with the prevailing inductivist temper of the age. Nowhere did Hartley attempt to 'deduce the ether from the phenomena'; nowhere could he use Baconian techniques of eliminative induction to establish the epistemic credentials of his enterprise. Neither could he point to any direct (i.e., noninferential) evidence for the existence of his ubiquitous subtle fluid; hence it is no *vera causa*. Rather, what Hartley had to settle for was a kind of post hoc confirmation. His arguments invariably have the structure of Peircean 'abductions':

Here is a phenomenon x
But if there were an ether, then x

(Probably) there is an ether

In short, straightforward hypothetico-deduction was the official methodology of the *Observations on man*.

I need hardly add that therein lay the trouble. For however tolerant later generations were to be about post hoc confirmations and the abductive schema, Hartley's contemporaries – as he knew full well – viewed hypothetical reasoning as inherently fallacious (after all, abduction *is* a form of the fallacy of affirming the consequent). Hartley's primary defence of his procedures involved a stress on the wide range of confirming instances that his theory could lay claim to. He suggested that its broad explanatory scope compensated for the unobservability of its explanatory agents and mitigated its failure to exhibit a traditional inductive warrant. As he remarks early in the *Observations on man:* 'Let us suppose the existence of the aether, with these its properties, to be destitute of all direct evidence, still, *if it serves to explain a great variety of phenomena, it will have an indirect evidence in its favour by this means*'.[6] Hartley likened the search for deep-structural theories like his own to the process of decoding a message. Just as the decypherer's task is to find a code that will render the encoded message intelligible, the natural philosopher's job is to find some hypothesis that will save the phenomena. The latter, like the former, must be content with indirect evidence: 'The decypherer judges himself to approach to the true key, in proportion as he advances in the explanation of the cypher; and this without any direct evidence at all'.[7] But Hartley must have known that this analogy, baldly stated, would not take him very far. If he was to establish that it was scientifically respectable to speak about unobservable entities for which there could be no direct evidence, then he would have to reorient the epistemological convictions of an age which took the view that, where hypotheses were concerned, indirect evidence was no evidence at all.

To that end, he composed a long section of the *Observations* dealing with 'Propositions and the nature of [rational] assent'. The unambiguous aim of that methodological excursus was to show that enumerative and eliminative induction are not the only routes to knowledge. He began the section on a then familiar note, to wit, that the methods of induction and analogy are the soundest methods of inquiry in natural philosophy. Indeed, he went so far as to give an associationist and vibrationist account of the mechanisms whereby repetitions of particular instances of a generalisation reinforce one another so as to habituate us to accept the generalisation of which they are instances.

Apart from the relatively strong assimilation of the method of induction to the calculus of probabilities, Hartley was here covering ground that was familiar territory to his contemporaries.

But Hartley went on to insist that the techniques of induction and analogy do not exhaust the methodological repertoire of the natural philosopher. There is, in addition, the *method of hypothesis*. Apropos of the Newtonian insistence that 'hypotheses have no place in experimental philosophy', Hartley was uncompromising: 'It is in vain', he explained, 'to bid an enquirer form no hypothesis. Every phenomenon will suggest something of this kind'.[8] Those who pretend to make no hypotheses are deceiving themselves and confuse their speculative hypotheses with 'genuine truths . . . from induction and analogy'.[9] Since the mind willy-nilly forms hypotheses when confronted by any phenomenon, it is far better to acknowledge tentative hypotheses than to acquiesce unwittingly in them: 'He that [explicitly] forms hypotheses from the first, and tries them by the facts, soon rejects the most unlikely ones; and, being freed from these, is better qualified for the examination of those that are probable'.[10]

Moreover, Hartley insisted, the examination and testing of hypotheses, even false ones, has the heuristic advantage of leading us quickly to the discovery of new facts about the world that we would be otherwise unlikely to discover:

> The frequent making of hypotheses, and arguing from them synthetically, according to the several variations and combinations of which they are capable, would suggest numerous phenomena, that otherwise escape notice, and lead to *experimenta crucis*, not only in respect of the hypothesis under consideration, but of many others.[11]

But even granting (as some of the most trenchant inductivists were willing to)[12] that hypotheses can be of heuristic value, Hartley was still confronted with the problem of explaining the circumstances under which it is legitimate, contra inductivism, to accept or believe a hypothesis involving unobservable entities. Unless it could be shown that there are some circumstances in which a hypothesis – not generated by inductive methods – warrants acceptance, Hartley's own programme for a hypothetical science of mind would be without foundation. Hartley himself propounded the conundrum:

> But in the theories of chemistry, of manual arts and trades, of medicine, and, in general, of the powers and mutual actions of the small parts of matter, the uncertainties and perplexities are as great, as in any part of science. For the small parts of matter, with their actions, are too minute to be the objects of sight; and we are neither possessed of a detail of the phenomena sufficiently copious and regular, whereon to ground an [inductive] investigation; nor of a method of investigation subtle enough to arrive at the subtlety of nature.[13]

It is disappointing that, after much fanfare, Hartley's defence for believing a speculative hypothesis that explains many phenomena took him no further than the early pages of *Observations on man* had done. Invoking again the cypher analogy, Hartley merely insisted that if a hypothesis is compatible with all the available evidence, then that hypothesis 'has all the same evidence in its favour, that it is possible the key of a cypher can have from its explaining that cypher'.[14] In a nutshell, Hartley's method of hypothesis boils down to the claim that a hypothesis warrants belief if it has a large number of known positive instances and no known negative instances. Confirmed explanatory scope thus functioned for Hartley as the decisive criterion for the acceptability of hypotheses.

Hartley immediately conceded that this criterion does not guarantee that the hypotheses it licenses will be true or even that they will stand up to further testing. They will possess none of the reliability (then) associated with the methods of induction and analogy. But, given the inevitability of hypotheses, what (he seems to ask) is the alternative? As he puts it, 'the best hypothesis which we can form, i.e. the hypothesis which is most conformable to all the phaenomena, will amount to no more than an uncertain conjecture; and yet still it ought to be preferred to all others, as being the best that we can form'.[15]

Not surprisingly, this epistemology carried little weight with most of Hartley's inductivist contemporaries. As they could point out, there were many rival systems of natural philosophy that – after suitable ad hoc modifications – could be reconciled with all the known phenomena. The physics of Descartes, the physiology of Galen, and the astronomy of Ptolemy would all satisfy Hartley's criterion. There was, in Hartley's approach to the epistemology of science, nothing that would discredit the strategy of saving a discarded hypothesis by cosmetic surgery or artificial adjustments to it.[16] As his critics pointed out, the great Newtonian epistemological innovation involved the insistence that 'saving the phenomena' or merely explaining the known data was an insufficient warrant for accepting a theory. Neither Newton nor his followers would quarrel with the view that it was a *necessary* condition of the acceptability of a theory that it must fit all the available data;[17] but they would not brook Hartley's transformation of this plausible *necessary* condition of theoretical adequacy into a *sufficient* condition for theory acceptance. Moreover, it was quite clear to any perceptive reader of Hartley's epistemological writings that they were meant to rationalise his ethereal neurophysiology.

Accepting the former meant acquiescing in the latter, which few were willing to do.

If Hartley chose to take on the inductivists somewhat obliquely, conceding to them that induction and analogy were sound modes of inference, expecting (but not receiving) in return an admission that the method of hypothesis, too, had its use, a more direct frontal attack on the prevailing epistemology came from another partisan of ether, Hartley's Swiss contemporary George LeSage. By his own account, LeSage discovered in 1747 the theory that was to make him alternatively acclaimed and notorious for well over a century. LeSage's approach involved postulating a medium surrounding all bodies. The corpuscles constituting this ether move in all directions and occasionally impact upon the particles constituting observable physical objects. The latter are 'semipermeable' to streams of ethereal particles; that is, most ethereal particles will pass completely through a macroscopic object (chiefly because the volume of its constituent particles is always a very small proportion of the space occupied by the body). Some, however, will collide with particles of the body; when they do, there will be appropriate transfers of momentum, with the ethereal particles rebounding and with the atoms of the body moving in the reverse direction. This kinematic ether was utilised by LeSage to explain a wide diversity of phenomena. In an article in *Mercure de France,* he utilised it to explain weight;[18] in his *Essai de chimie mécanique,*[19] it was invoked to explain many phenomena of chemical affinity; still more significantly, he used this approach in 1764[20] and again in 1784[21] to develop his famous kinematical model of gravitation.

The details of LeSage's ether model need not detain us here.[22] It was sympathetically, if critically, evaluated by Maxwell a century ago, and its mathematical and physical articulation was extensively explored by (among others) Preston, Kelvin, Croll, Farr, George Darwin, and Oliver about the same time. For our purposes a very brief summary of LeSage's model should suffice. LeSage explained gravity by assuming that there are streams of ethereal corpuscles flowing into the world from every direction. As already pointed out, ordinary bodies are highly porous and thus most of these *ultramondain* corpuscles move through a body with no interaction. A few, however, will impact with constituent corpuscles of the body. The result of such impacts is that some ethereal particles reverse their direction, while the body itself has a net force exerted on it by the collisions. In a one-body universe, there would be no resultant motion since there would be an equal number of collisions on all sides of the body. But if we introduce a second large object into this universe, each will act as a partial shield against the ultramundane corpuscles. This will lead to a pressure differential, with each body undergoing fewer

collisions on one side than the other; as a result, each body will tend to move towards the other. In this way, the qualitative character of gravitational attraction is attained via contact action rather than action at a distance. The quantitative features of gravity (namely, its relation to the square of the distance and to the masses of the bodies) are explained in LeSage's full-blown *Traité de physique mécanique* (published posthumously by Pierre Prevost).[23] Although many of his predecessors (e.g., Fatio, Daniell Bernoulli) had sought mechanical explanations for gravitational attraction, LeSage's was the only one to emerge from the eighteenth century as a prima facie physically adequate model of gravitational interaction. As James Clerk Maxwell remarked of LeSage's theory: 'Here, then, seems to be a path leading towards an explanation of the law of gravitation, which, if it can be shown to be in other respects consistent with facts, may turn out to be a royal road into the very arcana of science'.[24]

But that reasonably flattering pronouncement by Maxwell is a far cry from the almost universal reaction of LeSage's contemporaries to his ethereal models. Immediately upon publication of his theory (and, since LeSage widely precirculated his ideas, in many cases before publication), it was subjected to a steady stream of abuse. Roger Boscovich called LeSage's system a 'purely arbitrary hypothesis', for which there is no direct proof.[25] The French astronomer Bailly, protesting against LeSage's model because it postulated hidden or unobservable entities, insisted that 'nous ne connoissons la nature que par son extérieur, nous ne pouvons la juger que par celles de ses lois qu'elle nous a manifestées'.[26] During the 1760s, Leonhard Euler wrote several encouraging letters to LeSage about the latter's work. Nonetheless, Euler made it clear that 'je sens encore une trés-grande repugnance pour vos corpuscles ultramondains, et j'aimerois toujours mieux d'avouer mon ignorance sur la cause de la gravité que de recourir à des hypotheses si etranges'.[27]

As these few passages suggest, not only was the reaction to LeSage's model largely negative; the grounds for criticism were generally *epistemological* rather than substantive. LeSage's critics were claiming that hypotheses about unobservable entities were no part of legitimate science. As early as 1755, LeSage complained that the common objection to his system of ultramundane corpuscles was that 'mon explication ne peut être qu'une hypothèse'.[28] In 1770, he worried to a correspondent that unless his system was presented very carefully, 'on le jugeroit sur l'etiquette comme une de ces hypothèses en l'air'.[29]

By 1772, LeSage became convinced that his theory was not getting the hearing it deserved because no one would evaluate it on its scientific merits. Instead, they dismissed it as a mere hypothesis. He claimed it to be an 'almost universal prejudice' of his age that any theory dealing with unobservable entities could not be regarded as genuinely scientific.[30] A decade later, LeSage had become so convinced that he could not get a fair hearing for his views that he withdrew publication of what was to have been his magnum opus. As he bitterly explained:

> Puisque vos physiciens sont si prévenus contre le possibilité d'établir solidment l'existence de mes agens imperceptibles, très-propres cepedant à rendre intelligibles les attractions, affinités, et expansibilités, que constituent à present toute la physique, je suspendrai encore quelque temps la publication des ouvrages que je préparois sur ses agêns.[31]

If the story ended here, it would amount to just one more case of a scientist whose works were suppressed because out of tune with prevailing epistemological fashion. But the LeSage case is more interesting and more important than that, because the difficulties he encountered with his physics prompted him to respond in kind, that is, to articulate a rival methodology to the dominant inductivist one. LeSage did this in a number of philosophical essays, which were designed to show that hypotheses in general and ones involving imperceptible entities in particular may have a sound epistemic rationale. LeSage dealt with this matter at greatest length in a much-quoted essay published posthumously by Prevost. Originally but unsuccessfully intended for publication as an article in the great *Encyclopédie,* the essay was titled 'On the method of hypothesis'.[32] In it, LeSage sought to show that enumerative induction and the method of analogy – the two dominant methods advocated by the ubiquitous Newtonians – were not as foolproof as their partisans claimed and that the method of hypothesis[33] was not so weak as its critics insisted.

As LeSage shrewdly recognized, the core presupposition of the method of enumerative induction is that a clear distinction can be made between what is observable and what is not. The traditional contrast between inductive methods and speculative ones was that the former stayed very close to sensory experience, whereas the latter moved a long way from it, thereby acquiring a much greater degree of uncertainty. LeSage would admit that there is possibly a distinction to be made between what is observable and what is not, but he insisted that rigid and exclusive adherence to the former would produce a very emaciated science. 'Those', he said, 'who disparage the method of hypothesis do not allow us to make conjectures, *except those which follow naturally and immediately from experience*'.[34] Although this view is 'repeated superstitiously', there is no 'precise idea' that can be attached to it. For 'what on earth is an immediate consequence deduced from the observation of a fact? The existence of that fact and nothing more'.[35] But perhaps what was meant is that claims about the world that are closer to observation are better supported than those that are, as it were, several inferential steps away from sensory particulars. LeSage will have none of this. If, he says, the claims we make 'are hasty, what is gained if they are also immediate?' And if the claims we make are well evidenced, 'what does it matter whether [they are] immediate or as far removed from the phenomena as the last propositions of Euclid are from his axioms?'[36]

LeSage's point is that any form of theorising goes well beyond the available data; so there is no viable distinction to be made between theories that do and those that do not go beyond the evidence. Given that we must extrapolate from the known to the unknown, and that such a process is usually a risky business, what methodological rules can we use to ensure that such inferences are justified? The bulk of LeSage's essay is devoted to that task, and specifically to showing that the method of hypothesis is of greater moment in scientific reasoning than the rival methods of induction and analogy. As LeSage uses the term, 'the method of hypothesis' refers to any procedure that involves a comparison of the logical consequences of a theory with observations. The highest form of proof for a hypothesis would involve showing that all its consequences were true, that it exhibited 'exact correspondence with the phenomena'.[37] But, as he acknowledged, we are rarely if ever in a position to obtain such exhaustive evidence. Failing that, 'if the assumed cause [i.e., the hypothesis] is able to produce *all the presently known features* related to the principal effect, then it will have the highest degree of certitude that we can at present hope for in the circumstances'.[38] LeSage here added an important qualification. Before we accept a hypothesis that can explain all the available evidence, we must be sure that our evidence represents a large sample. Hypotheses that work well when the data are limited frequently break down as the data base is extended. LeSage insisted that our belief in a hypothesis should be a matter of degree: 'The greater the number of facts with which the [hypothesis] agrees, the more faith we should have in it'.[39] Pressing his analysis more closely, he stressed that mere agreement (i.e., logical consistency) between a hypothesis and the evidence is a very weak relation. If a hypothesis is to be solidly confirmed by a piece of evidence, then it must entail that evidential statement and none of its contraries. The greater the specificity with which the hypothesis correctly entails what we observe, the greater the extent to which it will be confirmed by those observations.[40]

The fact that hypotheses are not *generated* by an analysis of the evidence had been, for many of LeSage's contemporaries, a major strike against them. Inductivists, in particular, had stressed that the only legitimate theories are those that arise as generalisations from experience. LeSage replied that so long as we subject our hypotheses to a rigorous process of 'verification', it does not matter how they were generated initially. LeSage was as contemptuous as the inductivists of hypotheses that are not, or cannot be, verified. But he urged that we ought not to confuse the horrors of unverifiable hypotheses with the merits of verified ones. 'Thousands of times it has been said: the abuse of something, however universal, must never be taken as an argument against its legitimate usage'.[41] He maintained that Newton, in rejecting the method of hypothesis, 'never realized that one could utilize a method of research whose pitfalls he had recognized so well!'[42]

LeSage's next ploy consisted of a lengthy demonstration that, Newton's *hypotheses non fingo* notwithstanding, Newton's work in optics and mechanics was permeated by hypotheses and hypothetical reasoning.[43] Similarly, the works of Kepler, Copernicus, and Huygens rested on hypothetical modes of inference.[44]

He conceded that some hypotheses are spurious, singling out in particular the vortex theories of Descartes and his followers. 'The seventeenth century', he sagely observed, 'preferred to acknowledge every hypothesis, however implausible, while our century finds it more convenient to reject them all'.[45] Noting that epistemology, 'too, is subject to the rule of fashion and prejudice', LeSage maintained that there is a coherent middle way, which allows the use of hypotheses so long as strong empirical constraints are put on them. LeSage concluded his essay by pointing out the limited scope of analogical inference, insisting that it is 'inconclusive', 'arbitrary', 'impractical', and, in most cases, parasitic upon the method of hypothesis.[46]

Throughout this essay, as well as LeSage's other methodological writings, there are two levels of motivation. At one level, LeSage was genuinely concerned about epistemological issues in the abstract and, as a philosopher, felt it important to get as clear as he could about the logic of science. But lurking in the background (as in the case of Hartley's discussion) is LeSage the ether theorist, struggling desperately to get a fair hearing for a scientific theory that was being dismissed on what he regarded as flimsy methodological grounds. There is nothing untoward in all of this. We expect, after all, knowledge and the theory of knowledge to be closely intertwined. Nor are the excursions of LeSage and Hartley into epistemology merely self-serving apologiae. They are, of course, that; but they are substantially more than that as well. Chiefly, they are well-reasoned attempts to articulate and defend a hypothetico-deductive methodology in the face of inductivist criticisms.

There is another methodological theme common to the work of Hartley and LeSage that sets them apart from most of their contemporaries: a commitment to the progressive character of science. Throughout the earlier history of

epistemology, the prevailing view was that putative scientific theories or doctrines were to be judged as true or false *simpliciter* and that the authentic methodology of science should be one that would produce true theories more or less immediately; provided, of course, that appropriate rules of inquiry were obeyed. Hartley and LeSage both protested against this all-or-nothing view of scientific theories. Acknowledging that their own theories might be false (because they were at best only probable), they stressed the approximative character of scientific inquiry. They insisted that theories, once promulgated, could be amended and improved to bring them into ever closer agreement with their objects. To illustrate the point, both likened the development of scientific theories to certain mathematical methods of approximation (in fact, both singled out the rule of false position and the Newtonian technique for approximating roots as examples). As Hartley remarked:

> Here a first position is obtained, which, though not accurate, approaches, however, to the truth. From this, applied to the equation, a second position is deduced, which approaches nearer to the truth than the first . . . Now this is indeed the way, in which all advances in science are carried on.[47]

LeSage took the process of long division as a paradigm case of successful approximation.[48] Each step in the process brings us another digit closer to finding the true quotient. The testing and correction of hypotheses against experiment is, he maintained, a suitable parallel.

This ploy proved very useful for LeSage and Hartley. Confronted with the inductivists' insistence that the method of hypothesis is inconclusive, they could grant the point; confronted by the claim that their specific ethereal models might be false, they could concede the possibility. But as they saw it, neither the fallibility of the method of hypothesis nor the falsity of some theories produced by it need force one to the conclusion that it has no role to play in science. On the contrary, once one admits that science is approximative and self-corrective, it becomes possible to envisage a theory that is both false and an important step forward.

Ultimately, in fact, the epistemological views of these two thinkers turned out to be significantly more influential than the ether models they developed (as we shall see). But before we turn to look at the later fortunes of both the method of hypothesis and the ether hypothesis, we need to examine the opposition more closely.

The early critics of ether and its epistemology

Although ethereal explanations of electricity, magnetism, gravity, and heat became increasingly frequent and respectable in the course of the second half of the eighteenth century, there were still many natural philosophers who refused to countenance them. This was particularly true in Scotland, where, the Scottish 'enlightenment' notwithstanding, many of the major philosophers or scientists were unwilling to acknowledge the importance of such theories. If ethers were generally regarded by the Scots as unsavoury, Hartley's nervous ether was reserved for special abuse. In the early decades of the nineteenth century the Scottish philosopher Thomas Brown noted, with some pride, that 'it is chiefly in the southern part of the island that the hypothesis of Dr. Hartley has met with followers'.[49] What Brown says about Hartley's theories can be duplicated for almost all the other major ethereal doctrines of the late eighteenth century; they rarely made it past Hadrian's Wall. Many of Scotland's major natural philosophers (including Robison, Leslie, Reid, and Hutton) rejected the imperceptible fluid hypotheses that were, by the 1770s, being widely discussed (and, in some cases, widely accepted) in England and throughout much of Continental Europe.

The primary reason for opposition to ether theories was the widespread acceptance among Scottish philosophers and scientists of a trenchant inductivism and empiricism, according to which speculative hypotheses and imperceptible entities were inconsistent with the search for reliable science. This linking of opposition to ether theories with an inductivist philosophy is neatly summed up in the 'Aether' article for the first edition of the *Encyclopaedia Britannica* (1771), which, in spite of its title, was a predominantly Scottish production. Barely was *ether* defined ('the name of an imaginary fluid, supposed by several authors . . . to be the cause . . . of every phenomenon in nature')[50] before the author of the article launched into a virulent attack on the method of hypothesis and a spirited defence of the inductivism of Newton and Bacon: 'Before the method of philosophising by induction was known, the hypotheses of philosophers were wild, fanciful, ridiculous. They had recourse to aether, occult qualities, and other imaginary causes'.[51] The article insisted that 'the way of conjecture . . . will never lead any man to truth'.[52] These passages from the *Britannica* reflect a thoroughgoing inductivism that influenced much Scottish writing on science of the period. Most of it sprang from the very influential work of Thomas Reid, leader of the so-called commonsense philosophers. As I have shown elsewhere, Reid's works are replete with abusive attacks on the method of hypothesis.[53]

What is important for our purposes is the explicit linkage between Reid's repudiation of Hartley's ethereal speculations and his attack upon the method that undergirds them. In his *Essays on the intellectual powers of man* (1785), Reid discussed Hartley's vibratory hypothesis at length. In the course of rejecting Hartley's 'hypotheses concerning the nerves and brain',[54] Reid embarked on the lengthiest methodological discussion in his opus. Reid quickly perceived that Hartley's *Observations on man* ran directly counter to the Newtonian inductivism that Reid himself espoused. He noted, significantly, that the epistemological chapter of the *Observations on man* was written to justify the methodology utilised in Hartley's research: 'Having first deviated from [Newton's] method in his practice, [Hartley] is brought at last to justify this deviation in theory'.[55] Reid claimed that Hartley was the only author he knew who rejected the principles of Newton's inductivism, and 'Dr. Hartley is the only author I have met with who reasons against them, and has taken pains to find out arguments in defense of the exploded method of hypothesis'.[56] It seems natural to infer, then, that most of Reid's methodological tirades against hypotheses were directed chiefly at Hartley, since he alone, in Reid's view, had criticised 'the true method of philosophizing'.

That 'true method', so far as Reid was concerned, was some form of *enumerative* induction.[57] Reid nowhere spelled out the rules of his form of induction (as even his followers had to concede),[58] but some of its features are clear. As he wrote in the *Intellectual powers of man:* 'The true method of philosophizing is this: From real facts, ascertained by observation and experiments, to collect by just induction the laws of Nature, and to apply the laws so discovered, to account for the phenomena of Nature'.[59] The nub of the issue concerns the kinds of inductive generalisations that we can perform on 'observations and experiments'. Reid was adamant on two issues at this stage: (1) that any entities postulated in a putative law should really exist 'and not be barely conjectured to exist without proof'; and (2) *all* the known deductive consequences of the law must be true.[60]

Condition (1), which Reid intended as a gloss on Newton's idea of *verae causae,* was extensively discussed, both in his published work and in his correspondence.[61] More often than not, Reid's first condition amounted to the rule that the scientist is allowed to postulate *only those entities that are observable.* Ethers and other imperceptible fluids are thus, by their very nature, disqualified from legitimate scientific status.

It is sorely tempting to side epistemologically with Hartley and LeSage against Reid. After all, the method of enumerative induction is not rich enough to build science; speculations and conjectures are inevitable. It does count in favour of a theory that it can explain a wide variety of phenomena, even if that theory postulates imperceptible entities. But to see the debate solely in these terms is misleading. In large part, the epistemological issue at stake between the inductivists and the hypotheticalists in the eighteenth century is simply this: Does a confirming instance of a theory automatically count as a ground for accepting or believing the theory? Both Hartley and LeSage insisted that any and every confirming instance provides evidence for the theory that entails it; accumulate enough such instances and the theory becomes credible. Thomas Reid, like such later philosophers of science as Whewell, Peirce, the Bayesians, and Popper, maintained that 'mere' confirming instances are not enough, that some additional demand must be met before we can legitimately say that true deductive consequences of a theory count towards warranted belief in that theory.

What motivates this conviction in Reid's case is a sound intuition that many unsavoury theories have some true consequences. If we were to regard every theory with some true consequences as well established, then we could never judge one theory to have stronger evidence than another, since 'there never was an hypothesis invented by an ingenious man which has not this [kind of] evidence in its favor. The vortices of DesCartes, the sylphs and gnomes of Mr. Pope, serve to account for a great variety of phenomena'.[62] Reid believed that a theory should not be regarded as confirmed or well established merely because it was sufficient to save the appearances. Of course, he did not quarrel with the view that legitimate theories must be compatible with the available evidence. But what Hartley and LeSage were willing to regard as a *sufficient* condition for an acceptable theory Reid viewed as a *necessary* but not sufficient condition.[63]

But in a classic case of babies and bath water, Reid's requirement went too far. His very narrow observational construal of the grounds for warrantedly asserting the existence of a thing left him completely unable to give an account of the success of the many deep-structural theories of his time. Because it demanded too much, his epistemology was altogether unable to come to grips with the contemporary theoretical sciences. But if Reid and the inductivists demanded too much, Hartley and LeSage ran the risk of requiring too little. After all, many vacuous explanations – witness the classic *virtus dormitivus* in Molière – have true deductive consequences. Neither Hartley nor LeSage provided any plausible criterion for distinguishing vacuously true theories from legitimate scientific ones. In the absence of such a distinction, they had no grounds for claiming that their theories should be accepted in lieu of a multitude of equally well-confirmed but vacuous ones.

In sum, by the late eighteenth century, neither the inductivists nor the hypothetico-deductivists had yet constructed a plausible epistemology of science; equally, the fortunes of ether theories were still unsettled, precisely because it was unclear whether there was genuine evidence for them. All this was to change profoundly in the next half century. Philosophers of science would articulate new and more detailed criteria for determining which deductive consequences of a theory were genuinely confirmatory, and ether theories, at least certain ether theories, would pass these criteria with flying colours.

The second phase, 1820–1840

The early nineteenth-century debate about imperceptible fluids and the methods for establishing their existence focused chiefly upon the wave theory of light (with its seemingly attendant commitment to a luminiferous ether). On the epistemological side, most of the interest centered around the emergence of a new methodological criterion for evaluating hypotheses. In brief, this criterion, which was nowhere prominent in the late eighteenth-century debates about the methodological credentials of subtle fluids, amounted to the claim that a hypothesis that successfully predicts future states of affairs (particularly if those states are 'surprising' ones), or that explains phenomena it was not designed to explain, acquires thereby a legitimacy not possessed by hypotheses that merely explain what is already known generally. The major figures in this part of the story are Herschel, Whewell, and Mill. To put the matter concisely, John Herschel accepted the new criterion, saw that it provided a rationale for the wave theory of light, but did not recognise how that criterion threatened the traditional inductivist programme to which he was committed. Whewell accepted the new criterion and used it both to defend the wave theory of light and to attack traditional inductive procedures. Finally, John Stuart Mill, perceiving that the new criterion undermined induction, repudiated the former and, along with it, the vibratory theory of light. It is this set of interconnections with which I shall be concerned in this section.

The wave theory of light was revived chiefly, of course, by Thomas Young and Augustin Fresnel in the early years of the nineteenth century. Through the first three decades of that century, opinion on the merits of the wave or the corpuscular theory of light was very divided among natural philosophers.[64] If the wave theory could provide a more convincing account of interference than the corpuscular theory, the latter seemed better able to explain problems of stellar abberation and double refraction. (Some participants in the ether debates, such as Comte, actually believed that the two theories were observationally equivalent and that it was a matter of complete indifference which theory one utilised.[65])

By the late 1820s, however, the balance of opinion was shifting perceptibly towards the undulatory theory. There were many factors responsible, but most important among them was the Fresnel–Poisson experiment. The idea for the experiment arose in response to an essay written in 1816 by Fresnel on the nature of diffraction. In this paper, Fresnel elaborated the wave theory of light and applied it to the explanation of diffraction phenomena. Poisson, a confirmed corpuscularian and member of a panel refereeing Fresnel's paper, observed that, according to the analysis of light that Fresnel was using, it would follow that the centre of the shadow of a circular disc would exhibit a bright spot. This predicted result was highly unlikely; it contradicted both the corpuscular theory and the scientists' intuitive sense of what was 'natural'. Indeed, the fact that the wave theory possessed this bizarre consequence was seen, prior to performing the experiment, as a kind of reductio ad absurdum of it. But when the appropriate tests were performed, the wave theory was vindicated by an exact concordance between what it predicted and the observed results.

There are two obvious methodological construals of the outcome of this experiment. On one interpretation, the experiment functions as a Baconian *experimentum crucis,* proving the falsity of the corpuscular theory and the truth of the wave theory. (Interestingly, although precisely this construal was given to the later Foucault experiments on the speed of light in water and air, this was not the dominant interpretation of this earlier result.) On another interpretation, widely adopted at the time, the disc experiment can be viewed as providing convincing evidence for the wave theory by virtue of its successful prediction of a surprising (i.e., unexpected on the background knowledge) observational effect. The logic that undergirds the former interpretation is the familiar logic of eliminative induction. But what provides the epistemic rationale for the latter interpretation is a set of developments within the methodology of hypothesis evaluation.

As we have seen, throughout the eighteenth century, proponents of the method of hypothesis pointed to post hoc (and, according to the critics of that method, presumably ad hoc) explanations of known phenomena as the chief vehicle whereby a hypothesis proves its mettle. But none of the early proponents of the method of hypothesis could show what distinguished arbitrary and vacuous hypotheses from genuine and worthy ones, since both classes possessed large sets of post hoc confirming instances.

During the 1820s and 1830s, proponents of the method of hypothesis articulated machinery for distinguishing 'artificial' hypotheses from legitimate ones. Specifically, they insisted that a proper hypothesis is one that not only explains what is already known, but also can be extended beyond the initial range of phenomena it was designed to explain. Particularly if the hypothesis can predict results that are unusual or surprising, the hypothesis loses its artificiality and becomes a legitimate contender for rational belief. What is involved in this modification of the method of hypothesis (a modification especially prominent in the work of Herschel and Whewell) is *nothing less than a redefinition of what constitutes evidence*. In stressing that a hypothesis must

establish its credentials by going beyond its initial data base, Herschel and Whewell defined a promising *via media* between the earlier extremes of Hartley and LeSage, on the one hand, and Reid and the other inductivists, on the other. By claiming that nothing can be arbitrary about a hypothesis that is successfully tested against a body of evidence independent of the circumstances that the hypothesis was invented to explain in the first place, they thus defused the charge of arbitrariness that was traditionally directed against the method of hypothesis.

Given the familiarity of this idea to a modern reader, it is easy to underestimate how significant a shift in the history of methodology it represents.[66] Part of its importance lies in the stress it puts on testing claims against the unknown rather than the known. Before the method of hypothesis could be plausibly viewed as anything more than the logical fallacy which it had often been considered, something new was called for, something that would separate serious and legitimate hypotheses from bogus or specious ones. What suited the bill, at least so far as the 1830s were concerned, was what I shall call *the requirement of independent or collateral support*. In brief, this requirement amounted to the demand that before a hypothesis was credible, it had to explain (or predict) states of affairs significantly different from those that it was initially invented to explain. Evidence of this kind might come from one of two sources: either a surprising prediction of unknown effects or a successful explanation of phenomena that were already known but that did not serve as the original base for the formulation of the hypothesis.

This methodological requirement of independent support ought not to be confused with the earlier empiricist requirement that theories must involve *verae causae*. That earlier demand had nothing to do with the capacity of a theory to predict surprising results; it insisted, rather, that the entities postulated in a theory had to be directly observable or directly 'inferred from the phenomena'. In short, the methodological tradition of *verae causae* had rested upon a rigid distinction between directly observable entities and not-directly observable ones, endorsing the former and eschewing the latter. By contrast, what I am calling the requirement of independent support is indifferent to the question whether theoretical entities are observable. It focuses, rather, on the epistemic features of the sentences that can be deduced from a theory.

A theory like LeSage's gravorific ether, even if it had been entirely successful in its explanatory ambitions, did not seem to possess independent support in the sense defined. More to the point, no epistemologist in the eighteenth century would have been impressed if it had, for the notion of independent support in the sense under discussion here was very much a product of the early nineteenth century. By invoking this requirement, as we shall see, proponents of the optical ether were able to argue – as their eighteenth-century ethereal precursors were not – that there was some very impressive evidence available for a luminiferous ether, evidence that went well beyond the ability of that hypothesis merely 'to save the phenomena'. The Fresnel–Poisson circular disc experiment (a confirmed prediction), as well as the successful extension of the wave theory to polarisation, dipolarisation, and double refraction, betokened a degree of collateral support for the optical ether that earlier ethereal doctrines did not exhibit.

This set of issues was stressed with particular emphasis by both John Herschel and William Whewell. Herschel insisted that we cannot reasonably expect a theory to be a reliable predictor in the future unless it has also been so in the past. Unless we have seen that a theory enables 'us to extend our views beyond the circle of instances from which it is obtained', then 'we cannot rely on it'.[67] Before we accept any theory we must try 'extending its application to cases not originally contemplated . . . and pushing the application of [it] to extreme cases'.[68] Although Herschel frequently enunciated this requirement, and saw it as a vehicle for protecting us from 'the unrestrained exercise of imagination . . . arbitrary principles . . . [and] mere fanciful causes',[69] he did not discuss its rationale at any length.

What he did do, however, was to invoke this requirement repeatedly to show that the hypothesis of the optical ether was a sound one. Herschel pointed out that Young's wave theory, originally developed to explain reflection, refraction, and interference, was eventually applied with much success to the explanation of polarisation and double refraction.[70] He pointed out not only that (Fresnel's version of) the wave theory explained 'perhaps the greatest variety of facts that have ever yet been arranged under one general head'[71] but that Fresnel's theory also predicted 'a *fact* which had never been observed . . . and all opinion was against it'.[72] The confirmation of this surprising prediction did much, in Herschel's view, to establish the 'probability' of the undulatory hypothesis.

Like Herschel, Whewell found it necessary to supplement the weak demands of the traditional method of hypothesis by further constraints. He elaborated these in a lengthy section of the *Philosophy of the inductive sciences* (1840) devoted to 'Tests of hypotheses'. With LeSage and Hartley, Whewell insisted on the minimal condition: 'The hypotheses which we accept ought to explain phenomena we have observed'.[73] More precisely, he stipulated that every hypothesis had to be 'consistent with *all* the observed facts'.[74] But unlike Hartley and LeSage, who saw in such a requirement a sufficient condition for adequacy, Whewell argued that hypotheses 'ought to do more than this: our hypotheses ought [successfully] to *fortel* [*sic*] phenomena which

have not yet been observed'.[75] At a minimum, these predictions should be borne out by tests of the hypothesis against phenomena 'of the same kind as those which the hypothesis was invented to explain'.[76]

The rationale for this rather more exacting requirement was precisely that it dissipates the air of arbitrariness surrounding a hypothesis whose only known instances are those used in its generation: 'Men cannot help believing that the laws laid down by discoverers must be in a great measure identical with the real laws of nature, when the discoverers thus determine effects beforehand in the same manner in which nature herself determines them when the occasion occurs'.[77]

Successful prediction of effects similar to those already known does much to increase our confidence in a hypothesis. 'But the evidence in favour of our induction is of a much higher and more forcible character when it enables us to explain and determine [i.e., predict] cases of a kind different from those which were contemplated in the formation of our hypothesis'.[78] Whewell's technical term for this particular mode of evidencing (which involves testing a hypothesis against types of processes different *in kind* from those it was devised to explain) is the *consilience of inductions*. In his view, this is the most impressive type of evidence that theories can possess. Whewell's stress on the special confirmatory value of successful predictions, and the contrast it marks with the earlier eighteenth-century discussions of the method of hypothesis, comes out very clearly in his unusual utilisation of the decypherhing analogy. We saw this analogy already in Hartley (and it was used before him by Descartes and Boyle, inter alia). In its pre-Whewellian form, the analogy had suggested that if a certain hypothetical assignment of letters to an encoded cypher produces an intelligible message, then this constitutes evidence that the hypothetical assignment is correct. In this version of the analogy, it is assumed that the entire cypher is known in advance to the decoder. In Whewell's version of the analogy, however, a portion of the cypher is initially concealed from the decoder and the test of his decoding is whether he can predict the character of the concealed cypher.[79] As Whewell's variation on this traditional analogy makes clear, his concern is to assign differential weights to the confirming instances of a theory.

Barely had Whewell defined this notion of consilience before he cited the wave theory of light *in extenso* as one of the few theories to have passed this demanding test. In quick succession, he ticked off the explanatory successes of the wave theory: reflection, refraction, colours of thin plates, polarisation, double refraction, dipolarisation, and circular polarisation. By contrast, the emission theory exhibited 'what we may consider the natural course of things in the career of a false theory'.[80] It could well enough explain 'the phenomena which it was at first contrived to meet; but every new class of facts requires a new supposition . . . as observation goes on, these incoherent appendages accumulate, till they overwhelm and upset the original framework'.[81]

So impressed was Whewell by the predictive successes of the wave theory that he used it as one of his two paradigm cases of exemplary theory development (Newtonian gravitational theory was the other), and his elaborate 'Inductive table of optics' culminated in the wave theory. Even in his earlier *History of the inductive sciences* (1837), Whewell saw the wave theory as the optical equivalent of Newtonian mechanics. After a lengthy discussion of what he then called 'the undulatory theory', he remarked:

> We have been desirous of showing that the *type* of this progress in the histories of the two great sciences, Physical Astronomy and Physical Optics, is the same. In both we have many *Laws of Phenomena* detected and accumulated by astute and inventive men; we have *Preludial* guesses which touch the true theory . . . finally, we have the *Epoch* when this true theory . . . is recommended by its fully explaining what it was first meant to explain, and confirmed by its explaining what it was *not* meant to explain.[82]

This passage neatly foreshadows Whewell's philosophy of science (with its emphasis on the independent support requirement) and the key role that the wave theory played in its formulation.

The Mill–Whewell debate

But what of the opposition? As we have seen, both Herschel and Whewell based their endorsement of the wave theory on its satisfaction of an innovative and highly controversial methodological demand (i.e., the requirement of independent support). If my account of the connection between views towards optics and this methodological rule is correct, we should expect that those who did not accept the Herschel–Whewell methodology would not share their enthusiasm for the undulatory theory. Confirmation for this expectation is ready at hand in the works of John Stuart Mill. Mill was, of course, the arch-foe of Whewell during the 1840s and 1850s. Their respective philosophies of science exhibited divergences at almost every major point. Not the least of these differences was the disagreement of the two men about the relative confirmational value of different instances of a theory. As we have seen, Whewell maintained that a theory was better confirmed by predictive instances or by explaining phenomena it was not originally devised to explain than it was by explaining phenomena it had been devised to explain. By contrast, Mill insisted that predictive successes, every bit as much as post hoc explanatory ones, were highly inconclusive; both, in Mill's view, if taken as grounds for asserting the theories that achieved them, were highly fallacious.

Significantly, this controversy emerges in Mill's *System of logic* (1843) with specific reference to the luminiferous ether. Mill's central point was that no number of confirming instances of a theory can establish it conclusively, chiefly 'because we cannot have, in the case of such an hypothesis (viz., the optical ether), the assurance that if the hypothesis be false it must lead to results at variance with true facts'.[83] The fact that the wave theory of light 'accounts for all the *known* phenomena' does not warrant the view that it is 'probably true'.[84] He then turned to consider Whewell's (and Herschel's) claim that it is chiefly the predictive successes of the wave theory that render it likely. 'It seems to be thought', Mill observed, 'that an hypothesis of the sort in question is entitled to a more favourable reception, if, besides accounting for all the facts previously known, it has led to the anticipation and prediction of others which experience afterwards verified'.[85]

Mill was scathing in his insistence that such a view flagrantly confounds the psychology of surprise with the methodology of support. Referring specifically to one of the predictions of the wave theory, he said:

> Such predictions and their fulfillment are, indeed, well calculated to impress the uninformed,[86] whose faith in science rests solely on similar coincidences between its prophecies and what comes to pass. But it is strange that any considerable stress should be laid upon such a coincidence by persons of scientific attainments. If the laws of the propagation of light accord with those of the vibrations of an elastic fluid in as many respects as is necessary to make the hypothesis afford a correct expression of all or most of the phenomena known at the time, it is nothing strange that they should accord with each other in one respect more.[87]

Mill was not taking exception to Whewell's psychological observation that many people are impressed by a theory that successfully makes surprising predictions. What he was calling for was a logical or epistemological account of why we should regard such instances as being of a privileged, probative character. Like Popper a century later, Whewell did not ultimately meet this challenge; he never showed why a theory's novel predictions should count for so much more, in terms of its epistemic appraisal, than its successes at explaining what it was devised to explain.

Nonetheless, Whewell tried to restate his case in his *Of induction, with especial reference to Mr. J. S. Mill's System of Logic* (1849). Whewell rightly summarised Mill's attack on him by saying that it amounted to the traditional inductivist charge that one ought not allow 'hypotheses to be established, merely in virtue of the accordance of their results with the phenomena'.[88]

Whewell reiterated that his was not merely the old-fashioned method of hypothesis, for he added the demands of independent support and consilience. As for Mill's charge that successful predictions should impress only the 'uninformed', Whewell insisted that 'most scientific thinkers . . . have allowed the coincidence of results predicted by theory with fact afterwards observed, to produce the strongest effects upon their conviction'.[89] (In the same passage he referred to 'the curiously felicitous proofs of the undulatory theory of light'.) But most of Whewell's reply consisted chiefly of pious hand waving rather than cogent arguments; he never satisfactorily met Mill's challenge to produce a plausible epistemological rationale for the requirement of independent support.

At another level, however, Mill had perhaps missed the point. It is one thing to stress, as he rightly did contra Whewell, that one or two surprising, but confirmed, predictions do not prove the theory that produces them. But if I am right, the motivation for introducing the predictive requirement was not to transform the method of hypothesis into a proof technique. Rather, the concern had been to find some way of reducing the arbitrariness and the ad hoc nature of hypotheses. So long as the only confirming instances that a hypothesis could claim were those used in its generation, there was no reason to expect that applications of it to further instances would be successful. After all, the hypothesis might simply have fastened on some noncausal or non-nomic accidents of the cases so far surveyed, and generalised these into a universal theory. But, as Herschel and Whewell observed, if the hypothesis can be successfully extended to cases or even to domains that were not used in its development, then it can no longer be claimed that the reliability of the hypothesis is limited to the phenomena that it was devised to explain.[90]

Conclusion and postscript

By the 1850s, the wave theory of light (and the associated hypothesis of a luminiferous ether) enjoyed a degree of acceptance among natural philosophers that none of the subtle fluid ethers of the eighteenth century had possessed. I have tried to show in this chapter that the different receptions afforded to the two sets of theories are to be explained primarily by the wave theory's possession of certain epistemic or methodological features not present in the earlier ethers. But the moral of the tale is not merely to be found in this difference; for even if earlier ethers had been predictive (in the sense indicated here), it is not clear that they would have been any more widely accepted than they were. What was needed was a shift not only in physics but in epistemology as well, so that independent support could be recognised as a decided epistemic virtue. That shift came (I have argued) in the 1830s,

provoked in part by the wave theory itself, which served as an epistemological archetype for such philosophers as Herschel and Whewell. There is, of course, a circularity here. But, far from being of the vicious variety, it reflects the kind of mutual dependence between theory and praxis that has always characterised science and philosophy at their best.[91]

There remain many philosophers of science and theorists of scientific change who, though granting that substantive theories about the world do change, nonetheless adhere to the view that the canons of legitimate scientific inference are perennial and unchanging. (Included here are thinkers as diverse as Popper, Nagel, Carnap, Hesse, and Lakatos, among others.) The case we have before us stands as a vivid refutation of their claim that scientific standards of theory evaluation are immutable. It simply cannot be denied that, prior to the early nineteenth century, the ability of a theory to make successful, surprising predictions was no sine qua non for its acceptability; nor can it be denied that by the turn of the twentieth century, the requirement of predictivity was a commonplace in both scientific and philosophical circles.[92] Such a profound shift is irreconcilable with any philosophical dogma to the effect that scientific methodology has a fixed character.

The case before us poses an equally acute challenge to that group of relativists – associated with the work of Thomas Kuhn – who believe that new methods and new standards are paradigm-specific (and who argue that standards are retained or are rejected on the strength of the specific paradigms with which they are initially associated). Although the predictivity requirement may well have been ushered in by the wave theory of light (as I have argued here), it soon acquired a life of its own that dissociated its fortunes from the fate of the particular physical paradigm that had brought it to the fore.

Acknowledgments

The author is grateful to the National Science Foundation, which supported part of this research, and to the Librarian of the University of Geneva Library.

Notes

1 For a detailed discussion of the rationale for such interactions, see L. Laudan, *Progress and its problems* (London, 1977), chap. 2.
2 See, for instance, Laudan (1970); Cohen (1956).
3 That the prevailing eighteenth-century interpretation of Newton's first rule involves the demand that all theories restrict themselves to purely observable entities is slightly a matter of conjecture. What can be said with some confidence is that British and French glosses on Newton's *regulae* generally construe it in this way. (For a

typical and very influential discussion of Newton's first rule, see W. Hamilton (ed.) *The philosophical works of Thomas Reid,* 6th ed., 2 vols. Edinburgh, 1863), 1:57, 236, 261, 271–2.

4 J. Priestley, *History and present state of electricity,* 3rd ed., 2 vols. (London, 1775), 2:16.

5 See query 16 to Newton's *Opticks,* 2nd ed. (London, 1717).

6 D. Hartley, *Observations on man, his frame, his duty, and his expectations,* 2nd ed., 2 vols. (London, 1791), 1:15. A very similar methodological argument had been made two years earlier by Bryan Robinson: 'This *Aether* being a very general material Cause, without any Objection appearing against it from the Phaenomena, no Doubt can be made of its Existence: For by how much the more general any cause is, by so much the stronger is the Reason for allowing its Existence. The Aether is a much more general Cause then our Air: And on that Account, the Evidence from the Phaenomena, is much stronger in Favour of the Existence of the *Aether,* than it is in Favour of the Existence of the Air' [*A dissertation on the aether of Sir Isaac Newton* (London, 1747), preface, n.p.]

7 Hartley, *Observations,* 1:16. This decyphering analogy has a long prehistory among earlier proponents of the method of hypothesis. It can be found, for instance, in Descartes (*Oeuvres,* eds. C. Adam and P. Tannery, 12 vols. [Paris, 1897–1910], 10:323) and Boyle (Royal Society, Boyle papers, vol. 9, f. 63), among other seventeenth-century writers. As I shall show, it continued to be used by methodologists for well over a century after Hartley.

8 Hartley, *Observations,* 1:346.

9 Ibid.

10 Ibid.

11 Ibid., 347.

12 Many inductivists distinguished between what have subsequently been called the contexts of discovery and of justification. They were quite prepared to grant that hypothetical methods were useful for the former, but insisted that they had no role in the latter. As Thomas Reid succinctly put it, 'Let hypotheses . . . suggest experiments, or direct our inquiries; but let just induction alone govern our belief'. *Works,* 1:251.

13 Hartley, *Observations,* 1:364.

14 Ibid., 350.

15 Ibid., 341.

16 Indeed, Hartley even seems to endorse such a course of action. If, he says, our suppositions and hypotheses 'do not answer in some tolerable measure [to the real phenomena, we ought] to reject them at once; or, if they do, to add, expurge, correct, and improve, till we have brought the hypothesis as near as we can to an agreement with nature'. *Observations,* 1:345.

17 Indeed, Newton's first *regula philosophandi* insisted that theories must be 'sufficient to explain the appearances'.

18 [G. LeSage], 'Lettre à un academicien', *Mercure de France* (May 1756).

19 Published in Paris in 1758.

20 G. L. LeSage, 'Loi, qui comprend, malgré sa simplicité, toutes les attractions . . .', *Le Journal des Sçavans* (April 1764), 230–4. He wrote later papers on this same topic in the *Journal des Beaux-arts* (Nov. 1772) and (Feb. 1773) and in the *Journal de Physique* (Nov. 1773).

21 G. LeSage, 'Lucrèce Newtonien', in *Mémoires de l'Académie Royale des Sciences et Belles-lettres de Berlin* (Berlin, 1784), 1–28; reprinted in P. Prevost, *Notice de la vie et des écrits de George-Louis LeSage* (Geneva, 1805), 561–604.

22 Readers seeking details about LeSage's gravitational ether should consult either LeSage, 'Lucrèce', or Aronson (1964).

23 Cf. P. Prevost, *Deux traités de physique mécanique* (Geneva, 1818).

24 W. Niven (ed.), *Scientific papers of James Clerk Maxwell* (London, 1890), 2:474.

25 Prevost, *Notice,* 358.

26 Ibid., 300.

27 Ever an Enlightenment liberal, Euler added, 'Mais j'accorde très volontiers cette liberté à d'autres'. Euler to LeSage, 8 Sept. 1765, University of Geneva Library: MS Suppl. 512, f.314^r. (LeSage's transcription of Euler's letter is MS fr. 2063, f.141^r.)

28 Prevost, *Notice,* 464–5.

29 Ibid., 237.

30 He spoke of 'la prétendue impossibilité d'établir solidement un système, qui roule sur les objects essentiellement imperceptibles'. Ibid., 264).

31 Ibid., 242.

32 The full text of the essay was published in P. Prevost, *Essais de philosophie* (Geneva, 1804), 2:258 ff. The original can be found in the University of Geneva Library, Ms fr. 2019(2). Because there are some (largely minor) discrepancies between the printed version and the original, all my quotations shall be from the latter. I shall give in parentheses references to the appropriate section in Prevost's text.

33 By the term *method of hypothesis,* LeSage meant the view that science 'is conducted by the method of trial and error . . . by gropings followed by verification, by hypotheses which are then confirmed by their agreement with the phenomena'. University of Geneva Library, MS fr. 2019(2) (§5).

34 Ibid. (§26).

35 Ibid.

36 Ibid.

37 Ibid. (§7). Joseph Priestley would similarly insist that if a scientist can frame his theory so as to suit all the facts, 'then it has all the evidence of truth that the nature of things can admit'. Priestley, *History,* 16. For allusions to the very early history of this principle, see L. Laudan, 'Ex-Huming Hacking', *Erkenntnis 13* (1978), 417–35.

38 University of Geneva Library, MS fr. 2019(2) (§7) (my italics).

39 Ibid.(§15).

40 Cf. Ibid. LeSage even seemed to think, as did Hartley, that with a sufficiently large number of confirming instances, the hypothesis becomes virtually certain. As he wrote in his *Principes généraux de la téléologie:* 'Plus les phénomènes sont nombreux et plus la précision est grande; plus aussi ils jugent avec assurence qu'il ne sauroit y avoir d'autre hypothèse sur le même sujet, qui ait les mêmes avantages'. Prevost, *Notice,* 529–30.

41 University of Geneva Library, MS fr. 2019(2) (§18).

42 Ibid. (§19).

43 He claimed that 'almost all that the first two books of his *Principia* [i.e., Newton's] contain . . . is nothing more than a collection' of 'curious hypotheses'. Ibid. (§20).

44 Ibid. (§§23–5).

45 Ibid. (§29). Lest the wary reader may think LeSage was painting an exaggerated picture of the Newtonians' aversion to hypotheses, it is worth saying that his view is borne out by many of his contemporaries. Thus the *Encyclopédie* (in the article 'Hypothèse' observed that Newton 'et sur-tout ses disciples' were very opposed to hypotheses, regarding them as 'le poison de la raison et la peste de la philosophie'. Across the channel, the Newtonian Benjamin Martin noted in the 1750s: 'The Philosophers of the present Age hold [hypotheses] in vile Esteem, and will hardly admit the name in their Writings; they think that which depends upon bare Hypothesis and

Conjecture, unworthy the name of Philosophy'. *Philosophical Grammar*, 7th ed. (London, 1769), 19.

46 University of Geneva Library, MS fr. 2019(2) (§§30–8).

47 Hartley, *Observations*, 1:349.

48 University of Geneva Library, MS fr. 2019(2) (§5).

49 T. Brown, *Lectures on the philosophy of the human mind*, 20th ed. (London, 1860), 279.

50 *Encyclopaedia Britannica*, 3 vols. (Edinburgh, 1771), 1:31.

51 Ibid.

52 Ibid., 32.

53 Laudan (1970).

54 Reid, *Works*, 1:248–53.

55 Ibid., 250.

56 Ibid., 251.

57 Like Newton, in whose footsteps he sought to follow, Reid rejected *eliminative* induction, chiefly on the grounds that we cannot perform the exhaustive enumeration of possible hypotheses which that method requires. See, for instance, ibid., 250.

58 As Reid's successor Dugald Stewart observed: 'It were perhaps to be wished that [Reid] had taken a little more pains to illustrate the fundamental rules of that [inductive] logic the value of which he estimated so highly'. Ibid., 11.

59 Ibid., 271.

60 Ibid., 250.

61 See especially the exchange of letters between Reid and Lord Kames in ibid.

62 Ibid., 1:251.

63 That it is a necessary condition is made clear by Reid's condition (2), cited in the text.

64 I shall not discuss the first methodological debate that the wave theory provoked, namely, that between Young and Brougham. I skip over it for two reasons: (1) It has already been investigated at length by Cantor (1971); (2) it represents a more vituperative but less substantial replay of the earlier ether debates I have discussed, with Brougham playing Reid to Young's Hartley.

65 See especially A. Comte, *Cours de philosophie positive*, 2 vols. (Paris, 1924), 2:331–52. A general discussion of Comte's philosophy of science can be found in L. Laudan, 'Towards a reassessment of Comte's "méthode positive" ', *Philosophy of Science 38* (1971), 35–53.

66 Indeed, given the ubiquity of the requirement – or analogues of it – in contemporary philosophy of science, it is remarkable that its prehistory has not yet been explored. In an intriguing but false surmise, Karl Popper once observed that 'successful new prediction – of new effects – seems to be a late idea, for obvious [*sic*] reasons; perhaps it was first mentioned by some pragmatist'. *Conjectures and refutations* (London, 1965), 247.

67 For a useful discussion of related issues, see V. Kavaloski, 'The "vera causa" principle', unpublished doctoral dissertation, University of Chicago, 1974.

68 J. F. W. Herschel, *A preliminary discourse on the study of natural philosophy* (London, 1830), 167. See also ibid., 172, 203.

69 Ibid., 190.

70 Ibid., 259 ff.

71 Ibid., 32.

72 Ibid., 32–3.

73 W. Whewell, *The philosophy of the inductive sciences, founded upon their history*, 2 vols. (London, 1847), 2:62.

74 Ibid.

75 Ibid.

76 Ibid., 62–3.

77 Ibid., 64.

78 Ibid., 65. For a fuller discussion of these issues, see L. Laudan, 'William Whewell on the consilience of inductions', *Monist 55* (1971), 368–91.

79 'If I copy a long series of letters, of which the last half dozen are concealed, and if I guess these aright, as is found to be the case when they are afterwards uncovered, this must be because I have made out the import of the inscription'. *Philosophy of discovery* (London, 1860), 274.

80 Whewell, *Philosophy of the inductive sciences*, 2:72.

81 Ibid.

82 W. Whewell, *History of the inductive sciences from the earliest to the present time*, 3rd ed., 3 vols. (London, 1857), 2:370.

83 J. S. Mill, *System of logic, ratiocinative and inductive*, 8th ed. (London, 1961), 328.

84 Ibid.

85 Ibid.

86 In early editions of the *System*, he referred here to the 'ignorant vulgar' rather than the 'uninformed'.

87 Mill, *System*, 328–9.

88 Whewell, *Discovery*, 270.

89 Ibid., 273.

90 The Herschel – Whewell requirement of independent support has shown up in a new guise in the work of E. Zahar, especially his 'Why did Einstein's programme supersede Lorentz's?' in *Method and appraisal in the physical sciences*, ed. C. Howson (Cambridge, 1976), 211–76. Zahar is no more successful than Herschel and Whewell were in providing a philosophical justification for the requirement. (For a latter-day 'Millean' critique of recent work in this area, see Laudan, *Progress*, 114–18.)

91 Cantor (1975). There are many other dimensions of the ether debates that deserve serious exploration. For instance, some natural philosophers regarded the mathematical formalisms of the wave theory as well established but refused to regard this as evidence for the existence of a luminiferous ether. What was at stake was whether a theory could be 'accepted' without its ontological presuppositions being taken seriously.

I have argued for the symbiotic character of the general relationship between science and the philosophy in Laudan, *Progress*, chap. 2, and in Laudan, 'The sources of modern methodology', in *Historical and philosophical dimensions of logic, methodology and philosophy of science*, eds. R. E. Butts and J. Hintikka (Dordrecht, 1977), 3–20.

92 A minor caveat is in order here. Several philosophers and scientists before the nineteenth century (e.g., Boyle, Huygens, Liebniz) had claimed that the ability of a theory to make surprising predictions was an epistemic advantage. But prior to the 1820s no *systematic* arguments had been made to the effect that such an ability was a sine qua non for an adequate theory.

6

The electrical field before Faraday

J. L. HEILBRON

Office for History of Science and Technology, University of California at Berkeley, Berkeley, California 94720 USA

'Aether, being no object of our sense, but the mere work of imagination, brought only on the stage for the sake of hypothesis, or to solve some phenomenon, real or imaginary; authors take the liberty to modify it as they please'.[1] We have ethers continuous and discontinuous, material and immaterial, subject to and free from the laws of ordinary mechanics; ethers filling the heavens, pervading the atmosphere, penetrating hard bodies; 'ethers cluttered by a great variety of concepts'.[2] They elude classification by the historian almost as effectively as they have escaped detection by the physicist.

I define ethers as subtle substances that mediate interactions between gross bodies. *Subtle* means very tenuous or rare, highly penetrating, and undetectable directly by sense. *Mediation* signifies the power to propagate action or the potential for action without moving as a whole.[3] Neither light particles nor air is ether: The former are subtle but not a medium; the latter is a medium, at least for sound, but not subtle.

The subset of ethers under discussion here are representations of the electric and electromagnetic fields discussed by physicists from about 1780 to 1830. *Field* in general signifies a region of space considered in respect to the potential behaviour of test bodies moved about in it;[4] the electricians of 1780 lacked the word but not the concept, which they called 'sphere of influence', *sphaera activitatis*, or *Wirkungskreis*. The electric field or *Wirkungskreis* has been represented geometrically, by lines the number and direction of which indicate potential velocities, orientations, and directions of motion; and physically, by a medium whose local pressures, stresses, or other circumstances cause the behaviour of the test bodies. Such a medium, if subtle, qualifies as an electric ether.

Strictly speaking, a field and its representation must act by local forces, by forces exerted between contiguous elements of space or ether, or, in the case of a discrete ether, between a particle and its nearest neighbours. One might therefore reject as a representation of a field any medium the elements of which act not only on nearest neighbours but also on further elements. Such close discrimination is doubtless desirable in an analysis of classical field theories. For my purpose, however, it would not be useful. The older writers seldom developed their theories in sufficient detail to allow an exact judgement of the range or mode of action of an ether particle. I admit all subtle media the function of which is to propagate electric and magnetic action between ponderable bodies appearing to interact directly over macroscopic distances.

Mediation by a classical field takes time. The older writers seldom stated explicitly that the ethers with which they modelled the field worked in time; instead they made analogies to the spread of light and references to the progressive displacements of ether elements. No more did they explicitly declare that the field had energy. But again it is plain that they held the proposition implicitly, for their ethers could be strained or tensed, and moved internally. The ethers we examine here played the same role as the 'medium or substance' that transmitted electrical action in Maxwell's theory, and that he commended to the attention of physicists in the last paragraph of his *Treatise*.[5]

Discomfort with postulating actions over macroscopic distances is endemic among physicists. It afflicted Aristotelians, corpuscularians, Newtonians, and electricians, among others, long before it bothered Faraday. Seeking its etiology in a particular philosophy or epistemology is not promising. Nonetheless some historians, considering the concept of the electric field as an artifact of discomfort with distance forces, pursue its origin in metaphysics. A more obvious and, in the event, a more fruitful approach is to read the old electricians, and to try to grasp the problems that might have recommended the field concept to them. It turns out that the prime mover towards field representations was not a fundamental principle of metaphysics, epistemology, or natural philosophy, but nagging problems in electrical theory.

Both the existence of these representations and their source in standard problems in electricity counter a common opinion about the origin of field theory and Faraday's part in it. Although my main purpose is not to correct this opinion, but to call attention to a neglected class of electrical theories, scholarly civility demands mention of the points at issue.

It is held that Faraday was the 'father of field theory'[6] and that the seeds for his fathering came from Kant and Boscovich. The first assertion concerns field theories that attempt to make do without electrical matter or fluid. It is the theory at which Maxwell aimed in following out Faraday's later ideas, a

theory 'out of harmony' with the 'gross' conception of a 'molecule of electricity'.[7] This conception of the field rules out Lorentz's electrodynamics as well as pre-Faraday electric ethers.[8] In trying to do without accumulations or deficits of electrical matter, Faraday did take a fresh tack in his time, although one common in the eighteenth century; insofar as the result is *defined* to be field theory, Faraday must have been its father. But this is only to say that Faraday invented Faraday's theory.

The second assertion rests on poor historiography. About 1800, we are told, everyone subscribed to the 'dogma of action-at-a-distance physics'. To reach field theory, a 'metaphysical revolution' was required. Faraday alone made it, taking metaphysics from Kant and, especially, Boscovich; 'from his earliest productive years' he believed ponderable bodies to consist of seedless, multivalent, indefinitely extended, mutually interpenetrating 'atoms'.[9] By 1831, or perhaps a decade earlier, he had made a similar theory of electricity: Localised electrical fluids and magnetic poles had been exploded into the electromagnetic field.[10] This interpretation has been attacked on the pertinent ground that there is no good evidence for it.[11] The ground has been conceded without relaxing the interpretation. The very want of evidence has become a supportive argument. 'Hesitant to discuss his metaphysics', inhibited by 'a conspiracy of silence', Faraday did not write what he thought.[12] Here we leave metaphysics for metapsychics.[13]

The historiographical method that produced this interpretation is not uncharacteristic of recent writing in history of science. It appears to rest on the assumption that the innovative scientist acquires his basic and guiding conceptions from general philosophy and metaphysics. This preconception inspires a search for what is literally a philosopher's stone, a single key that unlocks the corpus, the philosophical thread, the leitmotiv, that runs through a life's work. The quest distracts the adept from the first task of the intellectual historian, reconstructing the public state of knowledge. Only after surveying the literature that his scientist may be presumed to have known can the historian identify what in the scientist's work might need explanation by intellectual biography. The explanation will invoke whatever then seems pertinent, including philosophical commitment. The alchemical historian, on the other hand, sublimes directly to the invention of the explicandum, and concocts pseudo-problems in place of historical ones. The alchemical analysts of Faraday have worked out answers to several pseudo-problems. It is pertinent that the latest of these adepts does not mention any of the ether theories with which we shall be particularly concerned, although he says that he has tried 'systematically to find the *problems* which led to the field theories'.[14]

I begin by outlining the consensus about electrical theory reached late in the eighteenth century. The consensus distinguished the roles of field and charge, which previous theories had conflated, and called forth speculations about ether mechanics. Next come accounts of ideas about electric ethers and galvanic currents from 1800 to 1820, when Ørsted's discovery augmented the load that both sorts of models had to carry. The third and last sections concern electromagnetic fields and ethers between 1820 and 1831, when Faraday found the inverse of Ørsted's effect. With that discovery, Faraday began his systematic work on electrodynamics, which led him far beyond the old speculations to the idiosyncratic field theory on which Maxwell built. The development of Faraday's ideas about fields must be reserved for a future essay.

Electrical theories of the late eighteenth and early nineteenth centuries

The physicists of the seventeenth century referred electrical attraction to the action of an effluvium or subtle vapour elicited from electric bodies by friction. The effluvium drew either by spearing chaff and rebounding with it or by mobilising the atmosphere, 'thinning' the near air to cause an inrush of that beyond.[15] The model does not qualify as an ether; it operates by projection and impact, not by mediation.

The discoveries of electrical repulsion and of sparking early in the eighteenth century made the model inadequate. Among the new mechanisms proposed was the 'atmosphere', an aura of electrical matter surrounding an electrified body. Pulsations of the atmosphere cause attractions and repulsions; ruptures of the electrical matter, which contains or is related to fire or light, create sparks and glows. In Franklin's version of the theory, the atmosphere stands quietly; positively electrified bodies repel one another via short-range forces between the particles of their respective atmospheres. Note that the atmosphere is the *charge:* The surplus of electrical matter that constitutes positive electricity spreads over and beyond the surface of the body it clothes. Such an atmosphere functions both as the source of the field and as an ether: As source it can be accumulated and conveyed; as ether it acts by pushes between neighbouring particles.

This model could not be extended to cover the air condenser, which can sustain a charge with its plates so close together that the grounded one lies within the atmosphere of the other. In this situation, according to Franklinist theory, the condenser should short internally; the atmosphere, as charge, should run to ground. Similar puzzles were presented by the dissectible condenser, or electrophorus, invented in 1775. Physicists escaped from their difficulty by dissociating the functions with which they had burdened the atmosphere. As source of the field it became a charge placed very close to the electrified surface. As the field, it lost its material connection with the charge, and ceased to be an ether: *Atmosphere* became synonymous with *Wirkungskreis* and *sphaera activitatis*, with the space in which electrical forces could be detected.

Consensus regarding the localisation of charge was established during the 1780s. The rules of the *Wirkungskreis* accounted for the phenomena: Accumulations of charge have atmospheres, or spheres of influence, within which the electrical matter normally contained in conductors segregates into plus and minus charges. 'Electrical motions' of ponderable bodies were usually represented as consequences of distance forces acting between the accumulated and segregated (or induced) charges, and between electrical and common matter.

The consensus included acknowledgment that no known experiment could decide whether electricity came in one fluid or two. In the former case one had to postulate a repulsion between particles of common matter or devise an ether that could otherwise account for repulsion between negatively charged bodies. In the latter case symmetry between positive and negative electrification was obtained at the cost of reasoning about a second electrical substance. One affected to choose on the basis of simplicity or convenience. The English mind tended to be singlist; the Continental, and especially the French, dualist. Although the singlists included authorities like Volta and van Marum, they probably made a minority among electricians by 1790, and certainly did so during most of the nineteenth century. The equal and contrary electrochemical powers of the poles of the pile, invented in 1800, strengthened the position of the dualists.[16]

The consensus of the 1780s referred to the localisation of charge, to the rules of the *Wirkungskreis*, and to peaceful coexistence between singlists and dualists. It did not bring agreement about the machinery that might account for the rules of electrical interaction. Many proposals were made, including some ether representations; the physicists of the eighteenth century were not content with the elementary scheme of distance forces with which historians credit them. Their disquiet arose not only from the usual and usually inconsequential scruple against action at a distance, but also, and primarily, from a non sequitur, or even a contradiction, in the rules of the *Wirkungskreis:* the connection between motions of electrified bodies and their supposed cause, attractions and repulsions between elements of electrical fluid(s).

Electrical matter placed inside a conductor spreads instantly to its surface. Similarly, the natural supply of electrical matter in a conducting body segregates immediately under the influence of an external charge. It appeared that no force exists between the electrical fluid(s) and the molecules of conductors, or at most a force very small compared with that between elements of the electrical matter. Two puzzles resulted: What retains a charge at the surface of a metal sphere? How do electrical attractions and repulsions develop the ponderomotive forces needed to account for electrical motions?

The usual answer to the first puzzle rested on accurate but misleading experiments. Rarified air does not begin to insulate until reduced to a pressure of about 10^{-3} atmospheres. The best air pumps of the eighteenth century did not reach pressures so low; not until after 1850, with the help of Geissler's mercury pump, did physicists find the region of increasing dielectric strength.[17] Throughout our period they reasoned that, since an electrified conductor would discharge immediately in a perfect vacuum, its electricity must be clamped upon it by the pressure of the surrounding air. As Biot explained, the surface layer of repellent charge exerts an outward force on each particle of the electric fluid proportional to the local depth of the layer (or to its surface density). These forces, when added together, produce a pressure proportional to the square of the thickness (or the density), and this pressure must be resisted by the weight of the adjacent atmosphere.[18] As Faraday remarked, Biot's theory assumes the existence of 'gross mechanical relations' between 'the ponderous air and the subtle . . . fluids of electricity'.[19]

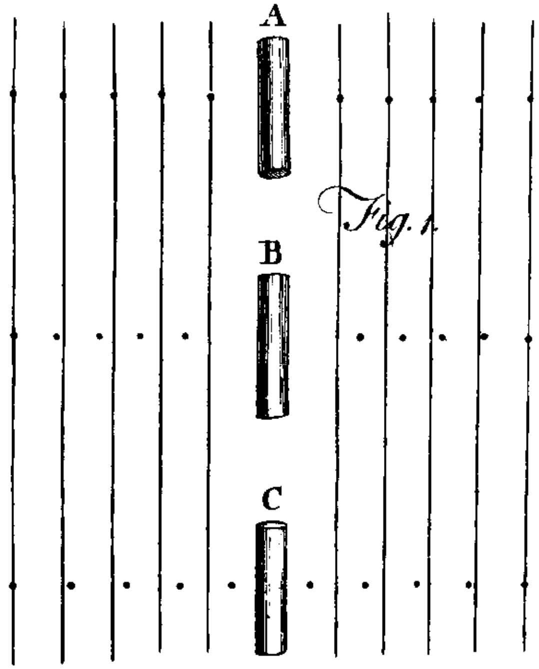
These gross mechanical relations provided an answer to the second question as well: The ponderomotive forces responsible for electrical motions arise from imbalances in air pressure set up by charges. For example, surplus fluid given a pair of touching suspended cork balls accumulates opposite to the point of contact, and creates there a large pressure against the atmosphere. 'Evidently,' that is, in order to save the phenomena, this excess in pressure drives off some of the air, allowing the normal air between the balls to force them apart. Attractions occur via the pressure gradient called up by charge accumulations on the facing surfaces of the balls.[20] In general, the fluids distribute themselves as required by their distance forces, and the resultant accumulations cause motions by mobilising the air.[21]

The account just sketched is not an ether theory or even a medium theory, since the air acts not by propagating pressures but by immediate contact. It is easy, however, to construct a rudimentary ether theory by substituting for the 'gross mechanical relations' forces between the charges accumulated on bodies and the electrical fluid(s) naturally present in the air. Perhaps the earliest example of this approach is John Canton's representation of electrical atmospheres as 'alterations of the state of the electrical fluid contained in and belonging to the air'. In Figure 6.1, A, B, and C are neutral, positive, and negative, respectively. Since B's charge repels surrounding electrical matter

(Canton was a singlist), the air near B has less, and beyond B more, than its usual amount, up to a certain distance, represented by the outermost lines in the figure, where the state returns to normal. Since the air's electrical matter moves towards C, the gradient about C is opposite to that about B. An electrified body polarises the medium around it; a conductor suffers induction under the influence of the displaced electrical matter of the air *adjacent* to it.[22]

Canton's suggestion propagated through a string of British singlists, including Tiberius Cavallo and Thomas Milner, who tried to base a theory of electrical motions upon it.[23] Milner thought to obtain ponderomotive forces as interactions between the electricities of bodies and the polarised air contiguous to them: 'The air thus charged is the medium which enables one electrified body to act upon another at some distance'.[24] Cavallo explained how the trick is played: Neither of two touching electrified balls can sit at the centre of the system of polarised air it creates. Hence, on the assumption that 'there is something on the surface of bodies which prevents the sudden incorporation' of the contrary electricities of the balls and the air, the balls will move

Figure 6.1. Canton's representation of the electric field. (From Priestley, 1775, *The history and present state of electricity*, London.)



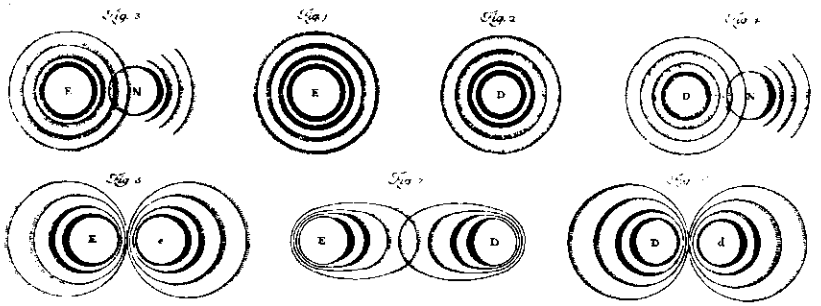
apart into the centres of their atmospheres under the attractive force assumed to act between electrical and common matter.[25]

The implausibility of Cavallo's special mechanism – the initial polarising of the air about 'centres' outside the balls – was granted by everyone, including himself. Other singlists tried to do better, inspired by a desire to avoid Aepinus's solution to the problems that recommended dualism: the infamous force of repulsion between particles of ponderable matter. Although they continued to use the analogy between air and insulators known to polarise, like glass and tourmaline,[26] they now emphasised the conductivity of the atmosphere.[27] The electrified balls polarise the air, but they also charge it by conduction. When the air between them has acquired a surplus of electricity of the same sign as theirs, they are pushed apart by the contiguous electrical matter, assisted, perhaps, by the polarisation of the remoter air.[28] Whatever the details, the general proposition that electrical motions and inductions over air gaps occur via electrical matter made active in the surrounding air was well represented in texts of the early eighteenth century.[29]

A notable form of the ether theory of electrostatic interactions was published in 1806 by Amedeo Avogadro, who began with Canton's ideas as mediated by Beccaria.[30] According to Beccaria, 'the electricity of a body resides within the superficial pores of it, and actuates the ambient air, not by diffusing itself into it, but by exciting either a tension or relaxation in the natural [electrical] fire inherent in it'. That he thought of these actuations as polarisations appears from Figure 6.2. Induction occurs via the tensed air. So do ponderomotive forces, if one allows the principle that bodies interacting electrically strive to move so as to minimize the tensions in the air.[31]

Avogadro goes forward from the proposition that no electricity can appear on any surface unless an equal and opposite electricity can develop on an-

Figure 6.2. Beccaria's representation of the electric field. (a) About a positive body, (b) about a negative one, (c) between positive bodies, (d) between negative ones, (e) between unlike ones. (From Beccaria, 1772, *Elettricismo artificiale*, Turin.)



other. This answering electrification does not occur in response to a force acting over macroscopic distances. Only the universal force, gravity, has such power. An electrified body acts by imposing an equal and opposite electrification on the contiguous insulating layer. To fix ideas, consider a horizontal air condenser with the positive plate uppermost. The top layer of air becomes 'charged' (a 'peculiar state' enjoyed only by dielectrics) in a certain manner, imposes the charge on the next layer, and so on, until the lower plate is reached, and rendered negative by the air adjacent to it.

Avogadro, a singlist, pictured the normal air molecule as a little sphere surrounded by a spherical shell of electrical fluid. The surplus fluid of the upper plate distorts the atmospheres of the molecules adjacent to it; a relative defect occurs above, a relative surplus beneath. This surplus in turn decentres (to use Avogadro's term) the atmosphere of the next lowest molecule, whence the distortion propagates to the lowest layer of air. Avogadro clearly stated that 'charge' consists of strain in the air or other insulator that preserves the 'electricity' of conductors, and that all electrical motions arise from very short-range forces propagated through dielectrics.

In support of his views, Avogadro pointed to the behaviour of dissectible condensers and to the complicated phenomena of absorbed or residual charge, which had inspired Beccaria to introduce the cumbersome concept of vindicating electricity.[32] One experiment may stand for them all. A glass plate sliced through longitudinally is armed with metal plates top and bottom and electrified as a parallel plate condenser. Connect the armatures and explode the condenser; remove them; separate the glass slices. The upper glass, the one that had touched the positive armature, shows negative on its top surface and positive on its bottom; the lower glass shows negative on top and positive on bottom. Avogadro inferred that every horizontal layer in the dielectric is polarised; and he used this model to explain what Beccaria and Volta called the 'oscillation of the electricities', the reversal of polarity of the coatings when a Leyden jar explodes. The effect arises from induction in the coatings occasioned by residual charges in the dielectric, much as Avogadro pictured it.

Avogadro's memoir had a good press. It appeared in English and German as well as in French.[33] But it did not receive the attention that its elegant solutions of electrostatic problems deserved, perhaps because most electricians were then fascinated with the Voltaic cell. Perhaps Faraday knew nothing of Avogadro's concept of the role of the dielectric medium. After the publication of Faraday's ideas about electrostatic induction, Avogadro himself called attention to his neglected work, 'for the history of science', and observed that induction could be understood to operate on the ether as well as on ponderable matter.[34]

Electric ethers and galvanic currents, 1800–1820

The electrician of the late eighteenth century, who served sometimes to short-circuit the Leyden jar, had an intimate acquaintance with the discharge: a rushing together through his body of the equal and opposite electricities unable to unite directly through the bottom of the jar. The explosion annihilated the electricities; to reproduce them required labour, such as the cranking of the electrical machine, no less personally felt than the discharge itself.

Physicists tried to apply this understanding to the pile. According to Biot's conception, which came to dominate, the apparently continuous current stemming from the pile in fact consists of a very large number of discharges. The pile has the capacity when open to acquire opposite electricities on its terminals, which show the usual signs of frictional electricity. Close the circuit by taking one terminal in each hand; the old signs disappear; you feel a 'commotion', similar to that caused by a feeble torpedo. Release the terminals; the old signs reappear, the commotion ceases. Apparently the pile has the power of recuperation; and it is easy to imagine that, when closed, it explodes and recuperates so quickly as to appear to act continuously. 'It is becoming more and more probable that [galvanism] is the successively repeated effect of a very feeble electricity possessed of a very great velocity'.[35]

How the pile acquired its polarity was a matter of lively debate. Did it work entirely, as Volta and Biot thought, by the contact of dissimilar metals, or, as William Wollaston and the English argued, by chemical combinations between the metals and the moist conductor separating them?[36] Despite their difference over this major point, both schools accepted Biot's conception of a discontinuous *external* current. Davy, for example, explained that the contact disturbs the electrical equilibrium; closure of the circuit restores it; chemical action revivifies the metals; their contact again generates electricity; and so on.[37] In trying to picture the action of this discontinuous current physicists once again flirted with ethers and mediated forces.

Consider first electrochemical processes such as those instigated by the current in acidulated water. How do the hydrogen particles come to the negative electrode and oxygen to the positive? Davy suggested that the electrodes might polarise the water just as charged conductors segregate the electricities in neighbouring pieces of wire. The molecules touching the electrodes are torn apart; one fragment escapes and the other combines with the oppositely charged portion of the neighbouring polarised molecule, whose liberated portion decomposes the next polarised particle, and so on. In the middle of the fluid, where the molecules remain neutral, the last fragments to be released in the progression from each electrode unite into a normal water particle.[38] The concept that a substance undergoing electrolysis first polarised like a conductor suffering induction and then decomposed as the result of contact forces propagated from particle to particle was not original with Davy, who indeed called it the 'received opinion'.[39] Perhaps its earliest proponent was Theodor von Grotthuss, who reasoned from a fanciful analogy between the polarised pairs in Volta's 'electrical magnet' to the state of the water molecules between the electrodes.[40]

Also outside the pile, in a circuit uninterrupted by an electrolytic cell, phenomena occur that focus attention on the medium. Foremost among them are the heat and light produced in wires of narrow gauge. Many investigators, adapting a research topic in frictional electricity, tried the lengths of wires that the current could melt, and the variety of refractory substances that it could fuse. Elaborate and expensive equipment was built for the purpose, such as the battery with plates of thirty-two square feet put into operation by J. G. Children in 1815.[41] To understand these effects the physicist needed to connect his representations of electricity and of heat and light. Many associations were proposed,[42] the variety arising from disagreement over the number of electrical fluids and over the nature of light and heat. A common opinion built upon dualism and the theory of caloric, which referred heat to a material substance spread through the spaces separating the particles of common matter.[43]

Recall that the current was pictured not as a continuous river, but as a sequence of squirts, which start from the terminals at every destruction of tension. On striking the electrical matter combined in the neutral wire, the squirts create disturbances that propagate in opposite directions towards mutual annihilation in the middle of the circuit. The propagation might be conceived as a wavelike motion in the wire's electrical matter,[44] or, in analogy to prevailing theories of the pile, as disruptions and recombinations of the two electrical fluids. (The possibility that oppositely directed squirts move freely until colliding in the middle of the wire was harder to credit.)[45] By 1820, if not before, the dominant theory of the external current, or rather its very definition, was a series of microscopic displacements of the two electricities: A current, says Ampère, is a 'succession, in all the particles of conducting wires, of decompositions and recombinations of the fluid formed by the union of the two electricities'.[46]

To explain the appearance of heat in and beyond the wire it was assumed that the conflict of the squirts put in motion or freed caloric. For example, if, as many speculated, the electrical fluids consisted of 'bases' plus heat and light, the latter might be detached during the disruptions and recombinations in the current-carrying wire.[47] The passage of an electric current might set up vibrations in this matter as well as disruptions; and these vibrations, communicated to the environing caloric ether, would constitute the propagation of heat.[48] For those who accepted the undulatory theory of light, the glow of high-resistance wires propagated in the same manner as their heat; for others, such as Biot, who retained the emission theory, the vibrations threw off light particles 'in a purely mechanical way'.[49]

A consequential form of this theory was worked out by Ørsted. In 1806, after descending from the heights of *Naturphilosophie* to a level more common to physicists, he published a paper on electricity to which he attached great importance.[50] His purpose was to assimilate the mechanisms of conduction and induction. He observed that insulators, including the air, polarise under an external charge whether the charge stands on or next to them. In the first case the insulator electrifies by conduction: The surface touched takes on electricity of the same type as the external charge; the layer just beneath the surface becomes oppositely electrified by induction; and so on. In the second case, the surface electrifies oppositely to the external charge and the adjacent layer does similarly. Now, according to Ørsted, in an imperfect insulator the situation just described holds for the first instant of induction only; in the next instant, the electricity induced in the second layer separates the normal electricities of the third, combines with its opposite to free, momentarily, its homologue; and so on. Exactly the same process occurs, although in an insensible interval, in conductors, which can be defined as very poor insulators. Ørsted calls the process of conduction and induction 'undulatory', in reference, apparently, to the sequence of unions and divisions of the electrical forces at each point in a body passing a current.

The wire closing the galvanic circuit enjoys a peculiar state. Undulations commence at each of its ends and propagate rapidly towards the centre, where, if they meet no resistance, they annul one another quietly. The pile recuperates instantly to produce the next set of undulations, which performs as the earlier. Ørsted notes that 'the electrical forces are conducted only by themselves', by which he means that the instrument of conduction is the electrical forces supposed to occupy all of space.[51] If the forces in the wire suffer impediment to their reunion after separation, as they do when the magnitude of the current exceeds their powers of recovery, heat results. If the impediment is strong, light may be produced, of a colour dependent upon the violence of the union. Evidently both light and (radiant) heat spread throughout the region around the conducting wire. This propagation also takes place via

'undulations' in the forces occupying space.[52] The electric spark propagates by a very rapid breakdown and recovery of the electric forces in the air;[53] a light ray is nothing but a 'series of immeasurably small electronic sparks'.[54] These views claimed the careful consideration of all physicists, when, in pursuing them, Ørsted discovered the action of a galvanic current on a magnetic needle, and so brought to a close a 'long, sorry time of endless, purposeless, indecisive and fruitless detail' in galvanic researches.[55]

In 1812, Ørsted suggested that to detect an interaction between electricity and magnetism one should employ electricity in a 'bound' or 'latent' state. He reasoned that in frictional electricity the opposed forces are free, as appears from the usual attractions and repulsions; in galvanic electricity, they give heat, light, and chemical action, but stay too tightly together to cause electrical motions; in magnetism they may be so closely bound as to have escaped detection.[56] He proposed to try electricity in the most bound state of galvanism, where large currents and resistances constrain the electric forces. He persevered desultorily with large piles and narrow wires, sustained by the hope that undulations sent forth by a glowing wire might excite a magnet as well as the eye. Success awaited placing the needle parallel to the wire. Ørsted did so early in 1820, and soon learned that the strongest deflections occurred with the least resistance.[57]

Ørsted advertised his results in a Latin pamphlet entitled 'Experiments on the effect of the electric conflict on the magnetic needle'. The 'conflict' in the wire is just the converging undulations starting from the poles of the pile; as expected, it also acted outside the wire, in the 'circumadjacent space'.[58] There, contrary to all analogy, the conflict 'performed circles', driving the needle in one direction when placed below the wire and in the opposite when placed above. Ørsted supposed that negative electricity moves in a right spiral about the wire and acts only upon north poles, while positive electricity moves in a left spiral and acts only on south poles.[59] The intricate dance of the conflict calls up a material image; and, indeed, by giving the 'impetus of the contending powers' as the cause of the needle's motion, Ørsted suggested a process more mechanical than usual in dynamical (*naturphilosophisch*) physics.[60] Many of his readers tried to express his conflict through a material ether.[61]

Electromagnetic fields and ethers in the 1820s

Ørsted's fundamental fact related an apparently heterogeneous pair, a magnetic pole and an electric current. Physicists tried to grasp this hybrid in three different ways: by following Ørsted and supposing a novel, circular force between currents and poles; by assimilating the wire to a magnet; and by assimilating magnets to currents.[62] The last two approaches differed essentially from the first. They aimed at reducing all electromagnetic phenomena to a single elementary *force* (italicized to indicate a hypothetical, unmeasurable attraction or repulsion between infinitesimal *elements*); the first aimed at an exact description of the motions of interacting poles and wires, or of their mutual forces, the measurable integrals of *force*.[63]

The only reductionist theory now remembered is Ampère's. Building on his demonstration that current-carrying wires exert ponderomotive forces upon one another, he argued that the fundamental *force* is a rectilinear push or pull between elements of current: attraction between elements moving parallel, repulsion between ones moving in opposite directions. On this assumption he obtained, by double integrals, the observed ponderomotive forces.[64] Ørsted's fundamental fact, as expressed in the law of Biot and Savart,[65] proved more difficult. To deduce it, Ampère required the bold hypothesis that a magnet owes its power to elementary current loops perpendicular to its axis.[66] Mathematical derivations of the Biot–Savart law from Ampère's hypotheses were accomplished in 1823 by F. Savart and J. F. Demonferrand, 'marking a sort of epoch in the history of electrodynamics'.[67] All understood perfectly the relation between laboratory forces and Ampère's undetectable, hypothetical, electrodynamic *force*.[68]

A second *force* theory, developed mainly by German physicists, reduced the wire to a sequence of magnetic poles placed symmetrically at its surface. They tried to do with two poles in each cross section of wire, then with four, finally with two for every molecule.[69] This last version of the theory of 'transverse magnetism' resembles Ampère's: Both derive elementary magnets from a current at right angles to the magnetic moments. But Ampère's *replaced* magnets by currents, whereas the German theories admitted both electricity and magnetic poles. The replacement enabled Ampère to deal with homogeneous interactions only, to reduce the interactions of currents, of poles and wires, and of magnets to one and the same principle. 'Is it not evident that one must look for the primitive fact in the action of two things of the same nature?'[70] To the more literal Germans, it seemed preferable to take as primitive a *force* of the same character as the force observed in Ørsted's experiments.[71]

A decisive argument against transverse magnetism was developed in England by physicists who also emphasised measurable, macroscopic force. Accepting from Ørsted's account that the chief electromagnetic phenomenon was the tendency of poles to 'perform circles' about wires, Wollaston tried to explain the interactions of Ampère's wires 'upon the supposition of an electromagnetic current passing round the axis of [each]'.[72] Although the public

notice of his ideas is unintelligible, he was understood to think that 'a kind of revolution of magnetism [occurred] around the axis of the wire', a 'vertiginous magnetism' that swept poles about in circles.[73] Both Davy, who had inclined toward transverse magnetism, and Faraday, who had tried to do with rectilinear attractions and repulsions, adopted Wollaston's manner of interpreting electromagnetic interactions.[74]

Further experiment convinced Faraday that the primitive fact, the elementary phenomenon, was the rotation of a pole about a wire, or of a wire about a pole. His demonstration of these continuous rotations by ingenious apparatus that, in contrast with the usual arrangements, allowed motions over 360°, made his reputation as a physicist. They also made the theories of transverse magnetism untenable, as Faraday, Ampère, and many others insisted.[75] Ampère had no difficulty in accounting for Faraday's rotations, or in obtaining revolutions of wires and magnets about their own axes, an effect Faraday had sought in vain.[76] Above all, Ampère and Faraday could simulate a magnet by a solenoid, which they could make rotate like a needle, whereas no practicable polygonal deployment of needles could be made to drive a pole in a circle.

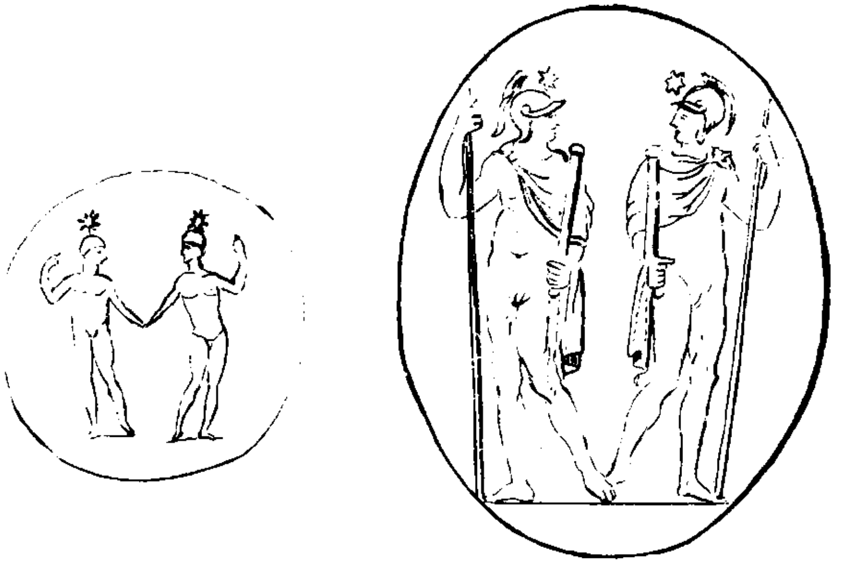
The explosion of the German theory left four principal alternatives. One could swallow one's objections to Ampère's 'unnatural assumptions', his arbitrary preference for electricity over magnetism, and his bizarre molecular currents, which have no apparent cause and do not, like other currents, spread through conductors.[77] Or one could patch up the German theory, making the *force* between the transverse magnetic poles and the poles of the needle act at right angles to the line joining them. Or, what comes to much the same thing, one could postulate a transverse *force* between elements of current and elements of boreal or austral fluids in the compass needle.[78] None of these three solutions had much appeal, among other reasons because of 'the difficulty of conceiving the mechanical principles by which such a tangential force . . . can operate', its want of analogy to the rest of physics, and its incompetence to explain all of electrodynamics.[79]

The fourth alternative eschewed these various *forces* in favour of observables. Here the difficulty of phenomena 'sufficient to perplex the mind' recommended the procedure to be followed. The memory needs a crutch, a picture or diagram; 'words alone will not do'.[80] Imagine yourself in the circuit, head towards the zinc end of the pile, face towards the needle: Its north pole will move towards your left hand when the current enters your feet.[81] Or place a picture of the Dioscorides (Figure 6.3) on the wire; the left hand of the right twin points to the direction in which a north pole moves.[82] This is no coincidence. According to the ingenious decypherer of these hieroglyphs, Castor and Pollux originally represented the contrary electricities. The physical knowledge of the builders of the pyramids may thus be made to serve the perplexed student of electromagnetism.[83]

Such representations put one either directly or vicariously in the space where electromagnetic interactions occur. The very complexity of the phenomena drew attention to the field, to the space considered with respect to test bodies moved about in it, and to the need for maps. T. J. Seebeck pictured concentric circles of iron filings in the 'magnetic atmosphere' of a wire.[84] English physicists, in keeping with their emphasis on macroscopic representations, consumed quantities of filings. Davy used them in his first set of electromagnetic experiments; Faraday sprinkled them about wires and helices, a 'very beautiful indication of [the] course of action'; Roget showed a figure similar to Seebeck's.[85] These 'magnetic curves' had the same significance as similar lines about magnets; namely, they were tangents to the 'positions which an infinitely small compass needle . . . will assume'.[86] Alternatively, physicists drew lines representing the paths of wires around poles, and of poles around wires, filling space with curves indicative of the directions of 'the forces that are drawing' the objects in motion.[87]

As in electrostatics, so in electrodynamics, the possession of clear rules and representations was not enough for physicists. They must know what

Figure 6.3. Schweigger's electromagnetic heiroglyphs. (From Schweigger, 1826, *Journal für Chemie und Physik 46.*)



goes on out there, in the field, where filings clump and the Dioscorides play. What is the 'essence of the Naturkräften'? What are the 'qualitative aspects of experience'?[88] P. Erman stressed the need for an accurate map of electromagnetic motions, which he planned to explain by hydraulics; Seebeck hinted that his filings lined up under a 'magnetische Strömung'; Faraday took Ampère to task for failing to provide a physics in the style of Ørsted.[89] In fact, many physicists of the 1820s, including Ampère, indulged conjectures deriving from Ørsted's conception of electric conflict. Two sets of these conjectures may be distinguished. One, emphasising the spiral force outside the wire, used analogies to fluid flow. The other, interpreting the conflict as a source of disturbance, used analogies to the propagation of waves.

Ørsted's circulating magnetic force, sweeping about poles as a whirlpool sweeps corks, recalled the comfortable physics of the Cartesian vortex. It also saved some phenomena. 'It is still scarcely possible', Schweigger wrote in 1826, 'to describe the chief features [of electrodynamics] more correctly than through the picture of a magnetic atmosphere', by which he meant 'rotatory magnetism', a vortex (*Wirbel*) extending outwards from the wire.[90] Davy and Wollaston were understood to have accepted Ørsted's maelstrom, and Faraday to have demonstrated it. 'Most discoveries made since the publication of Ørsted's works come to support his way of conceiving the general philosophy of nature'. The magnetic fluids that accompany the electric ones in motion rush in opposite vertiginous paths around the axis of the connecting wire.[91]

The other set of analogies deriving from Ørsted's conflict was ether models, mediums for the propagation of action. Ampère saw a deep parallel between electricity in motion and light.

> Everything that has been done in physics since the work of Dr. Young on light and the discovery of M. Ørsted is preparation for a new era . . . Explanations deduced from the effects produced by the motion of imponderable fluids will gradually replace those now accepted . . . I believe that we must look to the motions of fluids distributed in space for the explanation of general effects.[92]

The space-filling medium is formed by the union of the two electricities. Its vibration, or series of separations and combinations, constitute light; interference of disturbances from different wires accounts for electromagnetic motions; the electric spark and the Ørsted effect propagate from the conflict of electric current in the wire.[93] 'M. Ørsted regarded compositions and decompositions of electricity . . . as the unique cause of heat and light, i.e., of the vibrations of the fluid occupying all space . . . This opinion of the great physicist . . . agrees perfectly with all the phenomena'.[94]

Ampère's universal luminiferous and electromagnetic ether pervaded the

Annales de Chimie. Becquerel exploited it for chemistry and heat; Marianini's experiments on strengths of current were referred to the hypothesis that 'the electrical current is propagated by vibrations in the manner of sound'; A. de la Rive used Ampère's ether to elucidate calorific effects of currents; L. Nobili accounted for his rings by 'a law of interference' in the spread of electric currents; Savary traced pecularities in magnetism to attenuation of the 'vibrations transmitted from the wire to the environing media'; E. Becquerel reached a similar conclusion from the appearance of the wire itself.[95] 'Maybe Newton [the Newton of ether] glimpsed the true cause of electrical phenomena'.[96]

Our two sorts of mechanical analogies do not exhaust the range of pertinent proposals made during the 1820s. Were we to look further, we should find oddities like Erman's expansive flowing electrical *Thätigkeiten;* G. F. Pohl's oscillating antagonistic longitudinal electric and transverse magnetic tensions; and C. Hansteen's 'electric-fluid magnets', neutral polar products of the conflict, which surround the wire 'like the circular wave around a rock dropped in water'.[97] But we already have enough to suggest that the phenomena of electromagnetism caused many besides Faraday, including the paladins of mathematical reduction, to conjecture about an ether competent to propagate heat and light, and to mediate electrical interactions. It is striking that Ampère's protégé A. de la Rive, writing shortly after the discovery of magneto-electric induction, offered just such a conjecture as the conclusion to his history of electrodynamics during the 1820s.[98]

Faraday's response

The obvious inverse of the Ørsted effect, the induction of a current in a passive circuit by a current-carrying wire laid next to it, did not work. Nor did an electric current arise in a wire wrapped round a magnet, as one expected on Ampère's theory.[99] Ampère himself desired to know whether the molecular currents to which he attributed the power of magnets could be induced by external currents. He managed to deflect a thin copper disk suspended within a wire coil, an effect he interpreted as a consequence of the induction of one current by another. There he stopped. Although the phenomenon suffered from anomalies, he did not deem them important: There was no room in his *reductionist* theory for an induction of electricity by magnetism different from the straightforward, and spurious, induction of currents that he had detected.[100]

This proposition may be illustrated by Ampère's interpretation of Arago's discovery that a magnetic needle may follow the motion of a metal disk rotated beneath it. Ampère immediately tried whether a solenoid could substitute for the needle. The positive result satisfied his interest in the phenomenon: It fell under his general principle of the equivalence of magnets and current loops. He had only to postulate that the induction of one current by another took time. While the molecular currents grow or die, 'they exert forces that probably produce the singular effects that M. Arago discovered'.[101]

An alternative to Ampère's approach was to allow a magnetic body the power to induce poles in a metal plate rotating in its vicinity. The tangential component of the driving force might then be understood by supposing that the magnetism induced sequentially in each sector of the rotating disk reached a maximum after the sector had passed under the needle.[102] Still, the need for relative motion remained a mystery: Why could a magnet not induce poles in a stationary copper plate if it could do so in a moving one? To this difficulty, Arago added another: The vertical component of force on the needle turned out to be repulsive, not attractive as the magnetisers supposed.[103]

Faraday gave Arago's phenomenon sustained attention. The 'extraordinary character of [the] motion' made it a fine lecture demonstration; the impotence of philosophers to explain it made it a moral lesson as well as a challenge. 'Old obstacles removed – new ones raised – highly progressive'.[104] He puzzled over the facts that copper acted more powerfully than iron and that the drag (or push!) on the needle was much diminished by cutting radial slits in the disk. 'Very curious', he remarked of the first fact, 'and unlike ratios of ordinary magnetic power'. As to the second, it indicated 'a vortex or current'.[105]

One might think that from the base just sketched Faraday sprang to his grand discovery, the principle of magneto-electric induction.[106] A big clue had come to light: Metals acted the more powerfully the better they conducted electricity. Faraday had guessed that currents were implicated; he had only to recognise that the currents generated by the disk's motion would deflect the needle by the Ørsted effect. Then, examining the motion with his usual care, he would discover the principle for which everyone had been searching, the means of producing electricity from magnetism.

Two obstacles closed this line of reasoning. Firstly, the experimental arrangement, the deleterious effect of the radial slits, and a presumption in favour of closed currents would suggest that any current induced in the disk, whether electric or magnetic, would circulate *around* the axis of rotation. Faraday himself spoke of a 'vortex'. Such a current would not exert a tangential force on the needle. Only a radially directed current would do so; but where in the radial direction was there anything resembling a wire closing a Voltaic circuit or Ampère's molecular current? Secondly, a clear idea existed of the expected inverse of the Ørsted effect, and it did not involve motion of ponderable bodies.

Faraday tried three different ways to obtain electricity, or an electric force, from magnetism before he succeeded in 1831. All three were static arrangements interpretable as variations of Ampère's experiments on the induction of currents.[107] The experimental arrangement with which Faraday succeeded, however, the famous transformer ring, is more easily interpreted on Ørsted's than on Ampère's theory. The iron ring, wrapped on either side of a diameter with coils of wire, would be a conduit for circulating or undulating magnetic forces. Just as these forces were created in, and spread from, the conflict of the two electricities in the connecting wire in Ørsted's experiment, so an electric force might arise where oppositely propagating magnetic forces, concentrated in the iron ring, came into magnetic conflict. And so it was. When Faraday brought a pile into the circuit with one coil, say *A*, a galvanometer connected with the other, *B*, passed a current.[108] The inverse Ørsted effect had eluded physicists not because it is weak, but because in the usual experiments it was transient.

Faraday referred the transiency to the passage of a 'very short and sudden . . . wave of electricity' in *B*, 'caused at moments of breaking and completing contacts' in *A*. 'Hence', he concluded, '[there] is no permanent or peculiar state of [the] wire from *B*'; it responded passively to the pulse struck in it by the magnetic conflict.[109] Faraday set a high value on the discovery of this magnetic wave and its associated electric tension. On 12 March 1832 he sent a sealed note for deposit at the Royal Society. When opened a century later, this note seemed extraordinarily bold and prescient. 'Certain of [my] results', Faraday had written,

> lead me to believe that magnetic action is progressive, and requires time; i.e., when a magnet acts upon a distant magnet or piece of iron, the influencing cause . . . proceeds gradually from the magnetic bodies and requires time for its transmission which will probably be found to be very sensible. I think also, that I see reason for supposing that electric induction (of tension) is also performed in a similar progressive time.[110]

This progressive motion he compared to waves in the transmission of light and sound.

The informed electrician of 1832 would not have been astounded by these conjectures. The various ether theories of the Ørsted effect assumed the propagation of the conflict in time. Faraday knew these theories: He had studied Ørsted's work carefully and had received letters from Ampère detailing the hypothesis of the electromagnetic and luminiferous ether. Moreover, he

knew, and probably favoured, the old ether theory of electrical action. His first authority, the third edition of the *Encyclopaedia Britannica*, ascribed electrical action to stresses in the same medium that propagates light, and illustrated the models of Canton, Beccaria, and Cavallo;[111] a textbook he much admired, Singer's *Elements*, accounted for repulsion between like-charged objects in the manner of Cavallo, placing the seat of ponderomotive forces not in bodies, but in the medium between them.[112]

The claim for priority in Faraday's sealed note does not include a wavelike theory of *electromagnetism*, and quite properly, as many others were there before him. The note might best be construed as a hedge against those who might try to explain his new results on magneto-electric induction by an obvious generalisation of the usual ether theories. It bears strong witness that the ideas about the field and its representation with which Faraday *began* his systematic work on electromagnetism were familiar to the physicists of his time.[113]

Notes

Owing to restrictions in space, it has not been possible to give all pertinent citations to the primary literature examined.

1 'Aether', *Encyclopaedia Britannica*, 3rd ed., 18 vols. (Edinburgh, 1797), 1:216.

2 M. J. Brisson, *Dictionnaire raisonné de physique*, 2nd ed., 6 vols. (Paris, 1800), passage quoted from 3:158; Charles Hutton, *A philosophical and mathematical dictionary*, 2nd ed., 2 vols. (London, 1815), 1:487; J. T. S. Gehler, *Physikalisches Wörterbuch*, ed. H. W. Brandes et al., 20 vols. (Leipzig, 1825–45), 1:271.

3 Cf. M. A. Tonnelat (1959).

4 J. C. Maxwell, *Scientific papers*, ed. W. D. Niven, 2 vols. (Cambridge, 1890), 1:527; W. Thomson, quoted in H. Stein, 'On the notion of field in Newton, Maxwell and beyond', *Historical and philosophical perspectives of science*, ed. R. H. Stuewer (Minneapolis, 1970), 264–86, 299–310, on 306: 'Any space at every point of which there is a finite magnetic force is called . . . a magnetic field'. Cf. ibid., 266; M. B. Hesse (1961), 197–8.

5 J. C. Maxwell, *A treatise on electricity and magnetism*, 3rd ed., 2 vols. (Cambridge, 1891), 2:493 (§866). Cf. Stein, 'The notion of field', 299, 305.

6 L. P. Williams, *The origins of field theory* (New York, 1966), 67.

7 Maxwell, *Treatise*, 1:380 (§260).

8 Cf. H. A. Lorentz, *The theory of electrons*, 2nd ed. (Leipzig, 1915), 22–3, 30–3, 213–15.

9 L. P. Williams, *Origins of field theory*, pp. 31, 67, 78. Cf. L. P. Williams, 'Faraday and the structure of matter', *Contemporary Physics 2* (1960), 92–105, on 95, 100; Berkson (1974), 25, 28, 30, 33.

10 L. P. Williams, 'Michael Faraday and the evolution of the concept of the electric and magnetic field', *Nature 187* (1960), 730–3; Joseph Agassi, *Faraday as a natural philosopher* (Chicago, 1971), 79–80.

11 J. Brookes Spencer, 'Boscovich's theory and its relation to Faraday's researches: an analytic approach', *Archives for History of Exact Sciences 4* (1967), 184–202; T. H. Levere, 'Faraday, matter and natural theology: reflections on an unpublished manu-

script', *British Journal for the History of Science 4* (1968), 95–107; Heimann (1971).

12 Respectively, Berkson (1974), 53; and Agassi, *Faraday*, 28–9, 323.

13 Berkson (1974), 33, 40, 46, 48, 60, 62. 'Faraday's new metaphysical view . . . became clear to him in a rush'. Ibid., 72.

14 Ibid., x.

15 See Heilbron (1979).

16 Cf. C. H. Pfaff, 'Elektricität', in Gehler, *Physikalisches Wörterbuch*, 3:233–389, on 254.

17 Cf. S. P. Thompson, 'The development of the mercurial air pump', *Telegraph Journal 21* (1887), 556–8, 587–90, 610–13, 632–4, 656–65.

18 J. B. Biot, *Traité de physique*, 4 vols. (Paris, 1816), 2:213–14, 254–5, 265, 291–4, 312–13; A. C. Becquerel, *Traité expérimental de l'électricité et du magnétisme*, 7 vols. (Paris, 1834–40), 2:4, 169, 185.

19 M. Faraday, *Experimental researches in electricity*, 3 vols. (London, 1839–55), 1:439.

20 R. J. Haüy, *Traité élémentaire de physique*, 2 vols. (Paris, 1803), 1:361–4. Cf. Becquerel, *Traité expérimental* 2:192–3.

21 Biot, *Traité*, 2:218, 242–3, 281, 316–18, 326. Cf. P. M. Roget, 'Electricity', in *Library of useful knowledge: natural philosophy*, vol. 2 (London, 1832), 12.

22 Canton to Priestley, 5 Apr. 1766, in J. Priestley, *The history and present state of electricity*, 3rd ed., 2 vols. (London, 1775), 1:305–6.

23 T. Cavallo, *A complete treatise on electricity*, 3rd ed., 3 vols. (London, 1786–95), 3:195–7; T. Milner, *Experiments and observation in electricity* (London, 1783), 91–9. Cf. A. Bennet, *New experiments on electricity* (Derby, 1789), x–xi; G. Adams, *An essay on electricity*, 2nd ed. (London, 1785), 56, 76–80.

24 Milner, *Experiments*, 99.

25 Cavallo, *Treatise*, passage quoted from 3:192–5; T. Cavallo, *Elements of natural or experimental philosophy*, 4 vols. (London, 1803), 3:359, 404–5; adapted by J. J. Prechtl as 'Untersuchungen über die Modification des elektrischen Ladungszustandes', *Annalen der Physik (Gilberts Annalen, Poggendorffs Annalen) 35* (1810), 28–104, on 46 ff., and by M. van Marum as 'On the theory of Franklin', *Annals of Philosophy* 16 (1820), 441–53.

26 F. U. T. Aepinus, *Tentamen electricitatis et magnetismi* (St. Petersburg, 1759), 192; T. Young, *Course of lectures on natural philosophy and the mechanical arts*, 2 vols. (London, 1807), 1:664–5; A. Libes, *Histoire philosophique des progrès de la physique*, 4 vols. (Paris, 1810–13), 3:201, 205; Cavallo, *Elements*, 3:353; Roget, 'Electricity', 10, 29; Pfaff, 'Elektricität', 301–4.

27 Cf. G. Miller, 'Observations on the theory of electrical attraction and repulsion', *Transactions of the Royal Irish Academy 7* (1800), 139–50; 'Observations on [van Marum's] memoir', *Annals of Philosophy 1* (1821), 181–6.

28 E.g., G. J. Singer, *Elements of electricity and electrochemistry* (London, 1814), 24, 55–7, 65, 78–9; 'On the supposed repulsion of electricity', *Philosophical Magazine 49* (1817), 208–9; J. A. Deluc, *Traité élémentaire sur le fluide électrogalvanique*, 2 vols. (Paris, 1804), 1:55–9; J. B. van Mons, 'On . . . galvanism and electricity', *Journal of Natural Philosophy, Chemistry and the Arts (Nicholson's Journal) 24* (1809), 179–89; Prechtl, 'Untersuchungen', 46–50.

29 Cf. F. A. C. Gren, *Grundriss der Naturlehre* (Halle, 1820), §§1114, 1125, 1199, 1204–5, 1212, 1400; J. C. Fischer, *Anfangsgründe der Physik* (Jena, 1797), 650–3; E. G. Fischer, *Physique mécanique* (Paris, 1806), 263–9.

30 A. Avogadro, 'Considérations sur l'état . . . d'un corps non-conducteur', *Journal de Physique 63* (1806), 450–62, and *65* (1807), 130–45. Cf. M. Gliozzi, *Fisici piemontesi* (Turin, 1962), 14.

31 G. B. Beccaria, 'De atmosphaera electrica', *Philosophical Transactions 60* (1770), 277–301; G. B. Beccaria, *A treatise upon artificial electricity* (London, 1776), 186, 218, 220, 384–7. Cf. M. Gliozzi, 'Giambatista Beccaria nella storia dell'elettricità', *Archeion 17* (1935), 15–47, on 41–2.

32 Heilbron (1979), 409–12.

33 *Journal of Natural Philosophy, Chemistry and the Arts (Nicholson's Journal) 21* (1808), 278–90; *Journal für Chemie und Physique (Schweiggers Journal), 6* (1808), 53–83. Cf. J. J. Prechtl, 'Einige Bemerkungen zu Herrn Avogadros Abhandlung', ibid., 84–115; Prechtl, 'Untersuchungen', 30 ff.

34 A. Avogadro, 'Note sur la nature de la charge électrique', *Archives de l'Electricité 2* (1842), 102–10. Cf. Becquerel, *Traité expérimental*, 2:178–9; the editor's remarks in A. Avogadro, *Opere scelte* (Turin, 1911), 375–7.

35 J. B. Biot, 'Sur le mouvement du fluide galvanique', *Journal de Physique 53* (1801), 224–74; T. M. Brown, 'The electric current in early 19th-century French physics', *Historical Studies in the Physical Sciences 1* (1969), 61–103, on 72–5; M. L. Shagrin, 'Resistance to Ohm's law', *American Journal of Physics 31* (1963), 536–47.

36 J. Bostock, *An account of the history and present state of galvanism* (London, 1818), 100–50; W. Henry, 'On the theories of the excitement of galvanic electricity', *Memoirs of the Manchester Literary and Philosophical Society 2* (1813), 293–312; Singer, *Elements*, 378–85.

37 H. Davy, 'On some chemical agencies of electricity', *Philosophical Transactions 97* (1807), 1–56, on 45–7; H. Davy, *Elements of chemical philosophy*, vol. 1 (London, 1812), 168–70. Cf. J. W. Pfaff, *Der Voltaismus* (Stuttgart, 1803), 69; C. A. Russell, 'The electrochemical theory of Sir Humphry Davy', *Annals of Science 15* (1959), 1–13, 15–25, and *19* (1963), 255–71, on 6–9.

38 Davy, 'Some chemical agencies', 29–30; Davy, *Elements*, 170.

39 According to a notebook dating from *c.* 1810 (Russell, 'Electrochemical theory', *19*, 262). Cf. Pfaff, *Voltaismus*, 35, 40; Bostock, *Account*, 141–2.

40 T. von Grotthuss, 'Mémoire sur la décomposition de l'eau . . .', *Annales de Chimie et de Physique 58* (1806), 54–74. Cf. Faraday, *Researches*, 1:142–3; E. Hoppe, *Geschichte der Elektrizität* (Leipzig, 1884), 278–9.

41 J. G. Children, 'An account of some experiments with a large Voltaic battery', *Philosophical Transactions 105* (1815), 363–74.

42 J. T. Mayer, *Anfangsgründe der Naturlehre*, 3rd ed. (Göttingen, 1812), 476–82; Whittaker (1910), 78–80; Pfaff, 'Elektricität', 350–89.

43 Cf. A. Bennet, 'A new suspension of the magnetic needle . . .', *Philosophical Transactions 82* (1972), 81–98, on 87–8; *Encyclopaedia Britannica*, 3rd ed., 1:216.

44 E.g., Young, *Course*, 1:668–9, 684.

45 J. G. Children, 'An account of some experiments . . .', *Philosophical Transactions 99* (1809), 32–8. Cf. K. Meyer, 'The scientific life and works of H. C. Ørsted', in H. C. Ørsted, *Naturvidenskabelige skrifter*, ed. K. Meyer, 3 vols. (Copenhagen, 1920), 1:xiii–clxvi, on lv; J. F. Demonferrand, *Manuel d'électricité galvanique* (Paris, 1823), 7.

46 A. M. Ampère, in Société française de Physique, *Collection de mémoires relatifs à la physique*, pts. 2–3, *Mémoires sur l'électrodynamique*, 2 vols. (Paris, 1885–7), 1:192 n. Hereafter *Mémoires sur l'électrodynamique*. Cf. A. M. Ampère, in ibid., 1:6–10, 12, 216; A. M. Ampère, letter of 21 Feb. 1821, in A. M. Ampère, *Correspondance du grand Ampère*, ed. L. de Launay, 3 vols. (Paris, 1936–43), 2:566. Similar definitions may be found in Prechtl, 'Untersuchungen', 30–46; Demonferrand, *Manuel*, pp. 5–7; P. Prevost, 'Tentative faite dans le but de concilier deux principes fondementaux de la théorie d'électricité', *Bibliothèque Universelle 21*

(1822), 178–88, on 185; J. D. Colladon, 'Déviation de l'aiguille aimantée par le courant', *Annales de Chimie et de Physique 33* (1826), 62–75; A. de la Rive, 'Recherches sur la cause d'électricité voltaïque', ibid. *39* (1828), 297–324, on 304–5; Becquerel, *Traité expérimental*, 2:3, 214–16.

47 H. Davy, *A syllabus of a course of lectures* (1802), in H. Davy, *Collected works*, ed. J. Davy, 9 vols. (London, 1839–40), 2:327–436, on 390–1. Cf. A. C. Becquerel, *Traité de physique*, 2 vols. (Paris, 1842–4), 1:53–4.

48 E.g., Becquerel, *Traité expérimental*, 2:215–16. Cf. Gren, *Grundriss*, §§1403–4.

49 Biot, quoted in Pfaff, 'Elektricität', 380–1. Cf. Young, *Course*, 1:670–1; T. Young, 'Outlines of experiments and inquiries respecting sound and light', *Philosophical Transactions 90* (1800), 106–50, on 125–6.

50 Ørsted, *Skrifter*, 1:267–73. Cf. Meyer, 'Scientific life', xxv–xxxvi, cxiv; R. C. Stauffer, 'Speculation and experiment in the background to Ørsted's discovery of electromagnetism', *Isis 48* (1957), 33–50.

51 H. C. Ørsted, *Ansischt der chemischen Naturgesetze* (1812), in Ørsted, *Skrifter*, 2:90–100.

52 Ibid., 40–1, 110, 132–5; H. C. Ørsted, 'Theorie over lyset' (1815–16), in Ørsted, *Skrifter*, 2:433–5.

53 As this proposition conflicted with the usual explanation of electric light *in vacuo*, Ørsted hinted that vacuum might insulate. *Skrifter*, 2:126–7.

54 Skrifter, 2:435. Cf. T. Thomson, review of Ørsted's *Recherches* (1813), *Annals of Philosophy 13* (1819), 368–77, 456–63, and *14* (1819), 47–50 and esp. 459–63.

55 Thomson, review of Ørsted's *Recherches*, 369; F. W. J. von Schelling, 'Ueber Faradays neueste Entdeckung', in F. W. J. von Schelling, *Sämmtliche Werke*, 14 vols. (Stuttgart and Augsburg, 1856–61), 9:439–52, quoted passage on 443–4. Cf. Bostock, *Account*, 102; W. Whewell, *History of the inductive sciences*, 3rd ed., 2 vols. (New York, 1859), 2:243; Hoppe, *Geschichte*, 200, 222–3.

56 *Ansicht*, in Ørsted, *Skrifter*, 2:147–8.

57 Ørsted, *Skrifter*, 2:223–5, 356–8; J. S. C. Schweigger, 'Ueber Elektromagnetismus', *Journal für Chemie und Physik (Schweiggers Journal) 46* (1826), 1–72, and *48* (1826), 289–352, on 10–14; Meyer, 'Scientific life', xc, xcvii–xcviii; Stauffer, 'Speculation', 48–50; J. P. Gérard, 'Sur quelques problèmes concernant l'oeuvre d'Ørsted en électromagnétisme', *Revue d'Histoire des Sciences 14* (1961), 297–312.

58 H. C. Ørsted, *Experimenta* (1820), in Ørsted, *Skrifter*, 2:214–18; translated in *Annals of Philosophy 16*, 473–6, *Journal de Physique 91*, 72–80, *Annales de Chimie et de Physique 14*, 417–25, *Annalen der Physik (Gilberts Annalen, Poggendorffs Annalen) 66*, 295–304, *Journal für Chemie und Physique (Schweiggers Journal) 29*, 275–81, all in 1820.

59 *Skrifter*, 1:cxiii.

60 Ørsted had eschewed ordinary pictures such as atoms and fluids (*Ansicht*, 149–57), but some of his concepts, such as 'quantities' of forces, might best be understood on a mechanical representation, which he deemed to be compatible with his general theory (ibid., 129–30, 136). He occasionally used one himself. Meyer, 'Scientific life', lviii–lx.

61 E.g., M. Faraday, 'Historical sketch of electromagnetism', *Annals of Philosophy 18* (1821), 195–200, 274–90, and *19* (1822), 107–17, on 107–8; C. H. Pfaff, *Der Elektro-Magnetismus: eine historisch-kritische Darstellung der bisherigen Entdeckungen* (Hamburg, 1824), 206–9.

62 A similar classification occurs in G. T. Fechner, *Elementar-Lehrbuch des Elektromagnetismus* (Leipzig, 1830), 127. Whewell, *History*, 2:247, conflates the first two types.

63 The distinction between suppositious elementary *force* and measurable force is elaborated in Heilbron (1979), 449, 475–7.

64 A. M. Ampère, 'Suite du mémoire sur l'action mutuelle des deux courans', *Annales de Chimie et de Physique 15* (1820), 170–218, on 178.

65 [J. B. Biot and F. Savart], 'Note sur le magnétisme de la pile de Volta', *Annales de Chimie et de Physique 15* (1820), 222–3; J. B. Biot, *Précis* (1823), in *Mémoires sur l'électrodynamique*, 1:80–127.

66 Ampère, in *Mémoires sur l'électrodynamique*, 1:20. Initially, Ampère hesitated between a collection of elementary loops and a current circling the magnet. Dumonferrand, *Manuel*, 117–19; Hoppe, *Geschichte*, 396; Fresnel, in *Mémoires sur l'électrodynamique*, 1:141–7, 219–23; Ampère, letters of 18 Dec. 1820, 25 March 1821, in Ampère, *Correspondance*, 2:563, 568.

67 Ampère to his son, 4 Feb. 1823, in Ampère, *Correspondance*, 2:624. Cf. Ampère to A. de la Rive, 23 Mar., 21 Aug. 1823, and to Faraday, 18 Apr. 1823, in ibid., 628, 630, 635; Demonferrand, *Manuel*, 65–8.

68 Demonferrand, *Manuel*, 43–5. Cf. Fechner, *Lehrbuch*, 20–2, 102; Ampère, in *Mémoires sur l'électrodynamique*, 1:26, 185.

69 C. H. Pfaff, *Elektro-Magnetismus*, 245–51, 268–9; Fechner, *Lehrbuch*, 129–31.

70 Ampère to A. de la Rive, Oct. 1822, in Ampère, *Correspondance*, 2:605. Ampère insisted on this point: e.g., ibid., 570, 646–7, 653; *Mémoires, sur l'électrodynamique*, 1:27, 184, 187, 403–4. Cf. Pfaff, *Electro-Magnetismus*, 254–5, 261–2.

71 Cf. K. L. Caneva, 'From Galvanism to electrodynamics: the transformation of German physics in its social context', *Historical Studies in the Physical Sciences 9* (1978), 63–159, on 78–83, 91; Pfaff, *Electro-Magnetismus*, 254–5, 261–2.

72 Wollaston, quoted in Faraday, 'Sketch', 110. Cf. D. H. Conybeare to Faraday, 4 Apr. 1823, in M. Faraday, *The selected correspondence of Michael Faraday*, ed. L. P. Williams, 2 vols. (Cambridge, 1971), 1:142.

73 Respectively, H. Davy, 'On the magnetic phaenomena produced by electricity', *Philosophical Transactions 111* (1821), 7–19, 425–39, on 14; and M. Faraday, 'On some new electro-magnetical motions, and on the theory of magnetism', *Quarterly Journal of Science 12* (1822), 74–96, on 79; reprinted in Faraday, *Researches*, 2:127–47.

74 Davy, 'Magnetic phaenomena', 14, 17, 427; Faraday, 'Sketch', 198–9; Faraday, *Researches*, 2:161–2; M. Faraday, *Diary*, ed. T. Martin, 7 vols. (London, 1932–6), 1:49; M. Faraday, 'A course of lectures on the philosophy and practice of chemical manipulation' (1827), MS no. 13 at the Royal Institution, 54.

75 Faraday, 'Sketch', 109–10; Faraday to Marcet, 15 Jan. 1822, and to Ampère, 3 Sept. 1822, in Faraday, *Correspondence*, 1:129, 134–5; Ampère to Faraday, 10 July 1822, and to A. de la Rive, Oct. 1822, 23 Mar. 1823, in Ampère, *Correspondance*, 2:590, 606, 629–30; *Mémoires sur l'électrodynamique*, 1:243; Demonferrand, *Manuel*, 190–2; P. Barlow, *An essay on magnetic attractions*, 2nd ed. (London, 1824), 230; J. T. Seebeck, 'Ueber den Magnetismus der Galvanischen Kette', *Abhandlungen der Akademie der Wissenschaften, Berlin* (1820/1), 289–346, on 295–6. Cf. Fechner, *Lehrbuch*, 128; Hoppe, *Geschichte*, 220.

76 A. M. Ampère, 'Exposé sommaire' (1822), in *Mémoires sur l'électrodynamique*, 1:192–204. Cf. Ampère, *Correspondance*, 2:576, 579, 583, 605.

77 The objections, respectively, of Schweigger, 'Ueber Elektromagnetismus', 17–18, and Muncke as reported by Fechner, *Lehrbuch*, 116–18; and of Pfaff, *Elektro-Magnetismus*, 245–6, and Faraday, letters to Marcet, 15 Jan. 1822, in Faraday, *Correspondence*, 1:130, and to G. de la Rive, 12 Sept. 1821, in Williams, *Faraday*, 126. Cf. Davy to Ampère, 26 May 1821, in S. Ross, 'The search for electromagnetic

induction, 1820–1831', *Notes and Records of the Royal Society of London 20* (1965), 184–219, on 197: 'I wish you may be able to furnish some *direct* proof of the existence of electrical currents in the magnet'.

78 Pfaff, *Elektro-Magnetismus*, 261–3; G. F. Pohl, 'Versuch über die Einwirkung des Erd-Magnetismus . . . ,' *Annalen der Physik (Gilberts Annalen, Poggendorffs Annalen) 74* (1823), 389–409, and *75* (1823), 269–322; Barlow, *Essay*, 232–3.

79 Barlow, *Essay*, passage quoted from 255–6; Demonferrand, *Manuel*, 193–4.

80 Respectively, Whewell, *History*, 2:255; and Schweigger, 'Ueber Elektromagnetismus', 58–9, 290 n. Cf. Pfaff, *Elektro-Magnetismus*, 65; Caneva, 'Galvanism', 93–4.

81 This mnemonic is Ampère's: see *Mémoires sur l'électrodynamique*, 1:12; Ampère, 'Suite du mémoire', 203; Hoppe, *Geschichte*, 206–7. It became standard.

82 Schweigger, 'Ueber Elektromagnetismus', 63.

83 J. S. C. Schweigger, 'Ueber die elektrische Erscheinung, welche die Alten mit dem Namen Kastor und Pollux bezeichneten', *Journal für Chimie und Physik (Schweiggers Journal) 37* (1823), 245–342, on 252, 262, 268, 276, 320–1; Schweigger, 'Ueber Elektromagnetismus', 42–3, 66–7.

84 Seebeck, 'Magnetismus', 297–8. Cf. H. Schimank, 'Die Entdeckung der elektromagnetischen Induktion', *Beiträge zur Geschichte der Technik und Industrie 21* (1932), 1–11, on 10 n.

85 Davy, 'Magnetic phaenomena', 11, 14; Faraday, *Diary*, 1:57; Faraday, 'Sketch', 276, 283; Faraday, 'New motions', 88–9.

86 P. M. Roget, 'Magnetism', in *Library of useful knowledge: natural philosophy*, vol. 2 (London, 1832), 19–22; P. M. Roget, 'On the geometrical properties of the magnetic curve', *Journal of the Royal Institution 1* (1830/1), 311.

87 E.g., Faraday, 'New motions', plate 3, figs. 6–8, 11, 14; Faraday, *Diary*, 1:53.

88 Pfaff, *Elektro-Magnetismus*, 201.

89 P. Erman, *Umrisse zu den physikalischen Verhältnisse des von Herrn Oersted entdeckten elektrochemischen Magnetismus* (Berlin, 1821), 53, 58–63; Seebeck, 'Magnetismus', 298; Faraday, 'Sketch', 117.

90 Schweigger, 'Ueber Elektromagnetismus', 3, 54–5, 316–17.

91 A. C. Becquerel, 'Des effets électriques qui se développent pendant diverses actions chimiques', *Annales de Chimie et de Physique 23* (1823), 244–58, quoted passage on 245; P. M. Roget, 'Electromagnetism', in *Library of useful knowledge: natural philosophy*, vol. 2 (London, 1832), 80. Cf. Faraday, 'A course', 55: 'My rotatory apparatus is a striking illustration of the vertiginous nature of the power in the wire'.

92 Ampère to Faraday, 1825, in Ampère, *Correspondance*, 2:674–5. Cf. *Mémoires sur l'électrodynamique*, 1:250.

93 Ampère to A. de la Rive, 2 July 1824, 5 Aug. 1826, in Ampère, *Correspondance*, 2:215–19, 249–50. Cf. Becquerel, *Traité expérimental*, 2:3.

94 Ampère, 'Réponse à . . . M. van Beck' (1821), in *Mémoires sur l'électrodynamique*, 1:216.

95 Becquerel, 'Effets électriques', 248–9; review of S. Marianini, *Saggio di esperienze elettrometriche*, *Annales de Chimie et de Physique 33* (1826), 113–54, on 120; A. de la Rive, 'Recherches', 323; L. Nobili, 'Sur une nouvelle classe de phénomènes électrochimiques', *Annales de Chimie et de Physique 34* (1827), 280–92, 409–38, on 292 (cf. Hoppe, *Geschichte*, 249–50); F. Savary, 'Mémoire sur l'aimantation', *Annales de Chimie et de Physique 34* (1827), 5–57, 220–1, on 55–6, 220–1; E. Becquerel, according to A. Becquerel, *Traité de physique*, 1:71. Cf. Roget, 'Electromagnetism', 57.

96 Becquerel, *Traité de physique*, 1:44. Cf. Heilbron (1979), 223–6, 256–7, 271–2, 293.

97 Erman, *Umrisse*, 92; Pohl, 'Beiträge zur näheren Kenntnis des Elektromagnetismus', *Isis* (1820), 390–409, on 393–5; C. Hansteen, 'Zusätze und Berichtigungen zu den Bemerkungen über Polarlichter und Polarnebel', *Journal für Chemie und Physik (Schweiggers Journal) 48* (1826), 360–73, on 366–8.

98 'Esquisse historique des principales découvertes faites dans l'électricité depuis quelques années', *Bibliothèque Universelle 52* (1833), 225–64, 404–47, and *53* (1833), 70–125, 170–227, 315–51, on 225. Cf. Becquerel, *Traité expérimental*, 1:183, and 2:215.

99 Faraday, 'Sketch', 281; Fresnel, in *Mémoires sur l'électrodynamique*, 1:76. Cf. Ross, 'The search', 185–90; Faraday, *Researches*, 1:24.

100 A. de la Rive, in *Mémoires sur l'électrodynamique*, 1:328; Ampère, in ibid., 1:332–4; and Ampère, letter to G. de la Rive, 25 Sept. 1822, in Ampère, *Correspondance*, 2:602. Cf. ibid., 2:744, 760–7, 774. Ampère's disk should have jumped at make and break of the current in the coil; perhaps it also oscillated under the torsion in the wire of suspension.

101 Ampère, in *Mémoires sur l'électrodynamique*, 2:169, 173. Cf. K. R. Gardiner and D. L. Gardiner, 'André-Marie Ampère and his English acquaintances', *British Journal for the History of Science 2* (1965), 235–45; Ross, 'The search', 193–5.

102 C. Babbage and J. F. W. Herschel, 'Account of the repetition of M. Arago's experiments', *Philosophical Transactions 115* (1825), 467–96, on 471–4, 485–90; S. H. Christie, 'On the magnetism developed in copper and other substances during rotation', ibid., 497–509; P. Barlow to Faraday, 4 May 1825, in Faraday, *Correspondence*, 1:148–9; *Quarterly Journal of Science 20* (1825), 385–6.

103 M. Arago, 'Note concernant les phénomènes magnétiques auxquels le mouvement donne naissance', *Annales de Chimie et de Physique 32* (1826), 213–23, on 216–18; Roget, 'Magnetism', 91–6; A. de la Rive, 'Esquisse', 438–41.

104 M. Faraday, 'Friday evenings, 1825–9', MS no. 12 at the Royal Institution, 207 (1825), and 227 (26 Jan. 1827), respectively.

105 Faraday, 'A course', 57–8; and Faraday, 'Friday evenings', 227 (26 Jan. 1827), respectively.

106 Hoppe, *Geschichte*, 397–401.

107 Faraday, *Diary*, 1:178, 279, 310. Cf. T. Martin, *Faraday's discovery of electromagnetic induction* (London, 1949), 45–51.

108 Faraday, *Diary*, 1:367 (29 Aug. 1831); Faraday, *Researches*, 1:7–9; Martin, *Faraday's discovery*, 52–6. The connection of ideas is obscured in Faraday, *Researches*, 1:2–4, which introduces the discovery via an experiment with a *wooden*, not an iron, ring. Cf. Williams, *Faraday*, 177–83.

109 Faraday, *Diary*, 1:369 (30 Aug. 1831); Martin, *Faraday's discovery*, 57. The electrotonic state was therefore a second thought; an analysis of this concept will be given in the paper mentioned in n. 113.

110 The note is quoted in full in Williams, *Faraday*, 181, and in Faraday, *Correspondence*, 1:217. Cf. J. Larmor, 'Faraday on electromagnetic propagation', *Nature 141* (1938), 36.

111 H. Bence-Jones, *The life and letters of Faraday*, 2 vols. (London, 1870), 1:11; [J. Tytler], 'Electricity', *Encyclopaedia Britannica*, 3rd ed., 6:418–545, on plate 177, fig. 75, and pp. 443–5, 452, 455–61.

112 Singer, *Elements*, v, 56–8, 66–8.

113 In a paper in preparation, I show that Faraday's suspicion of material electrical currents, his recurrent positivism, and his usage of the terms *power* and *state* also had roots in contemporary physics. There is no reason to posit that Kant or Boscovich or the Sandemanians gave Faraday his basic concepts of electricity and magnetism.

7

The quantitative ether in the first half of the nineteenth century

JED Z. BUCHWALD

Institute for the History and Philosophy of Science and Technology, University of Toronto, Toronto, Canada M5S 1A1

Throughout the nineteenth century ethers played a major role in many physical theories, particularly the wave theory of light, since, it was thought, light waves travel in some ethereal medium. Between 1830 and 1890 wave optics was primarily concerned with the mathematical laws that govern ether waves. Much effort was expended on attempts to analyse various quantitative physical models for ether in order to determine its laws of motion. The first such model was developed by Augustin Fresnel in the early 1820s; by the mid 1830s it was being carefully investigated. The results of the investigations were frequently useful: Several optical phenomena that had previously escaped explanation were accounted for. However, difficulties also arose with the model, and one in particular led to important changes in it. In this chapter, I shall discuss the origins of this first theory of the optical ether and examine the broad outlines of its subsequent development. First it is essential to recall briefly the major developments in optics during the late seventeenth and the eighteenth centuries.

During the seventeenth century the hypothesis that light was a spherical pulse propagating at a finite rate in a universally present medium was created and developed mathematically, culminating in the *Traité de la lumière* (Leiden, 1690) of Christiaan Huygens. Huygens demonstrated that the common facts of refraction – Snell's law and total internal reflection – could be deduced by assuming that the pulse travelled more slowly in the refracting body than in the bounding medium. Moreover, Huygens was also able to explain the peculiar 'double refraction' of Iceland crystal[1] by using his principle of secondary waves and by assuming that, in the crystal, light was propagated in both spheroidal and spherical pulses.[2] What Huygens's theory lacked was a strong dynamical foundation; his explanation of how the universal medium propagated light was not quantitative – he could not show mathematically that the characteristics of the medium implied the kinds of motion that, his theory supposed, constituted light.[3]

In the 1670s, Isaac Newton created a quite different theory of light that agreed with Huygens's only in supposing that light takes time to travel (not everyone believed that it did until after 1728).[4] Using experiments on prismatic refraction, Newton argued, in the first instance, that light was not a form of motion, whether actual or potential, but had properties much more like those possessed by matter. White light, he argued, consisted of a large number of distinct kinds of light, each of which corresponded to a spectral colour. The passage of white light through a prism separated its component kinds, which could not thereafter be further decomposed. Each kind of light obeyed Snell's law of refraction, but with a unique index of refraction that increased from the red to the blue end of the spectrum.[5]

According to Newton each kind of light consisted of a single type of material 'corpuscle', and he explained refraction by an attractive force exerted between the refracting medium and the corpuscle.[6] Consequently, in Newton's theory the speed of light had to be greater in the refracting than in the bounding medium – a requirement that, of course, conflicted with the requirements of the pulse theory. Newton modified his theory in later years in order to explain partial reflection, the colours of thin films, and the related phenomenon of 'Newton's rings', and he then introduced, in addition to the corpuscles, a universal medium in which pulses were propagated and which acted dynamically upon the optical corpuscles.[7] However, Newton continued to identify light with the corpuscles and not with the pulses.

Comparing the two theories at the end of the seventeenth century, we find that they differed, in the first instance, in the velocity condition: Light moved faster in refracting media according to Newton; it moved more slowly according to Huygens. Neither Newton nor Huygens (who did not consider colours) provided a mathematical, deduced formula for dispersion. Newton, however, could explain the phenomenon, and, perhaps less convincingly, he could also explain the colours of thin films and related reflection phenomena; Newton could also explain Grimaldi's phenomenon, or what we would now call diffraction.[8] Both Newton and Huygens provided formulae for double refraction, but only Huygens's was deduced (from his hypothesis that light spreads in spheroidal, as well as in spherical, pulses in Iceland crystal). However, neither formula was widely accepted throughout most of the eighteenth century. Finally, whereas Newton could explain the polarisation effects associated with double refraction, Huygens could not.[9]

Corpuscular optics in the early nineteenth century

By the middle of the eighteenth century many of Newton's ideas, and other ideas that, though not his, were developed by people who thought of themselves as his disciples, were widely accepted on the Continent as well as in Britain. Corpuscular optics in particular achieved widespread assent, and Huygens's pulse theory was largely disregarded.[10] The powerful impact of Newtonian concepts led, during the Napoleonic era in France, to a programme of physical research that was institutionally organised and carried on by men who thought of themselves as professional scientists.[11] The guiding figure of the programme, Pierre Simon de Laplace, investigated several phenomena within a sharply defined theoretical context. Laplace and his followers assumed that electricity, magnetism, heat and light, as well as matter, consisted of particles between which centrally directed forces acted at a distance. They frequently referred to the particles as 'molecules', which meant simple, uncompounded units. Although they rarely discussed the nature of the molecule, for the most part they treated it as a material 'point' that, although it possessed mass and exerted force, was not extended. The molecules could, however, exist in stable groups, thereby composing a complex structure.[12] Rays of light, according to the Laplacians, consisted of particles that could be acted on by matter through central, short-range forces.

Laplace used this assumption to deduce a formula for atmospheric refraction by calculating the action that the atmosphere exerted on the optical corpuscle.[13] Then, between 1807 and 1811, Laplace and his pupil Etienne Louis Malus were able to show that Huygens's laws of double refraction (which were empirically confirmed by Malus) were consistent with corpuscular optics.[14] They based their demonstration on the principle of least action, which, they believed, applied uniquely to systems governed by forces acting at a distance.[15] In effect, Laplace and Malus assumed an expression for the velocity of the extraordinarily refracted corpuscle; they then demonstrated that, when this expression was substituted into the principle of least action, Huygens's laws resulted. Consequently, by 1811 corpuscular optics had already achieved a high degree of success in accounting for two phenomena that the pulse theory had never treated in detail (specifically, atmospheric refraction) or that had, until recently, been its sole province (double refraction).

During this period corpuscular optics achieved further successes after Malus discovered that polarisation could be produced by reflection, as well as by passing light through a doubly refracting crystal.[16] It was also observed that regular, coloured images were produced when light passed through a thin plate of crystal and then through a double refractor (chromatic polarisation).[17] No one was able to deduce a mathematical explanation of these phenomena from corpuscular principles in the way that Huygens's laws for double refraction could be deduced from them. Malus and another participant in this programme of research, Jean Baptiste Biot, were able, however, to provide convincing qualitative explanations and to predict certain phenomena by using experimentally justified (though not theoretically deduced) laws.[18] The foundation of their explanation was, of course, the optical corpuscle, or, rather, the complex grouping of material points that had replaced the corpuscle by about 1812.[19]

Thomas Young's wave optics

At about the time that the Laplacian programme took shape (*c.* 1800), but shortly before its first major successes in optics, Thomas Young revived Huygens's theory in England,[20] and he added to it the requirement that, like sound, light consisted of a series of regularly spaced troughs and peaks, or waves, which could therefore interfere with one another; the peak to peak, or trough to trough, distance – the wavelength – determined the colour. Aided by his principle of interference, Young was able to explain several diffraction phenomena. Moreover, he used results of eighteenth-century experimental work on sound to claim – without mathematical demonstration – that the waves of light would not diverge more than was necessary to explain diffraction.[21] However, like Huygens, Young did not develop a quantitative account of the properties of the medium in which light propagated, though he frequently speculated upon the nature of the medium (Cantor, 1970). Moreover, again like Huygens, Young initially ignored polarisation since he could not see how to introduce an asymmetry about the axis of a ray. In both France and Britain there was no compelling reason, especially after 1811, to pay close attention to Young's revival and development of the medium theory. Only a quantitative account of diffraction escaped corpuscular optics but not wave optics; however, as far as was then known, this phenomenon was nevertheless consistent with corpuscular hypotheses.[22]

Towards the end of the Napoleonic era (*c.* 1814), then, most French scientists would have had little reason to consider the wave theory a compelling alternative to corpuscular optics. If, however, we look dispassionately at the rival claims of the two theories at this time, we can see that, if quantitative prediction were our sole criterion, there would have been little to choose between them. Both could yield Huygens's construction for double refraction; both also yielded Snell's law for ordinary refraction. Both could be founded on a basic principle (least action for corpuscular optics and least time for wave optics).[23] The wave theory did provide a mathematics – crude though it then was – for diffraction,[24] but the corpuscular theory could at least explain the

phenomenon, and not enough was known empirically about diffraction to compel consideration of the wave theory's mathematics in this area. Physically, moreover, the corpuscular theory was unquestionably superior to the wave theory because it could explain polarisation, and it therefore at least held out the hope that it might one day be possible to deduce the laws of polarisation from corpuscular principles. And the physical hypotheses of corpuscular optics – material points with forces acting at a distance between them – were widely accepted in both France and Britain at the time, whereas the concept of a universal medium was rarely used in France, although it was occasionally discussed in Britain. Clearly the wave theory could only begin to command serious attention as a rival to corpuscular optics if, at the least, it could deal with polarisation and could provide a clear physical alternative to the optical corpuscle.

Augustin Fresnel and ether

In 1814 a young French civil engineer wrote to his brother concerning his increasing doubts about the received corpuscular theories of both light and heat. At the time, Augustin Fresnel had only a sketchy knowledge of contemporary optics, but, for reasons that remain unclear, he was disposed to believe that both light and heat were somehow connected with the 'vibrations of a special fluid which extended throughout space'.[25] Within little more than a year, Fresnel, independently of Young, had explained diffraction by means of the principle of interference. He also demonstrated experimentally that the pattern of fringes produced by a rectangular diffractor was inconsistent with what could reasonably be expected on corpuscular grounds, but that the pattern was well explained by the principle of interference. Fresnel soon strengthened his critique of corpuscular optics with an experiment in which interference occurred without the presence of a material body, which would have to be present according to the corpuscular theory in order to deflect the rays of light.[26]

Before 1818, however, Fresnel, like Young, had not provided for wave optics a sophisticated mathematical structure that permitted a completely general calculation of interference patterns.[27] This he accomplished by the end of 1818,[28] and, as is well known, he submitted his paper for the prize competition the Paris Académie des Sciences had announced for a mathematical account of diffraction. Fresnel's memoir won the prize, but it succeeded in convincing none of the judges that the wave theory should be accepted, despite the fact that the Laplacian programme of research had never itself provided a mathematical account of diffraction.

There were, no doubt, many reasons for the feeble impact of Fresnel's mathematically sophisticated treatment of diffraction on most of his contemporaries, but two in particular were probably of especial importance. First, at this time, Fresnel had not published a theory of polarisation, and the old objection to the wave theory in this area – that waves had to be symmetrical about the ray – still stood. Second, Fresnel's memoir was only superficially founded on the dynamics of the wave-propagating medium. In his prizewinning memoir on diffraction, Fresnel used an expression for the force on an ether particle in order to calculate the wave velocity.[29] However, he did not at this time deduce the expression from ether's mechanical properties; he simply assumed it. Indeed, Fresnel did not even describe ether's properties in detail. Consequently, he had not as yet provided a firm physical alternative to the optical corpuscle. Nevertheless, he had been working on polarisation since 1816, and his investigations here inevitably led him to the question of the dynamical properties of the medium.

Fresnel would probably not have penetrated very far into the vexing question of ether dynamics had it not been for the problems posed by polarisation, for his theory of diffraction accounted quite well for the phenomenon of diffraction without utilising ether dynamics. Certainly the existence of a dynamics for ether led, in the 1830s, to a thriving programme of theoretical research, and perhaps furthered the acceptance of the wave theory among those who would otherwise have refused to choose between it and corpuscular optics. Once the dynamics was available, lines of mathematical research arose that led to important progress in explaining phenomena that the kinematic wave theory had not been able to treat. But Fresnel's deep interest in the dynamics of ether was due neither to a clairvoyant perception of events that occurred after his death nor to the demands of his unconvinced contemporaries. Rather, the invention of the new science of ether dynamics – whose purpose was to deduce optical phenomena mathematically from the mechanical properties of the medium – was a solitary affair that derived from questions plaguing no one but Fresnel.

The primary problem that bothered Fresnel was the connection between the polarisation of a ray and the law governing its motion in a doubly refracting crystal.[30] In mid-1816, Fresnel and François Arago, Fresnel's earliest supporter, had shown experimentally that light rays polarised in planes at right angles to one another do not mutually interfere, but that rays polarised in parallel planes do.[31] Using these facts, Fresnel explained chromatic polarisation through the principle of interference (though he did not actually produce a quantitative theory of it). Now, the interference properties of polarised light clearly could not be explained unless a polarised ray possessed some kind of

asymmetry about its axis, and that was incompatible with the natural assumption that light waves were longitudinal, on analogy with sound waves.

Fresnel had early (1817) perceived that if an unpolarised ray consisted of two vibratory components, one along the ray (longitudinal), and the other at right angles to it (transverse), then at least the basic facts of polarisation could be explained if the act of polarisation destroyed the longitudinal component.[32] However, he could not understand how such a process of selective destruction could occur mechanically. Nor did he then imagine that all light might be purely transverse, because he could no more understand how a mechanical process could generate only a transverse component than he could understand how it could destroy only a longitudinal one. For several years, therefore, Fresnel tried to construct polarised light out of a combination of longitudinal and transverse components, but he was in this way unable to develop a satisfactory combination. Although he was not explicit about the difficulties of the theory, the major problem was almost certainly the connection between polarisation and the paths of the rays in double refraction.

New ideas frequently contain many elements of the concepts that they were created to replace, and the physical hypothesis Fresnel developed for ether by 1821, which solved the problem of polarisation, was no exception to the rule. He had, of course, completely rejected the optical corpuscle, and perhaps the matters of heat and electricity as well, but he had not rejected material points (molecules) and central forces that act directly at a distance, because these concepts were among the most fundamental in the contemporary store of common notions. Consequently, he argued that ether itself consisted of molecules in the Laplacian sense, between which forces acted.[33] By adopting the additional (and debatable) assumption that two parallel lines of molecules can be readily separated laterally, but strongly resist mutual approach, Fresnel was able to uncouple the transverse and longitudinal vibrations, since the former would now travel much more slowly than the latter. Consequently, all light, polarised or unpolarised, was, Fresnel argued, purely transverse, and the problem of the destruction of the longitudinal component was avoided. Moreover – and here were the tangible fruits of the hypothesis – the velocity of propagation of a transverse wave clearly had to depend upon the intensity of the reaction force that a molecular displacement produced. If, then, the force of reaction depended upon the direction of the displacement, then the velocity of propagation would depend upon it as well. That, of course, was precisely what Fresnel needed to construct an explanation of the polarisation of doubly refracted rays, because the direction of vibration determined the direction of polarisation.

In this momentous 1821 memoir, Fresnel provided his first, essentially qualitative theory of polarisation and propagatory velocity in double refraction. But he did not explicitly deduce the theory from the properties of the molecular ether. Rather, he merely assumed that the force of reaction could depend upon the direction of displacement without demonstrating that the postulated dependence was consistent with the dynamical properties of the medium. During the next six months, as Fresnel progressed towards his most stunning mathematical achievement – the wave surface of the biaxial crystal[34] – he probably did not rely directly on the dynamics of ether. But, in his last memoir, he finally deduced the wave surface from ether dynamics, albeit not without a crucial failure. That deduction was among his most influential analyses, for it inaugurated a new era in physics, the era of quantitative ether dynamics.

Consider a single molecule of Fresnel's ether. If there are no incident waves, then the molecule is in equilibrium, and there is zero force acting. When the molecule is displaced, a restoring force is exerted. Now, if the ether were not homogeneous – that is, if the distribution of molecules differed from one region to another – then, for a displacement of given magnitude and direction, different molecules would experience different forces. Thus calculations that are valid for one molecule would not apply to another; hence the course of a wave could not be predicted. Consequently, Fresnel assumed homogeneity, and, to simplify the calculation further, he also imposed a condition of symmetry on the distribution of the molecules. One further assumption was necessary. To carry out the calculation he had to hold still every molecule but the one that was displaced. That, he admitted, was an incorrect assumption, since it would forbid waves: The displacement could not be propagated from particle to particle if the displacement of one did not move the others. However, the manner in which Fresnel had formulated his mathematics for double refraction required this assumption, and he accordingly used it. As a result he was able to deduce the wave surface in biaxial crystals from ether dynamics; the properties of uniaxial crystals were simply degenerate cases of biaxial properties.

Cauchy's ether dynamics

Fresnel's method inaugurated a programme of research that was continued after his death in 1827 by the mathematician Augustin Louis Cauchy. From the fundamental hypothesis of Fresnel's molecular ether, Cauchy was able to deduce a general equation of motion for any ether particle that, unlike Fresnel's equation of motion, did not presuppose an incorrect calculation of the reaction.[35] Cauchy, that is, allowed every particle to be displaced, and he calculated the net force on any given particle that resulted from its altered

distances from the remaining ether particles. The equation of motion he obtained contained summations over the set of remaining ether particles, and these summations involved the differences between the displacements of the given ether particle and the particles included in the summations. This equation, unfettered by special conditions, became the foundation of all subsequent work in the dynamics of the molecular ether.

However, precisely because Cauchy's equation was exactly derived, whereas Fresnel had used a limiting and incorrect assumption, Cauchy could not obtain wave motions as readily as could Fresnel, whose equation yielded wave velocities without requiring integration: Cauchy's equation was not so malleable. That was a difficulty for which Cauchy was particularly well prepared. Intimately familiar, unlike Fresnel, with linear partial differential equations, and rapidly becoming more familiar with them, Cauchy realised that he could transform his basic molecular equation into a differential equation of wave motion by expanding the differences in the displacements into a Taylor's series. Having done so, he then imposed restrictive conditions upon the coefficients in the resulting equations, these conditions being equivalent to arranging the particles of ether in certain symmetrical patterns. As a result he obtained differential equations of motion that yielded Fresnel's wave surface almost, but not quite, exactly.[36] Cauchy's expression for the wave surface differed from Fresnel's by terms that were much too small to detect experimentally. Consequently, his formula was as acceptable as Fresnel's. In effect, Cauchy not only had corrected Fresnel's erroneous deduction, but had also introduced a new mathematical element – the differential equation – into wave optics. Subsequently, the aim of most theoretical research in optics, whether based on an ether dynamics or not, was to discover the differential equation of wave motion obeyed by light passing through matter.

Why, one might ask, was it necessary for Cauchy to embrace the hypothesis of a molecular ether if (as was almost certainly the case) he was primarily interested in discovering and solving differential equations for light? Fresnel was directly concerned with the physical nature of ether, primarily because he could not solve the problem of polarisation without investigating this question; the sequence of his researches betrays a continual, if occasionally uncertain, preoccupation with ether dynamics. Cauchy, on the other hand, was more concerned with the mathematics of the wave theory than he was with ether's physical nature. But without the hypothesis of a molecular ether there would, at the time, simply have been no route at all to the mathematics. For, although the ultimate aim for Cauchy was always a mathematical proposition from which calculable phenomena could be deduced, this aim could only be achieved throughout most of the 1830s by deductions founded on a molecular ether. Moreover, there was then little to object to, because the hypothesis fitted so well into contemporary physical ideas; that is, it utilised the widely accepted concepts of material points and central forces. Cauchy's theory gave more than just the laws of double refraction; it also provided a formula for dispersion – something that had escaped corpuscular optics. By 1835, Cauchy had developed a unified theory of double refraction and dispersion in which both phenomena were explained by assigning the requisite values to the constant coefficients in the differential equations of motion.[37]

Dispersion had previously been a problem for the wave theory. Since light, like any wave motion, must propagate at a speed determined by the 'elasticity' and the 'density' of the medium, and since these were assumed to be fixed quantities for the ether in a given body, dispersion, which required the speed of light to depend upon the wavelength, was difficult to explain. George Biddell Airy was among the first after Fresnel to suggest an explanation, and he recurred to the Laplacian concept of adiabatic compression of a gas, which had the disquieting consequence of granting ether something like a capacity for latent heat.[38] Airy's ideas were, however, of only passing importance, and it was an earlier suggestion of Fresnel's that provided the elements of a quantitative theory.

If, Fresnel remarked in 1822,[39] the range of the molecular force were of the order of the longest visible wavelength, then the velocity of a wave would increase with the wavelength. Fresnel did not prove his claim, nor could he have done so convincingly given the limitations of his ether dynamics. Cauchy's theory of dispersion built upon Fresnel's suggestion by referring dispersion to the values of the constant coefficients in the differential equations of motion; these coefficients, of course, are functions of the molecular force. Thus in 1835, Cauchy took his basic molecular equations and, solving them, obtained an expression for the wave velocity without making approximations (he had made approximations in his theory of double refraction of the late 1820s). The expression contained the constant coefficients of the original differential equation (combined to form other constants), as well as a factor that was a function of the wavelength. In a complicated analysis, Cauchy expanded both the coefficients and the wavelength function into a Taylor's series, and he obtained the following series for the square of the index of refraction, n^2, with a wavelength λ:

$$\frac{1}{n^2}=A_1+\frac{A_2}{\lambda^2}+\frac{A_3}{\lambda^4}+\frac{A_4}{\lambda^6}+. \; . \; .$$

The constants A_i were sums that contained the masses of ether molecules, the force between them, and their mutual distances in a state of equilibrium.

Now, Cauchy's formula did not directly explain the cause of dispersion, because dispersion depended upon all of the A_i except A_1, and there were several physical conditions that determined their values. Adopting Fresnel's suggestion, however, Cauchy recurred to the problem of the range of the molecular force. He offered a demonstration (unfortunately incorrect) that, if the molecular force were repulsive and of the form $1/r^4$, then dispersion would not occur.[40] That repulsive action therefore governed ether in the void, where dispersion did not occur, whereas, to account for dispersion in matter, the force on an ether particle had somehow to be changed. Cauchy did not here discuss the mechanism that produced the change, but in any case his (incorrect) deduction of the $1/r^4$ repulsion for nondispersive media remained generally unknown for some time, as, in fact, did his dispersion formula itself.

Cauchy's theory of dispersion (with its allied improvement of the theory of double refraction of the late 1820s) had little impact in France when it was published late in 1835.[41] His earlier theory of double refraction had, however, been widely discussed in France, almost certainly because he presented its principles in his lectures at the Collège de France. Cauchy, however, left France after the July Revolution (1830) for political reasons and did not return for eight years. Probably because his dispersion memoir was lithographed in Prague, it was apparently difficult to obtain in France and attracted little attention.

The reception of Cauchy's theory in Britain

In Britain, however, an active programme of research founded on Cauchy's theory was in existence by early 1836. In part, the British programme emerged because Baden Powell, who corresponded with Cauchy, published a series of detailed and favourable articles on the theory in mid-1835 in the widely read *Philosophical Magazine*. The existence of Powell's account explains why Cauchy's theory was well known in Britain, but it does not fully account for the highly favourable reception that the theory there received. One further reason, no doubt, was that the kinematic wave theory was just at this time achieving full recognition in Britain. But that cannot explain the large impact of Cauchy's theory. Indeed, quite possibly Cauchy's ether dynamics was itself a major factor in the acceptance of the wave theory in Britain. Cauchy's theory was influential because for the first time it provided the wave theory with a dynamical foundation to which the tools of modern analysis – themselves introduced in Britain only during the preceding twenty years – could be successfully applied to deduce quantitative propositions. Humphrey Lloyd well expressed the view of the mathematical scientist in 1834, one year before Cauchy's dispersion theory became known in Britain:

> In making this comparison [between the wave and the corpuscular theories of light] it is not enough to rest in vague explanations which may be molded to suit any theory. Whatever be the apparent simplicity of an hypothesis, – whatever its analogy to know laws, – it is only when it admits of mathematical expression, and when its mathematical consequences can be numerically compared with established facts, that its truth can be fully and finally ascertained. Considered in this point of view, the wave-theory of light seems now to have reached a point almost, if not entirely, as advanced as that to which the theory of universal gravitation was pushed by the single-handed efforts of Newton. Varied and comprehensive classes of phenomena have been embraced in its deductions; *and where its progress has been arrested, it has been owing in a great degree to the imperfections of that intricate branch of analysis by which it was to be unfolded.*[42]

The intricate branch of analysis Lloyd mentioned was probably the integration of differential equations, for he was well aware of Cauchy's early work on double refraction and the role that differential equations played in it. Although the analysis remained intricate in Cauchy's 1835 theory, nevertheless Powell's 'Abstract' of its principles gave the mathematical scientist a method for explaining optical phenomena.[43] Within a year several British and Irish mathematicians extended Cauchy's theory to new and previously unanalysed areas of optics. They soon produced an impressive body of results that for about six years dominated optical science. The first results were dictated by the phenomenon that Cauchy had spent half a decade explaining: dispersion. Their initial aim, of course, was to use the theory to calculate dispersive effects for comparison with experiment. This itself posed a number of problems, some theoretical, others experimental, which are worth describing because they are excellent examples of the difficulties frequently faced in nineteenth-century optics in linking theory and experiment.

Foremost among the theoretical problems was the actual deduction of a dispersion formula. Powell had not seen all of Cauchy's memoir (which was issued a year *after* his 'Abstract'), but Cauchy had evidently sent him enough to reproduce preliminary deductions. Unfortunately, the dispersion formula that Cauchy derived required many pages of complicated analysis, and Powell was not a sufficiently competent mathematician to create the analysis by himself. Consequently, he had recourse to an approximation that gave the following formula for the refractive index:

$$\frac{1}{n} = H \, \frac{\sin(\delta/\lambda)}{(\delta/\lambda)}$$

Here H is a constant determined by the properties of the medium (the force law and the molecular spacing), and δ is directly proportional to the distance between consecutive particles – or, rather (and herein lay one form of the approximation), δ is, in effect, an average molecular distance.

Even this formula was quite difficult to compare with Fraunhofer's measured values of the indices of refraction for the seven (solar) spectral lines in different media. To use the formula as it stands, Powell would have had to find an arc for each wavelength that satisfied two conditions: First, the ratio of the arcs for any two spectral lines in a given medium had to be inversely proportional to the ratio of the wavelengths of the two lines; second, the ratio of the sines of the two arcs had to be directly proportional to the ratio of the respective indices of refraction. To avoid this time-consuming method, Powell recurred to a simpler, though again approximate, method that required the indices of two spectral lines as data points, leaving five for comparison with experiment. The results were quite good for the ten media Fraunhofer had investigated, though, as was usual at the time, there was no error analysis.

Later in 1835, William Rowan Hamilton pointed out to Powell that his formula was based on a questionable approximation (that the molecular force extends only between contiguous molecules), and Hamilton deduced a more exact one (which was, however, still approximate).[44] During 1836, Powell embarked on a series of original experimental investigations aimed at confirming Hamilton's formula, for which Hamilton had developed a particularly simple method of calculation that was not approximate, as Powell's earlier method for his first formula had been. Hamilton's formula, however, required more data points than had Powell's: Three indices and the corresponding periods had to be specified for each medium. Powell and Hamilton calculated the indices for the D line in several different media, using, as data points, Fraunhofer's indices for the B, H, and F lines in these media. Powell then tabulated the differences between theory and experiment for Hamilton's formula, as well as the same differences for his first formula. Hamilton's formula agreed with experiment to the fifth decimal place in the index; Powell's agreed only to the fourth.[45]

Continuing his researches, however, Powell found that theory and experiment did not always agree quite so well. In the fall of 1836 he completed a series of experiments that revealed a number of discrepancies.[46] Although the formula remained well confirmed for media of low dispersive power, problems arose for large dispersions. Powell actually found a fairly constant discrepancy for several media in the E and G lines, theory always predicting somewhat too low an index for E and much too high an index for G, the latter even for several weakly dispersing substances. When the dispersion was large, the deviation from theory was quite marked and, Powell admitted, could not be attributed to observational error.[47] Nevertheless, he felt that neglected terms in Hamilton's formula could possibly explain the discrepancy, though he did not attempt the calculation.[48]

The intransigent indices for the E and G lines are particularly fascinating historically because they almost certainly were due to a phenomenon that, though discovered several times, was ignored until 1870: It is now called *anomalous* dispersion. Anomalous dispersion was later seen as a fundamental phenomenon, though the recognition both of the phenomenon and of its importance was accompanied by a profound change in the theoretical complexion of optics.[49] Powell had observed anomalous dispersion, but he had not discovered it. The theoretical perspective imposed by molecular optics, which implied a specific dispersion formula, impelled Powell to conclude that this phenomenon could be explained merely by manipulating the formula implicit in the molecular theory. In the end, Powell bypassed the difficulty altogether by assigning it to experimental problems associated with the measurement of refraction spectra.[50]

The power of the molecular ether theory: selective absorption

The tenacity with which Powell upheld his dispersion formula in the face of discrepant data is one indication of his continued, and growing, faith in molecular analysis. Others in Britain were also persuaded by the theory as early as 1835, for two scientists, John Tovey and Philip Kelland, also produced dispersion formulae – which were consistent with Hamilton's – between 1835 and 1837.[51] I shall shortly discuss Kelland's work, for it led to an important controversy, but first let us examine Tovey's work, which appeared between 1835 and 1842. Tovey extended the molecular theory to selective absorption, which, like dispersion, had not previously been subject to quantitative analysis. His theory is not complicated, but it is one of the most striking examples of the power of molecular analysis and is therefore worth examining as an epitome of the theory.

In selective absorption, discovered by David Brewster in 1833, the coloured substance absorbs numerous distinct spectral portions from a beam of white light.[52] Tovey began his theory with Cauchy's basic molecular equations of motion.[53] Consider a wave travelling along the x axis. Since it is transverse, it has components η, ζ, respectively, along the y,z axes. Cauchy's equations here have the form:

$$d^2\eta/dt^2 = \Sigma b\Delta\eta + \Sigma a(\Delta y\Delta\eta + \Delta z\Delta\zeta)\Delta y$$
$$d^2\zeta/dt^2 = \Sigma b\Delta\zeta + \Sigma a(\Delta y\Delta\eta + \Delta z\Delta\zeta)\Delta z$$

Here a and b are functions of the molecular force. The differences Δx, Δy, Δz over which the summations are carried out are the components of the distances between the ether molecules; $\Delta\eta$, $\Delta\zeta$ are the differences between the values of the components of the displacement for different molecules.

Since ether consists of discrete elements, we must calculate $\Delta\eta$, $\Delta\zeta$ by taking *finite* differences. If η, ζ are represented by the exponentials $ce^{(nt+kx)}$, $de^{(nt+kx)}$, respectively, in which k can be a complex number (if the medium is absorbent), and n is a pure imaginary, then we have, for the difference $\Delta\eta$ between two molecules, one at x_2 and the other at x_1, with $x_2 - x_1 = \Delta x$:

$$\begin{aligned}\Delta\eta = \eta(x_2) - \eta(x_1) &= ce^{(nt+kx_1)}(e^{k(x_2-x_1)} - 1)\\ &= ce^{(nt+kx_1)}(e^{k\Delta x} - 1)\end{aligned}$$

This is an *exact* equation.

Using a similar expression for $\Delta\zeta$, Tovey substituted the two into Cauchy's equation of motion, first defining several constants: s, equal to $\Sigma\left[b\,a(\Delta y)^2\right](e^{k\Delta x} - 1)$; s', equal to $\Sigma\left[b + a(\Delta z)^2\right](e^{k\Delta x} - 1)$; and s_1, equal to $\Sigma b\Delta y\Delta z(e^{k\Delta x} - 1)$. As a result he found that, setting the imaginary number n equal to $m(-1)^{1/2}$, where m is real:

$$(m^2 + s)(m^2 + s') = s_1 \tag{7.1}$$

Now k can be complex, say $q + r(-1)^{1/2}$, and, if it is, and if q is less than zero, then clearly the wave will be attenuated, since, taking real parts, it will decrease as e^{qx}. Moreover, r is equal to 2π divided by the wavelength. Consequently, according to equation (7.1), the wavelength, the frequency (n), and the coefficient of absorption (q) are linked by a quartic equation in m: Any two of q, r, m determine an equation for the other. The result is, obviously, that the absorption depends in a quite complicated fashion upon the colour of the light and upon the molecular characteristics of the medium.

The existence of a relation between colouration and absorption seems at first to be remarkably puzzling, because it appears to imply that the real and the imaginary parts of a complex number (q and r) must be related to one another. Herein lies the power of molecular theory, for the relationship is a direct result of the discrete character of ether. If the distances between the ether molecules were so small that the discreteness could be neglected, then Tovey might have calculated $\Delta\eta$ by taking the differential, $d\eta$, equal to $k\eta\Delta x$. This would have given him, instead of (7.1), an equation containing only real terms on the left, and the product of the complex number k and a real term on the right. Consequently, k would have to be real; r would accordingly be banished; and therefore no wave motion at all would be implied. In other words, Tovey had concisely demonstrated that, if ether is discrete, then some form of selective absorption *must* occur (though it could still be eliminated by properly adjusting the molecular constants for transparent media and for the void).

Although Tovey did not actually test the formula – it was too complicated and the experimental data were as yet meager – nevertheless the implication was clear and powerful: The molecular theory of ether could in principle even provide a quantitative explanation of a complex phenomenon. Tovey, and Powell, further demonstrated between 1837 and 1842 that, if the molecular constants are properly adjusted, then ether will propagate *only* elliptically polarised waves. Since reflection always produces elliptical polarisation (except, it was – incorrectly – thought, at the 'polarising' angle), they therefore had a physical account of the state of ether at the boundaries between media.[54]

The effect of matter upon ether

As molecular theory was increasingly exploited, controversies arose concerning the causes of particular phenomena. One controversy involved the cause of dispersion. Some – Kelland in particular – attributed dispersion to the finite distances between ether particles, a hypothesis requiring a particular force law. Others attributed dispersion to the direct action of material upon ethereal particles; for example, Matthew O'Brien proposed this kind of theory in 1842. However, Samuel Earnshaw, a Cambridge mathematician, vehemently rejected Kelland's theory; angry public exchanges between Kelland and both Earnshaw and O'Brien resulted.[55]

Although these differences occurred between scientists who, at the time, accepted the fundamental principles of the molecular theory, nevertheless they reflect an important internal difficulty in the theory that was extremely important after 1842. The differences involved the role of matter in optical phenomena. Although most scientists assumed that material particles exerted forces on the ether particles, they differed considerably about the details of the process. Kelland assumed that an extremely short-range force acted between ether and matter, its function being to separate ether from the matter particles and not to affect the motion of the ether particles. The major effect of this force was therefore to cause matter to displace ether, thereby altering the distances between ether particles. According to Kelland, dispersion depended solely upon these distances. Hence in his theory the force between ether and matter only *indirectly* produced dispersion. O'Brien and Earnshaw, on the other hand, used a longer-range force, which they insisted directly caused dispersion by affecting the motion of the ether particles. The issue was a complicated one, but it was generally agreed that the presence of matter had

one effect on ether that could not be treated exactly with the current tools of analysis: It rendered ether inhomogeneous.

Every molecular theory of optics in the 1830s and the early 1840s, whether or not it employed matter directly, presupposed that ether remained homogeneous throughout a given material body. That is, as noted previously, no single volume of ether differed in the number of molecules it contained from any other portion. The spatial arrangement of the molecules was, of course, affected by matter, but the density (number of molecules per unit volume) was constant for a given substance. However, it was universally admitted that, in fact, matter *must* render ether inhomogeneous, since in the vicinity of a material particle the ethereal distribution had to be different from the distribution at a distance from the particle, whether the ether–matter force was extremely short range or not. Since inhomogeneity would make ether's equations of motion impossibly difficult to integrate, it was ignored. This type of problem – which is purely internal to the structure of the theory – occurs frequently in the history of physics, and usually the partisans of an ongoing programme of research either claim that the problem does not affect the validity of the researches undertaken, without first solving it, or simply do not write about it. Both tacks were taken in molecular theory.

One kind of problem that cannot be ignored or easily dismissed, however, is an unavoidable conflict between theory and experiment. If the conflict seems to be manageable by including further terms in an approximate formula, or if it constitutes only one or two exceptions within a given class the other members of which are not aberrant, then the discrepancy may be explained by recurring to difficulties of calculation or of experiment. For example, Powell first attempted to explain the difficulty with the indices of the E and G lines by arguing that terms had been left out of his dispersion formula; if the extra terms were included, he claimed, the discrepancy would disappear. He later abandoned this explanation, which would not have worked, and instead blamed the failure on experimental difficulties. But this kind of difficulty is essentially a numerical one: For example, it was not that dispersion was totally incompatible with theory, since only a small proportion of the results did not fit the theory well. There is, however, a second kind of discrepancy between theory and experiment that is much deeper, for the theory may be fundamentally incompatible with an entire class of phenomena. Given the startling successes of ether theory in explaining phenomena as diverse as double refraction, dispersion, selective absorption, and the production of elliptical polarisation by reflection, it is not surprising that theorists strongly believed that the theory could be applied successfully to all optical phenomena. In 1841 they were proved decisively wrong.

The problem of optical rotation

When a beam of linearly polarised light passes through a crystal of quartz in a certain direction (or through certain types of solutions in any direction), it is split into two beams that travel at different speeds but in the same direction. One beam is left circularly polarised; the other is right circularly polarised. A single resultant beam emerges, and it is again linearly polarised, but its plane of polarisation has been rotated. The phenomenon is appropriately called optical rotation. It was widely assumed in the late 1830s and the early 1840s that Cauchy's molecular equations could explain the phenomenon. But in 1841 the irascible Irish mathematician James MacCullagh conclusively and simply demonstrated that optical rotation is *incompatible* with the molecular equations of motion.[56]

A problem of this depth requires a solution of equal depth, and a quantitative one, if the theory as a whole is to remain capable of motivating further research. But in this case the difficulty was so profound that only someone who possessed a penetrating familiarity with modern analysis could even hope to pursue it. Cauchy had that familiarity, and when he realised that the problem was linked to the basic form of his molecular equation, he began to reflect upon another problem, previously dismissed casually, which also affected the equation. This was the problem of the inhomogeneity of ether in the presence of matter. At some time after 1844, Cauchy realised that the latter problem contained the solution to the otherwise fatal conflict between his basic equation and optical rotation. In effect, Cauchy argued, one can imagine that, in crystals and in certain solutions, the inhomogeneity imposed upon ether by matter is not random – as had previously been assumed – but periodic: That is, the alteration in the density of ether repeats at certain small intervals throughout the body. This implies that the coefficients in Cauchy's basic equation are not constants, nor do they vary haphazardly; they are periodic functions of position, and thus the partial differential equations had to be integrated with periodic coefficients. This was a difficult task, and Cauchy never published a complete method of integration. However, he was able to show that periodicity contained the answer to the problem of optical rotation.[57] Uniting an old puzzle (inhomogeneity and the effect of matter upon ether) with a pressing quandary (optical rotation), Cauchy not only saved molecular theory; he gave it a new mathematical foundation. However, the new mathematics was extremely complicated and underdeveloped. During the years after Cauchy's death in 1857, two French scientists – Charles Briot and Emile Sarrau – continued Cauchy's optical work. Their studies resulted finally in Sarrau's concise and rigorous exposition of 1867,[58] which was the last innovative use of molecular principles in optics.

Conclusion

The function of the molecular theory of ether in the development of wave optics was fundamentally mathematical and empirical. It was initially used by Cauchy to create a differential equation for wave motions in anisotropic media. He further developed the mathematics of the theory to deduce a law for dispersion. During the late 1830s and the early 1840s, Cauchy's analysis was further applied to discover the law of selective absorption and to explain elliptical polarisation. During these years the molecular ether was utilised primarily because it yielded a wealth of mathematical and empirical consequences.

Like all hypotheses, however, this one was not complete, in that, in its early form, it failed to specify the link between ether and matter. Divergent opinions on this question arose, but all shared the fundamental hypothesis that both matter and ether consisted of material points between which central forces acted directly at a distance. During the first twelve years of the life of Cauchy's theory (1830–42) most scientists concentrated on deducing laws from the theory under the assumption that, whatever the true effect of matter on ether, one could nevertheless ignore it mathematically.

This situation might have continued for several more years if the theory had been altogether successful in yielding empirical laws. But it was not. It failed to explain optical rotation, and the failure pointed to a deep-seated difficulty. In circumstances like these, scientists often become acutely conscious of lacunae in their work that they had previously ignored, and Cauchy was so affected. The function of matter therefore became a pressing question; by introducing a quantitative hypothesis concerning the ether–matter link, he was able to solve the crisis precipitated by optical rotation.

When, later in the nineteenth century, the concept of ether was generally accepted, it served a number of purposes, many of which were neither mathematical nor empirical. It buttressed philosophical speculation;[59] it salved religious doubt by locating the supernatural in ethereal regions.[60] However, ether always retained a distinctly scientific function, and it frequently yielded important results. We have seen that, in a direct way, the hypothesis of the molecular ether gave birth to the science of the differential equations of wave optics.

Notes

1 In double refraction a ray of light entering the crystal is, in general, split in two. One of the two obeys Snell's law, whereas the other does not. If the two rays, after leaving the crystal, are incident on a second crystal, then there are two positions of the second crystal with respect to the first in which each ray produces only one

ray – rather than two rays – in the second crystal. In one of these two positions, each ray produces a ray that obeys the same law of refraction as the original ray; in the other position, each produces a ray that obeys the contrary law. The two-crystal phenomenon involves the property of light now called polarisation. Huygens was able to deduce a law of refraction for the ray that does not obey Snell's law, but he was unable to explain polarisation. Polarisation eluded him because, since his pulses of light were, like sound, compressions and rarefactions, they were entirely symmetrical about the axis of the ray, and therefore no phenomenon should have depended on the orientation of the refracting body.

2 Huygens assumed that each point on the surface of a wave was itself the source of an expanding front, and that these various secondary fronts determine by their common tangent at a given moment the resultant primary front. In Iceland spar, Huygens assumed, the secondary waves consisted of spheres (for the ray that obeys Snell's law) and spheroids (for the ray that does not).

3 A dynamical theory is one that, instead of simply assuming the possibility of certain motions, deduces them from more basic propositions. For a discussion of kinematic optics, which does not involve dynamics, see A. E. Shapiro, 'Kinematic optics: a study of the wave theory of light in the seventeenth century', *Archive for History of Exact Sciences 11* (1971), 134–226.

4 James Bradley's discovery of stellar aberration and his explanation of it in terms of the finite speed of light (1728) rapidly quelled any lingering opposition.

5 Newton's published articles on optics in the 1670s are conveniently collected in section 2 of *Isaac Newton's papers and letters on natural philosophy*, ed. I. B. Cohen (Cambridge, Mass., 1958).

6 I. Newton, *Principia*, bk. 1, props. 94, 95. The difficulties Newton faced in developing a consistent theory of dispersion are discussed in Z. Bechler, 'Newton's law of forces which are inversely as the mass: a suggested interpretation of his later efforts to normalise a mechanistic model of optical dispersion', *Centaurus 18* (1974), 184–222.

7 I. Newton, *Opticks*, 4th ed. (London, 1730; reprinted, New York, 1952), 347–54.

8 Ibid., bks. 2,3.

9 Ibid., 354–61. Newton's optical corpuscles could have a shape, in which case there would be an asymmetry about the ray.

10 One of the most striking theoretical successes of corpuscular optics in the eighteenth century was its account of stellar aberration, which Huygens's theory could not readily explain (though no one tried to use it for that purpose until Fresnel, who greatly modified it; Huygens had actually denied the possibility of anything like aberration in his *Traité* by arguing that the motion of the observer does not affect the visual position of the object).

11 R. Fox, 'The rise and fall of Laplacian physics', *Historical Studies in the Physical Sciences 4* (1975), 89–136.

12 The concept of the material point and the idea of grouping points to form complex structures were first developed by R. J. Boscovich, *A theory of natural philosophy*, trans. J. M. Child (Cambridge, Mass., 1966).

13 P. de Laplace, *Mécanique céleste*, trans. Nathaniel Bowditch, 4 vols. (Boston, 1829–39), 4:415.

14 E. Frankel, 'The search for a corpuscular theory of double refraction: Malus, Laplace and the prize competition of 1808', *Centaurus 18* (1974), 222–45.

15 Although usually referred to in the early nineteenth century as the principle of 'least' action, it should be called the principle of extreme action, since, under certain circumstances, the action is a maximum. The principle asserts that the integral of the velocity of light over distance is an extremum for the path actually traversed. The principle holds for any system that is governed by central forces acting at a distance.

16 E. Malus, 'Sur une propriété de la lumière réfléchie par des corps diaphanes', *Mémoires de la Société d'Arcueil 2* (1809), 254–67.

17 D. F. Arago, 'Mémoire sur une modification remarquable qu'éprouvent les rayons lumineux dans leur passage à travers certains corps diaphanes, et quelques autres nouveaux phénomènes d'optiques', *Mémoires de l'Institut*, pt. 1 (1811), 92–134.

18 E. Frankel, 'Jean-Baptiste Biot', unpublished doctoral dissertation, Princeton University, 1972, 231–86.

19 Material points had to be gathered together in groups to form the corpuscle in order to grant it a shape, which was necessary to explain polarisation phenomena.

20 T. Young, 'Outline of experiments and inquiries respecting sound and light', *Philosophical Transactions 90* (1800), 106–50; T. Young, 'On the theory of light and colours', ibid. *92* (1802), 12–48; T. Young, 'An account of some cases of the production of colours', ibid., 387–97; T. Young, 'Experiments and calculations relative to physical optics', ibid. *94* (1804), 1–16.

21 Newton had urged against all theories of light as a motion that, because the motion would diverge into the shadow, one would be able to see around corners (*Opticks*, 362–3). For Young's reply see the first and second articles cited in n. 20.

22 According to the corpuscular explanation of diffraction, a deflection occurs in the *immediate vicinity* of the diffracting body owing to the short-range force. Moreover, the deflection depends upon the colour, i.e., upon the type of corpuscle.

23 The principle of least time, like the principle of least action (n. 15), is in fact an extremum principle, for under certain circumstances the time can be a maximum. The principle asserts that the integral of the inverse of velocity over distance is an extremum. One can convert from least time to least action, and conversely, by replacing the velocity with its inverse – which, of course, corresponds to the opposite assumptions of corpuscular and wave optics concerning the velocity of light in refraction.

24 Young's calculations went no further than comparing the distances from two optical point sources to a given point; moreover, he could only predict positions of total destructive or total constructive interference: He could not calculate intermediate intensities.

25 Letter from A. Fresnel to L. Fresnel, 5 July 1814, in A. Fresnel, *Oeuvres complètes d'Augustin Fresnel, publiées par MM. Henri de Sénarmont, Emile Verdet et Léonor Fresnel*, 3 vols. (Paris, 1866–70), 2:821. Fresnel's work has yet to receive a comprehensive treatment. See, however, R. Silliman, 'Augustin Fresnel (1788–1827) and the establishment of the wave theory of light', unpublished doctoral dissertation, Princeton University, 1967.

26 A. Fresnel, 'Supplément au deuxième mémoire sur la diffraction de la lumière', in Fresnel, *Oeuvres*, 1:150–5. Fresnel set two plane mirrors facing one another at an acute angle and placed a light source within the angle. He obtained an interference pattern where the rays reflected from one mirror intersected those reflected from the other mirror. This was surprising, on corpuscular principles, because no matter was present to deflect the rays.

27 At this stage, Fresnel, like Young, could calculate only points of total destructive or total constructive interference produced by two point sources.

28 Fresnel used Huygens's principle to effect a general theory of diffraction: Fresnel, 'Mémoire sur la diffraction de la lumière, couronné par l'Académie des Sciences', in Fresnel, *Oeuvres*, 1:248.

29 Silliman, 'Fresnel', 189–90.

30 See n. 1.

31 A. Fresnel and D. F. Arago, 'Mémoire sur l'action que les rayons de lumière exercent les uns sur les autres', in Fresnel, *Oeuvres*, 1:509–22.
32 A. Fresnel, 'Nôte sur le calcul des teintes que la polarisation développe dans les lames cristallisées', in Fresnel, *Oeuvres*, 1:629.
33 Fresnel, 'Nôte sur le calcul', 629.
34 The 'wave surface' is a surface whose radii are each equal to the distance that a ray of light refracted along the radius travels in a unit of time. It is used to find the angle of refraction for a given angle of incidence.
35 A. Cauchy, 'Sur l'équilibre et le mouvement d'un système de points matériels sollicités par des forces d'attraction ou de répulsion mutuelle', in A. Cauchy, *Oeuvres complètes de Cauchy* (2 ser., Paris, 1882–1975), ser. 2, 8:227–52.
36 A. Cauchy, 'Mémoire sur la théorie de la lumière', *Mémoires de l'Académie des Sciences 10* (1830), 293–316.
37 A. Cauchy, *Mémoire sur la dispersion de la lumière* (Prague, 1836); reprinted in Cauchy, *Oeuvres*, ser. 2, 10:195.
38 G. B. Airy, 'On the nature of the light in the two rays produced by the double refraction of quartz', *Transactions of the Cambridge Philosophical Society 4* (1833), 79–123. Airy conceived that, if ether were heated on compression, then the wave velocity would depend on the frequency: The quicker the vibrations, the *less* time there would be for the heat to increase the elasticity of the medium. Hence the slower the vibration – with the implication of more time for the heating to increase the elasticity – the greater the wave velocity. Airy was not proffering the hypothesis as more than a suggestion, for it obviously introduced into optics all of the growing contemporary problems concerning the nature and effects of heat.
39 A. Fresnel, 'Extrait du second mémoire sur la double réfraction', in Fresnel, *Oeuvres*, 2:473.
40 Cauchy's deduction was flawed by a trivial mathematical error (he neglected a differentiation). The correct repulsion to forbid dispersion varies inversely as the sixth power of the distance.
41 See n. 37.
42 H. Lloyd, 'Report on the progress and present state of physical optics', *Annual Report of the British Association for the Advancement of Science 4* (1834), 295.
43 B. Powell, 'An abstract of the essential principles of M. Cauchy's view of the undulatory theory, leading to an explanation of the dispersion of light; with remarks', *Philosophical Magazine 6* (1835), 16–25, 107–13, 189–93, 262–7.
44 B. Powell, 'On the formula for the dispersion of light derived from M. Cauchy's theory', *Philosophical Magazine 8* (1836), 204–11.
45 B. Powell, 'Remarks on the formula for the dispersion of light', *Philosophical Magazine 9* (1836), 116–19.
46 B. Powell, 'Researches towards establishing a theory of the dispersion of light: no. III', *Philosophical Transactions 127* (1837), 19–24.
47 Hamilton carried out an elementary calculation to determine what changes in the wavelength were necessary to explain the discrepancies. The results were much too large for the accuracy of Fraunhofer's experiments.
48 In fact, if Powell had calculated further terms, the discrepancy would have been even greater, not smaller.
49 After the mid-1870s refraction of all kinds was attributed to characteristic vibrations within or of the molecules of bodies. If the frequency of the incident light was less than that of the characteristic vibration, then the index of refraction rose from unity to its maximum as a function of increasing frequency, reaching the maximum just below the characteristic frequency. The index decreased as the frequency of light approached the characteristic frequency from above, reaching a minimum (less than

one) just above it. 'Ordinary' dispersion was explained by characteristic frequencies in the ultraviolet; in anomalous dispersion the body contained, in addition to characteristic frequencies in the ultraviolet, others in the visible spectrum. Finally, absorption occurred at and near the characteristic frequencies.
50 He argued that the *G* line in the refraction spectrum was probably not single but an assemblage of narrow lines with consequently different wavelengths. The *G* line produced by an interference spectrum – which was Fraunhofer's technique for measuring wavelengths – was, on the other hand, extremely narrow, argued Powell, with a correspondingly precise wavelength. Consequently, index measurements that used wavelengths found from interference spectra would not be reliable for this line. See B. Powell, 'Researches towards establishing a theory of the dispersion of light: no. IV', *Philosophical Transactions 128* (1838), 67–72.
51 Kelland's and Tovey's memoirs are too numerous to list here. Kelland's work was published in the *Philosophical Magazine* at intervals from 1836 to 1842, and in the *Philosophical Transactions* in 1836. Tovey's work appeared exclusively in the *Philosophical Magazine* from 1836 to 1842.
52 D. Brewster, 'Observations of the absorption of specific rays, in reference to the undulatory theory of light', *Philosophical Magazine 2* (1833), 362–8.
53 J. Tovey, 'Researches in the undulatory theory of light, continued: on the absorption of light', *Philosophical Magazine 15* (1839), 450–5, and *16* (1840), 181–5.
54 See, e.g., B. Powell, 'Remarks on the theory of the dispersion of light as connected with polarization', *Philosophical Transactions 128* (1838), 253–64; B. Powell, 'A supplement', ibid. *132* (1842), 157–9.
55 Again, the exchanges are too numerous to list here; all were printed in the *Philosophical Magazine* between 1839 and 1842.
56 J. MacCullagh, 'Notes on some points in the theory of light', in J. MacCullagh, *Collected works of J. MacCullagh* (Dublin, 1880), 194–217.
57 A. Cauchy, 'Mémoire sur les perturbations produites dans les mouvements d'un système de molécules par l'influence d'un autre système', *Comptes Rendus 30* (1850), 17–24.
58 C. Briot, *Essais sur la théorie mathématique de la lumière* (Paris, 1864); E. Sarrau, 'Sur la propagation et la polarisation de la lumière dans les cristaux', *(Liouville) Journal de Mathématiques Pures et Appliquées, 12* (1867), 1–46, and *13* (1868), 59–110.
59 See, e.g., H. Spencer, *First principles*, 4th ed. (New York, 1880).
60 See, e.g., B. Stewart and P. G. Tait, *The unseen universe; or, physical speculations on a future state*, 3rd ed. (London, 1875).

8

Thomson, Maxwell, and the universal ether in Victorian physics

DANIEL M. SIEGEL

Department of the History of Science, University of Wisconsin, Madison, Wisconsin 53706 USA

By the middle of the nineteenth century the wave theory of light had become widely accepted in Britain, and with it had come belief in the existence of the luminiferous ether – an elastic solid that filled space and whose transverse undulations constituted light waves. In the decades after 1850 the luminiferous ether was assigned additional functions: James Clerk Maxwell showed how the optical ether could be fruitfully regarded as the seat of electrical and magnetic effects as well; and William Thomson (Lord Kelvin) argued that atoms of ordinary matter could be viewed as nothing but patterns of vortex motion in a ubiquitous, space-filling medium – 'the Universal Plenum' (Knudsen, 1972:200). These ideas of Maxwell and Thomson, as interpreted and elaborated in the closing decades of the nineteenth century by George Fitzgerald, Oliver Lodge, Joseph Larmor, and others, gave rise to the notion of a truly universal ether. This 'fundamental' and 'primordial medium' was 'assumed to be the ultimate seat of all phenomena', and all phenomena were then seen as 'dynamical', being referred to motions and associated forces in the universal medium. The programme of dynamical explanation based on the ether had great appeal because of the unification it promised, and it enjoyed great success in connection with electromagnetic theory and optics; thus, in spite of its very limited success in the realm of matter theory, the programme compelled allegiance through the end of the nineteenth century.[1]

Thomson, Maxwell, and Faraday's lines of force

That the ideas of Michael Faraday played a crucial role in the development of ether theory is not without irony, for he started out from a position of fairly radical scepticism about subtle fluids and ethers and remained reasonably true to that scepticism. It was Thomson and Maxwell who identified Faraday's lines of force with mechanical conditions in a material medium, thus providing the background for a unified ether theory of electromagnetic and optical phenomena.

Faraday, in brief, rejected the electrical and magnetic fluids and the action-at-a-distance forces associated with them, and instead proposed an approach that he considered less hypothetical. He characterised electrical and magnetic phenomena by the geometrical patterns in space of the vectorial forces that would be exerted on electrical or magnetic test bodies. These spatial patterns were graphically delineated by representative 'lines of force', whose directions represented the directions of the forces, while their spacing indicated the magnitudes of the forces (close spacing indicating strong force, and vice versa). Faraday managed to work out a rudimentary calculus of these lines of force, which enabled him to deal with known electrical and magnetic phenomena effectively, and which proved immensely fruitful in suggesting new experiments that led to the discovery of novel phenomena.[2]

Though Faraday's sceptical stance militated against the identification of these abstract lines of force with some concrete space-filling substratum or ether, he firmly believed in the relatedness of all natural phenomena, and he speculated in 1844 on a possible relationship between his lines of force and light. He proposed that 'radiant phaenomena' – light and radiant heat – be referred to transverse vibrations of lines of force rather than to the luminiferous ether; this 'notion [would] . . . dispense with the aether'. Faraday's speculations concerning a relationship between electromagnetic phenomena and light were given a convincing experimental foundation in the very next year, when he found that a beam of light propagating in a piece of glass situated in a strong magnetic field would experience a rotation of its plane of polarisation. This discovery of an action of magnetism on light profoundly influenced the later work of Thomson and Maxwell, resulting ultimately in the assimilation of Faraday's lines of force to the luminiferous ether. Indeed, Faraday himself later adopted a somewhat more positive attitude toward ether, granting that magnetic force, in particular, 'may be a function of the aether; for it is not unlikely that, if there be an aether, it should have other uses than simply the conveyance of radiations'.[3] This turn of Faraday's thought in the early 1850s was closely tied to what Thomson had done with Faraday's ideas and discoveries in the 1840s (Doran, 1975:162–79; Buchwald, 1977).

Thomson's scientific career spanned the Victorian era, from the 1840s to the opening years of the twentieth century, and throughout he was centrally occupied with questions relating to ether. In the 1840s, initially on an analogical basis, he began to associate Faraday's lines of force with conditions in a medium in space. Paradigmatic for Thomson and subsequently for Maxwell was an analogy that had arisen in the context of Thomson's studies of Joseph

Fourier's theory of the conduction of heat in solids. Thomson found that the relationship between heat sources and heat flow in this case was analogous to the relationship between electricity and electrical attraction in the conventional treatment of electrostatics, as developed by Coulomb, Laplace, and Poisson. Given this analogy, which Thomson developed in strict mathematical fashion in his initial publication of 1842, he was able simply by reinterpreting the symbols in the equations to transfer theorems and calculations from electrostatics to heat flow and vice versa.[4]

The relevance of this analogy to Faraday's ideas did not emerge until 1845, when Thomson spent some time in Paris, working in the laboratory of Victor Regnault. Faraday's work had been received with some bemusement on the Continent, and Thomson was asked if he could explicate Faraday's approach, giving particular attention to the relationship between Faraday's electrical lines of force and the traditional, action-at-a-distance electrostatics. Thomson took on the task and found that some of the mathematical tools he had used in his initial and continuing investigations of the analogy between heat flow and conventional electrostatics were relevant.[5] In particular, Thomson had developed, in part on his own and in part on the basis of his reading of works of George Green and Carl Friedrich Gauss, a mathematical approach to electrostatics that emphasised the spatial distribution and geometrical relationships of electrical forces. This approach made use of the language of partial differential equations and potential theory, including the electrostatic potential function and its derivatives, the associated families of orthogonal curves, and the relevant theorems of Gauss and Green. Using this formalism, Thomson was able to show that action-at-a-distance electrostatics and Faraday's calculus of electrical lines of force could both be cast in the same mathematical form, so that they were equivalent, at least in mathematical and operational terms.[6]

Thomson was now in a position to consider two alternative perspectives on Faraday's lines of force. On the one hand, because a complete equivalence between Faraday's treatment and action-at-a-distance electrostatics had been demonstrated, Faraday's work – including the work on dielectric materials, which Thomson had managed to treat by the methods Poisson had developed for magnetic materials – was no longer a threat to the continued acceptance of action at a distance. On the other hand, Faraday's calculus of electrical lines of force had now been mathematised in such a manner that it could be inserted as one term in Thomson's original analogy, which then became an analogy between Faraday's calculus of electrical lines of force and Fourier's treatment of heat flow. If one took the mathematical analogy to be suggestive of a physical similarity, the implication would be that, just as heat flow depends on an 'intervening medium', so also do Faraday's lines of force depend upon a medium:

> It is, no doubt, possible that . . . [electrical forces] may be discovered to be produced entirely by the action of contiguous particles of some intervening medium, and we have an analogy for this in the case of heat, where certain effects which follow the same laws are undoubtedly propagated from particle to particle. It might also be found that magnetic forces are propagated by means of a second medium, and the force of gravitation by means of a third.[7]

Faraday's discovery of the magnetic action on light in 1845 stimulated further development of Thomson's ideas concerning the dependence of electrical and magnetic forces on a medium or media. In a paper of 1847 entitled 'On a mechanical representation of electric, magnetic, and galvanic forces', Thomson considered states of linear and rotational strain in an elastic solid, showing that their distributions in space would be analogous respectively to distributions of electric force (for the case of a point charge) and magnetic force (for the cases of magnetic dipoles and electric currents).[8] Once again the basic thrust was analogical, but once again there was an undercurrent of physical significance. This was evident in the choice of a rotational mechanical strain to represent magnetic force, in consonance with Faraday's finding that the magnetic action on light was rotary in character. Furthermore, as Thomson wrote to Faraday at the time, though what he had published was merely a 'mathematical analogy', what he looked forward to was a 'physical theory':

> 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 . . . If such a theory could be discovered, it would also, when taken in connection with the undulatory theory of light, in all probability explain the effect of magnetism on polarized light.[9]

Thomson, then, presented his results to Faraday with diffidence; a real theory was a long way off. But it was already apparent that, when that theory came, it would exhibit a connection with the luminiferous ether. Looking back, half a century later, Thomson traced the onset of his 'fits of ether dipsomania', which had left him 'not . . . a moment's peace' over a period of fifty years, to the time of the composition of this paper of 1847.[10]

Thomson soon acquired a close follower in these endeavours, a fellow Scotsman seven years his junior named James Clerk Maxwell. The rudiments of a master–apprentice relationship centering on electromagnetic subjects were established even before Maxwell went up to Cambridge in 1850.[11] In 1854, opening a correspondence that was to continue until his death in 1879, Maxwell wrote to Thomson, asking him to prescribe a course of reading in electricity and magnetism: 'If [one] wished to read Ampère Faraday &c how should they be arranged, and at what stage and in what order might he read your articles?' In 1855, Maxwell reported on his reading, indicating that he had concentrated on the works of Thomson and Faraday and that he had been particularly impressed by Thomson's analogical method: 'Have you patented that notion with all its applications?' asked Maxwell; 'for I intend to borrow it for a season'.[12] Borrow it he did, making it the basis for a paper entitled 'On Faraday's lines of force'. In the introductory section of this paper he discussed 'physical analog[ies]' and indicated his reasons for using them. 'By a physical analogy', he wrote, 'I mean that partial similarity between the laws of one science and those of another which makes each of them illustrate the other'. As a prime example of a physical analogy, he described in some detail the parallels between heat flow and electrostatics that had been pointed out by Thomson. Thomson, it will be recalled, saw such analogies as merely steps along the way to a hoped-for 'physical theory'. Maxwell agreed, giving a clear exposition of this preparatory role of physical analogies, in particular as applied to electricity and magnetism. Maxwell felt that although much was known about the various branches of electricity and magnetism, there were also great gaps in existing knowledge, so that it would be folly to attempt a complete theory at this point.

In this situation, Maxwell felt that there were three possible approaches. First, one might try for a partial physical theory or hypothesis; but this would be dangerous, as it would tend to foster premature commitment: 'If . . . we adopt a physical hypothesis, we see the phenomena only through a medium, and are liable to that blindness to facts and rashness in assumption which a *partial* explanation encourages'. Of course one could avoid any premature conjecture concerning the actual physical basis of the phenomena by seeking refuge in a 'purely mathematical' description, devoid of 'physical conceptions', but such an approach was also to be avoided, since it would inevitably be unfruitful. The third possible approach – indeed, the approach of his choice – was that of physical analogy: Here one could avoid any commitment as to the actual physical nature of the phenomena, for the physical picture entertained was merely an analogy; on the other hand, one also avoided the sterility of the purely mathematical approach. The construction of physical analogies, then, was an appropriate first step in electrical science; the ultimate goal, however, for Maxwell as for Thomson, was a 'mature theory, in which physical facts will be physically explained'.[13]

In discussing physical analogies in general, Maxwell allowed for the possibility that such an analogy might ultimately form the basis for a physical theory. Thus, in the case of the wave theory of light, the original basis was an 'analogy . . . between light and the vibrations of an elastic medium' – a 'resemblance *in form* between the laws of light and those of vibrations' – and ultimately this 'resemblance in mathematical form . . . [gave] rise to a physical theory of light'. In the case of the particular analogies for electromagnetic phenomena that Maxwell was presenting in 'Faraday's lines', however, no basis for a physical theory was evidently being envisioned. Maxwell treated electrical lines of force, magnetic lines of force, and electric currents each by analogy with the flow of an incompressible fluid through a resistive medium, while insisting that this incompressible fluid

> is not even a hypothetical fluid which is introduced to explain actual phenomena. It is merely a collection of imaginary properties which may be employed for establishing certain theorems in pure mathematics in a way more intelligible to many minds and more applicable to physical problems than that in which algebraic symbols alone are used.

Intended as heuristic devices, these analogies in fact proved quite successful, enabling Maxwell to progress with the task begun by Thomson of expressing Faraday's calculus of lines of force in partial differential equations. The resulting equations, with some later modifications, have come to be known as 'Maxwell's equations'.[14]

Within a year of the publication of Maxwell's 'Faraday's lines', Thomson decided that the time had come to go beyond mere heuristic analogy in the mechanical representation of electromagnetic phenomena. In a paper of 1856 he announced that he was now proposing a description of 'reality', of the 'ultimate nature of magnetism'.[15] This new departure must be understood against the broader background of Thomson's developing commitment to a 'dynamical' understanding of physical phenomena. Looking back some years later, Thomson traced the origin of this programmatic commitment to an encounter with James Prescott Joule at the British Association meeting in 1847, when he 'learned from Joule the dynamical theory of heat, and was forced to abandon at once many, and gradually from year to year *all* other, statical preconceptions regarding the ultimate causes of apparently statical phenomena'. Thomson's conversion to the dynamical theory of heat was in fact not quite so abrupt; he continued to defend the caloric theory against Joule's novel views for some years after 1847. In essence, however, Thomson's recollection was correct: He did experience a dramatic conversion to the dynamical theory of heat – by 1851 if not in 1847 – and the general lesson that he abstracted from this conversion experience – that all phenomena are ultimately 'dynamical' – informed his thoughts concerning the ultimate nature of things for the next half a century.[16] For Thomson, a 'dynamical' theory, as opposed to a 'statical' theory, was one in which the forces – and hence effects in general – exerted by a given physical system were referred to internal motions within that system, rather than to primitive attractions or repulsions between its particles. Thus, in Thomson's paradigmatic case of the dynamical theory of heat and gases, gas pressure was the result of internal motions rather than the static repulsive forces between caloric particles within the gas.[17]

Thomson in his dynamical theory of heat followed Humphry Davy and W. J. M. Rankine in assuming that the motions that constitute heat were rotary motions associated with individual molecules – 'molecular vortices', in Rankine's terminology. Surrounding each 'molecular nucle[us]', then, there was vortical motion of the material medium that 'interpermeat[es] the spaces between molecular nuclei'. The nature of this material medium was not precisely specified: It might be 'electricity', a 'continuous fluid', or a molecular fluid.[18] In 1856, Thomson drew a connection between this picture of the motion that constitutes heat and a peculiar feature of the Faraday rotation – namely, that the handedness of the rotation of the plane of polarisation of the light beam depends on the direction of propagation. Thomson argued that this distinguishing feature of the Faraday rotation could be explained *only* on the assumption that the magnetic line of force corresponds to an axis of rotation of some of the material through which the light propagates. He concluded that the actual mechanical condition characterising a region traversed by magnetic lines of force would be one where the axes of the molecular vortices were all aligned in one direction, this being the direction of the line of force. This particular mechanical representation of the magnetic line of force was proposed not as an analogy, but as 'reality'. A few years later, in a talk at the Royal Institution, Thomson asserted that 'a certain alignment of axes of revolution in this [vortical] motion IS *magnetism*. Faraday's magneto-optic experiment makes this not a hypothesis, but a demonstrated conclusion'.[19]

There was more to Thomson's dynamical programme. He hoped that ultimately all magnetic forces would be fully explained in terms of the pressures associated with the centrifugal forces of the vortices, whereas electromagnetic induction would be understood in terms of the rotational inertia of the vortices. The material medium in which the vortical motions existed was identified in some notebook entries of 1858 as 'the Universal Plenum', and Thomson speculated that this 'universal fluid' might be the material substratum of the entire physical universe; ordinary matter would then be explained dynamically, in that its properties would be seen as arising from certain motions, 'vortical or other', in this universal plenum.[20] These speculations on vortical motion in a universal ethereal medium were vague and incomplete, but they were soon rendered quite definite and precise by Maxwell and used as the basis for a unified ether theory of electromagnetic and optical phenomena.

Maxwell's theory of molecular vortices

'Professor Thomson has pointed out that the cause of the magnetic action on light must be a real rotation going on in the magnetic field', Maxwell wrote approvingly; the time had come to go beyond the analogical approach and begin constructing, on the basis of Thomson's picture of molecular vortices oriented along the magnetic field lines, something like the 'mature theory' that Maxwell had envisaged in 1855. Maxwell may have begun thinking seriously along these lines as early as 1857,[21] and the results of his deliberations were published in a series of installments – constituting parts 1 through 4 of a paper entitled 'On physical lines of force' – over a period of eleven months in 1861–2. By the time he had finished, Maxwell had not only accounted for the whole known range of electromagnetic phenomena in terms of a mechanical medium in space – the 'magneto-electric medium' – he had also shown that this medium could be identified with the luminiferous ether.[22]

Part 1 of 'Physical lines' was published in March 1861 and was entitled 'The theory of molecular vortices applied to magnetic phenomena'. In the introductory passages, Maxwell contrasted his new theory with the analogies he had worked out earlier. Thus the theory of molecular vortices was not a 'mechanical illustration . . . to assist the imagination' or an 'analog[y]', but a theory with truth value: 'a theory, which if not *true*, can only be proved to be *erroneous* by experiments which will greatly enlarge our knowledge of this part of physics'. The new theory differed from the earlier analogies also in that it would 'account for' or 'explain' the observed forces: In 'Faraday's lines' the purpose had been to 'lay before the mind of the *geometer* a clear conception of the relation of the lines of force to the space in which they are traced' but 'not to account for the phenomena'. 'Now', however, he 'propose[d] . . . to determine what tensions in, or motions of, a medium are capable of *producing* the mechanical phenomena observed'. A parallel contrast was drawn with the analogies Thomson had presented in 1847. Thomson had 'not attempt[ed] to explain the origin of the observed forces', whereas Maxwell was going to 'consider the magnetic influence as existing in the form

of some kind of *stress* in the medium', from which he was going to derive the 'observed forces'.[23]

In part 1 of 'Physical lines', Maxwell conceived the magneto-electric medium as a fluid. In a region of nonzero magnetic field, this medium would be filled with innumerable small vortex tubes or filaments, corresponding in geometrical arrangement to the magnetic field lines; the angular velocities of these 'molecular vortices' would be taken proportional to the field intensity. These rotational motions of the molecular vortices would engender centrifugal forces, causing the vortex filaments to have a tendency to expand equatorially and contract along their lengths; the corresponding magnetic field lines would appear to repel each other and have a tendency to shorten – a behaviour that Faraday had already discussed qualitatively. Evaluating the associated stress tensor in the medium quantitatively, Maxwell was able to account precisely for all classes of magnetic forces. Thus, at least as far as magnetic forces were concerned, Maxwell had successfully shown how to construct an explanatory mechanical theory rather than an illustrative analogy.[24]

In part 2 of 'Physical lines', entitled 'The theory of molecular vortices applied to electric currents', Maxwell began by considering the system of molecular vortices as would a mechanical engineer, and he encountered the following problem:

> 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.

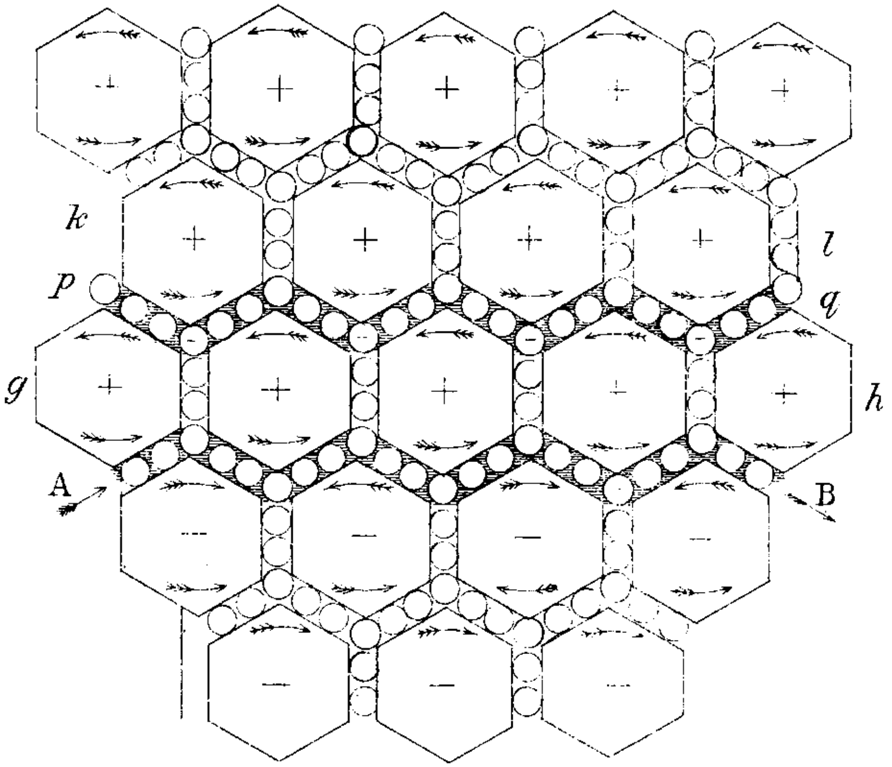
He went on to observe that this kind of problem had been encountered and solved by the designers of mechanical devices:

> In mechanism, when two wheels are intended to revolve in the same direction, a wheel is placed between them so as to be in gear with both, and this wheel is called an 'idle wheel'. The hypothesis about the vortices which I have to suggest is that a layer of particles, acting as idle wheels, is interposed between each vortex and the next, so that each vortex has a tendency to make the neighbouring vortices revolve in the same direction with itself.

The magneto-electric medium was now to be conceived as a cellular medium, each cell consisting of a molecular vortex – a rotating parcel of fluid – surrounded by a cell wall consisting of a monolayer of small spherical particles, which roll without slipping between adjacent vortices, functioning as idle wheels or ball bearings. The cells were taken to be pseudospherical in shape – perhaps dodecahedral – and a string of vortex cells, rather than an extended vortex filament, was now taken to correspond to a magnetic field line. The cells were pictured in cross section as hexagonal (Figure 8.1).[25]

The introduction of the idle-wheel particles had two purposes: It resolved a mechanical problem left over from part 1 of 'Physical lines', and it provided for the further extension of the scope of the theory in part 2. Thus, Maxwell noted that in a region of homogeneous magnetic field, where adjacent vortices had equal angular velocities, the particles making up the cell walls would behave as ordinary idle wheels, rotating but undergoing no spatial translation. In a region of inhomogeneous magnetic field, however, adjacent vortices would have slightly different angular velocities, giving rise to a translational motion of the idle-wheel particles. Characterising the angular velocities of the vortices by a vector field $\boldsymbol{\omega}^*$, and the motions of the idle-wheel particles by

Figure 8.1



an averaged flux density $\boldsymbol{\iota}$, Maxwell calculated on a purely kinematic basis that[26]

$$\boldsymbol{\iota} = (1/4\pi)\ \mathrm{curl}\ \boldsymbol{\omega}^* \tag{8.1a}$$

As Maxwell duly noted, this equation is similar in form to what he called the 'equation . . . of electric currents' (i.e., Ampère's circuital law in differential form), which relates the electric current density $\mathbf{J}$ to the magnetic field intensity $\mathbf{H}$:[27]

$$\mathbf{J} = (1/4\pi)\ \mathrm{curl}\ \mathbf{H} \tag{8.1b}$$

Having calculated equation (8.1a) on a mechanical basis, and having noted its similarity in form to Ampère's law, (8.1b), Maxwell immediately drew the following conclusion: 'It appears therefore, that according to our hypothesis, an electric current is represented by the transference of the moveable particles interposed between the neighbouring vortices'.[28] The logic of this conclusion can be more explicitly represented as follows:

Given the already established correspondence

$$\boldsymbol{\omega}^* \text{ corresponds to } \mathbf{H} \tag{8.1i}$$

and given the parallel mechanical and electromagnetic relationships

$$\boldsymbol{\iota} = (1/4\pi)\ \mathrm{curl}\ \boldsymbol{\omega}^* \tag{8.1a}$$

$$\mathbf{J} = (1/4\pi)\ \mathrm{curl}\ \mathbf{H} \tag{8.1b}$$

it is concluded that

$$\boldsymbol{\iota} \text{ corresponds to } \mathbf{J} \tag{8.1f}$$

The use of this kind of argument ensured that the coherence of the theory would be maintained as its scope was enlarged: The condition for the inclusion of a new mechanical variable in the theory was that it be properly related to the variables already treated, and this condition was imposed at each step in the extension of the theory.

Maxwell next directed attention to the question of changes in the angular velocities of the vortices, corresponding to changes in the magnetic field. Associated with such changes would be inertial forces, exerted by the vortices on the idle-wheel particles; these forces, tangential to the surfaces of the vortices, were characterised by a vector field $\boldsymbol{\tau}_1$. From energetic considerations, Maxwell calculated that $\boldsymbol{\tau}_1$ would be related to $\boldsymbol{\omega}^*$ as follows:[29]

$$-curl\ \boldsymbol{\tau}_1 = (\pi\rho_m)d\boldsymbol{\omega}^*/dt \tag{8.2a}$$

where ρ_m is the mass density of the fluid in the cells. This equation is similar in form to Faraday's law of electromagnetic induction in differential form, which relates the relevant part of the electric field, $\mathbf{E}_1$, to the time derivative of the magnetic intensity $\mathbf{H}$:[30]

$$-\text{curl } \mathbf{E}_1 = \mu(d\mathbf{H}/dt) \quad (8.2b)$$

where μ is the magnetic permeability. Given the established correspondence between $\boldsymbol{\omega}^*$ and $\mathbf{H}$ (and an associated correspondence between the fluid density ρ_m and the magnetic permeability μ), Maxwell concluded that the forces $\boldsymbol{\tau}_1$ could be taken to correspond to the electric field $\mathbf{E}_1$:[31]

Given the already established correspondences

$$\begin{aligned}&\boldsymbol{\omega}^* \text{ corresponds to } \mathbf{H}\\ &\pi\rho_m \text{ corresponds to } \mu\end{aligned} \quad (8.2i)$$

and given the parallel mechanical and electromagnetic relationships

$$-curl\ \boldsymbol{\tau}_1 = (\pi\rho_m)d\boldsymbol{\omega}^*/dt \quad (8.2a)$$

$$-\text{curl } \mathbf{E}_1 = \mu(d\mathbf{H}/dt) \quad (8.2b)$$

it is concluded that

$$\boldsymbol{\tau}_1 \text{ corresponds to } \mathbf{E}_1 \quad (8.2f)$$

Maxwell closed part 2 with what was obviously intended to be a final summary and conclusion to the whole paper. No further installments were contemplated and no extension to electrostatics was promised, though Maxwell was clearly dissatisfied with what he had presented. The problem that he faced and that he did ultimately resolve – further installments of the paper were eventually published – was as follows. The natural route to electrostatics, which Maxwell did ultimately use, can be represented schematically in this way:

Given the already established correspondence

$$\boldsymbol{\iota} \text{ corresponds to } \mathbf{J} \quad (8.3i)$$

and given the parallel mechanical and electromagnetic relationships

$$\text{div } \boldsymbol{\iota} + d\rho_p/dt = 0 \quad (8.3a)$$

$$\text{div } \mathbf{J} + d\rho/dt = 0 \quad (8.3b)$$

it is concluded that

$$\rho_p \text{ corresponds to } \rho \quad (8.3f)$$

where ρ_p is the excess density of idle-wheel particles and equation (8.3a) follows from the implicit assumption that the idle-wheel particles are neither created nor destroyed; ρ is the electric charge density; and equation (8.3b) is what Maxwell called 'the equation of continuity', expressing the conservation of electric charge.[32]

The problem here was that schema (8.3) was not consistent with schema (8.1): Nonzero values for the terms in equations (8.3a) and (8.3b) are inconsistent with equations (8.1a) and (8.1b), which are applicable only to closed circuits, and do not allow for the accumulation of charge. Maxwell had long

been aware that equation (8.1b), Ampère's circuital law in differential form, had restricted applicability, and this was a matter of concern to him.[33] Having incorporated this law into his mechanical theory in the form of equation (8.1a), he now had a theory of restricted applicability, which could not be extended to electrostatics in a straightforward way. This problem halted Maxwell's progress on the theory temporarily, but he eventually found a solution and published two further installments early in 1862.[34]

Just as in part 2, the innovations of part 3 – entitled 'The theory of molecular vortices applied to statical electricity' – were set in train by consideration of a problem in the mechanical functioning of the magneto-electric medium. In electromagnetic induction, when the angular velocities of the vortices are changing in response to torques exerted at their surfaces, how are the changes in motion transmitted from the outer strata of the vortices to the interiors? If the fluid medium were frictionless, neighbouring strata of the vortex would slip aginst each other without interacting, and changes in motion would not be transmitted. If the fluid were viscous, changes in motion would be transmitted but would be accompanied by conversion of some of the rotational energy of the vortex into heat; such frictional losses of magnetic field energy were unacceptable in a realistic theory. Maxwell's solution to this problem was to endow the material in the cells with elastic properties. He now envisaged each 'molecular vortex' as a rotating, approximately spherical parcel or blob of elastic material. The tendency of these rotating blobs to bulge equatorially and flatten at the poles would still give rise to the appropriate magnetic forces, and the account of electromagnetic induction was now much improved, as changes in rotational motion would be transmitted throughout a given vortex by nondissipative elastic shear stresses.[35]

The assignment of elastic properties to the magneto-electric medium also provided for the modification of Ampère's law and thereby made possible the extension of the theory to electrostatics. Consider equation (8.1b), relating the flux of idle-wheel particles $\boldsymbol{\iota}$ to the angular velocities of the vortices $\boldsymbol{\omega}^*$. This equation followed from the assumption that each vortex, approximated as spherical in shape, had uniform angular velocity throughout, and thus rotated as a rigid sphere. The molecular vortices were now, however, regarded as rotating elastic spheres, and, as Maxwell put it, the equation would now have to be 'correct[ed] . . . for the effect due to the elasticity of the medium'. This was accomplished by inserting a correction term in the equation:

$$\boldsymbol{\iota} = (1/4\pi) \text{ curl } \boldsymbol{\omega}^* + \boldsymbol{\iota}_{elas\ def} \quad (8.4a)$$

where $\boldsymbol{\iota}_{elas\ def}$ represents the flux of idle-wheel particles attributable to progressive elastic deformation of the elastic vortex spheres. These deformations entailed elastic reaction forces, exerted by the vortex spheres on the idle-wheel particles; characterising these by a vector field $\boldsymbol{\tau}_2$, and denoting the shear modulus of the elastic medium as m, Maxwell calculated the following correction term:

$$\boldsymbol{\iota}_{elas\ def} = -(1/4\pi^2 m) d\boldsymbol{\tau}_2/dt \tag{8.4b}$$

The final form for the equation giving the flux density of idle-wheel particles $\boldsymbol{\iota}$ was[36]

$$\boldsymbol{\iota} = (1/4\pi)\ [\text{curl}\ \boldsymbol{\omega}^* - (1/\pi m) d\boldsymbol{\tau}_2/dt] \tag{8.4c}$$

Maxwell apparently regarded this last as primarily a mechanical equation, rooted in the mechanical conceptions of the theory of molecular vortices, and finding its use in the extension of that theory to embrace electrostatics. (A corresponding electromagnetic equation was of course implied, but Maxwell did not stress this point. He did discuss the electromagnetic significance of the correction term itself, but this discussion was somewhat obscure, and its interpretation is problematic. Only later was the electromagnetic significance of this 'displacement current' term clarified [Whittaker, 1951:187–9; Bromberg, 1967, 1968*a*, 1968*b*; Hesse, 1973].)

The significance of equation (8.4c) for the theory of molecular vortices lay in the fact that the divergence of $\boldsymbol{\iota}$, the flux of idle-wheel particles, could now take on nonzero values, depending on the spatial pattern of the elastic forces $\boldsymbol{\tau}_2$:

$$\text{div}\ \boldsymbol{\iota} = -\frac{1}{4\pi^2 m}\frac{d}{dt}(\text{div}\ \boldsymbol{\tau}_2) \tag{8.4d}$$

Schema (8.3) could now be invoked, with the conclusion that the excess density of idle-wheel particles ρ_p would correspond to the electric charge density ρ. Furthermore, Maxwell was able to calculate, from equations (8.4d) and (8.3b), that ρ_p would be related to the pattern of elastic stresses $\boldsymbol{\tau}_2$ as follows:

$$\rho_p = \frac{1}{4\pi(\pi m)}\ \text{div}\ \boldsymbol{\tau}_2 \tag{8.5a}$$

where m again is the shear modulus of the medium. This equation is similar in form to the electromagnetic equation relating the electric charge density ρ to the electrostatic field $\mathbf{E}_2$:

$$\rho = \frac{1}{4\pi k^2}\ \text{div}\ \mathbf{E}_2 \tag{8.5b}$$

where k is an electromagnetic constant whose signficance will emerge presently. Given the established correspondence between ρ_p and ρ, the conclusion could then be drawn that the elastic stresses $\boldsymbol{\tau}_2$ correspond to the electrostatic field $\mathbf{E}_2$, with an associated correspondence between the shear modulus of the medium m and the electromagnetic constant k:[37]

Given the already established correspondence

$$\rho_p \text{ corresponds to } \rho \tag{8.5i}$$

and given the parallel mechanical and electromagnetic relationships

$$\rho_p = \frac{1}{4\pi(\pi m)}\ \text{div}\ \boldsymbol{\tau}_2 \tag{8.5a}$$

$$\rho = \frac{1}{4\pi k^2}\ \text{div}\ \mathbf{E}_2 \tag{8.5b}$$

it is concluded that

$$\begin{aligned} &\boldsymbol{\tau}_2 \text{ corresponds to } \mathbf{E}_2 \\ &\pi m \text{ corresponds to } k^2 \end{aligned} \tag{8.5f}$$

With this, Maxwell had accomplished his goal of formulating a comprehensive and coherent mechanical field theory: he had identified mechanical correlates for all of the principal electromagnetic variables and constants, and had done this according to a consistent procedure – as in schemata (8.1), (8.2), (8.3), and (8.5) – which ensured the coherence of the theory. Maxwell went on to show that the elastic stresses in the medium that were associated with the electrostatic field would give rise to the appropriate forces on electric charges.[38] Taken in conjunction with his earlier derivation of magnetic forces, this satisfied his goal of accounting for electric and magnetic forces in terms of stresses in the medium. The reward for the fulfilment of these two of Maxwell's basic goals was a theory that encompassed all electromagnetic phenomena and extended to the propagation of light as well.

The key to the treatment of the propagation of light in the theory of molecular vortices was the incorporation of the ratio of electrical units into the theory. This ratio expresses the relative strengths of electric and magnetic forces, as attributable to the presence or passage of a given amount of charge. Having developed a theory that treated both electric and magnetic forces and related them to each other through a chain of corresponding electromagnetic and mechanical equations – schemata (8.1), (8.2), (8.3), and (8.5) – Maxwell was able to connect the ratio of units with the mechanical parameters of that theory. This was accomplished in connection with the constant k in equation (8.5b), and the conclusion was as follows:[39]

$$\sqrt{m/\rho_m} = \textit{ratio of units} = 310{,}740{,}000{,}000\ \text{mm/sec} \tag{8.6a}$$

where m is the shear modulus of the medium, which controls the strength of electric forces; ρ_m is the mass density of the medium, which controls the strength of magnetic forces; the square root arises from the quadratic dependence of force on charge; and the numerical value in millimeters per second was taken from Wilhelm Weber's measurement of the ratio of units, as published in 1857.

Maxwell was now in a position to calculate the velocity, V, of transverse elastic waves propagating in the magneto-electric medium:

$$\begin{aligned} V = \sqrt{m/\rho_m} &= 310{,}740{,}000{,}000 \text{ mm/sec} \\ &= 193{,}088 \text{ mi/sec} \end{aligned} \tag{8.6b}$$

Given the agreement of this result with the measured velocity of light within about 1 percent, Maxwell concluded that these transverse undulations of the magneto-electric medium were to be identified as light waves: '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*'.[40]

This identification of the magneto-electric and luminiferous media was not yet an electromagnetic theory of light, for the 'transverse undulations' constituting light waves were not given any definite interpretation in terms of electromagnetic variables; what Maxwell had accomplished was not a reduction of optics to electricity and magnetism but rather a reduction of both to the mechanics of one ether. The unification of optics with electricity and magnetism on this basis bore immediate fruits in Maxwell's explanation of the Faraday rotation and his prediction of a relationship between refractive index and dielectric constant. The success of this unification also served to confirm Maxwell's belief in the reality of his molecular vortices, in three ways: First, the magneto-electric medium, having been identified with the luminiferous ether, now partook of the acknowledged realistic character of the latter; second, certain technical aspects of Maxwell's treatment of the Faraday rotation strengthened the conclusion that 'magnetism is really a phenomenon of rotation'; and finally, the general success of the whole theory, in scope and explanatory power, inspired confidence.[41] It was in this atmosphere of confidence in ether theory that Thomson's speculations concerning 'vortex atoms' were most favourably received.

Thomson's vortex atom and the universal ether

Thomson had broached the plan of understanding the properties associated with ordinary matter as arising in connection with certain patterns of motion, 'vortical or other', in the universal plenum, in the late 1850s, but certain difficulties had presented themselves: 'I see no possibility however of explaining the constancy of the qualities of particular substances on this hypothesis, and I see no opening for a successful investigation on dynamical principles of any of the motion that would result from the supposed circumstances'. One of the chief properties of atoms was their durability, and vortices or 'eddies' in a fluid did not seem to have the requisite permanence. Furthermore, the dynamics of rotational motion in fluids as hitherto developed was not sufficient for investigating this problem. The necessary theorems concerning vortex motion, which showed that certain vortex motions would have the kind of permanence that Thomson was looking for, were published by Hermann von Helmholtz in 1858 and reached Thomson by 1859, but it was not until 1867 that he perceived the relevance of Helmholtz's work to his own concerns. Peter Guthrie Tait, professor of natural philosophy at the University of Edinburgh and Thomson's close friend and collaborator, had taken a great interest in Helmholtz's paper on vortex motion, having translated it for his own use in 1858. Helmholtz had shown that in a perfect fluid having no viscosity or internal friction a vortex filament (i.e., a long thin region of vortical motion, the rotation being around the long axis) would have certain interesting properties. First, the vortex filament could not end within the fluid; it could have termini on the surface of the fluid or it could close back on itself, forming a vortex loop or ring. Second, a vortex filament or loop within such a perfect fluid would persist eternally, with no alteration in the strength of its vorticity. Third, vortex filaments or rings would in general move through the fluid, but vortex filaments would not be able to pass through each other – they would behave as if they repelled each other very strongly at close distances of approach. These laws of vortex motion in a perfect fluid are not subject to direct experimental test, because ordinary fluids have nonzero viscosity. Nevertheless, smoke rings in air exhibit the basic properties. They constitute vortex filaments closing back on themselves; though not eternal they persist for an impressively long time; and they move through the air in such a way that the filaments do not cross, so that two smoke rings approaching each other will be seen to rebound, being set in vibration but not destroyed by the interaction. Tait had devised a demonstration apparatus to display these effects, involving boxes containing smoke generators and fitted with rubber diaphragms, which when struck sharply propelled perfectly circular smoke rings a foot in diameter out of circular holes in the boxes. Thomson visited Tait in January of 1867 and was treated in Tait's lecture room to a 'magnificent display' of the properties of vortex rings. Apparently as a result of seeing this display, Thomson realised that Helmholtz's laws of vortex motion in a perfect fluid provided solutions to the problems of permanence and difficulty of dynamical treatment that had obstructed his earlier speculations about vortex atoms.[42]

Within a few days, Thomson had written to Helmholtz concerning the 'magnificent display' and the conclusions he had drawn:

> Just now, however, *Wirbelbewegungen* [vortex motions] have displaced everything else, since a few days ago Tait showed me in Edinburgh a magnificent way of producing them . . . The absolute permanence of the rotation . . . in a perfect fluid, shows that if there is a perfect fluid all through space, constituting the substance of all matter, a vortex-ring would be as permanent as the solid hard atoms assumed by Lucretius and his followers (and predecessors) to account for the permanent properties of bodies (as gold, lead, etc.) and the differences of their characters. Thus if two vortex-rings were once created in a perfect fluid, passing through one another like links in a chain, they could never come into collision, or break one another, they would form an indestructible atom; every variety of combinations might exist . . . a long chain of vortex rings . . . three rings, each running through each of the other . . .[43]

Expanding on this theme in a paper entitled 'On vortex atoms', which was presented to the Royal Society of Edinburgh and published widely, Thomson wrote that he was 'inevitably' led to 'the idea that Helmholtz's rings are the only true atoms'. Thomson also discussed three further aspects of the theory of vortex atoms. First, primitive elastic or distance forces between molecules, which had necessarily been assumed in the kinetic theory of gases, as developed most recently by August Krönig, Rudolph Clausius, and Maxwell, could now be replaced by the 'kinetic elasticity' of vortex rings, thereby removing anything of an arbitrary or ad hoc nature from the kinetic theory of gases. Second, the complex patterns in the optical spectra produced by heated gases were now to be explained by the vibrations of the vortex-atom rings. Finally, the solid state could also be explained by vortex atoms: An array of 'closely packed vortex-atoms . . . must produce in the aggregate an elasticity agreeing with the elasticity of real solids'.[44]

In the decades after 1867, Thomson and a growing band of followers sang the praises of the vortex atom and discussed various extensions of its explanatory power. Thomson's conjecture that 'closely packed vortex-atoms' would have the properties of an elastic solid was the key to the assimilation of the luminiferous ether to Thomson's universal plenum with its vortex motion, and he reported his continued application to this problem, as well as some progress, in 1887. Gravity would have to be comprehended within the theory if it were ever to achieve completeness, and Thomson in 1871 considered adapting LeSage's hypothesis of ultramundane corpuscles to a theoretical framework based on the vortex atom. Fruitful application of the theory to chemistry was promised by J. J. Thomson's work of 1882, in which he in-

vestigated the stability of groupings of vortex rings and found what appeared to be an explanation of the regular variation of valency displayed in Mendeleyev's periodic table. None of these developments of the theory was quantitatively conclusive or physically unassailable, but they were all suggestive of the possibilities of the approach, and were positively received in that vein.[45]

Maxwell was no shallow enthusiast, and he was at his most sober and critical in the articles on various physical subjects that he wrote for the ninth edition (1875) of the *Encyclopaedia Britannica*. It is thus most significant that in the article 'Atom', which was intended to transmit a considered, authoritative view of the subject to a broad scientific and lay public, Maxwell was very enthusiastic concerning the vortex atom, devoting fully one-quarter of the article to it, even while admitting that it was 'an infant theory', in fact more a programme than a theory. As far as Maxwell was concerned, the glory of this theory was its freedom from any arbitrary element; it represented a form of mechanical explanation that would start from the laws of mechanics and the assumption of the existence of an inviscid and incompressible universal plenum, and would proceed to explain all phenomena of the physical universe deductively, without the aid of any auxiliary assumptions. Maxwell is worth quoting at some length:

> But the greatest recommendation of this theory, from a philosophical point of view, is that its success in explaining phenomena does not depend on the ingenuity with which its contrivers 'save appearances', by introducing first one hypothetical force and then another. When the vortex atom is once set in motion, all its properties are absolutely fixed and determined by the laws of motion of the primitive fluid, which are fully expressed in the fundamental equations. The disciple of Lucretius may cut and carve his solid atoms in the hope of getting them to combine into worlds; the follower of Boscovich may imagine new laws of force to meet the requirements of each new phenomenon; but he who dares to plant his feet in the path opened up by Helmholtz and Thomson has no such resources. His primitive fluid has no other properties than inertia, invariable density, and perfect mobility, and the method by which the motion of this fluid is to be traced is pure mathematical analysis. The difficulties of this method are enormous, but the glory of surmounting them would be unique.[46]

Tait was another enthusiastic advocate of the vortex atom; less distinguished than Maxwell, he was also much more prolific in his popular writings and probably reached a much wider audience. A mathematician of some significance, Tait, considering the different ways in which vortex atoms could be linked and knotted, developed a classification of knots that stands as one of the pioneer efforts in modern topology. In his popular writings, Tait presented the idea of the vortex atom as 'by far the most fruitful in consequences of all the suggestions that have hitherto been made as to the ultimate nature of matter'. For Tait as for Thomson, the great goal was an ultimate theory in which there would be no primitive, static forces, but rather only the effects of motion. Tait expressed this by the dictum that all energy is ultimately kinetic, there being no real potential energy in nature. Tait supported this conclusion by a weak argument based on dimensional analysis, and he also invoked the LeSage–Thomson explanation of gravitation, which he apparently took as representative of the nature of all forces.[47]

By the 1880s, the doctrine of the universal ether had become firmly established in Victorian physics, and this doctrine was clearly stated by a number of spokesmen. George Fitzgerald, who was centrally concerned with ether theory through the 1880s and 1890s, saw both matter and ether as manifestations of motion in the universal medium. Two kinds of vortex motion were envisaged: On the one hand, there would be closed vortex rings – localised vortex motions – and these would constitute matter; on the other hand, there would be vortex filaments strung out all through the universal fluid, and these would confer on the fluid as a whole the properties of ether, the bearer of light and various forces. 'This hypothesis', wrote Fitzgerald, echoing Thomson's dynamical programme, 'explains the differences in Nature as differences of motion. If it be true, ether, matter, gold, air, wood, brains are but different motions'. Thus all of nature was to be reduced to motions in the universal plenum.[48]

Oliver Lodge, one of the most vociferous proponents of the universal ether, from the 1880s well into the twentieth century, had a somewhat different articulation of this theory. Lodge placed great emphasis on the requirement that the universal plenum be 'a perfectly homogeneous . . . continuous body incapable of being resolved into simple elements or atoms; it is, in fact, continuous, not molecular'. The treatment of ether at least phenomenologically as a continuous medium whose properties are not to be deduced from considerations of molecular structure had a substantial history in Britain, and Maxwell had argued specifically that the perfect fluid that is the substratum of vortex atoms cannot be molecular, for a molecular fluid would allow for internal diffusion of momentum, and hence viscosity. Now, as Lodge pointed out, a perfectly continuous material medium is something beyond the range of ordinary experience: 'There is no other body of which we can say this [that it is perfectly continuous], and hence the properties of ether must be somewhat different from those of ordinary matter'. Thus, although 'ether is often called a fluid, or a liquid, and again it has been called a solid . . . none of

these names are very much good; all of these are molecular groupings, and therefore not like ether'. The universal ether, then, was a material *sui generis*, not classifiable as either solid or fluid, and Lodge had no difficulty in imagining that this universal plenum could have the rigidity required for the transmission of light waves and at the same time the perfect fluidity required for the existence of permanent vortex atoms. Thus, whereas for Fitzgerald both ether and matter were emergent from the universal plenum as a result of its various motions, for Lodge ether was to be strictly identified with the universal plenum itself, and matter was a manifestation of vortex motion in the plenum. Strict definition, then, would allow us to denote only Lodge's as a theory of the *universal ether;* Fitzgerald's would be a theory of the *universal plenum* or *medium.* In any case, Lodge agreed with Fitzgerald that all nature was emergent from this universal medium:

> One continuous substance filling all space: which can vibrate as light; which can be sheared into positive and negative electricity; which in whirls constitutes matter; and which transmits by continuity, and not by impact, every action and reaction of which matter is capable. This is the modern view of the ether and its functions.[49]

In the United States, the scientific community and especially the popular science tradition were quite closely tied to British examples; A. A. Michelson proclaimed the universal ether basically in Lodge's form:

> Suppose that an ether strain corresponds to an electric charge, an ether displacement to the electric current, these ether vortices to the atoms – if we continue these suppositions, we arrive at what may be one of the grandest generalizations of modern science – of which we are tempted to say that it ought to be true even if it is not – namely, that all phenomena of the physical universe are only different manifestations of the various modes of motion of one all-pervading ether.[50]

The retreat

If Maxwell, in his work on the theory of molecular vortices in 1861–2, had been a leader in the movement towards a definite, concrete, and realistically intended mechanical account of the universal ether, he also very soon, beginning in 1864, played a central role in the subsequent retreat from that kind of ether theory. As we shall see, it was a tactical retreat that Maxwell undertook in 1864, but this soon developed into a long-term strategy, a strategy of concentration on *dynamical theories* of ether, which involved much less commitment to specific mechanical pictures.

The theory of molecular vortices had been extremely successful on a heuristic level, and Maxwell remained forever committed to the reality of the molecular vortices themselves, but he had some problems with the idle-wheel particles, whose existence he had been led to postulate in the context of the particular logic of theory construction he had employed in 'Physical lines'. Maxwell had felt, from the outset, that the conception of idle-wheel particles was perhaps too ingenious, too artificial; and he had to admit that this was a 'provisional', 'awkward', and 'temporary' hypothesis. Also, the idle-wheel particles, according to the correspondences established in the theory, were taken to be electrical particles or electrical matter, and the notion of electrical particles or substance was repugnant to the field theories of both Faraday and Maxwell.[51] Maxwell, then, was highly motivated to avoid somehow the hypothesis of idle-wheel particles. Beyond this, as he wrote to a friend at the time, he wanted to 'clear . . . the electromagnetic theory of light from any unwarrantable assumption, so that we may safely determine the velocity of light by measuring the attraction between bodies kept at a given difference of potential, the value of which is known in electromagnetic measure'.[52] Thus, irrespective of the combination of commitment and doubt that characterised his own personal stance towards the concrete mechanical theory of 'Physical lines', Maxwell felt that it would be advantageous to connect the experiments on the ratio of units with measurements of the velocity of light in the most direct way possible, without the intervention of any hypothetical elements. To accomplish this, Maxwell used what he called the approach of 'dynamical theory'; it was presented in rudimentary form in 'A dynamical theory of the electromagnetic field' (1864), and in mature form in the *Treatise on electricity and magnetism* (1873).[53]

In part, Maxwell followed Thomson's definition of a dynamical theory:

> The theory I propose may . . . be called a theory of the *Electromagnetic Field,* because it has to do with the space in the neighbourhood of the electric or magnetic bodies, and it may be called a *Dynamical* Theory, because it assumes that in that space there is matter in motion, by which the observed electromagnetic phenomena are produced.

But there was also another meaning of the term *dynamical,* as Maxwell used it, that is even more relevant in the present context: It referred to abstract or generalised dynamics as developed by Joseph Louis Lagrange and further elaborated by William Rowan Hamilton and also by Thomson and Tait in their *Treatise on natural philosophy* (1867). Lagrangian methods had been used earlier in studies of the luminiferous ether, most notably by James MacCullagh and George Green, whose approach had been to postulate an appropriate potential energy function for ether without specifying the details

of its mechanical structure; a wave equation was then derived using the Lagrangian variational principle. In this way, ether could be treated as a mechanical system, without a complete specification of the 'unseen . . . machinery' that gave it the characteristics exhibited in the potential energy function. The phenomenological aspect of this approach was reinforced by parallel trends in thermodynamics, and the energy approach was successfully integrated with the Lagrangian formalism by Thomson and Tait in their *Treatise on natural philosophy.*[54]

In the *Treatise on electricity and magnetism,* Maxwell used the formalism of Thomson and Tait, and expressed his aims as follows:

> What I propose now to do is to examine the consequences of the assumption that [electromagnetic] phenomena . . . are those of a moving system [ether], the motion being communicated from one part of the system to another by forces, the nature and laws of which we do not yet even attempt to define, because we can eliminate these forces from the equations of motion by the method given by Lagrange for any connected system.

Through the Lagrangian formalism, Maxwell was able to maintain his commitment to the mechanical world view and simultaneously avoid any 'unwarrantable assumption' concerning the specific mechanism of ether. He still felt that 'we have good evidence for the opinion that some phenomenon of rotation is going on in the magnetic field' and 'that this rotation is performed by a great number of very small portions of matter, each rotating on its own axis, this axis being parallel to the direction of magnetic force'. But he now preferred not to specify the picture any further, relying instead on the abstract Lagrangian formalism to give a 'dynamical' characterisation of the behaviour of ether. Thus, although he still regarded ether as a material system, he found that the possibility of knowing its detailed structure was receding.[55]

The detailed mechanical picture of ether that Maxwell had constructed still had a role to play, as a 'working model', which furnished a 'demonstration that mechanism may be imagined capable of producing a connexion mechanically equivalent to the actual connexion of the parts of the electromagnetic field'. This kind of model has recently been characterised by Peter Achinstein as an 'imaginary model'; such a model, in general, is not proposed with realistic intent, but rather to show that it is possible to imagine a specific mechanism that will display the appropriate behaviour.[56] Another such model of ether, closely related to Maxwell's, was proposed by Fitzgerald in 1885. It involved, in one version, spinning wheels connected by rubber bands, and in another, paddle wheels coupled by a fluid flowing in connecting canals. Fitzgerald carefully noted:

> I need hardly say that I do not intend it to be supposed that the ether is actually made up of wheels and indiarubber bands, nor even of paddle-wheels, with connecting canals. I think, however, that we may learn several things as to the conditions that the elements of the ether should fulfill if they are to represent Maxwell's equations by motions in ways analogous to those of my model.[57]

Perhaps most prolific in constructing these kinds of models of ether in the closing decades of the nineteenth century was William Thomson. Thomson encountered difficulties in fully developing his programme of explaining all physical phenomena realistically in terms of vortex motions in the universal substratum, and instead had recourse to unrealistic, though illustrative, mechanical models. One of the problems he encountered when attempting to understand the elasticity of ether in terms of vortex motion was that of stability. 'It is exceedingly doubtful,' he wrote in 1887, 'so far as I can judge after much anxious consideration from time to time during these last twenty years, whether the configuration . . . is stable . . . I am thus driven to admit, in conclusion, that the most favourable verdict I can ask . . . is the Scottish verdict of *not proven*'. (Stability problems impinging upon the vortex-atom theory of ordinary matter were also beginning to appear at about this time, and Thomson was beginning to harbour deep doubts concerning it as well.) Two years later, Thomson wrote that 'the difficulties in the way of proving a comprehensive dynamical theory of electricity, magnetism, and light are quite stupendous'.[58] In this situation, he turned to illustrative models to show how the elastic properties of ether could result from motion. In the words of his sympathetic Victorian biographer, 'he proceeded to describe an imaginary model ether or medium . . . To this end he imagined a sort of network, across the meshes of which were set minute gyrostats [gyroscopes] spinning each about its own axis, to give the structure the necessary immobility'. 'I do not admit that [this] is merely playing at theory', wrote Thomson, 'but it is helping our minds to think of possibilities, if by a model, however rough and impracticable, we show that a structure can be produced' having the requisite properties. He proposed various other crude models of ethereal functions, with the intention of providing illumination; however, he merely accentuated the growing gap between the needs of ether theory and the available mechanical models.[59]

Joseph Larmor was the last of the great Victorian ether theorists. A generation younger than Thomson and Maxwell, he synthesised the work of his predecessors and carried existing trends to their logical conclusions. In a series of papers entitled 'A dynamical theory of the electric and luminiferous medium', published between 1893 and 1897, Larmor set out to give an account of 'the primordial medium which is assumed to be the ultimate seat of

all phenomena'. Employing an ether similar to MacCullagh's (which had been given a concrete realisation in Thomson's gyrostatic ether) in conjunction with Maxwell's electromagnetic field theory, and incorporating the vortex atom as well, Larmor achieved a monumental synthesis. This synthesis, however, was achieved in the context of a methodological programme that represented an extreme development of traditional ether theory, and that in a sense negated traditional ether theory.[60]

Larmor followed Maxwell in his enthusiasm for the Lagrangian method, which 'allow[s] us to ignore . . . altogether the details of the mechanism; . . . it makes everything depend on a single analytical function representing the distribution of energy in the medium'. Traditionally in the Lagrangian approach to ether theory, the abstract treatment would be followed, sooner or later, by an attempt to give a concrete account of the hidden machinery that gives rise to the effects abstractly treated by the Lagrangian formalism – either in a realistic way or at least by means of an illustrative model. 'The chief representative' of the builders of such models, observed Larmor, 'has been Lord Kelvin'. Larmor, for his own part, however, suggested, 'It may be held that the [abstract treatment] really involves in itself the solution of the whole problem; . . . [the concrete model] is rather of the nature of illustration and explanation, by comparison of the intangible primordial medium with other dynamical systems of which we can directly observe the phenomena'. Thus, as so often happens in the history of ideas, a crucial change in outlook had been brought about by having the tail wag the dog: Whereas the abstract Lagrangian approach had before been regarded as a step towards the ultimate theory, which would display ether as a concrete mechanical system, the abstract theory was now to be regarded as the final and ultimate theory, concrete models being significant only for heuristic or pedagogical purposes.[61]

Larmor soon went even further in relinquishing the fundamental tenets of Victorian ether theory. In 1894 he gave up the vortex atom as the basis of ordinary matter, and instead espoused an electron theory, thus parting company with the tradition of Faraday, Maxwell, and Thomson.[62] By the close of the century, as observed by a sympathetic contemporary, 'chiefly under the influence of Larmor, it came to be generally recognized that the aether is an immaterial medium, *sui generis*, not composed of identifiable elements having definite locations in absolute space' (Whittaker, 1951:1:303).

Conclusion

Viewed in a broader context, Victorian ether theory can be seen as the embodiment of an extreme option within the mechanical world view. Mechanics deals with matter in motion, and mechanical theories could provide for a richness of possibilities by invoking either a variety of matters or a variety of motions. Newton's idea of endowing particles with active forces provided for the utilisation of a variety of matters in the eighteenth century. Thus there were electrical particles of two kinds, bearing specifically electrical forces; magnetic particles of two kinds with their specific forces; heat particles exerting short-range forces; and light particles. In the course of the nineteenth century, especially in the thinking of British physicists, this variety of particles and forces gave way to a variety of motions in one universal substratum. This substratum first appeared in the guise of the luminiferous ether, and, as a result of the work of Faraday, Thomson, and Maxwell, the phenomena of electricity and magnetism were also referred to the motions and strains of this medium. Heat had by midcentury been referred to the motions of particles of ordinary matter, and Thomson in turn viewed these as nothing but patterns of motion in the universal substratum. There resulted a form of the mechanical world view in which all natural phenomena – and perhaps the 'supernatural' as well – were to be explained by the dynamics of the universal ethereal medium.[63]

Acknowledgments

A much expanded version of the section of this chapter entitled 'Maxwell's theory of molecular vortices' will be published elsewhere; an earlier form was presented at the Dec. 1976 meeting of the History of Science Society, and discussions with coparticipants there in a session on the history of field theories, including Donald F. Moyer, Barbara Giusti Doran, and Howard Stein, have been very informative. John Heilbron, Norton Wise, and Geoffrey Cantor have read drafts of this chapter and made extremely helpful suggestions. The research was supported in part by a grant from the National Science Foundation.

Notes

1 J. Larmor, 'A dynamical theory of the electric and luminiferous medium' (1893), in J. Larmor, *Mathematical and physical papers*, 2 vols. (Cambridge, 1929), 1:389–413, on 389.

2 L. P. Williams, *Michael Faraday: a biography* (London, 1965).

3 M. Faraday, 'Thoughts on ray-vibrations' (1844), in M. Faraday, *Experimental researches in electricity* (1839–55; reprinted, 3 vols. in 2, New York, 1965), 3:447–52, first two quoted passages on 447; Williams, *Faraday*, 380–1; M. Faraday, 'On the magnetization of light and the illumination of magnetic lines of force' (1845), in Faraday, *Experimental researches*, 3:1–26; Williams, *Faraday*, 383–92; Spencer (1970); M. Faraday, 'On lines of magnetic force' (1852), in Faraday, *Experimental researches*, 3:328–70, final quoted passage on 331.

4 W. Thomson, 'On the uniform motion of heat in homogeneous solid bodies and its connection with the mathematical theory of electricity' (1842), in W. Thomson, *Reprint of papers on electrostatics and magnetism*, 2nd ed. (London, 1884), 1–14. See also Buchwald (1977).

5 J. Z. Buchwald, 'Sir William Thomson (Baron Kelvin of Largs)', in *Dictionary of scientific biography*, vol. 13 (1976), 374–88. See also S. P. Thompson, *The life of William Thomson, Baron Kelvin of Largs*, 2 vols. (London, 1910), 113–27, esp. 127. Buchwald (1977), 105, 122–3, and Thompson, *Kelvin*, 19–20, disagree in emphasis concerning Thomson's prior knowledge of and commitment to Faraday's approach.

6 W. Thomson, 'On the elementary laws of statical electricity' (1845), in Thomson, *Electrostatics and magnetism*, 15–37; Thomson 'On the uniform motion of heat', n. on 1–2; Buchwald (1977), esp. 106–7, 119–34.

7 Thomson, 'Elementary laws', 37. The nuances of Thomson's paragraph 50 have not been exhausted in my account here.

8 W. Thomson, 'On a mechanical representation of electric, magnetic, and galvanic forces' (1847), in W. Thomson, *Mathematical and physical papers*, 6 vols. (Cambridge, 1882–1911), 1:76–80.

9 Thompson, *Kelvin*, 203–4.

10 Ibid., 1062–5.

11 L. Campbell and W. Garnett, *The life of James Clerk Maxwell*, with a new preface and appendix with letters by R. H. Kargon, 2nd ed. (1882; reprinted, New York, 1969), 144–6; Thompson, *Kelvin*, 222–3.

12 J. Larmor (ed.), 'The origins of Clerk Maxwell's electrical ideas, as described in familiar letters to W. Thomson', *Proceedings of the Cambridge Philosophical Society 32* (1936), 695–750, esp. 697–705.

13 J. Clerk Maxwell, 'On Faraday's lines of force' (1855–56), in J. Clerk Maxwell, *The scientific papers of James Clerk Maxwell*, ed. W. D. Niven (1890; reprinted, 2 vols. in 1, New York, 1965), 1:155–229, on 155–9 (my italics). See also Siegel (1975); Hesse (1973); Kargon (1969); d'Agostino (1968); J. Turner (1955); Olson (1975), 299–302.

14 Maxwell, 'Faraday's lines', 156, 160 (Maxwell's italics). See also Olson (1975), 287–321; J. Clerk Maxwell, 'Are there real analogies in nature?' (presented to the Apostles at Cambridge in 1856), in Campbell and Garnett, *Maxwell*, 235–44.

15 W. Thomson, 'Dynamical illustrations of the magnetic and the helicoidal rotatory effects of transparent bodies on polarized light', *Proceedings of the Royal Society 8* (1856), 150–8; reprinted in *Philosophical Magazine 13* (1857), 198–204, on 199. The paper is discussed in Knudsen (1976), esp. 244–7 and 273–81.

16 Thomson, *Electrostatics and magnetism*, 423 ('Note added Jan. 1872') (my italics); W. Thomson, 'On an absolute thermometric scale founded on Carnot's theory of the motive power of heat, and calculated from Regnault's observations' (1848), in Thomson, *Mathematical and physical papers*, 1:100–6, on 102–3; W. Thomson, 'An account of Carnot's theory of the motive power of heat; with numerical results deduced from Regnault's experiments on steam' (1849), in ibid., 1:113–55, on 116–17; M. J. Klein, 'Gibbs on Clausius', *Historical Studies in the Physical Sciences 1* (1969), 127–49; Thompson, *Kelvin*, 263–83; W. Thomson, 'On the dynamical theory of heat, with numerical results deduced from Mr. Joule's equivalent of a thermal unit, and M. Regnault's observations on steam' (1851), in Thomson, *Mathematical and physical papers*, 1:174–232.

17 See esp. W. Thomson, 'On a universal tendency in nature to the dissipation of mechanical energy' (1852), in Thomson, *Mathematical and physical papers*, 1:511–14, on 511, for a direct opposition of 'dynamical' to 'statical'. See also, e.g., Thomson, 'Dynamical theory of heat', 174–5; Thomson, *Electrostatics and magnetism*, 423 ('Note added Jan. 1872'); and [W. Thomson], 'Dynamics', in *[Nichol's] Cyclopaedia of the physical sciences* (London, 1860), 212–15.

18 Thomson, 'Dynamical theory of heat', 174–5; W. J. M. Rankine, 'On the centrifugal theory of elasticity, as applied to gases and vapours', in *Miscellaneous scientific papers,* ed. W. J. Millar (London, 1881), 16–48, esp. 16–18; Thomson, 'Dynamical illustrations', quoted passages on 199–200. Joule at first espoused the hypothesis that heat was rotatory motion but then gave it up – see Brush (1976), 1:161.

19 Thomson, 'Dynamical illustrations', 198–200; W. Thomson, 'Atmospheric electricity (Royal Institution Friday Evening Lecture, May 18, 1860)', in Thomson, *Electrostatics and magnetism,* 208–26, quoted passage on 224 (Thomson's capitalisation and italics).

20 Thomson, 'Dynamical illustrations', 199–200, quoted passage on 200; Knudsen (1972), quoted passages on 47; Doran (1975), 179–90.

21 Campbell and Garnett, *Maxwell,* 199 n.

22 J. Clerk Maxwell, 'On physical lines of force' (1861–2), in Maxwell, *Papers,* 1:451–513, quoted passages on 505, 489.

23 Ibid., 452–3 (my italics, except for *'stress'*). In modern terminology – cf. P. Achinstein, *Concepts of science: a philosophical analysis* (Baltimore, 1968), 203–25 – the mechanical representation of 'Physical lines' might be denoted a *theoretical model,* in contrast to the *analogue models* of 'Faraday's lines'; Maxwell used the term *model* only in later work (see 'The retreat' in this chapter).

24 Maxwell, 'Physical lines', 454–66; M. Faraday, 'On the physical character of the lines of magnetic force' (1852), in Faraday, *Experimental researches,* 3:407–37, on 419, 435–6.

25 Maxwell, 'Physical lines', 467–9, quoted passages on 468; ibid., figure facing p. 488. Pseudosphericity of the cells is evidently assumed in the calculations in ibid., 469–71, and is explicitly used at 492 ff.

26 Ibid., 468–71, resulting in equations (33) and (34) on 471; my $\boldsymbol{\iota}$ represents Maxwell's p, q, r, and $\boldsymbol{\omega}^*$ his α, β, γ.

27 Ibid., 471, equations (33) and (34). This differential form of Ampère's law had appeared before in Maxwell, 'Faraday's lines', 194. Cf. Maxwell, 'Physical lines', 462, 496. My $\boldsymbol{\iota}$ and $\mathbf{J}$ both correspond to Maxwell's p, q, r, reflecting explicitly his shifting interpretation of this vector; similarly, both $\boldsymbol{\omega}^*$ and $\mathbf{H}$ correspond to his α, β, γ.

28 Maxwell, 'Physical lines', 471.

29 Ibid., 472–5, resulting in equation (54) on 475; my $\boldsymbol{\tau}_1$ represents Maxwell's P, Q, R, and ρ_m his ρ on 456, to within a multiplicative constant. Cf. ibid., 457.

30 Ibid., 475, equation (54). My $\boldsymbol{\tau}_1$ and $\mathbf{E}_1$ both correspond to Maxwell's P, Q, R, reflecting his shifting interpretation of this vector.

31 Ibid., 475–6: 'The forces exerted on the layers of particles between the vortices' represent, 'in the language of our hypothesis, . . . electromotive forces'.

32 Ibid., 485–8, quoted passage on 486. Schema (8.3) is adumbrated in ibid., 477: 'The [idle-wheel] particles . . . in our hypothesis represent electricity'. For equation (8.3b), see Maxwell, 'Faraday's lines', 191–2, where the symbol ρ was used for charge density; and Maxwell, 'Physical lines', 496, where e was used.

33 See Maxwell, 'Faraday's lines', 195; Siegel (1975).

34 Larmor, 'The origins', 728–9; C. W. F. Everitt, *James Clerk Maxwell: Physicist and natural philosopher* (New York, 1975), 98–9.

35 Maxwell, 'Physical lines', 489. Cf. ibid., 486.

36 Ibid., 490–6, quoted passage and final equation, equation (112), on 496. Cf. equation (108), ibid., 495.

37 Equation (8.4d) from equations (113) and (114), ibid., 496–7. Equation (8.5a) from equation (115), ibid., 497; cf. (108) on 495; my ρ_p corresponds to Maxwell's e. Equation (8.5b) also from (115), ibid., 497, interpreted electromagnetically, my k corresponds to Maxwell's E. Schema (8.5) esp. from (118), ibid., 497; cf. (108) on 495.

38 Ibid., 497–8.

39 Ibid., 495 (equation [108]), 499 (equation [133] and stipulation '$\mu = 1$').

40 Ibid., 499 (equation [136]), 500 (Maxwell's italics).

41 Ibid., 500–13, quoted passages on 500, 505. Cf. Bromberg (1968*a*), esp. 219, 229–30.

42 Thomson, 'Dynamical illustrations', 200; Knudsen (1972), 47–8; H. Helmholtz, 'Ueber Integrale der hydrodynamischen Gleichungen, welche den Wirbelbewegungen entsprechen', *Journal für die Reine und Angewandte Mathematik 55* (1858), 25–55; Thompson, *Kelvin,* 402; C. G. Knott, *Life and scientific work of Peter Guthrie Tait* (Cambridge, 1911), 127, 176–204; W. Thomson, 'On vortex atoms' (1867), in Thomson, *Mathematical and physical papers,* 4:1–12, quoted passage on 2; Thompson, *Kelvin,* 510–15; Silliman (1963).

43 Thompson, *Kelvin,* 513–15.

44 Thomson, 'Vortex atoms', 1–4, quoted passages on 1–2. Cf. a report of Thomson's talk in 'the *Scotsman* of February 19, 1867', in Thompson, *Kelvin,* 517–19, quoted passage on 519. 'Vortex atoms' was published in the *Philosophical Magazine* as well as the *Proceedings of the Royal Society of Edinburgh.*

45 W. Thomson, 'On the propagation of laminar motion through a turbulently moving inviscid fluid', *Philosophical Magazine 24* (1887), 342–53; W. Thomson, 'On the ultramundane corpuscles of LeSage . . .', ibid., *45* (1873), 321–45, discussed in Thompson, *Kelvin,* 1029; J. J. Thomson, *A treatise on the motion of vortex rings* (London, 1883).

46 J. Clerk Maxwell, 'Atom', in Maxwell, *Papers,* 2:445–84, discussion of vortex atom on 466–77, quoted passages on 471–2.

47 Knott, *Tait,* 105–9; P. G. Tait, *Lectures on some recent advances in physical science,* 2nd ed. (London, 1876), 290–300, 362–3, quoted passage on 290–1.

48 G. F. Fitzgerald, 'On a model illustrating some properties of the ether' (1885), in G. F. Fitzgerald, *The scientific writings of the late George Francis Fitzgerald,* ed. J. Larmor (Dublin, 1902), 142–62, on 154–6; G. F. Fitzgerald, 'Electromagnetic radiation' (1890), in ibid., 266–76, quoted passage on 276; G. F. Fitzgerald, 'Helmholtz memorial lecture' (1896), in ibid., 340–77, on 345–54.

49 O. Lodge, 'The ether and its functions', *Nature 27* (1883), 304–6, 328–30, quoted passages on 305, 330; Maxwell, 'Atom', 466–7; I. Todhunter and K. Pearson, *A history of the theory of elasticity and of the strength of materials,* 2 vols. (Cambridge, 1886–93), 1:496–505; E. M. Parkinson, 'George Gabriel Stokes', in *Dictionary of scientific biography,* vol. 12 (1976), 74–9; Wilson (1971).

50 A. A. Michelson, *Light waves and their uses* (Chicago, 1903), 161–2, cited in Silliman (1963), 473.

51 Maxwell, 'Physical lines', 486; J. Clerk Maxwell, *A treatise on electricity and magnetism,* 2 vols. (1873; 3rd ed., 1891; reprinted, New York, 1954), 1:380.

52 Campbell and Garnett, *Maxwell,* 340.

53 J. Clerk Maxwell, 'A dynamical theory of the electromagnetic field' (1864), in Maxwell, *Papers,* 1:526–97; Maxwell, *Treatise,* 2:199–262.

54 Quoted passages in Maxwell, 'Dynamical theory', 527 (Maxwell's italics); and J. Clerk Maxwell, 'On the dynamical evidence of the molecular constitution of bodies' (1875), in Maxwell, *Papers,* 2:418–38, on 419. A definitive treatment of this whole matter may be found in Moyer (1977). See also Klein (1973), 69–70.

55 Maxwell, *Treatise,* 2:198, 470. Cf. Maxwell, 'Dynamical theory', 533 (¶ [16]).

56 Achinstein, *Concepts of science.*

57 Maxwell, *Treatise,* 2:470; Achinstein, *Concepts of science,* 203–26, esp. 218–21;

Fitzgerald, 'A model', 142–51, quoted passage on 151. Cf. Fitzgerald, 'Electromagnetic radiation', 270–1.

58 Thomson, 'Propagation of laminar motion', quoted passage on 352 (Thomson's italics); Thompson, *Kelvin*, 1043, 1046–7, quoted passage on 1043.

59 Thompson, *Kelvin*, 1044–6, quoted passages on 1044–5. See also W. Thomson, 'On a gyrostatic adynamic constitution for "ether"' (1889–90), in Thomson, *Mathematical and physical papers*, 3:466–72; W. Thomson, 'Ether, electricity, and ponderable matter' (1889), in ibid., 484–511, esp. 500–11; W. Thomson, 'Steps toward a kinetic theory of matter' (1889), in W. Thomson, *Popular lectures and addresses*, 3 vols. (London and New York, 1889–91), 1:225–59, esp. 242–50; Whittaker (1951), 145; Schaffner (1972), 68–75, 194–203.

60 A. E. Woodruff, article "Joseph Larmor" in Dictionary of scientific biography; J. Larmor, 'A dynamical theory of the electric and luminiferous medium (*abstract of memoir following: and general discussion*)' (1893), in Larmor, *Papers*, 1:389–413, quoted passages on 389. Cf. J. Larmor, 'A dynamical theory of the electric and luminiferous medium: part I' (1894), in ibid., 414–535, on 414–15. For a somewhat different view of Larmor in particular and British ether theory in general, see Doran (1975).

61 Larmor, 'Dynamical theory (*abstract*)', 389–90. Cf. Larmor, 'Dynamical theory: part I', 417.

62 Larmor, 'Dynamical theory: part I', 514 ff.

63 Cf., e.g., Doran (1975); Schofield (1970); Hesse (1961), 126–225; Buchwald (1977), 134–6; B. Stewart and P. G. Tait, *The unseen universe; or, physical speculations on a future state*, 7th ed., (London, 1886); Heimann (1972); Wilson (1971), esp. 34–41.

9

German concepts of force, energy, and the electromagnetic ether: 1845–1880

M. NORTON WISE

Department of History, University of California at Los Angeles, Los Angeles, California 90024 USA

The goal of unity in nature

The notion that there should exist a single pervasive ether uniting all natural phenomena was not new in the middle of the nineteenth century. Descartes, Leibniz, and Kant long before had provided respectability for the idea of primitive matter, *Urstoff*, or the *Weltäther* as a possible ground for systematic natural philosophy. Their views may be taken to stand as well for the ultimate goals of some practising scientists of the eighteenth and early nineteenth centuries, as indicated for the British in Chapter 1 of this book. Nevertheless, precise description of empirically distinguishable phenomena, such as electricity, magnetism, and heat, had generated more specific referents than simply 'ether'. Many ethereal media – electric, magnetic, caloric, luminiferous, and gravitational – populated different regions of specialised investigation. Not until the late 1840s did the general philosophical and the specialised claims find a common foundation acceptable across the broad spectrum of physical scientists. And only then did 'ether' turn into a generally recognisable object of research.

No doubt a major factor in the acceptance of a single *Weltäther* was the explanatory power of the wave theory of light, which seemed to require a medium throughout space. By itself, however, the wave theory was too limited in its range of applications to compel general assent to, or even general concern with, a role for ether throughout nature. That situation changed rapidly around midcentury when the wave theory of light was coupled to the theory of conservation of force and the mechanical theory of heat. This newly interrelated set of ideas formed a qualitatively different context for physical theory.

Conservation of force arose from and further motivated the search for unity in nature, for the ultimate identity and interconvertibility of all natural powers. And this search had by 1845 been legitimised by manifold empirical interconversions.[1] Yet the only adequate foundation available for a physical theory of conservation was mechanics: conservation of *vis viva* (or, as the concept acquired independence from force, conservation of kinetic plus potential 'energy') and the mechanical theory of heat. The mechanical theory of heat proposed that heat in a body consists in internal *vis viva* and potential energy, and usually that heat radiated between bodies consists in some form of transmitted motion. The latter view presupposed an ether throughout space – already required by the wave theory of light – that could transmit the motions of heat as well as the closely related motions of light. As an ideal mechanical medium it would conserve *vis viva*, again a basic tenet of the wave theory. By postulating an ether, physicists could incorporate heat and light into mechanics under conservation of *vis viva*. Reversing the argument, if ether did not exist, then the mechanical foundation of light, heat, and conservation was destroyed, and with it the satisfying vision of a unified physical science. Many investigators considered the existence of a unifying medium throughout space a necessity for the rational comprehension of science as it stood at midcentury.

General assent to the existence of a single ether, however, did not imply agreement on its nature. Those concerned primarily with one set of phenomena – for example, heat, electricity, physical optics, or philosophy – constructed ether primarily to satisfy the demands of their own area, often without paying close attention to other areas but hoping eventually to subsume them. A comprehensive history of unifying ether theories after midcentury would have to begin by considering this wide variety of bases for constructing ether and then show how most of them lost significance; for by about 1880, electromagnetism, including the electromagnetic theory of light and radiant heat, had emerged as dominant. I shall not attempt that complex analysis here. I shall consider only the electromagnetic bases themselves and only as they developed in the German-language community, in which I include authors whose work was disseminated in German-language journals or who received their training in German universities. Restricting discussion of the search for unity in this way will allow a focus, first, on the reception and reformulation in Germany of French mathematical action-at-a-distance theories of electricity and magnetism and, second, on the development of mathematical field theories as alternatives to action at a distance. Both topics will serve to elucidate concepts of force that were unique to German natural philosophy. These general concepts will form the third, and overriding, theme of my discussion. A few introductory remarks on fields and on German physics at midcentury are necessary.

Fields defined

A distinction between action at a distance and field action, as it came to be generally understood in the nineteenth century, is basic to much that follows. The distinction, in the first instance, separates action directly over finite distances from action only between contiguous elements, that is, immediate (*unmittelbar*) from mediate (*mittelbar*) action between separated objects (Heilbron, this volume). In a field view, local action depends directly on local conditions. It is related to changes in conditions at a distance indirectly, through the mediating action of the field existing in the intervening space. The field, moreover, has an existence in space independent of its sources. It carries in itself the power to effect action, that is, quantity of force or energy, and propagates that power in time from point to point. A space in which force is defined at every point merely as a resultant of sources acting from all distances, such as in Laplace's gravitational theory or Poisson's electrostatics, does not qualify as a field. More stringently, an action between two objects that is merely modified (rather than mediated) by an intervening substance, where every point of the substance interacts directly with the objects, is not a field action. This stipulation is necessary in order that descriptions of polarised media in terms of forces acting directly at a distance should be distinguished from polarisation in a field by contiguous action. Examples of the former are Poisson's 1822 theory of induced magnetism,[2] William Thomson's 1845 theory of dielectric inductive capacity (Wise, 1977; Siegel, this volume), and Helmholtz's 1870 electromagnetic ether theory of light (described in a later section of this chapter). All of these authors presented their theories in opposition to contiguous action or field theories.

Fields in the nineteenth century were basically of two kinds: force fields and ether fields, depending on whether force itself was taken to be a power distributed in space (as in modern electromagnetic theory) or whether the power was carried in the state of a medium, ether. Faraday's mature theory of lines of force provides a classic example of a pure force field. Such theories, however, were less acceptable around midcentury than they had been earlier or would be later. This was, if anything, more true in Germany than elsewhere, partly for reasons that form an integral part of my story and that lead to a few general considerations on German concepts of force and ether.

Natural philosophy in German physics

German physicists[3] at midcentury continued to express intellectual concerns traditional in German Idealist philosophy while vehemently rejecting its excesses in speculative *Naturphilosophie*. The significance of this love–hate relationship with tradition is difficult to specify with precision. In an effort to sharpen its relevance for concepts of force and ether, I shall employ as a foil a man who expressed many of the values of physical science but who was not a physicist and whose work remained unacceptable to practising physicists. J. R. Mayer, a medical doctor, self-educated in physics, became well known after 1850 for his pioneering analysis of the mechanical equivalent of heat and of conservation of force.[4] His first paper on conservation was published in 1842 in Liebig's *Annalen der Chemie und Pharmacie,* following rejection by Poggendorf's physics-oriented *Annalen der Physik und Chemie.* Along with several subsequent papers, it remained largely unknown until Helmholtz began to announce widely the priority of Mayer's paper over his own classic conservation paper of 1847. In treating Mayer as a foil I shall consider first concepts that he and the practising physicists continued to employ, then aspects of their common rejection of *Naturphilosophie,* and finally the physicists' rejection of Mayer's theoretical style.

In company with many German scientists of the nineteenth century, Mayer concerned himself not merely with the physical coherence of material nature but also with the relation between matter and mind, or better, between *Natur* and *Geist,* where *Geist* implies both mind and soul. From Spinoza and Leibniz in the seventeenth century, through Kant and Hegel in the nineteenth, most German natural philosophers would have agreed with Mayer's rationalist judgement: 'What subjectively is correctly thought, is also objectively true'.[5] Matter occupied the world of nature as ideas occupied the world of mind. And where logic governed the rational relations of ideas, forces governed the interactions of matter. In the wide variety of constructions available by the mid-nineteenth century for these ideas, no single aspect is more common than the notion of force as an entity occupying a middle position between inert matter and *Geist,* between nonliving nature and the region of purposiveness and beauty, progress and freedom. Mayer posited 'three categories of existence: 1. matter, 2. force, and 3. soul, or the *geistig* principle'. Although he was among those physically oriented physiologists who separated sharply the conserved forces of nonliving nature from *Geist,* force was for him still a stepping-stone between matter and *Geist.* 'Having once attained the insight that there are not merely material objects, that there are also forces, forces in the narrow sense of the new science, just as indestructible as the matter of the chemists, then it is only a further step to the assumption and recognition of *geistig* existence'.[6]

Not all physical scientists, or even biological scientists, were willing to admit the existence of a separate spiritual realm; some, like Helmholtz, sidestepped the issue; others, like the materialist Büchner, explicitly attempted to reduce the realm of mind to that of matter and force.[7] Even such agnostic and

reductionist theorists, however, carried in Germany the weight of the tradition that Mayer more directly represented. Forces, they generally agreed, played the role in nature that relations of ideas played in mind. Forces expressed the rationality of nature. They expressed causality (Mayer, Riemann, Helmholtz) or law (Weber, Fechner, Helmholtz) or, generally, relation.

As relations, forces were often seen as something beyond the things related; they were taken not as seated in isolated matter but as arising only in the interrelations of matter. The significance of this begins to emerge when one considers that it refers force more nearly to chemical affinity – *Verwandschaft* = relationship – than to mechanical push or pull. Mayer provides an illustration for falling objects grounded in the principle of sufficient reason. 'A cause, which effects the raising of a weight, is a force; its effect, *the raised weight,* is therefore likewise *a force.* More generally expressed this is: *spatial separation [Differenz] of ponderable objects is a force*'.[8] The concept of force as a relation (and relationship) will provide the overriding theme in my discussion of German electromagnetic ether theories.

In his understanding of conservation of force, Mayer expressed another traditional idea common among his contemporaries: '*There is in truth only a single force*',[9] a force conserved through all the transformations of nature. In Germany before midcentury, however, this idea had its home in the speculative systems of *Naturphilosophie,* which Mayer, as well as most practising scientists, was at pains to combat. *Naturphilosophie* is often associated in its negative connotations with the natural philosophies of Schelling and Hegel, for whom nature did not exist independent of *Geist* and who therefore animated all nature with *Geist.*[10] But seen as the attempt to construct a priori a system of nature that would reflect the operations of mind, *Naturphilosophie* should include as well the less subjective philosophies of Kant and Herbart, who preserved an external – though unknowable – *Ding an sich* as a constraining ground for conceptualisations. As their means for construction of a system of nature the *Naturphilosophen* sought to employ an ultimate logic of mental activity, often a dialectical logic that would allow the mind to construct its concepts through irony, that is, through a mutual conflict of primitive opposites, leading to a synthesis at a higher level. The fundamental opposites in the realm of nature were powers or forces, typically attractive and repulsive, and there was given, as both product and ground of the original conflict of attraction and repulsion, a primordial ether filling space as a continuum. It served as the ground of all other forms of matter and force. In Kant's words:

> The elementary system of the moving forces of matter depends upon the existence of a *substance* which is the basis (the primordi-

ally originating moving force) of all moving forces of matter, and of which it can be said as a postulate (not as an hypothesis): There exists a *universally distributed all-penetrating* matter within the space it *occupies* or fills through repulsion, which *agitates* itself uniformly in all its parts and endlessly *persists* in this motion.[11]

The secondary forms of matter and moving force were typically conceived either as modifications of the primitive ether motions or as higher powers of the original conflict of forces. Thus they arose either dynamically or dialectically, with no clear separation between the two modes. Also, since matter and moving force were the grounds for each other, they too were not clearly distinct. Here then is a striking difference from action-at-a-distance theories. Primitive matter, like force, is distributed continuously – as matter in action-at-a-distance theories was not. Concomitantly, force, like matter, occupies space – as action-at-a-distance forces did not. One sees immediately the conceptual precedent for both continuous ether fields, filling space with energy of motion, and pure force fields, filling space with energy of attraction and repulsion.

In all systems of *Naturphilosophie*, finally, the dynamical processes continued uninterruptedly, as reflections of the rational operations of *Geist*. Kant, for example, claimed that 'the primordial forces of motion, as originally agitating, cannot bring themselves to rest, for a state of rest itself presupposes a counteraction of the agitating forces in actuality, not merely in potentiality, so that the hindrance of these motions in universal rest is self-contradictory'.[12] Ideas of this sort, widespread in early nineteenth-century Germany, form the basis in Idealist traditions for attempts such as Mayer's to enunciate conservation more precisely.

While continuing to seek the unity of nature through the unity and interconvertibility of forces, and while continuing to believe in a close parallel between rational relations of ideas and physical relations, Mayer and many of his contemporaries nevertheless rejected the notion that one could construct the true system of nature by merely developing the structure of thought. One could not know enough about *Geist*. They rejected in particular the dialectical process, with its conflict of opposites leading to a higher synthesis. And they required that all legitimate science be closely tied to empirical observations. Of particular relevance here, through an emphasis on the empirical materiality of objects, physicists separated matter from force and, like Mayer, required their independent conservation.

Though Mayer thus represents, in the traditions he rejected as well as in those he continued, a variety of common goals, he did not reject enough to suit the dominant mood of practising physicists. In its objectionable aspects

his work remained explicitly metaphysical. He employed empirical results primarily as confirming instances and attempted to establish conservation of force a priori, relying on the principle of sufficient reason, and without a rigorous treatment of *vis viva*. Helmholtz expressed the common attitude when he called this a 'metaphysically formulated pseudoproof'.[13]

Mayer fell outside the mainstream of physics in other ways. As noted earlier, most physical scientists sought to unify nature by constructing realistic physical models, usually on a mechanical basis; but even though Mayer had himself discovered the mechanical equivalent of heat, and even though he accepted a mechanical wave theory of light, he did not accept the theory that heat was mechanical. He argued that, although force in the form of heat could be interconverted with mechanical forces – with *vis viva* or a raised weight – the fact of interconversion did not justify taking any particular form of force as the fundamental one or the basis of transformations. The obvious advantage of a completely mechanical theory was that it provided just such an explanatory foundation for conversion and conservation of forces. In refusing, furthermore, to reduce heat (and electricity) to matter, motion, and the forces acting between parts of matter, Mayer maintained the *naturphilosophisch* notion of force as a sort of substance, now independent of matter but still having the same status as matter. To describe both light and heat transmitted through space, for example, he required not only a material ether to carry light waves but an independently transmitted and apparently immaterial force of heat.[14] That the rejection of *Naturphilosophie* meant precisely the rejection of such 'metaphysical' force substances in favour of the mechanics of matter helps to explain why space-occupying force fields did not receive serious consideration in Germany much before 1880. A mechanical ether provided the only legitimate basis for unity. In the last decades of the century, however, as electromagnetic fields and energy relations, described on a positivist basis, came increasingly to be considered the proper foundation for physical theory, Mayer's views were resurrected as precursors of the new trends, particularly by so-called Energeticists who made energy the basis of all reality.[15]

The following discussion of electromagnetic ether theories covers the period from 1845 to 1880, from enunciation of conservation of energy to the period just preceding full incorporation of the energy idea into pure force fields. I attempt to show how general philosophical concerns both conditioned and were conditioned by several specific electromagnetic ether theories. The result is less a coherent history of such theories in the period than a small collection of conceptual histories taken as representative of the full story.

The great questions of electromagnetism as they arose in the German context were two: How were electromagnetic forces, which seemed to require a velocity-dependent relation between portions of electrical matter, to be comprehended; and how were electromagnetism and light, apparently quite closely related, to be unified? Could ether actually perform its role as the unifying medium? Three primary examples serve to develop these questions and the major sorts of answers proposed for them. The questions arose through Wilhelm Weber's velocity- and acceleration-dependent law of electrical forces; and Weber attempted to resolve them through a comprehensive ether theory based on an action-at-a-distance interpretation of the force. Bernhard Riemann outlined, preliminarily, an equally comprehensive reinterpretation utilising a variety of ether field pictures and emphasising propagation of force in time. Hermann von Helmholtz, finally, attempted to explain apparent propagation of force within an action-at-a-distance framework, but his ether theory served instead only to mediate between the demise of action-at-a-distance theories and the rise of mature field theories. All three of these examples point up the significance in German physics of the concept of force as relation.

Weber's ether and Fechner's metaphysics

In the historiography of nineteenth-century physics we have become used to the notion of a Continental action-at-a-distance tradition formalised by early nineteenth-century Frenchmen, typically Laplace and Poisson, and adopted somewhat later by Germans such as Gauss, Wilhelm Weber, and Helmholtz (Woodruff, 1962). As the culmination of this tradition we are likely to think of Weber's law for forces acting at a distance between particles of electrical fluids. This is the law announced by Weber in 1846 that made the force between two electrical particles, e and e', depend not only on the inverse square of the distance r between them but also on their relative velocity v and relative acceleration a:

$$F = \frac{ee'}{r^2}\left(1 - \frac{1}{c^2}v^2 + \frac{2r}{c^2}a\right) \tag{9.1}$$

(Here c is a constant that Weber and Kohlrausch showed in 1857 to be approximately $\sqrt{2}$ times the velocity of light.)[16] The action-at-a-distance tradition indeed explains much about Weber's law, but in travelling from France to Germany the concept of action at a distance was transformed. Whereas Laplace and his associates described force as though it emanated from one particle of matter and acted on another particle at a distance – a description that implicitly tied force to a particle as its source – Weber insisted that force existed only as a pairwise relation between particles. The pair of particles, therefore, formed his fundamental unit of analysis. I shall begin to develop the implications of this idea through a description of Weber's programme for unity in physics, proceeding then to metaphysical foundations of the programme as presented by Gustav Theodor Fechner.[17]

Weber's electrical ether

From the first presentation of his force law in 1846, Weber conceived it as the core of an incomplete theory of an electrical ether that might eventually unify many or all natural phenomena, in the sense that it would reduce many forces to a single law of force. By 1848 he had shown that the force law could be derived from a potential, a function describing a system of particles that always acquires the same value when the particles acquire their initial position and velocities.[18] Existence of such a function guaranteed conservation of *vis viva* in the system and, therefore, conservation of all those natural powers that could be subsumed under the law of force. Thus unity and conservation were both early elements of Weber's programme. But for Weber and other empirically minded physicists unity under a specific law of force was the primary goal, conservation somewhat secondary. Only gradually did the conserved and generalised quantity of force, measured by *vis viva* and work, acquire conceptual independence as kinetic and potential 'energy'.

This view of Weber's project is consistent with the French tradition that he partially adopted. Laplace, among others, suggested that chemical affinities might possibly be explained as modifications of the inverse square force of gravity.[19] (As will appear later in this section, Weber's force law may be seen as such a modification.) Apparently following Boscovich in the attempt to make Newton agree with Leibniz, Laplace also reduced Newton's finite atoms of matter to material points so that they could never collide and so that only attractive and repulsive forces could act between them. That guaranteed conservation of *vis viva,* implying that the universe could never run down and that equality of cause and effect, or the principle of sufficient reason, would always be observed.[20] Ampère carried the argument considerably further when he attempted to construct chemical elements from geometrical arrangements of material points interacting through attractive and repulsive forces. After adopting the wave theory, Ampère also incorporated light and radiant heat as vibrations in a self-repulsive ether.[21] This ether formed an atmosphere around every point atom by attraction and extended through all space. It therefore served to transmit, as waves of heat and light, the *vis viva* of vibrating atoms, which he thought might constitute the internal heat of bodies. The same neutral ether could decompose to form positive and negative electrical fluids that, when flowing in opposite directions, would constitute an electrical current. Such double currents when flowing around an atom or molecule would form a little magnet or electrodynamic molecule, thereby accounting for magnetic materials.[22] In this physical picture, Ampère assumed that all of the ultimate forces were forces of material points that would conserve *vis viva*, but he did not attempt to formulate a theory based on interchanges of conserved quantities. His best-known law of force – the ponderomotive force between any two abstracted elements, or short sections, of current – he presented as a purely positive description of empirical observations.

Weber and his friend Fechner followed this French tradition when they reduced Ampère's abstract law of force between two currents to a physical action between point atoms of the two electrical fluids supposed to constitute both currents. The net force resulted from four interactions, all governed by the same law of force,[23] namely, the velocity- and acceleration-dependent law of force displayed in equation 9.1. The first term described Coulomb's inverse square electrostatic force between two separated electrical particles, attractive between unlike particles and repulsive between like particles. The remaining two terms modified the static law for electrodynamics, or electricity in motion, including both electromagnetic moving forces between constant velocity currents (Ampère's law) and electromotive forces induced between accelerating currents (Faraday's law). All major phenomena of electricity and magnetism were thus included under Weber's fundamental law of attraction and repulsion.

Again following Ampère, Weber believed in 1846 that the two electrical fluids in their normal unseparated state formed a neutral ether surrounding ponderable molecules and extending through all space. He hoped to be able to explain the wave theory of light on the basis of oscillations in this ether governed by the electrodynamic force law.[24] His project gained considerable credibility through association with Michael Faraday's discovery of diamagnetism in 1845 and Faraday's related discovery that diamagnetic bodies, when placed in a strong magnetic field, would rotate the plane of polarisation of light transmitted through them. Diamagnetism is the induction in normally nonmagnetic material, when it is subjected to a strong magnetic force, of a magnetic polarity opposite to that induced in normally magnetic substances, or paramagnetics. Weber explained the phenomenon successfully as induction of Ampèrian double currents in the neutral electrical atmosphere surrounding ponderable molecules.[25] These induced diamagnetic currents were similar, but opposite in direction, to the permanent currents around paramagnetic molecules. They offered a natural ground for explanation of magnetic rotation of light, particularly when the luminiferous ether was identified with the neutral electrical ether.

Following such early explanatory successes, Weber began to develop in the

1850s an increasingly comprehensive picture of the interaction of ponderable molecules with ether, a picture that became by 1880 a nearly complete electrical theory of matter, including even gravitation. It was in the course of these extensions that he departed markedly from French traditions, when he isolated and reified into a physical structure a notion that had been central even to his original force doctrine: the pairwise relation. This reified relation was the atompair, to which all actions were to be reduced. But the pair itself in Weber's conception could not be reduced; it was more than a sum of two atoms. A single 'physical point, or atom', Weber later insisted, could possess only mass and motion. More complex properties of matter, even extension, had to be considered as arising, not from properties attributed to individual atoms, but from the independent properties of pairs, existing only in their relation.[26] In 1846, Weber had thought that the acceleration-dependent term in his law of force might indicate a mediating action of ether between two atoms or even the existence of irreducible three-particle relations, because the relative acceleration of two atoms would depend directly, through Newton's second law of motion, on the action of any third particle on the first two.[27] In chemical reactions, Berzelius had named such forces 'catalytic forces', and their appearance in the electrical law of force heightened Weber's awareness of the possibility that his law might contain the reduction, long sought by Faraday and others, of chemical to electrical action. If three-particle and higher-order relations were fundamental, however, nature would be infinitely complex. Weber soon found relief from this complexity in his potential function, from which the force law could be derived in the conventional way as a gradient. The potential contained purely pairwise relations: relative position r and relative velocity v,

$$V = \frac{ee'}{r}\left(\frac{v^2}{c^2} - 1\right)$$

Believing already that the whole of a two-particle system was greater than the sum of its parts, Weber considered it extremely important that according to this potential function the 'totality of many bodies' would not give rise to new forces and new properties that could not be reduced to the properties of single particles and pairs.[28]

The simplicity alone of Weber's expression for potential probably argued for its precedence over the force law, but it is indicative of the changes taking place in physical theory around midcentury that he expressed that precedence as he assimilated his unifying force law to the general unity of energy conservation. Under the energy doctrine, potential was not only a simplifying expression for the force between two atoms; it was potential energy, the work expended against those forces in assembling the pair with given relative position and velocity. Total energy, thought of as contained in the pair, was this potential energy plus the *vis viva* (kinetic energy) of the pair. As energy was separated from moving force within the general concept of force or power, Weber came to recognise potential energy as the extra physical something in the abstract relation of a pair. In 1871 he quoted August Beer approvingly: 'In many respects one can speak with more justification of the *physical existence of work, expressed through the potential,* than of the *physical existence of a force,* of which one can only say that it *seeks to change physical relations of bodies*'.[29] A brief summary of the evolution of Weber's programme for ether will help clarify both this changing perspective and his grand plan for the unity of nature based on the atompair.

In 1852, Weber developed an explanatory model for resistance to electrical conduction in which he first described atompairs.[30] The model raised concrete problems in the relation between electrical and ponderable particles that ultimately suggested their unity. If free ether consisted of neutral pairs of positive and negative particles, Weber reasoned, then the particles of each pair should orbit about each other under the action of their attractive force. When subjected to an applied electromotive force, as inside a wire connected to a battery, such orbiting pairs would undergo successive breakup and recombination into new pairs, resulting in opposite motions of positive and negative particles along the wire. Resistance to this double current would derive from the force required to divide electrical pairs. (He did not yet discuss how the expended force disappeared.) Since all pairs were identical, the number of divisions per unit time would distinguish resistances in different materials. Conductors, nonconductors, and free ether, supposedly, would decrease in resistance with decreasing ether density. Though heuristically interesting, this model of double currents and resistance was clearly inadequate. It provided no ground for explaining different densities of ether in different materials. More problematically, it required that the double currents about ponderable molecules, those responsible for magnetism, have completely independent positive and negative components, for otherwise the component currents would resist each other and stop. Weber could only suggest that the positive and negative molecular currents moved in circles of different radii. He offered no explanation of the difference nor even an account of why electrical particles would orbit about ponderable molecules in the first place.

He did not stop long to ponder these obvious difficulties. By 1855 he had outlined the more profound problems that would occupy the remainder of his career.[31] Electrical resistance, he thought, must derive somehow from the connection of electrical particles to ponderable particles: What was the connection? Electrolysis showed a close relation between specific chemical elements and specific quantities of electrical charge: What then was the relation between chemical affinity and charge? One could add the velocity- and acceleration-dependent terms of the electrical force law to the gravitational force law without altering observable results (because $1/c^2$ was small): How close was the analogy thereby revealed between gravitational and electrical forces? By 1862, Weber had added to his unifying list the problem of converting work to heat, specifically, of converting the work done to produce electrical motions in a current into the heat of ponderable molecules, thereby conserving energy. He had also seen that this relation, as well as many other relations between ponderable molecules and electricity, might be discovered in a modified version of molecular currents.[32]

Rather than being surrounded by two currents in smaller and larger circles, molecules might have one kind of electricity, negative say, adhering to their central mass, while only positive electricity circulated about this negative nucleus and also constituted ether. That would eliminate the problem of resistance to molecular currents while simultaneously explaining resistance to conduction currents. To move positive particles from one molecule to another would require work from an applied electromotive force. Motion in conduction currents, furthermore, could be transferred to motions in the molecular currents, appearing as the heat of molecules or the heat of a ponderable body. Any disturbances of the molecular currents would cause oscillations in them at the frequency of the molecular orbits, and the oscillations would produce waves in the surrounding ether, or heat radiation. Similar wave disturbances would constitute light.

This was the model for which Weber, during the sixties and seventies, sought to build a mathematical foundation. Responding directly to the claims of energy conservation, he now made energy and energy exchange the basis of his analysis. By reexamining the states of motion of electrical pairs under the action of his fundamental force law, he found that not only two unlike particles, but also two like particles, could form a stable atompair.[33] In fact, any system of two like particles would possess two different states between which no transitions could occur except under the action of an external force. In a bound or attractive state (*Molekularbewegung*) the two particles remained always closer together than a small limiting value r_0, whereas in an alternate unbound or repulsive state (*Fernbewegung*) they moved between r_0 and infinity. Similar states existed no matter how large the masses of the particles. If, therefore, negative electrical particles were somehow united with ponderable atoms, pairs of the resulting atoms could form stable massive configurations (*Molekularbewegung*) under electrical forces alone. Similarly, positive pairs of negligible mass in the state of *Fernbewegung* could constitute ether.

With that rigorous result, Weber had broken open his set of problems, for now he could envision constructing all ponderable matter out of identical negative atoms, thereby uniting chemical atomic theory with electrical theory. The chemical identity of elements would depend on the number of negative atoms united in a massive nucleus and on their organization; magnetic, electrical, and thermal properties would depend on the states of binding between the nucleus and the light, positive atoms surrounding it.[34] Thermal radiation and light, as before, would interact with matter as waves in the positive ether. Weber conceived this ether first as a gas and later as a stable mass, but in either case it consisted of only positive pairs in the repulsive unbound state.[35] All phenomena of nature seemed to be subsumable under the new mathematical results; all, that is, but one – gravitation.

In 1875, Weber wrote the law of gravitational force in a form corresponding to his chemical atomic theory, where the mass m and charge e of ponderable atoms were proportional, $m = \alpha e$, so that

$$f = \frac{mm'}{r^2} = \alpha^2 \frac{ee'}{r^2}$$

This proportionality between the electrostatic and gravitational forces acting between all ponderable bodies suggested that the two might be integrally related, especially since the usual velocity- and acceleration-dependent terms could be added without affecting appreciably the long-range interactions of gravity. To reduce gravity to electrical forces, however, a force was required that at large distances would act attractively between *neutral* molecules, that at smaller distances would become repulsive (as in a gas), and that at yet smaller distances would again become attractive (as in chemical bonding). At the suggestion of a close associate at Leipzig, Karl Friedrich Zöllner, Weber attempted to encompass all of these factors within his force law, by employing two hypotheses: (1) that each of the identical ponderable atoms combining to form chemical elements was a neutral system consisting of a negative, but highly massive, central particle with a positive satellite of much smaller mass; and (2) that the attractive force between equal but opposite electrical particles was *greater than* the repulsive force between like particles, by a factor $(1+\alpha)$.[36] On these assumptions the net force between two neutral ponderable atoms (four interactions between two negative–positive pairs) would appear as a net attractive force between the negative component e of each pair and the positive component e' of the other:

$$F = 2\alpha \frac{ee'}{r^2}\left(1 - \frac{1}{c^2}v^2 + \frac{2r}{c^2}a\right)$$

This could be the 'gravitational' force at large distances r if α were the ratio between the large mass of a negative particle and the small mass of a positive particle.

Weber had here transformed his original neutral electrical ether of 1846 into a system consisting still of only positive and negative atoms but encompassing now all of the properties of both ether and matter. He was tantalisingly close to a single force law uniting all nature – with one qualification. The final scheme is coherent only with the understanding that force is a relation of two atoms, and its intensity depends on the character of both, for otherwise there is no basis for hypothesis (2).

Fechner's philosophical basis for action at a distance

The preceding discussion of Wilhelm Weber's evolving ether theory has displayed the coherence of form that marked his theorising throughout at least forty years of research. This coherence cannot be fully appreciated, however, without recognition of the philosophical goals he intended the theory to fulfil. Weber wrote with the traditions of German Idealism in mind. He apparently considered it a prerequisite of any valid theory that it make provision for a close connection between *Geist* and *Natur* but that it not degenerate into Materialism by reducing *Geist* to the mechanics of matter. The pairwise relation, which contributed properties to the whole beyond those of its parts, fulfilled this criterion.

> If [two] material essences, which are spatially and temporally separated, interact, then the ground of this interaction lies in the essence of both *as a whole*. The interdependent parts of this *whole* exist in different spatial and temporal points. If there are *material* essences which as *wholes* cannot be reduced to a point in space and time, that applies even more to *spiritual* [*geistig*] essences.[37]

Weber did not publish many of his thoughts on this connection between *Geist* and *Natur*, but throughout his life he associated himself most closely with people who made it their central concern. In the cases of his brother Ernst Heinrich Weber and of Fechner, both celebrated professors at Leipzig, it led to research in psychophysics. For Zöllner, also at Leipzig as professor of astronomy, it led eventually to experiments in spiritualism, experiments in which the Weber brothers and Fechner took part.[38] Weber's philosophical concerns seem to be reflected most closely in the writings of Fechner, with whom Weber first developed his ideas of an electrical double current and neutral ether, and whose 1855 defence of atomism, *Ueber die physikalische und philosophische Atomenlehre*, relied heavily on interaction with Weber.[39]

In the *Atomenlehre*, Fechner sought to found metaphysical atomism on contemporary experimental and mathematical physics, on the 'presentable connections of objects [*Weltdinge*], directly and compactly summarised in their ultimate points and knots [*letzte Spitzen und Knoten*]'.[40] The ultimate points and knots were, to Fechner, physical atoms and forces. Arrayed as enemies against this atomism of the physicists he saw the dynamicists and dialecticians described earlier: Kant, Schelling, Hegel, and Herbart. To them matter was a constructed concept. Even space was phenomenal, an appearance deriving from the process of perception, and time too they ascribed to perception. In that way, *Natur* had been filled with *Geist,* in contrast to the crude, materialistic, and spiritless conceptions of atomism.

The closest approach of the dialecticians to a physicist's atomism appeared in Herbart's 'monadology',[41] a metaphysics somewhat like Leibniz's, based on a hierarchy of discrete, simple essences, called reals or monads. A high-level monad would be the soul of an organism, whereas a low-level one would appear in the physical world as an atom. All of our knowledge, for Herbart, both of external (physical) and of internal (psychological) phenomena derived from the relations between monads. Modelling his view of these relations after a notion of force as chemical affinity (*Verwandschaft,* implying kinship or sympathy in a relationship), Herbart asserted that the relations arose as a process of conflict between monads, each striving to preserve itself. The dialectic of nature, therefore, occurred not between forces but between monads, and one could have knowledge of this conflict through the conflict between one's own soul and external monads, which gave rise to discrete presentations (*Vorstellungen*) in the mind. The relations between presentations mirrored the relations of external monads and produced one's perceptions of the objective world. Space and time appeared as general aspects of the process of relation, whereas the properties of bulk matter arose from more specific relations.

Into this, Fechner proposed to inject some commonsense reality, the reality of empirical appearances reduced to their essence. Most real of all, he argued, were material objects in space and time, which could be tasted, smelled, touched, and heard, not merely once but repeatedly, and not merely by one person but by many.[42] Matter, then, was an empirical reality and should lie at the foundation of any philosophy of nature. But additionally, physical research in its modern state required that matter be conceived as atomic. The successes of chemical atomic theory, particularly, established this view for ponderable matter; and the wave theory of light, to the degree that it explained on the basis of a particulate ether such phenomena as polarisation and differential diffraction of different wavelengths, established the atomic constitution of ether. These claims could hardly have been considered the consensus of physical scientists, but Fechner had in mind specifically as representatives of

the 'new mathematical physics' Weber and the French analysts: Laplace, Poisson, Fresnel, and Cauchy. Their work in particle mechanics was only beginning to receive a serious challenge from mathematically developed continuum mechanics, as Jed Buchwald (this volume) has shown for physical optics.

Modern physics established for Fechner the practical necessity of atomism, both for normal matter and for ether. But he had still to construct for the atomistic world view a satisfactory philosophical foundation in order to present it as a coherent idea, and ideal, in the context of German traditions. Most important, he wished to establish, contrary to the dialecticians and dynamicists, that atomism was neither materialistic nor spiritless. The materialism of traditional atomistic mechanics, Fechner reasoned, resided in two notions: that atoms consisted fundamentally of extended gross matter and that forces inhered in this matter, or that forces were properties of matter.[43] On that view all phenomena of nature would derive from independently acting atoms of matter. There would be no unifying tissue tying all of nature together as an organic whole, such that the whole was more than the parts; and there would be no room for the interrelation of mind and matter.

To elevate atomism towards the realm of *Geist,* Fechner began by supposing that material atoms stripped of their relations to other atoms were actually only physical points, or real physical monads, existing in real space and time. These atoms possessed no properties of their own other than mobility. All motions of atoms and all properties of assemblies of atoms, or matter, derived from relations between atoms.[44] To the degree, then, that the concept of matter referred to properties of matter, matter was constructed from immaterial relations of point atoms. That concept allowed an immediate translation from the realm of matter to the realm of mind. Relations between atoms in inorganic nature were laws, physical laws of force; these same laws, however, represented the rule of *Geist: 'Geist* steps up and asks, what have I to do with you? And the atoms say: we spread our individualities under your unity; the law [of force] is the commander of our band, but you are the king in whose service he leads'.[45]

In the new ordering of concepts – space, time, atoms, laws, *Geist,* and ultimately God – one moved from what one could know objectively to what one could know subjectively. At the interface between the external world and the perceiving subject, objective and subjective were only two references to the same thing.[46] That was the basis of Fechner's and Ernst Weber's well-known pioneering work in empirical psychology, labelled psychophysics by Fechner. Their investigations rested on the assumption that a determinate increase in objective stimulus should produce a corresponding intensity of subjective response (e.g., pain). The entire line of argument involved a simple inversion, similar to Herbart's, of the usual dialectic of conflicting attractive and repulsive forces, which had produced continuous matter throughout space. Fechner supposed instead that the 'dialectic' of discrete atoms gave rise to forces, that forces were the objective relations between atoms, and that this relation either constituted or measured subjective perception.[47]

Fechner's world view provided not so much a complete philosophy of nature as a programme for quantitative research, the programme that he and Ernst Weber pursued primarily in psychophysics and that Wilhelm Weber pursued in physics. Beginning with an empirically and theoretically derived law of pairwise atomic interactions, the force law or the law of potential energy, one would attempt to construct the empirical properties of matter, of ether, and of their interactions. Fechner believed that a single kind of atom, or monad, would serve to explain all of these phenomena; qualitative differences would depend only on different groupings of the atoms. Already in 1855 one could conceive of reducing magnetism, heat, and light to motions and relations of electrical matter; but were both attractive and repulsive forces, and both positive and negative electricities, necessary? What of gravitation; and what of mass itself as a quality of atoms? Fechner took hope in the increasing range of Weber's pairwise force law and attempted to uncover its ultimate implications. If the pairwise law alone proved insufficient, however, he was prepared, like Weber, to consider higher-order relations of three or more particles and higher-order relations of relations.[48]

I have described the Weber–Fechner action-at-a-distance theory of matter and ether at some length in order to show, first, the broad range of applications of the theory envisioned by Weber and, second, its metaphysical significance as presented by Fechner. Both considerations should indicate, third, that when critics of action at a distance charged that such action was incomprehensible, because one could not imagine how the action could reach from one body to another, they were missing much of what action at a distance meant to its German adherents. One body did not simply act on another – they interacted in a relationship – and the force existed as the law of relation. Nevertheless, Weber's law of force severely strained the plausibility of transforming an abstract relation into a real entity. It made of force a relation dependent not only on the distance between two atoms but also on the rate of change of that distance (velocity) and on the rate of change of the rate (acceleration). It is one thing to imagine a relation between two atoms at an instant; but it is quite another to imagine that relation depending implicitly on just prior and just following instants, as Weber's law required. Weber and Fechner found little

difficulty in this, apparently because they so closely identified force with a logical relation in which space and time had similar status. A number of other physicists, however, sought a more physical explanation for time dependence. It indicated to them that force did not act instantaneously between two bodies but took time to propagate from one to the other.

Already in 1846, upon reading Weber's first electrodynamics paper, Carl Friedrich Gauss, Weber's close associate at Göttingen, remarked that he had earlier made a similar investigation himself but had lacked what he regarded as the keystone: 'namely, *the derivation* of the additional forces [added to electrostatic force] from *non-instantaneous* actions, actions propagating in time (in a manner similar to that of light)'.[49] This association between electricity and light became all the more pressing with realisation that the constant c in Weber's law was close to the velocity of light. Alternatives to Weber's theory of the electrical ether, therefore, often proposed that electrical action propagated at the speed of light and, in fact, constituted light.

Propagation of force and early field theories

For force to propagate from one place to another, on analogy with the wave theory of light, it must be describable not simply as an abstract relation but as something existing objectively in space and distributed throughout space. That was the judgement of those who saw in Weber's implicitly time-dependent law of action at a distance the denial altogether of such action. Their alternatives were what soon came to be called field theories: both ether fields, in which force actually consisted in a form of motion in ether, and pure force fields, in which force itself propagated independently through space. Although physically quite different alternatives, ether fields and force fields arose in the same context and emphasised the same descriptive foundation, namely direct descriptions, dynamic relations, and partial differential equations.

A partial differential equation in space and time may be seen as describing the behaviour in time of any and all infinitesimal elements of space, elements taken to be characterised by one or more properties. Because the equation describes the change of these properties across an element it also describes their change between elements, and it traces that change in time. It therefore expresses naturally any continuous process of propagation from element to element through a field of properties, which might be properties of ether or properties of force. Partial differential equations, then, served as a broad path for describing directly the processes of nature while ostensibly avoiding speculation. When taken also as the most fundamental description, they served as the ground for continuum theories of ether and force. And they sometimes served as a mathematical rationale for maintaining the emphasis of *Naturphilosophie* on nature as continuous dynamic process.

Field theory was not established independently as an ongoing research tradition in Germany much before 1880; it existed more as an undercurrent that surfaced along with the notion of energy as an independent, conserved entity and in response to an action-at-a-distance interpretation by Helmholtz in 1870 of Maxwell's field theory. For that reason I present here only a cursory summary of several early field theories in order to establish their character. I consider Helmholtz's theory in a following section as a critical articulation that helped to motivate mature field theory.

Ether fields

The ether field for electromagnetism arose first, and with enduring characteristics, directly under the gaze of Weber himself in the work of his young student, assistant, and friend at Göttingen, Bernhard Riemann.[50] Riemann was also a student of Gauss and a friend of one of the most luminous exponents of differential equations applied to physics, Lejeune Dirichlet. Riemann shared Fechner's taste for metaphysics and for unifying all natural phenomena on a common physical basis, though from a different perspective. Basing his position on the epistemology and psychology of Herbart – as opposed to Herbart's *Naturphilosophie* of monads, which had been important for Fechner – Riemann argued that only relations or states could serve as causes of action, or forces, for only something subject to change of degree could itself be the cause of such change: 'What an agent strives to effect must be determined through the concept of the agent'.[51] Being (of things) was not subject to change of degree; therefore things could not be causes, only relations or states could be.

Some perspective on Riemann's view can be obtained by observing that Herbart had considered perception to consist of reception by the mind of discrete presentations, followed by 'sinking' of the presentations below the threshold of consciousness. This sinking derived from mutual suppression of opposed presentations through the forces they exerted on one another. Applying to perception the notion of force that he employed in his discussion of monads, Herbart regarded the forces between presentations as affinities, grounded not in the presentations but in their relations. He supposed also that the relations established themselves in time, approaching only asymptotically a static balance of forces. Perception was thus a dynamic process, and Herbart attempted to describe it through differential equations in time. Riemann seized on this aspect of Herbart's psychology and made continuity and differential equations the starting point for both psychology and physics. 'We observe', noted Riemann in 1853, 'a continuous activity of our soul'. This activity consisted in presentations continuously disappearing from consciousness and yet remaining as part of the substance of the soul. By analogy, he reasoned, gravitation might consist in a continuous flow of imponderable space-filling *Stoff* (substance or ether) into ponderable atoms. In fact, 'both hypotheses may be replaced by the one, that in all ponderable atoms substance perpetually enters the spiritual world from the world of body'.[52] In this way action at a distance would not exist; gravitation might be thought of as dependent on the pressure of ether immediately surrounding ponderable atoms, where pressure in turn depended on velocity of the ether. The same ether could serve to propagate the oscillations that we perceive as light and heat.

Riemann attempted initially to analyse the processes of gravitation and light in terms of resistance of a homogeneous ether to change of volume (gravitation) and to change of shape (light), the latter reducing to resistance to change of length in any physical line. He hoped also to be able to include electricity and electromagnetism in the schema, with electrostatic inverse square force depending on change of volume and electrodynamic force on change of length. All actions, therefore, were to be actions only between 'neighbouring' or 'immediately surrounding' elements. Although Riemann did not carry much further this first glimmer of a field theory, he preserved its essential characteristics in several later attempts. He always sought to unify nature on the basis of a geometrically conceived system of continuous dynamic processes in ether, founding his description on differential equations that described the processes of relation that one could perceive directly, that is, forces and interconversions of forces. His project was probably the first attempt at a mathematically founded unified field theory, much in the spirit of Einstein's later attempts, and he assigned it more weight even than his now-famous efforts in pure mathematics.[53]

Riemann's more sophisticated field descriptions in the late 1850s rested on his general assumption that the cause of both motion and change of motion (inertia and accelerating force) of a body at any point should be sought in 'the form of motion of a substance spread continuously through the entire infinite space . . . This substance can therefore be conceived as a physical space whose points move in the geometrical one'.[54] Neglecting now any explicit correlation with the sinking of presentations in perception, Riemann proposed certain motions of a homogeneous ether that would reproduce the partial differential equations of gravitation and of light propagation. The first of these equations was a continuity equation, with a nonrotational velocity system $\mathbf{u}$ standing for force:

Divergence $\mathbf{u} = -4\pi\rho$

It stated that the net flux of ether into any volume element should be given by the ponderable mass ρ per unit volume of the element. The second equation was a wave equation for transverse oscillations in velocity w (not oscillations in displacement) propagating at the speed of light c:

$$\frac{\partial^2 w}{\partial t^2} = c^2\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2}\right)$$

Combining the two motions for gravity and light produced a well-behaved velocity function, which confirmed the possibility of uniting the two processes.

Riemann had difficulty incorporating electrostatic and electromagnetic effects into the system, but in a series of lectures delivered in 1861 he indicated that variations of density in ether might be the key. If V were the usual electrostatic potential and $\mathbf{u}$ an electrodynamic potential (vector potential), the latter could be chosen to satisfy the equation

$$\frac{\partial V}{\partial t} = \text{Divergence } \mathbf{u}$$

In that case, 'V may be regarded as the density, $\mathbf{u}$ as the flux of this ether'.[55] A stable gradient in density would then correspond to electrostatic force, whereas a time rate of change in density, and associated fluxes, would correspond to electrodynamic effects. Actually, only the rotational part of the flux system in ether was required for electrodynamics, so that the nonrotational part was presumably still left over for gravitation.

Riemann never published his speculations on the unified ether field, no doubt because he did not succeed in fully integrating the various effects of gravity, light, electricity, and magnetism. He did, however, present a paper to the Göttingen Society of Science in 1858 (published only in 1867, after his death), which helps to complete this schema. Here he proposed an electromagnetic theory of light. It returns us to Weber's electrodynamic force law and to the problem of explaining velocity- and acceleration-dependence in a relation at a distance.

'I have found', Riemann announced,

> that the electrodynamic effects of galvanic currents may be explained if one assumes that the action of one electrical mass on the rest does not occur instantaneously but propagates to them with a constant velocity (equal to the velocity of light within the limits of observational error). The differential equation for the propagation of electrical force, according to this assumption, will be the same as that for the propagation of light and radiant heat.[56]

His assumption meant that one electrical mass experienced the action of another always as an electrostatic force (inverse square force or inverse first power potential), but it experienced at time t an action produced at an earlier time t', where the time lag was the propagation time from the position of the acting particle to that of the affected particle. The potential for such a 'retarded' force obeyed a wave equation rather than the usual continuity equation of electrostatics. And Riemann provided a proof that the retarded potential, when referred only to the time of action t, would correspond to the potential of Weber's force law. There are certain inadequacies in the proof, however, and perhaps for that reason Riemann withdrew the paper from publication.[57]

In summarising the general characteristics of Riemann's programme for unifying physics, we note that he sought to replace Weber's discrete relations at a distance with continuous action between neighbouring elements, which circumvented the problem of time-dependent forces. More important, he replaced the abstract relation, force, with states and processes in ether, providing a physical basis for the relation. Electrostatic potential, for example, he ascribed to density in ether. This made variations of density the reality of force, a shift of considerable importance for the energy concept, for it spread force throughout space as potential energy. Most profoundly, by reducing causality (forces or dynamics) in the observable world to the problem of describing motions in ether (kinematics), Riemann reduced physics to descriptive mathematics, or to the geometry of motions in a 'physical space'. That pushed the problem of causes one step farther back into metaphysics, and into the dynamical principles of ether. Riemann recognised this explicitly in regard to his idea of gravitation and light as two forms of motion in ether:

> The further development of this hypothesis divides into two parts insofar as one seeks:
>
> 1. The laws of ether motion [*Stoffbewegungen*] which must be assumed for explanation of phenomena,
> 2. The causes from which these motions can be explained.
>
> The first task is a mathematical one, the second a metaphysical.[58]

He employed as a foundation for the metaphysical dynamics of ether the principle of continuity of motion and a maximum–minimum condition on the velocity potential integrated over all space (the latter exhibited only mathematically). These ideas should be compared with the very similar ideas of William Thomson, Maxwell, and other midcentury British physicists. They formed the core of what the British called 'dynamical theory' (Moyer, 1978; Wise, 1980).

In this regard, Riemann's emphasis on differential equations as the basis for positive description, coupled with energy as the basis for force, is particularly significant. Both ideas made the study of fluid motion, or hydrodynamics, a pressing research topic, even when a theorist did not set out to reduce all forces to ethereal states and motions. Helmholtz, for example, established in 1858 several important theorems in hydrodynamics while rejecting such a reduction.[59] Nevertheless, his proof that a vortex in a frictionless fluid would last forever became the basis of William Thomson's ether theory of vortex atoms (Siegel, this volume). G. R. Kirchhoff developed hydrodynamics considerably further, apparently with the intention of ultimately ridding physics of forces. Adopting basic theorems of Thomson's and Tait's *Principles of natural philosophy,* he made maximum and minimum conditions on fluid motion the foundation of his hydrodynamics.[60] This suggested a quite general approach to the dynamics of an ether that would explain observable forces. At the same time, however, Kirchhoff spread the new positivism of differential equations, denying the need for specific hypotheses on the reality of nature.[61] He did not pursue directly a generalised ether theory.

Perhaps the most sustained programme of hydrodynamic reduction was carried on by a Norwegian mathematical physicist working in near isolation at the University of Oslo. Carl Anton Bjerknes attended Dirichlet's lectures at Göttingen in 1855–6 and studied there also with Riemann.[62] Impressed with Dirichlet's proof that a solid sphere in a uniform stream of frictionless, imponderable fluid would not be dragged along by the flow, Bjerknes set out to discover whether oscillations and pulsations of such spheres, producing oscillations and pulsations in the fluid, might be affected by forces of the type usually ascribed to action at a distance. In a series of mathematical papers stretching from 1863 to the end of the century, but largely unknown outside Scandinavia, he gradually extended the scope of his theory until by 1880 he could reproduce hydrodynamically the basic forces of magnetism (including induced diamagnetism and paramagnetism), electrostatics, and electromagnetism.[63] Isolated magnetic poles and electric charges appeared as pulsations in volume, polar magnets as linear oscillations, and electric currents as rotational oscillations about the axis of the current. One troublesome aspect of this description was that it produced forces opposite to those observed in nature, giving always repulsions where attractions should have appeared. More important, however, Bjerknes's description did not keep up with the demands of electromagnetic theory as it developed after about 1870, when Maxwell's electromagnetic theory of light became widely known on the Continent.

Although hydrodynamics offered a ready avenue for reducing forces to propagated motions in an ether field and, thereby, for unifying all forces, one could equally well begin, not with the nature of ether, but with the differential

equations one wished to represent, and then construct whatever kinds of ether motions the equations seemed to require. That was the approach of Maxwell in Britain and of his nearest competitor on the Continent, Ludwig Lorenz. A professor of engineering at Copenhagen and, like Bjerknes, relatively unknown, Lorenz developed during the 1860s sophisticated mathematical descriptions of processes in physical optics, particularly of reflection and refraction at boundaries. He conceived his optical equations originally on the basis of an elastic ether, but he soon became convinced that elasticity theory was incompatible mathematically with the boundary requirements of optics. Abandoning the elastic ether, and ostensibly all ethers, he developed phenomenologically a set of wave equations for the propagation of light that behaved correctly at boundaries. But Lorenz was not content with a mathematical theory of optics; like many others he sought the physical unity of nature.

'As is well known', Lorenz began a paper in 1867, 'science in our century has succeeded in demonstrating so many connections between the different forces, between electricity and magnetism, between heat, light, molecular, and chemical forces, that one is led with a certain necessity to regard them all as *manifestations of one and the same force*'. Despite this necessity, unity still seemed distant: 'Generally, the two electricities are still viewed as electrical *fluids,* light as vibrations of the *ether,* and heat as motions of the *molecules of bodies*'.[64] Lorenz presented his own first step towards remedy with a demonstration that the motions of light waves were actually motions of electric currents and that both were propagated by contiguous action.

Beginning his analysis with general equations directly relating electric currents and electromotive forces (equations of conduction, previously derived by Kirchhoff from Weber's law), Lorenz argued that no change in observable consequences would follow if these equations, representing instantaneous action at a distance, were replaced by a similar set incorporating the assumption that electrical action traveled at the velocity of light. That reproduced Kirchhoff's equations in terms of retarded potentials, like Riemann's. From the latter equations he then derived a set of partial differential equations of the same form as his phenomenological wave equations for light, but more comprehensive. Reversing the derivation, he obtained the conduction equations by integration from the wave equations, assuming boundary conditions that he had previously found necessary for light at an interface. This proof of mathematical identity Lorenz took to be sufficient evidence for identifying light physically with electric currents and also for assuming the differential wave equations for electric currents to be the 'original and generally valid ones'. With the characteristic emphasis of field theory, he argued on the basis of this priority of the differential description for the further priority of continuously propagated action over action at a distance: 'Every action of electricity and of electrical currents in reality depends only on the electrical condition of the *immediately surrounding* elements'.[65]

The form of the interconvertible equations of light and of currents, Lorenz continued, showed that light could be regarded as '*rotational* oscillations [of current] in the interior of bodies around axes whose direction is the same as that which we regard as the direction of vibration according to the elasticity theory',[66] that is, around the axis of linear polarisation, transverse to the direction of propagation. Whereas a steady current would be a steady rotation continued along its own axis, oscillations in the current would be propagated by electromagnetic induction perpendicular to the current axis, constituting light. One sees immediately the family resemblance of this hypothesis to Maxwell's picture, apparently unknown to Lorenz, of magnetic vortices in the ether. That raises the question of what medium Lorenz imagined to be rotating. He preempted the question unambiguously: 'This conception gives scarcely any basis for maintaining the hypothesis of an ether, since one may very probably assume that in so-called empty space there is contained so much material substance that it offers sufficient substratum for the motion'.[67] But of course this is just another ether hypothesis, one in which ether and normal matter differ only in aggregation or density and not in substance.

Force fields

Ether fields provided a means for spreading force in space and thereby for escaping the problems of Weber's velocity- and acceleration-dependent law of force. But they also transformed push–pull notions of force at a point into energy states in ether, reinforcing a more general recognition of energy as an independent, conserved, entity. Coupled with emphasis on direct, positive description, this suggested an alternative field conception: Treat energy as the reality and ignore the unobservable ether. As discussed previously, such views only gradually overcame the stigma of *Naturphilosophie*, but characteristics of the emergent transition appear in a very short paper presented by Carl Neumann in 1868 to the Göttingen Society of Science.[68]

Neumann developed in a new way Riemann's earlier theory of propagated forces and retarded potentials. 'I take the liberty', Neumann asserted, 'of regarding potential [potential energy] as primary, as the characteristic motive power (*Bewegungs-Antrieb*), while conceiving forces as secondary, as the form in which that power manifests itself'. He then demonstrated that if potential propagated with the velocity of light, then the effective action of an electrostatic 'emissive potential' of one electrical mass m on another m', given by mm'/r, would be changed when the two were in relative motion into

a 'receptive potential', corresponding exactly to Weber's law of force. All of the known laws of electromagnetic action (known only for *closed* currents) likewise followed. As an indication of the generality of his approach, Neumann showed that the results were independent of whether currents were considered from a 'unitary' perspective – Weber's view, with the positive electrical fluid in motion and the negative tied to the ponderable mass of the conductor – or from a 'dualistic' perspective, with both fluids in motion. This generality would not pertain, he noted, for open circuits, a consideration soon to be emphasised by Helmholtz.

Neumann's intent seems to have been not so much to replace Weber's atomistic electrical ether with a pure force field as to defend Weber and his law against a charge by Helmholtz that velocity-dependent potentials were unacceptable from the perspective of energy conservation. He had recently taken a position at Leipzig, where he came into close association with Weber's circle of friends, and in several later publications he fortified Weber's defences. Nevertheless, the thrust of Neumann's defence ran quite counter both to Weber's premises and to his style of analysis. By taking potential (energy) as fundamental and making it propagate through space he gave to it the independent status of a force field, somewhat like J. R. Mayer's force of heat. With good reason the later Energeticists, who took energy to be the only ultimate reality, looked back to Neumann as well as to Mayer as pioneers of their viewpoint.[69] In typical fashion also for all field theorists, Neumann usually based his descriptions on differential equations rather than force laws, justifying that approach as the most positive and nonhypothetical, and looked to maximum and minimum principles for the foundation of dynamics. Here that meant that he took Hamilton's principle in mechanics to have 'unlimited validity' and used it both to derive Weber's force law from the 'receptive potential' and to show that this potential would obey, approximately, conservation of *vis viva*. Even though Neumann employed his field description only to describe the force between particles of Weber's electrical ether, he made that ether largely superfluous for the propagation of light; and even though he may have had in mind a material basis, or ether field, for his propagating potential, he made no reference to it and thereby undermined the necessity of such a material foundation for a force field.

Helmholtz: energy and electrodynamics

Field theories of electromagnetism, as stressed previously, emerged to replace action at a distance for the general community of German physicists when energy was recognised as an entity different from and more fundamental than moving force. Field theories also emerged only when it seemed apparent that no action-at-a-distance theory could avoid satisfactorily the velocity- and acceleration-dependence of Weber's law. The problems of both conceptual shifts can be observed in two famous papers by Helmholtz, his 1847 formulation of conservation of force and his 1870 alternative to Maxwell's electromagnetic field theory.

In 1847, Helmholtz set out quite consciously to build a bridge between German metaphysics (Kant's) and French physics, much as Weber and Fechner were doing at the same time.[70] Helmholtz, however, sought to exclude from nature all considerations of *Geist*, in the sense of soul and purpose, which signalled to him the excesses of *Naturphilosophie*. He wished to establish the physics of point atoms and attractive and repulsive forces as necessary, on both rational and empirical grounds, and in so doing to unite all causes in nature under the determinism of conservation of *vis viva*. Helmholtz's primary conceptual tools in this effort, I shall argue, were the concept of force as relation and a distinction between quantity and intensity.

To Helmholtz, as to Weber and Fechner, forces existed as relations: 'Motive force, as the cause of change, can be predicated only in cases involving at least two bodies spatially related to one another and is then to be defined as the effort of the two bodies to change their relative positions' (cf. the discussion of Weber in the section 'Weber's electrical ether'). To Helmholtz again these relations were rational relations, but in a restricted sense. Following Kant in the view that our knowledge of nature is scientific only to the degree that we understand its laws as necessary, he argued that forces must be invariable: 'As ultimate causes, forces which do not vary in time should be found'.[71] Even implicit time dependence, he supposed, would vitiate the comprehensibility of nature. If force did not always return to the same value for the same spatial relation, *vis viva* would not necessarily be conserved, but might be produced continuously from nothing. That would violate not only the principle of sufficient reason but also, from the empirical side, the impossibility of a perpetual motion machine. The same sort of argument applied to extended atoms, for which continuously acting rotational forces could be imagined, so that only attractive and repulsive spatial forces were allowable.

To formulate his argument verbally and mathematically, Helmholtz had continually to move back and forth between concepts of force as moving force and as *vis viva* and to reconcile the two ideas. This he accomplished with a distinction, drawn in the Kantian tradition, between quantity and intensity as two different categories of understanding. We know concepts as quantities or intensities depending on whether they are considered to add up extensively – that is, side by side (*nebeneinander*), like space – or intensively – on top of one another (*aneinander*), like density. In his construction of continuous matter from forces, Kant applied this distinction in order to demonstrate how a given quantity of matter could completely fill different spaces with different intensities, in contrast to the full-or-empty conception of Newton and Descartes.[72] Helmholtz, conceiving forces already as rational relations, took the logical distinction over into the realm of forces acting between discrete, point atoms, making Newtonian moving force the measure of intensity of force and, under various circumstances, either *vis viva*, potential, or work the measure of the conserved quantity of force.

The intriguing aspect of Helmholtz's description is that he ascribed two aspects to the single notion of a tensional force between atoms at a given spatial separation. The relation possessed at any instant an intensity of tension producing changes in spatial relation, but it also possessed a quantity of tension connecting the past history of the relation to its future, to which the present action either added or subtracted.[73] This quantity was potential. Helmholtz described it as the quantity of tensional force, or work, available in a spatial relation for future consumption, thinking of the quantity consumed between two separations R and r as 'the sum total of the intensities of all the forces which act at all distances between R and r'.[74] Tensional force consumed became *vis viva*, and conversely, *vis viva* lost became a quantity of tensional force available again for consumption.

Helmholtz was able to maintain mathematically his dual conception of the single entity force only through a nonrigorous understanding of an integral. Development of the independent energy concept removed this problem while separating intensity and quantity into force and energy. The point here, however, is that already in Helmholtz's conception potential was a real physical quantity that became increasingly real through myriad theoretical and experimental applications of the conservation law. And as it became an entity the rational notion of force as an abstract relation of atoms became less acceptable. Weber's notion of velocity- and acceleration-dependent action at a distance was unacceptable to Helmholtz on rational grounds and Helmholtz's own notion of action at a distance proved unacceptable on physical grounds.

When, in 1870, Helmholtz turned his attention to electromagnetism he faced two successful challenges to his conception of force, each of which offered a conceivable physical basis for the unity of nature (Woodruff, 1968; Hirosige, 1969). On the one hand, implacably, sat Weber's law for forces acting at a distance, which Helmholtz had finally to admit could conserve *vis viva* formally in spite of its implicit time dependence. He set as his first task uncovering its deeper flaws. On the other hand loomed the propagation theories, particularly Maxwell's elegant mathematical system, but also the several less extensive field theories already discussed. Helmholtz could not accept ether fields like Maxwell's and Riemann's, which replaced force altogether with motions in ethereal matter, apparently because he considered that matter could act only through force;[75] but neither could he accept pure force fields, because a propagating pure force, taking time to act, would be again a time-dependent force like Weber's. Yet field theories of both kinds did succeed in explaining light propagation electromagnetically, as by now seemed necessary. The retarded potential theories reproduced correct empirical laws of electrodynamics from the assumption that inverse square forces travelled at the velocity of light, and Maxwell had derived the correct velocity of light from purely electrodynamic laws. Helmholtz's second and major task, therefore, was to provide an action-at-a-distance alternative to Weber's law that would rely on non–time-dependent forces and still provide an electromagnetic theory of light propagation.

Helmholtz began his search for flaws in Weber's conception by comparing Weber's potential function with a well-known phenomenological potential of Franz Neumann (Carl Neumann's father), describing interactions between closed circuits, and with Maxwell's theory. For this purpose he put Carl Neumann's propagating potential on an equal footing with Weber's mathematically equivalent, direct one. The differences among Weber's, Franz Neumann's, and Maxwell's expressions had not previously been easy to define, but Helmholtz showed in a lucid analysis that all three could be subsumed under a single potential describing the interaction between any two current elements of lengths ds and $d\sigma$, carrying currents i and j, and separated by a distance r:[76]

$$-\frac{ij}{2c^2r}[(1+k)\cos(d\sigma, ds)+(1-k)\cos(r, ds)\cos(r, d\sigma)]\,ds\,d\sigma$$

Here k is a variable parameter. When $k=-1$, Weber's potential results, assuming that a current consists of electrical particles in motion. Franz Neumann's potential corresponds to $k=+1$ when the expression is integrated over closed circuits; and $k=0$ leads to results similar to Maxwell's. When applied only to closed circuits, k vanishes from the general expression and all three cases reduce to Neumann's. Only an investigation of open circuits could distinguish among them.

As mentioned previously, Kirchhoff had derived from Weber's law general equations for currents in extended conductors. Helmholtz now found in the same way equations of electricity in motion corresponding to his generalised potential function. From the new equations he showed that Weber's potential, and only that one among the three choices, would in certain cases of open circuits lead to unstable motions of electricity. For example, if two like electrical particles approached each other at a very high speed, their relative velocity would become infinite in a finite distance. Weber and his supporters soon suggested ways of circumventing the difficulty, and a bitter dispute followed, but Helmholtz found no reason to doubt his original judgement that 'the inadequacy of the Weberian law here brought to light is founded deep in its nature'.[77]

Helmholtz still attributed the inadequacy of Weber's law to velocity dependence in the potential between any two electrical atoms. He attempted to establish his own potential solely on the basis of spatial separation between any two points where currents existed (phenomenologically rather than as moving atoms of electricity), this being 'the only spatial magnitude which is completely determined by two points'.[78] Current velocity, nevertheless, remained hidden in the strength of the current at any point. Helmholtz simply ignored it, even though he himself considered currents to be electricity in motion.

Such mathematical manoeuvres could not impress physical theorists like Weber; indeed, Helmholtz later stated that his law was 'no elementary expression of the ultimate acting forces'.[79] It did allow him, however, to employ an instantaneous action-at-a-distance force between currents while avoiding either explicit velocity-dependence or propagation of force.

The latter issue brings us to Helmholtz's second task, explaining propagation of light electromagnetically. For that he appropriated Maxwell's theory to his own ends. Maxwell had shown in 1863 that an elastic displacement in a medium possessing the electric and magnetic properties of free space would propagate at the velocity of light, through a process of mutual induction between electric and magnetic polarisations (Siegel, this volume). The polarisations were attributed by Maxwell to motions in ether; they constituted electric and magnetic forces, and no actions except between contiguous elements of ether occurred. Few on the Continent pretended to grasp how this contiguous action was supposed to function, nor what electricity might be, and Maxwell offered little aid. But his equations provided a thorough description of electromagnetic effects and of electromagnetic waves propagating at the speed of light. The critical element in those equations was the assumption that electric displacement, or polarisation, acted during the displacement exactly like an electric conduction current.

Considering these features from his own perspective, Helmholtz assumed ether to be an electrically and magnetically polarisable medium in which all forces – electrostatic, electromagnetic, and magnetic – acted directly at a distance between its parts (presumably atoms, but not explicitly). The polarisation at any point, an elastic response, was proportional to the sum of the polarising forces acting on that point from all distances. Helmholtz's basic equations, therefore, were two elasticity equations for electric and magnetic polarisations, both of the form $\mathbf{F}=a\mathbf{P}$, where $\mathbf{F}$ is force, $\mathbf{P}$ polarisation, and a an elastic constant. The forces, however, were complex and interconnected between the two equations. Electric forces resulted from free electricity, electricity of polarisation, changing conduction currents, changing polarisation currents, and changing magnetic polarisations, the last three being sources of electromagnetic induction. Magnetic forces, similarly, resulted from magnetic polarisation and from electric conduction currents and polarisation currents.

Employing his own electromagnetic potential, but including both conduction and polarisation currents, Helmholtz derived from the elasticity equations partial differential equations for purely electric and for purely magnetic polarisations, having eliminated cross terms. These new equations were wave equations with solutions describing both transverse and longitudinal waves. The choice $k=0$ in the potential made the longitudinal waves vanish (infinite velocity), leaving only transverse waves of electric and magnetic polarisation with perpendicular oscillations. Assuming the electric polarisability of ether to be very large, Helmholtz showed that the transverse waves would spread – *propagate* is not quite right – with the velocity of light. In the limit, therefore, and with the choice $k=0$, his theory reduced mathematically to Maxwell's theory of light.[80]

In this derivation the original assumption of instantaneous action at a distance disappeared, to be replaced by equations apparently describing successive action between contiguous elements of a medium. The result, however, was only mathematical and not physical; it did not imply that polarisations actually propagated, but only that they spread in time. In each original elasticity equation – through differentiations, transformations, and substitutions – Helmholtz had replaced the summed forces of currents and polarisations, acting *from all distances* on a local region, by differentials of their effects (polarisations) *in the local region*. Spreading of polarisation, or the wave equation, arose from the replacement process as a coincidental result of the relation between spatial and temporal derivatives in the laws of electromagnetic action. No approximations were necessary, not even neglect of long-range actions over short-range ones.

Helmholtz was fully aware of the coincidental nature of his agreement with Maxwell and turned the coincidence to support his own view: 'The remarkable analogy between the motion of electricity in a dielectric and that of the luminiferous ether does not depend on the particular form of Maxwell's hypothesis, but follows in essentially the same way if we maintain the older view of electrical action at a distance'.[81] But though he used his own formulation of spatial action at a distance against Maxwell's ether field, Helmholtz also turned Maxwell's field against Weber's time-dependent action at a distance. 'Maxwell's hypothesis appears to me to be very important', he said in 1872, 'because it furnishes proof that nothing is implied in electrodynamical phenomena that forces us to reduce them to an anomalous [*ganz abweichend*] kind of natural forces, to forces that depend, not merely on the positions of the corresponding masses but also on their motions'.[82] With strong physical theories on either hand, both alien, Helmholtz stood on the neutral territory of uninterpreted mathematics and argued that neither alternative possessed the force of necessity, the necessity he had attempted to establish in 1847 for direct spatial action between point atoms.

Conclusion

What was the historical significance of Helmholtz's ether theory? This question has sometimes been answered with the observation that Helmholtz translated Maxwell's theory into terms comprehensible to Continental physicists; and that is certainly correct so far as it goes. By providing a double-ended theory that displayed both poles from which previous electromagnetic theories of light propagation had been sought – atoms and forces *versus* differential field equations – Helmholtz made clear what propagation would have to mean from the perspective of forces acting at a distance. But in supplying this illumination, I suggest, the theory also illuminated the inadequacies of any action-at-a-distance theory, especially for energy considerations.

By 1870 propagation meant propagation of energy. Energy had become the symbol of unity in nature and the quantity that had to be followed through any series of conversions and transmissions: for example, from a chemical battery, to an electric current, to polarisation of the ether, to induced currents, to heat. Any acceptable electromagnetic theory of light, therefore, had to supply a direct explanation of the process of propagation of the energy of light. Ironically, Helmholtz, the most notable author of energy concepts on the Continent, could not meet this demand in his own theory, primarily because it did not distribute energy independently through space; it was not a field theory.

Wave theorists of light ever since Fresnel had considered light to be *vis viva* in ether. In the usual action-at-a-distance theories, which treated ether as an elastic solid, one supposed that direct actions over several molecular distances were negligible in comparison with actions between adjacent molecules. Even here, therefore, *vis viva* was effectively propagated by successive action. Time for transmission derived from inertia of the molecules, each taking time to respond to the force exerted by its neighbours, and the process was conceived mechanically under Newton's laws of motion. In Helmholtz's phenomenological ether theory, however, actions occurred at all distances, there were no masses to carry *vis viva*, all energies were potential energies, and no mechanics was supplied. Propagation of light could almost be conceived as a spreading of potential energy of polarisation, but one had to regard the energy of polarisation in any local element of ether as a temporary receptacle for energy of abstract spatial relation between this element and all other elements at all distances. The local polarisation, upon relaxation, was again transmitted to all distances. Energy did not propagate between contiguous elements of ether as seemed necessary for phenomena of light, particularly when the wave equation was considered fundamental.

Problems of this sort are not merely retrospective evaluations of Helmholtz's theory. They were recognised and discussed immediately and formed the background for the reception and development of Maxwell's field theory in Germany. Joseph Stefan, for example, published in 1874 a detailed analysis of the energy relations in Helmholtz's theory, intending to illuminate the question

> whether one does not in general have to conceive magnetic and electric phenomena as conditions of a medium, perhaps the luminiferous ether, and whether particularly one does not have to assume that magnetic and electric forces are only apparent actions at a distance, being in fact immediate actions of the medium, dependent on its momentary states and therefore also propagated, just as these states, with finite velocity.[83]

Although Stefan preferred the latter theory he recognised that it was not yet a necessity but only more coherent. He limited his critique to a development of the energy relations that would be required for translating between the two views. He obtained in this way a clear perception of a problem soon widely recognised as basic to the development of Maxwell's, or any other, field theory of electromagnetism: What was the relation between energy in the ether field and the source of that field? His own preference was to treat energy in ether as the independent reality and to make source strength dependent on the state of the ether.

Although Stefan's specific views were apparently not influential, the problems that he recognised in Helmholtz's interpretation of Maxwell's theory were also appreciated by the acknowledged giants of Continental electromagnetic theory, H. A. Lorentz and Heinrich Hertz. Lorentz originally followed Helmholtz's own path and tried to generate a more detailed mechanical description of the role of electricity in matter that would allow an explanation of phenomena of physical optics in terms of the interaction of an electromagnetic

wave of polarisation with electricity bound to molecules (Hirosige, 1969). Hertz, however, represented the new generation of physicists who believed that differential equations presented the most perspicuous view of reality and that theories should be based directly on them. After bemoaning the incomprehensibility of the relation in Maxwell's theory between sources and propagating effects he produced in 1890 his purely mathematical theory of an electromagnetic field.[84] He eliminated the distinction present in both Maxwell's and Helmholtz's theories between polarising forces and polarisation – thereby eliminating the distinction between forces and ether – and made of the field an entity unto itself. By employing this independent field and relating it to sources, Lorentz created his electron theory in the 1890s (Schaffner, 1969). Lorentz believed that some sort of ether had to be imagined as the basis of the field, but others, such as Wien and Abraham, actively sought a pure force field that would reduce even matter and mechanics to electromagnetic processes (McCormmach, 1970). This electromagnetic view of nature matched the goal of the Energeticists – Wilhelm Ostwald and Georg Helm are notable – who sought to reduce all of nature to energy alone. Although their programme foundered on the second law of thermodynamics, their rejection of atomism and belief in the ultimate continuity of nature was widely shared, even by critics such as Max Planck, to the degree that atomists like Ludwig Boltzmann despaired of having any impact in Germany at all.

Developments in the electromagnetic theory of light only partly caused this shift to continuity in nature, but they well exemplify general trends. The atomistic ethers of Weber and Helmholtz had given way to the electromagnetic ether field, which came increasingly to mean a field of electromagnetic energy. Energy and continuous flux seem best to symbolise late nineteenth-century views of nature in Germany.

Notes

1 T. S. Kuhn, 'Energy conservation as an example of simultaneous discovery', in *Critical problems in the history of science*, ed. M. Clagett (Madison, Wis., 1959), 321–56; P. M. Heimann, 'Conversion of forces and the conservation of energy', *Centaurus 18* (1974), 147–61.

2 S. D. Poisson, 'Mémoire sur la théorie du magnétisme', *Mémoire de l'Académie 5* (1820–2), 247–338, 488–533.

3 *Physicists* will refer here to practising scientists holding institutional positions in physics, mathematics, astronomy, or related fields when they devoted significant effort to *Physik*, as recognised in the abstracting journal *Fortschritte der Physik*.

4 R. S. Turner, 'Julius Robert Mayer', in *Dictionary of scientific biography*, vol. 9 (1974), 235–40.

5 'Ueber notwendige Konsequenzen und Inkonsequenzen der Wärmemechanik', a lecture delivered in 1869 to the Versammlung deutscher Naturforscher und Aerzte, in

Die Mechanik der Wärme in gesammelten Schriften von Robert Mayer, ed. J. J. Weyrauch, 3rd rev. ed. (Stuttgart, 1893), 357.

6 Ibid., 356.

7 H. von Helmholtz, 'The aim and progress of physical science', opening address in 1869 to the Versammlung deutscher Naturforscher und Aerzte, in H. von Helmholtz, *Selected writings of Hermann von Helmholtz*, ed. and trans. Russell Kahl (Middletown, Conn., 1971), 223–45; F. Gregory, *Scientific materialism in nineteenth-century Germany* (Boston, 1977), 100–21.

8 'Bemerkungen über die Kräfte der unbelebten Natur', *Annalen der Chemie und Pharmcie 42* (1842); reprinted in Weyrauch, *Mechanik der Wärme*, 24.

9 *Die organische Bewegung in ihrem Zusammenhange mit dem Stoffwechsel: ein Beitrag zur Naturkunde* (Heilbronn, 1845); reprinted in Weyrauch, *Mechanik der Wärme*, 48.

10 B. Gower, 'Speculation in physics: the history and practice of *Naturphilosophie*', *Studies in History and Philosophy of Science 3* (1973), 301–56.

11 I. Kant, *Kant's gesammelte Schriften*, ed. Preussischen Akademie der Wissenschaften, vols. 21–2, *Opus postumum* (Berlin and Leipzig, 1936), 21:593, lines 7–15, quoted in W. K. Werkmeister, 'The Critique of Pure Reason and Physics', *Kant Studien 68* (1977), 41.

12 *Opus postumum*, 22:583, lines 23–7, quoted in Werkmeister, 'The Critique and Physics', 45.

13 In *Zusatz* 5 (1881) to 'Ueber die Erhaltung der Kraft: eine physikalische Abhandlung' (1847), in H. von Helmholtz, *Wissenschaftliche Abhandlungen von Hermann von Helmholtz* (Leipzig, 1882), 1:73; reprinted in Helmholtz, *Selected writings*, 53.

14 *Beiträge zur Dynamik des Himmels: in populärer Darstellung* (Heilbronn, 1848); reprinted in Weyrauch, *Mechanik der Wärme*, 162, 176.

15 G. Helm, *Die Energetik: Nach ihrer geschichtlichen Entwicklung* (Leipzig, 1898), 16–27.

16 R. Kohlrausch and W. Weber, 'Elektrodynamische Maassbestimmungen insbesondere Zurückführung der Stromintensitäts – Messungen auf mechanisches Maass', *Abhandlungen der Königlichen Sächsischen Gesellschaft der Wissenschaften zu Leipzig 3* (1857); reprinted in *Wilhelm Weber's Werke*, 6 vols. (Berlin, 1892–4), 3:652. See also the 'Vorwort' to Kohlrausch and Weber's paper, *Berichte über die Verhandlungen der Königlichen Sächischen Gesellschaft der Wissenschaften zu Leipzig 17* (1855); reprinted in *Werke*, 3:594ff.

17 My description may be taken as complementary to Kenneth Caneva's presentation of Weber's and Fechner's methodology as hypothetico-deductive. K. Caneva, 'From galvanism to electrodynamics: the transformation of German physics and its social context', *Historical Studies in the Physical Sciences 9* (1978), 63–159. See also K. H. Wiederkehr, *Wilhelm Eduard Weber: Erforscher der Wellenbewegung und der Elektrizität, 1804–1891* (Stuttgart, 1967); A. P. Molella, 'Philosophy and nineteenth-century German electrodynamics: the problem of atomic action at a distance', unpublished doctoral dissertation, Cornell University, 1972.

18 'Ueber ein allgemeines Grundgesetz der elektrischen Wirkung', *Abhandlungen bei Begründung der Königlichen Sächsischen Gesellschaft der Wissenschaften zu* (Leipzig 1, 1846), reprinted in Weber, *Werke*, 3:157; 'Auszug', *Annalen der Physik und Chemie 73* (1848), reprinted in ibid., 245.

19 'Reflections on the law of universal gravitation', in P. de Laplace, *The system of the world*, trans. H. H. Harte (Dublin, 1830), bk. 4, chap. 15.

20 'Of the motion of a system of bodies', in Laplace, *System*, bk. 3, chap. 5; *Essai philosophique sur les probabilités*, 6th ed. (Paris, 1840), 2–4.

21 'Lettre de M. Ampère à M. le comte Berthollet, sur la détermination des proportions

dans lesquelles les corps se combinent d'après le nombre et la disposition respective des molécules dont leurs particules intégrantes sont composées', *Annales de chimie 90* (1814), 45–86; 'Note de M. Ampère sur la chaleur et sur la lumiere considérées comme résultant de mouvemens vibratoires', *Annales de chimie 58* (1835), 432–44.

22 Summarised by L. Pearce Williams, *Michael Faraday, a biography* (New York, 1971), 142–51.

23 G. T. Fechner, 'Ueber die Verknupfung der Faraday'schen Inductions-Erscheinungen mit den Ampère'schen elektro-dynamischen Erscheinungen', *Annalen der Physik und Chemie 64* (1845), 337–45.

24 'Allgemeines Grundgesetz', 3:213 ff.

25 'Ueber die Erregung und Wirkung des Diamagnetismus nach den Gesetzen inducirter Ströme', *Annalen der Physik und Chemie 73* (1848); reprinted in Weber, *Werke*, 3:255–68.

26 'Ueber das Aequivalent lebendiger Kraft', *Annalen der Physik und Chemie 152* (1874); reprinted in Weber, *Werke*, 4:302. See also fragment of letter from Weber to Fechner in G. T. Fechner, *Ueber die physikalische und philosophische Atomenlehre* (Leipzig, 1855), 73.

27 'Allgemeines Grundgesetz', 3:212 ff.

28 'Aequivalent lebendiger Kraft', 4:303.

29 'Electrodynamische Maassbestimmungen insbesondere über das Princip der Erhaltung der Energie', *Abhandlungen der Königlichen Sächsischen Gesellschaft der Wissenschaften zu Leipzig 10* (1871); reprinted in Weber, *Werke*, 4:255 n.

30 'Elektrodynamische Maassbestimmungen insbesondere Widerstandsmessungen', *Abhandlungen der Königlichen Sächsischen Gesellschaft der Wissenschaften zu Leipzig 1* (1852); reprinted in Weber, *Werke*, 3: esp. 400–5.

31 'Vorwort', 3:595 ff.; Kohlrausch and Weber, 'Zurückführung der Stromintensitäts-Messungen auf mechanisches Maass', 3:652–67.

32 'Zur Galvanometrie', *Abhandlungen der Königlichen Gesellschaft der Wissenschaften zu Göttingen 10* (1862); reprinted in Weber, *Werke*, 4:91–6.

33 'Princip der Erhaltung der Energie', 4:268–78.

34 Ibid., 278 ff.; 'Ueber die Bewegungen der Elektricität in Körpern von molekularer Konstitution', *Annalen der Physik und Chemie 156* (1875); reprinted in Weber, *Werke*, 4:334–57.

35 'Elektrodynamische Maassbestimmungen insbesondere über die Energie der Wechselwirkung', *Abhandlungen der Königlichen Sächsischen Gesellschaft der Wissenschaften zu Leipzig 11* (1878), reprinted in Weber, *Werke*, 4:389–95; 'Elektrodynamische Maassbestimmungen insbesondere über den Zusammenhang des elektrischen Grundgesetzes mit dem Gravitationsgesetze', first published in ibid., 4:524 ff.

36 See Zöllner's editorial contributions to *Abhandlungen zur atomistischen Theorie der Elektrodynamik von Wilhelm Weber*, vol. 1, bk. 1, of J. C. F. Zöllner, *Principien einer elektrodynamischen Theorie der Materie* (Leipzig, 1876); Weber, 'Zusammenhang des elektrischen Grundgesetzes mit dem Gravitationsgesetze', 4:481–5.

37 'Aphorismen', in Weber, *Werke*, 4:630 ff.

38 J. C. F. Zöllner, *Transcendental physics*, trans. C. C. Massey (London, 1880).

39 *Atomenlehre*, 23, 53, 73, 187, 206–9.

40 Ibid., ix.

41 See J. F. Herbart, *Johann Friedrich Herbart's sämmtliche Werke* (Leipzig, 1850–2), ed. G. Hartenstein, vol. 5, *Lehrbuch zur Psychologie* (1816), and vols. 3–4, *Allgemeine Metaphysik, nebst den Anfängen der philosophischen Naturlehre* (Köngsberg, 1828–9), esp. pt. 5, 'Umrisse der Naturphilosophie'.

42 *Atomenlehre*, 90–1, 95.

43 Ibid., 63–73.
44 Ibid., basic viewpoint developed 106–18.
45 Ibid., 65.
46 Ibid., 95.
47 Ibid., 113. Cf. G. T. Fechner, *Elemente der Psychophysik* (Leipzig, 1860), esp. chaps. 1, 5.
48 *Atomenlehre*, 181–210.
49 C. F. Gauss, *Werke* (Göttingen, 1877), 5:629.
50 Riemann's associations and interests at Göttingen are described by R. Dedekind in 'Lebenslauf', *Bernhard Riemann's gesammelte mathematische Werke*, ed. H. Weber, 2nd ed. (Leipzig, 1892), 539–58.
51 'Fragmente philosophischen Inhalts', in H. Weber, *Riemann's Werke*, 524.
52 Ibid., 528 ff.
53 Ibid., 503.
54 Ibid., 533.
55 *Schwere, Elektricität, und Magnetismus: nach den Vorlesungen von Bernhard Riemann*, ed. K. Hattendorff (Hannover, 1876), 330.
56 'Ein Beitrag zur Elektrodynamik', *Annalen der Physik und Chemie 131* (1867); reprinted in H. Weber, *Riemann's Werke*, 288.
57 See note by H. Weber in *Riemann's Werke*, 293.
58 Ibid., 533 ff.
59 'Ueber Integrale der hydrodynamischen Gleichungen, welche den Wirbelbewegungen entsprechen', *Journal für die reine und angewandte Mathematik 55* (1858); reprinted in Helmholtz, *Wissenschaftliche Abhandlungen*, 1:101–34.
60 *Vorlesungen über mathematische Physik: Mechanik* (Leipzig, 1876), lectures 15–26.
61 Ibid., 'Vorrede'.
62 Bjerknes's career is described by V. Bjerknes in C. A. Bjerknes, *Hydrodynamische Fernkräfte: fünf Abhandlungen über die Bewegung kugelförmiger Körper in einer inkompressiblen Flüssigkeit (1863–1880)*, trans. A. Korn, ed. A. Korn and V. Bjerknes (Leipzig, 1915), 212–23. See also Dedekind, 'Lebenslauf', *Riemann's Werke*, 551.
63 'Hydrodynamische Analogien zu den elektrostatischen und magnetischen Kräften', *Naturen* (1880), reprinted in Bjerknes, *Hydrodynamische Fernkräfte*, 176–211; G. Forbes, 'Hydrodynamic analogies to electricity and magnetism', *Nature 24* (1881), 360 ff.
64 'Ueber die Identität der Schwingungen des Lichts mit den electrischen Strömen', *Annalen der Physik und Chemie 131* (1867), 243 ff.
65 Ibid., 261–2.
66 Ibid., 262.
67 Ibid., 263.
68 'Resultate einer Untersuchungen über die Principien der Elektrodynamik', *Nachrichten von der Königlichen Gesellschaft der Wissenschaften und der Georg Augustus Universität zu Göttingen* (1868), 223–35. For Neumann's earlier attempt to derive the magnetic rotation of light from Weber's law of force see Knudsen (1976).
69 E.g., Helm, *Energetik*, 229–31.
70 P. M. Heimann, 'Helmholtz and Kant: the metaphysical foundations of "*Ueber die Erhaltung der Kraft*" ', *Studies in History and Philosophy of Science 5* (1974), 205–38; Y. Elkana, 'Helmholtz's "Kraft": an illustration of concepts in flux', *Historical Studies in the Physical Sciences 2* (1970), 263–98.
71 'Erhaltung der Kraft', in Helmholtz, *Wissenschaftliche Abhandlungen*, 1:14; reprinted in Helmholtz, *Selected writings*, 5.
72 *Metaphysical foundations of natural science*, trans. J. Ellington (Indianapolis and New York, 1970), esp. chap. 2.

73 'Erhaltung der Kraft', in Helmholtz, *Wissenschaftliche Abhandlungen*, 1:17–19; reprinted in Helmholtz, *Selected writings*, 6–8.
74 Ibid., in Helmholtz, *Wissenschaftliche Abhandlungen*, 1:22; reprinted in Helmholtz, *Selected writings*, 11.
75 Ibid., in Helmholtz, *Wissenschaftliche Abhandlungen*, 1:14; reprinted in Helmholtz, *Selected writings*, 4.
76 'Ueber die Bewegungsgleichungen der Elektricität für ruhende leitende Körper', *Journal für die reine und angewandte Mathematik 72* (1870); reprinted in Helmholtz, *Wissenschaftliche Abhandlungen*, 1:567.
77 Ibid., 553.
78 Ibid., 565.
79 'Ueber die Theorie der Elektrodynamik', *Monatsberichte der Berliner Akademie* (1872); reprinted in Helmholtz, *Wissenschaftliche Abhandlungen*, 1:637.
80 'Bewegungsgleichungen der Elektricität', 1:625–8.
81 Ibid., 558.
82 'Theorie der Elektrodynamik', 1:639.
83 'Ueber die Gesetze der magnetischen und elektrischen Kräfte in magnetischen und dielektrischen Medien und ihre Beziehung zur Theorie des Lichtes', *Sitzungsberichte der Kaiserlichen Königlichen Akademie der Wissenschaften zu Wien 70* (1874), 596.
84 R. McCormmach, 'Heinrich Hertz', in *Dictionary of scientific biography*, vol. 6 (1972), 340–50.