Philosophical Issues in Classical Electrodynamics

1. Overview

Traditionally research in the philosophy of physics has focused predominantly on interpretative puzzles in quantum mechanics and, perhaps to a slightly lesser extent, on the revolution in our understanding of the structure of space and time ushered in by the Special and General Theories of Relativity. This overly narrow focus may give the impression that classical physics is metaphysically and methodologically unproblematic, with the consequence that many philosophically interesting issues in the foundations of physics have been largely ignored by philosophers of physics. Fortunately, this situation has begun to change.

In this essay I will survey several philosophical issues arising in the context of classical electrodynamics. I will focus on the following topics, some of which overlap with more traditional concerns. First, we can ask what role theories of classical physics play in contemporary physics, given that these theories have in a sense been superseded by quantum theories. What, for example, is the status of the ontologies postulated by classical theories? Can we accept quantum mechanics and nevertheless rationally remain realists about classical fields and charges? Second, historically, the development of microscopic classical electrodynamics and 'theories of the electron' has been intimately linked with the development of the Special Theory of Relativity. But how exactly did our understanding of space and time change from Lorentz's theory of the electron to Einstein's theory of relativity? According to one traditional way of understanding the change, Lorentz's theory of the electron proposed dynamical explanations for phenomena that Einstein's theory has shown to be purely geometric and kinematic effects. Yet this orthodox view has recently been challenged.

Third, at the heart of classical electrodynamics are the Maxwell equations and the Lorentz force law.[1] The microscopic Maxwell equations allow us to determine the

---

[1] The standard graduate- and research-level textbook on classical electrodynamics is (Jackson 1999). Books focusing on the problem of arriving at a classical particle equation of motion are (Rohrlich 2007; Parrott 1987; Spohn 2004).

electromagnetic fields, the electric field **E** and the magnetic induction **B**, associated with a given microscopic charge and current configuration. The equations are usually solved in terms of a modified initial value problem: The initial fields at some time t_0 together with the full trajectory of the charges are posited as input into the Maxwell equations, which then determine the evolution of the fields. According to the Lorentz-force law, fields exert a force on charged particles. For continuous distributions of charges (so-called 'charged dusts') the Lorentz force law can be derived from the Maxwell equations together with the standard formulation of the principle of energy-momentum conservation. The situation is far more complicated in the case of discrete particles coupled to an electromagnetic fields. While we can model both kinds of interactions separately for discrete charges and the resulting models are empirically highly successful within the domain of classical physics, it is unclear whether there is a conceptually fully satisfactory theory of the interaction between localized microscopic charges and fields that takes both modes of interaction into account simultaneously. Thus, despite its tremendous empirical success, the theory is beset by serious, and perhaps irresolvable, foundational problems.

Fourth, electromagnetic waves in the presence of wave sources exhibit a striking temporal asymmetry. We observe waves that coherently diverge from a source, but we generally do not observe the temporal inverse—waves coherently converging onto a wave source—despite the fact that the underlying dynamical equations of the theory are time-reversal invariant and permit both kinds of behavior. What is the source of this asymmetry? Existing answers fall into two classes: Some physicists and philosophers argue that the radiation asymmetry can ultimately be reduced to thermodynamic and statistical considerations, while others argue that the asymmetry is an instance of a general causal asymmetry. This is an issue where the philosophy of classical electrodynamics makes contact both with the philosophy of thermodynamics and with a core issue in metaphysics—the place of causation in our conception of the world.

Fifth, according to the standard interpretation, the ontology of classical electrodynamics consists of discrete charged objects and continuous electromagnetic fields transmitting the influence from one charge on another. The fact that interactions between charged objects are mediated by fields ensures that the theory satisfies various locality conditions—conditions which are violated, for example, by Newtonian gravitational theory. Indeed, arguably the very fact that eletromagnetic fields act locally provides an important reason for interpreting fields realistically. However, the traditional interpretation comes under pressure from the semi-classical Bohm-Aharanov effect, which has been taken to suggest either that fields can act non-locally or that the so-called electromagnetic potentials, which in the traditional interpretation are viewed as mere calculational devices, ought to be interpreted realistically.

2. Classical physics in a 'quantum world'

An immediate question that arises when we discuss theories of classical physics is what the present status of these theories is in light of our acceptance of quantum physics. Does our best current physics not tell us that we 'live in a quantum world'? Has classical electrodynamics, thus, not been replaced by quantum theories and is today of merely historical interest? An affirmative answer to these questions is suggested by many of the major methodologists of science of the twentieth century, including by Thomas Kuhn in his famous book *Structure of Scientific Revolutions* (Kuhn 1996), but also by many of Kuhn's critics. According to this view, when scientists adopt a new theory, such as quantum mechanics, the new theory replaces its predecessors, which are abandoned. While Kuhn and his critics disagree on whether the process of theory acceptance and rejection is rational and while there is widespread disagreement on the circumstances that lead to a theory's demise, there seems to be broad agreement that the history of physics does indeed consist of theories that are accepted only to be eventually discarded in favor of more successful rivals.

This view is seriously flawed, however. To be sure, there are examples of theories that do fit the standard view and which were completely abandoned in favor of their successors. Phlogiston theory is one such case. Yet neither classical electrodynamics nor Newtonian physics fit this traditional picture and both theories continue to play an important role in contemporary physics. A more adequate way of thinking about theories such as these has been proposed by Fritz Rohrlich, among others (Rohrlich and Hardin 1983). On Rohrlich's view, quantum mechanics has not fully replaced classical electrodynamics but rather has helped to establish the limits of the latter theory's validity. Despite the fact that classical electrodynamics has been superseded by quantum mechanics in some sense—we take quantum theories to be more fundamental and there are phenomena which cannot be modeled classically—the classical theory remains the most appropriate theory for modeling and explaining phenomena in its domain of validity. In Rohrlich's terminology, classical electrodynamics is an "established theory": it is well-confirmed within a certain domain of applications but with clear limits of validity given by a more fundamental quantum theory.

Yet this view of classical theories also raises a number of questions that deserve a fuller discussion. Does classical electrodynamics indeed provide us with the best explanations of phenomena within its domain of application, and if 'yes', what theory of scientific explanation can account for this fact? How can we characterize in more detail the relation of classical electrodynamics to the putatively more fundamental quantum theories? What, for example, is the relation between the different ontologies postulated by classical and quantum theories and what ought our ontological commitment to the classical theory be? Can we coherently be committed both to the existence of classical fields and to quantum field theory as a more fundamental theory? Or is our commitment better described as a commitment to the claim that the world is such that on a certain level it is best represented as consisting of classical fields interacting with charged material objects? (Rohrlich 2004) argues for a non-reductionism that allows us to be realists about the classical *ontology* even while we acknowledge that the classical theory's mathematical *formalism* can be shown to arise from the more fundamental quantum mechanical formalism in the limit. Whether there is room for such a (weak) non-reductive position, which collapses neither into a full-blown ontological reductivism nor into a pragmatism or instrumentalism about the higher-level ontology is an issue that deserves to be explored further.

3. Classical Electrodynamics and Special Relativity

The development of microscopic electrodynamics at the turn of the twentieth century was intimately bound up with that of the special theory of relativity. Lorentz showed that from an assumption about how intermolecular electromagnetic interactions transform from one frame to another one can derive a result that we today would interpret as the claim that the Maxwell equations are Lorentz invariant: if the Maxwell equations allow a certain configuration of charges and fields in a system at rest, then they allow the same configuration of charges in an inertial system moving through the ether. For Lorentz, however, only measurements in frames that are at rest with respect to the ether reveal the real fields; the fields satisfying the Maxwell equations in a moving frame are 'fictitious'. Lorentz was also able to derive equations equivalent to the relativistic equation for length contraction (and eventually also for time dilation) from dynamical considerations.[2]

One of Einstein's main motivations in developing the special theory of relativity appears to have been the fact that even though the Maxwell equations allow the same configurations of charges and fields at rest and in motion, there is no frame-independent answer as to what the electric and magnetic fields are. Rather electric fields in one frame can appear as magnetic fields in another inertial frame and vice versa. Thus, according to relativity theory, the object that exists frame-independently is the electromagnetic field, which can be represented in terms of the electromagnetic field tensor $F^{\mu\nu}$ which has electric and magnetic field strengths as coordinate-dependent components.

It is generally held today that Einstein's special theory of relativity has changed our understanding of space and time and of the motion of objects in two important ways. First, we no longer think that there is a privileged class of inertial frames, the 'ether rest frames'; and, second, we think that relativistic phenomena, such as length contraction, do not require a dynamical explanation in terms of electromagnetic forces or quantum mechanical interactions, but are purely kinematic effects and a consequence of the geometry of spacetime—of the fact that spacetime is Minkowskian. This orthodox view has recently been challenged by Harvey Brown (Brown 2005; Brown and Pooley 2006), who agrees with the first part of the orthodox view but not with the second. According to what Brown takes to be the orthodox view, the structure of spacetime—that is, the fact that spacetime is Minkowskian—explains the fact that the laws of our theories are Lorentz invariant and accounts for the universal behavior of rods and clocks. But Brown argues that this view has the arrow of explanation backward. On his view, relativistic phenomena ultimately require a dynamical explanation; it is a brute fact, which itself

---

[2] A summary of Lorentz's theory of the electron can be found in (Lorentz 1916).

permits of no further explanation, that the laws are Lorentz invariant and it is this fact that explains that Minkowski spacetime is the proper arena in which to represent non-gravitational physical phenomena.

One of Brown's targets is the view that spacetime substantivalism plays an important role in explaining relativistic effects. The substantivalist takes spacetime to be an entity in its own rights. Its affine structure determines the geodesics, along which free particles move. Once we know that objects 'live' in Minkowski spacetime, and satisfy the constraints of Minkowski geometry, there is a simple well-known geometric construction that allows us to derive length contraction and time dilation. Thus, length contraction and time dilation appear to be purely geometric effects, which are a straightforward consequence of the structure of Minkowski spacetime. But Brown argues that merely appealing to the structure of spacetime does not answer the question as to *why* objects obey the constraints of Minkowski geometry. In positing that force-free objects follow the geodesics of Minkowski spacetime the substantivalist needs to assume that objects are somehow able to sense 'the ruts and grooves' of spacetime and Brown finds this assumption utterly mysterious. What is needed instead, Brown argues, is a "constructive" explanation that appeals to "the details of the bodies' microphysical constitution" (Brown and Pooley 2006)

Brown's account has been forcefully challenged by Michel Janssen, who takes himself to defend the orthodox view that the special theory of relativity has shown length contraction and time dilation to be purely kinematic phenomena (Janssen 2008; see also Balashov and Janssen 2003). Janssen agrees with Brown in his rejection of spacetime substantivalism but argues that even within a relationalist framework, which takes spatiotemporal relations to have no existence independent of physical objects instantiating them, the explanatory arrow is from the structure of Minkowski space to the Lorentz invariance of the dynamical laws, rather than the other way around.

On first sight there seems to be a rather stark disagreement between Brown's and Janssen's views. Brown, it seems, believes that only an account of the particular forces pushing and pulling the microscopic constituents of a rod can explain length contraction, while Janssen maintains that the structure of Minkowski spacetime explains the phenomenon. But on closer inspection the differences might strike one as rather more subtle. While Brown stresses that the behavior of rods and clocks requires a dynamical and constructive explanation—invoking Einstein's distinction between constructive theories and theories of principle[3]—he also appears to concede that the only feature of the dynamics that is required to explain length contraction or time dilation is its Lorentz invariance. The full dynamics may be needed to account for the existence of stable solid objects and of time-keeping devices, but given that there are stable macroscopic objects, no more than the Lorentz-invariance of the laws is required to account for their behavior in different inertial frames. And while Janssen seems to insist that the proper explanation of length contraction is geometrical, he also suggests that the explanation proceeds by appealing to Lorentz invariance as a "meta-law", and that the role of Minkowski spacetime is to "encode" this nomological fact. Thus, on one reading of the two competing views, the disagreement reduces to the question whether Lorentz invariance is postulated as brute constraint, as Brown maintains, or as meta-law or overarching nomic constraint, as Janssen claims. But then the disagreement would perhaps be best captured neither in terms of the kinematics-dynamics distinction nor in terms of the distinction between principle- and constructive theories, but rather as a disagreement about the role of laws in science and about the explanatory force of appeals to nomic constraints.

4. The coherence of the theory

The fundamental equations at the heart of microscopic classical electrodynamics are the microscopic Maxwell equations, which determine the fields in the presence of a charge and current configuration, and the Lorentz force law, which specifies the force acting on a charge in an electromagnetic field and is used as input into Newton's second law to give the momentum change of a charge. Together they constitute a coupled set of quasi-linear equations that, one would hope, determine the temporal evolution of charges and fields specified at some initial time *t*. In the case of continuous charge distributions or 'charged dusts' there are indeed existence and uniqueness proofs for the coupled Maxwell-Lorentz equations. The only problem arising in this case is that these proofs

---

[3] While the distinction between principle and constructive theories is usually traced to Einstein, a closely related distinction had earlier been drawn by H.A. Lorentz (see Frisch 2005b).

only guarantee the existence of *local* solutions. Moreover, there are intuitively plausible initial charge distributions for which local solutions cannot be arbitrarily extended and for which the equations have no global solutions (see Frisch 2004a).

But many of the applications of microscopic classical electrodynamics concern the interaction of discrete charged particles, such as electrons, with electromagnetic fields and it is much less clear in this case whether there is a coherent and fully satisfactory way of integrating the two ways in which charges and fields interact. At the heart of the problem lies the fact that a completely satisfactory treatment should include so-called self-interaction effects—the effects that the field of a charged particle has on the motion of the particle itself—and arguably there is no conceptually completely unproblematic way of incorporating self-interactions into the theory.[4]

The standard way of modeling interactions between electromagnetic fields and charges, familiar to physics students from both undergraduate and graduate courses in electrodynamics, is simply to ignore self-interactions—usually without making explicit that this is being done. The resulting set of equations—the Maxwell equations and a Newtonian equation of motion for charged particles that takes into account only the Lorentz force due to fields *external* to the charge is inconsistent with the principle of energy-momentum conservation, even though the theory is taken to satisfy this principle. (Frisch 2004b; Frisch 2005a) argues that this inconsistency has consequences for how we think about scientific theories and that classical electrodynamics does not fit standard philosophical accounts of scientific theorizing that take theories to delineate conceptually coherent possible worlds. These arguments are criticized in (Belot 2007), (Muller 2007), and (Vickers 2008). Fred Muller—incorrectly, I believe—criticizes the inconsistency argument, while Gordon Belot and Peter Vickers agree that the assumptions in question are indeed inconsistent but question what conclusions can be drawn from this fact. Belot, in particular, argues that the inconsistency is merely an instance of the common phenomenon that approximations that we make in modeling physical systems are strictly speaking inconsistent with the underlying fundamental equations.

---

[4] Since self interactions are infinitesimal for a charged dust, the problem does not arise in this case.

The overarching philosophical issue at the heart of this debate is whether standard formal philosophical accounts of scientific theories—be it as deductively closed sets of sentences or as classes of models representing the possible worlds allowed by the theory—allow us adequately to capture the conceptual structure of classical electrodynamics. The problem is that any theory with self-interactions has to posit a model for charged particles and arguably there is no physically 'well-behaved' model of a discrete finitely charged particle that does not involve what *by the theory's own lights* are idealizations.

Models of charges can be grouped into two classes, depending on whether charges are treated as point particles or as extended particles. Point-particle models fit better with relativistic constraints, yet an immediate problem faced by these models is that, according to the Maxwell equations, the field of a finitely charged point particle is infinite at the location of the charge and the field-energy diverges for any volume that includes the charge. Extended-particle models avoid the infinities, but run afoul of relativistic constraints such as the prohibition against superluminal propagation. (Note, also, that there can be no purely electromagnetic extended charged particles, since such a particle would explode due to the repulsive forces between its different parts.) There is a rich history of attempts to overcome these problems or to arrive at physically reasonable particle equations of motion despite these problems, many of which are reviewed in (Rohrlich 2007; Parrott 1987; Spohn 2004).

Even though this question is far from being completely settled, it appears that none of these solution attempts result in a treatment of the self-energy problem that do not involve what by the theory's own lights are idealizations of some form or other. But if that is indeed the case, then classical electrodynamics does not merely involve 'benign' kinds of approximations to a fully coherent and in-principle axiomatizable formalism, and hence may only ill fit with traditional philosophical accounts of scientific theorizing. Rather, idealizations and approximations are present at the very first step and in the very construction of a classical model of the electron. Classical electrodynamics, thus appears to be an example of what Mark Wilson calls a "theory façade": a patchwork of laws that are "stitched together" in ways that do not easily submit to formal axiomatization and that contain weak spots—in our case, the self-interactions—which are not readily covered by any of the patches, even though the resources of the theory can in often ingenious ways be extended beyond the patches and into the weak spots (see Wilson 2008).

5. The asymmetry of radiation

A further issue concerns possible explanations of an apparent temporal asymmetry that characterizes electromagnetic radiation phenomena. It is a consequence of the Maxwell equations that accelerated charges radiate. Radiation fields exhibit a temporal asymmetry, which intuitively can be characterized as follows: We observe electromagnetic fields coherently diverging from sources but we do not, or only rarely seem to observe the time-reverse of such phenomena—that is, fields coherently converging into a source. This asymmetry has struck many as puzzling, since the Maxwell equations are time-reversal invariant. One should note, however, that there is nothing puzzling about the fact that most models of fields in the presence of charged particles exhibit a temporal asymmetry. What is puzzling, it seems, is that all or most of the situations we observe exhibit the *same* temporal asymmetry.

The first surprisingly controversial issue is the question as to what exactly the asymmetry consists in. In order to get a sense of the issue, a little bit of background is needed. The Maxwell equations imply an inhomogenous wave equation for the fields in the presence of sources. This equation defines an initial value problem and the general solution can be written as a sum of diverging fields associated with the sources—the so-called *retarded fields* F_{ret}—and source-free incoming fields F_{in}, propagating from the initial value surface. But the fields in a given region can equally as well be represented in terms of a final value problem and the total field then consists of source-free outgoing fields F_{out}, propagated backward from the final value surface in accord with the source-free, homogeneous wave equation, together with fields converging on the source, the so-called *advanced fields* F_{adv}. That is, the same total field F_{total} can be represented as a combination of source-free incoming and retarded fields or of source-free outgoing and advanced fields:

$$F_{total} = F_{ret} + F_{in} = F_{adv} + F_{out}.$$

Thus, it seems that the asymmetry cannot consist in the putative fact that the fields associated with sources are retarded, since *both* retarded and advanced representations are possible (as well as linear combinations of the two).

However, if the incoming field is approximately equal to zero, then the total field exhibits an asymmetry: in this case the total field is *fully* retarded, but generally will not be fully advanced. Unless the retarded and advanced fields are equal, the advanced field representation will involve a non-zero source-free outgoing field $F_{out}= F_{ret} - F_{adv}$, which will destructively interfere with the advanced field before the source is turned on and will be equal to the retarded field after the source turns on. This fact has led Zeh to characterize the radiation asymmetry as follows: "Why does the condition $F_{in} = 0$ (in contrast to $F_{out} = 0$) approximately apply in most situations?" (Zeh 2001, 21)

But it is not clear that there is an asymmetry of the kind postulated by Zeh. First, if we assume that all forces are electromagnetic, then a charge cannot accelerate, and hence cannot radiate, unless it experiences a non-zero external field. Second, the incoming field is not zero in 'most' situations. For example, the cosmic background radiation is always present (see North 2003), and unless we are conducting an experiment in a darkened room there will also be incoming fields in the visual spectrum. Finally, every hyperplane, on which F=0 and which we can use as initial value surface to determine the fields in its future, can equally be used as a final value surface to determine the fields in its past. Thus, there seem to be exactly as many situations in which F_{in}=0 as there are situations with F_{out}=0.

One might reply that in many situations in which we are interested the cosmic background radiation or background radiation from the sun can be ignored and hence the models physicists use to represent these situations posit zero incoming fields. Thus, Zeh's formulation of the asymmetry might apply to the models that are appropriate in representing systems of radiating charges rather than to the complete physical situations modeled. More importantly, perhaps, paradigmatic situations exhibiting the asymmetry usually only involve only a small number of charges. Thus, (Frisch 2006) proposes a way of capturing an asymmetry that characterizes our representations of systems involving small numbers of charges.

Another way of characterizing the asymmetry might be this. Consider once more a single radiating charge in a zero incoming field. Not only is the free outgoing field generally not zero in such cases, but it also has a rather special shape in that it consists of a linear combination of the retarded and advanced fields associated with the charge in question. Now, in most realistic situations the incoming field will not be zero, but there nevertheless will be a difference between the free incoming and outgoing fields in retarded and advanced representations, respectively: the kind of incoming fields we observe do not have collapsing and rediverging waves as components that are centered on the trajectory of the source, while the outgoing field does have diverging waves as its components. (Thus, given what we know about realistic fields, we could in principle use knowledge of the 'free' outgoing field in an advanced field representation to retrodict where and when a radiating source was present, but we could not similarly use our knowledge of the free incoming fields alone to predict where a source will radiate.) Again, this asymmetry is restricted to situations involving a small number of charges, for in the case of a very large number of charges, the retarded and advanced fields of these charges can combine to result, for example, in approximately constant fields that are not centered on any of the charges. (One can see this intuitively by thinking about the fields in the presence of an absorber. In the future of a perfectly absorbing region the field is zero. Hence, the incoming field F_{in}, whatever it may be, is equal to the sum over all absorber particles of F_{adv} - F_{ret}.)

How can we explain the asymmetry of radiation fields? Broadly, the solutions that have been proposed fall into two classes: those that take the radiation asymmetry to be ultimately reducible to considerations from thermodynamics and statistical physics and those that take the asymmetry to be an expression of a causal asymmetry.

Frequently writers cite a debate between Ritz and Einstein as an early instance of the controversy, claiming that Einstein took the asymmetry to have a thermodynamical origin while Ritz thought that the radiation asymmetry was more fundamental than that of thermodynamics. But, as (Frisch 2005a) shows, Einstein's position is considerably more complex than the standard view suggests. While it is true that in his joint paper with Ritz Einstein wrote that he "believes that the irreversibility [of radiation processes] is exclusively based on reasons of probability" (Ritz and Einstein 1909, 324), he also in the very same year claimed that according to classical wave theory, “an oscillating ion produces a diverging spherical wave” and that “the reverse process does not exist as elementary process. […] The elementary process of the emission of light is, thus, not reversible.” (Einstein 1909, 819) The latter quote suggests that Einstein, too, believed that there was an asymmetry characterizing elementary radiation processes that was not of thermodynamic origin.

Many writers who try to explain the radiation asymmetry by appealing to thermodynamic considerations invoke Wheeler and Feynman’s time-symmetric absorber theory of radiation (Wheeler and Feynman 1945). Wheeler and Feynman propose that the universe is surrounded by a perfectly absorbing region and that field associated with a charge is half-retarded and half-advanced. The fully retarded, diverging fields we observe are the result of the half-retarded field’s stimulating an advanced response field in the absorber, which, as Wheeler and Feynman argue, combines with the half-retarded field of the source to give a fully retarded field. Fields do not similarly look fully advanced and the symmetry between retarded and advanced representations is broken, according to them, since a coherent absorber action in the past is extremely improbably. Huw Price (1996) has criticized this argument for committing what he calls “the temporal double-standard fallacy”: since Wheeler and Feynman postulate a time-symmetric wave associated with a charge, a coherent absorber response in the past seems to be no more or less improbable than the future response wave their theory requires.

There have been attempts to address this problem within Wheeler and Feynman’s framework (Hogarth 1962; Hoyle and Narlikar 1995), which are criticized in (Frisch 2005a), and there are other suggestions for how thermodynamic considerations might be marshaled without an appeal to an infinite absorbers and time-symmetric particle fields (North 2003; Price 2006). But all these attempts, it seems to me face the following dilemma. Either they are in danger of committing the double standard fallacy: any account that argues that incoming coherent radiation is extremely improbable has to be able to explain why outgoing coherent radiation is not similarly improbable. Or the account ends up presupposing the very asymmetry they are meant to explain. Thus, North appears to argue that the asymmetry is due to the fact that the universe began in a state in which there are large hot sources in a background of fully thermalized, approximately zero fields. Thermodynamic considerations are then meant to show that such sources are overwhelmingly likely to be associated with retarded radiation. But the assumption that initial fields are approximately equal to zero is already sufficient for the fact that total fields are approximately fully retarded and there remains no work to be done for thermodynamic considerations (see Frisch 2006).

It is also not clear that a thermodynamic explanation of the asymmetry could be general enough. For one, the asymmetry characterizes not only macroscopic systems of fields and sources, but microscopic systems as well; and thermodynamic explanations proceed by arguing that a certain macroscopic evolution is overwhelmingly likely, given that the system in question is an microstate that is in some sense typical for the system's macro-state. Moreover, the radiation asymmetry appears to be analogous to the following asymmetry that does not have a thermodynamic origin. Consider a box with a small hole containing a dense gas in an otherwise empty space. What we expect to observe is a stream of gas particles exiting through the hole, rather than the time-reverse: a stream of particles coming in from infinity 'carefully' aimed at the hole, resulting in a net influx of particles into the box. Yet the latter phenomenon is not excluded on thermodynamic grounds, since we are not dealing with a closed system.[5]

Instead of invoking thermodynamic considerations, physics textbooks usually explain the preference for a retarded representation of the fields by appealing to considerations of causality (see, e.g., Jackson 1999; Griffiths 1999). According to a causal interpretation of the interaction between fields and their sources, retarded and advanced representations of the total field are *mathematically* equivalent, but only the retarded representation represents the physical situation correctly: sources physically contribute retarded fields to the total field, and thus, only retarded fields are physically real, even though the linearity of the wave equation ensures that the total field can always be mathematically represented in terms of advanced fields as well (Fritz Rohrlich 2006;

---

[5] Recent Boltzmannian accounts of the thermodynamic asymmetry begin with the assumption that the world began in a state that is macroscopically special—a state of extremely low entropy—but microscopically normal. Yet such an assumption does not disallow the incoming stream of particles, since the system's initial state with an incoming stream of particles is a macroscopically distinguishable from one without incoming particle stream—that is, the state can be excluded on the grounds that it is not microscopically 'normal'.

Frisch 2005a). How does a causal account fare with respect to the dilemma I presented above? The account has an obvious way of avoiding the temporal double standard fallacy, since coherently diverging disturbances of the field in the future of a radiating charge are explained by the charge's action as common cause of the disturbances. By contrast, coherently converging radiation is extremely implausible, unless it, too, has a common cause explanation in *its* past. (For example, a coherently converging wave might be produced by a radiating source in the center of a spherical mirror, which reflects the diverging wave due to the source back towards the center.) Moreover, the account is not simply question begging. It explains the asymmetry between different mathematical representations of the total field involving free incoming and free outgoing fields, respectively, in terms of the physical asymmetry that field sources produce diverging field disturbances.

The main criticism of causal accounts is that the notions of a source 'causing', 'producing' or 'physically contributing' retarded fields are ill-defined and are an instance of scientifically illegitimate philosophy-speak. In response, defenders of causal accounts can try to show how the notion can be related to other concepts, such as an asymmetric counterfactual, and how these notions can account for asymmetries characterizing our experimental interactions with field sources. What advocates of an explanation of the radiation asymmetry that appeals to an irreducible notion of causation cannot do, however, is provide an account of the notion of cause that will only invoke notions that a "Humean" will find acceptable. Rather, a causal explanation can only be defended against the criticism of being ill-defined, by showing that the causal notions are related to a cluster of other concepts, including those of counterfactuals, interventions, and experimentation in rigorous and illuminating ways.

There is a second *prima facie* asymmetry in the theory: particle-equations of motion that include radiation reaction effects seem, on first sight, not to be time-reversal invariant. Whether appearances are correct in this case is a question that is controversially debated in the literature (see Rohrlich 2000; Rovelli 2004; Rohrlich 2006).

6. Locality

There is one final philosophical issue that I want to mention briefly. It is a widely held view that it is a desideratum that our best scientific theories postulate only local actions. Most intuitively, perhaps, this demand can be expressed and motivated in causal terms. There seems to be something intuitively objectionable about the idea that causes can act where they are not, and hence, it seems that a cause and a distant effect ought to be connected by a spatiotemporally continuous sequence of causes. Newton's theory of gravity violates this demand, even though Newton himself said that action-at-a-distance is a "great absurdity" (quoted in Lange 2002, 94). Classical electrodynamics, by contrast, seems to satisfy the demand, if we interpret electromagnetic fields realistically.

Indeed, interpretations of classical electrodynamics provide an ideal case study of how general methodological and metaphysical demands can influence the interpretation of our theories. Interpreting the fields realistically can be motivated by considerations of locality. Yet there is also a second entity that plays a role in describing fields, the electromagnetic potential A^{μ}. In the standard interpretation the potential is not interpreted realistically and is understood to be a mere calculational device, since it is only determined up to a so-called "gauge transformation" and the same observable fields can be represented by many different potentials.

This short gloss ignores a number of philosophically interesting questions and must be qualified in various important ways. First, one can ask what more precisely we mean when we demand that a theory be local. (Lange 2002) distinguishes several different senses in which a theory can be causally local. But philosophers and physicists have also proposed locality conditions that are not couched in causal terms and it is an interesting question to ask how various causal and non-causal senses of locality might be related (see Frisch 2002). Second, if the assumption that our best theories ought to be local, in some sense, is used to derive metaphysical implications, we should ask what reasons we can cite in support of this assumption. Lange (2002) argues that it is difficult to find a fully convincing argument in favor of a causal locality condition, and it seems to me that the case is even more difficult to make for non-causal locality conditions, which cannot, for example, appeal to the intuition that something cannot act where it is not.

Third, it is not entirely clear, whether introducing electromagnetic fields is enough to ensure that the theory is local. For one, a point-particle equation of motion that includes radiation reaction effects, the Lorentz-Dirac equation, on its standard interpretation is causally non-local in that it predicts that a charge accelerates in response to *future* fields. Moreover, the standard picture of locally acting fields comes under pressure from the semi-classical Bohm-Aharonov effect. This is the effect that quantum mechanical interference patterns of a beam of electrons passing a very long solenoid vary with the magnetic field inside of the solenoid, even though the magnetic field is zero in the regions through which the electron beam passes. Many philosophers have suggested that this effect forces us to give up either locality or determinism. One might take the effect to be due to the fields acting where they are not, and hence give up locality; or one might argue that the effect shows that potentials ought to be interpreted realistically, even though the 'true' potential is only determined up to a gauge transformation; and finally one could take the Bohm-Aharanov to point to the reality not of localized values of the potential, but of non-local loop integrals of the potential, which turn out to be gauge-invariant (see Healey 2007 for a detailed discussion of various interpretive options).

References:

Balashov, Yuri, and Michel Janssen. 2003. Presentism and Relativity. *Br J Philos Sci* 54, no. 2 (June 1): 327-346. doi:10.1093/bjps/54.2.327.

Belot, Gordon. 2007. Is Classical Electrodynamics an Inconsistent Theory? *Canadian Journal of Philosophy* 37, no. 2 (June): 263-282. doi:Article.

Brown, Harvey, and Pooley, Oliver. 2006. Minkowski space-time: a glorious non-entity. In *The Ontology of Spacetime*. Amsterdam: Elsevier.

Brown, Harvey R. 2005. *Physical Relativity: Space-time Structure from a Dynamical Perspective*. Oxford: Oxford University Press.

Einstein, Albert. 1909. Über die Entwicklung unserer Anschauung über das Wesen und die Konstitution der Strahlung. *Physikalische Zeitschrift* 10: 817-825.

Frisch, Mathias. 2002. Non-Locality in Classical Electrodynamics. *Br J Philos Sci* 53, no. 1 (March 1): 1-19. doi:10.1093/bjps/53.1.1.

---. 2004a. Laws and Initial Conditions. *Philosophy of Science* 71, no. 5: 696-706. doi:10.1086/425056.

---. 2004b. Inconsistency in Classical Electrodynamics. *Philosophy of Science* 71, no. 4: 525-549. doi:10.1086/423627.

---. 2005a. *Inconsistency, Asymmetry, and Non-Locality: A Philosophical Investigation of Classical Electrodynamics*. Oxford: Oxford University Press.

---. 2005b. Mechanisms, principles, and Lorentz's cautious realism. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics* 36, no. 4 (December): 659-679. doi:10.1016/j.shpsb.2005.07.002.

---. 2006. A tale of two arrows. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics* 37, no. 3 (September): 542-558. doi:10.1016/j.shpsb.2005.03.004.

Griffiths, David J. 1999. *Introduction to Electrodynamics*. 3rd ed. Upper Saddle River, N.J: Prentice Hall.

Healey, Richard. 2007. *Gauging What's Real: The Conceptual Foundations of Gauge Theories*. Oxford University Press, USA.

Hogarth, J. E. 1962. Cosmological Considerations of the Absorber Theory of Radiation. *Proceedings of the Royal Society of London* A267: 365-383.

Hoyle, F., and J. V. Narlikar. 1995. Cosmology and Action-at-a-Distance Electrodynamics. *Reviews of Modern Physics* 67: 113-155.

Jackson, John David. 1999. *Classical Electrodynamics*. 3rd ed. New York: Wiley.

Janssen, Michel. 2008. Drawing the Line Between Kinematics and Dynamics in Special Relativity. http://philsci-archive.pitt.edu/archive/00003895/.

Kuhn, Thomas S. 1996. *The Structure of Scientific Revolutions*. 3rd ed. University Of Chicago Press, December 15.

Lange, Marc. 2002. *An Introduction to the Philosophy of Physics: Locality, Fields, Energy and Mass*. Oxford: Blackwell.

Lorentz, H. A. 1916. *The Theory of Electrons and Its Applications to the Phenomena of Light and Radiant Heat*. 2nd ed. Leipzig: B.G. Teubner.

Muller, F. A. 2007. Inconsistency in Classical Electrodynamics? *Philosophy of Science* 74, no. 2 (April): 253-277. doi:10.1086/520942.

North, Jill. 2003. Understanding the Time-Asymmetry of Radiation. *Philosophy of Science* 70, no. 5: 1086-1097. doi:10.1086/377391.

Parrott, Stephen. 1987. *Relativistic Electrodynamics and Differential Geometry*. New York: Springer-Verlag.

Price, Huw. 1996. *Time's Arrow & Archimedes' Point: New Directions for the Physics of Time*. New York: Oxford University Press.

---. 2006. Recent work on the arrow of radiation. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics* 37, no. 3 (September): 498-527. doi:10.1016/j.shpsb.2006.03.004.

Ritz, Walter, and Albert Einstein. 1909. Zum gegenwärtigen Stand des Strahlenproblems. *Physikalische Zeitschrift* 10: 232-324.

Rohrlich, Fritz. 2000. Causality and the Arrow of Classical Time. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics* 31, no. 1 (March): 1-13. doi:10.1016/S1355-2198(99)00030-1.

---. 2004. Realism Despite Cognitive Antireductionism. *International Studies in the Philosophy of Science* 18, no. 1: 73. doi:10.1080/0269859041233128926O.

---. 2006. Time in classical electrodynamics. *American Journal of Physics* 74, no. 4: 313-315.

---. 2007. *Classical Charged Particles*. 3rd ed. Singapore: World Scientific.

Rohrlich, Fritz, and Larry Hardin. 1983. Established Theories. *Philosophy of Science* 50 (December): 603-617.

Rovelli, Carlo. 2004. Comment on: "Causality and the arrow of classical time", by Fritz Rohrlich. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics* 35, no. 3 (September): 397-405. doi:10.1016/j.shpsb.2004.02.005.

Spohn, Herbert. 2004. *Dynamics of Charged Particles and Their Radiation Field.* Cambridge, UK: Cambridge University Press.

Wheeler, John A., and Richard P. Feynman. 1945. Interaction with the Absorber as the Mechanism of Radiation. *Reviews of Modern Physics* 17: 157-181.

Wilson, Mark. 2008. *Wandering Significance: An Essay on Conceptual Behaviour*. New York: Oxford University Press.

Zeh, H. D. 2001. *The Physical Basis of the Direction of Time.* 4th ed. Berlin: Springer.