Time in classical electrodynamics

Fritz Rohrlich[a)]
Syracuse University, Syracuse, New York 13244-1130

(Received 22 August 2005; accepted 27 January 2006)

There is considerable misunderstanding in the literature concerning time in classical electrodynamics. The following conceptual exposition focuses on three points: advanced radiation exists mathematically but not physically; there exists an asymmetry of time due to radiation; and the equations of motion of charged particles are invariant under time reversal. These three points are mutually consistent.
[DOI: 10.1119/1.2178847]

I. UP AND DOWN THE LIGHT CONE

The role of time in classical electrodynamics is treated here conceptually rather than mathematically in order to provide intuitive understanding. A mathematical treatment can be found elsewhere.[1] Because classical electrodynamics involves electromagnetic radiation, it belongs to the special theory of relativity. One of the axioms of relativity is that in a vacuum, light moves with speed c, a universal constant. Radiation is produced by accelerated charges. But unaccelerated charges, although they produce no radiation, also produce electromagnetic fields, that is, Coulomb fields. These also propagate with the speed of light, contrary to their static appearance in the rest frame.[2]

Four-dimensional Minkowski space-time provides a convenient representation for motion in classical electrodynamics. Given a charge located at a point Q with coordinates t_Q and $\boldsymbol{x}_Q$, we can draw a light cone centered at that point. This double cone is given by the equation $(\boldsymbol{x}-\boldsymbol{x}_Q)^2=c^2(t-t_Q)^2$. The future cone, $V_+(Q)$, is characterized by $t>t_Q$, the past cone, $V_-(Q)$, is characterized by $t<t_Q$. All events on these two cones are in the future or past, respectively, of the space-time point Q.

In a vacuum any electromagnetic field produced at Q will arrive at a space-time point P if and only if P is located on $V_+(Q)$. This location expresses the fact that fields propagate only in the positive time direction and implies that as a consequence of *causality*, no point P on $V_-(Q)$ can ever receive electromagnetic fields from Q. The fields that travel on $V_+(Q)$ are called *retarded* fields, and the ones on $V_-(Q)$ are called *advanced* fields. The latter cannot be produced physically but *do* exist mathematically.

For some reason the fact the solutions of Maxwell's equations that describe advanced fields do not exist physically in the real world because of causality has not been fully appreciated and has therefore caused considerable confusion in the literature (see the Appendix). The confusion arose between the mathematical existence of solutions and their physical existence. Such confusion is surprising because there are many equations that have multiple solutions some of which are not realized in nature. An example is the nonrelativistic equation of motion[3] of a charge q, $ma-m\tau\, da/dt=F$ where $\tau=2q^2/(3mc^3)$. For a free charge, $F=0$, and there are two solutions. One is $a=0$ in agreement with the law of inertia, the other is $a(t)=a(0)e^{t/\tau}$ which violates that law and is devoid of physical meaning.

So far, only fields that were *emitted* at Q have been considered. It is also possible for retarded fields produced elsewhere to be absorbed at Q. Because causality tells us that such fields must have been emitted sometime before t_Q, they must have been emitted at a point P' on $V_-(Q)$. They are retarded fields incoming at Q. Thus, fields can be both emitted and absorbed at Q, both kinds being retarded.

There also exist advanced fields that are incoming at Q. Such fields arrive at Q from a point on $V_+(Q)$, that is, from a space-time point in the future. They move in the negative time direction and are therefore not physically realizable.

In summary, there exist four kinds of fields, retarded incoming and outgoing fields, and advanced incoming and outgoing fields. The former move in the positive time direction on the null cone and can exist physically; the latter move in the negative time direction and exist only mathematically.

II. AN ARROW OF TIME

There is a big difference between radiation that is emitted and radiation that is absorbed at Q. If the physical system includes several particles, Maxwell's equations provide quantitative information about the properties of the radiation produced by each of these sources of radiation. A source, P' on $V_-(Q)$ can be a charged particle, a current, or a multipole, for example. For each such emitter, we know how to calculate the energy, momentum, and angular distribution of its radiation.[3] But only one ray of all the radiation emitted by P' travels in the direction of the absorber Q. Having gone a considerable distance, what reaches Q is only a minute fraction of the momentum and energy of the total radiation emitted at P'. Even if many other sources exist on $V_-(Q)$, the energy and momentum of the total radiation absorbed at any instant at Q is generally small compared to what is being emitted. It cannot replenish all the loss due to radiation by Q. Thus, there is on average a net steady flow of radiation energy and momentum from a radiating charge to infinity.

In an open system of charges, the source of the radiation energy and momentum emitted by a particular charge at a point Q on its worldline is an *external* force that comes in part from outside the system and in part from the other charges in the system. There also exist closed systems in which no external forces act, for example, two colliding charged particles.

The asymmetry in time due to the net emission of electromagnetic radiation establishes a preferred time direction sometimes called "the electromagnetic arrow of time." Other arrows of time are well known in physics—the arrow of time due to an increase in entropy of thermodynamic systems, and the cosmological arrow of time that describes the way the universe evolves. There is also a biological arrow of time.

The electromagnetic arrow in special relativity is due to the dissipative nature of radiating systems; it is qualitatively different from all the other time arrows.

III. TIME REVERSAL INVARIANCE

A clear distinction must be made between "time symmetry" and "time reversal invariance." We can formulate classical electrodynamics in a Lagrangian form using an action integral that is completely time symmetric: the future and the past are treated equally.[4] Such a formulation of classical electrodynamics has the advantage of avoiding the static self-energy problem (self-energy due to the Coulomb field). However, time symmetry is not the same as time reversal invariance. In the former, the future and the past are treated simultaneously in a symmetric way. In particular, retarded and advanced fields are treated equally. Such a treatment is physically unsatisfactory in view of the nonexistence of advanced fields: only retarded fields should occur in a physically meaningful formulation of the theory.

Time reversal invariance is a property with respect to a particular transformation. It is an invariance property like space inversion, or (for continuous symmetry groups) Lorentz invariance and Galilean invariance. This invariance means that the *form* of the equations remains unchanged. After the transformation, the equations have the same form with respect to the new variables that they had with respect to the old ones.

The time reversal transformation T is defined in Minkowski space-time by the inversion of a small time interval, $\Delta t \to \Delta t' = -\Delta t$ while the position remains unchanged. This transformation is local; it takes place at each space-time point. When a charged particle is represented in Minkowski space by a timelike line with an arrow in the positive time direction, t, it is transformed into the same line with an arrow in the opposite time direction, t'. If the charge has position $\boldsymbol{x}(t)$ and velocity $\boldsymbol{v}(t)$, then the transformation is $\boldsymbol{x}(t) \to \boldsymbol{x}'(t') = \boldsymbol{x}(t')$ and $\boldsymbol{v}(t) \to \boldsymbol{v}'(t') = -\boldsymbol{v}(t')$ because the time derivative causes a sign change. The transformation T makes the charge move on the same worldline but in the opposite direction. Time reversal means motion reversal. The motion from Q_1 to Q_2 during the time interval $t_2 - t_1 > 0$ is reversed so that it becomes a motion from Q_2 to Q_1 during the time interval $t_1' - t_2' = -t_1 + t_2 > 0$.

The electromagnetic fields of a charge are produced at the location of the charge. The presence of the charge ensures the propagation of a Coulomb field emanating from the charge. In addition, if the charge is accelerated, it also produces a radiation field. The locality of the transformation is crucial. Mathematically, both retarded and advanced fields can be produced. The light cones of these fields are toward the future and the past, respectively. But future and past are characterized by the direction of the velocity four-vector at the emission point Q. In particular, a retarded field will propagate along the future light cone from Q. After the time reversal transformation, the direction of the velocity four-vector is reversed. Therefore, the future and the past of the light cones change places and retarded radiation emission follows suit. The transformed velocity four-vector and the future light cone are both in the new direction opposite to their original direction, and the retarded field will be emitted accordingly. Thus, a charge will send out retarded fields both before and after T and always in the time direction that is positive for that charge. It is a consequence of the locality of T: retarded fields transform into retarded fields and advanced fields transform into advanced fields.[5]

Note that the positive time direction must be known before one can speak of radiation. We should think of the direction of the future light cone as being carried along with the transformation, specifically with the inversion of direction of the four-velocity. Radiation "knows" which time direction is positive because of the way it is emitted, that is, always into the future rather than the past. Therefore, radiation remains retarded after time inversion.

For a single charged particle, Newton's equation of motion no longer applies because it ignores the existence of radiation. Neither the energy nor the momentum of radiation is contained in Newton's equation, either nonrelativistically or relativistically. Additional terms are necessary. These additional terms comprise the self-force, the force that the particle exerts on itself because it is charged. Relativistically, this self-force is a four-vector. Because we deal with physical systems, the self-force is always due to retarded fields. Thus, the equation of motion of a charged particle reads $ma = F_{\text{ext}} + F_{\text{ret}}$. The new force, F_{ret}, is the Lorentz force that the charge exerts on itself by means of its own retarded field.

A detailed calculation confirms that this equation remains invariant[1] under T. It is almost obvious: given that Newton's equation for an electrically neutral particle, $ma = F_{\text{ext}}$, remains invariant under T, the additional fact we have established that retarded fields remain retarded, also ensures the invariance of the equation of motion for a charged particle, $ma = F_{\text{ext}} + F_{\text{ret}}$.

When the retarded field is expanded in powers of the radius r of the charged particle, and when all positive powers of r are neglected, the resulting self-force in the equation of motion (the Lorentz-Abraham-Dirac equation)[5] contains both first and second time derivatives of the velocity. How can this expression possibly be invariant under T?

The answer is that this expansion is specific to the time direction t. We must first reverse the time and then expand F_{ret}. Then we obtain the same expansion but the independent time variable is now t' and the derivatives are now with respect to t'. We must expand after the positive time direction is specified, not before.

IV. THE COHERENCE OF PHYSICAL THEORIES

Physical theories on different scales are mathematically interrelated even if they involve qualitatively different concepts. For example, for small velocities, relativistic dynamics becomes nonrelativistic dynamics in the sense that the relativistic equations reduce to the nonrelativistic ones. Similarly, in the limit as the charge of a particle goes to zero (or becomes negligibly small), its equation of motion becomes the equation of motion of a neutral particle. This fact is called the "principle of undetectability of small charges" (see Ref. 5, p. 214).

Theories on a "deeper" level must reduce mathematically to theories on a "shallower" level. An example is quantum electrodynamics (QED). It must reduce to classical electrodynamics in a suitable limit, which implies that classical electrodynamics has only a finite domain of validity. Similarly, general relativity reduces to Newtonian gravitation theory and wave optics becomes ray optics. These limits are all well known and can be discussed on a more philosophical level.[6] The boundary of the validity domain of a theory is not sharp but depends on the desired accuracy of prediction.

In QED the distinction between retarded and advanced fields is also made. Indeed, advanced fields are explicitly excluded in QED. The proof of this assertion can be found in the first treatise of QED that was published after the great progress in QED made after World War II by Dyson, Feynman, Schwinger, and Tomonaga.[7] In perturbation theory of QED, electromagnetic radiation is described by photons. These are either emitted or absorbed by charged particles or else travel between them providing electromagnetic interaction. In the latter case, they are the "internal photon lines" of Feynman diagrams. These photon lines are described by "propagators"[8] that are essentially the quantum generalizations of classical Green functions. Analysis of these photon lines shows that in the classical limit, they become the *retarded* Green functions. The QED equivalents of advanced electromagnetic fields do not occur. A similar analysis also shows that advanced electron fields do not occur. There are no advanced fields in QED.

APPENDIX

The question of the physical meaning of advanced electromagnetic fields has been discussed for many years and deserves a detailed historical study. We only indicate here how some texts on classical electrodynamics have dealt with it.

In a 1938 paper on the derivation of the equation of motion of a charged particle, Dirac used the advanced field as a formal trick to separate the retarded field into a time-symmetric and a time-antisymmetric part,

$$F_{\text{ret}} = \tfrac{1}{2}(F_{\text{ret}} + F_{\text{adv}}) + \tfrac{1}{2}(F_{\text{ret}} - F_{\text{adv}}). \quad (1)$$

In this formal manipulation, no physical meaning needs to be ascribed to the advanced field itself because the two parts always occur together. However, the two parts differ physically. The first part is symmetric in time and falls off with distance as $1/r^2$; it is the (generalized) Coulomb field. The second part is antisymmetric in time and falls off as $1/r$; it is the radiation field. Because any source always emits both parts simultaneously, the advanced field does not enjoy an existence separate from the retarded field. But it is convenient to treat the Coulomb field and the radiation field separately.

After Dirac, Wheeler and Feynman developed an action-at-a-distance theory.[9] They postulated a cosmology in which all the radiation emitted is also completely absorbed so that the second part of Eq. (1) is effectively absent leaving only the Coulomb field. The latter is then treated as an action-at-a-distance. But such an action violates a basic assumption of relativity theory: no particle or field can move faster than the speed of light. Also, the cosmological assumption of complete absorption requires a very large number of absorbers. Although the complete absorber theory enjoyed a certain period of interest, it also gave the advanced fields a physical role on par with retarded fields.

Panofsky and Phillips wrote in the second edition (not in the first) of their text book: "… in electrodynamics, the field equations are basically symmetric in time, the asymmetry is introduced only by our arbitrary rejection of the advanced solution".[10] Later textbooks realized that advanced fields violate causality.[11] However, the claim that advanced fields are physically meaningful still survives in some of the literature.

Apart from the confusion between the time symmetry of classical electrodynamics (which does not hold in nature) and time reversal invariance (which does), another error entered the literature: The transformation of time reversal was done incorrectly in my own textbook.[5] It was claimed incorrectly that this transformation turns retarded into advanced fields and *vice versa*. When the retarded electromagnetic field is expressed by integration of its source over a retarded Green function, time reversal changes the retarded Green function into an advanced Green function yielding an advanced field. But this time inversion is incorrect because it is *not a local* transformation: the whole light cone produced by the current is thereby time reversed. A local transformation inverts time at a single point on the particle's world line thus causing *motion reversal*. The future light cone is reversed as a consequence. The retarded field produced by the velocity $\mathbf{v}$ and the acceleration $\mathbf{a}$ at one single space time point on the worldline of the charge is time-inverted into $\mathbf{v}'$ and $\mathbf{a}'$, which again produce a retarded field but relative to the new time direction. Fields are always emitted into the future (the positive time direction) which must therefore be specified *before* Maxwell's equations can be applied. When the motion is inverted before the fields are computed, classical electrodynamics is indeed found to be invariant under the local time reversal transformation.[1]

[a]Electronic address: rohrlich@syr.edu

[1]F. Rohrlich, "Time reversal invariance and the arrow of time in classical electro-dynamics," Phys. Rev. E **72**, 057601-1–4 (2005).

[2]F. Rohrlich, "Causality, the Coulomb field, and Newton's law of gravitation," Am. J. Phys. **70**(4), 411–414 (2002).

[3]See, for example, J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1999).

[4]F. Rohrlich, "Solution of the classical electromagnetic self-energy problem," Phys. Rev. Lett. **12**, 375–377 (1964).

[5]F. Rohrlich, *Classical Charged Particles* (Addison-Wesley, New York, 1965, 1990). The time reversal transformation given on pp. 248ff, is incorrect. See the Appendix.

[6]F. Rohrlich, "Realism despite cognitive antireductionism," Int. Stud. Phil. Sci. **18**, 73–88 (2004).

[7]J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons*, 2nd ed. (Addison-Wesley, New York, 1955) (Springer-Verlag, New York, 1976).

[8]See Ref. 7, Appendix A1. These photon propagators have various names: Feynman propagator D_F or causal propagator D_c are the most common.

[9]J. A. Wheeler and R. P. Feynman, Rev. Mod. Phys. **17**, 157–181 (1945); **21**, 425–433 (1949).

[10]W. Panofsky and M. Phillips, *Classical Electricity and Magnetism*, 2nd ed. (Addison-Wesley, New York, 1962), p. 395.

[11]See, for example, D. J. Griffith, *Introduction to Electrodynamics*, 3rd ed. (Prentice-Hall, Englewood, NJ, 1999), p. 425.