Presenting electromagnetic theory in accordance with the principle of causality

Oleg D Jefimenko

Eur. J. Phys. **25** (2004) 287–296 PII: S0143-0807(04)69749-2

Presenting electromagnetic theory in accordance with the principle of causality

Oleg D Jefimenko

Physics Department, West Virginia University, PO Box 6315, Morgantown, WV 26506-6315, USA

Received 1 October 2003
Published 28 January 2004
Online at stacks.iop.org/EJP/25/287 (DOI: 10.1088/0143-0807/25/2/015)

Abstract

A method of presenting electromagnetic theory in accordance with the principle of causality is described. Two 'causal' equations expressing time-dependent electric and magnetic fields in terms of their causative sources by means of retarded integrals are used as the fundamental electromagnetic equations. Maxwell's equations are derived from these 'causal' equations. Except for the fact that Maxwell's equations appear as derived equations, the presentation is completely compatible with Maxwell's electromagnetic theory. An important consequence of this method of presentation is that it offers new insights into the cause-and-effect relations in electromagnetic phenomena and results in simpler derivations of certain electromagnetic equations.

1. Introduction

One of the most important tasks of physics is to establish causal relations between physical phenomena. No physical theory can be complete unless it provides a clear statement and description of causal links involved in the phenomena encompassed by that theory. In establishing and describing causal relations it is important not to confuse equations which we call 'basic laws' with 'causal equations'. A 'basic law' is an equation (or a system of equations) from which we can derive most (hopefully all) possible correlations between the various quantities involved in a particular group of phenomena subject to the 'basic law'. A 'causal equation', on the other hand, is an equation that unambiguously relates a quantity representing an effect to one or more quantities representing the cause of this effect. Clearly, a 'basic law' need not constitute a causal relation, and an equation depicting a causal relation may not necessarily be among the 'basic laws' in the above sense.

Causal relations between phenomena are governed by the *principle of causality*. According to this principle, all present phenomena are exclusively determined by past events. Therefore equations depicting causal relations between physical phenomena must, in general, be equations where a present-time quantity (the effect) relates to one or more quantities (causes) that existed at some previous time. An exception to this rule are equations constituting causal

0143-0807/04/020287+10$30.00

relations by definition; for example, if force is defined as the cause of acceleration, then the equation $\mathbf{F} = m\mathbf{a}$, where $\mathbf{F}$ is the force and $\mathbf{a}$ is the acceleration, is a causal equation by definition.

In general, then, according to the principle of causality, an equation between two or more quantities simultaneous in time but separated in space cannot represent a causal relation between these quantities. In fact, even an equation between quantities simultaneous in time and not separated in space cannot represent a causal relation between these quantities because, according to this principle, the cause *must precede* its effect. Therefore the only kind of equations representing causal relations between physical quantities, other than equations representing cause and effect by definition, must be equations involving 'retarded' (previous-time) quantities.

Let us apply these considerations to the basic electromagnetic field laws. Traditionally these laws are represented by the four Maxwell's equations, which, in their differential form, are

$$\nabla \cdot \mathbf{D} = \rho, \tag{1}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{2}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{3}$$

and

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}, \tag{4}$$

where $\mathbf{E}$ is the electric field vector, $\mathbf{D}$ is the displacement vector, $\mathbf{H}$ is the magnetic field vector, $\mathbf{B}$ is the magnetic flux density vector, $\mathbf{J}$ is the current density vector, and ρ is the electric charge density. For fields in a vacuum, Maxwell's equations are supplemented by the two constitutive equations,

$$\mathbf{D} = \varepsilon_0 \mathbf{E} \tag{5}$$

and

$$\mathbf{B} = \mu_0 \mathbf{H}, \tag{6}$$

where ε_0 is the permittivity of space, and μ_0 is the permeability of space.

Since none of the four Maxwell's equations is defined to be a causal relation, and since each of these equations connects quantities simultaneous in time, none of these equations represents a causal relation. That is, $\nabla \cdot \mathbf{D}$ is not a consequence of ρ (and vice versa), $\nabla \times \mathbf{E}$ is not a consequence of $\partial \mathbf{B}/\partial t$ (and vice versa), and $\nabla \times \mathbf{H}$ is not a consequence of $\mathbf{J} + \partial \mathbf{D}/\partial t$ (and vice versa). Thus, Maxwell's equations, even though they are basic electromagnetic equations (since most electromagnetic relations are derivable from them), do not depict cause-and-effect relations between electromagnetic phenomena and leave the question of causality in electromagnetic phenomena unanswered.

The purpose of this paper is to show how Maxwellian electromagnetic theory can be reformulated and presented in a classroom in compliance with the principle of causality so that the causal relations between fundamental electromagnetic phenomena are clearly revealed.

2. Causal equations for electric and magnetic fields

A reformulation and presentation of Maxwell's electromagnetic theory in accordance with the principle of causality must be based on causal electromagnetic equations that are at least as general as Maxwell's equations and are in complete accord with the latter. What should be the form of such causal electromagnetic equations? Since an effect can be a combined or cumulative result of several causes, it is plausible that in causal equations a quantity representing an effect should be expressed in terms of *integrals* involving quantities representing the various causes of that effect. And since, by the principle of causality, the cause must precede its effect, the integrals in causal equations must be *retarded*, that is, the integrands in these integrals must involve quantities as they existed at a time prior to the time for which the quantity representing the effect is being computed.

The following equations for the electric fields **E** and the magnetic field **H** in a vacuum satisfy the above requirements for causal electromagnetic equations, and we shall use them as basic laws for presenting Maxwell's electromagnetic theory in compliance with the principle of causality:

$$\mathbf{E} = \frac{1}{4\pi\varepsilon_0}\int\left\{\frac{[\rho]}{r^3} + \frac{1}{r^2 c}\frac{\partial[\rho]}{\partial t}\right\}\mathbf{r}\,\mathrm{d}V' - \frac{1}{4\pi\varepsilon_0 c^2}\int\frac{1}{r}\left[\frac{\partial\mathbf{J}}{\partial t}\right]\mathrm{d}V', \tag{7}$$

and

$$\mathbf{H} = \frac{1}{4\pi}\int\left\{\frac{[\mathbf{J}]}{r^3} + \frac{1}{r^2 c}\left[\frac{\partial\mathbf{J}}{\partial t}\right]\right\}\times\mathbf{r}\,\mathrm{d}V'. \tag{8}$$

The square brackets in these equations are the retardation symbol indicating that the quantities between the brackets are to be evaluated for the 'retarded' time $t' = t - r/c$, where t is the time for which **E** and **H** are evaluated, ρ is the electric charge density, **J** is the current density, r is the distance between the field point x, y, z (point for which **E** and **H** are evaluated) and the source point x', y', z' (volume element $\mathrm{d}V'$), and c is the velocity of light. The integrals are extended over all space.

According to equation (7), the electric field has three causative sources: the retarded charge density $[\rho]$, the retarded time derivative of the charge density $\partial[\rho]/\partial t$, and the retarded time derivative of the current density $[\partial\mathbf{J}/\partial t]$. Likewise, according to equation (8), the magnetic field has two causative sources: the retarded current density $[\mathbf{J}]$ and the retarded time derivative of the current density $[\partial\mathbf{J}/\partial t]$.

As we shall presently see, in order to be equivalent to Maxwell's electromagnetic theory, the electromagnetic theory based on the causal equations (7) and (8) needs a third basic equation: the familiar continuity equation representing the conservation of electric charge

$$\nabla\cdot\mathbf{J} = -\frac{\partial\rho}{\partial t}. \tag{9}$$

Furthermore, equations (7)–(9) need to be supplemented by the constitutive relations, equations (5) and (6), unless only the fields **E** and **H**, without the fields **D** and **B**, are used.

In the mode of presentation of electromagnetic theory described here, the three laws, equations (7)–(9), are postulated, and their correctness is proved by demonstrating that they are in complete agreement with Maxwellian electrodynamics, that is, by demonstrating that Maxwell's equations can be derived from them. Therefore it is not necessary to discuss the original considerations that led to the formulation of equations (7)–(9). It may be noted, however, that equations (7) and (8) have been originally obtained from inhomogeneous equations for electromagnetic waves [1] and can also be obtained from the retarded electromagnetic potentials [2] as well as from the wave equation for the retarded electromagnetic potentials [3].

3. Deriving Maxwell's equations from the causal equations for E and H

To derive Maxwell's equations from the causal equations (7) and (8), we first transform these equations to a somewhat different form with the help of vector identities listed in the appendix. Using vector identity (A.4), we replace the two terms in the integrand of the first integral of equation (7) by a single term, obtaining

$$\mathbf{E} = -\frac{1}{4\pi\varepsilon_0}\int\nabla\frac{[\rho]}{r}\,\mathrm{d}V' - \frac{1}{4\pi\varepsilon_0 c^2}\int\frac{1}{r}\left[\frac{\partial\mathbf{J}}{\partial t}\right]\mathrm{d}V'. \tag{10}$$

Transforming now the integrand in the first integral of equation (10) by means of vector identity (A.5), we obtain (note that the ordinary operator ∇ operates upon the field-point coordinates, whereas the primed operator ∇' operates upon the source-point coordinates)

$$\mathbf{E} = -\frac{1}{4\pi\varepsilon_0}\int \frac{[\nabla'\rho]}{r}\,\mathrm{d}V' + \frac{1}{4\pi\varepsilon_0}\int \nabla'\frac{[\rho]}{r}\,\mathrm{d}V' - \frac{1}{4\pi\varepsilon_0 c^2}\int \frac{1}{r}\left[\frac{\partial \mathbf{J}}{\partial t}\right]\mathrm{d}V'. \tag{11}$$

The second integral in the last expression can be transformed into a surface integral by means of vector identity (A.2). But this surface integral vanishes, because ρ is confined to a finite region of space, while the surface of integration is at infinity. We thus have

$$\mathbf{E} = -\frac{1}{4\pi\varepsilon_0}\int \frac{[\nabla'\rho]}{r}\,\mathrm{d}V' - \frac{1}{4\pi\varepsilon_0 c^2}\int \frac{1}{r}\left[\frac{\partial \mathbf{J}}{\partial t}\right]\mathrm{d}V'. \tag{12}$$

Similarly, applying vector identity (A.4) to equation (8), we obtain

$$\mathbf{H} = \frac{1}{4\pi}\int \nabla \times \frac{[\mathbf{J}]}{r}\,\mathrm{d}V'. \tag{13}$$

Transforming equation (13) by means of vector identity (A.5) and eliminating $\nabla' \times \{[\mathbf{J}]/r\}$ by means of vector identity (A.3) (see the explanation below equation (11); note that $\mathbf{J}$ is confined to a finite region of space), we obtain for the magnetic field

$$\mathbf{H} = \frac{1}{4\pi}\int \frac{[\nabla' \times \mathbf{J}]}{r}\,\mathrm{d}V'. \tag{14}$$

Maxwell's equations can now be obtained from equations (12), (14) and (9) as follows:

The first Maxwell's equation. From equation (12) we have

$$\begin{aligned}\nabla\cdot\mathbf{E} &= -\frac{1}{4\pi\varepsilon_0}\nabla\cdot\int \frac{[\nabla'\rho]}{r}\,\mathrm{d}V' - \frac{1}{4\pi\varepsilon_0 c^2}\nabla\cdot\int \frac{1}{r}\left[\frac{\partial \mathbf{J}}{\partial t}\right]\mathrm{d}V' \\ &= -\frac{1}{4\pi\varepsilon_0}\int \nabla\cdot\frac{[\nabla'\rho]}{r}\,\mathrm{d}V' - \frac{1}{4\pi\varepsilon_0 c^2}\int \nabla\cdot\frac{1}{r}\left[\frac{\partial \mathbf{J}}{\partial t}\right]\mathrm{d}V',\end{aligned} \tag{15}$$

(the operator ∇ can be placed under the integral sign because it operates on the field-point coordinates x, y, z, while the integration is over the source-point coordinates x', y', z'). Applying vector identity (A.5) to equation (15) and eliminating $\nabla' \cdot \{[\nabla'\rho]/r\}$ by means of vector identity (A.1) (see the explanation below equation (11)), we obtain

$$\nabla\cdot\mathbf{E} = -\frac{1}{4\pi\varepsilon_0}\int \frac{[\nabla'\cdot\nabla'\rho]}{r}\,\mathrm{d}V' - \frac{1}{4\pi\varepsilon_0 c^2}\int \frac{1}{r}\left[\frac{\partial(\nabla'\cdot\mathbf{J})}{\partial t}\right]\mathrm{d}V', \tag{16}$$

and, using now equation (9) to replace $\nabla'\cdot\mathbf{J}$ by $-\partial\rho/\partial t$, we get

$$\nabla\cdot\mathbf{E} = -\frac{1}{4\pi\varepsilon_0}\int \frac{\left[\nabla'\cdot\nabla'\rho - \frac{1}{c^2}\frac{\partial^2\rho}{\partial t^2}\right]}{r}\,\mathrm{d}V'. \tag{17}$$

According to vector identity (A.6), the right side of equation (17) is simply $(1/\varepsilon_0)\rho$, so that, replacing $\mathbf{E}$ in equation (17) by $\mathbf{D}$ with the help of equation (5), we obtain the first Maxwell's equation

$$\nabla\cdot\mathbf{D} = \rho. \tag{1}$$

The second Maxwell's equation. From equation (14) we have

$$\nabla\cdot\mathbf{H} = \frac{1}{4\pi}\nabla\cdot\int \frac{[\nabla'\times\mathbf{J}]}{r}\,\mathrm{d}V' = \frac{1}{4\pi}\int \nabla\cdot\frac{[\nabla'\times\mathbf{J}]}{r}\,\mathrm{d}V' \tag{18}$$

(the operator ∇ can be placed under the integral sign because it operates on the field-point coordinates x, y, z, while the integration is over the source-point coordinates x', y', z'). Applying vector identity (A.5) to equation (18) and eliminating $\nabla' \cdot \{[\nabla' \times \mathbf{J}]/r\}$ by means of vector identity (A.1) (see the explanation below equation (11); note that **J** is confined to a finite region of space), we obtain

$$\nabla \cdot \mathbf{H} = \frac{1}{4\pi} \int \frac{[\nabla' \cdot (\nabla' \times \mathbf{J})]}{r} \, \mathrm{d}V', \tag{19}$$

and since $\nabla' \cdot \nabla' \times \equiv 0$,

$$\nabla \cdot \mathbf{H} = 0, \tag{20}$$

which, by equation (6), yields the second Maxwell's equation

$$\nabla \cdot \mathbf{B} = 0. \tag{2}$$

The third Maxwell's equation. From equation (12) we have

$$\begin{aligned} \nabla \times \mathbf{E} &= -\frac{1}{4\pi\varepsilon_0} \nabla \times \int \frac{[\nabla'\rho]}{r} \, \mathrm{d}V' - \frac{1}{4\pi\varepsilon_0 c^2} \nabla \times \int \frac{1}{r} \left[\frac{\partial \mathbf{J}}{\partial t} \right] \mathrm{d}V' \\ &= -\frac{1}{4\pi\varepsilon_0} \int \nabla \times \frac{[\nabla'\rho]}{r} \, \mathrm{d}V' - \frac{1}{4\pi\varepsilon_0 c^2} \int \nabla \times \frac{1}{r} \left[\frac{\partial \mathbf{J}}{\partial t} \right] \mathrm{d}V', \end{aligned} \tag{21}$$

where we have placed the operator ∇ under the integral sign (this can be done because ∇ operates on unprimed coordinates, while the integration is over the primed coordinates). Applying vector identity (A.5) to equation (21) and eliminating $\nabla' \times \{[\nabla'\rho]/r\}$ by means of vector identity (A.3) (see the explanation below equation (11)), we obtain

$$\nabla \times \mathbf{E} = -\frac{1}{4\pi\varepsilon_0} \int \frac{[\nabla' \times \nabla'\rho]}{r} \, \mathrm{d}V' - \frac{1}{4\pi\varepsilon_0 c^2} \int \nabla \times \frac{1}{r} \left[\frac{\partial \mathbf{J}}{\partial t} \right] \mathrm{d}V', \tag{22}$$

and since $\nabla' \times \nabla' \equiv 0$,

$$\nabla \times \mathbf{E} = -\frac{1}{4\pi\varepsilon_0 c^2} \int \nabla \times \frac{1}{r} \left[\frac{\partial \mathbf{J}}{\partial t} \right] \mathrm{d}V'. \tag{23}$$

Let us now multiply equation (14) by μ_0 and differentiate it with respect to time. Since, by equation (6), $\mathbf{B} = \mu_0 \mathbf{H}$, we have

$$\frac{\partial \mathbf{B}}{\partial t} = \mu_0 \frac{\partial \mathbf{H}}{\partial t} = \frac{\mu_0}{4\pi} \frac{\partial}{\partial t} \int \frac{[\nabla' \times \mathbf{J}]}{r} \, \mathrm{d}V'. \tag{24}$$

Transforming the integrand in equation (24) by means of the vector identities (A.5) and eliminating $\nabla' \times \{[\nabla' \times \mathbf{J}]/r\}$ by means of vector identity (A.3) (see the explanation below equation (11); note that **J** is confined to a finite region of space), we obtain

$$\frac{\partial \mathbf{B}}{\partial t} = \mu_0 \frac{\partial \mathbf{H}}{\partial t} = \frac{\mu_0}{4\pi} \frac{\partial}{\partial t} \int \nabla \times \frac{[\mathbf{J}]}{r} \, \mathrm{d}V'. \tag{25}$$

Differentiating under the integral sign and taking into account that $\partial[\mathbf{J}]/\partial t = [\partial \mathbf{J}/\partial t]$ and that $\mu_0 = 1/\varepsilon_0 c^2$, we obtain

$$\frac{\partial \mathbf{B}}{\partial t} = \frac{1}{4\pi\varepsilon_0 c^2} \int \nabla \times \frac{1}{r} \left[\frac{\partial \mathbf{J}}{\partial t} \right] \mathrm{d}V', \tag{26}$$

which together with equation (22) yields the third Maxwell's equation

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}. \tag{3}$$

The Fourth Maxwell's Equation. From equation (14) we have

$$\nabla \times \mathbf{H} = \frac{1}{4\pi} \nabla \times \int \frac{[\nabla' \times \mathbf{J}]}{r} \, \mathrm{d}V' = \frac{1}{4\pi} \int \nabla \times \frac{[\nabla' \times \mathbf{J}]}{r} \, \mathrm{d}V', \tag{27}$$

where we have placed the operator ∇ under the integral sign (this can be done because ∇ operates on unprimed coordinates, while the integration is over the primed coordinates). Applying vector identity (A.5) to equation (27) and eliminating $\nabla' \times \{[\nabla' \times \mathbf{J}]/r\}$ by means of vector identity (A.3) (see the explanation below equation (11); note that $\mathbf{J}$ is confined to a finite region of space), we obtain

$$\nabla \times \mathbf{H} = \frac{1}{4\pi}\int \frac{[\nabla' \times (\nabla' \times \mathbf{J})]}{r}\,\mathrm{d}V'. \tag{28}$$

Let us now find the time derivative of $\mathbf{D}$ by using equation (12). Since, by equation (5), $\mathbf{D} = \varepsilon_0\mathbf{E}$, we have

$$\begin{aligned}\frac{\partial \mathbf{D}}{\partial t} = \varepsilon_0\frac{\partial \mathbf{E}}{\partial t} &= -\frac{1}{4\pi}\int \frac{\partial}{\partial t}\frac{[\nabla'\rho]}{r}\,\mathrm{d}V' - \frac{1}{4\pi c^2}\int \frac{\partial}{\partial t}\frac{1}{r}\left[\frac{\partial \mathbf{J}}{\partial t}\right]\mathrm{d}V' \\ &= -\frac{1}{4\pi}\int \frac{\left[\nabla'\frac{\partial\rho}{\partial t}\right]}{r}\,\mathrm{d}V' - \frac{1}{4\pi c^2}\int \frac{1}{r}\left[\frac{\partial^2 \mathbf{J}}{\partial t^2}\right]\mathrm{d}V' \end{aligned} \tag{29}$$

and, making use of the continuity equation, equation (3), we obtain

$$\frac{\partial \mathbf{D}}{\partial t} = \frac{1}{4\pi}\int \frac{[\nabla'(\nabla' \cdot \mathbf{J})]}{r}\,\mathrm{d}V' - \frac{1}{4\pi c^2}\int \frac{1}{r}\left[\frac{\partial^2 \mathbf{J}}{\partial t^2}\right]\mathrm{d}V'. \tag{30}$$

Next, let us subtract equation (30) from (28). Placing the derivative $\partial\mathbf{D}/\partial t$ on the right side of the resulting equation and replacing the three integrals by a single integral, we obtain

$$\nabla \times \mathbf{H} = -\frac{1}{4\pi}\int \frac{\left[\nabla'(\nabla' \cdot \mathbf{J}) - \nabla' \times (\nabla' \times \mathbf{J}) - \frac{1}{c^2}\frac{\partial^2 \mathbf{J}}{\partial t^2}\right]}{r}\,\mathrm{d}V' + \frac{\partial \mathbf{D}}{\partial t}. \tag{31}$$

But, according to vector identity (A.7), the first term on the right in equation (30) is simply the current density $\mathbf{J}$. Replacing this term by $\mathbf{J}$, we obtain the fourth Maxwell's equation

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}. \tag{4}$$

4. Discussion

Although the presentation of electromagnetic theory on the basis of the causal equations for electric and magnetic fields, equations (7) and (8), is somewhat more complex than the traditional presentation based directly on Maxwell's equations, such a presentation, as we shall presently see, simplifies the derivation of some electromagnetic formulae, offers important new insights into certain electromagnetic phenomena and dispels certain erroneous views on electromagnetic cause-and-effect relations. In particular, as is explained below, the presentation based on the causal electromagnetic equations shows that the traditional explanation of the very important phenomenon of electromagnetic induction is incorrect and reinforces the original explanation of this phenomenon provided by Faraday and Maxwell.

But let us first demonstrate how the presentation of electromagnetic theory based on the causal electromagnetic equations simplifies the derivation of some representative electromagnetic formulae.

To start with, let us note that in time-independent systems there is no retardation and the time derivatives vanish. Therefore, for time-independent systems we immediately obtain from equation (7) the Coulomb field equation,

$$\mathbf{E} = \frac{1}{4\pi\varepsilon_0}\int \frac{\rho}{r^3}\mathbf{r}\,\mathrm{d}V', \tag{32}$$

and from equation (8) we immediately obtain the Biot–Savart law,

$$\mathbf{H} = \frac{1}{4\pi}\int \frac{\mathbf{J}}{r^3} \times \mathbf{r}\,\mathrm{d}V'. \tag{33}$$

Next, let us quickly derive some formulae whose derivation requires considerable effort in the conventional presentation of electromagnetic theory.

Factoring out the operator ∇ from under the first integral of equation (10), we immediately obtain the relation for the retarded electric scalar potential ϕ

$$\mathbf{E} = -\nabla\varphi - \frac{1}{4\pi\varepsilon_0 c^2}\int \frac{1}{r}\left[\frac{\partial \mathbf{J}}{\partial t}\right] \mathrm{d}V', \tag{34}$$

where φ is

$$\varphi = \frac{1}{4\pi\varepsilon_0}\int \frac{[\rho]}{r}\,\mathrm{d}V', \tag{35}$$

reducing to the ordinary scalar potential

$$\varphi = \frac{1}{4\pi\varepsilon_0}\int \frac{\rho}{r}\,\mathrm{d}V', \tag{36}$$

for time-independent electric field.

Likewise, factoring out the operator ∇ from under the integral of equation (13) and using equation (6), we immediately obtain the equations for the retarded vector potential $\mathbf{A}$

$$\mathbf{B} = \nabla \times \mathbf{A}, \tag{37}$$

with

$$\mathbf{A} = \frac{\mu_0}{4\pi}\int \frac{[\mathbf{J}]}{r}\,\mathrm{d}V', \tag{38}$$

which reduces to the ordinary vector potential

$$\mathbf{A} = \frac{\mu_0}{4\pi}\int \frac{\mathbf{J}}{r}\,\mathrm{d}V' \tag{39}$$

in the case of time-independent magnetic field.

Next, taking into account that $\mu_0 = 1/\varepsilon_0 c^2$, and noting that the last term in equation (34) is the partial time derivative of the retarded vector potential $\mathbf{A}$ given by equation (38), we obtain the equation expressing time-dependent electric field in terms of the retarded scalar electric potential and the retarded magnetic vector potential:

$$\mathbf{E} = -\nabla\varphi - \frac{\partial \mathbf{A}}{\partial t}. \tag{40}$$

To complete these representative derivations, we shall now quickly obtain the 'Lorenz condition'. Evaluating $\nabla \cdot \mathbf{A}$ by using equation (38), using vector identity (A.5), and then eliminating $\nabla' \cdot \{[\mathbf{J}]/r\}$ by means of vector identity (A.1), we have

$$\nabla \cdot \mathbf{A} = \frac{\mu_0}{4\pi}\int \nabla \cdot \frac{[\mathbf{J}]}{r}\,\mathrm{d}V' = \frac{\mu_0}{4\pi}\int \left\{\frac{[\nabla' \cdot \mathbf{J}]}{r} - \nabla' \cdot \frac{[\mathbf{J}]}{r}\right\}\mathrm{d}V' = \frac{\mu_0}{4\pi}\int \frac{[\nabla' \cdot \mathbf{J}]}{r}\,\mathrm{d}V'. \tag{41}$$

Using the continuity equation, equation (9), we replace $\nabla' \cdot \mathbf{J}$ by $-\partial\rho/\partial t$, obtaining

$$\nabla \cdot \mathbf{A} = -\frac{\mu_0}{4\pi}\int \frac{1}{r}\left[\frac{\partial \rho}{\partial t}\right]\mathrm{d}V' = -\frac{\mu_0}{4\pi}\frac{\partial}{\partial t}\int \frac{[\rho]}{r}\,\mathrm{d}V'. \tag{42}$$

Replacing the integral in equation (42) with the help of (35), we obtain the Lorenz condition (frequently erroneously referred to as the 'Lorentz condition')

$$\nabla \cdot \mathbf{A} = -\varepsilon_0\mu_0\frac{\partial \varphi}{\partial t}. \tag{43}$$

And now let us discuss the very important consequence of the causal equation, equations (7) and (8), pertaining to the phenomenon of electromagnetic induction. There is a widespread belief that time-variable electric and magnetic fields can cause, 'induce', each other.

It is traditionally asserted that, according to Maxwell's equation (3), a changing magnetic field produces an electric field ('Faraday induction') and that, according to Maxwell's equation (4), a changing electric field produces a magnetic field ('Maxwell induction'). The very useful and successful method of calculating induced voltage (emf) in terms of changing magnetic flux appears to support the reality of Faraday induction. And the existence of electromagnetic waves appears to support the reality of both Faraday induction and Maxwell induction. Note, however, that as explained in section 1, Maxwell's equation (3), which is usually considered as depicting Faraday induction, does not represent a cause-and-effect relation because in this equation the electric and the magnetic field is evaluated for the same moment of time. Note also that in electromagnetic waves electric and magnetic fields are in phase, that is, simultaneous in time, and hence, according to the principle of causality (which states that the cause always precedes its effect), the two fields cannot cause each other (by the principle of causality, the fields should be out of phase if they create each other).

Maxwell's equations by themselves do not provide an answer to whether or not the 'Faraday induction' or 'Maxwell induction' are real physical phenomena. In Maxwell's equations electric and magnetic fields are linked together in an intricate manner, and neither field is explicitly represented in terms of its sources. It is true, of course, that whenever there exists a time-variable electric field, there also exists a time-variable magnetic field. This follows from our equations (7) and (8) as well as from Maxwell's equations (3) and (4). But, as already mentioned, according to the causality principle, Maxwell's equations do not reveal a causal link between electric and magnetic fields. On the other hand, equations (7) and (8) show that in time-variable systems electric and magnetic fields are always created simultaneously, because these fields have a common causative source: the changing electric current $[\partial\mathbf{J}/\partial t]$ (the last term of equation (7) and the last term in the integral of equation (8)).

It is important to note that neither Faraday (who discovered the phenomenon of electromagnetic induction) nor Maxwell (who gave it a mathematical formulation) explained this phenomenon as the generation of an electric field by a magnetic field (or vice versa).

After discovering the electromagnetic induction, Faraday wrote in a letter of November 29, 1831, addressed to his friend Richard Phillips [4]:

'When an electric current is passed through one of two parallel wires it causes at first a current in the same direction through the other, but this induced current does not last a moment notwithstanding the inducing current (from the Voltaic battery) is continued. . ., but when the current is stopped then a return current occurs in the wire under induction of about the same intensity and momentary duration but in the opposite direction to that first found. Electricity in currents therefore exerts an inductive action like ordinary electricity (electrostatics, ODJ) but subject to peculiar laws: the effects are a current in the same direction when the induction is established, a reverse current when the induction ceases and a peculiar state in the interim. . ..'

Quite clearly, Faraday speaks of an inducing *current*, and not at all of an inducing *magnetic field*. (In the same letter Faraday referred to the induction by magnets as a 'very powerful proof' of the existence of Amperian currents responsible for magnetization.)

Similarly, Maxwell wrote in his *Treatise* [5]:

'It is only since the definitions of electromotive force. . . and its measurement have been made more precise, that we can enunciate completely the true law of magneto-electric induction in the following terms: the total electromotive force acting round a circuit at any instant is measured by the rate of decrease of the number of lines of magnetic force which pass through it. . .. Instead of speaking of the number of lines of magnetic force, we may speak of the magnetic induction through the circuit, or the surface integral of magnetic induction extended over any surface bounded by the circuit.'

As we see, Maxwell, too, considered the electromagnetic induction as a phenomenon in which a current (or electromotive force) is induced in a circuit, but not as a phenomenon in which a changing magnetic field causes an electric field. He clearly says that the induced electromotive force is *measured* by, not *caused* by, the changing magnetic field. Just like Faraday, he made no allusion to any causal link between magnetic and electric fields.

And there is one more fact that supports the conclusion that what we call 'electromagnetic induction' is not the creation of one of the two fields by the other. In the covariant formulation of electrodynamics, electric and magnetic fields appear as components of one single entity—the electromagnetic field tensor. Quite clearly, a component of a tensor cannot be a cause of another component of the same tensor, just like a component of a vector cannot be a cause of another component of the same vector.

We must conclude therefore that the true explanation of the phenomenon of electromagnetic induction is provided by the causal electromagnetic equations, equations (7) and (8). According to these equations, in time-variable systems electric and magnetic fields are always created simultaneously, because they have a common causative source: the changing electric current $[\partial \mathbf{J}/\partial t]$. Once created, the two fields coexist from then on without any effect upon each other. Hence electromagnetic induction as a phenomenon in which one of the fields creates the other is an illusion. The illusion of the 'mutual creation' arises from the facts that in time-dependent systems the two fields always appear prominently together, while their causative sources (the time-variable current in particular) remain in the background[1].

Thus, even though a presentation of electromagnetic theory on the basis of the causal electromagnetic equations is somewhat more complicated than the traditional presentation on the basis of Maxwell equations, such a presentation is well justified by the new possibilities that it offers and by the important new results revealed by it.

Appendix

Vector identities

In the vector identities listed below, U is a scalar point function; $\mathbf{V}$ is a vector point function; X is a scalar or vector point function of primed coordinates (source-point coordinates) and incorporates an appropriate multiplication sign (dot or cross for vectors); the operator ∇ operates upon unprimed coordinates (field-point coordinates); the operator ∇' operates upon primed coordinates (source-point coordinates).

Identities for the calculation of surface and volume integrals

$$\int \nabla' \cdot \mathbf{A}\, \mathrm{d}V' = \oint \mathbf{A} \cdot \mathrm{d}\mathbf{S}' \qquad \text{(Gauss's theorem)} \tag{A.1}$$

$$\int \nabla' U\, \mathrm{d}V' = \oint U\, \mathrm{d}\mathbf{S}' \tag{A.2}$$

$$\int \nabla' \times \mathbf{A}\, \mathrm{d}V' = -\oint \mathbf{A} \times \mathrm{d}\mathbf{S}'. \tag{A.3}$$

Identities for operations with retarded quantities

$$\frac{\mathbf{r}[X]}{r^3} + \frac{\mathbf{r}}{r^2 c}\left[\frac{\partial X}{\partial t}\right] = -\nabla \frac{[X]}{r} \tag{A.4}$$

$$\nabla \frac{[X]}{r} = \frac{[\nabla' X]}{r} - \nabla' \frac{[X]}{r} \tag{A.5}$$

$$U = -\frac{1}{4\pi} \int \frac{\left[\nabla' \cdot \nabla' U - \frac{1}{c^2}\frac{\partial^2 U}{\partial t^2}\right]}{r}\, \mathrm{d}V' \tag{A.6}$$

$$\mathbf{A} = -\frac{1}{4\pi} \frac{\left[\nabla'(\nabla' \cdot \mathbf{A}) - \nabla' \times (\nabla' \times \mathbf{A}) - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2}\right]}{r}\, \mathrm{d}V'. \tag{A.7}$$

[1] The author has been unable to determine by whom, where and why it was first suggested that changing electric and magnetic fields create each other. One thing appears certain however—the idea did not originate with either Faraday or Maxwell.

References

[1] Jefimenko O D 1989 *Electricity and Magnetism* 2nd edn (Star City: Electret Scientific) pp 514–6 (the same pages in the first edition of the book (Appleton-Century-Crofts, New York, 1966))
[2] Rosser W G V 1997 *Interpretation of Classical Electromagnetism* (Dordrecht: Kluwer) pp 82–4
[3] Jackson J D 1999 *Classical Electrodynamics* 3rd edn (New York: Wiley) pp 243–7
[4] Thompson S P 1898 *Michael Faraday, His Life and Work* (New York: Macmillan) p 115, 116
[5] Maxwell J C 1891 *A Treatise on Electricity and Magnetism* 3rd edn, vol 2 (New York: Dover) pp 176–7 (reprint)