the general science journal

Radio Waves – Part I

Ionel DINU, M.Sc.,
Member of Natural Philosophy Alliance
http://www.worldsci.org/people/Ionel_Dinu
e-mail: *dynuionel@yahoo.com*

(Dated: March 14, 2012. 1st Revision: March 25, 2012. 2nd Revision: May 13, 2013)

NOTE: This paper was intended to be a preamble to an article presenting author's own view regarding the nature of radio waves and the mechanism by which radio waves are produced in a radio antenna. While this original aim is still in author's view and will be realized in Part II, this first part (Part I) had to be developed into an article aimed at providing more arguments as to why Maxwell's theory of electromagnetic waves is not as reliable as textbooks lead us to believe, even in the realms of classical electromagnetism, i.e. outside the realm of quantum mechanics or relativity. The author chose to keep this paper at its original length and use it as a base for more detailed argumentation against Maxwell's theory in a separate paper titled "Trouble with Maxwell's Electromagnetic Theory: Can Fields Induce Other Fields in Vacuum?". The readers are kindly invited to read it.

Abstract

This article is about radio waves. In it, the author tries to show that the present understanding of the nature of radio waves is erroneous because it is based on a false method of theoretical investigation that leads to ideas unsupported by experimental facts.

The *production* of radio waves is unquestionably due to large currents surging back and forth along a metallic conductor, called antenna. These electric currents give rise to magnetic fields in the space around the antenna. In the overwhelming majority of cases the currents oscillating in the antenna are produced by other electric currents. The author has no knowledge of any instance in which the currents in the antenna, or the radio waves themselves, are produced by varying voltages.

These facts describe the circumstances in which radio waves are produced, but do not say anything specifically about *how* these waves are generated from these original causes and *what* is that which is waving when radio waves are said to travel through space.

Present day physics tries to answer the above questions by making some wild guesses. To the '*what* is waving' question the answer it gives is 'changing electric and magnetic fields that propagate on their

own in space away from the antenna'. To make the whole view consistent, the question '*how* are these waves generated' is answered 'these fields induce (create) each other' and travel on their own as waves.

Careful analysis shows that these answers contradict other facts in electromagnetism and are untenable. The most important idea of this study - and which actually sparked author's interest in this subject - is the realization that *neither electric fields can induce magnetic fields, nor vice-versa.*

If the fields cannot induce each other, then to the student who has studied electromagnetism, another possible answer might appear to be that the radio waves are just the waving lines of electric and magnetic field still bound to the charges. This is, however, contradicted by experience, which shows that *the radio waves can be detected at distances very far from the antenna, thousands of kilometers away, where the intensity of the fields produced by the charges in the antenna, whether waving or not, is undetectable*.

The theory advanced in this article is that what travel away from the antenna are waves produced in the aether by the charges surging in the antenna. It follows then that, if we wish to call the radio waves by the cause producing them, we can continue to call them electromagnetic. But if we wish to call them by what they really are, we should call them aether waves. By comparison, in acoustics the wave is called by what it really is – sound – and not by the cause producing it (vibrations).

Introduction

This study is addressed to that small percent of students and researchers who suspect that there is something wrong with the way in which we understand nowadays *how radio waves are generated and how they propagate in space*.

I know that there is always a feeling of distrust when university professors obtain the equation of a wave from the four Maxwell's equations. I felt that myself as a student and I have seen it again in the open courses made available on the Internet by prestigious universities of the world. Students ask pertinent questions but the professor fails to address the issue. [See http://www.youtube.com/watch?v=JJZkjMRcTD4&feature=endscreen, min. 0:35:0].

When still a student, I promised myself that, someday, I will get back to the subject of radio waves and analyze it piece by piece, statement by statement, equation by equation, and I will not declare myself in agreement with the theory if I discover unfounded assumptions, guesswork, or things contrary to experimental observations. I can say that I have found a few of these.

What I consider most controversial in all the present conception regarding radio waves is the belief that the electric and magnetic fields produced in and around the antenna by the charges moving in it induce each other and create new fields at other points of space, *even in regions of space where there are no electric charges*, and that these fields become self-sustaining 'electromagnetic waves'. The majority of physicists and engineers agree with this description. No wonder, since they were good students and learnt what they could from their teachers and the textbooks available to them, all expounding the same doctrine.

In this work I will argue that the *idea of electric and magnetic fields inducing each other without the mediation of electrical charges* is false because it is not based on experimental evidence. Pure electric fields, varying or not, do not produce pure magnetic fields in regions of space where electrical charges do not exist. Neither pure magnetic fields can produce, in regions of space where electrical charges do not exist, pure electric fields. It is only through the mediation of electric charges and currents that these phenomena can happen. I will take excerpts from the works of authors who support the present day theory and I will point out where their argument fails. Among other works, textbooks will be given priority.

To conclude this short introduction, I would like to repeat the main problem of this study, namely that we do know *what produces* the radio waves – rapidly changing electric currents in a conductor – but what we do not know for sure is *how exactly radio waves are generated* from these electric currents, and *what radio waves really are* when traveling through space. These, I contend, are problems still open for argument and will be discussed here in detail.

My alternative explanation is that these waves are waves in the aether produced by the charges (electrons) surging in the antenna. But even if you don't agree with my view, I hope that what I have to say about present day theory of radio waves will make you eager to study the subject yourself with more attention than when you did when you were a student and develop a personal opinion on what is believed to be one of the most important theories of physics.

Mainstream science considers these matters settled beyond question and I do not expect great interest in this work from professional scientists. My hope is only that the young student, the young researcher at the beginning of his career or scientists who want to remain true to their profession will have enough time to ponder on these questions. My intention is not to demolish something that is valuable, but to find the true answers to the questions posed above and avoid the perpetuation of false ideas and flawed reasoning in physical science by turning a blind eye to what I believe is inaccurate. I see it as a duty towards science itself as well as towards its present and future scientists.

SECTION I. What standard textbooks say

I know that it may be some time since you have graduated high school, but I want to remind you how little standard textbooks for secondary grades have to say about *how* radio waves are generated. So I will start with some excerpts that deal with this topic.

<u>A. First category : GCSE (and IGCSE) textbooks</u>

These textbooks are written for secondary students (Grades 9 and 10). The two examples chosen below give, in only one sentence, some information about *what* produces the radio waves. Nothing is said about *how* these waves are generated.

1. Tom Duncan, Heather Kennett, *GCSE Physics*, 4th Ed., Hodder Murray, 2001, p. 52:

> *Radio waves*
>
> Radio waves have the longest wavelengths in the electromagnetic spectrum. They are radiated from aerials and used to 'carry' sound, pictures and other information over long distances.

"They [radio waves] are radiated from aerials […]."

2. Stephen Pole, *Complete Physics for IGCSE*, Oxford University Press, 2007, p. 162:

> **Radio waves**
>
> Stars are natural emitters of radio waves. However, radio waves can be produced artificially by making a current oscillate in a transmitting aerial (antenna). In a simple radio system, a microphone controls the current to the aerial so that the radio waves 'pulsate'. In the radio receiver, the incoming pulsations control a loudspeaker so that it produces a copy of the original sound. Radio waves are also used to transmit TV pictures.

"[…] radio waves can be produced artificially by making a current oscillate in a transmitting aerial (antenna)."

B. Second category : Advanced Level (A-Level) Physics and IB Physics textbooks

These textbooks are written for secondary students (Grades 11 and 12) taking a Physics course after finishing GCSE. They discuss more technicalities but are still silent about *how* are the *waves* generated by the current (or the charges) oscillating in the antenna.

1. M. Nelkon and P. Parker, *Advanced Level Physics*, 3rd Ed., Heinemann Educational Books, 1970, p. 986:

Radiation of Electromagnetic Waves into Space

Consider an oscillator with a connected transmission line, and suppose the transmission line is bent as shown in Fig. 39.23 (i).

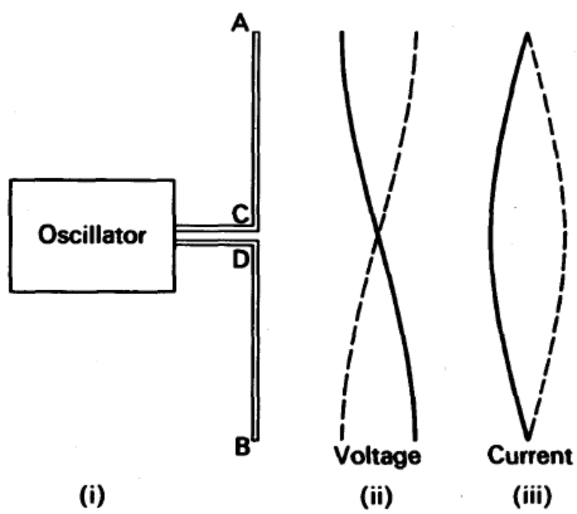


FIG. 39.23. Aerial. Half-wave dipole.

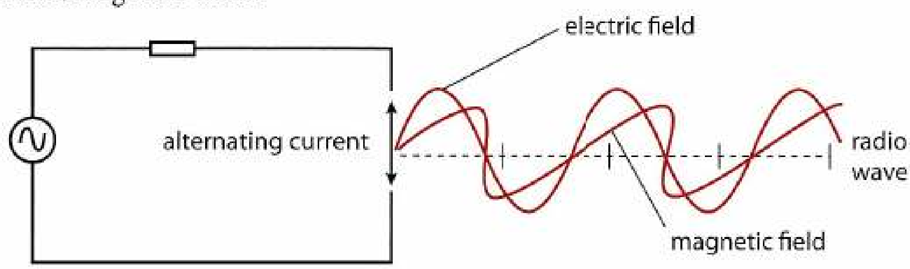
The charges moving along AB are forced up to A during one half cycle of oscillation and then down to B during the next half cycle. The charge, therefore, oscillates between A and B. *This accelerating charge radiates energy in the form of electromagnetic waves.* In contrast, charges moving in a line with a steady speed create a static magnetic field, and no electromagnetic wave is radiated.

"*This accelerating charge radiates energy in the form of electromagnetic waves*."

2. Chris Hamper, Keith Ord, *Standard Level Physics Developed Specifically for the IB Diploma*, Pearson Education Limited, 2007, p. 387:

Creating an electromagnetic wave

An electromagnetic wave can be created by passing an alternating current through a wire as shown in Figure 15.2. Waves created in this way are called *radio waves*. James Maxwell found that it was not the moving charge that caused the magnetic field but the changing electric field that caused the charge to move. This explains how electromagnetic waves can travel through a vacuum: the changing fields induce each other. Maxwell also calculated that the speed of the wave in a vacuum was approximately $3 \times 10^8\ \mathrm{m\,s^{-1}}$. This value was about the same as the measured value for the speed of light, so close in fact, that Maxwell concluded that light was an electromagnetic wave.



Here we find, for the first time, two statements that seem to me *inconsistent* with one another.

The first is:

"An electromagnetic wave can be created by passing an alternating current through a wire […]. Waves created in this way are called *radio waves*."

The second is:

"James Maxwell found that it was not the moving charge that caused the magnetic field, but the changing electric field that caused the charge to move."

The inconsistency rests in the fact that the electric currents are not seen any more as the primary cause producing the radio waves. The primary cause for the production of radio waves has been shifted to the *changing electric field* that produces the oscillating current.

Thus the textbook tells the student something new: that *a changing electric field generates a changing magnetic field*. But is it true? Can a field produce another field? The textbook says that this was "found" by Maxwell. But did Maxwell prove what the textbook says he "found"? The answer turns out to be no. Not only that Maxwell did not prove it by any experiment but nobody proved it experimentally in the 150 years that have passed since then. What Maxwell did was a mathematical manipulation, which we shall discuss later.

Why is this important? It is important because Maxwell's "finding" is then used to explain why 'electromagnetic waves' can travel through vacuum, where there are no electric charges. The explanation is: "the changing fields induce each other". This means that, after being created by the original charges that oscillated in the antenna, the electric and magnetic fields continue to create (induce) each other even in regions of space far from the antenna, where there are no electrical charges whatsoever. In what follows, I will argue that this picture is inaccurate.

Even before we discuss what mathematical manipulation Maxwell did and why he did it, there is an obvious thing that shows that electromagnetic waves are not produced by changing electric fields. Look at the antennas that we use: they are all conductors. If the primary source of radio waves would be the varying electric fields (which would then induce magnetic fields, which would then in their turn induce another new electric field further away, and so on) we would use for our antennas huge *capacitors* and not conductors. Our antennas would look like two huge metal plates separated by a dielectric (air) and connected to a source of oscillating high voltage. But this is not the case in practice: even since the times of Hertz and Marconi, radio waves have been produced by *discharges* (sparks) between the knobs of the induction coil. [See, J. J. Fahie, *A History Of Wireless Telegraphy 1838-1899*, William Blackwood and Sons, 1899]. All past experimentation comes to demonstrate that if an *electric current* is not made to move violently in a conductor, no radio waves can be released into space.

SECTION II. Changing fields cannot induce each other. Where is Maxwell not correct?

Since the previous textbook did not say *how James Clerk Maxwell found that a changing electric field can produce a magnetic field*, we will take another, more advanced, textbook, designed for undergraduate students: David J. Griffiths, *Introduction to Electrodynamics*, Prentice Hall, 1999. It is a well known standard textbook and many physics students have used it in their studies. This section makes heavy reference to it.

We discover from this textbook that Maxwell introduced the idea that a changing electric field can produce a changing magnetic field by modifying the experimentally found Ampere's law. At pages 321 and 326, we read:

7.3 Maxwell's Equations

7.3.1 Electrodynamics Before Maxwell

So far, we have encountered the following laws, specifying the divergence and curl of electric and magnetic fields:

$$\text{(i)} \quad \nabla \cdot \mathbf{E} = \frac{1}{\epsilon_0}\rho \quad \text{(Gauss's law)},$$

$$\text{(ii)} \quad \nabla \cdot \mathbf{B} = 0 \quad \text{(no name)},$$

$$\text{(iii)} \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad \text{(Faraday's law)},$$

$$\text{(iv)} \quad \nabla \times \mathbf{B} = \mu_0 \mathbf{J} \quad \text{(Ampère's law)}.$$

These equations represent the state of electromagnetic theory over a century ago, when Maxwell began his work.

This set of equations has been changed by Maxwell into:

7.3.3 Maxwell's Equations

In the last section we put the finishing touches on Maxwell's equations:

$$\text{(i)} \quad \nabla \cdot \mathbf{E} = \frac{1}{\epsilon_0}\rho \quad \text{(Gauss's law)},$$

$$\text{(ii)} \quad \nabla \cdot \mathbf{B} = 0 \quad \text{(no name)},$$

$$\text{(iii)} \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad \text{(Faraday's law)},$$

$$\text{(iv)} \quad \nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \quad \text{(Ampère's law with Maxwell's correction)}.$$

Observe that Ampere's original law, which was a mathematical description of *experimental findings* relating the magnetic field $\mathbf{B}$ to the current density $\mathbf{J}$ producing it, has been changed by Maxwell by adding a supplementary term to the right-hand side of the equation.

Maxwell's addition, $\mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}$, has received the name "Maxwell's displacement current".

Ampere's original law allows the calculation of the magnetic field $\mathbf{B}$ produced at a point in space by currents $\mathbf{J}$ flowing along other curves in space. It has its experimental roots in Oersted's great discovery that an electric current produces a magnetic field in the space around it. If another term is added to this equation, it follows that the magnetic field can be produced also in the manner described by this new term. Adding $\mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}$ to Ampere's original equation is equivalent to saying that a changing electric field $\mathbf{E}$ can produce a magnetic field $\mathbf{B}$.

Why is Maxwell not correct?

Maxwell is not correct for the following reasons:

(i) Such an effect (that a changing electric field $\mathbf{E}$ can produce a magnetic field $\mathbf{B}$) has not been observed experimentally. Therefore, adding the term $\mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}$ to Ampere's original equation is pseudo science.

To see how absurd the matters can get, observe that you obtain a magnetic field even if there are no electric currents at all. For $\mathbf{J} = 0$, Ampere's law (modified by Maxwell) becomes:

$$\nabla \times \mathbf{B} = \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}$$

Since the electric charges, static or in motion, do not appear in the equation, this equation says that a pure electric field $\mathbf{E}$ varying in time can create a pure magnetic field $\mathbf{B}$.

Why is this pseudo science? Because experiments show that fields are created by charges. The electric field is created by a static charge and a magnetic field by a moving charge. Every time there is a field, this field can be traced to an electrical charge, at rest or in motion.

According to Coulomb's law $\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0}$, the electric field $\mathbf{E}$ can change only if the charge density ρ changes, but this is not apparent any more in Maxwell's modification of Ampere's law.

Faraday's law seems to indicate another way in which the electric field $\mathbf{E}$ can be changed, but this is only apparent. As will be discussed later, Maxwell has modified Faraday's law by making the same conceptual mistake as he did when he modified Ampere's original law. For what Faraday observed was that a changing magnetic field induces *an electric current* and not an electric field. So the mathematical rendering of Faraday's law is also questionable and will be discussed later.

(ii) Maxwell's "displacement current" is not a current. If there are supplementary currents to be added in Ampere's law (and we will see later that one supplementary current must indeed be added), *these currents must be added as currents*, not as something else, because this is what observations show: moving electric charges produce a magnetic field around them. A current (more accurately, *current density*, because Ampere's law is written in terms of $\mathbf{J}$ - the current density) is defined as

$$\mathbf{J} = \rho \cdot \mathbf{v}$$

where ρ is the charge density and $\mathbf{v}$ is the velocity of the charges.

How should Maxwell have corrected Ampere's law?

Maxwell introduced his "displacement current" in Ampere's law in an attempt to make it more general. Look at the explanations below, which will start with a repetition of the excerpt shown above:

7.3 Maxwell's Equations

7.3.1 Electrodynamics Before Maxwell

So far, we have encountered the following laws, specifying the divergence and curl of electric and magnetic fields:

$$\text{(i)} \quad \nabla \cdot \mathbf{E} = \frac{1}{\epsilon_0}\rho \quad \text{(Gauss's law)},$$

$$\text{(ii)} \quad \nabla \cdot \mathbf{B} = 0 \quad \text{(no name)},$$

$$\text{(iii)} \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad \text{(Faraday's law)},$$

$$\text{(iv)} \quad \nabla \times \mathbf{B} = \mu_0 \mathbf{J} \quad \text{(Ampère's law)}.$$

These equations represent the state of electromagnetic theory over a century ago, when Maxwell began his work. They were not written in so compact a form in those days, but their physical content was familiar. Now, it happens there is a fatal inconsistency in these formulas. It has to do with the old rule that divergence of curl is always zero. If you apply the divergence to number (iii), everything works out:

$$\nabla \cdot (\nabla \times \mathbf{E}) = \nabla \cdot \left(-\frac{\partial \mathbf{B}}{\partial t}\right) = -\frac{\partial}{\partial t}(\nabla \cdot \mathbf{B}).$$

The left side is zero because divergence of curl is zero; the right side is zero by virtue of equation (ii). But when you do the same thing to number (iv), you get into trouble:

$$\nabla \cdot (\nabla \times \mathbf{B}) = \mu_0(\nabla \cdot \mathbf{J}); \tag{7.35}$$

the left side must be zero, but the right side, in general, is *not*. For *steady* currents, the divergence of **J** is zero, but evidently when we go beyond magnetostatics Ampère's law cannot be right.

[…]

Of course, we had no right to *expect* Ampère's law to hold outside of magnetostatics; after all, we derived it from the Biot-Savart law. However, in Maxwell's time there was no *experimental* reason to doubt that Ampère's law was of wider validity. The flaw was a purely theoretical one, and Maxwell fixed it by purely theoretical arguments.

7.3.2 How Maxwell Fixed Ampère's Law

The problem is on the right side of Eq. 7.35, which *should be* zero, but *isn't*. Applying the continuity equation (5.29) and Gauss's law, the offending term can be rewritten:

$$\nabla \cdot \mathbf{J} = -\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial t}(\epsilon_0 \nabla \cdot \mathbf{E}) = -\nabla \cdot \left(\epsilon_0 \frac{\partial \mathbf{E}}{\partial t}\right).$$

It might occur to you that if we were to combine $\epsilon_0(\partial \mathbf{E}/\partial t)$ with **J**, in Ampère's law, it would be just right to kill off the extra divergence:

$$\boxed{\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}.} \tag{7.36}$$

[…]

Such a modification changes nothing, as far as magneto*statics* is concerned: when **E** is constant, we still have $\nabla \times \mathbf{B} = \mu_0 \mathbf{J}$. In fact, Maxwell's term is hard to detect in ordinary electromagnetic experiments, where it must compete for recognition with **J**; that's why Faraday and the others never discovered it in the laboratory. However, it plays a crucial role in the propagation of electromagnetic waves, as we'll see in Chapter 9.

In my opinion, Maxwell should have "fixed" Ampere's law to comply with the equation of continuity by adding another current density $\mathbf{J}'$ such that:

$$\nabla \times \mathbf{B} = \mu_0 \cdot (\mathbf{J} + \mathbf{J}')$$

Then the vector calculus identity used by Maxwell, which says that, for *any* vector $\mathbf{B}$, the expression $\nabla \cdot (\nabla \times \mathbf{B})$ must be zero, gives:

$$\nabla \cdot (\nabla \times \mathbf{B}) \equiv 0 \Rightarrow \mu_0 \cdot \nabla \cdot (\mathbf{J} + \mathbf{J}') = 0 \Rightarrow \nabla \cdot (\mathbf{J} + \mathbf{J}') = 0 \Rightarrow \nabla \cdot \mathbf{J} + \nabla \cdot \mathbf{J}' = 0$$

The equation of continuity $\nabla \cdot \mathbf{J} + \dfrac{\partial \rho}{\partial t} = 0$ and the above result then show that the extra current $\mathbf{J}'$ that must be added to Ampere's law must be such that $\nabla \cdot \mathbf{J}' = \dfrac{\partial \rho}{\partial t}$.

How is this modification different from Maxwell's?

The above modification is different from Maxwell's in that Ampere's law still contains currents and only currents (current densities, actually), as observed experimentally. No other physical quantities are added artificially – only currents.

In Maxwell's modification, the supplementary current $\mathbf{J}'$ is not left as above, but it is expressed further through purely mathematical manipulations, starting from Coulomb's law. This was shown in the previous page in the excerpt from David J. Griffiths, *Introduction to Electrodynamics*, Prentice Hall, 1999, p.323, but I rewrite it here:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0} \Rightarrow \rho = \varepsilon_0 \cdot \nabla \cdot \mathbf{E} \Rightarrow \frac{\partial \rho}{\partial t} = \varepsilon_0 \cdot \frac{\partial}{\partial t}(\nabla \cdot \mathbf{E}) = \varepsilon_0 \cdot \nabla \cdot \left(\frac{\partial \mathbf{E}}{\partial t}\right) = \nabla \cdot \left(\varepsilon_0 \cdot \frac{\partial \mathbf{E}}{\partial t}\right)$$

Comparing $\nabla \cdot \mathbf{J}' = \frac{\partial \rho}{\partial t}$ with $\frac{\partial \rho}{\partial t} = \nabla \cdot \left(\varepsilon_0 \cdot \frac{\partial \mathbf{E}}{\partial t}\right)$ found above, Maxwell observed that $\mathbf{J}' = \varepsilon_0 \cdot \frac{\partial \mathbf{E}}{\partial t}$. He then *introduced it directly in Ampere's law*, obtaining:

$$\nabla \times \mathbf{B} = \mu_0 \cdot (\mathbf{J} + \mathbf{J}') \Rightarrow \nabla \times \mathbf{B} = \mu_0 \cdot \mathbf{J} + \mu_0 \varepsilon_0 \cdot \frac{\partial \mathbf{E}}{\partial t}$$

The difference between Maxwell's modification of Ampere's law, and the one which I consider correct, is summarized in the table below:

Maxwell's modification of Ampere's law	Amperes' law modified correctly
$\nabla \times \mathbf{B} = \mu_0 \cdot \mathbf{J} + \mu_0 \varepsilon_0 \cdot \frac{\partial \mathbf{E}}{\partial t}$	$\nabla \times \mathbf{B} = \mu_0 \cdot (\mathbf{J} + \mathbf{J}')$ and $\nabla \cdot \mathbf{J}' = \frac{\partial \rho}{\partial t}$

The difference between the two is enormous because, in physics in general, the equations connecting different physical quantities are interpreted phenomenologically, that is, they must correspond to effects observed in nature, experimentally.

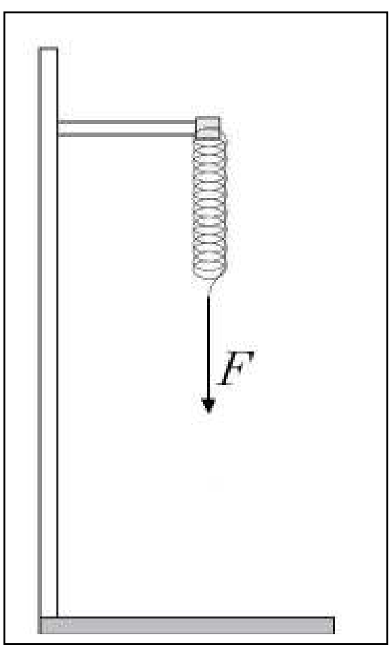
As stated above, Maxwell's version of Ampere's law implies that a magnetic field can be produced by a changing electric field, *even for cases when there are no charges flowing* ($\mathbf{J} = 0$), and this is not observed experimentally.

If correctly modified, Ampere's law states that there must always be *electric currents* to produce a magnetic field. Even if $\mathbf{J} = 0$, it is the supplementary current $\mathbf{J}' \neq 0$ that produces a magnetic field. This supplementary current $\mathbf{J}'$ is produced through the change of charge density ρ, such that $\nabla \cdot \mathbf{J}' = \frac{\partial \rho}{\partial t}$. The equations always *link the fields with the charges producing them* and never omit them as important intermediaries between the fields. The correctly modified Ampere's law does not predict absurd, never observed, phenomena such as that according to which a magnetic field can be produced by a changing electric field. Even if $\mathbf{J}'$ equals $\varepsilon_0 \cdot \frac{\partial \mathbf{E}}{\partial t}$ (as Maxwell showed), this is, at most, *an equality of magnitude* and has to be kept as a separate equation.

As a conclusion, Maxwell is not correct because, in science, the equations we write should not be correct only dimensionally and quantitatively, but they must also correspond to observed phenomena. *Substituting* $\varepsilon_0 \cdot \frac{\partial \mathbf{E}}{\partial t}$ *for* $\mathbf{J}'$ *in Ampere's law, although correct mathematically and dimensionally, is not correct phenomenologically, because the interpretation of the law thus modified leads to absurdities not observed in real world*.

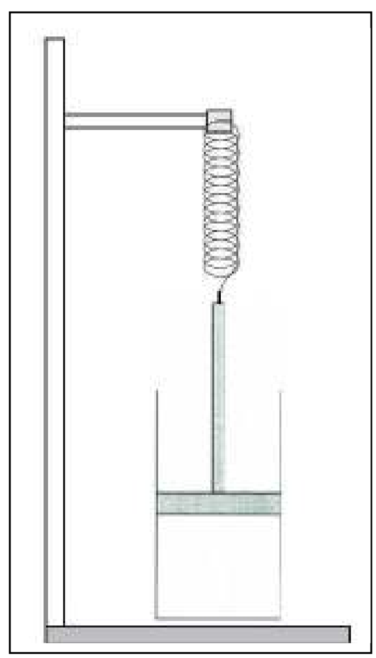
There are many situations in physics when we replace physical quantities in different equations, obtain other equations that are correct dimensionally and quantitatively, and use them to calculate unknown physical quantities. But we cannot expect these manipulated equations to make sense phenomenologically, to see in them a true, direct cause-effect relationship between the physical quantities that appear in it. As is the case with Maxwell's modification of Ampere's law, the equations manipulated by mathematical operations, even if correct, bring together mathematical expressions corresponding to physical phenomena that have no direct cause-effect relationship and turn out to be absurd statements if interpreted phenomenologically.

To give you an example, consider a spring hung vertically. We know experimentally that the spring stretches because there is a force F acting on it and we express the extension of the spring x in terms of the force F as $x = \frac{1}{k} \cdot F$



But we can apply a force to the spring in another way. For example, consider a piston attached to the spring and the cylinder fixed to the ground.

The gas in the cylinder contracts when cooled and pulls the piston downwards.



So besides pulling forces F that may act on the spring, we have to consider another force F' that produces the same effect. The original formula giving the extension of the spring becomes

$x = \frac{1}{k} \cdot (F + F')$, according to the law of addition of forces, verified experimentally.

Then we can measure experimentally how F' changes with the temperature. Suppose that experiments yield:

$$F' = -R \cdot \Delta T$$

where R is a constant and ΔT is the change in the temperature of the gas in the piston, showing that a negative temperature change produces a positive force F' that stretches the spring.

Now, equations $x = \frac{1}{k} \cdot F$, $x = \frac{1}{k} \cdot (F + F')$, and $F' = -R \cdot \Delta T$ have been obtained experimentally and can be interpreted phenomenologically.

But if we replace F' in the equation for extension x, we obtain

$$x = \frac{1}{k} \cdot (F - R \cdot \Delta T)$$

which, although correct mathematically, leads to absurdities when interpreted phenomenologically, for it says that *a spring can be stretched by a decrease in the temperature*.

Maxwell's modification of Ampere's law has been obtained by a similar *false method of theoretical investigation* and this is why it cannot be considered correct.

The problems are more serious than thought: Faraday's law

As mentioned earlier, I will discuss in this second section of the study another serious logical inaccuracy that I observed in the accepted laws of electricity and magnetism. It refers to the interpretation of Faraday's law of electromagnetic induction.

Let us refer this time to the textbook: John David Jackson, *Classical Electrodynamics*, John Wiley and Sons, 1962, a well known textbook designed, at its time, for beginning graduate students . At page 170, we read:

170 *Classical Electrodynamics*

6.1 Faraday's Law of Induction

The first quantitative observations relating time-dependent electric and

> magnetic fields were made by Faraday (1831) in experiments on the behavior of currents in circuits placed in time-varying magnetic fields. It was observed by Faraday that a transient current is induced in a circuit if (*a*) the steady current flowing in an adjacent circuit is turned on or off, (*b*) the adjacent circuit with a steady current flowing is moved relative to the first circuit, (*c*) a permanent magnet is thrust into or out of the circuit. No current flows unless either the adjacent current changes or there is relative motion. Faraday interpreted the transient current flow as being due to a changing magnetic flux linked by the circuit. The changing flux induces an electric field around the circuit, the line integral of which is called the *electromotive force*, $\mathcal{E}$. The electromotive force causes a current flow, according to Ohm's law.

It is clear, I think, to everyone, that the sentence:

"Faraday interpreted the transient current flow as being due to a changing magnetic flux linked by the circuit."

means that the observed cause-effect – which can be also observed by any student in the laboratory – is that *a changing magnetic flux causes an electric current.*

Then, what is the reason for which we have to invoke the existence of an *electric field*? Quote again from the excerpt above:

"The changing flux induces an *electric field* around a circuit […]. The *electromotive force* causes a current to flow, according to Ohm's law."

So the production of an electric field is invoked to account for the movement of charges. This comes to show that Faraday's law is not written according to observations, but has been modified and now contains an *unfounded assumption*: that electric charges can be made to move only if there is an electric field.

What was done here was pseudo science because instead of faithfully encoding in mathematical formulas the effects as they are observed in reality, we introduced guesses as to what is happening in that observed process. I think it can be said that Faraday's law has been spoiled and is not an accurate description of the observed phenomena.

Exactly the same ideas can be found in the more recent work of David J. Griffiths, *Introduction to Electrodynamics*, Prentice Hall, 1999. At pages 301-302, we find the flagrant: the author admits that Faraday observed an *electric current* induced in the circuit and that, before it was codified mathematically, the law was *interpreted* in terms of *electric field*.

7.2 Electromagnetic Induction

7.2.1 Faraday's Law

In 1831 Michael Faraday reported on a series of experiments, including three that (with some violence to history) can be characterized as follows:

Experiment 1. He pulled a loop of wire to the right through a magnetic field (Fig. 7.20a). A current flowed in the loop.

Experiment 2. He moved the *magnet* to the *left*, holding the loop still (Fig. 7.20b). Again, a current flowed in the loop.

Experiment 3. With both the loop and the magnet at rest (Fig. 7.20c), he changed the *strength* of the field (he used an electromagnet, and varied the current in the coil). Once again, current flowed in the loop.

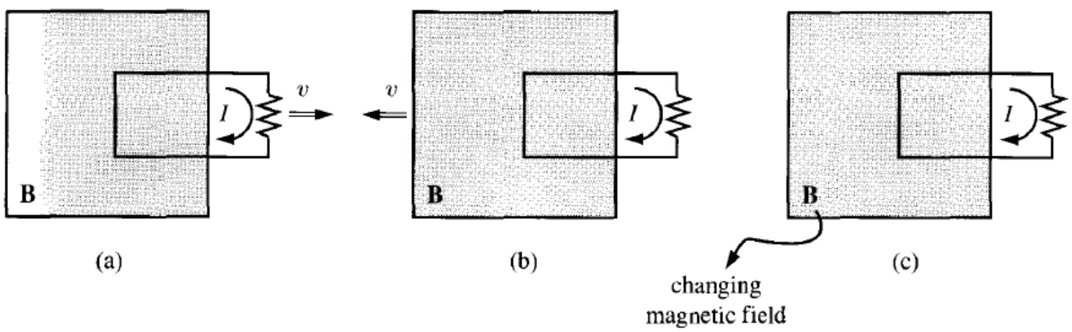


Figure 7.20

The first experiment, of course, is an example of motional emf, conveniently expressed by the flux rule:

$$\mathcal{E} = -\frac{d\Phi}{dt}.$$

I don't think it will surprise you to learn that exactly the same emf arises in Experiment 2—all that really matters is the *relative* motion of the magnet and the loop. Indeed, in the light of special relativity is *has* to be so. But Faraday knew nothing of relativity, and in classical electrodynamics this simple reciprocity is a coincidence, with remarkable implications. For if the *loop* moves, it's a *magnetic* force that sets up the emf, but if the loop is *stationary*, the force *cannot* be magnetic—stationary charges experience no magnetic forces. In that case, what *is* responsible? What sort of field exerts a force on charges at rest? Well, *electric* fields do, of course, but in this case there doesn't seem to be any electric field in sight.

Faraday had an ingenious inspiration:

A changing magnetic field induces an electric field.

It is this "induced" electric field that accounts for the emf in Experiment 2.[6]

[6]You might argue that the magnetic field in Experiment 2 is not really *changing*—just *moving*. What I mean is that if you sit at a *fixed location*, the field *does* change, as the magnet passes by.

Whatever explanations the author may offer, there is one thing that remains: that Faraday's law was interpreted in terms of *known facts* before it was codified mathematically. As such, it contains a *new effect linked with old knowledge* acquired from other experiments and cannot be considered to be a true representation of what occurs in Faraday's experiments.

In my opinion, rather that trying to explain Faraday's observations by invoking the magnetic force for the case moving loop / stationary magnetic field and the creation of an electric field for the case stationary loop / moving magnetic field, the law should have been translated in mathematical language in a form that expressed the fact that *any relative motion of the magnetic field and the charges create a force on the charges*. A possible way to achieve this would be to generalize the formula for the magnetic force

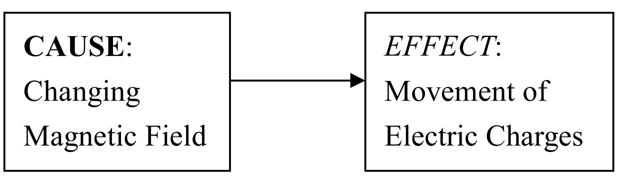
$$\mathbf{f} = q \cdot \mathbf{v} \times \mathbf{B}$$

through the addition of new terms that account for these phenomena.

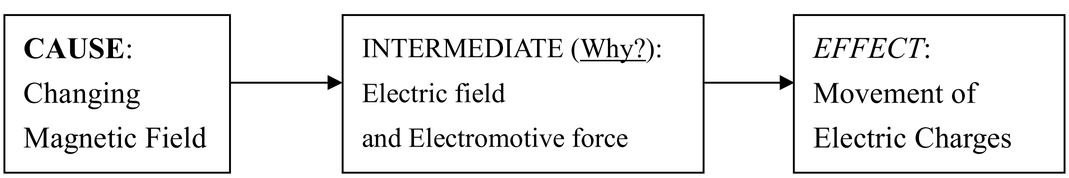
In conclusion, Faraday's law, strictly speaking, states that *there is a force acting on an electric charge whenever there is a relative motion between the charge and a magnetic field*. Even if the circuit is broken, the movement of charges still takes place. The charges moving in a broken circuit under the action of the changing magnetic field causes them to separate and gather at the ends of the gap in the circuit: the electrons gather at one end making it negative and leave the other end charged positively.

Below I tried to show diagrammatically the difference between Faraday's original discovery and its mathematical rendering:

Faraday's discovery:



Faraday's discovery was reinterpreted by the *artificial insertion* of an electric field and e.m.f. in the cause-effect chain of the observed phenomenon:



The equation below (David J. Griffiths, *Introduction to Electrodynamics*, Prentice Hall, 1999, p. 302) corresponds to the interpreted version of Faraday's law:

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

Observe that no induced current appears in this equation: Faraday's law has been transformed into an equation between two fields that does not mention any induced current and this not related to what was observed experimentally.

Maxwell's equations and his wave equations – a different interpretation

It is often stated that Maxwell's equations yield the equations of electromagnetic waves in vacuum, which means that they are valid for regions of space where there are no charges or currents.

In truth, the said equations can be obtained for regions where *there are* charges and currents, and no reason can be given why they should be valid for vacuum as well.

Since textbooks never mention the fact that Maxwell's famous equations for electromagnetic waves can be obtained even without the conditions $\rho = 0$ and $\mathbf{J} = 0$, few students suspect that they are being *lied by omission*.

Look at their derivation as given by one of the standard textbooks (David J. Griffiths, *Introduction to Electrodynamics*, Prentice Hall, 1999, p. 375):

9.2 Electromagnetic Waves in Vacuum

9.2.1 The Wave Equation for E and B

In regions of space where there is no charge or current, Maxwell's equations read

$$\left.\begin{array}{llll} \text{(i)} & \nabla \cdot \mathbf{E} = 0, & \text{(iii)} & \nabla \times \mathbf{E} = -\dfrac{\partial \mathbf{B}}{\partial t}, \\ \text{(ii)} & \nabla \cdot \mathbf{B} = 0, & \text{(iv)} & \nabla \times \mathbf{B} = \mu_0 \epsilon_0 \dfrac{\partial \mathbf{E}}{\partial t}. \end{array}\right\} \tag{9.40}$$

They constitute a set of coupled, first-order, partial differential equations for **E** and **B**. They can be *de*coupled by applying the curl to (iii) and (iv):

$$\begin{aligned} \nabla \times (\nabla \times \mathbf{E}) &= \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = \nabla \times \left(-\frac{\partial \mathbf{B}}{\partial t}\right) \\ &= -\frac{\partial}{\partial t}(\nabla \times \mathbf{B}) = -\mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}, \\ \nabla \times (\nabla \times \mathbf{B}) &= \nabla(\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B} = \nabla \times \left(\mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}\right) \\ &= \mu_0 \epsilon_0 \frac{\partial}{\partial t}(\nabla \times \mathbf{E}) = -\mu_0 \epsilon_0 \frac{\partial^2 \mathbf{B}}{\partial t^2}. \end{aligned}$$

Or, since $\nabla \cdot \mathbf{E} = 0$ and $\nabla \cdot \mathbf{B} = 0$,

$$\boxed{\nabla^2 \mathbf{E} = \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}, \quad \nabla^2 \mathbf{B} = \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{B}}{\partial t^2}.} \tag{9.41}$$

We now have *separate* equations for **E** and **B**, but they are of *second* order; that's the price you pay for decoupling them.

In vacuum, then, each Cartesian component of **E** and **B** satisfies the **three-dimensional wave equation**,

$$\nabla^2 f = \frac{1}{v^2} \frac{\partial^2 f}{\partial t^2}.$$

You can interpret the two differential equations for **E** and **B** in any way you wish. The ambiguity is so great that you can consider them to be the vibrations of a line of electric or magnetic field fixed at its ends, or of a line with one free end, or even without ends (closed loops); or you can consider that they are waves that travel in space at infinite distances. What criteria should we use when we choose between these possibilities?

The fact that the electric and magnetic fields cannot induce each other in vacuum *where there are no electric currents and no electric charges*, would prevent an honest scientist from interpreting them as being waves propagating freely in empty space.

However, the significance of the famous expression that yields the speed of light in vacuum

$$c = \frac{1}{\sqrt{\mu_0 \cdot \varepsilon_0}}$$

is not lost. This is *because the wave equations for* **E** *and* **B** *can be obtained even in regions where there are charges and currents*. Here is the proof, following the same method as shown in the above excerpt from David J. Griffiths, *Introduction to Electrodynamics*, Prentice Hall, 1999, p. 375:

We consider a region of space in which there is a charge density ρ and a current density $\mathbf{J}$. The equations are:

(i) $\nabla \cdot \mathbf{E} = \dfrac{\rho}{\varepsilon_0}$ (iii) $\nabla \times \mathbf{E} = -\dfrac{\partial \mathbf{B}}{\partial t}$

(ii) $\nabla \cdot \mathbf{B} = 0$ (iv) $\nabla \times \mathbf{B} = \mu_0 \cdot \mathbf{J} + \mu_0 \cdot \varepsilon_0 \cdot \dfrac{\partial \mathbf{E}}{\partial t}$

We proceed in the same way as in the said textbook and apply curl to (iii) and (iv).

Curl of (iii) yields:

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla \cdot (\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = \nabla \times \left(-\frac{\partial \mathbf{B}}{\partial t} \right) = -\frac{\partial}{\partial t}(\nabla \times \mathbf{B}) = -\frac{\partial}{\partial t}\left(\mu_0 \cdot \mathbf{J} + \mu_0 \cdot \varepsilon_0 \cdot \frac{\partial \mathbf{E}}{\partial t} \right)$$

So we have, $\nabla \cdot (\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\mu_0 \cdot \frac{\partial \mathbf{J}}{\partial t} - \mu_0 \cdot \varepsilon_0 \cdot \frac{\partial^2 \mathbf{E}}{\partial t^2}$

or, by using (i),

$$\frac{1}{\varepsilon_0} \cdot \nabla \rho - \nabla^2 \mathbf{E} = -\mu_0 \cdot \frac{\partial \mathbf{J}}{\partial t} - \mu_0 \cdot \varepsilon_0 \cdot \frac{\partial^2 \mathbf{E}}{\partial t^2} \qquad \text{(Eq. M1)}$$

Curl of (iv) yields:

$$\nabla \times (\nabla \times \mathbf{B}) = \nabla \cdot (\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B} = \nabla \times \left(\mu_0 \cdot \mathbf{J} + \mu_0 \cdot \varepsilon_0 \cdot \frac{\partial \mathbf{E}}{\partial t} \right)$$

or,

$$\nabla \cdot (\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B} = \mu_0 \cdot \nabla \times \mathbf{J} + \mu_0 \cdot \varepsilon_0 \cdot \frac{\partial}{\partial t}(\nabla \times \mathbf{E}) = \mu_0 \cdot \nabla \times \mathbf{J} - \mu_0 \cdot \varepsilon_0 \cdot \frac{\partial^2 \mathbf{B}}{\partial t^2}$$

So we have, $\nabla \cdot (\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B} = \mu_0 \cdot \nabla \times \mathbf{J} - \mu_0 \cdot \varepsilon_0 \cdot \frac{\partial^2 \mathbf{B}}{\partial t^2}$

or, by using (ii)

$$\nabla^2 \mathbf{B} = -\mu_0 \cdot \nabla \times \mathbf{J} + \mu_0 \cdot \varepsilon_0 \cdot \frac{\partial^2 \mathbf{B}}{\partial t^2} \qquad \text{(Eq. M2)}$$

Observe that Eq.M2 becomes the equation of a wave $\nabla^2 \mathbf{B} = \mu_0 \cdot \varepsilon_0 \cdot \frac{\partial^2 \mathbf{B}}{\partial t^2}$ for the magnetic field **B** if $\nabla \times \mathbf{J} = 0$ *without being necessary to use the condition for free space with no charge and no current* ($\rho \neq 0$ and $\mathbf{J} \neq 0$). Since the equation was obtained from the normal set of Maxwell's equation *with* charges and currents, it follows that even in Maxwell's theory *we cannot say that this is a wave corresponding to vacuum*.

Also observe that Eq.M1 becomes the equation of a wave $\nabla^2 \mathbf{E} = \mu_0 \cdot \varepsilon_0 \cdot \frac{\partial^2 \mathbf{E}}{\partial t^2}$ for the electric field **E**, if $\nabla \rho = 0$ and $\frac{\partial \mathbf{J}}{\partial t} = 0$. These equations tell that there can be charges ($\rho \neq 0$) but no charge gradient ($\nabla \rho = 0$) and there can be currents ($\mathbf{J} \neq 0$) but no time changing currents ($\frac{\partial \mathbf{J}}{\partial t} = 0$) for this wave equation to obtain.

It can be seen that the "electromagnetic wave equations" are valid for matter containing charges and currents and *no reason can be given for considering that they represent waves in vacuum*.

What is then the significance of the speed $c = \frac{1}{\sqrt{\mu_0 \cdot \varepsilon_0}}$?

I think that it must be sought in the dynamics involving the changes electric field – movement of electric charge – magnetic field *inside or in the vicinity of matter*. It may correspond to the speed with which disturbances in the magnetic field in and around matter propagate along a field line or from charge to charge.

In conclusion, my opinion is that this first part of the study on radio waves showed that Maxwell's theory of electromagnetic waves is untenable because it contains unacceptable ambiguities, false methods of theoretical investigation, unfounded assumptions and even predictions contrary to observations. In the absence of a correct theory, other mechanisms must be sought to explain the propagation in space of radio waves, light, and other disturbances. This will be attempted in the second part of this study.

--------------------- End of Part I ---------------------