Chapter 8
Solving Maxwell's equations

June 21, 2021

Potentials and gauge transformations are introduced for time-dependent situations. These lead to wave equations relating the potentials to the source. Plane wave solutions in vacuum are presented and the analog between EM waves and elastic waves is introduced through their respective Lagrangians.

Contents

1 The potentials

Much of classical electrodynamics is about solving Maxwell's equations for the fields from the given ρ and $\mathbf{J}$. The standard method proceeds in the following steps, in parallel with the way magnetostatics is handled.

- The homogeneous Maxwell equations are used to introduce potentials.
- The inhomogeneous Maxwell equations then lead to second-order partial differential equations for the potentials.
- A choice of gauge decouples these equations, leading to independent equations, one for each component of the potential.

The analysis is presented in an elementary way; later, the same steps will be repeated in the covariant formalism, regarding the scalar potential Φ and the vector potential $\mathbf{A}$ as components of a 4-vector $\vec{A} = (\Phi, \mathbf{A})$.[1]

1.1 The vector potential

Since $\nabla \cdot \mathbf{B} = 0$ remains unchanged even when the fields are time-dependent, the vector potential can be introduced as before:

$$\boxed{\mathbf{B}(\mathbf{r}, t) = \nabla \times \mathbf{A}(\mathbf{r}, t)} \tag{1}$$

The time-dependence has been indicated explicitly: the same procedure as in magnetostatics is repeated independently at each t.

1.2 The scalar potential

In electrostatics the scalar potential is introduced by $\mathbf{E} = -\nabla\Phi$, possible because $\nabla \times \mathbf{E} = 0$. But the latter condition fails when there is time-dependence, so Φ cannot be defined this way. The condition on $\nabla \times \mathbf{E}$ is now replaced by Faraday's law:

$$\nabla \times \mathbf{E} \quad = \quad -\frac{\partial \mathbf{B}}{\partial t} \tag{2}$$

Put (1) into (2), interchange the order of differentiation, and move both terms to the same side:

$$\nabla \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \right) \quad = \quad 0$$

The expression in parentheses has zero curl, so can be written as the gradient of a scalar, thus introducing Φ:

[1] Up to some factors of c which we shall be more careful with later.

$$\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t} \quad = \quad -\boldsymbol{\nabla}\Phi$$

The above defines Φ at each t. The minus sign accords with the convention in electrostatics. Rewrite in the more usual form, and explicitly indicate the spatial and temporal arguments:

$$\boxed{\mathbf{E}(\mathbf{r},t) = -\boldsymbol{\nabla}\Phi(\mathbf{r},t) - \frac{\partial \mathbf{A}(\mathbf{r},t)}{\partial t}} \qquad (3)$$

generalizing the expression in electrostatics.

1.3 The voltage in circuits

The problem

Here we digress to discuss several related points often glossed over in elementary courses — simple yet a common source of confusions.

Consider any circuit with inductors; these carry magnetic flux linkages φ_B proportional to the current I through them:

$$\varphi_B \quad = \quad LI \qquad (4)$$

When there is a changing flux, $\oint \mathbf{E}\cdot d\boldsymbol{\ell} \neq 0$. Then the usual definition of the voltage V_P (in electrostatics, the same as Φ) at any point P in the circuit (with respect to a reference point O)

$$V_P \quad = \quad -\int_O^P \mathbf{E}\cdot d\boldsymbol{\ell}$$

becomes problematic: the RHS depends on path. Question 1: What path is to be chosen when we talk about the voltage or potential difference (pd)?

Next, Kirchhoff's law states that the algebraic sum of pd's around a closed loop is zero. That is just another way of stating $\oint \mathbf{E}\cdot d\boldsymbol{\ell} = 0$ — which is clearly untrue in general! Question 2: How should we understand Kirchhoff's law in these circumstances?

The inductance L is defined in terms of a flux (or flux linkage), through the formula (4). But a flux can only be defined for a *closed* loop. An inductor is shown schematically in **Figure 1a**, with two terminals B and C. Going from B to C through the inductor does *not* constitute a closed loop, and implicitly we must complete the loop by adding another path γ, as illustrated schematically in **Figure 1b**. Question 3: Which path γ should be used to complete the loop and hence define φ_B and L? This ambiguity is accentuated for an inductor of only one turn (**Figure 1c**); the two ways to complete the loop would give totally different values of flux and L.

An example

The circuit shown in **Figure 2a** is made up of a battery with EMF $\mathcal{E}_0$, a resistor R and an inductor L (a solenoid), drawn off to one side for clarity. The magnetic field is assumed to be confined to the shaded region. Choose the direction indicated as positive I, and hence magnetic field and flux are counted as positive if pointing into the page.

Various points A, B, C are labelled on the circuit. For convenience, points not separated by a component, which are therefore necessarily at the same potential, are labelled by the same letter. Define the following pd's, with the convention $V_{XY} = V_X - V_Y$.

$$\begin{aligned} V_{BA} &= \text{pd across battery} \\ V_{CB} &= \text{“pd across inductor”} \\ V_{AC} &= \text{pd across resistor} \end{aligned}$$

Voltage across inductor

The innocuous looking term "pd across inductor" has several different meanings which are by no means the same. By definition,

$$V_{CB} \quad = \quad -\int_B^C \mathbf{E}\cdot d\boldsymbol{\ell}$$

However, this integral is path-dependent, and we consider the integral for three paths γ_i, to be denoted at $V_{CB}(\gamma_i)$; see **Figure 2b**.

- γ_1 is to the left of the flux.
- γ_2 is along the wires in the inductor.
- γ_3 is to the right of the flux.

What can be said about these versions of the pd?

Statement 1. In applying Kirchhoff's law, we must use $V_{CB}(\gamma_1)$ because it is related to the statement

$$\oint_{\Gamma_1} \mathbf{E}\cdot d\boldsymbol{\ell} \quad = \quad 0 \qquad (5)$$

where Γ_1 is the closed loop shown in **Figure 2c**, in which the last leg from B to C is along γ_1. This loop does not enclose any flux, so (5) is valid. But (5) is *not* true if we use γ_2 or γ_3 instead of γ_1.

Statement 2.

$$V_{CB}(\gamma_1) - V_{CB}(\gamma_3) = -\frac{d\varphi'_B}{dt} \qquad (6)$$

where φ'_B is the flux enclosed by the closed loop[2] $\gamma_1 - \gamma_3$. Since there is only *one* turn, φ'_B is much smaller compared to the flux linkage φ_B through the solenoid; in fact, if there are N turns, $\varphi_B \sim N\varphi'_B$.

Statement 3. Likewise

$$V_{CB}(\gamma_1) - V_{CB}(\gamma_2) = -\frac{d\varphi_B}{dt} \qquad (7)$$

[2] The notation means the loop going along γ_1 and then in the reverse direction along γ_3.

where φ_B is the flux linkage through the closed loop $\gamma_1 - \gamma_2$. This is the closed loop whose flux defines L.

Statement 4. Along the wire

$$V_{CB}(\gamma_2) = 0 \tag{8}$$

because the integral is over a conductor (with resistance assumed negligible) in which the electric field must be zero.

Given all these, in particular the combination of (7) and (8), we have, for the pd that appears in Kirchhoff's law

$$V_{CB}(\gamma_1) = -\frac{d\varphi_B}{dt} = -L\frac{dI}{dt} \tag{9}$$

The other pd's are

$$V_{BA} = \mathcal{E}_0 \qquad , \qquad V_{AC} = -IR$$

So adding up the pd's, Kirchhoff's law gives

$$\mathcal{E}_0 - L\frac{dI}{dt} - IR = 0 \tag{10}$$

which is the correct circuit equation.[3]

What would a voltmeter show?
Consider a voltmeter VM1 connected as in **Figure 3a**. The voltage it shows is obviously

$$V_{CB}(\gamma_1) = L\frac{dI}{dt}$$

On the other hand, a voltmeter VM3 connected as in **Figure 3b** would show

$$V_{CB}(\gamma_3) = (1-\varepsilon)\, L\frac{dI}{dt}$$

where $\varepsilon = \varphi'_B/\varphi_B$ because of the difference of φ'_B in (6). This prefactor is unimportant for practical inductors with $N \gg 1$.

2 Gauge transformations

There is considerable freedom in defining the potential. Given any function $\Lambda(\mathbf{r},t)$, the fields $\mathbf{E}$ and $\mathbf{B}$ are unchanged if the potentials are altered by

$$\begin{aligned} \Phi &\mapsto \Phi - \frac{\partial \Lambda}{\partial t} \\ \mathbf{A} &\mapsto \mathbf{A} + \boldsymbol{\nabla}\Lambda \end{aligned} \tag{11}$$

as can be checked directly from (1) and (3). This is evidently a generalization of the gauge transformation introduced in magnetostatics.

[3] Two properties should be checked. (a) At steady state, I is positive: it flows in the direction driven by the battery. (b) For the homogeneous equation, i.e., removing the battery, we should have $dI/dt \propto I$ with a negative proportionality constant, in order to have a damped rather than exponentially growing solution.

3 Equations for the potential

The two homogeneous equations have already been used up in defining the potentials. Next put the potentials into the two inhomogeneous equations.

3.1 Arbitrary gauge

Gauss' law
Put (3) into Gauss' law:

$$\boldsymbol{\nabla}\cdot\left(-\boldsymbol{\nabla}\Phi - \frac{\partial \mathbf{A}}{\partial t}\right) = \frac{\rho}{\epsilon_0}$$

Upon simplification

$$-\nabla^2\Phi - \frac{\partial}{\partial t}(\boldsymbol{\nabla}\cdot\mathbf{A}) = \frac{\rho}{\epsilon_0} \tag{12}$$

Ampere's law
Ampere's law with displacement current can be written as (putting all the fields on the LHS)

$$\boldsymbol{\nabla}\times\mathbf{B} - \mu_0\epsilon_0\frac{\partial \mathbf{E}}{\partial t} = \mu_0\mathbf{J}$$

Put (1) and (3) into this to get

$$\begin{aligned} &\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{A}) - \mu_0\epsilon_0\frac{\partial}{\partial t}\left(-\boldsymbol{\nabla}\Phi - \frac{\partial \mathbf{A}}{\partial t}\right) \\ &= \mu_0\mathbf{J} \end{aligned}$$

Simplifying, we get

$$\begin{aligned} &\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{A}) - \nabla^2\mathbf{A} \\ &+ \mu_0\epsilon_0\frac{\partial}{\partial t}(\boldsymbol{\nabla}\Phi) + \mu_0\epsilon_0\frac{\partial^2 \mathbf{A}}{\partial t^2} = \mu_0\mathbf{J} \end{aligned}$$

which can be written as

$$-\nabla^2\mathbf{A} + \mu_0\epsilon_0\frac{\partial^2\mathbf{A}}{\partial t^2} + \boldsymbol{\nabla}\mathcal{G} = \mu_0\mathbf{J} \tag{13}$$

where

$$\boxed{\mathcal{G} = \boldsymbol{\nabla}\cdot\mathbf{A} + \mu_0\epsilon_0\frac{\partial \Phi}{\partial t}} \tag{14}$$

Thus we are left with solving four *coupled* equations: one in (12) and three in (13).

3.2 Decoupling

Decoupled equation
Now we claim that we can make a gauge transformation such that

$$\mathcal{G} = 0 \tag{15}$$

If this is the case, then (13) simplifies to the three decoupled equations

$$\left(-\nabla^2 + \mu_0\epsilon_0 \frac{\partial^2}{\partial t^2} \right) \mathbf{A} \quad = \quad \mu_0 \mathbf{J} \qquad (16)$$

with each component of $\mathbf{A}$ related only to the corresponding component of $\mathbf{J}$.

At the same time, the condition (15) implies that the second term in (12) becomes

$$-\frac{\partial}{\partial t}\left(-\mu_0\epsilon_0 \frac{\partial \Phi}{\partial t}\right) \quad = \quad \mu_0\epsilon_0 \frac{\partial^2 \Phi}{\partial t^2}$$

and when this is put into (12) we find

$$\left(-\nabla^2 + \mu_0\epsilon_0 \frac{\partial^2}{\partial t^2} \right) \Phi \quad = \quad \frac{\rho}{\epsilon_0} \qquad (17)$$

which has the same structure as (16).

In fact, we can put all four equations together:

$$\boxed{D \begin{pmatrix} \Phi \\ \mathbf{A} \end{pmatrix} = \begin{pmatrix} \rho/\epsilon_0 \\ \mu_0 \mathbf{J} \end{pmatrix}} \qquad (18)$$

where D is the operator for the wave equation (also called the d'Alembertian)

$$\boxed{D = -\nabla^2 + \mu_0\epsilon_0 \frac{\partial^2}{\partial t^2}} \qquad (19)$$

The condition $\mathcal{G} = 0$ defines the *Lorenz gauge*.[4]

It is an interesting point of history that Maxwell's original derivation of these results did not make use of a gauge choice, and the derivation was strictly not valid. He circumvented the difficulty with an intuitive (but not mathematically valid) assumption. An account can be found in Ref. [1].

Freedom to choose gauge

We need to go back and check that it is indeed possible to make $\mathcal{G}$ vanish by a gauge transformation.

Under (11), we have

$$\begin{aligned} \mathcal{G} \quad &\mapsto \quad \mathcal{G} + \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla}\Lambda) - \mu_0\epsilon_0 \frac{\partial}{\partial t}\left(\frac{\partial \Lambda}{\partial t}\right) \\ &= \quad \mathcal{G} - D\Lambda \end{aligned}$$

which can be made to vanish if we can guarantee a solution to

$$D\Lambda \quad = \quad f \qquad (20)$$

for any given f. This will be shown in a later chapter.

By the way, given one potential in the Lorenz gauge, we can perform another transformation with Λ satisfying $D\,\Lambda = 0$. The result is *another* potential in the Lorenz gauge. In other words, the Lorenz condition does not select one potential, but a class of potentials.

4 Solution in free space

This Section deals with the case of vacuum, for which the homogeneous version of (18) applies. The case with sources is handled in the next few Chapters.

It is possible, indeed in some ways simpler, to derive plane wave solutions directly from Maxwell's equations in vacuum using $\mathbf{E}$ and $\mathbf{B}$, without even mentioning potentials. But we choose to present a formalism using potentials, first to align with the case where there are sources (for which potentials are normally used), and second to introduce the idea that EM energy can be thought of the sum of a kinetic energy and a potential energy.

4.1 Scalar analog

First start with the analog of a scalar field $\phi(\mathbf{r}, t)$ satisfying

$$D\phi \quad = \quad 0$$

Try a plane-wave solution[5]

$$\phi(\mathbf{r}, t) \quad = \quad \phi_0 \exp\left[i(\mathbf{k}\cdot\mathbf{r} - \omega t)\right]$$

Acting on such a wave

$$\begin{aligned} \boldsymbol{\nabla} \quad &\mapsto \quad i\mathbf{k} \\ \frac{\partial}{\partial t} \quad &\mapsto \quad -i\omega \end{aligned}$$

so the differential operator D in effect becomes a multiplicative factor

$$D \quad = \quad k^2 - \mu_0\epsilon_0\omega^2$$

and the condition $D = 0$ gives

$$c^2 \quad = \quad \frac{\omega^2}{k^2} = (\mu_0\epsilon_0)^{-1} \qquad (21)$$

where $c = \omega/k$ is obviously the speed of the wave.

[4] From the Wikipedia entry on Lorenz gauge condition: The Lorenz condition is named after Ludvig Lorenz. It is a Lorentz invariant condition, and is frequently called the "Lorentz condition" because of confusion with Hendrik Lorentz, after whom Lorentz covariance is named.

[5] Always with the understanding that the real part is to be taken.

4.2 EM waves

The same argument applies to each component of (18), and we conclude that the potential behaves as a plane wave in free space, with wave velocity given by (21).

Putting in numbers, we find

$$c = 3.0 \times 10^8 \text{ m s}^{-1}$$

the same as the velocity of light. Thus Maxwell concluded that light is a kind of EM waves.

To emphasize how remarkable this result is (but see Section 5), note that electrostatic and magnetostatic parameters ϵ_0 and μ_0 can be determined without reference to any time-dependent phenomena.

Henceforth make use of (21) and write D as

$$\boxed{D = -\nabla^2 + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}} \tag{22}$$

In 3D space, the Pythagorian invariant interval is $x^2+y^2+z^2$, so the natural second-order derivative

$$\nabla^2 = \partial_x^2 + \partial_y^2 + \partial_z^2$$

is also an invariant. In 4D spacetime, the invariant interval is $x^2+y^2+z^2-(ct)^2$, so the natural second-order derivative is (up to an overall sign which is merely a matter of convention) the operator D, which is likewise an invariant.

4.3 Polarization

The defining equations are, from (18) and specializing to vacuum,

$$D\begin{pmatrix}\Phi \\ \mathbf{A}\end{pmatrix} = 0 \tag{23}$$

and the solution is, up to an arbitrary overall amplitude and with the understanding that ω and k are related as $\omega = ck$:

$$\begin{pmatrix}\Phi \\ \mathbf{A}\end{pmatrix} = \begin{pmatrix}e_0 \\ \mathbf{e}\end{pmatrix} \exp\left[i(\mathbf{k}\cdot\mathbf{r} - \omega t)\right] \tag{24}$$

The column factor in the prefactor can be regarded as the relevant eigenvector, or polarization for the wave. The eigenvector is characterized (to start with) by four numbers (a scalar e_0 and a 3-vector $\mathbf{e}$). What can be said about these four numbers?

First, the above is derived for a particular gauge, in which $\mathcal{G} = 0$, or

$$\begin{aligned}
\boldsymbol{\nabla}\cdot\mathbf{A} + \frac{1}{c^2}\frac{\partial\Phi}{\partial t} &= 0 \\
i\mathbf{k}\cdot\mathbf{e} + c^{-2}(-i\omega)e_0 &= 0
\end{aligned}$$

which gives

$$\boxed{e_0 = \frac{\mathbf{k}\cdot\mathbf{e}}{k}\, c} \tag{25}$$

Thus e_0 is completely determined by the 3-vector $\mathbf{e}$ — in fact by its longitudinal component (i.e., the component along the wave vector).

Second, consider $\mathbf{E}$ and $\mathbf{B}$.

$$\begin{aligned}
\mathbf{E} &= -\boldsymbol{\nabla}\Phi - \frac{\partial\mathbf{A}}{\partial t} \\
&= (-i\mathbf{k}e_0 + i\omega\mathbf{e}) \exp\left[i(\mathbf{k}\cdot\mathbf{r} - \omega t)\right] \\
&= i\omega\left(\mathbf{e} - \mathbf{k}\,\frac{\mathbf{k}\cdot\mathbf{e}}{k^2}\right) \\
&\qquad \times \exp\left[i(\mathbf{k}\cdot\mathbf{r} - \omega t)\right]
\end{aligned} \tag{26}$$

Transverse projection
Consider the vector in brackets in (26):

$$\begin{aligned}
\mathbf{e}^{\mathrm{T}} &\equiv \mathbf{e} - \mathbf{k}\,\frac{\mathbf{k}\cdot\mathbf{e}}{k^2} \\
e_i^{\mathrm{T}} &= T_{ij}\, e_j
\end{aligned} \tag{27}$$

where

$$T_{ij} = \delta_{ij} - \frac{k_i k_j}{k^2} = \delta_{ij} - n_i n_j$$

where $\hat{\mathbf{n}}$ is the unit vector in the direction of $\mathbf{k}$. The matrix T is the transverse projection operator, in the sense of the following Problem.

Problem 1
Consider T as an operator or matrix.
(a) Show that $T^2 = T$. This is the definition of a projection operator.
(b) Show that for any operator satisfying $T^2 = T$, the allowed eigenvalues are 0 and 1. This means some eigenvectors are annihilated and other eigenvectors are exactly preserved.
(c) Show that if $\mathbf{v}$ is a vector parallel to $\mathbf{k}$, then $T\mathbf{v} = 0$; i.e., $T_{ij}v_j = 0$.
(d) Show that if $\mathbf{v}$ is a vector perpendicular to $\mathbf{k}$, then $T\mathbf{v} = \mathbf{v}$, i.e., $T_{ij}v_j = v_i$. §

We shall adopt the schematic notation

$$\mathbf{e}^{\mathrm{T}} = (T\mathbf{e})$$

so (26) is written simply as

$$\mathbf{E} = i\omega\,(T\mathbf{e})\, \exp\left[i(\mathbf{k}\cdot\mathbf{r} - \omega t)\right] \tag{28}$$

In short, only the transverse components of $\mathbf{e}$ matter as far as $\mathbf{E}$ is concerned. Next

$$\begin{aligned}
\mathbf{B} &= \boldsymbol{\nabla}\times\mathbf{A} \\
&= i(\mathbf{k}\times\mathbf{e})\, \exp\left[i(\mathbf{k}\cdot\mathbf{r} - \omega t)\right]
\end{aligned} \tag{29}$$

Again only the transverse components of $\mathbf{e}$ matter as far as $\mathbf{B}$ is concerned.

Since the longitudinal component of $\mathbf{e}$ is irrelevant, we shall henceforth set it to zero.[6] Then from (25), $e_0 = 0$. So we are left with only two components in (24) — the two spatial components transverse to $\mathbf{k}$. Also, since $\Phi = 0$ (because we have chosen $e_0 = 0$),

$$\begin{aligned} \mathbf{E} &= -\frac{\partial \mathbf{A}}{\partial t} \\ \mathbf{B} &= \boldsymbol{\nabla} \times \mathbf{A} \end{aligned} \tag{30}$$

which is valid not just for a solution such as (24), but also for a superposition of such solutions.

Transverse waves

Thus we have shown that EM waves are transverse. A more direct proof comes from applying, in vacuum,

$$\begin{aligned} \boldsymbol{\nabla} \cdot \mathbf{E} &= 0 \\ i\mathbf{k} \cdot \mathbf{E} &= 0 \end{aligned} \tag{31}$$

Note that we have two conditions on e_0 and $\mathbf{e}$. One corresponds to this physical transverse condition; the other is a gauge choice.

Problem 2

Consider a plane wave propagating in the $+z$ direction, with $\mathbf{E} = E\,\hat{\mathbf{e}}_x$.
(a) Find the direction of $\mathbf{B}$. Hint: Apply Faraday's law to the plane wave.
(b) Show that $\mathbf{k}$ is in the direction of $\mathbf{E} \times \mathbf{B}$.
(c) Show that the magnitude of $\mathbf{B}$ is $B = E/c$. Thus when a plane wave impinges on a non-relativistic particle, the ratio of the magnetic force to electric force is $F_B/F_E \sim v/c \ll 1$. §

5 Analog

The last Section showed wave solutions, with the wave speed determined by two *static* parameters, μ_0 and ϵ_0. The latter property would have seemed remarkable, as indeed it did to Maxwell and his contemporaries. But from another point of view, this result is not surprising at all, as this Section explains through an analogy.

5.1 1D elastic waves

Consider a 1D medium. If there is a longitudinal displacement $\xi(x)$ which is not uniform, then there will be a strain. In fact, for a small segment from x to $x+\Delta x$, there is an elongation

$$\Delta\xi = \xi(x+\Delta x) - \xi(x)$$

and hence a strain

$$e(x) = \frac{\Delta\xi}{\Delta x} = \frac{\partial\xi}{\partial x}$$

(A partial derivative appears to indicate that the differentiation is done at constant t.) There is then a strain energy density[7]

$$\mathcal{U}(x) = \frac{M}{2} e(x)^2 = \frac{1}{2} M \left(\frac{\partial\xi}{\partial x}\right)^2$$

This can be taken to be the definition of the elastic modulus M, just as, for a Hookean spring with displacement ξ, $U = (1/2)k\xi^2$ defines the force constant k.

Likewise, the kinetic energy density is

$$\mathcal{K} = \frac{\rho}{2} \left(\frac{\partial\xi}{\partial t}\right)^2$$

where ρ is the mass density (i.e., mass per unit length). Thus the Lagrangian is

$$L = \int \mathcal{L}\, dx$$

where the Lagrangian density $\mathcal{L}$ is given by

$$\boxed{\mathcal{L} = \frac{\rho}{2} \left(\frac{\partial\xi}{\partial t}\right)^2 - \frac{M}{2} \left(\frac{\partial\xi}{\partial x}\right)^2} \tag{32}$$

consisting of a time derivative squared, and a spatial derivative squared.

For this system, the speed of elastic waves is well known to be

$$c^2 = \frac{M}{\rho} \tag{33}$$

Problem 3

Review the derivation of (33) by working out the Newtonian equation of motion, and also derive it by minimizing $S = \int L\, dt$. §

5.2 EM waves

The EM field in vacuum has a magnetic energy density

$$\mathcal{U} = \frac{1}{2\mu_0} \mathbf{B}^2 = \frac{1}{2\mu_0} (\boldsymbol{\nabla} \times \mathbf{A})^2$$

and an electric energy density

$$\mathcal{K} = \frac{\epsilon_0}{2} \mathbf{E}^2 = \frac{\epsilon_0}{2} \left(\frac{\partial \mathbf{A}}{\partial t}\right)^2 \tag{34}$$

[6] Or apply the operator T as necessary.

[7] We use script letters to denote the densities.

where we have used a specific gauge, namely (30), and denoted the electric term as $\mathcal{K}$ because it depends on the time derivative.

Thus we have

$$\boxed{\mathcal{L} = \frac{\epsilon_0}{2}\left(\frac{\partial \mathbf{A}}{\partial t}\right)^2 - \frac{1}{2\mu_0}(\boldsymbol{\nabla}\times\mathbf{A})^2} \tag{35}$$

similar to the case of the string, except that the spatial derivative has a more complicated vectorial form.

Thus we see the analogy

$$\begin{aligned} \rho &\mapsto \epsilon_0 \\ M &\mapsto \mu_0^{-1} \\ \frac{M}{\rho} &\mapsto (\mu_0\epsilon_0)^{-1} \end{aligned}$$

In other words, ϵ_0 and μ_0 determine the speed of light just as ρ and M determine the speed of elastic waves.

Problem 4
Derive the equation of motion for EM waves by minimizing $S = \int L\,dt = \int \mathcal{L}(\mathbf{r})\,d^3r\,dt$. §

The above involves a taste of the Lagrangian and the action for EM waves — but only in the context of free fields, and only for the purpose of illustrating the analogy with elasticity. A more complete treatment of the action formalism will be presented later.

But the above analogy is imperfect (and therefore potentially misleading) in one respect. For an elastic medium, the density and the modulus can be thought of as independent variables, and the wave speed is a consequence. For EM in vacuum, the logic is actually reverse: the universal speed c comes first, as a property of spacetime,[8] and a relationship between the two variables ϵ_0 and μ_0 emerges as a result — the latter two variables are not mutually independent. The above argument going from ϵ_0 and μ_0 to c^2 is only the historical sequence, not the modern understanding of the logical relationship.

A Supplement: Beams of finite lateral extent

Each supplement will consist of one or more problems, usually more challenging and of an open nature, to stretch the better students.

[8] Or more properly as a property of the units we choose to describe spacetime.

This Chapter has derived plane waves from Maxwell's equations, and it is assumed that the elementary properties of plane waves are well known from previous study of the subject. A *plane* wave by definition has an infinite lateral extent. This Supplement gives a brief introduction to waves with a *finite* lateral extent, say a, with $a \gg \lambda$. This limitation is good enough for describing optical beams with a lateral dimension of say 1 mm. For simplicity, we only construct such a beam limited in the x direction, leaving the y direction infinite.

A tilted plane wave
Consider a plane wave with frequency ω and hence wavenumber $k = \omega/c$, but travelling slightly off the z-axis. Thus the wavevector is assumed to be

$$\mathbf{k} = q\,\hat{\mathbf{e}}_x + k'\,\hat{\mathbf{e}}_z \tag{36}$$

where

$$\begin{aligned} k' &= \sqrt{k^2 - q^2} \\ &\approx k - \frac{q^2}{2k} \end{aligned} \tag{37}$$

Because the vector $\mathbf{k}$ is *slightly* off the z-axis, we are entitled to assume $|q| \ll k$, which justifies the approximation in (37).

Its polarization
If the polarization is

$$\mathbf{e} = u\,\hat{\mathbf{e}}_x + v\,\hat{\mathbf{e}}_z$$

the transverse condition leads to

$$0 = \mathbf{k}\cdot\mathbf{e} = qu + k'v$$

Thus

$$v = -\frac{q}{k'}\,u \approx -\frac{q}{k}u$$

Thus, up to normalization (i.e., choosing $u = 1$), and putting in the time-dependent factor, we have

$$\begin{aligned} \mathbf{E} = {} & \left(\hat{\mathbf{e}}_x - \frac{q}{k}\hat{\mathbf{e}}_z\right)\exp(iqx)\exp[-iq^2z/(2k)] \\ & \cdot\exp[i(kz-\omega t)] \end{aligned} \tag{38}$$

Now we can replace

$$q \mapsto -i\frac{\partial}{\partial x}$$

with which we then have

$$\begin{aligned} E_x &= \exp(iqx)\exp[-iq^2z/(2k)] \\ &\quad \cdot\exp[i(kz-\omega t)] \end{aligned} \tag{39}$$

$$E_z = \frac{i}{k}\frac{\partial E_x}{\partial x} \tag{40}$$

From this point onwards, we omit the common plane-wave factor

$$\exp[i(kz - \omega t)]$$

Superpose many such waves
We now assume

$$\begin{aligned}\mathbf{E}_x &= \int \frac{dq}{2\pi} f(q) \\ &\quad \times \exp(iqx) \exp[-iq^2 z/(2k)] \quad (41)\end{aligned}$$

for some weight function $f(q)$; the formula (40) for E_z, because it has been written without any q, survives superposition over different values of q.

The function $f(q)$ must fall off rapidly as $|q|$ increases, since we limit to small q. We continue the analysis with the example

$$f(q) = \exp(-a^2 q^2/2)$$

Profile at the waist
First take the case of $z = 0$. (We shall see later why we call this the waist.) Then (41) reduces to

$$E_x = \int \frac{dq}{2\pi} e^{-a^2q^2/2} e^{iqx}$$

The exponent is

$$\begin{aligned} & -\frac{a^2}{2} q^2 + iqx \\ = & -\frac{a^2}{2} \left(q - ix/a^2\right)^2 - x^2/(2a^2)\end{aligned}$$

By changing variable to

$$q' = q - ix/a^2$$

we get

$$\begin{aligned}E_x &= C \exp[-x^2/(2a^2)] \\ |E_x|^2 &= |C|^2 \exp[-(x/a)^2]\end{aligned}$$

showing that the beam profile is gaussian, with a width a. We leave the overall constant C unevaluated.

The intensity also contains a small contribution due to E_z. Thus a finite beam propagating in the z direction has a small z-polarization.

Profile further downstream
For $z \neq 0$, we have to include the term

$$-\frac{iq^2 z}{2k}$$

in the exponent. Thus the coefficient of the $-q^2$ term in the exponent becomes

$$\frac{a^2}{2} \quad \mapsto \quad \frac{a^2}{2} + \frac{iz}{2k} \equiv \frac{b^2}{2}$$

Thus the previous result for E_x applies with the change $a^2 \mapsto b^2$, so the x-dependence becomes $E_x \propto \exp(-F)$, where

$$\begin{aligned}F &= \frac{x^2}{2b^2} = \frac{x^2}{2(a^2 + iz/k)} \\ &= x^2 \, \frac{a^2 - iz/k}{2(a^4 + z^2/k^2)}\end{aligned}$$

and $|E_x|^2 \propto \exp(-2\,\Re F)$, with

$$2\,\Re F = x^2 \frac{a^2}{a^4 + z^2/k^2} \equiv \frac{x^2}{a(z)^2}$$

where

$$\boxed{a(z)^2 = a^2 + \frac{z^2}{a^2k^2}} \quad (42)$$

is the lateral dimension at distance z.

Far away, we have approximately

$$a(z) \approx \frac{z}{ak} \quad (43)$$

The interpretation is as follows. By limiting to $|x| \sim a$ at $z = 0$, we must have $q \sim 1/a$ by the uncertainty principle, and hence the wave vector is inclined to the z-axis by an angle $\theta \sim q/k \sim 1/(ka)$. Thus, after propagating longitudinally for a distance z, the photon would have a transverse position $x \sim z\theta \sim z/(ka)$.

Incidentally, the way we have constructed the solution guarantees that the lateral dimension $a(z)$ in (42) is minimum at $z = 0$, which is therefore called the "waist".

Summary
So the qualitative lessons are as follows for a beam of finite lateral extent.

- There would have to be a bit of longitudinal polarization, of order $1/(ka)$.
- The beam would have to spread out, at an angle $\sim 1/(ka)$.

References

[1] PT Leung, "On Maxwell's discovery of electromagnetic waves and the gauge condition", Eur. J. Phys., **36**, 025002 (2015). DOI:10.1088/0143-0807/36/2/025002.

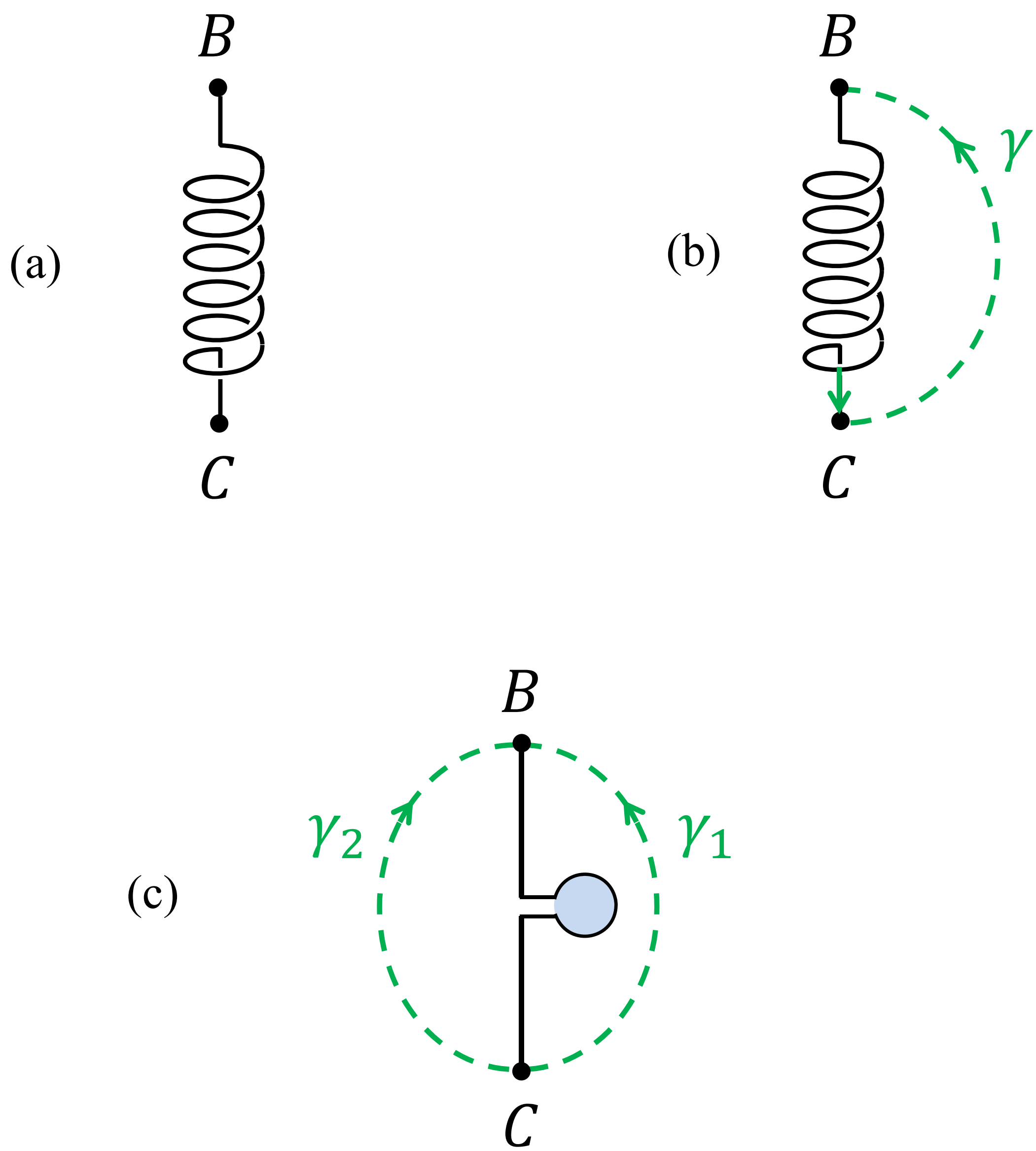


Figure 1 (a) An inductance
(b) The flux can only be defined if a path γ is added to form a closed loop
(c) An inductance with only one loop and two ways of adding a path γ

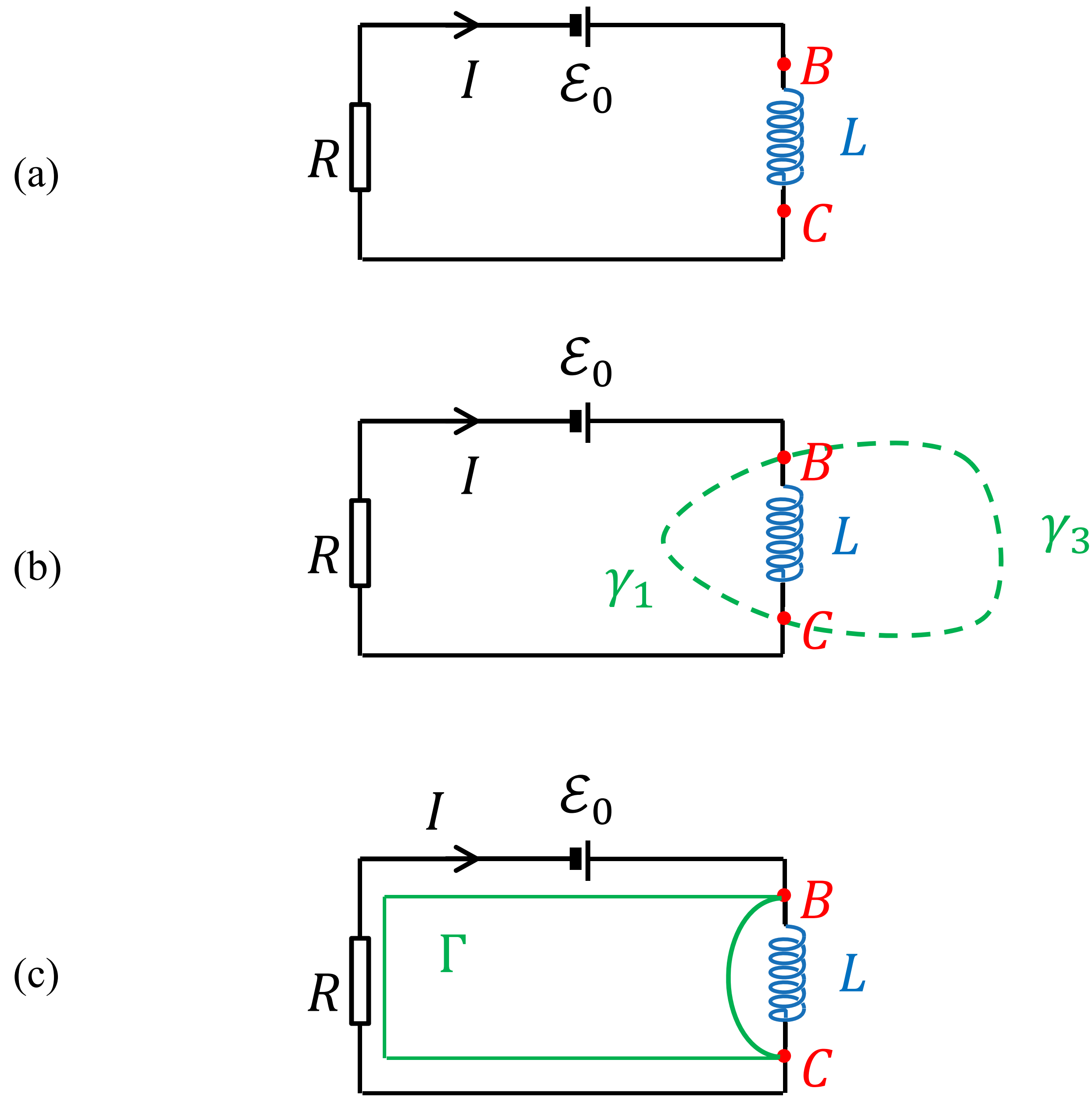
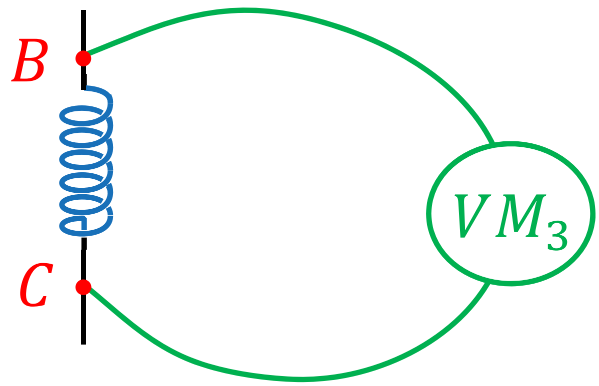


Figure 2 (a) The circuit with an inductance
(b) Completing the closed loop. γ_2 is along the wires of the solenoid
(c) The closed loop Γ involves γ_1 from B to C. Kirchhoff's law is applied to Γ

(a)



(b)

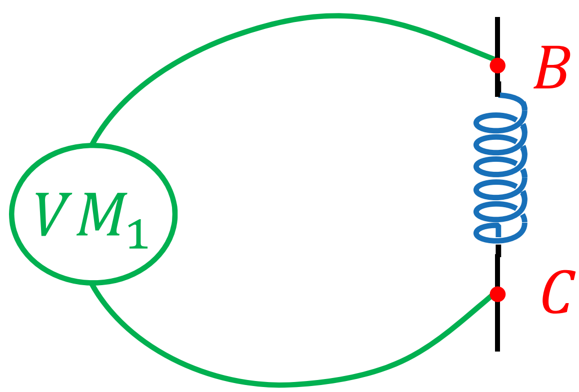


Figure 3   What does a voltmeter measure?