Electromagnetism

Jung Hoon Han

June 13, 2013

Contents

Chapter 1

Charges in Electromagnetic Fields

The subject of electrodynamics is composed of several parts. The first part deals with the way the charged particles behave under, or couple to, the influence of given electromagnetic (EM) fields. The second part concerns the way the electromagnetic fields propagate in space-time. The third part deals with the way the fields are generated by the charged particles themselves. From the theoretical perspective it means that the total action governing the electromagnetic behavior should consist of the action for the particle, its coupling to EM fields, and another action governing the dynamics of EM fields themselves. We will deal with them one by one, beginning with the action of a charged particle.

1.1 Action of a charged article

In this lecture the equation of motion of a charged particle will be presented from the perspective of its action. As in the point-particle classical mechanics, equation of motion will follow once the appropriate Lagrangian has been defined.

The amount of coupling of a given particle to a given EM environment is characterized by a single scalar quantity, called the charge q. Since the particle is assumed to be point-like, the charge's dynamics is characterized completely by the coordinate vector $\mathbf{r}$ and momentum (velocity) $\mathbf{p}$ ($\mathbf{v}$). Each particle has a unique charge, as well as mass, associated with it. The action of a free particle in relativistic mechanics is given by (m=matter)

$$S_m = -mc^2 \int_{t_1}^{t_2} \sqrt{1 - \frac{\mathbf{v}^2}{c^2}}\, dt, \tag{1.1}$$

which reduces to the familiar form $S = \int (m\mathbf{v}^2/2)\, dt$, up to an irrelevant constant, in the limit $\mathbf{v}^2/c^2 \ll 1$. Unlike in classical mechanics, the action is no longer the integral of the kinetic energy, which in relativistic mechanics takes the form

$$K = m\gamma c^2,\ \gamma = (1 - \mathbf{v}^2/c^2)^{-1/2}. \tag{1.2}$$

The manner in which the particle of charge q couples to the EM environment is dictated by the action (mf=matter-field)

$$S_{mf} = -q \int A_\mu dx^\mu. \tag{1.3}$$

The use of the four-vector A_μ ($\mu = t, x, y, z$) indicates that the theory of EM coupling to matter can only be done properly within the relativistic framework. The space component of the four-vector is called the vector potential $\mathbf{A}$, whereas the time component is called the scalar potential A_0. One thus writes $A_\mu = (A_0, \mathbf{A})$. Noting the convention $dx^\mu = (dt, -d\mathbf{r})$, the matter-field coupling can be written out like

$$S_{mf} = q \int (\mathbf{A} \cdot d\mathbf{r} - A_0 dt) = q \int (\mathbf{A} \cdot \mathbf{v} - A_0)\, dt. \tag{1.4}$$

We have introduced the velocity vector $\mathbf{v} = \dot{\mathbf{r}}$ in the above. Finally we obtain the combined action for matter and its coupling to EM field as

$$S = \int L\ dt,\ L = -mc^2\sqrt{1 - \frac{\mathbf{v}^2}{c^2}} + q\mathbf{A} \cdot \mathbf{v} - qA_0. \tag{1.5}$$

This is the action governing the relativistic dynamics of a point charge q in the presence of *fixed* EM backgrounds $(\mathbf{A}, A_0)$.

According to the recipe of Hamiltonian mechanics, we call the functional derivative of the Lagrangian L with respect to the velocity $\mathbf{v}$ as the momentum:

$$\mathbf{P} = m\gamma\mathbf{v} + q\mathbf{A} = \mathbf{p} + q\mathbf{A}. \tag{1.6}$$

The one denoted in upper-case $\mathbf{P}$ is the canonical momentum, while the lower-case momentum $\mathbf{p}$ is the mechanical momentum. We may also find the Hamiltonian of the matter-field system by the recipe

$$H = \mathbf{v} \cdot \frac{\partial L}{\partial \mathbf{v}} - L = m\gamma c^2 + qA_0. \tag{1.7}$$

Apparently, it seems as though the vector potential $\mathbf{A}$ have miraculously disappeared from the Hamiltonian! It must be noted, however, that the Hamiltonian is to be expressed in terms of the momentum, not velocity. The velocity-to-momentum conversion takes place by noting that $\mathbf{P} - q\mathbf{A} = m\gamma\mathbf{v}$, hence

$$(\mathbf{P} - q\mathbf{A})^2 = \frac{m^2\mathbf{v}^2}{1 - \mathbf{v}^2/c^2} \rightarrow m\gamma c^2 = \sqrt{m^2c^4 + c^2(\mathbf{P} - q\mathbf{A})^2}. \tag{1.8}$$

Substitution yields the full relativistic Hamiltonian of a point charge

$$H = \sqrt{m^2c^4 + c^2(\mathbf{P} - q\mathbf{A})^2} + qA_0. \tag{1.9}$$

In the non-relativistic limit this Hamiltonian reduces to a more familiar form

$$H = \frac{1}{2m}(\mathbf{P} - q\mathbf{A})^2 + qA_0, \tag{1.10}$$

again up to an irrelevant constant.

1.1.1 Geometric action

Unlike the particle motion in classical mechanics, we have all sorts of potentials that influence the charge's dynamics. All together it forms a relativistic four-vector A_μ, but for historical reason one distinguishes the time component A_0 (called the scalar potential) from the vector part $\mathbf{A}$ (called the vector potential). The action governing the coupling of the vector potential to the particle's velocity is particularly intriguing when one notes that

$$q\int \mathbf{A}\cdot\dot{\mathbf{r}}dt = q\int \mathbf{A}\cdot d\mathbf{r} \tag{1.11}$$

is independent of time, but only on the path taken by the particle. When thrown in the context of quantum mechanics this quantity appears as the phase factor of a charged particle moving in the background of the vector potential $\mathbf{A}$. For a particle moved around a closed loop $\mathcal{C}$ the phase picked up during the cycle is

$$\exp\left(i\frac{q}{\hbar}\oint \mathbf{A}\cdot d\mathbf{r}\right) = \exp\left(i\frac{q}{\hbar}\int \boldsymbol{\nabla}\times\mathbf{A}\cdot d\mathbf{S}\right) \tag{1.12}$$

by use of Stokes' theorem. Hence the phase ϕ is equal to the flux $\Phi = \int \mathbf{B}\cdot d\mathbf{S}$, divided by the flux quantum h/e and multiplied by 2π assuming $q = e$. In quantum mechanics this is known as the Aharonov-Bohm (AB) phase.

1.2 Equation of motion of a charged particle

A moving charge inevitably generates electric and magnetic fields. For one charge this effect is infinitesimal compared to the influence of fields acting it, so to a first approximation one can simply focus on how a charge moves in the background of *fixed* EM fields. Equation of motion of a charged particle can be derived from the Lagrangian or the Hamiltonian of the previous section. Euler-Lagrange equation

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \mathbf{v}}\right) = \frac{\partial L}{\partial \mathbf{r}} \tag{1.13}$$

gives the result

$$\frac{d\mathbf{p}}{dt} = q\left(-\boldsymbol{\nabla}A_0 - \frac{\partial \mathbf{A}}{\partial t}\right) + q\mathbf{v}\times(\boldsymbol{\nabla}\times\mathbf{A}). \tag{1.14}$$

Denoting $\boldsymbol{\nabla} \times \mathbf{A} \equiv \mathbf{B}$ and $-\boldsymbol{\nabla} A_0 - \frac{\partial \mathbf{A}}{\partial t} \equiv \mathbf{E}$ yields the familiar Lorentz's force law:

$$\frac{d\mathbf{p}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \tag{1.15}$$

Lorentz's law gives the time evolution of the mechanical momentum $\mathbf{p}$ of the particle, not the canonical momentum $\mathbf{P}$. In this regard $\mathbf{p}$ is to be regarded as the physical quantity whereas $\mathbf{P}$ is not. Another way of saying this is that $\mathbf{p}$ is gauge-invariant even though $\mathbf{P}$ is not.

The gauge transformation of the EM potential A_μ is implemented as

$$\mathbf{A}' = \mathbf{A} + \boldsymbol{\nabla} f, \; A_0' = A_0 - \frac{\partial f}{\partial t} \tag{1.16}$$

for an arbitrary space-time dependent function $f(\mathbf{r}, t)$. Under such transformation it is easy to verify that $\mathbf{B}$ and $\mathbf{E}$ are not changed at all. That means the r.h.s. of the Lorentz's equation is unchanged by the gauge transform (1.16), and neither should the l.h.s. involving the mechanical momentum $\mathbf{p}$. If $\mathbf{p}$ remains gauge invariant, $\mathbf{P}$ cannot be simultaneously so $\mathbf{P} = \mathbf{p} + q\mathbf{A}$ will change to $\mathbf{P}' = \mathbf{p} + q\mathbf{A}'$.

The physically independent degrees of freedom associated with $(\mathbf{B}, \mathbf{E})$ are not six, as the naive counting of field components suggests. Actually there are only four independent components embodied by $(A_0, \mathbf{A})$. Even this statement is not quite correct due to the gauge choice degree of freedom associated in selecting a particular set of $(A_0, \mathbf{A})$. The true number of independent variables associated with the electromagnetic field is thus only three. The procedure by which one goes from the seemingly four independent variables of the electromagnetic potential A_μ to the physically relevant three variables is called gauge-fixing. Examples of gauge-fixing will be given in subsequent chapters.

From symmetry point of view, l.h.s. of the Lorentz's equation is odd under the inversion, $\mathbf{r} \to -\mathbf{r}$, as it would alter $\mathbf{v} \to -\mathbf{v}$ and therefore $\mathbf{p} \to -\mathbf{p}$. The r.h.s. must also share this property, so that $\mathbf{E} \to -\mathbf{E}$, but $\mathbf{B}$ must remain $\mathbf{B}$ under inversion if $(\mathbf{v} \times \mathbf{B})$ is to change the sign to $-(\mathbf{v} \times \mathbf{B})$. Such a vector $\mathbf{B}$ is called axial, while all other vectors in the Lorentz equation are called polar. Under time reversal, $t \to -t$, $\mathbf{p} \to -\mathbf{p}$, $\mathbf{v} \to -\mathbf{v}$, both sides of Lorentz's equation will remain invariant provided that we require $\mathbf{E} \to \mathbf{E}$ and $\mathbf{B} \to -\mathbf{B}$. Thus the same motion for a given $(\mathbf{E}, \mathbf{B})$ is possible to execute in reverse time provided the EM environment is changed to $(\mathbf{E}, -\mathbf{B})$.

In non-relativistic mechanics the kinetic energy of a charged particle is only changed by the electric field, but not by the magnetic field. We comment before closing this section that the same statement holds true even for the relativistic case as the relativistic kinetic energy is given by $K = m\gamma c^2$ and, using the Lorentz equation of motion, the time rate of change of K is worked out to be

$$\frac{dK}{dt} = q\mathbf{v} \cdot \mathbf{E}. \tag{1.17}$$

1.3 Gauge invariance in quantum mechanics

The EM fields $(\mathbf{E}, \mathbf{B})$ are examples of gauge fields, satisfying a certain invariance property under the transformation of the underlying potentials. Loosely put all of the physical results must remain the same no matter what sort of gauge fields one chooses in the equation one solves, as long as the different choices lead to the same physical quantities in the EM fields. The gauge invariance can be seen at the level of the particle's action. Choosing a new set of potentials as in Eq. (1.16) will modify the Lagrangian (1.5) to

$$L' = L + q\boldsymbol{\nabla} f \cdot \dot{\mathbf{r}} + q\frac{\partial f}{\partial t} = L + q\frac{df}{dt}. \tag{1.18}$$

The equation of motion, as derived from the variation of the corresponding action, will be invariant under the change of the Lagrangian by a total derivative.

The principle of gauge invariance is such a general one, applicable to both classical and relativistic particles alike, of classical and quantum-mechanical particles alike. In this section we briefly discuss how the dynamics of a quantum-mechanical particle is ensured of its gauge invariance. A quantum-mechanical charged particle moving in the background of electromagnetic fields is not described by the Lorentz's force equation. Rather it is described by the Schrödinger equation

$$i\hbar\frac{\partial \psi}{\partial t} = \left[\frac{(\mathbf{P} - q\mathbf{A})^2}{2m} + qA_0\right]\psi. \tag{1.19}$$

For the same electromagnetic field backgrounds, however, someone might have chosen a different set of potentials $(\phi', \mathbf{A}')$ as given in Eq. (1.16) and worked with a different Schrödinger equation

$$i\hbar\frac{\partial \psi'}{\partial t} = \left[\frac{(\mathbf{P} - q\mathbf{A} - q\boldsymbol{\nabla} f)^2}{2m} + q\left(A_0 - \frac{\partial f}{\partial t}\right)\right]\psi'. \tag{1.20}$$

The wave functions ψ' satisfying the equation for the new potential field A'_μ is easily obtained from ψ obeying Eq. (1.19) by the phase factor multiplication

$$\psi'(\mathbf{r}, t) = \psi(\mathbf{r}, t)e^{iqf(\mathbf{r},t)/\hbar}. \tag{1.21}$$

It means two different scientists solving the same physical problem using different choices of potentials in their Schrödinger equations can nevertheless find a way to compare their respective solutions and find a match. Such is the statement of the *local* gauge invariance of quantum-mechanical wave function. We emphasize the local nature of the gauge transformation because when compared point-by-point, the two wave functions ψ' and ψ might seem to be completely different quantities even though they are not.

An aspect of quantum-mechanical gauge invariance which is not emphasized often enough is that the statement is only strictly applicable when there is no degeneracy of the eigenstates. Take, for example, the problem of uniform magnetic field $\mathbf{B} = (0, 0, B)$ for which two possible gauge field choices are

$\mathbf{A}_1 = (0, Bx, 0)$ and $\mathbf{A}_2 = (-By, 0, 0)$. The difference is expressed as the gradient $\mathbf{A}_1 - \mathbf{A}_2 = \boldsymbol{\nabla}(Bxy)$. In the first gauge choice $\mathbf{A} = \mathbf{A}_1$ one obtains a set of wave functions corresponding to a given Landau level index n as $\psi_{1,n,k_y}(x,y) = e^{ik_y y}\phi_n(x - k_y)$. A series Gaussian-enveloped functions $\phi_n(x - k_y)$ are obtained for each guiding-center momentum k_y. For $\mathbf{A} = \mathbf{A}_2$ one has a series of wave functions $\psi_{2,n,k_x}(x,y) = e^{ik_x x}\phi_n(y - k_x)$. Between any pair of $\psi_{1,n,k_y}(x,y)$ and ψ_{2,n,k_x} one readily discovers $\psi_{2,n,k_x}(x,y) \neq \psi_{1,n,k_x}(x,y)e^{iqBxy/\hbar}$. The reason is that an arbitrary linear combination of the degenerate states such as $\sum_{k_x} c_{k_x}\psi_{1,n,k_y}(x,y)$ is an equally valid eigenstate for the vector potential $\mathbf{A}_1$. Accordingly the statement of gauge invariance must appropriately be generalized:

$$\psi_{2,n,k_x}(x,y) = e^{iBxy/\hbar}\left(\sum_{k_y} c_{k_x,k_y}\psi_{1,n,k_y}(x,y)\right), \tag{1.22}$$

with a suitable choice of coefficients c_{k_x,k_y}. A similar strategy must be employed to ensure gauge invariance between any pair of degenerate set of eigenstates. The important statement is that one can always find a suitable linear combination of the degenerate states, such that a one-to-one correspondence of the type (1.21) can be established between the solutions using the vector potential $\mathbf{A}_1$ and $\mathbf{A}_2$.

1.4 Electromagnetic field tensor

Although $\mathbf{B}$ and $\mathbf{E}$ remain invariant under the gauge transform, they do not do so under the Galilean boost of the reference frame. What is seen as pure electric or pure magnetic field is seen as part electric and part magnetic field by another observer moving relative to the original one at constant velocity $\mathbf{V}$. In this sense, $(\mathbf{E}, \mathbf{B})$ are like components of some larger vector - tensor, to be exact - whose projected components onto given axes depend on the choice of the coordinate frame itself. The whole idea is embodied by the EM field tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{1.23}$$

The four derivatives are $\partial_\mu = (-\partial_t, \boldsymbol{\nabla})$. By definition $F_{\mu\nu}$ is anti-symmetric, $F_{\mu\nu} = -F_{\nu\mu}$, so out of 16 components only 6 will be independent. The 6 component can be grouped as

$$\begin{aligned} F_{0i} &= -\partial_t A_i - \partial_i A_0 = E_i, \\ F_{ij} &= \partial_i A_j - \partial_j A_i = \varepsilon_{ijk} B_k. \end{aligned} \tag{1.24}$$

Symbolically we may write $F_{\mu\nu} = (\mathbf{E}, \mathbf{B})$. Introducing the four-velocity $v_\mu = (-1, \mathbf{v})$, the Lorentz's equation is written in the simplified form

$$\frac{dp_i}{dt} = qF_{i\nu}v_\nu. \tag{1.25}$$

It is tempting to extend this equation to the fully four-dimensional form $dp_\mu/dt = qF_{\mu\nu}v_\nu$. In fact the time component of the four-vector p_μ is nothing but the kinetic energy K, $p_\mu = (K, \mathbf{p})$, then the first equation will be no other than the earlier result $dK/dt = q\mathbf{E}\cdot\mathbf{v}$.

An observer in a reference frame R' moving at velocity $\mathbf{V}$ relative to the initial frame R sees electric and magnetic field $\mathbf{E}'$ and $\mathbf{B}'$ related to those seen by the observer R, $(\mathbf{E}, \mathbf{B})$ according to[1]

$$\mathbf{E}' = \mathbf{E} - \mathbf{B}\times\mathbf{V},\ \mathbf{B}' = \mathbf{B} + \mathbf{E}\times\mathbf{V}/c^2. \tag{1.26}$$

If the magnetic field was absent in the first frame, $\mathbf{B} = 0$, the moving observer sees the non-zero magnetic field $\mathbf{B}'$ related to $\mathbf{E}$ by $\mathbf{B}' = \mathbf{E}\times\mathbf{V}/c^2$. There are two scalar quantities which remain invariant under the Lorentz transformation of the fields: $\mathbf{B}^2 - \mathbf{E}^2$ and $\mathbf{E}\cdot\mathbf{B}$. In fact we can introduce a second kind of field tensor $F^{\mu\nu} = (-\mathbf{E}, \mathbf{B})$ in addition to $F_{\mu\nu} = (\mathbf{E}, \mathbf{B})$, and write the two invariants as

$$2(\mathbf{B}^2 - \mathbf{E}^2) = F_{\mu\nu}F^{\mu\nu}, \quad \mathbf{E}\cdot\mathbf{B} = \varepsilon_{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}. \tag{1.27}$$

A nice application of the Lorentz transformation formula for the EM fields is the derivation of the Hall effect. When electrons in two dimensions are subject to an in-plane electric field $\mathbf{E}$ and a perpendicular magnetic field $\mathbf{B}$, a Hall current develops in the direction perpendicular to both $\mathbf{E}$ and $\mathbf{B}$. This is the Hall effect. From the relativistic transformation laws we see that to a moving observer with velocity $\mathbf{V} = c^2\mathbf{E}\times\mathbf{B}/E^2$, the magnetic field will appear to have vanished:

$$\mathbf{B}' = \mathbf{B} + \mathbf{E}\times\mathbf{V}/c^2 = 0, \tag{1.28}$$

leaving only the electric field

$$\mathbf{E}' = \mathbf{E} - \mathbf{B}\times\mathbf{V} \approx \mathbf{E} \tag{1.29}$$

to act on the electrons. Thus the electrons in the new reference frame will move only along the direction of the $\mathbf{E}$ field. In order for this statement to hold, we must deduce that in the original frame the electrons had a net drift velocity $\mathbf{v}_d = \mathbf{V}$, directed along the $\mathbf{E}\times\mathbf{B}$ direction. This is the explanation of the Hall effect, without invoking any details of the electron motion in the solid, involving only the relativity principle.

1.5 Problems

1. Verify Eq. (1.8).

2. Verify Eq. (1.14). Derive the same result using Hamilton's equations.

[1] Formulas are valid to first order in V/c. General formulas are much more complicated and dropped from the present text.

3. Verify Eq. (1.17).

4. Verify that $\mathbf{B}^2 - \mathbf{E}^2$ and $\mathbf{E} \cdot \mathbf{B}$ remain invariant under the Lorentz transformation, Eq. (1.26), to first order in V/c.

5. Establish the gauge invariance (1.22) for the lowest Landau level ($n = 0$) problem by finding a suitable set of coefficients c_{k_x,k_y}.

6. Verify Eq. (1.27).

7. Lorentz equation itself must be invariant under the Lorentz transformation (1.26). Prove this statement up to first order in V/c.

Chapter 2

EM Field Equations

The previous chapter deals with the motion of a charged particle under a fixed EM field environment, as summarized in the Lorentz's equation of motion. The celebrated Maxwell's equations, on the other hand, address the contrasting issue of how the fields behave, even even generated, in the presence of fixed charged backgrounds.

2.1 Maxwell's equation: the first pair

Some identities immediately follow from writing down the **E** and **B** fields in terms of the underlying potentials,

$$\mathbf{E} = -\boldsymbol{\nabla} A_0 - \frac{\partial \mathbf{A}}{\partial t}, \quad , \mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}. \tag{2.1}$$

The first one is obtained by taking the divergence of the **B** field:

$$\boldsymbol{\nabla} \cdot \mathbf{B} = \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{A}) = 0. \tag{2.2}$$

Although mathematically speaking a trivial consequence of vector calculus, the enormous implication of it is that magnetic monopoles do not exist in the universe, at least within the established theory of electromagnetism. The second theorem relates **E** with **B** through

$$\boldsymbol{\nabla} \times \mathbf{E} = -\boldsymbol{\nabla} \times \boldsymbol{\nabla} A_0 - \frac{\partial}{\partial t}(\boldsymbol{\nabla} \times \mathbf{A}). \tag{2.3}$$

The first term on the r.h.s. is again zero from vector identity, while the second one is just the time derivative of the magnetic field. Hence,

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{2.4}$$

which is precisely Faraday's law. Two of the four equations formulated by Maxwell are in fact nothing more than mathematical identities straight from

the definitions of EM fields in terms of the potentials. An alternative way to put it is that the two "empirical laws" as written down in Eq. (2.2) and (2.4) *constrain* the character of the EM fields themselves. In order to obey the two constraints imposed by Eqs. (2.2) and (2.4) the EM fields must be written in the manner of Eq. (2.1).

The other two of the Maxwell's equation address the way the charged matter density and current are related to the EM fields. Such equations would follow from the action involving the matter (S_m), the coupling of matter to field (S_{mf}), and finally, the action of the field itself (S_f). Among these the first two have been already written down in the previous chapter. In the following section we discuss how to write down the field action S_f.

2.2 Electromagnetic field action

The action for the EM fields must be a space-time integral of some quantity, a scalar, which is also invariant under the Lorentz transformation. Only two such quantities are known, $\mathbf{B}^2 - \mathbf{E}^2$ and $\mathbf{B} \cdot \mathbf{E}$. One can further show that

$$\begin{aligned} \mathbf{B} \cdot \mathbf{E} &= (\varepsilon_{ijk}\partial_j A_k)(-\partial_i A_0 - \partial_t A_i) \\ &= -\varepsilon_{ijk}\partial_i[A_0 \partial_j A_k] - \frac{1}{2}\varepsilon_{ijk}\partial_t[A_i \partial_j A_k] - \frac{1}{2}\varepsilon_{ijk}\partial_i[A_j \partial_t A_k] \\ &= -\boldsymbol{\nabla} \cdot (A_0 \mathbf{B}) - \frac{1}{2}\partial_t(\mathbf{A} \cdot \mathbf{B}) - \frac{1}{2}\boldsymbol{\nabla} \cdot [\mathbf{A} \times \partial_t \mathbf{A}] \end{aligned} \tag{2.5}$$

is a total derivative. In an infinite medium such surface terms will vanish, leaving only the $\mathbf{E}^2 - \mathbf{B}^2$ term as the Lorentz-invariant scalar. Up to an overall constant, the EM field action must be given by

$$S_f = k \int (\mathbf{E}^2 - \mathbf{B}^2) d^3\mathbf{r} dt. \tag{2.6}$$

Note the field action is given as the integral over the whole space-time with $k(\mathbf{E}^2 - \mathbf{B}^2)$ acting as the "density" of action.

In the previous chapter the action for the particle and its coupling to field was given as an integral only over time. This is so only because we assumed a single point charge interacting with the EM environment. For a collection of charges, or even a continuous distribution of charge densities, the matter as well as the matter-field actions have to be generalized to a space-time continuum form. For the derivation of Maxwell's equation all we need to know is how to generalize the matter-field coupling action S_{mf} to the continuum case. Recall that a single charge coupling to EM field has the action

$$S_{mf} = \int (q\mathbf{A} \cdot \mathbf{v} - qA_0)\, dt. \tag{2.7}$$

In the continuum this action will be generalized to

$$S_{mf} = \int \rho\, (\mathbf{A} \cdot \mathbf{v} - A_0)\, d^3\mathbf{r} dt = \int (\mathbf{j} \cdot \mathbf{A} - \rho A_0) d^3\mathbf{r} dt. \tag{2.8}$$

With the aid of four-current $j_\mu = (-\rho, \mathbf{j})$ one could also write the above as the space-time integral of $j_\mu A_\mu$. In any case, the combined action governing the dynamics of EM field and its coupling to matter current is given, in the continuum, by

$$S = k \int (\mathbf{E}^2 - \mathbf{B}^2) d\Omega + \int (\mathbf{j} \cdot \mathbf{A} - \rho A_0) d\Omega, \; d\Omega \equiv d^3\mathbf{r}dt. \tag{2.9}$$

The importance of gauge potentials A_μ is apparent in the action since it is to them that the matter four-current j_μ couples, not to the fields themselves.

2.3 Maxwell's equation: the second pair

Having introduced the field action in Eq. (2.9) and thus completed the action of the matter-field system, variational procedure will yield the remaining two equations of the Maxwell's equations. Since the scalar potential A_0 only enters in the electric field, $\mathbf{E} = -\boldsymbol{\nabla} A_0 - \partial_t \mathbf{A}$, it is easier to work out the variation of the action $S = S_{mf} + S_f$ with respect to A_0 first:

$$\delta S = \int (-2kE_i \partial_i [\delta A_0] - \rho \delta A_0) = \int (2k \boldsymbol{\nabla} \cdot \mathbf{E} - \rho) \, \delta A_0 = 0. \tag{2.10}$$

By taking $2k = \varepsilon_0$ we obtain the Coulomb's law:

$$\boldsymbol{\nabla} \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0}. \tag{2.11}$$

The unknown constant k is related to the empirical value of the strength of Coulomb force fixed in this way. The complete action reads

$$S = \frac{\varepsilon_0}{2} \int (\mathbf{E}^2 - \mathbf{B}^2) d\Omega + \int (\mathbf{j} \cdot \mathbf{A} - \rho A_0) d\Omega. \tag{2.12}$$

Variation of the action *w.r.t.* $\mathbf{A}$ yields the Ampere-Maxwell equation,

$$\boldsymbol{\nabla} \times \mathbf{B} - \frac{\partial \mathbf{E}}{\partial t} = \frac{\mathbf{j}}{\varepsilon_0}. \tag{2.13}$$

2.4 Conservation laws

In the theory of electromagnetism one cannot destroy or create charges. In the realm of relativistic quantum mechanics creation/annihilation of charges become possible, but only in such a manner that the total charge should always remain conserved. The statement of charge conservation is one of the most important cornerstones of physical concepts. With every conservation law there must follow a continuity equation. For example, the time variation of the local

charge density $\dot{\rho}$ must be accompanied by the corresponding flow of current as dictated by the continuity equation[1]:

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \mathbf{j} = 0. \tag{2.14}$$

When both sides are integrated over a space volume V the results are ($Q \equiv \int \rho dV$)

$$\frac{dQ}{dt} + \int (\boldsymbol{\nabla} \cdot \mathbf{j}) dV = \frac{dQ}{dt} + \int \mathbf{j} \cdot d\mathbf{S} = 0. \tag{2.15}$$

The quantity $\int \mathbf{j} \cdot d\mathbf{S}$ indicates the flux of current $\mathbf{j}$, the amount of charge leaving the volume over a unit time, and must be equal to minus the changes in the charge contained within the same volume.

At first it might seem surprising that a similar conservation law applies to the EM field itself. Take two of the Maxwell's equations and multiply,

$$\begin{aligned} \mathbf{B} \cdot \boldsymbol{\nabla} \times \mathbf{E} &= -\mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial t}, \\ \mathbf{E} \cdot \boldsymbol{\nabla} \times \mathbf{B} &= \frac{1}{\varepsilon_0} \mathbf{E} \cdot \mathbf{j} + \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t}, \end{aligned} \tag{2.16}$$

The difference of the two equations gives

$$\frac{1}{2} \frac{\partial}{\partial t} (\mathbf{B}^2 + \mathbf{E}^2) = \mathbf{E} \cdot \boldsymbol{\nabla} \times \mathbf{B} - \mathbf{B} \cdot \boldsymbol{\nabla} \times \mathbf{E} - \frac{1}{\varepsilon_0} \mathbf{j} \cdot \mathbf{E}. \tag{2.17}$$

Using vector analysis we can re-write $\mathbf{E} \cdot \boldsymbol{\nabla} \times \mathbf{B} - \mathbf{B} \cdot \boldsymbol{\nabla} \times \mathbf{E} = -\boldsymbol{\nabla} \cdot (\mathbf{E} \times \mathbf{B})$,

$$\frac{1}{2} \frac{\partial}{\partial t} (\mathbf{B}^2 + \mathbf{E}^2) + \boldsymbol{\nabla} \cdot (\mathbf{E} \times \mathbf{B}) = -\frac{1}{\varepsilon_0} \mathbf{j} \cdot \mathbf{E}. \tag{2.18}$$

When the current is absent, $\mathbf{j} = 0$, the electromagnetic fields obey a continuity equation of its own, with the roles of field density and the field current taken up by $\rho_f = (\mathbf{B}^2 + \mathbf{E}^2)/2$ and $\mathbf{j}_f = \mathbf{E} \times \mathbf{B}$, respectively. Several implications of this equation are:

- Electromagnetic field carries the associated energy density $\rho_f = \varepsilon_0(\mathbf{B}^2 + \mathbf{E}^2)/2$.
- EM field carries the associated energy flux $\mathbf{j}_f = \varepsilon_0(\mathbf{E} \times \mathbf{B})$. The energy flux vector is called the Poynting vector.
- EM field alone does not satisfy the energy conservation. The sum of the field energy density and the matter kinetic energy density does. It means that energy can be transferred between field and matter.

[1] Similar equation of continuity must exist for any conserved quantities in nature, not just the charges.

2.5 Conservation of momentum in electrodynamics

In the previous section we saw how the EM field carries an energy density and that the combined energy of matter and field are conserved through a conservation law. We also know, from relativity, that energy is but the zeroth component of a four-component energy-momentum tensor. So if there exists a conservation law of energy, there should also be a conservation of momentum. For Newtonian mechanics the momentum density $\mathbf{p} = \rho\mathbf{v}$ will have the time dependence proportional to the local force density $\mathbf{f}$: $\partial\mathbf{p}/\partial t \sim \mathbf{f}$. Before proceeding to write down a similar equation for the EM fields, we must first ask: what is the meaning (definition) of momentum for EM fields?

There is basically only one vector quantity one can construct out of $\mathbf{B}$ and $\mathbf{E}$ fields and we have already seen it in the previous section. It is the Poynting vector: $\mathbf{E} \times \mathbf{B}$. Assuming therefore $\mathbf{p}_f \propto \mathbf{E} \times \mathbf{B}$, the task at hand will be to write down the equation for $\partial_t \mathbf{p}_f$. That's exactly what we are about to do now. By repeatedly inserting Maxwell's equations one obtains

$$\begin{aligned} \frac{\partial}{\partial t}(\mathbf{E} \times \mathbf{B}) &= \left(\boldsymbol{\nabla} \times \mathbf{B} - \frac{\mathbf{j}}{\varepsilon_0}\right) \times \mathbf{B} + \mathbf{E} \times (-\boldsymbol{\nabla} \times \mathbf{E}) \\ &= (\mathbf{E} \cdot \boldsymbol{\nabla})\mathbf{E} + (\mathbf{B} \cdot \boldsymbol{\nabla})\mathbf{B} - \frac{1}{2}\boldsymbol{\nabla}(\mathbf{E}^2 + \mathbf{B}^2) - \frac{1}{\varepsilon_0}\mathbf{j} \times \mathbf{B}. \end{aligned} \tag{2.19}$$

Terms like $(\mathbf{E} \cdot \boldsymbol{\nabla})\mathbf{E}$ has the i-th component $E_j \partial_j E_i = \partial_j(E_i E_j) - E_i(\boldsymbol{\nabla} \cdot \mathbf{E})$. With this trick we can write the r.h.s. as

$$\begin{aligned} &\partial_j \left(E_i E_j + B_i B_j - \frac{1}{2}\delta_{ij}[\mathbf{E}^2 + \mathbf{B}^2]\right) - E_i(\boldsymbol{\nabla} \cdot \mathbf{E}) - \frac{1}{\varepsilon_0}(\mathbf{j} \times \mathbf{B})_i \\ &\quad = \frac{1}{\varepsilon_0}\partial_j T_{ij} - \frac{1}{\varepsilon_0}[\rho\mathbf{E} + \mathbf{j} \times \mathbf{B}]_i. \end{aligned} \tag{2.20}$$

In the above we have introduced the Maxwell stress tensor

$$T_{ij} = \varepsilon_0 \left(E_i E_j + B_i B_j - \frac{1}{2}\delta_{ij}[\mathbf{E}^2 + \mathbf{B}^2]\right) \tag{2.21}$$

and used the Coulomb's law $\boldsymbol{\nabla} \cdot \mathbf{E} = \rho/\varepsilon_0$. The quantity $\rho[\mathbf{E} + \mathbf{j} \times \mathbf{B}]$, the force density, must equal the rate of change of the momentum density of the matter: $d\mathbf{p}_m/dt$. With the field momentum density $\mathbf{p}_f = \varepsilon_0 \mathbf{E} \times \mathbf{B}$, we obtain the conservation law of the momentum of the combined matter-field system:

$$\frac{\partial}{\partial t}(\mathbf{p}_f + \mathbf{p}_m)_i = \partial_j T_{ij}. \tag{2.22}$$

2.6 Problems

1. Verify Eq. (2.5).

2. Suppose we modify the action of the EM field to include the $\mathbf{E}\cdot\mathbf{B}$ term:

$$S = \frac{1}{2}\int(\mathbf{E}^2 - \mathbf{B}^2)d\Omega + \int \alpha\mathbf{E}\cdot\mathbf{B}d\Omega + \int(\mathbf{j}\cdot\mathbf{A} - \rho A_0)d\Omega. \tag{2.23}$$

We have set $\varepsilon = 1$ to simplify the notation. Prove that the Maxwell's equations are modified as follows:

$$\begin{aligned} \boldsymbol{\nabla}\cdot\mathbf{E} &= \rho - \boldsymbol{\nabla}\alpha\cdot\mathbf{B}, \\ \boldsymbol{\nabla}\times\mathbf{B} &= \frac{\partial\mathbf{E}}{\partial t} + \mathbf{j} + (\dot{\alpha}\mathbf{B} + \boldsymbol{\nabla}\alpha\times\mathbf{E}). \end{aligned} \tag{2.24}$$

The other two relations, Faraday's law and $\boldsymbol{\nabla}\cdot\mathbf{B} = 0$, remain unmodified. Only the space- and time-dependent α can modify the Maxwell equation.

Using the modified Maxwell's equations, consider a magnetic monopole of unit charge surrounded by a thin shell (of some small thickness). Inside the shell we have $\alpha = 0$, and outside it equals some constant value, $\alpha = \theta$. Show that there is electric charge induced in the shell whose magnitude is $q = -\theta\Phi$. The magnetic flux generated by the monopole is denoted Φ. The field represented by α is called the "axion", and the presence of axion induces a magnetic monopole paired with an electric charge. The composite object is called a "dyon". [F. Wilczek, Phys. Rev. Lett. **58**, 1799 (1987).]

3. Verify that Eq. (2.13) follows from the variational principle.

4. Prove $\boldsymbol{\nabla}\cdot(\mathbf{E}\times\mathbf{B}) = \mathbf{B}\cdot(\boldsymbol{\nabla}\times\mathbf{E}) - \mathbf{E}\cdot(\boldsymbol{\nabla}\times\mathbf{B})$.

5. Prove that Schrödinger's equation

$$i\hbar\frac{\partial\psi}{\partial t} = \left(-\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2 + V(\mathbf{r})\right)\psi \tag{2.25}$$

gives rise to the continuity equation $\partial_t\rho + \boldsymbol{\nabla}\cdot\mathbf{j} = 0$ for the probability density $\rho = |\psi|^2$. Derive the expression of the current density $\mathbf{j}$.

6. Assume a quantum-mechanical particle described by the two-component wave function $\Psi = \begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}$ and obeying the equation

$$i\hbar\frac{\partial\Psi}{\partial t} = \left(-i\hbar v_0\boldsymbol{\sigma}\cdot\boldsymbol{\nabla} + V(\mathbf{r})\right)\Psi. \tag{2.26}$$

Here v_0 is a constant having the dimension of velocity, and $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the three Pauli matrices. A particle obeying such an equation is called a Weyl particle. Derive the conservation equation for the Weyl matter density $\rho = \Psi^\dagger\Psi = |\psi_1|^2 + |\psi_2|^2$.

7. A parallel-plate capacitor of circular shape (radius = a), with two plates separated by a distance h, is being charged slowly. The total electric field energy contained in the space between the plates is $U = (E^2/2)(\pi a^2 h)$. Show that the rate of change of the field energy equals to the energy flow through the sides of the plates governed by the Poynting vector $\mathbf{S} = \mathbf{E} \times \mathbf{B}$. Here $\mathbf{B}$ is induced by the time-varying electric field $\mathbf{E}$ inside the plates.

Chapter 3

Motion of Charged Particles

In this chapter we discuss the motion of a charged particle under a variety of electric and magnetic field configurations. Focus will be on development of methods to tackle non-uniform, but gradually varying EM fields in space and time and the derivation of associated, if any, invariants.

3.1 Uniform magnetic field

The motion of non-relativistic, charged particle is governed by Lorentz's equation of motion

$$\frac{d\mathbf{p}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \tag{3.1}$$

where $\mathbf{E}$ and $\mathbf{B}$ are arbitrary space- and time-dependent fields. In the case of zero electric field and uniform magnetic field (directed along the z-direction), the motion is that of uniform drift along the field direction with velocity $\mathbf{v}_\parallel$ superimposed on a circular motion in the xy-plane with velocity $\mathbf{v}_\perp$. The center of the circular orbit is commonly called the guiding center. The circular motion about it is called the cyclotron motion, whose radius will become the cyclotron radius. In general the combined motion becomes "helical". The kinetic energy $m\mathbf{v}^2/2 = m(\mathbf{v}_\parallel^2 + \mathbf{v}_\perp^2)/2$ is conserved during the motion, implying that the parallel and perpendicular speeds $v_\parallel$ and $v_\perp$ are separately conserved. The pitch angle α defined as

$$\tan\alpha = \frac{v_\perp}{v_\parallel} \tag{3.2}$$

is therefore a constant of motion for uniform magnetic field. In the reference frame that moves with the velocity $\mathbf{v}_\parallel$ we obtain the pitch angle $\pi/2$. Such a reference frame wherein the guiding center appears to remain stationary will be called the guiding center frame.

One can apply the concept of the guiding center frame even for the non-uniform magnetic field, provided the spatial variation of the field occurs on a length scale much larger than the cyclotron radius and the temporal variation is much slower than the the inverse of the cyclotron frequency, $\omega_c = qB/m$, equal to the Larmor frequency ω_L[1]. In this case one can still decompose the motion as the sum of the constant drift along the local field direction (forming the lines of force) plus a cyclotron motion about the local magnetic field axis. In this frame there will be a magnetic moment $\boldsymbol{\mu}$ associated with charge q given by

$$\boldsymbol{\mu} = \frac{1}{2} q \mathbf{r}_\perp \times \mathbf{v}_\perp = \frac{q}{2m} \mathbf{L}_\parallel. \tag{3.3}$$

The last equality relates the magnetic moment to the classical angular momentum $\mathbf{L}$. Both $\mathbf{r}_\perp$ and $\mathbf{v}_\perp$ are the components of the position and velocity vectors in the plane perpendicular to the local field direction. Recall that an electron ($q = -e$) executes a cyclotron motion according to the right-handed sense, therefore produces a magnetic moment in opposite direction to the field. For positive charges both the motion and the charge are reversed, giving the same orientation of the magnetic moment pointing opposite to the magnetic field. This situation of the opposite magnetic moment being induced by the external magnetic field is called diamagnetism.

The magnetic moment $\boldsymbol{\mu}$ subject to the magnetic field $\mathbf{B}$ contains the magnetic energy $E_m = -\boldsymbol{\mu} \cdot \mathbf{B}$. For the circular motion of the charge one can prove easily that

$$E_m = |\boldsymbol{\mu} \cdot \mathbf{B}| = \frac{1}{2} m \mathbf{v}^2, \tag{3.4}$$

where $\mathbf{v}$ is the velocity of the circular motion. Now imagine a particular entering into a space filled with magnetic field region with constant velocity $\mathbf{v}$. Since magnetic field cannot do any work on the particle, all of the kinetic energy of the linear motion must be converted to the magnetic energy E_m: $K \rightarrow E_m$. The reason why the kinetic energy due to the translational motion appears to be absent in the magnetic field case is that the guiding center of the circular motion itself remains stationary. There is no motion of the center-of-mass (CM) itself. Then one can equally well regard the kinetic energy as the magnetic energy due to the coupling of magnetic moment (generated by the circular motion) to the external field.

3.2 Uniform electric and magnetic fields

In the previous section we saw that the charged particle under the constant magnetic field becomes "trapped". It will be an interesting engineering question to ask if there is a way to get the CM moving again, even with the magnetic field on. Obvious choice is to apply the electric field to accelerate the charge in the desired direction. Another situation of common interest before moving on

[1] Larmor frequency refers to the frequency of rotation of spin under the constant magnetic field. The two frequencies, Larmor and cyclotron frequencies, coincide.

to the case of inhomogeneous fields is one of uniform, orthogonal electric and magnetic fields, $\mathbf{E} \cdot \mathbf{B} = 0$:

$$\frac{d\mathbf{p}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \tag{3.5}$$

Recall that we were able to employ the Lorentz transformation to the frame with vanishing $\mathbf{B}$ field to explain the Hall effect phenomena. In a frame moving with velocity $\mathbf{V}$ the fields transform as

$$\mathbf{E}' = \mathbf{E} - \mathbf{B} \times \mathbf{V}, \quad \mathbf{B}' = \mathbf{B} + \mathbf{E} \times \mathbf{V}. \tag{3.6}$$

This time we choose the frame's velocity $\mathbf{V}$ to cancel the electric field, $\mathbf{V} = (\mathbf{E} \times \mathbf{B})/B^2$. The magnetic field is reduced to $\mathbf{B}' = (1 - E^2/B^2)\mathbf{B}$. One sees that this kind of cancelation of the electric field is possible only when we had $E^2/B^2 < 1$ in the original frame.[2] In such a moving frame with vanishing electric field, the charged particle should execute a spiral motion as described earlier. Back in the original frame, such a helical motion will be superimposed on top of the drift motion with velocity $\mathbf{V} = (\mathbf{E} \times \mathbf{B})/B^2$. Although being acted on with a constant electric field, the average kinetic energy of the charged particle will not increase over time. Roughly put, the particle loses the energy during the half-cycle wherein the motion is going opposite to the $\mathbf{E}$-field direction. The drift velocity $\mathbf{V}$ is independent of the sign of the charge or its mass. All particles subjected to crossed $\mathbf{E}$ and $\mathbf{B}$ fields will drift in the same direction with the same velocity.

3.3 Adiabatic invariant

In this section we examine the question of what happens to the charged particle dynamics when the magnetic field is slowly varying, either in space or in time. The notion of "adiabatic invariant" will emerge naturally in the discussion, referring to the conservation of the flux enclosed by the cyclotron orbit despite the changes in magnetic field over time and/or space. Specifically we first consider uniform, but time-dependent magnetic field pointed along the $\hat{z}$-direction, focusing on the instance where the time rate of change is much slower than the cyclotron frequency. To simplify the discussion somewhat we only consider the case of two-dimensional cyclotron motion with $\mathbf{v}_{\parallel} = 0$. Lorentz's equation of motion becomes

$$\frac{d\mathbf{v}_{\perp}}{dt} = \frac{q}{m}(\mathbf{E} + \mathbf{v}_{\perp} \times \mathbf{B}). \tag{3.7}$$

Time dependence of the $\mathbf{B}$-field induces electric field by Faraday's law: $\boldsymbol{\nabla} \times \mathbf{E} = -\partial \mathbf{B}/\partial t$. Taking $\mathbf{B} = B_z \hat{z}$ we arrive at $\boldsymbol{\nabla} \times \mathbf{E} = -(dB_z/dt)\hat{z}$, which is solved, approximately, by

[2] In general the Lorentz transformation can be employed to cancel the weaker of the two fields $\mathbf{E}$ and $\mathbf{B}$ in the original frame.

$$\mathbf{E} = -\frac{1}{2}\left(\frac{dB_z}{dt}\right)(x\hat{y} - y\hat{x}) = -\frac{1}{2}\left(\frac{dB_z}{dt}\right) r\hat{\theta}. \tag{3.8}$$

The tangential unit vector $\hat{\theta}$ is introduced. Substituted into the Lorentz's equation, the tangential electric field provides acceleration or deceleration along the direction of the motion, causing increase or decrease of the radius according to $v_c = \omega_c r_c$. The change in the cyclotron radius r_c is just enough to compensate for the change in the magnetic field B_z, so that the flux $\Phi \equiv r_c^2 B_z$ is maintained a constant. The proof goes as follows.

Because the field is changing slowly, we may regard that a constant magnitude of electric field acts on the particle over the course of one cyclotron motion. One can approximately write the electric field as $\mathbf{E} = -(1/2)(dB_z/dt) r_c \hat{\theta}$,

$$\frac{d\mathbf{v}}{dt} = \frac{q}{m}\left(-\frac{1}{2}\left(\frac{dB_z}{dt}\right) r_c\hat{\theta} + B_z \mathbf{v} \times \hat{z}\right). \tag{3.9}$$

Substituting $\mathbf{v} = -v\hat{\theta}$ into the above, we find $d\hat{\theta}/dt = \omega_c \hat{r}$, and

$$-\frac{dv}{dt}\hat{\theta} - v\omega_c \hat{r} = \frac{q}{m}\left(-\frac{1}{2}\left(\frac{dB_z}{dt}\right) r_c\hat{\theta} + B_z \mathbf{v} \times \hat{z}\right). \tag{3.10}$$

Matching the radial and tangential components separately, one finds

$$\frac{dv}{dt} = \frac{1}{2}\frac{q}{m}\frac{dB_z}{dt} r_c = \frac{q}{m}\frac{d}{dt}(B_z r_c). \tag{3.11}$$

It follows that

$$\frac{1}{2}\frac{dB_z}{dt} r_c + B_z \frac{dr_c}{dt} = 0 \rightarrow \frac{d}{dt}(B_z r_c^2) = 0. \tag{3.12}$$

The time-dependent magnetic field $B_z(t)$ and cyclotron radius $r_c(t)$ cooperate in such a way as to maintain the same flux through its orbit, $\Phi = r_c^2(t) B_z(t)$.

Next, one can consider a situation where the magnetic field strength increases along the z-direction: $dB_z/dz > 0$. Due to the general theorem $\boldsymbol{\nabla} \cdot \mathbf{B} = 0$ this increase must be accompanied by the change of the field in the perpendicular direction, $\boldsymbol{\nabla}_\perp \cdot \mathbf{B}_\perp = -dB_z/dz < 0$. It creates an overall shape of the flux line that is shrinking in radius as z increases. The equation of motion becomes

$$\begin{aligned} \frac{d\mathbf{v}_\perp}{dt} &= \frac{q}{m}\left(\mathbf{v}_\parallel \times \mathbf{B}_\perp + \mathbf{v}_\perp \times \mathbf{B}_\parallel\right), \\ \frac{d\mathbf{v}_\parallel}{dt} &= \frac{q}{m}\mathbf{v}_\perp \times \mathbf{B}_\perp. \end{aligned} \tag{3.13}$$

We take the inner product of the first line with $\mathbf{v}_\perp$,

$$\frac{1}{2}\frac{d\mathbf{v}_\perp^2}{dt} = \frac{q}{m}\mathbf{v}_\perp \cdot \mathbf{v}_\parallel \times \mathbf{B}_\perp. \tag{3.14}$$

It is further assumed that motion of the guiding center along the z-direction is much slower than the cyclotron motion. At every z, we can treat $\mathbf{B}_\perp$ as the radial field given by $\mathbf{B}_\perp = -(dB_z/dz)(x\hat{x}+y\hat{y})/2$ sufficiently close to the guiding center position. We further assume that the cyclotron motion has been executed more than once before the z-coordinate of the guiding center has drifted a sufficient amount. The three vectors on the r.h.s. of the above equation are all mutually orthogonal, $dv_\perp/dt = (q/m)v_\parallel B_\perp = v_\parallel dv_\perp/dz$. It follows that

$$\frac{dv_\perp}{dz} = -\frac{q}{m}B_\perp. \tag{3.15}$$

The strength of the $B_\perp$ field on the r.h.s. is to be evaluated at the cyclotron radius, $r_c = mv_\perp/qB_\parallel$, also related to the z-dependent perpendicular velocity to the cyclotron radius as $v_\perp = (q/m)B_z r_c$:

$$\frac{d}{dz}(B_z r_c) = \frac{dB_z}{dz}\frac{r_c}{2} \rightarrow \frac{d}{dz}(B_z r_c^2) = 0. \tag{3.16}$$

What the last formula shows is that the flux through the cyclotron orbit, $\Phi(z) = B_z r_c^2$, remains an invariant irrespective of the z-coordinate. Such quantity that remains constant in the presence of a slow variation of the physical condition is called the adiabatic invariant. In this section we derived the adiabatic invariant without resorting to the general theory of it.

Information can be squeezed about the velocity along the z-direction from the second equation in Eq. (3.13). Under the adiabatic approximation the two vectors on the r.h.s. are orthogonal to each other and to $\mathbf{v}_\parallel$. Hence we arrive at a scalar equation

$$\frac{dv_\parallel}{dt} = \frac{q}{m}v_\perp B_\perp = \frac{q}{m}\cdot\frac{q}{m}B_z r_c \cdot \frac{dB_z}{dz}\frac{r_c}{2} = \frac{1}{2}\left(\frac{q}{m}\right)^2 (B_z r_c^2)\frac{dB_z}{dz}. \tag{3.17}$$

The r.h.s. is proportional to dB_z/dz. The l.h.s. becomes $v_\parallel dv_\parallel/dz = (1/2)dv_\parallel^2/dz$. The proportionality relation becomes $dv_\parallel^2/dz \propto dB_z/dz$.

3.4 Motion under non-uniform magnetic field

In the most general case of space- and time-dependent magnetic field the equation of motion of a charge becomes

$$\frac{d\mathbf{v}}{dt} = \mathbf{v}\times\left(\frac{q}{m}\mathbf{B}(\mathbf{r},t)\right) = \mathbf{v}\times\boldsymbol{\omega}(\mathbf{r},t). \tag{3.18}$$

At first the problem appears quite similar to that of a rotating rigid body, $d\mathbf{r}/dt = \mathbf{r}\times\boldsymbol{\omega}$. In fact, the problem we are facing is more challenging because $\boldsymbol{\omega}(t)$ in the rigid-body rotation will only depend on time t, whereas the effective angular velocity vector $\boldsymbol{\omega}(\mathbf{r}(t),t)$ will depend on the position of the particle itself, as well as on time, and the position $\mathbf{r}(t)$ is the quantity we are trying to solve! It is like saying we need to know the answer for $\mathbf{r}(t)$ first, in order to solve the equation for $\mathbf{r}(t)$...

In this section we restrict ourselves to the case of time-independent magnetic field whose variation over space is slow compared to the typical radius of cyclotron motion. The slowness condition has a mathematical expression that $r_c|\boldsymbol{\partial}_\alpha \mathbf{B}|/|\mathbf{B}|$ ($\alpha = x, y, z$) is much less than unity, r_c being the cyclotron radius. Alternatively, we assume that there is a well-defined guiding center $\mathbf{r}_g$ at any particular instant, and the motion of the guiding center coordinate is only visible after several cyclotron motions have been executed. Due to variations of the magnetic field, the charged particle senses a different field strength in the course of its orbit. To a good approximation, variations in the **B**-field is small enough to be treated as a Taylor expansion:

$$\mathbf{B}(\mathbf{r}) \simeq \mathbf{B}(\mathbf{r}_g) + [(\mathbf{r} - \mathbf{r}_g) \cdot \boldsymbol{\nabla}]\mathbf{B}(\mathbf{r}_g). \tag{3.19}$$

It will be convenient to separate the "big" and the "small" fields as $\mathbf{B}_0 = \mathbf{B}(\mathbf{r}_g)$ and $\mathbf{B}_1 = [(\mathbf{r}-\mathbf{r}_g)\cdot\boldsymbol{\nabla}]\mathbf{B}(\mathbf{r}_g)$, respectively. The big field leads to "fast" cyclotron motion whereas the small field will lead to slow changes in it. One can make a corresponding dissociation of the scales in the velocity vector as $\mathbf{v} = \mathbf{v}_0 + \mathbf{v}_1$. The equation of motion can be matched at the lowest order as

$$\frac{d\mathbf{v}_0}{dt} = \frac{q}{m}\mathbf{v}_0 \times \mathbf{B}_0, \tag{3.20}$$

which gives rise to the cyclotron motion with the solution

$$\begin{aligned} \mathbf{r}(t) - \mathbf{r}_g &= r_c(\hat{x}\cos\omega_c t + \hat{y}\sin\omega_c t) \equiv (X(t), Y(t)), \\ \mathbf{v}_0(t) &= r_c\omega_c(-\hat{x}\sin\omega_c t + \hat{y}\cos\omega_c t). \end{aligned} \tag{3.21}$$

Here r_c is the cyclotron radius, $\omega_c = |q||\mathbf{B}(\mathbf{r}_g)|/m$ is the cyclotron frequency, and the sign of q was chosen to be negative, $q < 0$. The three axes $\hat{x}, \hat{y}, \mathbf{B}(\mathbf{r}_g)$ form the right-handed local orthogonal frame.

At the next order the equation becomes

$$\frac{d\mathbf{v}_1}{dt} = \frac{q}{m}\left(\mathbf{v}_0 \times \mathbf{B}_1 + \mathbf{v}_1 \times \mathbf{B}_0\right). \tag{3.22}$$

The magnetic field at the guiding-center location will be chosen to point in the z-direction, $\mathbf{B}_0 = B_0(\mathbf{r}_g)\hat{z}$. Insertion of the zeroth-order solutions found in Eq. (3.21) gives, after some manipulation,

$$\mathbf{v}_0 \times \mathbf{B}_1 = r_c\omega_c(\cos\omega_c t, \sin\omega_c t, 0)[X(t)\partial_x + Y(t)\partial_y]B_0. \tag{3.23}$$

The r.h.s. contains terms that oscillates at the frequency $2\omega_c$, and terms that are constant in time. When both terms are fed back into Eq. (3.22), the solution will also involve oscillating and constant components. The quantities of interest to us is the constant one, which may be obtained as the average of $\mathbf{v}_1$ over one cyclotron period. Implicitly taking the time average renders the simplified equation of motion

$$\frac{d\langle \mathbf{v}_1 \rangle}{dt} = \frac{q}{m} \left(\langle \mathbf{v}_0 \times \mathbf{B}_1 \rangle + \langle \mathbf{v}_1 \rangle \times \mathbf{B}_0 \right) = \frac{q}{m} \left(\frac{1}{2} r_c^2 \omega_c \boldsymbol{\nabla} B_0 + \langle \mathbf{v}_1 \rangle \times \mathbf{B}_0 \right). \quad (3.24)$$

As promised earlier, we are looking for a solution that is time-dependent, or $d\langle \mathbf{v}_1 \rangle / dt = 0$. The answer is found in the vector form as

$$\langle \mathbf{v}_1 \rangle = \frac{r_c^2 \omega_c}{2} \frac{\hat{z} \times \boldsymbol{\nabla} B_0(\mathbf{r}_g)}{B_0(\mathbf{r}_g)}. \quad (3.25)$$

To get an idea of both its magnitude and direction, divide both sides by the cyclotron speed $v_c = r_c \omega_c$, which is the speed of the charged particle in the cyclotron motion.

$$\frac{\langle \mathbf{v}_1 \rangle}{v_c} = \frac{r_c}{2} \frac{\hat{z} \times \boldsymbol{\nabla} B_0(\mathbf{r}_g)}{B_0(\mathbf{r}_g)}. \quad (3.26)$$

The speed of the drift motion relative to the cyclotron speed is in precise proportion to the magnetic field difference measured from the guiding center to the edge of the cyclotron motion, divided by the field strength at the guiding center.

Had we chosen a different sign of charge $q > 0$ we would have obtained the same form with the opposite overall sign. The gradient of the inhomogeneous magnetic field thus generates drift motion along the direction of $\boldsymbol{\nabla}|B|$ which points in opposite directions for the opposite sign of charges. The positive and negative charges will thus drift apart ways under such non-uniform field.

3.5 Problems

1. Show that the motion under the constant in-plane **E**-field and perpendicular **B**-field is given as the superposition of constant drift **V** and a circular motion, by solving the Lorentz's equation

$$\frac{d\mathbf{v}}{dt} = \frac{q}{m} \left(\mathbf{E} + \mathbf{v} \times \mathbf{B} \right). \quad (3.27)$$

2. Solve for the charged particle's trajectory when $\mathbf{B} \parallel \mathbf{E}$. Both fields are constant. Refer to Jackson, Problem 12.6 for hints.

Chapter 4

Electromagnetic Waves

Chapter two was concerned with derivation of fundamental laws of electromagnetism from the action principle and the examination of some conservation laws. In this chapter, we will derive solutions of the Maxwell's equation in the absence of matter field. The result is the well-known electromagnetic wave.

4.1 Wave equation in free space

In vacuum the Maxwell's equations simplify to

$$\begin{aligned} \boldsymbol{\nabla}\times\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t}, \qquad & \boldsymbol{\nabla}\times\mathbf{B} = \frac{\partial\mathbf{E}}{\partial t}, \\ \boldsymbol{\nabla}\cdot\mathbf{E} = 0, \qquad & \boldsymbol{\nabla}\cdot\mathbf{B} = 0. \end{aligned} \tag{4.1}$$

From undergraduate courses we are familiar with the derivation of the wave function starting from the first line of the above and applying some partial differentiation. Let us now try to derive the same wave equation using the potential field. Inserting $\mathbf{E} = -\boldsymbol{\nabla}A_0 - \partial_t\mathbf{A}$ and $\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}$ into the Maxwell's equations yields (temporarily restoring the matter fields)

$$\begin{aligned} \left(\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2\right)\mathbf{A} + \boldsymbol{\nabla}\left(\boldsymbol{\nabla}\cdot\mathbf{A} + \frac{\partial A_0}{\partial t}\right) = \frac{1}{\varepsilon_0}\mathbf{j}, \\ \boldsymbol{\nabla}^2 A_0 + \frac{\partial}{\partial t}(\boldsymbol{\nabla}\cdot\mathbf{A}) = -\frac{\rho}{\varepsilon_0}. \end{aligned} \tag{4.2}$$

They look quite complicated, until we realize that one has the freedom to choose a gauge in which certain terms in the above equations can be made to disappear. One such convenient gauge is the Lorenz gauge[1],

$$\boldsymbol{\nabla}\cdot\mathbf{A} + \frac{\partial A_0}{\partial t} = 0. \quad \text{(Lorenz)} \tag{4.3}$$

[1]This Lorenz is not the same as the person who wrote down the equation of motion for charged particles. Lorenz is a contemporary of Maxwell and lived much earlier than Lorentz. He is also Danish, not Dutch like Lorentz.

In such a gauge the scalar and vector potentials will also satisfy the wave equation (with a source term):

$$\left(\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2\right)\mathbf{A} = \frac{1}{\varepsilon_0}\mathbf{j},$$
$$\left(\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2\right)A_0 = \frac{\rho}{\varepsilon_0}. \tag{4.4}$$

In free space each equation obeys the wave equation and Lorentz invariance is known to hold for it. Such might be one of the sources of the wrongful association of Lorentz's name with this particular gauge.

In another gauge, called the Coulomb gauge, we choose

$$\boldsymbol{\nabla}\cdot\mathbf{A} = 0 \quad \text{(Coulomb)}. \tag{4.5}$$

The equations reduce to

$$\left(\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2\right)\mathbf{A} + \boldsymbol{\nabla}\left(\frac{\partial A_0}{\partial t}\right) = \frac{1}{\varepsilon_0}\mathbf{j},$$
$$\boldsymbol{\nabla}^2 A_0 = -\frac{\rho}{\varepsilon_0}. \tag{4.6}$$

The reason for the nomenclature is that the second equation looks exactly like the usual Coulomb's law expressed in terms of the scalar potential A_0, even though our consideration extends beyond the static charge distribution. Coulomb gauge is also known as the transverse, or radiation gauge, due to the fact that $\boldsymbol{\nabla}\cdot\mathbf{A} = 0$ will impose $\mathbf{k}\cdot\mathbf{A} = 0$ in the momentum coordinates. Irrespective of the gauge choice, the final result for $\mathbf{E}(\mathbf{r},t)$ and $\mathbf{B}(\mathbf{r},t)$ must turn out to be equal for the same choice of $(\rho(\mathbf{r},t),\mathbf{j}(\mathbf{r},t))$.

4.2 Plane waves

Solutions of Maxwell's equations in vacuum are obtained by taking $\rho = \mathbf{j} = 0$. Choosing the plane-wave forms for both scalar and vector potentials as

$$\mathbf{A} = \mathbf{A}_0 e^{i\mathbf{k}\cdot\mathbf{r}-i\omega t}, A_0 = \mathcal{A}_0 e^{i\mathbf{k}\cdot\mathbf{r}-i\omega t}, \tag{4.7}$$

inserting them into the Maxwell's equation yields the relations

$$\left(-\omega^2 + k^2\right)\mathbf{A}_0 + \mathbf{k}\left(-\mathbf{k}\cdot\mathbf{A}_0 + \omega\mathcal{A}_0\right) = 0,$$
$$-k^2\mathcal{A}_0 + \omega(\mathbf{k}\cdot\mathbf{A}_0) = 0. \tag{4.8}$$

Anticipating the dispersion relation in vacuum $\omega^2 = k^2$ one can simplify the first equation as $\omega\mathcal{A}_0 = \mathbf{k}\cdot\mathbf{A}_0$, identical to the second equation. At this point we can utilize the gauge degree of freedom to simply fix the gauge such that $\mathcal{A}_0 = 0$. In that same gauge one would also have $\mathbf{k}\cdot\mathbf{A}_0 = 0$, the condition that

requires the vector potential to be orthogonal, or transverse, to the propagation direction $\mathbf{k}$. Electric and magnetic fields are obtained as

$$\mathbf{E} = i\omega\mathbf{A}, \quad \mathbf{B} = i\mathbf{k}\times\mathbf{A}. \tag{4.9}$$

Apart from the same space-time factor $e^{i\mathbf{k}\cdot\mathbf{r}-i\omega t}$ one can also write the the constants of electric and magnetic fields as

$$\mathbf{E}_0 = i\omega\mathbf{A}_0, \ \mathbf{B}_0 = i\mathbf{k}\times\mathbf{A}_0. \tag{4.10}$$

The two vectors, $\mathbf{E}_0$ and $\mathbf{B}_0$, are orthogonal to each other.

Earlier we defined the Poynting vector $\mathbf{S} = \mathbf{E}\times\mathbf{B}$. In the complex notation just adopted the two fields are averages of $\mathbf{E}$ and its complex conjugate $\mathbf{E}^*$ (or $\mathbf{B}$ and $\mathbf{B}^*$), or the Poyting vector is written as

$$\begin{aligned}
\mathbf{S} &= \frac{1}{4}(i\omega)(\mathbf{A}-\mathbf{A}^*)\times[i\mathbf{k}\times(\mathbf{A}-\mathbf{A}^*)] \\
&= -\frac{1}{4}\omega\mathbf{k}(\mathbf{A}-\mathbf{A}^*)\cdot(\mathbf{A}-\mathbf{A}^*) \\
&= -\frac{1}{4}\omega\mathbf{k}[\mathbf{A}\cdot\mathbf{A}+\mathbf{A}^*\cdot\mathbf{A}^*]+\frac{1}{2}\omega\mathbf{k}(\mathbf{A}\cdot\mathbf{A}^*).
\end{aligned} \tag{4.11}$$

In the last line, the first two terms contain fast-oscillating components proportional to $e^{2i\mathbf{k}\cdot\mathbf{r}-2i\omega t}$. Given that typical frequency of lights is extremely high, one would hardly have a reason to be interested in the behavior of light on time scales comparable to one cycle. What is of more practical importance is the averaged quantity, obtained as the temporal average over many cycles. Averages of $\mathbf{A}\cdot\mathbf{A}$ and its complex conjugate over many cycles are zero. Hence, the averaged Poynting vector is obtained as

$$\langle\mathbf{S}\rangle = \frac{1}{2}\omega\mathbf{k}\langle\mathbf{A}\cdot\mathbf{A}^*\rangle = \frac{1}{2}\omega\mathbf{k}\mathbf{A}_0\cdot\mathbf{A}_0^* = \frac{1}{2}\omega^2(\mathbf{A}_0\cdot\mathbf{A}_0^*)\hat{k}. \tag{4.12}$$

In the last expression we have used the relation $\mathbf{k} = |\mathbf{k}|\hat{k} = \omega\hat{k}$. Similar averaging applies to electric and magnetic field energies.

$$\begin{aligned}
\mathcal{U}_E &= \frac{1}{4}\langle(\mathbf{E}+\mathbf{E}^*)^2\rangle = \frac{1}{2}\mathbf{E}\cdot\mathbf{E}^* = \frac{1}{2}\omega^2\mathbf{A}_0\cdot\mathbf{A}_0^* \\
\mathcal{U}_M &= \frac{1}{4}\langle(\mathbf{B}+\mathbf{B}^*)^2\rangle = \frac{1}{2}\mathbf{B}\cdot\mathbf{B}^* = \frac{1}{2}\omega^2\mathbf{A}_0\cdot\mathbf{A}_0^*.
\end{aligned} \tag{4.13}$$

There is equal amount of average energies stored in the form of electric and the magnetic fields. The Poyting vector is related to the field energy density by the relation

$$\langle\mathbf{S}\rangle = \frac{1}{2}(\mathcal{U}_E+\mathcal{U}_M)\hat{k}. \tag{4.14}$$

Spinor wave function

A brief digression to spin-1/2 spinor wave function is in order. The electron whose spin is pointing "up" is denoted as $\begin{pmatrix}1\\0\end{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}-i\omega t}$. The first, spinor part refers to the spin orientation while the plane-wave, orbital-part gives out the orbital content of the particle's state. Similarly the spin-down electron is given the wave function assignment $\begin{pmatrix}0\\1\end{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}-i\omega t}$. For spin state whose orientation is along an arbitrary direction $\mathbf{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, the corresponding spinor wave function is given by

$$\Psi_{\mathbf{k},\mathbf{n}} = \begin{pmatrix}\cos[\theta/2]\\ e^{i\phi}\sin[\theta/2]\end{pmatrix} e^{i\mathbf{k}\cdot\mathbf{r}-i\omega t} \equiv e^{i\mathbf{k}\cdot\mathbf{r}-i\omega t}|\mathbf{n}\rangle. \tag{4.15}$$

The general spinor wave function is a linear superposition of the up- and down-spin wave functions, with coefficients that tell us which way the spin is pointing on average:

$$|\mathbf{n}\rangle = \cos\left[\frac{\theta}{2}\right]|\uparrow\rangle + e^{i\phi}\sin\left[\frac{\theta}{2}\right]|\downarrow\rangle. \tag{4.16}$$

Polarization vector

One can show that the polarization vector $\mathbf{A}_0$ has a remarkable resemblance to the quantum-mechanical spin-1/2 spinor wave function just discussed. The transverse constraint $\mathbf{k}\cdot\mathbf{A} = 0$ dictates that we choose the vector $\mathbf{A}_0$ to lie orthogonal to the direction of propagation, $\mathbf{k}$. A pair of real, orthogonal unit vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ can be introduced with $\mathbf{e}_1\times\mathbf{e}_2 = \hat{k}$, to span such a polarization vector

$$\mathbf{A}_0 = |\mathbf{A}_0|(z_1\mathbf{e}_1 + z_2\mathbf{e}_2). \tag{4.17}$$

The pair of complex coefficients (z_1, z_2) are constrained by the condition that the square of the r.h.s. must equal $|\mathbf{A}_0|^2$, or

$$|z_1|^2 + |z_2|^2 = 1. \tag{4.18}$$

Any such pair of complex numbers can be parameterized as

$$z_1 = e^{i\phi_1}\cos(\theta/2),\ z_2 = e^{i\phi_2}\sin(\theta/2). \tag{4.19}$$

Taking out the phase factor of z_1 in front (and re-labeling it as $e^{i\varphi}$) and re-defining $\phi_2 - \phi_1$ as ϕ, we arrive at the parametrization of the arbitrary polarization vector $\mathbf{A}_0$ as

$$\mathbf{A}_0 = |\mathbf{A}_0|e^{i\varphi}\left[\cos\frac{\theta}{2}\mathbf{e}_1 + e^{i\phi}\sin\frac{\theta}{2}\mathbf{e}_2\right]. \tag{4.20}$$

Note the exceptional similarity of this expression to the spin-1/2 spinor wave function given in Eq. (4.16). Cases where $|\cos\theta/2| \neq |\sin\theta/2|$, that is when $\theta \neq \pi/2$, are known as elliptic polarization. At $\theta = \pi/2$ one has the circular polarization. In particular $e^{i\phi} = \pm i$ are known as right(left)-circularly polarized. For $\phi = 0$ the polarization vector is a real vector $\mathbf{e}_\theta = \cos(\theta/2)\mathbf{e}_1 + \sin(\theta/2)\mathbf{e}_2$ defining the axis of linear polarization. One can see there is one point on the unit sphere (called the Bloch sphere) corresponding to each polarization vector $\mathbf{A}_0$. The unit vector $\mathbf{n}$ has the polar angle θ and the azimuthal angle ϕ and may be obtained from the spinor $\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ through the matrix operation

$$\mathbf{n} = \mathbf{z}^\dagger \boldsymbol{\sigma} \mathbf{z} \tag{4.21}$$

with respect to the Pauli matrices $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$.

4.3 Mixing of polarizations and the use of density matrix

It is not easy to produce light wave consisting solely of a single wavevector or a single frequency. In general what one has is a superposition of plane waves over a reasonably small range of momentum and frequency that forms a wavepacket:

$$\mathbf{A}(\mathbf{r}, t) = \sum_{\mathbf{k}} \mathbf{A}_0(\mathbf{k}) e^{i(\mathbf{k}\cdot\mathbf{r} - \omega_\mathbf{k} t)}, \ \omega_\mathbf{k} = |\mathbf{k}|. \tag{4.22}$$

Light is characterized by more than its direction of propagation and the wavelength. At each $\mathbf{k}$ the polarization vector $\mathbf{A}_0(\mathbf{k})$ satisfies the transversality condition: $\mathbf{A}_0(\mathbf{k}) \cdot \mathbf{k} = 0$. The two basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ defined in the previous section span the possible orientations of the polarization vector for a given $\mathbf{k}$. There are circumstances in which lights of (nearly) same $\mathbf{k}$, but with different polarizations, will mix in the same spatial location. For example one light source produces RCP (right circularly polarized) light while another source produces LCP lights only. The two lights propagation in nearly identical directions and merge at some spatial position. When the two lights meet, certain fraction of light quanta (or photons) will have RCP character, the remaining fraction LCP character. However, it would be wrong to describe such a mixture of lights as a superposition of polarization vectors like

$$\mathbf{A} = \mathbf{A}_R + \mathbf{A}_L. \tag{4.23}$$

The addition of two polarization vectors will merely produce another polarization vector $\mathbf{A}$, corresponding to another, completely different point on the Bloch sphere. This is not what happens when one mixes, say, 30% of LCP and 70% of RCP lights in one space. The mixture is a classical one, no different than when we mix 30% of beans and 70% of rice in a cooking pot. The appropriate language of the mixed state is the density matrix.

First construct the density matrix of a single polarization vector (the pure state). For pure state whose polarization vector is $|\mathbf{A}_0|e^{i\phi}(z_1\mathbf{e}_1 + z_2\mathbf{e}_2) = \psi_1\mathbf{e}_1 + \psi_2\mathbf{e}_2$, we may adopt a quantum-mechanical expression

$$|P\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \tag{4.24}$$

and the associated density matrix would be

$$\hat{\rho}(P) = \frac{|P\rangle\langle P|}{\langle P|P\rangle} = \frac{1}{|\psi_1|^2 + |\psi_2|^2} \begin{pmatrix} \psi_1^*\psi_1 & \psi_2^*\psi_1 \\ \psi_1^*\psi_2 & \psi_2^*\psi_2 \end{pmatrix}. \tag{4.25}$$

Actually density matrix is a useful device to extract information about the content of a particular polarization orientation in a given polarization vector. Out of $P\rangle$, say, one wishes to know how much content of $|1\rangle$ polarization (directed along the $\mathbf{e}_1$ direction) there is. The answer, intuitively, is

$$1-\text{fraction} = \frac{|\psi_1|^2}{|\psi_1|^2 + |\psi_2|^2}. \tag{4.26}$$

If one asked how much of the R-fraction there is in a given $|P\rangle$, one can find the answer by first re-writing the given state vector $|P\rangle$ in the new basis, $\mathbf{e}_R$ and $\mathbf{e}_L$, extract the new components ψ_R and ψ_L, and provide the answer as

$$\text{R}-\text{fraction} = \frac{|\psi_R|^2}{|\psi_R|^2 + |\psi_L|^2}. \tag{4.27}$$

A short algebra shows $\psi_R = (\psi_1 + i\psi_2)/\sqrt{2}$, while $\psi_L = (\psi_1 - i\psi_2)/\sqrt{2}$. It can be verified readily that the same answer is obtained by taking the average of the density matrix $\hat{\rho}$ given in Eq. (4.25) with respect to the $|R\rangle$ itself. In the linear basis the $|R\rangle$ state has the wave function $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}$. Carrying out the necessary matrix multiplication, $\langle R|\hat{\rho}(P)|R\rangle$ yields the same answer as Eq. (4.27). The great utility of the density matrix is that it is trivial, a matter of automated multiplication, to work out the fraction of any given polarization.

Suppose we have two sources of monochromatic lights, and each light has different polarizations. The nature of the combined beam is the statistical average of the individual light. The average property of the mixture of lights is described by

$$\hat{\rho} = \sum_{\alpha=1,2} w(P_\alpha)\hat{\rho}(P_\alpha). \tag{4.28}$$

The probability function $w(P_\alpha)$ adds up to unity, $\sum_\alpha w(P_\alpha) = 1$. Suppose we wish to identify how much "right-polarization" there is in a given mixture of light. The answer can be found by taking the average of the density matrix $\hat{\rho}$, corresponding to the mixture, with the basis vector $|R\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ on either side

$$\langle R|\hat{\rho}|R\rangle = \sum_\alpha w(P_\alpha)\langle R|\hat{\rho}(P_\alpha)|R\rangle. \tag{4.29}$$

4.4 Problems

1. Show that the spinor wave function $|\mathbf{n}\rangle = \begin{pmatrix} \cos[\theta/2] \\ e^{i\phi}\sin[\theta/2] \end{pmatrix}$ given in Eq. (4.15) obeys the expectation value

$$\langle \mathbf{n}|\boldsymbol{\sigma}|\mathbf{n}\rangle = \mathbf{n}. \tag{4.30}$$

The three Pauli matrices are combined as one vectorial quantity, $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$. Further prove that $|\mathbf{n}\rangle$ satisfies the eigenvalue problem

$$\mathbf{n}\cdot\boldsymbol{\sigma}|\mathbf{n}\rangle = +|\mathbf{n}\rangle. \tag{4.31}$$

Finally, derive the eigenstate $|-\mathbf{n}\rangle$ that satisfies the equation $\mathbf{n}\cdot\boldsymbol{\sigma}|-\mathbf{n}\rangle = -|-\mathbf{n}\rangle$. What is the spin expectation value of $|-\mathbf{n}\rangle$?

2. Check that the matrix given in the second line of Eq. (4.25) satisfies $\hat{\rho}^2 = \hat{\rho}$. What does this imply for the eigenvalues of the density matrix $\hat{\rho}$?

3. Given the density matrix (4.25), calculate the fraction of the pure state whose polarization direction is along $\mathbf{n}$. You can find the answer by taking the average of the with respect to the state $|\mathbf{n}\rangle$. Define another unit vector $\mathbf{m}$ in terms of the elements of $|P\rangle$ as $\mathbf{m} = \langle P|\boldsymbol{\sigma}|P\rangle/\langle P|P\rangle$. Express your answer in terms of the two unit vectors thus defined, $\mathbf{m}$ and $\mathbf{n}$. Finally, work out the fraction of the pure state whose polarization is along $-\mathbf{n}$. Verify that the sum of the two fractions is one.

Chapter 5

Radiation of Electromagnetic Waves

5.1 Green's function

Most equations of physics are written in the form

$$[\text{differential operator}][\text{effect}] = [\text{cause}]. \tag{5.1}$$

Some examples are Newton's equation of motion relating force (cause) to the change of the position (effect) through

$$\frac{d^2}{dt^2}\mathbf{r}(t) = \frac{1}{m}\mathbf{F}(t), \tag{5.2}$$

and Poisson's equation relating the potential (the "effect") to the charge distribution (the "cause") through

$$\boldsymbol{\nabla}^2 A_0 = -\frac{1}{\varepsilon_0}\rho(\mathbf{r}). \tag{5.3}$$

As we are interested in quantitatively working out the "effect" for a given "cause", what we need to achieve mathematically is to "invert" the differential operator and re-write the equation as

$$[\text{effect}] = [\text{differential operator}]^{-1}[\text{cause}]. \tag{5.4}$$

Such inverse of the differential operator is called the Green's function.

There is a clear analogy between the inversion of the differential operator discussed above, and the inversion of the matrix in the linear equation

$$\boldsymbol{M}\boldsymbol{x} = \boldsymbol{y}. \tag{5.5}$$

The two column vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ are obviously the analogues of "effect" and "cause" in the equations, while $\boldsymbol{M}$ would be the differential operator. In fact the matrix equation shares a significant characteristic with the differential equation known as the principle of superposition. Suppose we know the effect$_1$ for cause$_1$ and also effect$_2$ for cause$_2$, then the effect arising from the combined causes, cause$_1$ + cause$_2$, is the sum of the effects, effect$_1$ + effect$_2$. The superposition principle is useful in reducing the solution of a complicated cause as a superposition of the solutions for simple sources. In the matrix problem, $\boldsymbol{M}\boldsymbol{x}_1 = \boldsymbol{y}_1$ and $\boldsymbol{M}\boldsymbol{x}_2 = \boldsymbol{y}_2$ automatically implies

$$\boldsymbol{M}(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \boldsymbol{y}_1 + \boldsymbol{y}_2. \tag{5.6}$$

Suppose then that a simple differential operator d/dt was really like a matrix, and some sort of inverse matrix $(d/dt)^{-1} = G$ did exist. The multiplication of a matrix with its inverse, in the case of linear algebra, gives out the unit matrix. For differential operators, which are like matrices of infinite dimensions, the analogue of the unit matrix is the Dirac delta function:

$$\left(\frac{d}{dt}\right) G(t, t') = \delta(t - t'). \tag{5.7}$$

We have introduced t and t' as two indices of the infinite dimensional matrix $D(t, t')$. The function obeying such a relation is trivial to find. It is the step function!

$$G(t, t') = \theta(t - t'). \tag{5.8}$$

Having found the inverse of the differential operator d/dt, one can proceed to invert the Newton's law

$$\frac{dp}{dt} = F(t) \tag{5.9}$$

to find

$$p(t) = \int_{-\infty}^{\infty} dt' G(t, t') F(t') = \int_{-\infty}^{t} F(t') dt'. \tag{5.10}$$

The solution automatically obeys the causality, i.e. forces acting on the particle can only influence its motion after its application.

As a second example, consider the Poisson's equation given in Eq. (5.3). Following the same strategy, we would first look for the function $G(\mathbf{r}, \mathbf{r}')$ which obeys the relation

$$\boldsymbol{\nabla}^2 G(\mathbf{r}, \mathbf{r}') = -\frac{1}{\varepsilon_0} \delta(\mathbf{r} - \mathbf{r}'). \tag{5.11}$$

Having found such a function, the solution of the Poisson's equation will be automatically found as

$$A(\mathbf{r}) = \int d^3\mathbf{r}' G(\mathbf{r}, \mathbf{r}')\rho(\mathbf{r}'). \tag{5.12}$$

Fortunately, we already know the Green's function that satisfies (5.11). It is

$$G(\mathbf{r} - \mathbf{r}') = \frac{1}{4\pi\varepsilon_0}\frac{1}{|\mathbf{r} - \mathbf{r}'|}. \tag{5.13}$$

This solution, the Coulomb potential, is another example of the Green's function we mentioned earlier. General solution is immediately found,

$$A_0(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}d\mathbf{r}'. \tag{5.14}$$

We need to understand why the Green's function solution, corresponding to a delta-function source, is easy to find. Making the assumption that $G(\mathbf{r}, \mathbf{r}')$ only depends on the difference of the coordinates as $G(\mathbf{r} - \mathbf{r}')$, the Fourier transform of both sides of Eq. (5.11) yields

$$-\mathbf{k}^2 G(\mathbf{k}) = -\frac{e^{-i\mathbf{k}\cdot\mathbf{r}'}}{\varepsilon_0} \rightarrow G(\mathbf{k}) = \frac{e^{-i\mathbf{k}\cdot\mathbf{r}'}}{\varepsilon_0}\frac{1}{\mathbf{k}^2}. \tag{5.15}$$

The Fourier transform of a delta function is a constant, making it possible to obtain the Fourier-transformed $G(\mathbf{k})$ by a simple division. Now that we obtained $G(\mathbf{k})$ in the $\mathbf{k}$-space, it is a matter of doing the inverse Fourier transform to obtain the real-space solution, the Green's function:

$$G(\mathbf{r} - \mathbf{r}') = \frac{1}{\varepsilon_0}\int \frac{d^3\mathbf{k}}{(2\pi)^3}\frac{1}{\mathbf{k}^2}e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} = \frac{1}{4\pi\varepsilon_0}\frac{1}{|\mathbf{r} - \mathbf{r}'|}. \tag{5.16}$$

Let us return to the wave equation, Eq. (4.4), written in the Lorentz gauge. All equations are of the form

$$\left(\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2\right) A_\mu(\mathbf{r}, t) = \frac{1}{\varepsilon_0} j_\mu(\mathbf{r}, t). \tag{5.17}$$

To solve this partial differential equation by Green's function technique we will need to introduce the source term localized in both space and time:

$$\left(\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2\right) A_\mu(\mathbf{r}, t) = \frac{1}{\varepsilon_0}\delta(\mathbf{r} - \mathbf{r}')\delta(t - t'). \tag{5.18}$$

Expand the vector potential as

$$\begin{aligned} A_\mu(\mathbf{r}, t) &= \int \frac{d^3\mathbf{k}}{(2\pi)^2}\frac{d\omega}{2\pi} A_\mu(\mathbf{k}, \omega)e^{i\mathbf{k}\cdot\mathbf{r} - i\omega t}, \\ \delta(\mathbf{r} - \mathbf{r}')\delta(t - t') &= \int \frac{d^3\mathbf{k}}{(2\pi)^2}\frac{d\omega}{2\pi} e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}') - i\omega(t-t')}. \end{aligned} \tag{5.19}$$

On substituting we get an algebraic relation for $A_\mu(\mathbf{k},\omega)$ as

$$(\mathbf{k}^2-\omega^2)A_\mu(\mathbf{k},\omega)=\frac{1}{\varepsilon_0}e^{-i\mathbf{k}\cdot\mathbf{r}'+i\omega t'}\rightarrow A_\mu(\mathbf{k},\omega)=\frac{1}{\varepsilon_0}\frac{e^{-i\mathbf{k}\cdot\mathbf{r}'+i\omega t'}}{\mathbf{k}^2-\omega^2}. \quad (5.20)$$

Fourier transforming back to real space will require the evaluation of the integral

$$A_\mu(\mathbf{r}-\mathbf{r}',t-t')=\frac{1}{\varepsilon_0}\int\frac{d^3\mathbf{k}}{(2\pi)^2}\frac{d\omega}{2\pi}\ \frac{e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')-i\omega(t-t')}}{\mathbf{k}^2-\omega^2}. \quad (5.21)$$

Upon the successful evaluation of the integrals over the three momentum and one frequency variables one will obtain the Green's function for A_μ satisfying the Lorenz' gauge equation (5.18). Let's do the ω-integration first, since it is only one variable and one-dimensional integrals can be done usually in closed form. In doing this, however, one encounters a singularity at $\omega=\pm|\mathbf{k}|$. This singularity must be avoided one way or other, or we will never be able to find an exact answer for the integration. The procedure to avoid the singularity is not unique. Different boundary conditions lead to Green's functions of very different characters.

Let's then see first what happens when ω is displaced slight off the real-axis, by writing it as $\omega+i\delta$, where δ is an infinitesimally small positive number. We also choose $t'=0$ to simplify the notation. Now the pole lies at the lower half-plane, albeit very close to the real-axis, $\omega=\pm k-i\delta$. For $t<0$ we can close the contour through the upper semi-infinite circle which contains no singularity. The integral vanishes. For $t>0$ we need to close the contour through the lower half-plane, which contains two resides at $\omega=\pm k-i\delta$. Doing the residue integral yields

$$\int\frac{d\omega}{2\pi}\frac{e^{-i\omega t}}{\mathbf{k}^2-[\omega+i\delta]^2}=\theta(t)\frac{\sin(kt)}{k}=G^R(\mathbf{k},t). \quad (5.22)$$

Such Green's function, nonzero only for $t>t'$, is called the retarded Green's function. For the replacement $\omega\rightarrow\omega-i\delta$ we obtain the advanced Green's function,

$$\int\frac{d\omega}{2\pi}\frac{e^{-i\omega t}}{\mathbf{k}^2-[\omega-i\delta]^2}=-\theta(-t)\frac{\sin(kt)}{k}=G^A(\mathbf{k},t). \quad (5.23)$$

Carrying out the remaining $\mathbf{k}$-integral yields

$$\begin{aligned}G^R(\mathbf{r},t) &= \frac{1}{\varepsilon_0}\int\frac{d^3\mathbf{k}}{(2\pi)^3}G^R(\mathbf{k},t)e^{i\mathbf{k}\cdot\mathbf{r}}=\frac{1}{4\pi\varepsilon_0}\frac{1}{r}\theta(t)\delta(|\mathbf{r}|-t)\\ G^A(\mathbf{r},t) &= \frac{1}{\varepsilon_0}\int\frac{d^3\mathbf{k}}{(2\pi)^3}G^A(\mathbf{k},t)e^{i\mathbf{k}\cdot\mathbf{r}}=\frac{1}{4\pi\varepsilon_0}\frac{1}{r}\theta(-t)\delta(|\mathbf{r}|+t). \quad (5.24)\end{aligned}$$

Both $G^R(\mathbf{r}-\mathbf{r}',t-t')$ and $G^A(\mathbf{r}-\mathbf{r}',t-t')$ are solutions of Eq. (5.18).

As we mentioned, there is no reason why one choice of the displacement of ω ought to be superior to the other. From mathematical point of view, then,

the solution to the four-potential A_μ in the presence of the source term j_μ must be written as a general form

$$A_\mu(\mathbf{r},t) = \int \Big[\alpha G^R(\mathbf{r}-\mathbf{r}',t-t') + \beta G^A(\mathbf{r}-\mathbf{r}',t-t')\Big] j_\mu(\mathbf{r}',t')d^3\mathbf{r}'dt', \quad (5.25)$$

with a pair of coefficients α and β for the amount of mixing of retarded and advanced Green's functions. For the retarded part the interpretation is clear. A source at space time $(\mathbf{r}',t')$ will give rise to the potential at space time $(\mathbf{r},t)$ separated from the source exactly by the sphere satisfying $|\mathbf{r}-\mathbf{r}'| = t-t'$. The non-causal solution involving G^A is still admitted mathematically[1].

Recall that radiation laws can be written also in the Coulomb gauge $\boldsymbol{\nabla}\cdot\mathbf{A} = 0$, in which the equations look

$$\left(\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2\right)\mathbf{A} + \boldsymbol{\nabla}\left(\frac{\partial A_0}{\partial t}\right) = \frac{1}{\varepsilon_0}\mathbf{j},$$
$$\boldsymbol{\nabla}^2 A_0 = -\frac{\rho}{\varepsilon_0}. \quad (5.26)$$

The second of these equations admit a ready solution (because we have solved this problem before for static distribution of charges):

$$A_0(\mathbf{r},t) = \frac{1}{4\pi\varepsilon_0}\int \frac{\rho(\mathbf{r}',t)}{|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}'. \quad (5.27)$$

According to the above formula the charge distribution at a given time t has an instantaneous influence on the Coulomb potential elsewhere at the same time, in apparent violation of the causality condition that no information can travel faster than the speed of light. We are reminded, and re-assured, however, that the scalar potential itself is not physical and may look counter-intuitive, as long as the physical electromagnetic fields obey the causality faithfully. To check that, we also need to solve for the vector potential $\mathbf{A}(\mathbf{r},t)$ by solving the first of the equations in (5.26), then using $\mathbf{E} = -\boldsymbol{\nabla}A_0 - \partial_t\mathbf{A}$ to calculate $\mathbf{E}(\mathbf{r},t)$. The final expression will be free of causality violation. This is left as a homework exercise.

5.2 Dipole and quadrupole radiation

To work out the radiation field emanating from an oscillating charge and current, we assume the simple periodic time dependence

$$\mathbf{J}(\mathbf{r},t) = \mathbf{J}(\mathbf{r})e^{-i\omega t}. \quad (5.28)$$

[1]There is an interesting anecdote mentioned in Zangwill's book. Einstein argued that the advanced Green's functions are also physically admissible, if one only chose to look at a small space region. Another Swiss physicists name Walther Ritz argued instead that the advanced Green's function should be banned from the solution; otherwise the causality and the second law of thermodynamics will be in trouble. See p.723 of Andrew Zangwill, *Modern Electrodynamics* (Cambridge University Press, 2013)

In the Lorenz gauge this gives rise to the vector potential

$$\begin{aligned}\mathbf{A}(\mathbf{r},t) &= \frac{1}{4\pi\varepsilon_0}\int d^3\mathbf{r}'dt'\frac{\mathbf{J}(\mathbf{r}')e^{-i\omega t'}}{|\mathbf{r}-\mathbf{r}'|}\delta(t-t'-|\mathbf{r}-\mathbf{r}'|) \\ &= \frac{e^{-i\omega t}}{4\pi\varepsilon_0}\int d^3\mathbf{r}'\frac{\mathbf{J}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}e^{ik|\mathbf{r}-\mathbf{r}'|}. \qquad (5.29)\end{aligned}$$

We have replaced ω by $k = |\mathbf{k}|$ in the final expression. Using $\mathbf{A}$ the magnetic field is worked out, $\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}$, and from this, in the region away from the current source, the electric field as

$$\frac{\partial\mathbf{E}}{\partial t} = \boldsymbol{\nabla}\times\mathbf{B} \rightarrow -i\omega\mathbf{E} = \boldsymbol{\nabla}\times\mathbf{B}. \qquad (5.30)$$

Often, the source of radiation is confined over a distance d much less than the wavelength λ of the electromagnetic field being emitted from it, $kd \ll 1$. Take the radiation of optical waves from most atoms we see around us. The atomic dimension is on the order of one Angstrom, $\sim 10^{-10}$m, while the wavelength of the radiation being emitted from the atoms are several thousand Angstroms. The reason lies with the big number for the speed of light in determining the wavelength: $\lambda = c/f$. In the following we will always consider $\lambda \gg d$. There is another length scale involved in the integral of Eq. (5.29), which is the distance r from the source to the point of observation. As extremes, one can either imagine situations where this distance is much less than the wavelength λ, or much greater.

In the first case we may say the inequality $|\mathbf{r}-\mathbf{r}'| \ll \lambda$ holds, and the vector potential approximated as

$$\mathbf{A}(\mathbf{r},t) \approx \frac{e^{-i\omega t}}{4\pi\varepsilon_0}\int d^3\mathbf{r}'\frac{\mathbf{J}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \qquad (5.31)$$

The resulting expression is quite easy to appreciate. Apart from the overall periodic time dependence $e^{-i\omega t}$, this is the same as the instantaneous vector potential for the stationary current distribution. Because the distance from the source to the point of observation $|\mathbf{r}-\mathbf{r}'|$ is assumed much less than one wavelength, the retardation effect due to the finite speed of light has no place in the final expression.

The other extreme case, namely the far-field region where the inequality $|\mathbf{r}-\mathbf{r}'| \gg \lambda$ holds, is more interesting in fact. In this limit we can approximate

$$|\mathbf{r}-\mathbf{r}'| = r\sqrt{1+r'^2-2\mathbf{r}\cdot\mathbf{r}'/r^2} \approx r - \mathbf{r}\cdot\mathbf{r}'/r. \qquad (5.32)$$

Taking $e^{ik|\mathbf{r}-\mathbf{r}'|} \approx e^{ikr}e^{-ik\hat{r}\cdot\mathbf{r}'}$, we can expand the vector potential as

$$\mathbf{A}(\mathbf{r},t) = \frac{1}{4\pi\varepsilon_0}\cdot\frac{e^{ikr-i\omega t}}{r}\cdot\int\mathbf{J}(\mathbf{r}')e^{-ik\hat{r}\cdot\mathbf{r}'} = \frac{1}{4\pi\varepsilon_0}\cdot\frac{e^{ikr-i\omega t}}{r}\cdot\mathbf{J}(k\hat{r}). \qquad (5.33)$$

It is given by an out-going wave in three-dimensional space multiplied by the Fourier transform of the current distribution. The spherical wave factor in Eq.

(5.33) is due to the fact that the source, squeezed so tightly over a small region of size d, look no better than point-like to an observer far from it.

In carrying out the integration for the Fourier transform $\int \mathbf{J}(\mathbf{r}')e^{-ik\hat{r}\cdot\mathbf{r}'}$ we are always in the region where $k|\hat{r}\cdot\mathbf{r}'| \sim d/\lambda$ is much less than unity. As a first approximation, then, the number can be set to unity, or else we may set the argument of the Fourier-transformed function to zero: $\mathbf{J}(k\hat{r}) \to \mathbf{J}(0) = \int \mathbf{J}(\mathbf{r}')$. The vector potential $\mathbf{A}(\mathbf{r},t)$ for the field generated from the confined source of current somehow depends on the space integral of the entire current distribution. A series of manipulations can prove that this integral is actually related to the dipole moment being carried by the source. A glimpse of this relation is seen through a simple example of an oscillating charge, with $q = +1$. Such charge has the position $x(t) = d\cos[\omega t]$ and the dipole moment $p(t) = qx(t) = x(t)$. Because the charge is oscillating, it always generates a current $j(t) = \dot{x} = \dot{p}$. Namely, the current is proportional to the time derivative of the dipole moment.

Now we return to the integral $\int \mathbf{J}(\mathbf{r}')d\mathbf{r}'$ and formally re-write it as an equivalent expression

$$\int J_\alpha(\mathbf{r}')d\mathbf{r}' = \int \Big[\boldsymbol{\nabla}'\cdot(r'_\alpha \mathbf{J}(\mathbf{r}')) - r'_\alpha(\boldsymbol{\nabla}'\cdot\mathbf{J}(\mathbf{r}'))\Big]d\mathbf{r}', \tag{5.34}$$

the first of which will give zero by virtue of Gauss' theorem. In handling the second term we recall the continuity equation $\boldsymbol{\nabla}\cdot\mathbf{J}(\mathbf{r}',t) + \partial_t\rho(\mathbf{r}',t) = 0$. Assuming both quantities have the harmonic time dependence, $\mathbf{J}(\mathbf{r},t) = \mathbf{J}(\mathbf{r})e^{-i\omega t}$ and $\rho(\mathbf{r},t) = \rho(\mathbf{r})e^{-i\omega t}$, we easily arrive at a version of the continuity equation $\boldsymbol{\nabla}\cdot\mathbf{J}(\mathbf{r}) - i\omega\rho(\mathbf{r}) = 0$. Divergence of the current density becomes the charge density in this way,

$$\int r'_\alpha(\boldsymbol{\nabla}'\cdot\mathbf{J}(\mathbf{r}'))d\mathbf{r}' = i\omega\int r'_\alpha\rho(\mathbf{r}')d\mathbf{r}'. \tag{5.35}$$

The final expression is exactly the electric dipole $\mathbf{p} = \int \mathbf{r}'\rho(\mathbf{r}')d\mathbf{r}'$. To this order of approximation the vector potential becomes

$$\mathbf{A}(\mathbf{r},t) = \frac{-i\omega}{4\pi\varepsilon_0}\frac{e^{ikr-i\omega t}}{r}\mathbf{p}. \tag{5.36}$$

A given dipolar charge distribution $\mathbf{p}$ oscillating at the frequency ω gives rise to the vector potential of the form, Eq. (5.36). Electric and magnetic fields can be worked out,

$$\begin{aligned}
\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A} &= \frac{\omega^2}{4\pi\varepsilon_0}(\hat{r}\times\mathbf{p})\frac{e^{ikr}}{r}\left(1-\frac{1}{ikr}\right), \\
\mathbf{E} = \frac{i}{\omega}\boldsymbol{\nabla}\times\mathbf{B} &= \frac{\omega^2}{4\pi\varepsilon_0}\frac{e^{ikr}}{r}\Big[(\hat{r}\times\mathbf{p})\times\hat{r} + [3\hat{r}(\hat{r}\cdot\mathbf{p}) - \mathbf{p}]\left(\frac{1}{(kr)^2} - \frac{i}{kr}\right)\Big].
\end{aligned} \tag{5.37}$$

Magnetic field is strictly perpendicular to the radial direction, but the electric field has components both perpendicular and parallel to $\hat{r}$. Strict orthogonality

among $\mathbf{E}$, $\mathbf{B}$, and $\mathbf{k}$ holds only in the case of the plane wave. For large distance we recover the typical behavior of radiation field,

$$\begin{aligned} \mathbf{B} &\approx \frac{\omega^2}{4\pi\varepsilon_0}(\hat{r}\times\mathbf{p})\frac{e^{ikr}}{r}, \\ \mathbf{E} &\approx \frac{\omega^2}{4\pi\varepsilon_0}(\hat{r}\times\mathbf{p})\times\hat{r}\frac{e^{ikr}}{r} \approx \mathbf{B}\times\hat{r}. \end{aligned} \tag{5.38}$$

In the so-called near field region $kr \ll 1$ the fields become

$$\begin{aligned} \mathbf{B} &\approx \frac{i\omega}{4\pi\varepsilon_0}(\hat{r}\times\mathbf{p})\frac{1}{r^2}, \\ \mathbf{E} &\approx \frac{1}{4\pi\varepsilon_0}[3\hat{r}(\hat{r}\cdot\mathbf{p})-\mathbf{p}]\frac{1}{r^3}. \end{aligned} \tag{5.39}$$

The electric field is just that of a dipole. The magnetic field is roughly a factor kr times less than the strength of the electric field at the same position.

Higher-order terms in $\mathbf{J}(\mathbf{k})$ are obtained by expanding $e^{-i\mathbf{k}\cdot\mathbf{r}'}$ to first order,

$$\int \mathbf{J}(\mathbf{r}')e^{-ik\hat{r}\cdot\mathbf{r}'}d\mathbf{r}' \approx \int \mathbf{J}(\mathbf{r}')d\mathbf{r}' - ik\int(\hat{r}\cdot\mathbf{r}')\mathbf{J}(\mathbf{r}')d\mathbf{r}'. \tag{5.40}$$

Again we play some tricks with $(\hat{r}\cdot\mathbf{r}')\mathbf{J}(\mathbf{r}')$ to re-write it in a more recognizable form:

$$(\hat{r}\cdot\mathbf{r}')\mathbf{J}(\mathbf{r}') = \frac{1}{2}[(\hat{r}\cdot\mathbf{r}')\mathbf{J}(\mathbf{r}') + \mathbf{r}'(\hat{r}\cdot\mathbf{J}(\mathbf{r}'))] + \frac{1}{2}[\mathbf{r}'\times\mathbf{J}(\mathbf{r}')]\times\hat{r}. \tag{5.41}$$

Integral of the second term leads to the magnetic dipole moment,

$$\left(\frac{1}{2}\int[\mathbf{r}'\times\mathbf{J}(\mathbf{r}')]d\mathbf{r}'\right)\times\hat{r} = \mathbf{m}\times\hat{r}. \tag{5.42}$$

Integral of the first term leads to the electric quadrupole moment,

$$\begin{aligned} \int\frac{1}{2}[(\hat{r}\cdot\mathbf{r}')\mathbf{J}(\mathbf{r}') + \mathbf{r}'(\hat{r}\cdot\mathbf{J}(\mathbf{r}'))]_\alpha &= \frac{1}{2}(\hat{r})_\beta\int\left(r'_\alpha J_\beta + r'_\beta J_\alpha\right), \\ \int\left(r'_\alpha J_\beta + r'_\beta J_\alpha\right) &= -i\omega\int r'_\alpha r'_\beta\rho(\mathbf{r}') = -i\omega Q_{\alpha\beta}. \end{aligned} \tag{5.43}$$

Altogether the vector potential expansion gives

$$\mathbf{A}(\mathbf{r},t) = \frac{-i\omega}{4\pi\varepsilon_0}\frac{e^{ikr-i\omega t}}{r}\Big(\mathbf{p} + \mathbf{m}\times\hat{r} + \frac{i}{2}\omega(\hat{r})_\alpha Q_{\alpha\beta}\Big). \tag{5.44}$$

5.3 Problems

1. Verify that integration over $\mathbf{k}$ in Eq. (5.16) yields the Coulomb potential.
2. Derive Eqs. (5.22) and (5.23) by using the residue theorem.
3. Derive Eq. (5.24) by completing the integral.
4. Derive Eq. (5.37).
5. Prove that the electric field $\mathbf{E}(\mathbf{r}, t)$ obtained from the Coulomb gauge obeys the causality.
6. A uniform current I is established around a ring of radius R for $t > 0$. Calculate the electric and magnetic fields generated from such a ring at a distance much larger than the ring radius.

Chapter 6

Propagation of EM Waves

6.1 Position-dependent index of refraction

For electromagnetic waves passing through inhomogeneous matter, the dielectric constant ε_0 becomes a function of position $\mathbf{r}$ which we write $\varepsilon_0\varepsilon(\mathbf{r})$. The inhomogeneous Maxwell's equations in the absence of sources read

$$\boldsymbol{\nabla}\times\mathbf{E}=-\frac{\partial\mathbf{B}}{\partial t},\ \boldsymbol{\nabla}\times\mathbf{B}=\varepsilon(\mathbf{r})\frac{\partial\mathbf{E}}{\partial t}. \tag{6.1}$$

Taking the curl of the Faraday's law gives

$$\boldsymbol{\nabla}^2\mathbf{E}=\frac{\partial}{\partial t}(\boldsymbol{\nabla}\times\mathbf{B})=\varepsilon(\mathbf{r})\frac{\partial^2\mathbf{E}}{\partial t^2}. \tag{6.2}$$

The speed of propagation $c(\mathbf{r})=c_{\text{vac}}/\sqrt{\varepsilon(\mathbf{r})}$ is also a function of the position now. Taking the curl of the Ampere-Maxwell equation is slightly more complicated,

$$-\boldsymbol{\nabla}^2\mathbf{B}=\boldsymbol{\nabla}\times\left(\varepsilon(\mathbf{r})\frac{\partial\mathbf{E}}{\partial t}\right)=\frac{1}{\varepsilon(\mathbf{r})}\boldsymbol{\nabla}\varepsilon(\mathbf{r})\times(\boldsymbol{\nabla}\times\mathbf{B})-\varepsilon(\mathbf{r})\frac{\partial^2\mathbf{B}}{\partial t^2}. \tag{6.3}$$

If the medium is slowly varying, we can treat $\boldsymbol{\nabla}\varepsilon(\mathbf{r})\approx 0$ and obtain a set of wave-like equations

$$\boldsymbol{\nabla}^2\mathbf{E}=\varepsilon(\mathbf{r})\frac{\partial^2\mathbf{E}}{\partial t^2},\quad \boldsymbol{\nabla}^2\mathbf{B}=\varepsilon(\mathbf{r})\frac{\partial^2\mathbf{B}}{\partial t^2}. \tag{6.4}$$

For the time dependence $e^{-i\omega t}$, the problem is reduced to that of solving $\boldsymbol{\nabla}^2\psi(\mathbf{r})+\omega^2\varepsilon(\mathbf{r})\psi(\mathbf{r})=0$. For a small enough region the dielectric function is roughly constant, allowing for a plane-wave solution $\sim e^{i\mathbf{k}\cdot\mathbf{r}}$. More generally, we say that the complex function $\psi(\mathbf{r})$ has a constant amplitude but varying phase, $\psi(\mathbf{r})=\psi_0 e^{i\omega S(\mathbf{r})}$. Substitution yields the equation for S,

$$(\boldsymbol{\nabla}S)^2-n^2(\mathbf{r})-\frac{i}{\omega}\boldsymbol{\nabla}^2S=0. \tag{6.5}$$

We are writing $\varepsilon(\mathbf{r}) = n^2(\mathbf{r})$ as the index of refraction $n(\mathbf{r})$. Just the structure of the equation says that S must be complex, $S = S_r + iS_i$. The imaginary part S_i is responsible for the amplitude modulation of the wave, $\psi = e^{i\omega S_r} e^{-\omega S_i}$. To solve this equation, we will ignore $\boldsymbol{\nabla}^2 S$ temporarily, and later justify it. With $\boldsymbol{\nabla} S \cdot \boldsymbol{\nabla} S = n^2(\mathbf{r})$ we deduce

$$\boldsymbol{\nabla} S = \mathbf{k}(\mathbf{r}) = n(\mathbf{r})\hat{k}(\mathbf{r}), \tag{6.6}$$

where $\hat{k}(\mathbf{r}) = \boldsymbol{\nabla} S/|\boldsymbol{\nabla} S|$. The amplitude of S is fixed, but the direction $\hat{k}$ is not. From Snell's law we know that the propagation direction is related to the gradient of the index of refraction $\boldsymbol{\nabla} n(\mathbf{r})$. One can derive such a law, but one must first introduce the variable s for the distance traveled by the wave in the sense of geometrical optics. The orientation vector $\hat{k}$ is just $d\mathbf{r}/ds$, since $ds = \sqrt{d\mathbf{r} \cdot d\mathbf{r}}$. One can further note that $d/ds = \hat{k} \cdot \boldsymbol{\nabla}$, which therefore gives

$$\frac{d}{ds}\boldsymbol{\nabla} S = \boldsymbol{\nabla}\left(\hat{k} \cdot \boldsymbol{\nabla} S\right) = \boldsymbol{\nabla} n(\mathbf{r}). \tag{6.7}$$

Further expanding $d(\boldsymbol{\nabla} S)/ds$,

$$\frac{d}{ds}\boldsymbol{\nabla} S = \hat{k}(\hat{k} \cdot \boldsymbol{\nabla})n + n(\hat{k} \cdot \boldsymbol{\nabla})\hat{k} = \boldsymbol{\nabla} n. \tag{6.8}$$

This is the different equation to be obeyed by the orientation vector $\hat{k}$. Assuming there is only the z-dependence in the index of refraction we can reduce the equation to

$$\hat{k}_z\left(\hat{k}\frac{dn}{dz} + n\frac{d\hat{k}}{dz}\right) = \hat{k}_z \frac{d}{dz}\left(n\hat{k}\right) = \hat{z}\frac{dn}{dz}. \tag{6.9}$$

Projection to the xy-plane gives the Snell's law,

$$\frac{d}{dz}\left(n\hat{k}_\parallel\right) = 0. \tag{6.10}$$

Projection to the $\hat{z}$-axis gives

$$\hat{k}_z \frac{d}{dz}\left(n\hat{k}_z\right) = \frac{dn}{dz} = \hat{k}_z^2 \frac{dn}{dz} + \frac{1}{2} n \frac{d\hat{k}_z^2}{dz}. \tag{6.11}$$

With the identity $\hat{k}_z^2 = 1 - \hat{k}_\parallel^2$ one can reduce the equation to Snell's law again.

In the vicinity of a particular position $\mathbf{r}_0$ we can expand

$$\psi(\mathbf{r}) \approx e^{i\omega S(\mathbf{r}_0)} e^{i\omega \boldsymbol{\nabla} S(\mathbf{r}_0)\cdot(\mathbf{r}-\mathbf{r}_0)}. \tag{6.12}$$

The propagation can be described as a plane wave with the wave vector $\mathbf{k}(\mathbf{r})$ that varies with the position. The magnitude $|\mathbf{k}| = |\boldsymbol{\nabla} S| = n(\mathbf{r}) = 2\pi/\lambda(\mathbf{r})$

defines the local wavelength of the propagating wave. The orientation of the propagation is dictated by the generalized Snell's law.

The neglect of the second derivative $\boldsymbol{\nabla}^2 S$ from the equation (6.5) can be justified now. Assume that the length L over which $\varepsilon(\mathbf{r})$ varies is much larger than the wavelength of the wave, λ. This can be conveniently parameterized by writing $n(\mathbf{r}) = n(\mathbf{r}/L)$. Then from $\boldsymbol{\nabla} S = n(\mathbf{r}/L)\hat{k}$ we find

$$\boldsymbol{\nabla}^2 S = \boldsymbol{\nabla} n(\mathbf{r}/L)\hat{k} + n(\mathbf{r}/L)\boldsymbol{\nabla}\cdot\hat{k}. \tag{6.13}$$

The term $\nabla n(\mathbf{r}/L)$ must be proportional to $1/L$, and from Snell's law we also expect the magnitude of $\boldsymbol{\nabla}\hat{k}$ to scale as $1/L$. Hence, $|\boldsymbol{\nabla}^2 S|/\omega \sim 1/\omega L \sim \lambda/L \ll 1$.

Insisting on obtaining the correction, we may write S as a complex function $S = S_r + iS_i$ and insert it into Eq. (6.5). Approximately we obtain

$$[(\boldsymbol{\nabla} S_r)^2 - n^2] + i[2\boldsymbol{\nabla} S_r\cdot\boldsymbol{\nabla} S_i - \frac{1}{\omega}\boldsymbol{\nabla}^2 S_r] = 0. \tag{6.14}$$

The imaginary part of the equation gives

$$2n\hat{k}\cdot\boldsymbol{\nabla} S_i = \frac{1}{\omega}\boldsymbol{\nabla}\cdot(n\hat{k}). \tag{6.15}$$

We take $\boldsymbol{\nabla}\cdot\hat{k} = 0$, and note $(\hat{k}\cdot\boldsymbol{\nabla})S_i = dS_i/ds$,

$$2\omega\frac{dS_i}{ds} = \frac{1}{n}\frac{dn}{ds} = \frac{d}{ds}[\ln n]. \tag{6.16}$$

We obtain the wave function

$$\psi = e^{-\omega S_i}e^{i\omega S_r} \simeq \sqrt{\frac{1}{n(\mathbf{r})}}e^{i\omega S_r(\mathbf{r})}. \tag{6.17}$$

6.2 Scattering of EM waves

The influence of inhomogeneous medium can be understood from the perspective of scattering of incident light. The dielectric constant $\varepsilon(\mathbf{r})$ is written $\varepsilon(\mathbf{r}) = 1 + \delta\varepsilon(\mathbf{r})$, $|\delta\varepsilon(\mathbf{r})| \ll 1$. In the previous section we required the variation $\delta\varepsilon(\mathbf{r})$ to be slow, in the sense that the typical variation occurs on a length scale much larger than the wavelength of light passing through, but imposed no restriction on the magnitude of its variation. Here we drop the adiabatic (slow-variation) requirement, but instead focus on cases where the amplitude of the variation is small.

Initially let's say we have the incident light of wave vector $\mathbf{k}$, $\mathbf{E}_0(\mathbf{r}) = \mathbf{E}_0 e^{i\mathbf{k}\cdot\mathbf{r}}$, $\mathbf{B}_0(\mathbf{r}) = \hat{k}\times\mathbf{E}_0(\mathbf{r})$. Due to the inhomogeneity of $\varepsilon(\mathbf{r})$ this cannot be the solution of the Maxwell equation, but also cannot be too far from the true solution. We write

$$\begin{aligned}\mathbf{E}(\mathbf{r}) &= \mathbf{E}_0(\mathbf{r}) + \delta\mathbf{E}(\mathbf{r}),\\ \mathbf{B}(\mathbf{r}) &= \mathbf{B}_0(\mathbf{r}) + \delta\mathbf{B}(\mathbf{r}).\end{aligned} \tag{6.18}$$

Inserting them into Maxwell's equations gives

$$\left(\varepsilon(\mathbf{r})\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2\right)\mathbf{B} = \left(\frac{\boldsymbol{\nabla}\varepsilon(\mathbf{r})}{\varepsilon(\mathbf{r})}\right) \times (\boldsymbol{\nabla}\times\mathbf{B}) \approx \boldsymbol{\nabla}\delta\varepsilon \times (\boldsymbol{\nabla}\times\mathbf{B}_0). \tag{6.19}$$

Inserting $\mathbf{B}(\mathbf{r}) = \mathbf{B}_0(\mathbf{r}) + \delta\mathbf{B}(\mathbf{r})$ on the l.h.s. gives

$$\begin{aligned}\left(\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2\right)\delta\mathbf{B} &= \boldsymbol{\nabla}\delta\varepsilon(\mathbf{r}) \times (\boldsymbol{\nabla}\times\mathbf{B}_0(\mathbf{r})) - \delta\varepsilon(\mathbf{r})\frac{\partial^2\mathbf{B}_0(\mathbf{r})}{\partial t^2}\\ &= \left[\boldsymbol{\nabla}\delta\varepsilon(\mathbf{r}) \times (i\mathbf{k}\times\mathbf{B}_0) + \omega^2\delta\varepsilon(\mathbf{r})\mathbf{B}_0\right] e^{i\mathbf{k}\cdot\mathbf{r}-i\omega t}.\end{aligned} \tag{6.20}$$

The whole r.h.s. is a source term, written $\delta\mathbf{S}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}-i\omega t}$. We have encountered differential equations of this sort before when considering the dynamics of the vector potential in the Lorenz gauge. We also found the solution there, which read

$$\begin{aligned}\delta\mathbf{B}(\mathbf{r},t) &= \frac{1}{4\pi}\int d^3\mathbf{r}'dt'\ \delta(|\mathbf{r}-\mathbf{r}'| - (t-t'))\frac{\delta\mathbf{S}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}e^{i\mathbf{k}\cdot\mathbf{r}'-i\omega t'}\\ &= \frac{e^{-i\omega t}}{4\pi}\int d^3\mathbf{r}'\ \frac{\delta\mathbf{S}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}e^{i\mathbf{k}\cdot\mathbf{r}'+i\omega|v r-\mathbf{r}'|}.\end{aligned} \tag{6.21}$$

For points very far from the source of inhomogeneity we may further approximate the solution,

$$\delta\mathbf{B}(\mathbf{r},t) \approx \frac{e^{ikr-i\omega t}}{4\pi r}\int d^3\mathbf{r}'\delta\mathbf{S}(\mathbf{r}')e^{i\mathbf{k}\cdot\mathbf{r}'} = \frac{e^{ikr-i\omega t}}{4\pi r}\delta\mathbf{S}(\mathbf{k}). \tag{6.22}$$

There are two terms forming $\delta\mathbf{S}(\mathbf{r})$, both of which will give quantities proportional to $\mathbf{k}^2$ upon the Fourier transform. In turn the intensity of scattered light depends on the square $|\mathbf{S}(\mathbf{k})|^2 \propto k^4$. This is why there is more scattering light at longer wavelengths, and we see more of the blue portion of the incident light from the sun ("Why the sky is blue").