Chapter 2.

Electrodynamics

"When you arrive at a fork in the road, take it." ... Yogi Berra
Chapter 2. Electrodynamics
## Contents

Introductory Comments ................................................................. 121

1. The Equations of James Clerk Maxwell ........................................ 126
   Conserved Quantities ............................................................... 127
   Charge Conservation and Lorentz Transformation ............................... 130
   Potentials .................................................................................. 131
   Gauge Transformation .................................................................. 132
   Lorenz Gauge ........................................................................... 134
   Coulomb Gauge ......................................................................... 136
   Four-Divergences ...................................................................... 137
   Taking Stock ............................................................................. 138

2. Electric and Magnetic Fields ........................................................ 139
   Electromagnetic Field Strength Tensor ........................................... 140
   Elements of the Field Strength Tensor ........................................... 142
   Lorentz Force ............................................................................ 145
   Reminder .................................................................................... 147
   Example 1. An Accelerated Charge Radiates ...................................... 148

3. Quantization of Transverse Waves ............................................... 152
   Electromagnetic Energy Density .................................................... 153
   Quantization .............................................................................. 154
   Commutator Relations ................................................................... 155
   Summary of the Quantized Electromagnetic Field ............................... 157
   Example 2. Photon Spin .................................................................. 158
   Spin Density ................................................................................ 162

4. Absorption and Emission of Photons .......................................... 165
   Bosons, Lasers, and so on .............................................................. 168
   Spontaneous Emission .................................................................. 169
   Perturbation ................................................................................ 169
   Electric Dipole Approximation ....................................................... 171
"From a long view of the history of mankind the most significant event of the nineteenth century will be judged as Maxwell's discovery of the laws of electrodynamics."

Richard P. Feynman

Introductory Comments

We shall begin with the set of coupled partial differential equations formulated by the brilliant Scottish scientist James Clerk Maxwell. Recognized at an early age as a mathematical prodigy, his meteoric rise saw him appointed Professor of Physics at Marischal College (which later became part of the University of Aberdeen) at age twenty-four. He was elected Fellow of the Royal Society of Edinburgh that same year. He moved to King's College, London in 1860, at which point he was elected Fellow of the Royal Society (FRS) at age twenty-nine. He left King's College in 1865 to return with his wife Katherine to his family's estate in Glenlair, Scotland, where he continued his work. From there he went to the University of Cambridge in 1871 as the Cavendish Professor of Physics, where he immediately set to work founding the Cavendish Laboratory. It staggers the imagination that James Clerk Maxwell's appointment was in Experimental Physics. He remained at Cambridge until his untimely death at age forty-eight.

Maxwell commenced research on what came to be known as the theory of electromagnetism (later referred to as Maxwell's equations) in his late twenties. By age twenty-nine he had developed the concept of traveling electromagnetic waves: "...we can scarcely avoid the inference that light consists of the transverse undulations of the same medium which is the cause of electric and magnetic phenomena."

Aspects and consequences of the theory fell neatly into place, and did so rapidly, culminating in Maxwell's 1864 oral presentation to the Royal Society and his seminal 1865 masterpiece: A Dynamical Theory of the Electromagnetic Field. This was the formal introduction of Maxwell's equations. Maxwell's theory has withstood the test of time with essentially no modification. Indeed, it served as the cornerstone of twentieth century physics, and it spawned the field of electrical engineering. It even anticipated the theory of special relativity, which followed by four decades. This is noteworthy, in fact downright amusing, in light of the fact that Maxwell's theory was developed in the context of the "luminiferous aether theory" that Einstein's 1905 paper put to rest. One detects his adherence to the aether theory in the quote at the end of the previous paragraph. The impact of
Maxwell's equations on the world we live in cannot be overstated, for example, they are the basis of present day communication and information technologies.

Maxwell's contributions extend well beyond electromagnetism. He published his first paper at age fourteen in the field of mathematics. Despite being the sole author, he was considered too young to present the work to so august an audience as the Royal Society of Edinburgh, so this was done "for him" by a more senior mathematician, James Forbes. Again, at age eighteen he was considered too young for the Royal Society of Edinburgh, so his tutor presented his work. At age twenty-six, he proved that the rings of Saturn are composed of small solid objects that collectively behave as a fluid, ending a 200-year-old puzzle: What accounts for the uncanny stability of the rings? The case that Maxwell presented was so ironclad that he was awarded the 1857 Adams Prize, and the matter was considered settled. His model was verified one-and-a-quarter centuries later by the Voyager spacecraft. And of course we must not overlook his contributions to thermodynamics and the kinetic theory of gases: Maxwell's Thermodynamic Relations, the Maxwell demon, the Maxwell-Boltzmann distribution, and so on.

Despite the best intentions of the framers of the United States Constitution, in fact, not all people are created equal.

Another major contributor in the early days of electromagnetic theory was the Danish physicist Ludvig Lorenz. His contributions, particularly in the area of gauge transformations, which he originated, have been undervalued and even ignored. His work is often mistakenly attributed to the Dutch physicist Hendrik Antoon Lorentz, in part because gauge transformations are central to electrodynamics. For example, the Lorenz gauge ensures that electrodynamics is Lorentz covariant.

The main goal of Section 1 is a recasting of Maxwell's equations into relativistic format, wherein all vectors are four-vectors, and all elements of the theory are form invariant under Lorentz transformation. We know from Chapter 1 that this property is referred to as Lorentz covariance. This exercise brings to the forefront the gauge fields $\phi$ and $A$. They exhibit redundancy in the sense that they can be altered without affecting the so-called force fields $E$ and $B$. This flexibility is referred to as gauge freedom. It is helpful insofar as ensuring that the relativistic wave equation is Lorentz covariant, and it will prove to be even more central when the gauge principle is introduced in Section 5. The gauge fields $\phi$ and $A$ were introduced originally as scalar and vector potentials that serve to simplify calculations. It turned out that they have far greater significance.

It is worth noting that Maxwell's original work assigned privileged status to the potentials. It was Oliver Heaviside who rewrote the equations in terms of $E$ and $B$ – what we know today as Maxwell's equations. Heaviside disparaged the potentials as inferior mathematical objects.
Chapter 2. Electrodynamics

Section 2 deals with electric and magnetic fields: how they transform into one another via Lorentz transformation, how to create a Lorentz covariant creature referred to as the electromagnetic field strength tensor (sometimes referred to as the Faraday tensor), and how this tensor relates to the original Maxwell equations. This section is more mathematical than Section 1. It is worth a look, but not an exhaustive one at this point. It comes into its own when relativistic quantum field theory is introduced in Chapter 4.

Section 3 is devoted to quantization of the electromagnetic field. This is carried out using the vector potential in the Coulomb (transverse) gauge to write an expression for the electromagnetic energy density, which we then take to be a Hamiltonian density. The desired result follows quickly. The math is just a matter of quantizing harmonic oscillators following essentially the same procedure that was carried out in Part IV, Chapter 1, where phonons – the quanta of periodic-lattice vibrations – were obtained. A "dual use model" was introduced that is applicable to the quantization of bosonic fields other than that of lattice vibrations.

Quantization in the Coulomb gauge is a fast route to the photon that we know and love, the quantum of electromagnetism that carries energy $\hbar \omega$. However, it avoids a number of interesting and worthwhile subtleties. For example, the photons that emerge via this route are transverse. That is, their electric and magnetic fields are transverse to their direction of propagation, except for a minor edge effect where the fields drop to zero. After all, the Coulomb gauge condition is $\nabla \cdot \vec{A} = 0$, which is consistent with transverse plane waves. The two transverse photons are the ones that propagate and carry energy. This leaves aside quantization of the other two field polarizations that are necessary for quantization of the full electromagnetic field.

These other photons (gauge bosons) – a longitudinal one and a timelike one – mediate electromagnetic forces. For example, they are central to quantum electrodynamics. Were one so ambitious (in effect, crazy) as to attempt to formulate electronic structure theory in terms of fully quantized fields, longitudinal and timelike photons would be there in full armor. Luckily, classical fields work quite well. Longitudinal and timelike photons are, however, important in the quantum theory of core electrons in heavy atoms, and this will be discussed when we get to relativistic quantum mechanics in Chapter 4.

In the present chapter, the Coulomb gauge brings us to an important result quickly, without getting bogged down in such details. A more comprehensive derivation will be carried out later, time permitting, with attention paid to nuance. Section 3 ends with a derivation of photon spin, namely, the values of $\pm \hbar$ for the clockwise and counterclockwise polarizations.

The absorption and emission (annihilation and creation) of photons is presented in Section 4. The electric dipole approximation is discussed and examples are given. Everyone knows that spontaneous emission is responsible for lighting up the world we live in. Yet, this phenomenon cannot be explained theoretically without quantization of the electromagnetic...
field. You will see how this works. Namely, spontaneous emission is nothing more than emission that is stimulated by the zero point fluctuations of the quantized field.

For those of you who took the Math Methods course from me last semester, you undoubtedly recall the part on nanoparticles. There, we discussed *ad nauseam* the difference between retarded electromagnetic waves and the non-retarded regime. The latter is a manifestation of the electric dipole approximation that will be examined in Section 4, and certainly more rigorously than was done in the Math Methods course. We also will see how electric quadrupole and magnetic dipole (and yet higher) terms enter in the expansion of the electromagnetic wave.

Gauge field theory is introduced in Section 5. It is a beautiful theory. We will see how it melds electrodynamics and quantum mechanics into a single package. Whether approached from the electrodynamics or quantum mechanics end, the result is the same: A gauge invariant theory demands the union of quantum mechanics and electrodynamics. Application of gauge symmetry reveals how all of the electrodynamics forces manifest in quantum mechanics.

You might think everything is known about electromagnetism: how it is generated and how it behaves. After all, it has been around for a long time. Yet, some of the most powerful sources of radiation in the universe are poorly understood. For example, Jocelyn Bell discovered pulsars in 1967 while a Physics doctoral student at the University of Cambridge. Susan Jocelyn Bell was born in Belfast, Northern Ireland. She was forbidden to study science in primary school because she was a girl, though her parents' intervention eventually overcame this obstacle. She failed the infamous 11+ exams, but then caught on to physics due to the inspiration of an excellent physics teacher. She subsequently blossomed as a scientist.

Jocelyn Bell and her advisor Anthony Hewish labeled their discovery LGM-1, for *little green men*, as the source of the radiation they detected is $2 \times 10^8$ light years away. It turns out that the radiation emanates from a rotating neutron star. This is truly dense matter: a sphere of neutrons with a diameter of $\sim 20$ km. The density is a staggering $10^{15} \text{ g cm}^{-3}$, and its gravitational field is $10^{11}$ times larger than that of earth. In graphic terms, a mass that is a million times that of the earth is contained in a sphere whose diameter is the length of Manhattan. The amount of radiation is also enormous, with a spectrum that ranges from radio waves to gamma rays. Interestingly, the mechanism responsible for this radiation is still unknown.

Anthony Hewish was awarded the 1974 Nobel Prize in Physics for this discovery. It is unfortunate that Jocelyn Bell did not share this prize. This was at the time, and still is, considered a scandal.
A neutron star is formed when a regular star implodes from gravity after it has burned its fuel. Were the parent star larger, a black hole would be formed. The fact that a black hole radiates at all has only been settled recently, to say nothing of establishing the mechanism. There are also gamma ray bursts and other mysteries out there in the cosmos. Even here on earth, unresolved issues remain, from the infinite energy of the zero-point field to vagaries of the Aharonov-Bohm effect, plus things that appear less sophisticated. For example, a student in Professor Reisler's Chemistry 115 class asked something to the effect: If an atom in an excited state emits a photon, does that mean the photon was inside the atom trying to get out?
1. The Equations of James Clerk Maxwell

In what follows, vectors will be used because everyone can follow the manipulations. You may encounter a kind of mathematics called differential forms if you peruse texts on electromagnetism. This mathematics, though well suited to electromagnetism, has not yet found its way into chemical physics. Besides, we seek registry with special relativity, which is expressed in terms of four-vectors and the like.

We shall start with Maxwell's equations including sources (charge and current densities), and derive a number of relationships and conservation laws. The main goal in this section is to establish the registry between special relativity and electrodynamics, which turns out to be remarkable. They are not separate theories, but a single theory that subsumes both aspects. We will see that the potential fields φ and A emerge as central objects, serving together as a four-vector. Likewise, current density and gradients are cast as four-vectors.

When it comes to quantizing the electromagnetic field (Section 3), delicate issues are avoided because quantization is carried out in the Coulomb gauge using a conjured Hamiltonian. The Coulomb gauge is incompatible with Lorentz invariance, but it gets the job done. This yields the pair of transverse photons with which we are familiar. As mentioned earlier, the other two photons, which are necessary if all components of the four-vector (φ, A) are to be quantized, shall be ignored for the time being. Two excellent texts that cover subtleties encountered when quantizing massless fields are by Guidry [4], and by Aitchison and Hey [3].

Maxwell's equations for vacuum are given below in Heaviside-Lorentz rationalized units. These are like Gaussian units, but with factors of $4\pi$ suppressed.

\begin{align*}
\nabla \cdot \vec{E} &= \rho \\
\nabla \times \vec{E} &= -\partial_t \vec{B} \\
\nabla \cdot \vec{B} &= 0 \\
\nabla \times \vec{B} &= \vec{J}/c + \partial_t \vec{E}
\end{align*}

where $\rho$ is charge density and $\vec{J}$ is current density.

This is the microscopic version of Maxwell's equations. All charges and currents are treated explicitly. Consequently, there are no constitutive relations such as $\vec{D} = \varepsilon \vec{E}$. With Heaviside-Lorentz units, $\vec{H}$ and $\vec{D}$ appear as $\vec{B}$ and $\vec{E}$, and the only constant is $c$. There are no additional constants such as the permeability of free space $\mu_0$ and the permittivity of free space $\varepsilon_0$, nor are there factors of $4\pi$. When convenient, $c$ and $t$ are kept together as $ct$, in keeping with four-vector notation, where $ct = x^0 = x_0$ for a $(+---)$ metric signature.
A particle that carries electric charge experiences the Lorentz force

\[ \vec{F} = q\left(\vec{E} + \left(\vec{v} / c\right) \times \vec{B}\right), \]  

(1.5)

where \( \vec{v} \) is the particle's velocity relative to the frame in which \( \vec{E} \) and \( \vec{B} \) are measured. This force accounts for most of the particle-field interactions that lie at the heart of classical electrodynamics. The self-educated, enigmatic, and wonderfully eccentric British electrical engineer Oliver Heaviside derived eqn (1.5) in 1889, stimulated by a letter correspondence between him and the Irish physicist George Fitzgerald, of "Lorentz-Fitzgerald length-contraction" fame. It was not until 1892 that Hendrik Antoon Lorentz derived the same equation. Nowadays eqn (1.5) is derived straightforwardly starting from special relativity and the definition of electric field that underlies all of electricity and magnetism, namely, \( \vec{F} = q\vec{E} \) [17-21].

**Conserved Quantities**

Let us begin with charge conservation. The Maxwell equation:

\[ \nabla \times \vec{B} = \vec{J} / c + \partial_t \vec{E}, \]

seems like an appropriate place to start. If we take its divergence, its left hand side is automatically equal to zero, because the divergence of the curl of any vector field vanishes identically. On its right hand side, we note the appearance of the term \( \nabla \cdot \partial_t \vec{E} = \partial_t \nabla \cdot \vec{E} = \partial_t \rho \). Thus, taking the divergence of eqn (1.4) yields what is referred to as the continuity equation

\[ \nabla \cdot \vec{J} = -\partial_t \rho . \]  

(1.6)

In fact, this is an example of a general conservation law that states that minus the time rate of change of a scalar volume density is equal to the divergence of its associated flux density. Here it is applied to electric charge density and its associated current density.

When integrated over a volume, the right hand side of eqn (1.6) is equal to minus the time rate of change of the enclosed charge. The left hand side is converted to a surface integral by using the divergence theorem. This shows that the left hand side is equal to the rate with which charge passes through the boundary surface of the volume.\(^1\) Note that all of the vector identities and theorems enlisted in this chapter can be found in the hand-out entitled: Miscellaneous Math.

---

\(^1\) The divergence theorem and eqn (1.6) yield:

\[ \int_V \, dV (\nabla \cdot \vec{J}) = \oint_S \, d\vec{S} \cdot \vec{J} = -\int_V \, dV \partial_t \rho = -q . \]
The above language conjures an image of charge conservation in the non-relativistic regime. Let us now turn to the relativistic version starting with the fact that electric charge is a relativistic scalar. That is, a given amount of charge is the same regardless of the reference frame in which it is measured. The charge density of course varies from one reference frame to another due to length contraction, but not the amount of charge. Indeed, we shall see that the charge density is the timelike component of the four-vector current density, \( J^\nu \).

Referring to Fig. 1, suppose an observer in an unprimed reference frame that we take to be stationary finds that a charge density \( \rho \) passes her in a steady stream. This observer records a larger density than is measured by an observer in the charge density's rest frame. Namely, \( \rho \) is equal to \( \gamma \rho' \), where \( \rho' \) is the density measured in the rest frame.

![Figure 1. Observers measure the same amount of charge in a small volume from reference frames that move at constant speed with respect to one another.](image)

The current density measured by the observer in O is \( \rho \) times the velocity of the moving frame. This is the three-vector \( \vec{J} \) in eqn (1.6). It is interesting that this quantity also can be expressed as \( J^i = \rho' u'^i \), where \( u'^i \) denotes the space part of the relativistic four-velocity introduced in Chapter 1: \( u^\nu = \gamma \partial_\nu x^\nu \), with \( x^i \) the location of the charge density rest frame. Taken together, including the zero component, these facts give the following expression for the relativistic four-vector current density

\[
J^\nu = \rho' u^\nu. \tag{1.7}
\]

This identifies \( J^\nu \) as a four-vector. Again, \( J^\nu \) is measured relative to the stationary frame, \( \rho' \) is measured in the charge density's rest frame, making \( \rho' \) the proper charge density, and the components of the relativistic velocity are \( u^\nu = \gamma \partial_\nu x^\nu \). Length contraction causes a volume element in the moving frame to shrink when observed in the stationary frame. The charge density increases in proportion to this shrinkage, ensuring that the amount of charge remains constant.

Equation (1.7) is usefully written
Chapter 2. Electrodynamics

\[ J^\nu = (\gamma \rho') \partial_\nu x^\nu \]

\[ = \rho \partial_\nu x^\nu \quad (1.8) \]

Recall that the time component, \( J^0 \), is equal to \( \rho c \) [use \( x^0 = ct \) in eqn (1.8)], while the space components are the charge density measured in the stationary frame times the velocity components, relative to the stationary frame, of the frame in which the charge density is at rest.

Taking the four-divergence of eqn (1.8) yields

\[ 0 = \partial_i J^i + \partial_0 (c \rho) \quad (1.9) \]

\[ = \partial_i J^i + \partial_0 J^0 . \quad (1.10) \]

Thus, the conservation law in the relativistic case assumes the form

\[ \partial_\nu J^\nu = 0 . \quad (1.11) \]

The current density \( J^i \) is a contravariant component of a three-vector, whereas \( J^\nu \) is a contravariant component of a four-vector. Minor manipulation has yielded an interesting conclusion. A fundamental conservation principle contained in Maxwell's equations fits neatly into the format of special relativity. The four-divergence \( \partial_\nu J^\nu \) vanishes, and we will see that other four-divergences vanish as well. After all, when considering 4D spacetime, as opposed to 3D space, there is no place for 4D flux to pass into. Thus, it is natural for four-divergences to vanish.

In the following subsection, charge conservation is addressed in a manner that complements the above. We will see that Lorentz transformation not only returns the above result, but it also immediately yields the charges and currents in the two frames. My guess is that you will prefer the following approach.
Chapter 2. Electrodynamics

Charge Conservation and Lorentz Transformation

We have seen that the amount of charge in a given volume is independent of the inertial frame from which it is viewed. It turns out that charge is conserved with respect to any reference frame, inertial or otherwise, though we shall deal only with inertial ones. Referring to Fig. 2, a volume of uniform charge density \( \rho \) is at rest in reference frame \( O \). An observer in \( O \) measures the amount of enclosed charge and finds it to be \( \rho dx dy dz \). Now consider a frame \( O' \) moving with velocity \( \vec{v} \) relative to \( O \). An observer in \( O' \) finds that \( dx \) is contracted by \( \gamma^{-1} \) in which case the volume element \( dx dy dz \) is reduced by the factor \( \gamma^{-1} \). Consequently, the charge density measured by an observer in \( O' \) is larger than that measured by the observer in \( O \). In addition, the observer in \( O' \) finds that current flows.

To see what is going on in more quantitative terms, the current density four-vector is Lorentz transformed from the unprimed frame to the primed frame:

\[
\begin{pmatrix}
J^0 \\
J^i \\
0 \\
0
\end{pmatrix}
=
\begin{pmatrix}
\gamma & -\beta \gamma & 0 & 0 \\
-\beta \gamma & \gamma & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
\rho \\
0 \\
0 \\
0
\end{pmatrix}
=
\begin{pmatrix}
\gamma \rho \\
-\beta \gamma \rho \\
0 \\
0
\end{pmatrix}.
\] (1.12)

In the charge density's rest frame \( O \), there is no three-vector current density, so the four-vector scalar product on Minkowski space is \( J^\nu J_\nu = J^0 J_0 = c^2 \rho^2 \). However, from the perspective of the observer in \( O' \), the charge density's rest frame is moving. Thus, the observer in \( O' \) finds \( J^\nu J_\nu = (c \rho \gamma)^2 - (c \rho \beta \gamma)^2 = c^2 \rho^2 \gamma^2 (1 - \beta^2) = c^2 \rho^2 \), which is the same as the value of \( J^\nu J_\nu \) determined in the charge density's rest frame. This illustrates the constancy of \( J^\nu J_\nu \) with respect to Lorentz transformation. In the present case, \( J^\nu J_\nu \) has the value \( c^2 \rho^2 \). The charge density measured relative to \( O' \) is larger than that measured relative to \( O \), and a three-vector current density is present relative to \( O' \), whereas there is none relative to \( O \).

The charge density measured from \( O' \) has increased by a factor of \( \gamma \). The observer in this frame finds that the amount of charge contained in a given volume, \( i.e. \), the product
of the charge density, $\rho \gamma$, and the volume, $dx'dy'dz' = \gamma^{-1} dx dy dz$, is unchanged from the amount of charge measured by the observer in $O$. Once again we see that the amount of charge is the same whether viewed from within $\rho$'s rest frame or from a frame in which $\rho$'s rest frame is seen as moving.

The above example assumed a stationary charge density in one frame, and it followed that $J^0 J_\nu = c^2 \rho^2$ was conserved. This is not general. If we start with a stationary charge density plus a current density in one frame, it is not, in general, possible to transform to a frame in which only stationary charge is present. This could be achieved if the current density has a specific value relative to the stationary charge density, namely, the $J^1$ entry relative to the $J^0$ entry in the rightmost column vector in eqn (1.12). Except for this special case, it cannot be done. You should convince yourself of this by adding an arbitrary current density $J^1$ to the column vector to the immediate right of the $4 \times 4$ matrix in eqn (1.12) and writing out the transformation.

**Potentials**

Let us now introduce some things called potentials. Because the divergence of any curl is identically zero, the Maxwell equation $\nabla \cdot \vec{B} = 0$ indicates that perhaps $\vec{B}$ can be expressed as the curl of a vector field. We are dealing with three-vector fields, but we shall refer to them as fields, dropping "three-vector" unless required by context. The field $\vec{B}$ appears in two other Maxwell equations: $\nabla \times \vec{E} = -\partial_t \vec{B}$, and $\nabla \times \vec{B} = \vec{J}/c + \partial_t \vec{E}$. These are compatible with expressing $\vec{B}$ as the curl of a vector, so we are free to proceed. In other words, we can define a vector potential $\vec{A}$ through

$$\vec{B} = \nabla \times \vec{A}. \quad (1.13)$$

It is interesting that $\vec{A}$ need not be equal to zero in regions where $\vec{B} = 0$: only $\nabla \times \vec{A}$ must be equal to zero in these regions. Though $\nabla \times \vec{A}$ yields a Lorentz force, we will find that $\vec{A}$ is more important than one might surmise from its role in eqn (1.13).

Using $\vec{B} = \nabla \times \vec{A}$ with the Maxwell equation: $\nabla \times \vec{E} = -\partial_t \vec{B}$, leads to an expression for $\vec{E}$ in terms of potentials.

$$\nabla \times \vec{E} = -\partial_t (\nabla \times \vec{A}) \Rightarrow \nabla \times (\vec{E} + \partial_t \vec{A}) = 0. \quad (1.14)$$

The parenthetic term on the far right can be expressed as minus the gradient of a scalar: $-\nabla \phi$, because the curl of a gradient is identically zero. The minus sign simply adheres to a convention. Thus, the electric field is given by

$$\vec{E} = -\nabla \phi - \partial_t \vec{A}. \quad (1.15)$$

It is usually the case that the potentials $\phi$ and $\vec{A}$ can be obtained more easily than the fields $\vec{E}$ and $\vec{B}$. This bold statement will be justified soon. With the potentials in hand,
the electric and magnetic fields are obtained through differentiation, which is a lot easier than solving a set of coupled partial differential equations to obtain $\vec{E}$ and $\vec{B}$.

A wave equation for $\vec{A}$ is obtained by putting $\vec{B} = \nabla \times \vec{A}$ and $\vec{E}$ from eqn (1.15) into the Maxwell equation: $\nabla \times \vec{B} = \vec{J} / c + \partial_{ct} \vec{E}$, yielding

$$\nabla \times \nabla \times \vec{A} = \vec{J} / c - \partial_{ct}^2 \vec{A} - \partial_{ct} \nabla \phi .$$  \hfill (1.16)

Now use the vector identity: $\nabla \times \nabla \times \vec{A} = \nabla (\nabla \cdot \vec{A}) - \nabla^2 \vec{A}$, to write

$$\partial_{ct}^2 \vec{A} - \nabla^2 \vec{A} = \vec{J} / c - \left( \partial_{ct} \nabla \phi + \nabla (\nabla \cdot \vec{A}) \right) .$$  \hfill (1.17)

Switching the left hand side to four-vector notation yields

$$\partial^\nu \partial_\nu \vec{A} = \vec{J} / c - \nabla \left( \partial_{ct} \phi + \nabla \cdot \vec{A} \right) .$$  \hfill (1.18)

Likewise, inserting $\vec{E} = -\nabla \phi - \partial_{ct} \vec{A}$ into $\nabla \cdot \vec{E} = \rho$ yields

$$\nabla^2 \phi + \partial_{ct} \nabla \cdot \vec{A} = -\rho .$$  \hfill (1.19)

On the previous page it was stated: "It is usually the case that the potentials $\phi$ and $\vec{A}$ can be obtained more easily than the fields $\vec{E}$ and $\vec{B}$." This bold statement needs justification. On what basis can we assume that $\phi$ and $\vec{A}$ will usually be easier to obtain than $\vec{E}$ and $\vec{B}$? After all, were we to stop here and solve the coupled partial differential equations for $\vec{A}$ and $\phi$, it is not obvious that anything would be easier because of the introduction of the potentials $\vec{A}$ and $\phi$. Surely there is more to it.

**Gauge Transformation**

Equation (1.18) would be nice were it not for the parenthetic term on its right hand side, which creates a mathematical mess. Gauge transformation saves the day, however. This is possible because the potentials $\vec{A}$ and $\phi$ are not unique insofar as how they relate to the fields $\vec{B}$ and $\vec{E}$. Namely, they can be altered as long as $\vec{B}$ and $\vec{E}$ are unaffected.

The complete specification of a vector field requires that both its curl and its divergence are known. As mentioned earlier, because $\vec{A}$ is introduced via $\vec{B} = \nabla \times \vec{A}$, the curl of $\vec{A}$ has physical meaning insofar as Lorentz force, whereas its divergence does not. In other words, as far as $\vec{B}$ is concerned, $\vec{A}$ can be adjusted as we see fit, as long as this adjustment leaves $\nabla \times \vec{A}$ alone. Pursuant to this, the gradient of any non-singular scalar function [say $\nabla \zeta (r,t)$] can be added to $\vec{A}$ without affecting $\vec{B}$, because $\vec{B}$ becomes: $\nabla \times \vec{A} + \nabla \nabla \zeta$, and the second term vanishes identically.
Likewise, if \( \partial_{ct} \zeta \) is subtracted from \( \phi \), the electric field is unaffected. Namely, \( \vec{E} \), which started as \( -\nabla \phi - \partial_{ct} \vec{A} \), becomes:

\[
\vec{E} = -\nabla (\phi - \partial_{ct} \zeta) - \partial_{ct} (\vec{A} + \nabla \zeta)
\]

\[= -\nabla \phi + \nabla \partial_{ct} \zeta - \partial_{ct} \vec{A} - \partial_{ct} \nabla \zeta. \tag{1.20}\]

The terms containing \( \zeta \) cancel, leaving \( \vec{E} = -\nabla \phi - \partial_{ct} \vec{A} \), as promised.

Thus, when \( \vec{A} \) is augmented by the addition of the term \( \nabla \zeta \), while at the same time \( \phi \) is augmented by the addition of the term \( -\partial_{ct} \zeta \), neither \( \vec{E} \) nor \( \vec{B} \) is affected.

\[
\vec{A} \rightarrow \vec{A} + \nabla \zeta \quad \text{and} \quad \phi \rightarrow \phi - \partial_{ct} \zeta \tag{1.22}\]

Together these operations are referred to as a gauge transformation – a term introduced by Hermann Weyl. This flexibility, which is referred to as gauge freedom, turns out to be extremely useful, reaching well beyond electromagnetism. In the present context it enables eqn (1.18) to be simplified. The most obvious simplification is to eliminate the parenthetic term. In this case, the wave equation is the d’Alembertian operator, \( \partial^\nu \partial_{\nu} \), acting on the three-vector \( \vec{A} \), with a three-vector source term \( \vec{J} / c \). Referring to eqn (1.18), the condition that needs to be imposed is

\[
\partial_{ct} \phi + \nabla \cdot \vec{A} = 0. \tag{1.23}\]

Let us now see how this works. The issue at hand is whether eqn (1.23) can always be satisfied, thereby ensuring that the parenthetic term in eqn (1.18) can be set to zero. In other words, can it be proven that it is always possible to find a solution to eqn (1.23)? As you might imagine, the answer is yes. The math is given in the box below.

The important point is this: As long as eqn (1.23) can always be satisfied, we need not actually solve it. Instead, eqn (1.18) is solved with the parenthetic term set to zero. This is how the gauge condition enters. If the gauge transformation defined by eqn (1.23) is used, the wave equation given by eqn (1.18) simplifies considerably. The gauge transformation given by eqn (1.23) is called the Lorenz gauge.

When eqn (1.18) is solved with the right hand side equal to \( \vec{J} / c \), eqn (1.23) is automatically subsumed into the solution. There are other gauge transformations, each designed to simplify a class of problems.\(^2\)

\[^2\] The Coulomb gauge: \( \nabla \cdot \vec{A} = 0 \), though not Lorentz invariant, is also used frequently. It is useful when there are no free charges, \( i.e., \rho = 0 \), and when quantization of Maxwell’s equations is carried out to obtain the two transverse photons that we encounter in everyday life.
Lorenz Gauge

Suppose that instead of eqn (1.23) we have

$$\partial_c t \phi_0 + \nabla \cdot \vec{A}_0 = f(\vec{r},t), \quad (i)$$

where $f(\vec{r},t)$ is an arbitrary (non-singular) function, and $\phi_0$ and $\vec{A}_0$ are the original fields, i.e., before any transformation is applied to them. Remember that we are free to transform the potentials as long as $\vec{E}$ and $\vec{B}$ are not affected.

Equation (i) is not equation (1.23) because of the presence of $f(\vec{r},t)$. However, perhaps it can be transformed into eqn (1.23) without upsetting anything. In other words, can $\phi_0$ and $\vec{A}_0$ be altered in a way that maintains $\vec{E}$ and $\vec{B}$, but results in a new pair of potentials $\phi$ and $\vec{A}$ that satisfy eqn (1.23)?

Let us see what happens if $\vec{A}$ is equal to $\vec{A}_0 + \nabla \zeta$, where $\zeta$ is an arbitrary scalar function, while at the same time $\phi$ is equal to $\phi_0 - \partial_c t \zeta$. This leaves $\vec{E}$ and $\vec{B}$ unaffected and eqn (i) becomes

$$\partial_c t (\phi + \partial_c t \zeta) + \nabla \cdot (\vec{A} - \nabla \zeta) = f(\vec{r},t). \quad (ii)$$

This will be equal to eqn (1.23) as long as

$$\partial_c t^2 \zeta - \nabla^2 \zeta = f(\vec{r},t). \quad (iii)$$

This is a standard differential equation, for which a solution must exist as long as $f(\vec{r},t)$ is well behaved. It is important that we do not have to actually solve eqn (iii). Rather, we need only ensure that its solution exists, which we have done, assuming that no physics is encountered that results in $f(\vec{r},t)$ being singular.

Thus, eqn (1.23) can be satisfied, and we are free to toss out the parenthetic term in eqn (1.18).

Notice that this gauge condition is referred to as Lorenz, whereas the relativistic transformation is referred to as Lorentz. This is not a spelling error. As mentioned earlier, the Danish physicist Ludvig Lorenz introduced gauge, though not by this name. The Dutch physicist Hendrik Lorentz (see photo of the 1927 Solvay Conference, which hosted quite a group), together with George Fitzgerald, introduced the transformation used in special relativity that bears his name. Interestingly Fitzgerald and Lorentz carried out their work in support of the "luminiferous aether theory," not Einstein's theory of special relativity.
Putting eqn (1.23) into eqn (1.18) yields the suggestive form

\[ \partial^\nu \partial_\nu \vec{A} = \vec{J} / c. \]  

(1.24)

To go from \( \vec{A} \) and \( \vec{J} \) to four-vector counterparts, an equation is needed that relates \( A^0 \) to the current density component introduced earlier, \( J^0 = c \rho \). Moreover, we know \textit{a priori} that the needed equation must involve \( \phi \), as this is the only potential term that remains. The desired equation is obtained from \( \nabla \cdot \vec{E} = \rho \) by writing

\[ \rho = \nabla \cdot \left( -\nabla \phi - \partial_{ct} \vec{A} \right) \]  

(1.25)

\[ = -\nabla^2 \phi - \partial_{ct} (\nabla \cdot \vec{A}). \]  

(1.26)

Using the Lorenz condition: \( \nabla \cdot \vec{A} + \partial_{ct} \phi = 0 \), on the right hand side of eqn (1.26) yields

\[ \rho = \partial_{ct}^2 \phi - \nabla^2 \phi \]

\[ = \partial^\nu \partial_\nu \phi. \]  

(1.27)
Thus, eqn (1.24) accounts for the three-vector components $A^i$ and $J^i$, whereas $\partial^v \partial_v \phi = \rho$ accounts for $A^0$ and $J^0$. Specifically, defining $\phi = A^0$, and recalling that $J^0 = c\rho$, yields: $\partial^v \partial_v A^0 = J^0 / c$. Putting all of this together yields

$$\partial^v \partial_v A^\mu = J^\mu / c.$$  \hfill (1.28)

Victory at last: Maxwell's equations have been recast in Lorentz covariant form, and the Lorenz gauge is a statement of the vanishing of the four-divergence of $A^v$:

$$\partial_v A^v = 0.$$  \hfill (1.29)

### Coulomb Gauge

The gauge that we will use for quantization of the electromagnetic field is called the Coulomb gauge. It is also called the transverse gauge. It is given by

$$\nabla \cdot \vec{A} = 0.$$  \hfill (1.30)

Applying this to eqns (1.18) and (1.19) yields

$$\partial_c^2 \vec{A} - \nabla^2 \vec{A} + \partial_c \nabla \phi = \vec{J} / c$$  \hfill (1.31)

$$\nabla^2 \phi = -\rho.$$  \hfill (1.32)

In the Coulomb gauge $\vec{A}$ points in the plane perpendicular to the direction of propagation. Thus, the $\vec{A}$ obtained using the Coulomb gauge is referred to as transverse. Recall that a uniform plane wave satisfies $\nabla \cdot \vec{A} = 0$.

All vector fields can be decomposed into transverse and longitudinal components. The transverse component has zero divergence, while the longitudinal component has zero curl (Helmholtz theorem). For example, the current density in eqn (1.31) can be written $\vec{J} = \vec{J}_L + \vec{J}_T$, where $\nabla \times \vec{J}_L = 0$ and $\nabla \cdot \vec{J}_T = 0$. The longitudinal and transverse components are sometimes referred to as irrotational and solenoidal, respectively. Another way of saying this is that to completely specify a vector, it is necessary to specify its curl and its divergence. Note that the Maxwell equations do this, whereas the defining relations for the potentials do not.

The curl of any gradient is zero, so the $\nabla \phi$ term in eqn (1.31) is longitudinal. The benefit derived from decomposing eqn (1.31) into its transverse and longitudinal parts is...
that the vector potential is identified as being entirely transverse and uncoupled from the scalar potential. Thus, the wave equation for transverse $A$ in the Coulomb gauge is

$$\varepsilon \vec{J} / c = \partial_\nu^2 \vec{A} - \nabla^2 \vec{A}$$

(1.33)

$$= \partial_\nu \partial_\nu \vec{A}. \quad (1.34)$$

Exercise: Show that the longitudinal component of the current density in eqn (1.31) together with eqn (1.32) yield the continuity relation.

### Four-Divergences

It remains to show how quantities like four-divergences ($\partial_\nu J^\nu$, $\partial_\nu A^\nu$), the fields $\vec{E}$ and $\vec{B}$, and equations (Maxwell, Lorenz gauge condition, Lorentz force) behave under Lorentz transformation. We shall now approach these one at a time: the four-divergences in this subsection, the others in Section 2. To begin, recall the continuity condition

$$\partial_\nu J^\nu = 0. \quad (1.35)$$

This indicates that there is no net transfer of four-vector current density through a closed spacetime surface. Given that the system is conservative and there is no place for the current density to go, this is hardly surprising.

To demonstrate that the four-divergence of the current density, $\partial_\nu J^\nu$, is Lorentz covariant, the relationship between $\partial_\nu J^\nu$ and $\partial_\nu J^\nu$ is examined. The four-divergence is a scalar, which hints that it may be Lorentz covariant. We just need to make sure it is a Lorentz scalar and not some other kind of scalar.

Given that $J^\nu$ is Lorentz covariant, only $\partial_\nu J^\nu$ needs to be considered. Using the chain rule for differentiation it is seen that $\partial_\nu J^\nu = \partial_\nu J^\nu$. Thus, the four-divergence (continuity equation) is invariant with respect to Lorentz transformation.

We have seen that the Lorenz condition, $\partial_\nu A^\nu = 0$, results in an enticing form of Maxwell’s equations: $\partial_\nu A^\mu = J^\mu / c$. The proof of the invariance of $\partial_\nu A^\nu$ with respect to Lorentz transformation is the same as the one for $\partial_\nu J^\nu$. Because $A^\nu$ and $J^\nu$ are each four-vectors they have the same transformation properties, so if $\partial_\nu J^\nu$ is invariant, then $\partial_\nu A^\nu$ is also invariant. Thus, a modest amount of effort has confirmed that the four-divergences $\partial_\nu A^\nu$ and $\partial_\nu J^\nu$ are each Lorentz invariant, and each vanishes.

---

The partial derivative transforms according to: $\partial_\nu = \frac{\partial}{\partial x^\nu} = \frac{\partial x^\mu}{\partial x^\nu} \partial_\mu = \Lambda^\mu_\nu \partial_\mu$. Also: $J^\nu = \Lambda^\nu_\mu J^\mu$. Putting these together yields: $\partial_\nu J^\nu = \Lambda^\mu_\nu \Lambda^\nu_\mu J^\mu = \partial_\mu J^\mu$.  

137
Taking Stock

From the manipulations carried out so far we see that the potential and the current density are each four-vectors whose divergences vanish. A second order partial differential equation has emerged whose differential operator transforms as a Lorentz scalar, $\partial^\nu \partial_\nu$. This operator acts on the four-vector potential, $A^\mu$. The Maxwell equations were presented at the outset as a set of coupled partial differential equations without much apparent symmetry. They ended as a set of compact expressions that fall neatly into relativistic format.

The original set of Maxwell equations has undergone a remarkable distillation by virtue of their "fit" with special relativity. In Chapter 1 we saw that special relativity is premised on the homogeneity and isotropy of empty space. A universal speed limit follows from this ansatz, and its value of $c$ is obtained through the realization that the speed of light is the same in all inertial frames.

Spacetime is meaningless in the absence of matter, so the term empty space is understood to mean that we are operating in a regime where gravity has a negligible effect, or at least an effect that can be dealt with. But what about $c$ and its constancy, what could be surmised about electromagnetism on this basis?

The existence of light demands waves, and the fact that its speed of propagation is the same in all inertial reference frames requires a mathematical form that is not related to any particular frame. In other words, the wave operator must be a Lorentz invariant. This implicates $\partial^\nu \partial_\nu$ operating on a four-vector.

… finish later
2. Electric and Magnetic Fields

We have seen that the inhomogeneous equation: \(\partial^\nu \partial_\nu A^\mu = J^\mu / c\), is Lorentz covariant, the operator \(\partial^\nu \partial_\nu\) is Lorentz invariant, and \(A^\mu\) and \(J^\mu\) are Lorentz covariant. In this section we will see how to embed \(\vec{E}\) and \(\vec{B}\) into a Lorentz covariant framework, though neither of them is Lorentz covariant.

In transforming \(\vec{E}\) and \(\vec{B}\) between frames, electric fields become magnetic fields and vice versa. For example, in a frame in which the charge density is stationary, there is only an electric field. But if this frame is seen as moving by an observer in another frame, this observer senses a magnetic field in addition to an electric field (Fig. 3).

To obtain expressions for \(\vec{E}\) and \(\vec{B}\) following Lorentz transformation, let us do a bit of algebra and then jump to the result. The system starts in unprimed coordinates and ends in primed ones, with the velocity of the primed frame relative to the unprimed frame given by \(\vec{\nu} = v\hat{x}\). In the unprimed frame, the \(x\)-component of \(\vec{E}\) is given by

\[
E_x = -\partial_1 \phi - \partial_t A^1
\]

\[
= -\partial_t A^0 - \partial_0 A^1. \tag{2.1}
\]

Let us now see what \(E_x\) looks like to the observer in the primed frame. The relevant transformations are
Chapter 2. Electrodynamics

\[
\begin{pmatrix}
A^0' \\
A^1' \\
A^2' \\
A^3'
\end{pmatrix} =
\begin{pmatrix}
\gamma & -\beta \gamma & 0 & 0 \\
-\beta \gamma & \gamma & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
A^0 \\
A^1 \\
A^2 \\
A^3
\end{pmatrix}
\]

(2.2)

Lorentz transformations of contravariant and covariant objects differ only in the signs of the off-diagonal elements.

The above transformations are applied to \( E_x \), yielding

\[
E_x' = -\partial_1 A^0' - \partial_0 A^1'
\]

(2.3)

\[
= - (\gamma \partial_1 + \beta \gamma \partial_0) (\gamma A^0 - \beta \gamma A^1) - (\gamma \partial_0 + \beta \gamma \partial_1) (\gamma A^1 - \beta \gamma A^0).
\]

(2.4)

\[
= - \partial_1 A^0 - \partial_0 A^1
\]

(2.5)

\[
= E_x.
\]

(2.6)

The component of the electric field in the direction of the relative velocity is unchanged. The same kind of manipulation can be carried out for the other components of \( \vec{E} \) and the components of \( \vec{B} \). The results are summarized in eqns (2.7). You should work out the details. Subscripts denoting parallel and perpendicular refer to directions relative to the velocity \( \vec{v} \). Clearly, the direction of \( \vec{v} \) is arbitrary.

\[
\vec{E}_\parallel' = \vec{E}_\parallel \\
\vec{E}_\perp' = \gamma (\vec{E} + (\vec{v} / c) \times \vec{B}) \perp
\]

\[
\vec{B}_\parallel' = \vec{B}_\parallel \\
\vec{B}_\perp' = \gamma (\vec{B} - (\vec{v} / c) \times \vec{E}) \perp
\]

(2.7)

If the charge is stationary in the unprimed frame (\( \vec{B} = 0 \)), eqns (2.7) become:

\[
\vec{E}_\parallel' = \vec{E}_\parallel \\
\vec{E}_\perp' = \gamma \vec{E}_\perp
\]

\[
\vec{B}_\parallel' = 0 \\
\vec{B}_\perp' = - \gamma (\vec{v} / c) \times \vec{E} \perp.
\]

(2.7')

140
Electromagnetic Field Strength Tensor

A useful mathematical object is the anti-symmetric tensor referred to as the electromagnetic field strength tensor, $F^\mu\nu$. It is used to form the Lagrangian for the free electromagnetic field: $-\frac{1}{4} F^{\mu\nu} F_{\mu\nu}$. Its elements are defined as

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu = -F^{\nu\mu}. \quad \text{(2.8)}$$

This form is suggestive. One sees that the $(\mu,\nu)$ combinations $(1,2), (2,3), \text{and} (3,1)$ each give a component of a curl. Further inspection indicates that the elements of the tensor are the six Cartesian field components $E_x, E_y, E_z, B_x, B_y, \text{and} B_z$. You should plug in a few combinations of $\mu$ and $\nu$ to verify that this is true and see how eqn (2.8) works. These facts are nice, but how does one arrive at the form given by eqn (2.8)? What is the reasoning that underlies it?

To see how $F^{\mu\nu}$ comes about, consider the transformed field components and (with no loss of generality) a primed frame moving with velocity $\vec{v} = v \hat{x}$ relative to a stationary unprimed frame. Equations (2.7) in component form are

$$
\begin{align*}
E_x' &= E_x & E_y' &= \gamma (E_y - \beta B_z) & E_z' &= \gamma (E_z + \beta B_y) \\
B_x' &= B_x & B_y' &= \gamma (B_y + \beta E_z) & B_z' &= \gamma (B_z - \beta E_y). 
\end{align*}
\quad \text{(2.9)}
$$

The six independent field quantities $(E_x, E_y, E_z, B_x, B_y, B_z)$ need to be included in the tensor, which can be assigned a $4 \times 4$ matrix representation. Though a $4 \times 4$ matrix has 16 elements, because the tensor is anti-symmetric each diagonal element is zero and there remain six independent off-diagonals. Thus, a means of introducing the six field components into an anti-symmetric tensor is sought. Of course, Lorentz invariance is required: $F^{\mu'\nu'} = \Lambda^\mu_\mu \Lambda^\nu_\nu F^{\mu\nu}$, and we shall write this out for the independent elements: $(0,1), (1,2), (0,3), (1,2), (1,3), \text{and} (2,3)$. This tedious exercise (box on next page) yields expressions for the primed elements in terms of the unprimed elements. These are then brought into registry with eqns (2.9). The results from the box on the next page (plus its extensions) are

$$
\begin{align*}
F^{01'} &= F^{01} & F^{02'} &= \gamma (F^{02} - \beta F^{12}) & F^{03'} &= \gamma (F^{03} + \beta F^{31}) \\
F^{23'} &= F^{23} & F^{31'} &= \gamma (F^{31} + \beta F^{03}) & F^{12'} &= \gamma (F^{12} - \beta F^{02}).
\end{align*}
\quad \text{(2.10)}
$$
Elements of the Field Strength Tensor

Here we see how the elements of the field strength tensor are related to their Lorentz transformed counterparts: \( F^{\mu' \nu'} = \Lambda_{\mu}^{\mu'} \Lambda_{\nu}^{\nu'} F^{\mu \nu} \). Matrix representations of \( F^{\mu \nu} \) and the transformation from unprimed to primed quantities are listed below to help with algebraic manipulations. Keep in mind that \( \Lambda_{\mu}^{\mu'} \Lambda_{\nu}^{\nu'} F^{\mu \nu} \) is not evaluated using matrix multiplication. Let's start with \( F^{0'1'} \) and work it out step-by-step. Fortunately, many terms are zero.

\[
F^{0'1'} = \Lambda_{\mu}^{0'} \Lambda_{\nu}^{1'} F^{\mu \nu}
\]

\[
= \Lambda_{\mu}^{0'} \left( \Lambda_{0}^{1'} F^{\mu 0} + \Lambda_{\nu}^{1'} F^{\mu \nu} \right)
- \beta \gamma
\]

\[
= \Lambda_{\mu}^{0'} (-\beta \gamma) F^{10} + \Lambda_{0}^{0'} \gamma F^{01}
\]

Using \( F^{10} = -F^{01} \), this becomes

\[
F^{0'1'} = \gamma^2 (1 - \beta^2) F^{01}
\]

\[
= F^{01}.
\]

Bookkeeping was not difficult. Let's do another, this time with more stuff on the right hand side at the end. Again, many terms are zero.

\[
F^{0'2'} = \Lambda_{\mu}^{0'} \Lambda_{\nu}^{2'} F^{\mu \nu}
\]

\[
= \Lambda_{\mu}^{0'} \Lambda_{2}^{2'} F^{\mu 2}
\]

\[
= \Lambda_{0}^{0'} F^{02} + \Lambda_{\nu}^{0'} F^{12}
\]

\[
= \gamma F^{02} - \beta \gamma F^{12}
\]

This is enough. Working out the remaining terms is left as an exercise. The six equations that relate primed and unprimed tensor elements are listed in eqns (2.14).
The correspondence between eqns (2.9) and (2.10) is uncanny. Equating terms yields: \( F^{01} = E_x, \quad F^{02} = E_y, \quad F^{03} = E_z, \quad F^{23} = B_z, \quad F^{12} = B_z, \) and \( F^{31} = B_z. \) All fields, primed and unprimed, can be multiplied by a constant factor without changing the transformation properties. Therefore, the above fields are multiplied by \(-1\) to make them consistent with the phase convention introduced with the potentials. Putting all of this together yields the tensor \( F^{\mu \nu} \) and its \( F_{\mu \nu} \) counterpart, whose matrix representations are

\[
F^{\mu \nu} = \begin{pmatrix}
0 & -E_x & -E_y & -E_z \\
E_x & 0 & -B_z & B_y \\
E_y & B_z & 0 & -B_x \\
E_z & -B_y & B_x & 0
\end{pmatrix}, \quad F_{\mu \nu} = \begin{pmatrix}
0 & E_x & E_y & E_z \\
-E_x & 0 & -B_z & B_y \\
-E_y & B_z & 0 & -B_x \\
-E_z & -B_y & B_x & 0
\end{pmatrix}.
\]

(2.11)

The correspondence between field components and tensor elements that enabled us to obtain eqn (2.11) is not unique. The following correspondence also works: \( F^{01} = B_x, \quad F^{02} = B_y, \quad F^{03} = B_z, \quad F^{23} = -E_x, \quad F^{12} = -E_z, \) and \( F^{31} = -E_y. \) This is referred to as the dual tensor, \( G^{\mu \nu}. \) I am not aware of any other correspondences.

\[
G^{\mu \nu} = \begin{pmatrix}
0 & -B_x & -B_y & -B_z \\
B_x & 0 & E_z & -E_y \\
B_y & -E_z & 0 & E_x \\
B_z & E_y & -E_x & 0
\end{pmatrix}, \quad G_{\mu \nu} = \begin{pmatrix}
0 & B_x & B_y & B_z \\
-B_x & 0 & E_z & -E_y \\
-B_y & E_z & 0 & E_x \\
-B_z & E_y & -E_x & 0
\end{pmatrix}.
\]

(2.12)

Among other things, these tensors can be used to form the Lagrangian for free electromagnetic fields, and to express Maxwell's eqns. For example, the inhomogeneous Maxwell equations (with sources) are identical to

\[
\partial_\mu F^{\mu \nu} = J^\nu / c.
\]

(2.13)

Notice that summation is carried out over the index \( \mu, \) but not over the index \( \nu, \) as this latter index appears on each side of the equation. It is not difficult to verify eqn (2.13), but again, be careful with the bookkeeping. Let us work through a couple of these. For \( \nu = 0, \) eqn (2.13) gives

\[
J^0 / c = \rho = \partial_0 F^{00} + \partial_1 F^{10} + \partial_2 F^{20} + \partial_3 F^{30}
= \nabla \cdot \vec{E},
\]

as required. Next, with \( \nu = 1, \) eqn (2.13) gives

\[
J^1 / c = \partial_0 F^{01} + \partial_1 F^{11} + \partial_2 F^{21} + \partial_3 F^{31}
\]

143
\[ = -\partial_{ct}E_\alpha + \left(\nabla \times \vec{B}\right)_\alpha , \]
as required. And so on.

What about the homogeneous Maxwell equations: \[ \nabla \times \vec{E} = \partial_{ct} \vec{B} \text{ and } \nabla \cdot \vec{B} = 0 \] ? These are represented using either the field strength tensor or its dual, \( G^{\mu \nu} \):

\[
\partial_\alpha F_{\beta \gamma} + \partial_\gamma F_{\alpha \beta} + \partial_\beta F_{\gamma \alpha} = 0 \quad (2.14)
\]
\[
\partial_\nu G^{\mu \nu} = 0 \quad (2.19)
\]

In summary, we see that Maxwell's equations are Lorentz covariant even though neither \( \vec{E} \) nor \( \vec{B} \) is a four-vector.

In this section the form of the field strength tensor has been surmised on the basis of a definition, correspondence between transformed fields and transformed tensor elements, and reasonable arguments. The approach is algebraically tedious. A faster route enlists differential geometry. It provides deeper meaning, and removes guesswork and the seemingly \textit{ad hoc} maneuvers of the present section. However, though differential geometry is not unduly difficult, it is not something one picks up in a few days.
Lorentz Force

The Lorentz force law

\[
\vec{F} = q \left( \vec{E} + \frac{\vec{v}}{c} \times \vec{B} \right)
\]  

(2.20)

emerges when the transformation rules of special relativity are applied to the definition of the electric field as the force per unit charge: \( \vec{E} = \vec{F} / q \), and one of the inertial frames is taken to be the charge's rest frame. There are a number of derivations of the Lorentz force in the literature, for example [17-21], not all of which are easy reading. I recommend a relatively recent one that uses familiar concepts and straightforward math [17]: J. H. Field, Derivation of the Lorentz force law, the magnetic field concept and the Faraday-Lenz and magnetic Gauss laws using an invariant form of the Lorentz transformation, Phys. Scr. 73, 639-647 (2006). This paper is well written and accessible to people of varying backgrounds and levels of mathematical preparation. I do not see the point in working through it here, as one of you will be asked to present this paper as a project.

Let us turn instead to the electromagnetic field strength tensor of the previous section. Here we will see how it is used to obtain the Lorentz force, and some important results will be interpreted. The electromagnetic field strength tensor is defined as the exterior (anti-symmetric) derivative of the gauge field

\[
F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu .
\]  

(2.21)

As mentioned earlier, delving into its geometric properties and meaning would require a level of mathematical sophistication that would be out of place in the present context. The best description I know of is given in Doran and Lasenby [22]. They deal with Euclidean metrics on spaces of arbitrary dimension, so you will have to adjust for Minkowski space. Because the field strength tensor is defined rather than derived, it is up to us to relate it to things like Lorentz force, Maxwell’s equations, and so on. The Lorentz four-force \( F^\nu \) is obtained using

\[
F^\nu = \frac{q}{c} U_\alpha F^{\nu\alpha} .
\]  

(2.22)

It is useful to have expressions for \( U_\alpha \) and \( F^{\nu\alpha} \) near at hand for what follows, so these are given below.

\[
F^{\nu\alpha} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \quad U_\alpha = \gamma \begin{pmatrix} \frac{c}{\gamma v_x} \\ -\gamma v_y \\ -\gamma v_z \end{pmatrix}
\]

To see how eqn (2.22) works, let us write out eqn (2.22) meticulously, term-by-term:
\begin{align*}
F^0 &= \frac{q}{c} \gamma \left( U_0 F^{00} + U_1 F^{01} + U_2 F^{02} + U_3 F^{03} \right) \\
&= \frac{q}{c} \gamma \left( v_x E_x + v_y E_y + v_z E_z \right) \\
&= \frac{q}{c} \gamma \tilde{v} \cdot \tilde{E} \tag{2.23}
\end{align*}

\begin{align*}
F^1 &= \frac{q}{c} \gamma \left( U_0 F^{10} + U_1 F^{11} + U_2 F^{12} + U_3 F^{13} \right) \\
&= \frac{q}{c} \gamma \left( cE_x + v_y B_z - v_z B_y \right) \tag{2.24}
\end{align*}

\begin{align*}
F^2 &= \frac{q}{c} \gamma \left( U_0 F^{20} + U_1 F^{21} + U_2 F^{22} + U_3 F^{23} \right) \\
&= \frac{q}{c} \gamma \left( cE_y + v_z B_x - v_x B_z \right) \tag{2.25}
\end{align*}

\begin{align*}
F^3 &= \frac{q}{c} \gamma \left( U_0 F^{30} + U_1 F^{31} + U_2 F^{32} + U_3 F^{33} \right) \\
&= \frac{q}{c} \gamma \left( cE_z + v_x B_y - v_y B_x \right). \tag{2.26}
\end{align*}

**Exercise:** Use the metric $\eta_{\alpha\mu}$ to lower one of the indices of $F^{\nu\alpha}$, yielding the mixed tensor $F^{\nu\mu}$. If you use matrix multiplication be careful of order, as the tensor elements commute whereas matrix multiplication does not. Specifically, when contracting tensor elements, $F^{\nu\alpha} \eta_{\alpha\mu}$ is the same as $\eta_{\alpha\mu} F^{\nu\alpha}$, as $\eta_{\alpha\mu}$ and $F^{\nu\alpha}$ are numbers (as opposed to operators) in the classical theory, whereas multiplication of the corresponding matrices matters. Now contract this with $U^\mu$ to obtain $F^\nu$. Next, create the covariant tensor by applying the metric $\eta_{\nu\epsilon}$ to $F^{\nu\mu}$ in order to obtain $F^{\epsilon\mu}$. Contract this with $U^\mu$ to obtain $F^\epsilon$. Does this $F^\epsilon$ make sense to you? Explain.

Combining eqns (2.24) – (2.26) gives

\[ \vec{F} = q \gamma \left( \vec{E} + \frac{\tilde{v} \times \vec{B}}{c} \right). \tag{2.27} \]

Now write the relativistic three-force $\vec{F}$ as $d\vec{P} / d\tau$, where $\vec{P}$ is the relativistic three-momentum. Using this with $d\tau = dt / \gamma$, eqn (2.27) becomes

\[ \frac{d\vec{P}}{dt} = q \gamma \left( \vec{E} + \frac{\tilde{v} \times \vec{B}}{c} \right). \tag{2.28} \]

Again, keep in mind that the $\vec{P}$ in the above equation is the relativistic momentum $\gamma m \tilde{v}$. The zeroth component of the four-force is $dP^0 / d\tau$. Combining this with eqn (2.23) and using $P^0 = E / c$ and $d\tau = dt / \gamma$ yields
\[ \frac{dE}{dt} = q \vec{v} \cdot \vec{E} . \quad (2.29) \]

The time rate of change of the energy of a particle of charge \( q \) is the power delivered to the particle when it is in the presence of an external electric field \( \vec{E} \). Canceling the two \( dt \)'s (one from \( \vec{v} = d\vec{r} / dt \)) and integrating gives the change in energy

\[ \int dE = \Delta E = \int d\vec{r} \cdot q \vec{E} \quad (2.30) \]

**Reminder**

In Chapter 1 it was shown that the energy of a particle whose speed is \( v \) is \( E = \gamma mc^2 \). This is obtained here using the contraction \( F^\nu U_\nu = 0 \). That this contraction vanishes follows from \( F^\nu = m U^\nu \). In other words,

\[ F^\nu U_\nu = m U^\nu U_\nu = m \frac{1}{2} \frac{d}{dt} (U^\nu U_\nu) = m \frac{1}{2} \frac{d}{dt} c^2 = 0 . \quad (2.31) \]

Expressing \( F^\nu U_\nu \) in terms of momentum (\( F^\nu = dP^0 / d\tau \)) gives

\[ \frac{dP^0}{d\tau} c \gamma = \frac{d\tilde{P}}{d\tau} \gamma \vec{v} . \quad (2.32) \]

Canceling the \( d\tau \)'s, using \( P^0 = E / c \) and \( \tilde{P} = \gamma m \vec{v} \), and integrating yields

\[ E = m \int d(\gamma \vec{v}) \cdot \vec{v} . \quad (2.33) \]

Integrating by parts gives

\[ E = \gamma m \vec{v}^2 - \int \frac{d\vec{v} \cdot \vec{v}}{\sqrt{1-v^2/c^2}} . \quad (2.34) \]

After minor fiddling (see Chapter 1), we get the usual expression

\[ E = \gamma mc^2 . \quad (2.35) \]

This tells us how a particle's speed and energy are related. Equation (2.30) tells us how a charged particle gains energy in the presence of an external electric field.
Example 1. An Accelerated Charge Radiates

A fundamental aspect of electrodynamics is the coupling between electrically charged particles and electromagnetic fields. There are many ways to approach this. A traditional one is to take Maxwell's equations as given and introduce and/or derive particle-field interaction terms such as $\mathbf{p} \cdot \mathbf{A}$, $\mathbf{μ} \cdot \mathbf{E}$, and so on.

In the present example, a toy model is solved in which acceleration (deceleration) of a charged particle results in the emission of a traveling electromagnetic wave that expands outward from the region where the acceleration takes place.\(^4\) The model is classical, albeit with an implicit assumption about the quanta of the freely propagating electromagnetic field, photons. To simplify matters it is assumed that the region where acceleration occurs is small. The goal is a conceptual understanding of how electromagnetic waves arise in a classical context. Purcell gives an excellent treatment [15], and the material is presented in mathematical detail in Jackson [8]. I suggest you start with Purcell.

Figure 4 depicts a situation in which a particle with charge $q$ that has been traveling unimpeded in the +x-direction for a long time is accelerated rapidly in the immediate vicinity of $x = 0$. The mathematical form assigned to this acceleration is not important in the present case, so let's assume that the particle's velocity, $\mathbf{v} = v \hat{x}$, goes to zero linearly with time over a short interval $\tau_0$ (not shown in the figure). In other words, acceleration is constant over the interval $\tau_0$. An alternate scenario, in which a charged particle initially at rest is accelerated over a time interval $\tau_0$ to velocity $\mathbf{v} = v \hat{x}$, yields the same result.

Now consider the electric field that the observer detects. We have seen earlier that the observer will record the presence of a magnetic field in addition to the electric field because the charge $q$ is moving with respect to the observer. This motion occurs in the region $x < 0$ before the intense acceleration brings the particle's motion to a halt. If $v$ is small, the observer finds that the magnetic field due to the moving charge in the region $x < 0$ is sufficiently small that it can be taken as negligible relative to the electric field. However, it turns out that the traveling wave that is produced by the acceleration contains both electric and magnetic components, and the latter is not at all insignificant. In fact, these components contain equal amounts of energy. This is easily proved using Maxwell's equations. Here we shall focus on the electric field.

The electric field lines indicated in Fig. 4 are drawn for the non-relativistic regime, $v << c$. This is easier than drawing them for a relativistic case, and little is lost conceptually.\(^5\) Referring to Fig. 4, the observer has been tracking the particle during its transit along the x-axis on its way to $x = 0$. When the particle arrives at $x = 0$, it ceases its forward motion. Think of it as hitting a wall. However, at the moment of impact the field lines are still present. The observer will not know that something happened at $x = 0$ until a field line emanating from the stationary charge at $x = 0$ arrives. This requires a time $\tau_1 = l_1 / c$. The relativistic property of importance here is the finite speed of light, as manifest in the signals the observer receives before versus after acceleration.

---

\(^4\) The term acceleration includes deceleration.

\(^5\) Recall that I make nearly all of the figures. These are, with few exceptions, sketches, none of which are to scale.
To visualize the situation that exists immediately following impact at $x = 0$, think of the particle as having "cut its ties" with its field lines. The field lines propagate outward at the speed of light. Were the velocity large, the angles of the field lines would reflect this. As it is (small $v$), the observer sees, to a high degree of accuracy, the isotropic distribution of field lines shown on the left in Fig. 4. The observer, being fixed in space, perceives the field lines as rotating toward the vertical. This reflects the fact that the field lines originate from a moving particle. The observer detects the last field line from this group arriving at her location, and she then encounters the precipitous change brought about by the acceleration, where field lines from the stationary charge at $x = 0$ take over. This happens within the relatively short time interval $\tau_0$ over which acceleration takes place.

To visualize the situation that exists immediately following impact at $x = 0$, think of the particle as having "cut its ties" with its field lines. The field lines propagate outward at the speed of light. Were the velocity large, the angles of the field lines would reflect this. As it is (small $v$), the observer sees, to a high degree of accuracy, the isotropic distribution of field lines shown on the left in Fig. 4. The observer, being fixed in space, perceives the field lines as rotating toward the vertical. This reflects the fact that the field lines originate from a moving particle. The observer detects the last field line from this group arriving at her location, and she then encounters the precipitous change brought about by the acceleration, where field lines from the stationary charge at $x = 0$ take over. This happens within the relatively short time interval $\tau_0$ over which acceleration takes place.

Let us now examine the effect brought about through the rapid acceleration the particle experiences near $x = 0$. Specifically, we are interested in how the acceleration is manifest in the electric field that reaches observer. As mentioned earlier, acceleration takes place over a brief period $\tau_0$ with the constant value: $a = v / \tau_0$. Referring to Fig. 5, this creates a thin shell of electric field lines that advances outward at the speed of light.
Within this shell, something important happens to the electric field. In Region I, the electric field is radial: \( E_r = q / r^2 \). It arises from the static charge \( q \) at \( x = 0 \). In Region II, the electric field is also radial, but it arises from charge \( q \) that is in motion and has not yet arrived at \( x = 0 \). At the thin spherical shell, it appears to arise from \( x = v\tau_1 \).

Let us now apply Gauss's law to the surface created by rotating, about the \( x \)-axis, the line that goes from the origin to the inner part of the shell (labeled \( r = c\tau_1 \)). This yields the amount of flux in the solid angle defined by \( \theta \). Gauss's law tells us that \( 4\pi r^2 E_r = q \), and the solid angle defined by \( \theta \) just takes a fraction of this. The value of \( \theta \) is the same for this line as for the one starting at \( x = v\tau_1 \) and going to the outer part of the shell (dashed line in the figure). Consequently, these two field-line segments (red arrow heads) are part of the same field line. In other words, they connect to one another through the shell, as indicated in the box in Fig. 5.

It is now straightforward to obtain an expression for the field component \( E_\theta \). From Fig. 5, we see that

\[
\frac{E_\theta}{E_r} = \frac{v\tau_1 \sin \theta}{c\tau_0}.
\]  

(1)

Referring to Fig. 5 (particularly the white box), we see that \( E_r \) is proportional to the thickness of the shell, \( c\tau_0 \). The component \( E_\theta \) is proportional to \( v\tau_1 \sin \theta \), with the same
proportionality constant as with \( E_r \). Introducing the substitutions: \( a = v / \tau_0 \); \( r = c \tau_1 \); and \( E_r = q / r^2 = q / (c \tau_1)^2 \), into eqn (1) yields

\[
E_\theta = \frac{qa \sin \theta}{c^2 r}. \tag{2}
\]

This is the main result. The field \( E_\theta \) varies as \( r^{-1} \), whereas \( E_r \) varies as \( r^{-2} \). Therefore, integrating \( E_\theta^2 \) over the surface of a sphere gives a value that is independent of \( r \). On the other hand, \( E_r^2 \) varies as \( r^{-4} \), so at large distances it plays no role, whereas the radiated energy associated with the \( E_\theta \) component is constant. The initial kinetic energy of the particle has been converted to the energy contained in the transverse electromagnetic wave. Radiation is peaked at right angles to \( \vec{v} \) uniformly about the azimuth, and going to zero along the direction of \( \vec{v} \).

As mentioned earlier, we have neglected the magnetic field in this calculation. In the next section, however, we will see that a photon contains identical amounts of energy in its electric and magnetic fields. Thus, we have established the photon production rate as a function of the model's parameters.
3. Quantization of Transverse Waves

The equation for a classical transverse electromagnetic wave propagating in source-free space is a homogeneous partial differential equation in which the operator \( \partial^\nu \partial_\nu \) acts on a vector field that satisfies the Coulomb gauge: \( \nabla \cdot \vec{A} = 0 \). This gauge is not Lorentz covariant, as this would require \( \nabla \cdot \vec{A} = -\partial_t \phi \) (Lorenz gauge). Nonetheless, the polarizations of most importance to us – the ones whose quanta have energy \( \hbar \omega \) – are obtained. The other two arise with quantization of the four-vector \( A^\nu \), not just two spatial components. They mediate force but do not transport energy.

It will prove convenient to deal with explicit time and space dependencies. Thus, we shall work with

\[
\partial_t^2 \vec{A} - \nabla^2 \vec{A} = 0. \tag{3.1}
\]

Quantization follows along lines that are essentially the same as those used in Part IV to obtain phonons, plasmons, and polaritons. This earlier work constitutes a roadmap.

To begin, consider a field contained inside a cubic box whose sides have length \( L \). Later we can let \( L \) go to infinity. Right now it is assumed that the field obeys periodic boundary conditions. Except for small boxes, the size and shape of the box is immaterial. The most general form of the field can be expressed as a Fourier expansion over all allowed values of wave vector, as well as over the two independent spatial polarizations of the field. Unit vectors for orthogonal linear polarizations are often labeled \( \hat{e}_1 \) and \( \hat{e}_2 \). Circular or some other polarization could be used as well. For example, \( \hat{e}^\pm = (\hat{e}_1 \pm i \hat{e}_2) 2^{-1/2} \) describes circularly polarized fields, with the usual understanding that real parts are taken at the end of a calculation.

I am going to suppress the use of two polarization labels. Only two independent polarizations are needed, and it is easy to keep this in mind without explicit notation. Later we will enlist a pair of polarizations, but here we will treat one linear polarization component, with the understanding that the other linear polarization component is on equal footing. Two circular polarizations (clockwise and counterclockwise) is another option. To denote this single linear polarization component, \( \hat{e} \) shall be used rather than \( \hat{e}^\pm \), because cares are reserved for operators. Thus, for the field inside the cubic box, the vector potential has the form

\[
\vec{A}(\vec{r},t) = V^{-1/2} \sum_k \left( c_k e^{i(k \cdot \vec{r} - \omega_k t)} + c_k^* e^{-i(k \cdot \vec{r} - \omega_k t)} \right) \hat{e}, \tag{3.2}
\]

where \( c_k \) is an expansion coefficient, \( \vec{k} \) is the wave vector in the direction of propagation, and \( V^{-1/2} \) (\( V \) is volume) is for normalization.

In writing eqn (3.2), it has been possible to move \( \hat{e} \) outside the parentheses because it is common to each term inside the parentheses, \( i.e., \hat{e} = \hat{e}^\pm \). This would not have been possible had circularly polarized waves with \( \hat{e}^\pm = (\hat{e}_1 \pm i \hat{e}_2) 2^{-1/2} \) been used, because \( \hat{e}^\pm \)
≠ ē±*. Equation (3.2) is close in appearance to eqn (39) in Part IV, Chapter 8 [reproduced here as eqn (3.3)], which is the elastic limit displacement field \( \hat{\phi}(\vec{r}, t) \), with explicit \( \hbar \). There, it was promised that the expression for \( \hat{\phi}(\vec{r}, t) \) would prove useful when we got to quantization of the electromagnetic field. Here, we shall make good on that promise.

\[
\hat{\phi}(\vec{r}, t) = V^{-1/2} \sum_{k} (\hbar / 2\omega_{k})^{1/2} \left( \hat{a}_{k} e^{i(\vec{k} \cdot \vec{r} - \omega_{k} t)} + \hat{a}_{k}^{*} e^{-i(\vec{k} \cdot \vec{r} - \omega_{k} t)} \right) . \tag{3.3}
\]

Only minor differences exist between eqns (3.2) and (3.3). These equations provide a preview of how quantization will be carried out. Equation (3.3) is already in operator form, which tells us where we are headed.

To proceed with quantization, we write the energy density of the electromagnetic field for a given mode (\( k \) value) and a single polarization. For example, eqn (3.2) is for all of the allowed \( k \) values, each having one linear polarization. For a single mode, one of the \( k \) values is chosen. The energy density of an electromagnetic wave in vacuum can be expressed in terms of just its electric field (box on the right), as the electric and magnetic contributions have equal time-averaged values. Thus, the single-mode, time averaged energy \( E_{k}^{\epsilon} \) is given by

\[
E_{k}^{\epsilon} = V \langle E_{k}^{2} \rangle_{t} , \tag{3.4}
\]

where \( \langle \cdots \rangle_{t} \) indicates time average and we have used the fact that \( \langle E_{k}^{2} \rangle_{t} = \langle B_{k}^{2} \rangle_{t} \). Using \( \vec{E}_{k} = -\partial_{t} \vec{A}_{k} \), and considering just one of the \( \vec{A}_{k} \) modes on the right hand side of eqn (3.2), yields

\[
\vec{E}_{k} = i(\omega_{k} / c) V^{-1/2} \left( c_{k} e^{i(\vec{k} \cdot \vec{r} - \omega_{k} t)} - c_{k}^{*} e^{-i(\vec{k} \cdot \vec{r} - \omega_{k} t)} \right) \vec{e} . \tag{3.5}
\]

Thus,

\[
\vec{E}_{k} \cdot \vec{E}_{k} = (\omega_{k} / c)^{2} V^{-1} \left( 2c_{k} c_{k}^{*} - c_{k} c_{k} e^{i2(\vec{k} \cdot \vec{r} - \omega_{k} t)} - c_{k}^{*} c_{k}^{*} e^{-i2(\vec{k} \cdot \vec{r} - \omega_{k} t)} \right) . \tag{3.6}
\]
We have yet to invoke quantum mechanics, so $2c_k^*c_k$ could also be written $2c_k^*c_k$ or $c_k^*c_k + c_k^*c_k^*$. Maxwell's equations are based on real fields. Indeed, $\bar{E}_k$ in eqn (3.8) is real, as the parentheses contain a function and its complex conjugate. Averaging $\bar{E}_k \cdot \bar{E}_k$ over time results in the second and third terms in the large parentheses in eqn (3.6) vanishing, assuming that the time interval of concern exceeds $\omega_k^{-1}$, which is a quite short time. Note that for a wavelength of $\lambda = 500 \text{ nm}$, $\omega^{-1} = 2.65 \times 10^{-16} \text{s}$. Taking all of these things into consideration yields a compact expression for the single mode energy $\mathcal{E}_k^*$:

$$\mathcal{E}_k = 2(\omega_k / c)^2 c_k^*c_k^*.$$  \hspace{1cm} (3.7)

**Quantization**

The total energy is obtained by summing the $\mathcal{E}_k^*$ over all allowed values of $k$. We now equate this energy to a Hamiltonian. Formally, a Hamiltonian is obtained from a Lagrangian, so we are taking a shortcut. I have no problem with this, though a purist might object. The expression for the Hamiltonian is

$$H = \sum_k 2(\omega_k / c)^2 c_k^*c_k^*.$$  \hspace{1cm} (3.8)

The expansion coefficient $c_k^*$ is, in general, complex, and it appears together with $c_k^*$ as the product $c_k^*c_k^*$ in the expression for $\mathcal{E}_k^*$. The term $c_k^*c_k^*$ is reminiscent of a number operator, and this is what it becomes when the transition to quantum mechanics is made. Of course, in eqn (3.8) there is no additional $\frac{1}{2}$ (zero point energy) to go along with $c_k^*c_k^*$ because everything is still classical. Let us now carry out the manipulations that yield the quantum mechanical Hamiltonian expressed in second quantization, starting with the definitions

$$Q_k = (c_k + c_k^*) / c$$  \hspace{1cm} (3.9)

$$P_k = -i\omega_k (c_k - c_k^*) / c.$$  \hspace{1cm} (3.10)

Now express $c_k^*$ in terms of $Q_k$ and $P_k$

$$c_k^* = (c / 2\omega_k)(\omega_k Q_k + iP_k).$$  \hspace{1cm} (3.11)

Referring to eqn (3.8), it is straightforward to use the $c_k^*$ given by eqn (3.11) to express the Hamiltonian in terms of $Q_k$ and $P_k$. Substituting $c_k^*$ and $c_k^*$ into eqn (3.8) gives
\[ H = \sum_k \frac{2(\omega_k^2/c^2)(2\omega_k)^2}{2} \left( \omega_k Q_k - iP_k \right) \left( \omega_k Q_k + iP_k \right) \]

\[ = \sum_k \frac{1}{2} \left( \omega_k^2 Q_k^2 + P_k^2 \right). \]

As mentioned above, this expression is obtained regardless of whether we use \( 2c_k^* c_k \), \( 2c_k c_k^* \), or \( c_k^* c_k + c_k c_k^* \), as everything is still classical. Equation (3.13) identifies the system as a harmonic oscillator, with \( Q_k \) and \( P_k \) being the coordinate and momentum fields. Note that Hamilton's equations are satisfied

\[ \frac{\partial H}{\partial Q_k} = -\dot{P}_k, \]
\[ \frac{\partial H}{\partial P_k} = \dot{Q}_k. \]

These yield the equation for a harmonic oscillator

\[ \ddot{Q}_k + \omega_k^2 Q_k = 0. \]

The system is made quantum mechanical by promoting \( Q_k \) and \( P_k \) to operator status. Recall that the math for the Fock space representation was carried out in Part IV, where phonons were introduced and discussed. Let us now examine the commutation relations and the field operator that results from quantization.

**Commutation Relations**

So far the physics has been entirely classical. For example, \( \hbar \) is nowhere to be seen except for eqn (3.3), which has been borrowed from Part IV in anticipation of where the present development is headed. To turn on quantum mechanics, the canonical coordinate and its conjugate momentum are made to satisfy the commutation relation

\[ \left[ \hat{Q}_k, \hat{P}_{k'} \right] = i\hbar \delta_{k, k'}. \]

Equation (3.14) precludes the possibility of having \( \left[ \hat{c}_k, \hat{c}_{k'}^\dagger \right] \) equal to \( \delta_{k, k'} \). Namely, when \( \left[ \hat{c}_k, \hat{c}_{k'}^\dagger \right] \) is evaluated by introducing the expressions for \( Q_k \) and \( P_k \) given by eqns (3.9) and (3.10) into the commutator relation given by eqn (3.14), we find that

\[ \left[ \hat{c}_k, \hat{c}_{k'}^\dagger \right] = \left( e^{2\hbar / 2\omega_k} \right) \delta_{k, k'}. \]
This is a far cry from \( \delta_{\mathbf{k}, \mathbf{k}'} \). Consequently, annihilation and creation operators, \( \hat{a}_{\mathbf{k}} \) and \( \hat{a}_{\mathbf{k}}^\dagger \), respectively, are defined according to

\[
\hat{c}_{\mathbf{k}} = c \left( \frac{\hbar}{2 \omega_{\mathbf{k}}} \right)^{1/2} \hat{a}_{\mathbf{k}}.
\]

In this case, \([\hat{a}_{\mathbf{k}}, \hat{a}_{\mathbf{k}}^\dagger] = \delta_{\mathbf{k}, \mathbf{k}'}\), which is exactly what is needed for an uncluttered Fock space algebra. Let us now introduce the above items into the expression for \( \hat{A}(\bar{r}, t) \) that is obtained by converting eqn (3.2) into its quantum mechanical counterpart. Specifically, using eqn (3.15) with eqn (3.2) yields

\[
\hat{A}(\bar{r}, t) = V^{-1/2} \sum_{\mathbf{k}} \left( \frac{\hbar}{2 \omega_{\mathbf{k}}} \right)^{1/2} \left( \hat{a}_{\mathbf{k}} e^{i(\mathbf{k} \cdot \bar{r} - \omega_{\mathbf{k}} t)} + \hat{a}_{\mathbf{k}}^\dagger e^{-i(\mathbf{k} \cdot \bar{r} - \omega_{\mathbf{k}} t)} \right) c \hat{e} \tag{3.16}
\]

In comparing the above to eqn (3.3)

\[
\hat{\phi}(\bar{r}, t) = V^{-1/2} \sum_{\mathbf{k}} \left( \frac{\hbar}{2 \omega_{\mathbf{k}}} \right)^{1/2} \left( \hat{a}_{\mathbf{k}} e^{i(\mathbf{k} \cdot \bar{r} - \omega_{\mathbf{k}} t)} + \hat{a}_{\mathbf{k}}^\dagger e^{-i(\mathbf{k} \cdot \bar{r} - \omega_{\mathbf{k}} t)} \right) \tag{3.3}
\]

we see that everything has fallen into place.

It seems uncanny that these equations differ only by the factor \( c \hat{e} \). We have, in just a few short pages, obtained the Hamiltonian and the quantum mechanical operator for the transverse electromagnetic field.
Summary for the Quantized Electromagnetic Field

The field is contained in a cube. Sides have length $L$. Periodic boundary conditions are applied.

Classical vector potential, using the Coulomb gauge and a single linear polarization:

$$\vec{A}(\vec{r}, t) = V^{-1/2} \sum_k \left( c_k e^{i(\vec{k} \cdot \vec{r} - \omega_k t)} + c_k^* e^{-i(\vec{k} \cdot \vec{r} - \omega_k t)} \right) \hat{e}$$

$$c_k = (c / 2\omega_k) \left( \omega_k Q_k + iP_k \right)$$

$$Q_k = (c_k + c_k^*) / c$$

$$H = \sum_k \frac{1}{2} \left( \omega_k^2 Q_k^2 + P_k^2 \right)$$

$$P_k = -i\omega \left( c_k - c_k^* \right) / c$$

Single mode electromagnetic energy:

$$\mathcal{E}_k = L^3 \langle \frac{1}{2} (E_k^2 + B_k^2) \rangle_t = L^3 \langle E_k^2 \rangle_t = 2(\omega_k / c)^2 c_k^* c_k$$

$$\hat{H} = \sum_k \frac{1}{2} \left( \omega_k^2 \hat{Q}_k^2 + \hat{P}_k^2 \right) = \sum_k (\omega_k / c)^2 \left( \hat{c}_k^\dagger \hat{c}_k + \hat{c}_k \hat{c}_k^\dagger \right) = \sum \hbar \omega_k \left( \hat{a}_k^\dagger \hat{a}_k + \frac{1}{2} \right)$$

Number valued representation: $|n_k\rangle$, where $n_k$ is number of photons in $k^{th}$ mode.

Photon field operator: $\hat{A}(\vec{r}, t) = \sum_k \sqrt{\frac{\hbar}{2\omega_k V}} \left( \hat{a}_k e^{i(\vec{k} \cdot \vec{r} - \omega_k t)} + \hat{a}_k^\dagger e^{-i(\vec{k} \cdot \vec{r} - \omega_k t)} \right) c \hat{e}$

It is interesting to compare $\hat{A}(\vec{r}, t)$ to the displacement field operator of Part IV:

$$\hat{\phi}(\vec{r}, t) = \sum_k \sqrt{\frac{\hbar}{2\omega_k V}} \left( \hat{a}_k e^{i(\vec{k} \cdot \vec{r} - \omega_k t)} + \hat{a}_k^\dagger e^{-i(\vec{k} \cdot \vec{r} - \omega_k t)} \right)$$
Example 2. Photon Spin

We shall derive and discuss photon spin in this example. The derivation is straightforward and transparent. There are faster ways to arrive at the answer, but they invariably require more sophisticated understanding up front. In this sense, you are unlikely to find them transparent, and for all practical purposes not even faster.

The route followed below starts from a conjured field whose polarization is circular and whose spatial distribution is well suited to the mathematical manipulations. A particularly efficient route will be revealed at the end through an interesting enlistment of hindsight. Namely, a spin density will be identified, and it will be shown that the spatial distribution of the circularly polarized field has nothing to do per se with the wave's spin. Integration of the spin density over the volume of the field works for any spatial distribution, as it must. After all, unruly spatial distributions are common when dealing with beams of radiation.

Let us begin with the ansatz that $\mathbf{A}$ is a uniform plane wave over the circular area indicated in Fig. 6, except of course in the region of the periphery, where it goes to zero in a short distance, namely, the thickness of the blue line. Think of a laser beam of circular cross-section with reasonably uniform intensity over its cross-section except in its edge region. We shall consider a section of the beam whose length is much larger than the wavelength of the radiation – akin to the blue solid cylinder I brought to class to illustrate this volume, though of course much smaller.

As mentioned above, the field's polarization is taken to be circular, and we shall choose the field to circulate in the counterclockwise sense as viewed from the positive $z$-axis looking back at $z = 0$. Its mathematical expression is

$$\mathbf{A} = A_0 \left( \cos(\omega t - kz) \hat{x} + \sin(\omega t - kz) \hat{y} \right).$$

Circular polarization is an obvious choice, given that we are going to derive the spin of a photon. Moreover, the two circular polarizations, clockwise, and counterclockwise, are a convenient basis. We need only solve for one and then write down the solution for the other. Rationalized Lorentz-Heaviside units (no factors of $4\pi$, also $c = 1$) shall be used throughout. This uncluttered notation is fine for what we seek. These missing factors can be imported whenever they are needed.

The electric field is equal to $-\partial_t \mathbf{A}$. Consequently, it lies in the same plane as does $\mathbf{A}$, and it is given by

$$\mathbf{E} = A_0 \omega \left( \sin(\omega t - kz) \hat{x} - \cos(\omega t - kz) \hat{y} \right).$$

Figure 6. A convenient field distribution is used to illustrate photon spin.
The magnetic field is equal to \( \nabla \times \vec{A} \), so it is given by

\[
\vec{B} = \left( \partial_x A_y - \partial_y A_x \right) \hat{z} + \left( \partial_y A_z - \partial_z A_y \right) \hat{x} + \left( \partial_z A_x - \partial_x A_z \right) \hat{y}. \tag{3}
\]

The first term is special in the sense that it points in the z-direction rather than in the xy-plane. It is zero everywhere except on the periphery of the circle. There, the rapid variation of \( \vec{A} \), in which its magnitude drops from \( A_0 \) at the inside edge of the periphery to zero at the outside edge of the periphery, results in a nonzero z-component of \( \vec{B} \). We shall see in due course that the result is independent of the details of how the field approaches zero in the peripheral region.

In Section 3, the properties of transverse photons were derived through quantization of the Maxwell field using the Coulomb gauge: \( \nabla \cdot \vec{A} = 0 \). It is interesting that we are now dealing with a situation in which \( \nabla \cdot \vec{A} \neq 0 \), as indicated in Fig. 7. Perhaps it could be said that we have introduced a different gauge than \( \nabla \cdot \vec{A} = 0 \). Though correct in a literal sense, such pursuit most likely would prove to be a “blind alley,” so we will not go there.

We shall return to the term \( \vec{B}_z \hat{z} \) and the physics it engenders, but right now the Poynting vector \( \vec{E} \times \vec{B} \) is evaluated for the term \( \vec{B}_{xy} = -\partial_x A_y \hat{x} + \partial_y A_x \hat{y} \), just to see that everything is in order. First, \( \vec{B}_{xy} \) is obtained:

\[
\vec{B}_{xy} = -\partial_x A_y \hat{x} + \partial_y A_x \hat{y}
\]

\[
= A_0 \omega \left( \cos(\omega t - kz) \hat{x} + \sin(\omega t - kz) \hat{y} \right) \tag{5}
\]

\[
= \omega \vec{A}. \tag{6}
\]

Note that \( k = \omega \) has been used (i.e., \( c = 1 \)). Next, \( \vec{E} \times \vec{B}_{xy} \) is calculated:

\[
\vec{E} \times \vec{B}_{xy} = A_0^2 \omega^2 \left( \sin(\omega t - kz) \hat{x} - \cos(\omega t - kz) \hat{y} \right) \times \left( \cos(\omega t - kz) \hat{x} + \sin(\omega t - kz) \hat{y} \right) \tag{7}
\]

\[
= A_0^2 \omega^2 \hat{z}. \tag{8}
\]

Notice that \( E^2 = \vec{E} \cdot \vec{E} = A_0^2 \omega^2 \), and recall that the energy density of a wave having a specific value of \( k \) is \( E^2 \). This is as good a time as any to normalize \( E^2 \) to one photon. Namely, the energy density times the volume under consideration (\( V = \pi r^2 L \), where \( L \) is the length of the solid cylinder) is set equal to \( h \omega \). Combining these facts gives
\[
\vec{E} \times \vec{B}_{xy} = E^2 \hat{z}
\]
\[
= \frac{\hbar \omega}{V} \hat{z}.
\] (9)

Given that we are using \( c = 1 \), this expression is the same as that of the wave's momentum density. Try to avoid being confused by the use of \( c = 1 \).

So far everything is as expected. Next, we calculate the subtler part, namely \( \vec{E} \times \vec{B} \hat{z} \). It is nonzero only on the periphery, and it circulates on the periphery. This flow of energy density will be converted to a momentum density, and by operating from the left with \( \vec{r} \times \), to an angular momentum density.

Let us now calculate the Poynting vector \( \vec{E} \times \vec{B} \hat{z} \) for energy density flowing around the periphery of our solid cylinder of radiation:

\[
\vec{E} \times \vec{B} \hat{z} = A_0 \omega (\sin(\omega t - k z) \hat{x} - \cos(\omega t - k z) \hat{y}) \times (\partial_x A_y - \partial_y A_x) \hat{z}.
\] (10)

The rapid drop in \( \tilde{A} \) in passing from inside to outside the peripheral region is contained in the rightmost parentheses.

The \( z \)-component of the magnetic field can be evaluated on the \( x \)-axis without loss of generality, as the system has azimuthal symmetry. In addition, we can assume that the magnitude of the field drops from \( A_0 \) to zero linearly (\( i.e. \), constant slope) over the width of the periphery. This width is labeled \( \delta r \), which is taken as a positive quantity. Thus, on the periphery at the \( x \)-axis we have \( \partial_x A_y = -A_0 \sin(\omega t - k z) / \delta r \) and \( \partial_y A_x = 0 \). Notice that the magnitude of \( B \hat{z} \) is constant on the periphery. Generality is maintained because any smooth variation of how \( \tilde{A} \) approaches zero can be simulated using piecewise linear segments. When the above facts are taken into consideration, eqn (10) becomes

\[
\vec{E} \times \vec{B} \hat{z} = \frac{A_0^2 \omega}{\delta r} (\sin(\omega t - k z) \hat{x} - \cos(\omega t - k z) \hat{y}) \times (-\sin(\omega t - k z) \hat{z})
\] (11)

\[
= \frac{A_0^2 \omega}{\delta r} (\sin^2(\omega t - k z) \hat{y} + \sin(\omega t - k z) \cos(\omega t - k z) \hat{x}).
\] (12)

Taking the time average of \( \vec{E} \times \vec{B} \hat{z} \) yields zero for the term \( \sin(\omega t - k z) \cos(\omega t - k z) \) and \( 1/2 \) for the term \( \sin^2(\omega t - k z) \). Thus, the time-averaged version of eqn (12) is

\[
\vec{E} \times \vec{B} \hat{z} = \frac{A_0^2 \omega}{2 \delta r} \hat{y}.
\] (13)
The $\delta r$ in the denominator of eqn (13) reflects the fact that the $z$-component of the magnetic field depends on how sharply the magnitude of $\vec{A}$ drops from $A_0$ to zero within the peripheral region. Do not be alarmed by a small quantity in the denominator. It will cancel when the peripheral region volume multiplies the angular-momentum density on the periphery.

As mentioned earlier, the Poynting vector $\vec{E} \times \vec{B} \hat{z}$ is converted first to a momentum density and then to an angular momentum density, $\vec{l}_{\text{density}}$, the latter by operating from the left with $\vec{r} \times$. Given that $\vec{r}$ is perpendicular to $\vec{E} \times \vec{B} \hat{z}$ on the positive $x$-axis, the $\vec{r} \times$ operation yields

$$\vec{l}_{\text{density}} = \frac{r A_0^2 \omega}{2 \delta r} \hat{z}.$$  \hspace{1cm} (14)

We now introduce the energy density for a single photon within the solid cylindrical volume: $\hbar \omega / \pi r^2 L$. This is equal to $A_0^2 \omega^2$, so $A_0^2 \omega$ in eqn (14) is replaced with $\hbar / \pi r^2 L$. In addition, we need to use the fact that the volume occupied by the angular momentum density $\vec{l}_{\text{density}}$ (of course confined to the periphery) is equal to $2 \pi r \delta r L$. Combining all of these things yields the angular momentum $\vec{l}$ associated with a single (counterclockwise) circularly polarized photon:

$$\vec{l} = \vec{l}_{\text{density}} (2 \pi r \delta r L)$$

$$= \frac{r \hbar}{2 \delta r \pi r^2 L} (2 \pi r \delta r L) \hat{z}.$$  \hspace{1cm} (15)

Most of the stuff on the right hand side cancels, leaving

$$\vec{l} = \hbar \hat{z}.$$  \hspace{1cm} (16)

Had we started instead with a clockwise circularly polarized photon, $-\hbar \hat{z}$ would have been obtained.

This is a great result. We have derived photon spin through a straightforward application of Maxwell’s equations and the quantum nature of the photon that was derived earlier in this chapter.

$$\vec{s}_{\text{photon}} = \pm \hbar$$  \hspace{1cm} (17)
Recall that at the outset it was stated: "...a spin density will be identified, and it will be shown that the spatial distribution of the circularly polarized field has nothing to do with its spin. Integration of the spin density over the volume of the field works for any spatial distribution, as it must." Let's now see how this works.

**Spin Density**

An important quantity emerges when $\vec{E} \times \vec{A}$ is evaluated using the relations given in eqns (1) and (2). It turns out that $\vec{E} \times \vec{A}$ is equal to the density $A_0^2 \omega \hat{z}$. Multiplication by the volume $\pi r^2 L$ and normalizing the field to one photon (i.e., replacing $A_0^2 \omega$ by $\hbar/\pi r^2 L$) yields the spin, $\hbar \hat{z}$. Thus, $\vec{E} \times \vec{A}$ is identified as the volumetric spin density that exists throughout the field. My guess is that identification of $\vec{E} \times \vec{A}$ as a spin density is not obvious unless you have encountered this material before.

An important feature of the spin density is that it is present *throughout* the wave. It has nothing to do with the periphery. Integration of the spin density over the volume of the field yields the spin, whether the field is uniform or not. The edge circulation we have identified is a *manifestation* of the spin density. It is not itself the spin, but a neat way to visualize a property of the spin density.

Figure 9 shows how this works. It illustrates the fact that the spin density, though non-zero throughout the volume manifests as a current at the edge. Even important scientists have overlooked this point.

![Figure 9](image-url)

Figure 9. A circularly polarized field of uniform amplitude has a square cross section that is divided into 16 smaller squares. (a) Identical currents flow along the edges of the small squares. The arrows are displaced from the edges of the smaller squares for viewing convenience. Each square can be assigned a circulation that we call its spin. Spin density is uniform throughout. There is no net current in the horizontal or vertical directions for the inside edges because of cancellation. For example, red and blue have the same magnitudes but opposite directions. (b) There is no cancellation at the outer edges, so current flows on the periphery. As the small squares become infinitesimal, the spin density remains constant and current continues to flow on the periphery.
Photon spin density follows directly from Noether’s theorem. However, unless you are already familiar with Noether’s theorem, this statement means nothing. We have been fortunate in that the derivation that culminates in eqn (17) yields the spin density without delving into complicated math.

Next, the angular momentum density for one of the small squares in Fig. 9 is calculated by taking \( \vec{r} \times \vec{p}_{density} \). Referring to Fig. 10, we see that \( r = r_0 / \cos \theta \), and the magnitude of the angular momentum density, \( |\vec{r} \times \vec{p}_{density}| \), is equal to \( r p_{density} \cos \theta \). In replacing \( r \) with \( r_0 / \cos \theta \), the two \( \cos \theta \) terms cancel, leaving \( |\vec{r} \times \vec{p}_{density}| = r_0 p_{density} \). It is interesting that this is the same result as would be obtained using an inscribed circle.

No matter how the large square in Fig. 9 is divided, the answer is always the same. To illustrate this, suppose we sum the contributions from the 16 small squares. Figure 10 shows that for each one the circulating momentum density gives an angular momentum density of

\[
\vec{r} \times \vec{p}_{density} = \vec{r} \times (\vec{E} \times \vec{B}_z \hat{z})
\]

\[
= a \frac{A_0^2 \omega}{2} \frac{\hbar}{\delta r} \sin^2(\omega t - k z) \hat{z}.
\]

Averaging the \( \sin^2(\omega t - k z) \) term over time gives \( \frac{1}{2} \), and multiplying by the number of small squares and the peripheral volume of a small square yields

\[
\bar{I} = a \frac{\hbar}{2\delta r} \frac{16 a^2 L}{16 a^2} \sin^2(\omega t - k z) \left( 4a \delta r L \right) \hat{z}.
\]

Thus, the anticipated result is obtained:

\[
\bar{I} = \hbar \hat{z}.
\]

I could continue driving home the point that nuanced differences yield the same result. The last such example is shown in Fig. 11, where the intensity distribution is assigned a
horizontal gradient. It is a simple matter of adding squares to obtain the total spin, again underscoring the consistency of the approach.

In all cases, spin angular momentum is obtained straightaway by integration of the spin density. This is the method of choice.

Figure 11. Two ways to assess total spin: (a) Currents flow along edges of squares. Line thickness indicates currents (magnitudes 1 to 4). Corresponding spins magnitudes range from 1 to 4. Summing the spins for the 16 squares gives a total of 40. (b) Net currents along edges are shown. Four superposed rectangles are identified, each having a current of 1 unit. From right to left their areas are 4 (blue shading), 8, 12, and 16 squares, again giving total spin of 40.
4. Absorption and Emission of Photons

The interaction of radiation with matter means different things to different people. It can be approached from different (complementary) perspectives, applied to broad ranges of environments and phenomena, and modeled using vastly different levels of mathematical rigor. A person who lacks even a cursory understanding of the interaction of radiation with matter cannot be considered educated. In this section we shall interpret the term "matter" to mean a molecule or atom whose eigenstates are obtained either using a Hamiltonian with no externally applied field or through experimental endeavor. The incident electromagnetic field shall be treated as quantized according to the prescription presented and discussed in the previous section. The goal is to introduce the quantized field and note regimes where its quantum nature is important or unimportant.

The annihilation and creation of photons is one of the most important phenomena in the natural sciences. The most common coupling mechanism for energy being exchanged between an electromagnetic field and an isolated molecule is examined below, namely, the electric dipole interaction. It will be made clear how multipole moments higher than dipole (electric quadrupole, magnetic dipole, etc.) arise, including when and why they become important. A central mathematical object of the theory is the matrix element that couples an initial state to a final state. We will write an expression for this matrix element, discuss it, manipulate it into a form you might encounter in your research, and evaluate it in a few examples.

Let us begin by considering the operator for the kinetic energy of an electron in the presence of an electromagnetic field

\[ \hat{T} = \frac{1}{2m} \left( \hat{p} - (q/c) \hat{A} \right) \left( \hat{p} - (q/c) \hat{A} \right), \]

where \( \hat{p} \) is the canonical momentum, \( i.e., -i\hbar \nabla \) in quantum mechanics. The carats on \( \hat{p} \) and \( \hat{A} \) denote vector operators. It is appreciated, of course, that the molecule contains many electrons that interact with one another and with the nuclei, and approximations are inevitable, for example, within the molecular orbital description. Nonetheless, equation (4.1) indicates the coupling of an electron to a transverse electromagnetic field that is responsible for creating and annihilating molecular excitations. Think of laser radiation incident on some unsuspecting molecule.

If you are unfamiliar with eqn (4.1), you can read through Appendix 2: Classical Mechanics Including the Vector Potential, or simply accept eqn (4.1) for the time being. In the next section, we will discuss how the gauge field \( A^\mu \) enters particle dynamics via the elegant route of gauge field theory, culminating in a simple procedure referred to as minimal coupling. Alternatively, maybe I will take a detour and present the minimal coupling stuff before proceeding.

For an electron, \( q = -e \), where \( e \) is \( +1.6 \times 10^{-19} \) Coulombs. The correct sign of the charge in eqn (4.1) is one of those things that can be confusing even though it is not terri-
bly important. When eqn (4.1) is written out, the terms that involve the field (leaving aside the $p^2/2m$ term) are

$$
\hat{H}' = -\frac{q}{2mc} \left( \hat{p} \cdot \hat{A} + \hat{A} \cdot \hat{p} \right) + \frac{q^2}{2mc^2} \hat{A} \cdot \hat{A}.
$$

(4.2)

The term involving the field is written as a perturbation to the Hamiltonian. It is almost always much smaller than terms like $p^2/2m$, electron-electron repulsion, and so on. The part that is quadratic in the field, $\hat{A} \cdot \hat{A}$, is interesting in its own right, but it makes no contribution to a single photon process, so we will not deal with it. To see why this is so, note that $\hat{A} \cdot \hat{A}$ creates or annihilates either 0 or 2 photons, but never 1 photon.

We shall now write a mathematical expression for the matrix element that couples a pair of molecular states and a correlated pair of photon numbers. Recall that photons are field quanta described using a Fock-space representation. In other words, we keep track of photons by counting how many are in each mode. The basic idea that underlies the matrix element is not complicated.

An electromagnetic field that contains $n_k$ photons is combined with some initial molecular state to form the starting point. We stuff everything we know about this starting point into a single ket for bookkeeping. It is hybrid in the sense that it contains a molecular state composed of orbitals together with the number of photons that have a specified energy and wave vector. The perturbation acts on this starting point, creating amplitude in a system composed of $n_k \pm 1$ photons and either a lower or higher energy molecular state. Energy is conserved within some natural linewidth.

The matrix element that corresponds to photon absorption (annihilation) uses the $\hat{H}'$ given by eqn (4.2), but without the $\hat{A} \cdot \hat{A}$ term. This matrix element is strictly for photon absorption:

$$
H_{abs} = \langle \hat{B}, n_k - 1 | \hat{H}' | \hat{A}, n_k \rangle
$$

(4.3)

$$
= \frac{e}{2mc} \langle \hat{B}, n_k - 1 | \left( \hat{p} \cdot \hat{\tilde{A}} + \hat{\tilde{A}} \cdot \hat{p} \right) | \hat{A}, n_k \rangle.
$$

(4.4)

where $m$ is electron mass. Here and hereafter the arrow atop $k$ is to be understood.

The molecular states are assigned labels $\hat{A}$ and $\hat{B}$, with $\hat{A}$ having the lower energy, as indicated in Fig. 12. The ket in eqn (4.4) represents a molecule in state $\hat{A}$ and $n_k$ photons in the electromagnetic field. Again, we will see that the photon energy $\hbar \omega_k$ needs to be nearly equal to the energy difference $E_{\hat{B}} - E_{\hat{A}}$

![Figure 12. Two energy levels and their interactions with electromagnetic fields](image)

166
in order for a transition to occur. The bra in eqn (4.4) represents a molecule in state $\hat{B}$ and $n_k - 1$ photons in the field. As mentioned above, the matrix element in eqn (4.4) describes photon absorption. In the semi-classical theory, in which the electromagnetic field is treated classically, the matrix element for absorption is equal to the matrix element for emission. My guess is that you are more familiar with the semi-classical treatment than the one under consideration here. In most cases that you are likely to encounter there is little difference, though there is one remarkable exception: so-called spontaneous emission, as discussed below.

It is assumed that a single electron is involved in the transition. The spatial extent of the relevant molecular orbitals is nearly always less than a few nm for small molecules. On the other hand, the spatial extent of a typical photon that might be used to cause the transition is much larger than any molecular dimension. Interaction occurs in a region whose spatial extent is that of the molecular orbitals.

The moniker "absorption" is a graphic way to describe the phenomenon in which a light wave passes through a sample and creates molecular excitations. The intensity of the light wave is smaller upon exiting the sample than it was when it entered the sample. Therefore we say that some of the radiation has been absorbed. This terminology is less appealing at the molecular level because a molecule is small compared to the spatial extent of the photon that is annihilated. Moreover, the absorption cross-section for an allowed atomic transition can be much larger than the size of the atom. For example, the absorption cross-section for atomic hydrogen at Lyman-$\alpha$ (121.6 nm) is $\sim 10^3 \text{ Å}^2$. Compare this to the Bohr radius of 0.53 Å.

In eqn (4.4) the term $\hat{p} \cdot \hat{A}$ can be replaced by $\hat{A} \cdot \hat{p}$ because the Coulomb gauge is used. In other words, $\hat{p}$ acting on $\hat{A} \psi_{\alpha}$ gives $(\hat{p} \cdot \hat{A}) \psi_{\alpha} + \hat{A} \cdot (\hat{p} \psi_{\alpha})$. However, $\hat{p} \cdot \hat{A} = -i\hbar \nabla \cdot \hat{A} = 0$ when the Coulomb gauge is enforced. Thus, eqn (4.4) becomes

$$H_{\text{abs}} = \frac{e}{mc} \langle \hat{B}, n_k - 1 | \hat{p} \cdot \hat{A} | \hat{A}, n_k \rangle. \quad (4.5)$$

We now introduce the field operator $\hat{A}$ given by eqn (3.16) in the previous section, which is reproduced here as eqn (4.6).

$$\hat{A}(\vec{r}, t) = \sum_k \sqrt{\frac{\hbar}{2 \omega_k V}} \left( \hat{a}_k e^{i(\vec{k} \cdot \vec{r} - \omega_k t)} + \hat{a}_k^\dagger e^{-i(\vec{k} \cdot \vec{r} - \omega_k t)} \right) \hat{c} \hat{e} \quad (4.6)$$

Notice that this expression contains annihilation and creation operators. Only the annihilation operator gives a nonzero result in the present case, because a photon needs to be removed from the field when $\hat{A}$ operates to its right. Also, a single field mode (i.e., a single photon energy and $k$ value) is being treated. The resulting expression for the coupling matrix element is

$$H_{\text{abs}} = \frac{e}{m} \sqrt{\frac{\hbar}{2 \omega_k V}} \sqrt{n_k} \langle \hat{B} | e^{i\vec{k} \cdot \vec{r}} \hat{p} \cdot \hat{e} | \hat{A} \rangle e^{-i\omega_k t}. \quad (4.7)$$
Chapter 2. Electrodynamics

The field part is no longer present in the bra and ket (except for \( \vec{c} \)), as this part of the matrix element has been evaluated, yielding \( n_k^{1/2} \) and constants. As mentioned above, eqn (4.7) describes photon absorption with concomitant excitation of the molecule from state \( \tilde{A} \) to state \( \tilde{B} \).

Now consider the alternate situation in which the molecule is initially in state \( \tilde{B} \) and again there are \( n_k \) photons in the field. This corresponds to emission – the creation of a photon. The matrix element that couples the initial state \( \tilde{B} \) to the final state \( \tilde{A} \) is

\[
H_{em} = \frac{e}{m} \sqrt{\frac{\hbar}{2\omega_k V}} \sqrt{n_k + 1} \langle \tilde{A} | e^{-i\mathbf{k} \cdot \mathbf{r}} \hat{\mathbf{p}} \cdot \vec{c} | \tilde{B} \rangle e^{i\omega_{k}t}.
\] (4.8)

The difference between the matrix elements \( H_{abs} \) and \( H_{em} \) given by eqns (4.7) and (4.8) is important. The first thing you notice is that they are not related by complex conjugation. This stands in marked contrast to the semi-classical theory, where the field is treated classically and the molecule is treated quantum mechanically. There, \( H_{em} \) is equal to \( H_{abs}^* \). The difference in the quantum case under consideration is due to the factors \( n_k^{1/2} \) and \( (n_k + 1)^{1/2} \). Of course when \( n_k \) is large these are, for all practical purposes, the same. But this is not so when \( n_k \) is small, notably when it is zero.

**Exercise:** Discuss the situation in which \( \hat{a}^\dagger \) operates to the left in eqn (4.5).

**Bosons, Lasers, and so on**

When an ensemble of atoms or molecules has more population in an excited state than in one or more of its lower energy states, there is said to be a population inversion. If there is a nonzero matrix element (for example, electric dipole) between these states an incident electromagnetic field whose frequency matches the separation between the states can stimulate the excited state to emit a photon having the same frequency and wave vector as the incident field.

This is how lasers work. A population inversion is created, and a few stray photons get the process going. Stimulated emission then takes over, feeding photons into one or more cavity (resonator) modes. This can create radiation having many desired characteristics depending on the laser medium and the experimental arrangement: high spectral purity, short temporal pulse duration, high intensity, and so on. The stimulated photons arise from the energy levels of a system of fermions: electrons, protons, and neutrons. However, photons are bosons.
They have no difficulty being in the same mode; in fact they prefer it. The more photons present in a given mode, the larger the stimulated emission rate.

Recent advances in laser technology are impressive: (1) Power in excess of $10^{15}$ Watts has been demonstrated using a "table-top" device. If this radiation is focused to a spot whose diameter is 10 µm, which is not difficult, the electric field at the focus is a staggering $10^5$ Volts nm$^{-1}$. (2) Pulse durations shorter than 70 attoseconds (1 attosecond = $10^{-18}$ seconds) have been achieved. (3) Spectral purity of ~1 Hz out of $10^{15}$ Hz has been achieved. (4) A laser has been constructed using a single atom as the medium. (5) In medicine, miniature robots armed with intense laser radiation that is transported using fiber optics stand to revolutionize surgery. The list goes on and on, and the end is not in sight.

### Spontaneous Emission

Equation (4.8) indicates that the system still makes a transition from the upper state to the lower state when no photons are present in the field. What is called spontaneous emission is emission stimulated by the zero-point fluctuations of the electromagnetic field.

In the Introduction, it was stated: "Everyone knows that spontaneous emission is responsible for lighting up the world we live in. Yet, this phenomenon cannot be explained theoretically without quantization of the electromagnetic field." Equation (4.8) shows that spontaneous emission is a stimulated process that is present when $n_k = 0$. Classically, when the field vanishes the interaction vanishes. Quantum mechanically, there is a non-zero matrix element between states having $n_k = 0$ and $n_k = 1$.

### Perturbation

The spontaneous emission rate is obtained by using the matrix element given by eqn (4.8) with $n_k = 0$ together with Fermi's golden rule. Do not worry if you have not encountered Fermi's golden rule before. It is one of those results that can be obtained easily using time dependent perturbation theory. The expression for the spontaneous emission rate is

$$
\Gamma = \frac{2\pi}{\hbar} |H_{em}(n_k = 0)|^2 \rho(\hbar \omega_k).
$$

(4.9)

It is understood that $E_B - E_A = \hbar \omega_k$. That is, there is a good match between the frequency of the electromagnetic wave and the energy difference between states $A$ and $B$. The density of states, $\rho(\hbar \omega_k)$, is the same as the mode density $g(\omega)$ that was calculated for phonons, but with $\hbar$ inserted to transform the expression from frequency space to energy space. Namely, $\rho(\hbar \omega_k)$ can be taken from Part IV: Chapter 3, Section 5, eqn (15). This result is reproduced below:

$$
g(\omega) = 4\pi V \frac{k^2}{(2\pi)^3} \frac{dk}{d\omega}.
$$

(15)
Chapter 2. Electrodynamics

Here we write $\rho(h\omega_k)$ instead of $g(\omega)$, and we use $dk/dE = 1/\hbar c$. In addition, we do not carry out the integration over angles because the dot product in eqn (4.8) is taken into explicit account. In other words, $d\Omega$ is retained rather than using $4\pi$. Thus, $\rho(h\omega_k)$ is

$$\rho(h\omega_k) = \frac{V\omega^2}{(2\pi)^3\hbar c^3} d\Omega. \quad (4.10)$$

The matrix element in eqn (4.9) is given by eqn (4.8) with $n_k = 0$. It yields an expression that is compact except for constants:

$$\Gamma = \frac{\omega e^2}{m^2 8\pi^2\hbar c^3} \left| \langle \hat{A} | e^{-ik\cdot\hat{r}} \hat{p} \cdot \hat{e} | \hat{B} \rangle \right|^2 d\Omega. \quad (4.11)$$

This can be simplified slightly by introducing the fine structure constant: $\alpha = e^2 / 4\pi\hbar c \approx 1/137$. Further simplification will come soon.

$$\Gamma = \frac{\omega\alpha}{2\pi(mc)^2} \left| \langle \hat{A} | e^{-ik\cdot\hat{r}} \hat{p} \cdot \hat{e} | \hat{B} \rangle \right|^2 d\Omega \quad (4.12)$$

Important features of the emitted radiation can be surmised through examination of the term $\hat{p} \cdot \hat{e}$. It indicates how the polarization $\hat{e}$ of the electric field of the emitted radiation is related to the momentum of the electron. The direction of the radiation's wave vector of course is perpendicular to $\hat{e}$. The radiation propagates outward, entering the solid angle $d\Omega$ in a sine-squared distribution. The torus shape shown on the right is not far off. It is all I could come up with on short notice.

The electron momentum $\hat{p}$ indicates the movement of charge that takes place as the molecule undergoes a transition from one electronic state to another. Suppose the molecule starts in the excited state $\hat{B}$. As amplitude develops in the lower state $\hat{A}$, the system is no longer in an eigenstate, and it acquires an oscillating dipole moment.

The interaction term $\hat{p} \cdot \hat{A}$ can also be expressed as $\tilde{\mu} \cdot \tilde{E}$ (aside from constants), as shown in the subsection that follows. Thus, we see that the electric field of the emitted radiation is aligned with the molecule's transition dipole moment according to $\cos^2 \theta$. It follows that the radiation will be emitted into a lobe that is peaked in directions that are perpendicular to the transition dipole moment. In other words, the radiation intensity follows a $\sin^2 \theta$ distribution.
In summary, expressions for the spontaneous emission rate and the intensity pattern of the photon flux have been derived from first principles. Approximations have been introduced, but these are easily understood, and in most cases they are accurate. In general, terms that can be safely ignored relative to some dominant term take over if for some reason (e.g., symmetry) the otherwise dominant term is excluded. For example, neglect of spin-orbit interaction would become problematic were a transition forbidden except for this term.

When calculating absorption and emission rates, semi-classical formalism can usually be used with fields whose respective strengths are proportional to \( n_k^{1/2} \) and \( (n_k + 1)^{1/2} \). For example, the matrix element for absorption

\[
\langle \hat{B}, n_k - 1 | \hat{p} \cdot \hat{A} | \tilde{A}, n_k \rangle
\]  

(4.5)

can be replaced with

\[
\sqrt{\frac{2\hbar c^2 n_k}{\omega_k V}} \langle \hat{B} | \hat{p} e^{i\hat{k} \cdot \hat{r}} | \tilde{A} \rangle e^{-i\omega t}. 
\]  

(4.13)

When \( n_k \) is large this is accurate, and even for modest \( n_k \) values it still works quite well. There is no difference between \( n_k \) and \( n_k \pm 1 \). As \( n_k \) approaches zero, though, quantum effects appear.

**Electric Dipole Approximation**

Earlier it was mentioned that photon wavelengths that correspond to electronic transitions in molecules of modest size are much larger than the molecular dimensions. This leads to what is called the electric dipole approximation, which is a cornerstone of many different kinds of spectroscopy. These different kinds of spectroscopy span a broad spectral range: from radio waves to the vacuum ultraviolet. Think of it this way. The electromagnetic field is approximately constant over the dimension of the molecule. Therefore, from the perspective of the field, the molecule looks like a small assembly of charges to which it has no choice except to couple.

The electric dipole approximation arises through the terms \( e^{i\hat{k} \cdot \hat{r}} \) and \( e^{-i\hat{k} \cdot \hat{r}} \) that appear in the matrix elements \( H_{abs} \) and \( H_{em} \), respectively. For example, \( e^{-i\hat{k} \cdot \hat{r}} \) can be expanded in a series that converges rapidly because \(|\hat{k} \cdot \hat{r}| \ll 1\):

\[
e^{-i\hat{k} \cdot \hat{r}} = 1 - i\hat{k} \cdot \hat{r} - \frac{1}{2}(\hat{k} \cdot \hat{r})^2 \ldots
\]  

(4.14)

In the electric dipole approximation, only the leading term in this expansion \((i.e., 1)\) is retained, the other terms being of progressively less importance. They give rise to electric quadrupole and magnetic dipole (each arising from the \( \hat{k} \cdot \hat{r} \) term) and higher order
transitions. The unit vector \( \hat{e} \) is moved outside the matrix elements because it has nothing to do with anything inside the matrix elements. For example, in the electric dipole approximation the spontaneous emission rate given by eqn (4.12) becomes

\[
\Gamma = \frac{\omega^2 \alpha}{2\pi(mc)^2} \left| \langle \hat{A} | \hat{p} | \hat{B} \rangle \cdot \hat{e} \right|^2 d\Omega. \tag{4.15}
\]

It is easier to evaluate the matrix element in eqn (4.15) if it is replaced by a matrix element of \( \hat{r} \). To achieve this, use the fact that \( \hat{p} = -(im/\hbar)[\hat{r}, \hat{H}_{\text{mol}}] \), where \( \hat{H}_{\text{mol}} \) acts on the molecule. This equivalence is probably not a priori obvious, but it can be worked out in a few steps. Its use in the matrix element is straightforward:

\[
\langle \hat{A} | \hat{p} | \hat{B} \rangle = -(im/\hbar) \langle \hat{A} | [\hat{r}, \hat{H}_{\text{mol}}] | \hat{B} \rangle \\
= -(im/\hbar)(E_B - E_A) \langle \hat{A} | \hat{r} | \hat{B} \rangle \\
= -im\omega \langle \hat{A} | \hat{r} | \hat{B} \rangle \\
= -im\omega \hat{r}_{AB}. \tag{4.16}
\]

The quantity \( \hat{r}_{AB} \) is often referred to as the electric dipole matrix element, even though it does not contain the charge associated with the dipole. The resulting expression for the emission rate is

\[
\Gamma = \frac{\omega^3 \alpha}{2\pi c^2} \left| \langle \hat{A} | \hat{r} | \hat{B} \rangle \cdot \hat{e} \right|^2 d\Omega \tag{4.13}
\]

Or, using eqn (4.16),

\[
\Gamma = \frac{\omega^3 \alpha}{2\pi c^2} |\hat{r}_{AB} \cdot \hat{e}|^2 d\Omega. \tag{4.13}
\]

You have heard me refer frequently to \( -\hat{\mu} \cdot \hat{E} \), where \( \hat{\mu} = \hat{q} = -e\hat{r} \) (for an electron) is the dipole moment operator and \( \hat{E} \) is the electric field, as the electric dipole operator. We are now in a position to see the correspondence between \( \hat{\mu} \cdot \hat{E} \) and the \( \hat{p} \cdot \hat{A} \) term we have been examining. Let's start with the matrix element:

---

6 In 1D: \([x, \hat{H}_{\text{mol}}] = [x, \frac{1}{2m} p^2 + V(x)]\). However, \(V(x)\) commutes with \(x\), in which case,

\([x, \hat{H}_{\text{mol}}] = \frac{1}{2m} [x, p^2] = \frac{1}{2m}(xp - px) = \frac{1}{2m} 2ihp = ihp / m \).

Multiplication by \(-(im/\hbar)\) gives \(p\).
\[(e / mc) \langle \hat{B} | \hat{p} | \hat{A} \rangle \cdot \hat{A} \quad (4.14)\]

The field is taken as classical, and therefore it is outside the matrix element. It has negligible spatial variation over the dimension of the molecule. Using \( \hat{E} = -\partial_{ct} \hat{A} \), together with \( e^{-i \omega t} \) time dependence, gives \( \hat{A} = (-ic / \omega)\hat{E} \). Now use \( \hat{p} = -(im / h)[\hat{r}, \hat{H}_{\text{mol}}] \).

The matrix in eqn (4.14) becomes

\[
\left( \frac{e}{mc} \right) \left( \frac{-ic}{\omega} \right) \left( \frac{-im}{\hbar} \right) (E_{\hat{A}} - E_{\hat{B}}) \langle \hat{B} | \hat{r} | \hat{A} \rangle \cdot \hat{E} = e \langle \hat{B} | \hat{r} | \hat{A} \rangle \cdot \hat{E} \quad (4.15)
\]

\[
= -\langle \hat{B} | -e\hat{r} | \hat{A} \rangle \cdot \hat{E} \quad (4.16)
\]

\[
\hat{\mu} \cdot \hat{E}
\]

The right hand side is recognized as the matrix element of \( -\hat{\mu} \cdot \hat{E} \). Thus, the correspondence between \( \hat{p} \cdot \hat{A} \) and \( \hat{\mu} \cdot \hat{E} \) is established.
5. Introduction to Gauge Field Theory

According to the rules of non-relativistic quantum mechanics the probability of finding a particle whose wave function is $\psi(\vec{r}, t)$ in an infinitesimal 3D volume $d^3r$ is given by $|\psi(\vec{r}, t)|^2 d^3r$. It follows that the wave function $\psi(\vec{r}, t)$ can be multiplied by a phase factor $e^{i\zeta(\vec{r}, t)}$, where $\zeta(\vec{r}, t)$ is an arbitrary real scalar function of position and time, without affecting this probability. Although it is not obvious a priori, we shall see that $\zeta(\vec{r}, t)$ needs to be sufficiently well behaved to be differentiable. Other than meeting this condition, which is not asking much, there are no restrictions on $\zeta(\vec{r}, t)$.

At the same time, Hermitian operators that act on $\psi(\vec{r}, t)$ yield observables that are likely to be affected profoundly by the phase transformation $\psi(\vec{r}, t) \rightarrow \psi(\vec{r}, t)e^{i\zeta(\vec{r}, t)}$, notably the particle's momentum. Therefore, if acting alone, this transformation is unacceptable, and egregiously so. Nonetheless, the simple fact remains that in order to locate the particle $\psi(\vec{r}, t)e^{i\zeta(\vec{r}, t)}$ does just as well as $\psi(\vec{r}, t)$. The fact that $|\psi(\vec{r}, t)e^{i\zeta(\vec{r}, t)}|$ is equal to $|\psi(\vec{r}, t)|$ is a redundancy in the description of nature. An infinite number of wave functions each yield the same result when it comes to locating the particle.

The fact that multiplication of a particle's wave function by $e^{i\zeta(\vec{r}, t)}$ affects the mathematical determination of the particle's momentum is a clue to the strategy that will be used to accommodate the phase transformation without compromising the physics. The idea that underlies this strategy is that nature has no such problem with phases, and we are simply catching up with our mathematical models. In the derivation that follows, we shall accept the redundancy that is manifest in $|\psi(\vec{r}, t)| = |\psi(\vec{r}, t)e^{i\zeta(\vec{r}, t)}|$ as something that is inherent to quantum mechanics. Our task therefore is to ensure that the wave function’s phase transformation $\psi(\vec{r}, t) \rightarrow \psi(\vec{r}, t)e^{i\zeta(\vec{r}, t)}$ incurs no problems with the physics. In other words, it should be possible to carry out this transformation without the mathematical description going crazy. In turn, whatever remedy is enlisted is expected to reveal important features of the system that might have been overlooked otherwise. Another way to say this is that the phase transformation participates in a symmetry operation, but there must be another part. We seek the overall symmetry operation that leaves the physics invariant.

To ensure that nothing goes awry, a field that is already present – and of course embedded in the mathematical description – must change synchronously with the phase transformation $\psi(\vec{r}, t) \rightarrow \psi(\vec{r}, t)e^{i\zeta(\vec{r}, t)}$. We shall see that this field turns out to be the four-vector potential $A^\mu$ that yields the magnetic and electric fields according to $\vec{B} = \nabla \times \vec{A}$ and $\vec{E} = -\nabla \phi - \partial_t \vec{A}$.\footnote{Here we shall use the convention $c = \hbar = 1$.}

The field $A^\mu$ is referred to in the present context as the gauge field. This marriage of quantum mechanics and electrodynamics is the goal of this section. This will be achieved using a classical four-vector $A^\mu$ and the Schrödinger equa-
tion. It is equally valid for the fully quantized electromagnetic field (including longitudinal and timelike photons) and the Dirac field. We shall not go there, but it is nice to know that our brief foray opens the door to these more advanced topics, and even beyond, to the gauge groups of the strong, weak, and electroweak interactions. The importance of gauge symmetry and the gauge principle cannot be overstated.

Recall that electromagnetism is also a redundant theory. Adding the gradient of a scalar (say $\nabla \chi$) to $\tilde{A}$ leaves $\tilde{B}$ unaltered because the curl of any gradient is identically zero. In the corollary scalar field transformation, adding $-\partial_\tau \chi$ to $\phi$ leaves $\tilde{E}$ unaltered because $\tilde{E} = -\nabla \phi - \partial_\tau \tilde{A}$ is changed to $-\nabla (\phi - \partial_\tau \tilde{A}) - \partial_\tau (\tilde{A} + \nabla \chi)$, which is equal to $-\nabla \phi - \partial_\tau \tilde{A}$. In other words, before and after are the same.

The phase transformation of the wave function: $\psi(\vec{r},t) \rightarrow \psi(\vec{r},t)e^{i\zeta(\vec{r},t)}$, is referred to as a gauge transformation of the first kind, while the transformation of the gauge fields $\tilde{A}$ and $\phi$ (together $A^\mu$) is referred to as a gauge transformation of the second kind – labels introduced by Wolfgang Pauli. We shall see that the redundancy of the wave function and the redundancy of the gauge field $A^\mu$ fit together perfectly. As long as they act synchronously, an infinite number of wave functions and an infinite number of gauge fields yield correct answers.

The first use of gauge transformations in physical theory was when the Danish physicist Ludwig Lorenz applied them to electromagnetism in the 1860's. Hermann Weyl introduced the term gauge transformation in a model that proved to be incorrect. Yet, the label gauge transformation survived.

As stated by Guidry [4]: "The root of a symmetry principle is the assumption that certain quantities are unobservable; this in turn implies an invariance under a related mathematical transformation, and the invariance under this transformation (if it is unitary, as is usually the case in quantum mechanics) implies a conservation law or selection rule." The symmetry principle under consideration here – that of gauge symmetry – differs in an important way. Rather than a conservation law or selection rule, it reveals the forces of electrodynamics, and it tells us something about the strange substance called electric charge. The latter turns out to be conserved for a given particle type, and only manifests as an integer multiple of a unit charge. It is one of nature's truly good quantum numbers.

Gauge field theory is central to the theory of fundamental particles referred to as the standard model of physics. In this section we shall introduce and examine the synchronous gauge transformations of the first and second kind that arise in particle dynamics. A wave function is multiplied by a phase factor whose argument depends on spacetime coordinates, while a gauge field that is present undergoes a transformation that cancels nonphysical effects incurred through multiplication by the phase factor. This is the basis of quantum electrodynamics (QED), with the understanding that QED is a fully quantized theory – both Maxwell and fermion fields. The basic strategy can be extended to other gauge symmetries, notably in the description of the strong and weak nuclear forces.
Consequences of Redundancy

Given a particle wave function \( \psi(\vec{r}, t) \), it seems intuitively obvious that multiplication by \( e^{i\alpha} \), where \( \alpha \) is a real constant, does not change the calculated outcome of a measurement. Indeed, this invariance with respect to a global phase transformation is taught as unassailable dogma. Yet, the operation of multiplication by \( e^{i\alpha} \) takes effect instantaneously throughout all of space. One might ask how a wave function’s phase convention can be established at different points in space at the same time, as this seems to be inconsistent with special relativity. After all, the wave function describes a particle that leads a causal existence, so simultaneity cannot be invoked. Thus, something as apparently innocuous as multiplication by \( e^{i\alpha} \) is seen to be not entirely trivial. Such issues led scientists to think about local phase transformations.

Now consider the expression \( |\psi(\vec{r}, t) e^{ia\zeta(\vec{r}, t)}| = |\psi(\vec{r}, t)| \). The real constant \( a \) is appended to \( \zeta(\vec{r}, t) \). For the time being, think of this constant as a matter of (future) notational convenience. Later we will see that there is more to it than simply notational convenience. Hereafter, the parentheses on \( \psi(\vec{r}, t) \) and \( \zeta(\vec{r}, t) \) are dropped (to be understood) in favor of a more compact notation. As mentioned earlier, in order to ensure that \( \psi \rightarrow \psi e^{ia\zeta} \) leaves the description of the physical system unaltered, simultaneous change must occur elsewhere in the mathematics, as multiplication by \( e^{ia\zeta} \) itself incurs dire consequence. Moreover, this simultaneous change must reflect an entity that is of physical significance. The bottom line is that without a partner transformation \( \psi \) cannot be gauge transformed.

We shall now obtain a Schrödinger equation that is compatible with the gauge transformation \( \psi \rightarrow \psi e^{ia\zeta} \). To begin, consider the Schrödinger equation for a free particle

\[
\left( -\frac{1}{2m} \nabla^2 - i \partial_t \right) \psi = 0. \tag{5.1}
\]

Despite the fact that an incorrect result is destined to follow, we shall nonetheless replace \( \psi \) with \( \psi e^{ia\zeta} \), just to see what we are up against. This yields:

\[
e^{ia\zeta} \left( -\frac{1}{2m} (\nabla + i a \nabla \zeta)^2 + a \partial_t \zeta - i \partial_t \right) \psi = 0. \tag{5.2}
\]

---

8 The brute force approach is used starting with

\[
\nabla^2 \psi = \nabla \cdot \nabla (\psi e^{ia\zeta}) = \nabla \cdot \left( e^{ia\zeta} \left( \nabla + i a \nabla \zeta \right) \psi \right).
\]

The vector identity: \( \nabla \cdot (f(\vec{r}) \vec{Z}) = \vec{Z} \cdot \nabla f(\vec{r}) + f(\vec{r}) \nabla \cdot \vec{Z} \), yields

\[
\nabla \cdot \left( e^{ia\zeta} (\nabla + i a \nabla \zeta) \psi \right) = e^{ia\zeta} \left( \nabla \psi \cdot \nabla \zeta + \nabla^2 \psi + i a \nabla \zeta \cdot (i a \nabla \zeta) \psi + \nabla \psi + (\nabla \cdot i a \nabla \zeta) \psi \right).
\]

When \( (\nabla + i a \nabla \zeta)^2 \psi \) is multiplied out, we find that it is equal to the contents of the large parentheses on the right side.
At this point, the multiplicative factor $e^{ia\zeta}$ can be cancelled. However, we shall see that the correct Schrödinger equation is multiplied by $e^{ia\zeta}$ (transforms gauge covariantly) when the wave function is gauge transformed according to: $\psi \rightarrow \psi e^{ia\zeta}$. Thus, the factor $e^{ia\zeta}$ shall be retained.

Because of the terms $iaV\zeta$ and $a\partial_t\zeta$, eqn (5.2) describes a different physical situation than does eqn (5.1). Therefore these terms need to be eliminated. This requires that the Schrödinger equation already contains terms that make it possible to bring about the needed cancellation. The phase transformation of the wave function cannot simply be accompanied by additions of $-ia\nabla\zeta$ inside the small parentheses, and $-a\partial_t\zeta$ inside the large parentheses, as this simply eliminates the phase transformation altogether. The only option is that separate entities that are already present in the Schrödinger equation are altered in concert with the phase transformation.

Referring to eqn (5.2), to eliminate $ia\nabla\zeta$ a vector field must be present, whereas to eliminate $a\partial_t\zeta$ a scalar field must be present. Dealing first with $ia\nabla\zeta$, it is necessary that $\psi \rightarrow \psi e^{ia\zeta}$ be accompanied by a change in a vector field whose redundancy is in registry with that of the wave function. Thus, $-ia\nabla\zeta$ is added to $\nabla$. This yet-to-be-specified vector field is a function of $r$ and $t$. The multiplicative factor $-ia$ is included here for convenience, as seen below.

The Laplacian $\nabla^2$ in eqn (5.1) is now replaced with $(\nabla - ia\nabla)^2$. We see that the $ia\nabla\zeta$ term in eqn (5.2) is eliminated through the addition of $\nabla\zeta$ to $\nabla$. In other words, $\psi \rightarrow \psi e^{ia\zeta}$ must go hand-in-hand with $\nabla \rightarrow \nabla + \nabla\zeta$. Applying these transformations to $(\nabla - ia\nabla)\psi$ results in the needed cancellation:

\[
(\nabla - ia(\nabla + \nabla\zeta))\psi e^{ia\zeta} = e^{ia\zeta}(\nabla - ia\nabla)\psi .
\]

Multiply wave function by phase factor.

Add gradient to existing vector field.

Operating again with $(\nabla - ia(\nabla + \nabla\zeta))$ yields $e^{ia\zeta}(\nabla - ia\nabla)^2\psi$.

The same strategy is now applied to the $a\partial_t\zeta$ term in eqn (5.2). Namely, in order to eliminate it, a scalar field must be present. Thus, $a\eta$ is added to the parentheses in eqn (5.1), and $\eta$ is altered according to: $\eta \rightarrow \eta - \partial_t\zeta$. This results in the needed cancellation of the term $a\partial_t\zeta$ in eqn (5.2). In summary, when $-ia\nabla$ and $a\eta$ are included in eqn (5.1) it becomes

\[
\frac{1}{2m}(\nabla - a\nabla)^2 + a\eta - ia\partial_t \psi = 0 .
\]

This equation is said to be gauge covariant (form invariant) because it transforms as the wave function, i.e., when $\psi \rightarrow \psi e^{ia\zeta}$, the entire equation is multiplied by $e^{ia\zeta}$. Note that gauge covariant is not the same as Lorentz covariant.

Our mathematical description of nature cannot be such that viewing a physical situation from a different spacetime reference frame yields truly different results. For exam-
ple, we have seen that charge density in one frame can appear as current density in another frame. There is no difference other than perspective, and everything fits together nicely with four-vectors. For example, if the phase \( a\zeta(r,t) \) that is added through a phase factor to the wave function in eqn (5.2) is taken as \( a\xi(r,t) = \vec{k} \cdot \vec{r} \), the wave function has apparently acquired momentum. However, we see that the kinetic momentum is altered accordingly such that it experiences no net change.

**Link with Electromagnetism**

The assignments \( \vec{V} = \vec{A} \), \( \eta = \phi \), and \( a = q \) are now made. The magnitude of the charge \( q \) is proportional to the strength of the coupling between the particle and the electromagnetic field, and its sign determines the direction of the force. If the particle does not have electric charge this coupling is zero, so forget about applying the gauge transformation discussed above to the wave function.

As mentioned at the outset, the registry between quantum mechanics and electromagnetism is uncanny. The fields \( \vec{B} \) and \( \vec{E} \) are obtained from \( \vec{A} \) and \( \phi \) according to \( \vec{B} = \vec{\nabla} \times \vec{A} \) and \( \vec{E} = -\vec{\nabla} \phi - \partial_t \vec{A} \). Adding \( \nabla \zeta \) to \( \vec{A} \) leaves \( \vec{B} \) unaffected, while at the same time adding \( -\partial_t \zeta \) to \( \phi \) leaves \( \vec{E} \) unaffected. Thus, the redundancy of the fields needed to satisfy quantum mechanical gauge covariance matches perfectly the redundancy of electromagnetism. How cool is that.

The group under consideration is referred to as \( \text{U}(1) \), which stands for unitary group of dimension one. It is commutative (referred to by mathematicians as Abelian) because \( e^{i\xi_1} e^{i\xi_2} = e^{i\xi_2} e^{i\xi_1} \). You are already familiar with generators of unitary transformations. For example, \( H \) is the generator of time evolution via \( e^{-iHt} \). Here we have the unitary transformation \( e^{i\xi} \). Thus, it can be said that electric charge is the generator of the \( \text{U}(1) \) electrodynamics gauge transformation. This conclusion about charge being a generator of gauge transformation is, in fact, more general. It applies to weak and strong charges, though their gauge symmetries are more involved than multiplication by a phase factor.

The charge \( q \) is sometimes referred to as the electrodynamics gauge-coupling constant, because of the role it plays. In the present context this would confuse rather than enlighten, so we will stick to charge. Nonetheless, it is significant that charge arises naturally in the context of a gauge transformation. It follows from the symmetry principle that links particle dynamics to fields through their complementary redundancies. The net effect is that the gauge covariant Schrödinger equation for the particle is

\[
\left( \frac{1}{2m} \left( -i\hbar \nabla - (q/c)\vec{A} \right)^2 + q\phi - i\hbar \partial_t \right) \psi = 0, \tag{5.5}
\]

where explicit \( \hbar \) and \( c \) have now been included.

The overall gauge transformation affects both the wave function and the gauge field \( A^\mu \). Not only have we arrived at the four-vector \( A^\mu \), but again – as in relativistic electrodynamics – it has emerged as being more fundamental than \( \vec{E} \) and \( \vec{B} \). This is not surprising in light of its registry with relativity. For a long time the potentials \( \vec{A} \) and \( \phi \) were
believed to be mere conveniences for obtaining $\vec{E}$ and $\vec{B}$. This is what I was taught in school, and I fell for it.

The above result is profound though in hindsight not surprising. Particles that carry electric charge emanate fields, so it is not possible to have such particles in the absence of electromagnetism. Yet, the facile way in which the gauge field enters the Schrödinger equation and the synchrony of the gauge transformations of the first and second kind is impressive. The requirement of gauge covariance has eliminated the possibility that the particle can be free. Said differently, in order that a particle's wave function can undergo a U(1) gauge transformation the particle must carry electric charge.

No fundamental particle can be free. It must couple to at least one field if we are to know of its existence. All known fundamental particles carry charge of one or more kinds (electric, weak, strong) that enable them to couple to their respective fields. The gauge symmetry that has been discussed in this section establishes all of the forces of electrodynamics. Using this as a roadmap enables the same idea to be extended, on the basis of the principle of gauge invariance, to all of the known forces of nature other than gravity.

**Visual Aid**

A schematic illustration of the gauge field and how gauge invariance works is given below. The plane denoted $x^\mu$ is a schematic representation of 4D spacetime. From each spacetime point a strand called a fiber extends upward. In (a) the blue dots represent the gauge field associated with spacetime points. The gauge field in general varies in spacetime, so the points are shown at different heights. It is understood that the spacetime points are densely packed, as they must be if they are to represent a gauge field that is continuous on spacetime. The collection of fibers is called a bundle. This nomenclature is amusing. The mathematical term fiber bundle brings to mind the bundles of fibers that are used in fiber-optic communications.

This figure is highly schematic. (a) The gauge field $A^\mu$ is distributed over spacetime. Its value at each spacetime point is indicated by a blue dot above the spacetime plane, $x^\mu$. (b) A local gauge transformation changes, in a correlated manner, both the gauge field and the phase factor of a state vector. The former slides on its fiber; the latter is represented by red dots and yellow circles. Adapted from K. Huang: *Fundamental Forces of Nature: The Story of Gauge Fields*. 
Example 3. Aharonov-Bohm

It is straightforward to verify (for example, by direct substitution and using the Leibniz rule for differentiation of an integral) that the solution to eqn (5.5) has the general form

\[ \psi = \psi_{A=0} \exp \left( iq \int_{\vec{r}_0}^{\vec{r}} d\vec{r} \cdot \vec{A} \right), \tag{1} \]

where \( \psi_{A=0} \) is the solution to eqn (5.5) with \( \vec{A} \) set equal to zero, and \( c = \hbar = 1 \) is used. The analogous exponential factor that contains the integral of \( dt \phi \) has been subsumed into \( \psi_{A=0} \) for the time being. Right now we shall deal with the magnetic part. It is assumed that \( \vec{A} \) is static relative to a laboratory reference system, and that the particle travels slowly enough that it does not perceive significant time variation of \( \vec{A} \).

Suppose we choose to eliminate \( \vec{A} \) through the gauge transformation \( \nabla \zeta = -\vec{A} \). This sounds attractive, but we need to know whether it is possible. For this choice to be viable \( \vec{A} \) must not be associated with a local magnetic field. In mathematical terms, if \( \vec{A} \) is to be expressed as the gradient of a scalar, it is necessary that \( \nabla \times \vec{A} = 0 \). In regions where \( \nabla \times \vec{A} = 0 \), the value of the integral of \( d\vec{r} \cdot \vec{A} \) between two points is independent of the path. It depends only on the end points. Namely, \( d\vec{r} \cdot \nabla \zeta = d\zeta \) integrates trivially, yielding \( \zeta(\vec{r}) - \zeta(\vec{r}_0) \).

Figure 1 illustrates the fact that in regions where \( \nabla \times \vec{A} = 0 \) paths can be distorted without affecting values of line integrals. Alternatively, in regions where \( \nabla \times \vec{A} \neq 0 \) (\( \vec{B} \) is nonzero) the value of a given line integral depends on the path because closed circuits enclose flux.

Referring to eqn (1) the phase for a closed circuit \( C \) is

\[ \gamma(C) = q \oint_{\vec{C}} d\vec{r} \cdot \vec{A}. \tag{2} \]

This is the geometric phase of the magnetic version of the Aharonov-Bohm effect. Referring to Fig. 2(a), \( \vec{B} \) is confined to (and distributed uniformly over) a solenoid of cir-
cular cross section. Because there is no magnetic flux outside the solenoid, the path can be distorted such that the path indicated in (b) can be chosen. Note that the upper arrow in (a) has reversed direction in going from (a) to (b) because it is the difference between the phases acquired on the upper and lower paths that is sought.

You might ask why the paths in Fig. 2 are drawn as classical trajectories. Should we not sum over many paths, as in the path integral formulation of quantum mechanics? The answer lies with the ability to distort paths in regions of zero flux. However many paths are chosen, say, above the solenoid, they can all be distorted to the same single path for the purpose of computing the phase.

Adiabatic separation can be introduced by placing the particle in a small box (e.g., assign it to the ground state box eigenfunction), followed by slow transport of the box around the solenoid. Placing the particle in the box drives home the issue of adiabaticity, as a fast degree of freedom is present: the particle's location within the box.

The above case dealt with the magnetic Aharonov-Bohm effect, which is more popular than the electric version, possibly because it was the first to be verified experimentally. The scalar potential \( \phi \) that was suppressed when dealing with the magnetic version is now taken into account. Again, the workload is transferred to the wave function, and eqn (1) becomes

\[
\psi = \psi_{\vec{A}=0, \phi=0} \exp \left( i q \int_{\vec{r}_0, t_0}^{\vec{r}, t} (d\vec{r} \cdot \vec{A} - dt \phi) \right). \tag{3}
\]

As before, it is easily verified that this is the general solution by substituting it into eqn (5.5), and using the Leibniz rule for differentiation of an integral. Thus, the overall geometric phase is expressed as
Chapter 2. Electrodynamics

\[ \gamma = -q \int dx^\mu A_\mu . \quad (4) \]

In summary, the potentials \( \phi \) and \( \vec{A} \) have emerged as the central objects of electrodynamics, entering together as the four-vector gauge field \( A^\mu = (\phi, \vec{A}) \). There is no temptation to use \( \vec{E} \) and \( \vec{B} \), nor would they suffice were this attempted. The gauge field multiplied by the electric charge \( q \) is referred to as the gauge connection, though sometimes the field alone is referred to as the gauge connection. We shall see that it is responsible for the parallel transport of the wave function from one point to the next along a path.

**Covariant Derivative**

The above results can be distilled into a compact algorithm. To see how this works, consider a charged particle that is present in an electromagnetic field. With no particle-field coupling, the Hamiltonian is \( H = p^2/2m + H_{EM} \), where \( H_{EM} \) is an expression for the electromagnetic energy. Interaction is now turned on by using the transformation:

\[ \partial^\mu \rightarrow \partial^\mu + iqA^\mu . \]

You might wonder why \( D^\mu \) is called a covariant derivative, as it is obviously contravariant. The term covariant has two meanings. When an equation or expression does not change its form under a transformation it is said to be Lorentz covariant, or covariant for short. For example, Maxwell's equations and the Dirac equation are Lorentz covariant, the Schrödinger equation is not Lorentz covariant, and so on. The covariant derivative \( D^\mu \) falls into this category. However, we also deal with covariant and contravariant components of tensors. In this context, these terms denote the transformation properties of the tensor components. Thus, both \( D^\mu = \partial^\mu + iqA^\mu \) and \( D_\mu = \partial_\mu + iqA_\mu \) are Lorentz covariant derivatives, even though they are expressed in terms of contravariant and covariant components, respectively. It would possibly make more sense to use the term form invariant rather than covariant when referring to how an expression or equation transforms. This will be done if it is likely to lessen confusion, but in general I will stick to the common usage described above.

The covariant derivative converts the system from one of global gauge invariance to one of local gauge invariance. The covariant derivative mixes electric and magnetic interactions according to the requirements of electrodynamics. Without it the system is not invariant with respect to a local gauge transformation. Local gauge invariance is assured through the substitution given by eqn (5).
Because $D_\mu$ is Lorentz covariant, so is the commutator $[D_\mu, D_\nu]$. Expanding this commutator yields an important tensor.

$$[D_\mu, D_\nu] = [\partial_\mu + iqA_\mu, \partial_\nu + iqA_\nu]$$

$$= iq (\partial_\mu A_\nu - \partial_\nu A_\mu)$$

$$= iq F_{\mu\nu}. \quad (6)$$

The term $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the (gauge covariant) electromagnetic field strength tensor we encountered in Section 3. Maxwell's equations are expressed in terms of it, and $-\frac{1}{4} F^{\mu\nu} F_{\mu\nu}$ is the Lagrangian density for the free electromagnetic field. And so on.
Bibliography and References

Chapter 2. Electrodynamics

Exercises

1. A laser pointer that produces 50 mW of green light can be purchased for a modest amount of money. You probably do not know it, but these often have narrow linewidths (a few MHz) and good quality beam profiles (phase and shape). With a bit of care, this radiation can be focused to a spot of ~ 1 μm diameter. Estimate the magnitude of the electric field in the focal region.

2. It is found by an experimentalist working with a certain molecule that the spontaneous emission lifetime between two molecular states separated by 12000 cm⁻¹ is 300 ns. A theorist friend calculates properties of this molecule and finds that there is another transition, also allowed by symmetry, in which, roughly speaking, the same amount of charge moves the same distance. This transition has never been measured. The theorist predicts an emission wavelength of ~ 200 nm. Assuming that the theoretical estimate is of good accuracy, what do you predict for the spontaneous emission lifetime?

3. Estimate the spontaneous emission lifetime for the v = 1 vibrational level of HCl. How does this compare to the spontaneous emission lifetime of higher vibrational levels? You will have to express the dipole moment operator using a Taylor expansion about the equilibrium internuclear separation. Using this you can then estimate the transition dipole moment. This will require an estimate on your part. You will not be able to calculate everything needed.

4. Consider a high Rydberg state of atomic hydrogen, i.e., the principle quantum number is n = 100. What is the spontaneous emission lifetime of the transition from n = 100, l = 1 to n = 99, l = 0? What is wavelength of this emission? What is the diameter of the Rydberg level? Now put the excited atom in a hollow cube with perfectly conducting walls. How large must this cube be to permit spontaneous emission on the above transition? Explain your reasoning.

5. If the electric dipole term vanishes due to symmetry, we must consider the next term in the expansion of e^{±ik·r}. Work out the matrix element

\[ \langle \hat{B} | \langle \vec{k} \cdot \vec{r} \rangle \vec{p} \cdot \vec{e} | \hat{A} \rangle. \]

To begin, write \((\vec{k} \cdot \vec{r})\vec{p} \cdot \vec{e} = \frac{1}{2}((\vec{k} \cdot \vec{r})\vec{p} + (\vec{k} \cdot \vec{p})\vec{r}) + (\vec{k} \cdot \vec{r})\vec{p} - (\vec{k} \cdot \vec{p})\vec{r}) \cdot \vec{e} \) and note that \((\vec{k} \cdot \vec{r})\vec{p} + (\vec{k} \cdot \vec{p})\vec{r} = \vec{k} \cdot (\vec{r} \vec{p} + \vec{p} \vec{r}) \).

The product of two vectors (e.g., \(\vec{r} \vec{p}\)) is a second rank tensor, also called a dyadic, e.g., a dot product acting from the left on the term \(\vec{r} \vec{p}\) with a vector \(\vec{k}\) gives \((\vec{k} \cdot \vec{r})\vec{p}\).

Show that the anticommutator: \(\vec{r} \vec{p} + \vec{p} \vec{r}\), is equal to \((im / \hbar)[H_{\text{mol}}, \vec{r} \vec{p}]\). Use this to express \((\vec{k} \cdot \vec{B}) (\vec{r} \vec{p} + \vec{p} \vec{r}) | \hat{A} \rangle \cdot \vec{e}\) in terms of matrix elements of \(\vec{r}\). This is called the electric quadrupole term. Next, carry out the analogous set of exercises for the term: \((\vec{k} \cdot \vec{r})\vec{p} - (\vec{k} \cdot \vec{p})\vec{r}\). This is called the magnetic dipole term.

185