7320 Electromagnetic Theory II

University of Colorado, Boulder (Spring 2020)

Lecturer: Prof. Tom DeGrand
Scribe: Reuben R.W. Wang

These notes are the second in a series of 2 graduate level documents on classical electrodynamics. These notes are intended to accompany the relevant texts for this class, which are Jackson's "*Classical Electrodynamics*", Landau and Lifshitz' "*Classical Field Theory*" and "*Electrodyamics of Continuous Media*", and Ryder's "*Quantum Field Theory*". The topics covered in this course this semester will be as follows.

1. Non-Relativistic Radiation;
2. Scattering and Diffraction;
3. Special Relativity;
   (a) 4 Vectors,
   (b) Covariant Formalism,
   (c) Particle Motion in External Fields,
   (d) Classical Field Theory (Goldstone Bosons, Higgs effect, Meissner effect),
4. Relativistic Radiation;
   (a) Larmor Formula,
   (b) Synchrotron Radiation,
   (c) Bremsstrahlung Radiation,
5. Quantum Field Theory (Tentative).

For information on the class and relevant resources, refer to this link. All text in blue colored font are hyperlinks to ease navigation through these notes or for quick reference. Before beginning, it would be worth reading V.F. Weisskopf's "*How Light Interacts with Matter*".

Instructor: *Professor Thomas DeGrand.*
Instructor office hours: *Wednesday (2pm - 4pm) and Thursday (1pm - 5pm).*
Instructor email: `degrand@aurinko.colorado.edu`.

Contents

Chapter 1

Introduction

Let's start of with establishing the unit system we will be adopting in this class. As per usual, we always want a set of units that makes life easiest for you (in context of the physical theory you are working with). For this course, we will be adopting Gaussian units. This allows us to keep track of physical dimensions which keeps us sane when going back to check the derivations we go through. In fact, a lot of the quantities we are concerned with in this class are innately dimensionless (e.g. antenna patterns, scattering cross-sections, etc), so the unit system we adopted does not play that huge a role. Nonetheless, we shall start with a lightning introduction to Gaussian units.

§1.1 Gaussian Units

We start some with some relevant equations to juxtapose the equations between MKS and Gaussian (CGS) units.

$$
\begin{array}{ll}
\underline{\textbf{MKS}} & \underline{\textbf{CGS}} \\
\boldsymbol{\nabla}\cdot\vec{B}=0 & \boldsymbol{\nabla}\cdot\vec{B}=0 \\
\boldsymbol{\nabla}\cdot\vec{E}=\dfrac{\rho}{\epsilon_0},\quad \boldsymbol{\nabla}^2\Phi=-\dfrac{\rho}{\epsilon_0} & \boldsymbol{\nabla}\cdot\vec{E}=4\pi\rho,\quad \boldsymbol{\nabla}^2\Phi=-4\pi\rho \\
\boldsymbol{\nabla}\times\vec{B}-\mu_0\epsilon_0\dfrac{\partial\vec{E}}{\partial t}=\mu_0\vec{J} & \boldsymbol{\nabla}\times\vec{B}-\dfrac{1}{c}\dfrac{\partial\vec{E}}{\partial t}=\dfrac{4\pi}{c}\vec{J} \\
\boldsymbol{\nabla}\times\vec{E}+\dfrac{\partial\vec{B}}{\partial t}=0 & \boldsymbol{\nabla}\times\vec{E}+\dfrac{1}{c}\dfrac{\partial\vec{B}}{\partial t}=0 \\
\vec{B}=\boldsymbol{\nabla}\times\vec{A} & \vec{B}=\boldsymbol{\nabla}\times\vec{A} \\
\vec{E}=-\boldsymbol{\nabla}\Phi-\dfrac{\partial\vec{A}}{\partial t} & \vec{E}=-\boldsymbol{\nabla}\Phi-\dfrac{1}{c}\dfrac{\partial\vec{A}}{\partial t}
\end{array}
\tag{1.1}
$$

We see that in CGS, $\vec{E}$ and $\vec{B}$ have the same units, along with Φ and $\vec{A}$. Turning to macroscopic electrodynamics, we have the equations:

$$\begin{array}{ll} \underline{\textbf{MKS}} & \underline{\textbf{CGS}} \\ \vec{D} = \epsilon\vec{E} = \epsilon_0\vec{E} + \vec{P} & \vec{D} = \epsilon\vec{E} = \vec{E} + 4\pi\vec{P} \\ \vec{B} = \mu\vec{H} = \mu_0\vec{H} + \vec{M} & \vec{B} = \mu\vec{H} = \vec{H} + 4\pi\vec{M} \end{array} \tag{1.2}$$

For the problems in this course, all we are going to really need out of Maxwell's equations is getting the wave equation out, then getting the energy densities and Poynting vector out of that. As recap, we have: electrodynamics, we have the equations:

$$\begin{array}{ll} \underline{\textbf{MKS}} & \underline{\textbf{CGS}} \\ u = \frac{\epsilon_0}{2}\vec{E}^2 + \frac{1}{2\mu_0}\vec{B}^2 & u = \frac{1}{8\pi}\left(\vec{E}^2 + \vec{B}^2\right) \\ \vec{S} = \vec{E}\times\vec{H} & \vec{S} = \frac{c}{4\pi}\left(\vec{E}\times\vec{H}\right) \end{array} \tag{1.3}$$

§1.1.1 Radiation in Gaussian Units

The notion of radiation can most simply be treated as electromagnetic waves propagating in free-space. For this simple system, we have that the governing equations are written as:

$$\boldsymbol{\nabla}\cdot\vec{E} = 0, \quad \boldsymbol{\nabla}\cdot\vec{B} = 0 \tag{1.4}$$

$$\boldsymbol{\nabla}\times\vec{E} + \frac{1}{c}\frac{\partial\vec{B}}{\partial t}, \quad \boldsymbol{\nabla}\times\vec{B} - \frac{1}{c}\frac{\partial\vec{E}}{\partial t}, \tag{1.5}$$

Taking time-derivatives and curls on the second 2 equations above grants us the wave equation given as

$$\left(\boldsymbol{\nabla}^2 - \frac{1}{c^2}\frac{\partial^2}{\partial^2 t}\right)\begin{bmatrix}\vec{E}\\ \vec{B}\end{bmatrix} = \vec{0} \tag{1.6}$$

We then have the ansatz for the equation above are indeed planewaves which is written as

$$\begin{aligned} \vec{E} &= \hat{\varepsilon}E_0\exp\left\{i\left(\vec{k}\cdot\vec{x} - \omega t\right)\right\} \\ \vec{B} &= \left(\frac{ck}{\omega}\right)\hat{n}\times\vec{E} \\ \text{s.t.}\quad k^2 - \frac{\omega^2}{c^2} &= 0, \quad \vec{k}\cdot\vec{\varepsilon} = 0 \end{aligned} \tag{1.7}$$

where $\hat{\varepsilon}$ is the direction orthogonal to that of propagation. Nicely, we have that the magnitudes of the electric and magnetic fields are exactly the same in CGS (since $c = \omega/k$)! This grants us that the time-averaged Poynting vectors and energy densities have the nice forms:

$$\langle\vec{S}\rangle = \frac{c}{8\pi}\left\|\vec{E}_0\right\|^2\hat{n} \tag{1.8}$$

$$\langle u\rangle = \frac{1}{8\pi}\left\|\vec{E}_0\right\|^2 \tag{1.9}$$

Now, this allows us to talk about radiation from a source. A quantity of interest for radiation from a point source is the power incident on some solid angle in space:

$$\frac{dP}{d\Omega} = r^2 \hat{n} \cdot \langle \vec{S} \rangle \tag{1.10}$$

which allows us to compute the *antenna patterns* defined as $\frac{1}{P}\frac{dP}{d\Omega}$. We will now dive into radiation formally and properly understand these quantities.

Chapter 2

Radiation

We now get into radiation, which was alluded to at the end of the previous chapter. In the previous course (Electromagnetic Theory I), we learnt about how electromagnetic fields propagate as waves in space. In this chapter, we will be discussing about how these waves are in fact generated, and the mathematical formalism to describe such system.

§2.1 Solutions to the Wave Equation

Electromagnetic radiation results from the acceleration of charges in space. More specifically, since waves have a periodic structure in-built into their construction, a series of periodic time-dependent terms in a system of charges generates electromagnetic multipole moments which thus lead to non-trivial sources of radiation. To start talking about radiation, we need to solve the wave equation.

$$\boldsymbol{\nabla}^2\psi - \frac{1}{c^2}\frac{\partial}{\partial t}\psi = -4\pi f(\vec{x}, t) \tag{2.1}$$

An adequate means of solving such equations is again via the use of *Green's functions*.

§2.1.1 Green's Functions (Quick Refresher)

Given any dynamical system described by a system of linear time-invariant (LTI) equations, we can describe the response of the solution to inputs $f(\vec{x}, t)$ via a *response function* $\psi(\vec{x}, t)$ which solves the equation:

$$\mathcal{L}[\psi(\vec{x}, t)] = -4\pi f(\vec{x}, t) \tag{2.2}$$

where $\mathcal{L}$ denotes a *linear operator* (time-invariant) determined by the physical system and the added factor of -4π is tacked on here so as to suit our context.

Note: So as not to obscure any meaning, we go over some of the terminologies used. The term "inputs" here refer to external perturbations to the system (e.g. hitting an oscillator with a hammer, shooting a trapped BEC with a laser pulse, etc...). The term "response", then describes the system's reaction to such external inputs.

Green's functions $G(\vec{x}, t; \vec{x}', t')$, are a class of response solutions to such LTI system under the a δ-function input $(f(\vec{x}, t) = \delta^3(\vec{x} - \vec{x}')\delta(t - t'))$. Explicitly, a Green's function is a solution to the LTI system of equations:

$$\mathcal{L}\left[G(\vec{x}, t; \vec{x}', t')\right] = -4\pi\delta^3(\vec{x} - \vec{x}')\delta(t - t'). \tag{2.3}$$

Response solutions for any input function $f(\vec{x}, t)$ then follows by applying a *convolution*, in which we simply have to convolve the input with the Green's function to obtain this solution. To see this, first recall that a convolution of 2 functions f and g is defined as:

$$(f * g)(t) \equiv \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau \tag{2.4}$$

As such, the convolution between a δ-function and some input function $f(\vec{x}, t)$ is given as:

$$\begin{aligned}(f * \delta)(\vec{x}, t) &= \int f(\vec{x}\,', t')\delta^3(\vec{x} - \vec{x}\,')\delta(t - t')d^3x'dt' \\ &= -4\pi f(\vec{x}, t).\end{aligned} \tag{2.5}$$

What we have done above is essentially used the right-hand side of equation (2.3) to convolve $f(\vec{x}, t)$, so if we do the same for the left-hand side, we get:

$$\int f(\vec{x}, t)\mathcal{L}\left[G(\vec{x}, t; \vec{x}', t')\right] d^3x'dt' = -4\pi f(\vec{x}, t), \tag{2.6}$$

for which we can utilize the commutativity between operators $\int d^3x'dt'$ and $\mathcal{L}$ to get:

$$\begin{aligned}\mathcal{L}\left[\int f(\vec{x}, t)G(\vec{x}, t; \vec{x}', t')d^3x'dt'\right] &= -4\pi f(\vec{x}, t) \\ \Rightarrow \quad \psi(\vec{x}, t) &= \int f(\vec{x}, t)G(\vec{x}, t; \vec{x}', t')d^3x'dt'\end{aligned} \tag{2.7}$$

§2.1.2 Green's Functions for Wave Equations

Having derived the result above, we can adopt the use of a Green's function to construct wave-equation solutions since we can treat:

$$\mathcal{L} = \boldsymbol{\nabla}^2 - \frac{1}{c^2}\frac{\partial}{\partial t} \tag{2.8}$$

as the linear operator. As such, we replace the input function $f(\vec{x}, t)$ with a δ-function in the wave equation which gives:

$$\left(\boldsymbol{\nabla}^2 - \frac{1}{c^2}\frac{\partial}{\partial t}\right)G(\vec{x}, t; \vec{x}', t') = -4\pi\delta^3(\vec{x} - \vec{x}')\delta(t - t') \tag{2.9}$$

$$\text{where} \quad \psi(\vec{x}, t) = \int d^3x'dt'G(\vec{x}, t; \vec{x}', t')f(\vec{x}', t'). \tag{2.10}$$

To suitably work with these equations, it will be very useful to be able to jump back and forth between momentum-frequency space and spatial-time space $\{\vec{x}, t\} \to \{\vec{k}, \omega\}$. First, we define $\vec{R} = \vec{x} - \vec{x}'$ and $T = t - t'$, which grants us the relation:

$$G(\vec{R}, T) = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} d^3k \int_{-\infty}^{\infty} d\omega e^{i\vec{k}\cdot\vec{R}} e^{-i\omega T} \tilde{G}(\vec{k}, \omega) \tag{2.11}$$

We also have that the δ-functions can be written as:

$$\delta^3(\vec{R})\delta(T) = \frac{1}{(2\pi)^2} \int d^3k e^{i\vec{k}\cdot\vec{R}} \int d\omega e^{-i\omega T} \tag{2.12}$$

Plugging these back into the original wave equations grants us:

$$\left[\nabla^2 - \frac{1}{c^2}\frac{\partial}{\partial t}\right] G(\vec{R}, T) = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} d^3k \int_{-\infty}^{\infty} d\omega e^{i\vec{k}\cdot\vec{R}} e^{-i\omega T} \left[-k^2 + \frac{\omega^2}{c^2}\right] \tilde{G}(\vec{k}, \omega) \tag{2.13}$$

$$\Rightarrow \quad \tilde{G}(\vec{k}, \omega) = -\frac{4\pi}{\left[-k^2 + \frac{\omega^2}{c^2}\right]}. \tag{2.14}$$

Now plugging this into the integral defining $G(\vec{R}, T)$, we get:

$$G(\vec{R}, T) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega e^{-i\omega T} \left[\frac{1}{(2\pi)^3} \int \frac{4\pi d^3k e^{i\vec{k}\cdot\vec{R}}}{k^2 - \frac{\omega^2}{c^2}}\right] \tag{2.15}$$

This integral can be done by first noticing that the integral over k runs over all k, which allows us to rotate in k-space arbitrarily such that we are in the most convenient coordinate frame. As such, we rotate such that $\vec{k}$ is aligned with $\vec{R}$ which grants us that the integral just over k becomes:

$$\begin{aligned} I_k(R, \omega) &= \frac{4\pi}{(2\pi)^3} \int_0^{\infty} \frac{2\pi k^2 dk}{k^2 - \frac{\omega^2}{c^2}} \int_{-1}^{1} d(\cos\theta) e^{ikR\cos\theta} \\ &= \frac{1}{i\pi R} \int_{-\infty}^{\infty} \frac{k dk e^{ikR}}{k^2 - \frac{\omega^2}{c^2}} \\ &= \frac{1}{i\pi R} \int_{-\infty}^{\infty} \frac{k dk e^{ikR}}{\left(k - \frac{\omega}{c}\right)\left(k + \frac{\omega}{c}\right)} \end{aligned} \tag{2.16}$$

The integral above has poles at $k = \pm\frac{\omega}{c}$, so we want to convert this integral into a contour integral, which forces us to evaluate the residues (we pick the poles, one being outside the contour and the other at $k = \frac{\omega}{c} + i\epsilon$):

$$\begin{aligned} \lim_{\epsilon \to 0^+} I_k^{(+)}(R, \omega) &= \lim_{\epsilon \to 0^+} \left(\frac{2\pi i}{R\pi i}\right) \left(\frac{\frac{\omega}{c} + i\epsilon}{\frac{2\omega}{c} + i\epsilon}\right) e^{iR\left(\frac{\omega}{c} + i\epsilon\right)} \\ &= \frac{e^{i\omega R/c}}{R} \end{aligned} \tag{2.17}$$

Putting this back into the Green's function in $\vec{R}$ and T gives us:

$$\boxed{G(R, T) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega e^{-i\omega T} \left[\frac{e^{i\omega R/c}}{R}\right] = \frac{\delta\left(T - \frac{R}{c}\right)}{R}} \tag{2.18}$$

The $1/R$ tells us that the Green's function decays away from the source, which is why it is often referred to as a *retarded Green's function.* Furthermore, we see that this Green's function admits solutions that propagate forward in time limited by the speed of light c. This explains why we chose the poles the way we did (if we had flipped them, the Green's function we would get would violate causality called an *advanced wave*). As such, this grants us the wavefunction solution as:

$$\boxed{\psi(\vec{x},t) = \int \frac{d^3x'}{\|\vec{x}-\vec{x}'\|} f\left(\vec{x}', t' = t - \|\vec{x}-\vec{x}'\|/c\right)} \tag{2.19}$$

where we noted that the argument in the δ-function of the Green's function vanishes when:

$$T = \frac{R}{c} \quad \Rightarrow \quad t - t' = \frac{\|\vec{x}-\vec{x}'\|}{c} \tag{2.20}$$

It is worth noting that one could imagine solving the wave-equation with boundary conditions that are specified at late times. Asking what happened to the wave at early times would then require the use of advanced wave solutions (refer to Jackson), but otherwise, the retarded wave solution is the most physical and what we will be using for the rest of the class.

§2.1.3 Waves in the Lorenz Gauge

Starting with the simplest gauge, we consider the wave equation in the Lorenz gauge which gives us the wave equation:

$$\left(\boldsymbol{\nabla}^2 - \frac{1}{c^2}\frac{\partial}{\partial t}\right)\begin{bmatrix}\vec{A}\\ \Phi\end{bmatrix} = -\frac{4\pi}{c}\begin{bmatrix}\vec{J}\\ c\rho\end{bmatrix} \tag{2.21}$$

Using again the Green's function technique we did above, the vector potential is then:

$$\begin{aligned}\vec{A}(\vec{x},t) &= \frac{1}{c}\int d^3x' dt' G(\vec{x},t;\vec{x}',t')\vec{J}(\vec{x}',t') \\ &= \frac{1}{c}\int \frac{d^3x'}{\|\vec{x}-\vec{x}'\|}\vec{J}(\vec{x}', t' = t - \|\vec{x}-\vec{x}'\|/c)\end{aligned} \tag{2.22}$$

This is the starting point for most of our analysis but many approximations will be made to this since it is hard to do anything with this exact form. There is a special case where we have $\vec{J}(\vec{x},t) = e^{-i\omega t}\vec{J}(\vec{x})$, that it the time-dependence is of the current is purely harmonic. This occurs when we also have the charge density being $\rho(\vec{x},t) = e^{-i\omega t}\rho(\vec{x})$ which in fact gives that:

$$\vec{A}(\vec{x},t) = e^{-i\omega t}\vec{A}(\vec{x}) \tag{2.23}$$

$$\Rightarrow \quad \boxed{\vec{A}(\vec{x},t) = e^{-i\omega t}\frac{1}{c}\int d^3x' \frac{\vec{J}(\vec{x}')}{\|\vec{x}-\vec{x}'\|}e^{ik\|\vec{x}-\vec{x}'\|}} \tag{2.24}$$

$$\Rightarrow \quad \vec{A}(\vec{x}) = \frac{1}{c}\int d^3x' \frac{\vec{J}(\vec{x}')}{\|\vec{x}-\vec{x}'\|}e^{ik\|\vec{x}-\vec{x}'\|} \tag{2.25}$$

where $k = \omega/c$. It turns out we don't need to know Φ in harmonic time-dependencies and observers far away from the current, since for a localized distribution gives us that $\vec{J} \to 0$ far away, this gives us that:

$$\boldsymbol{\nabla} \times \vec{B} = \boldsymbol{\nabla} \times \left(\boldsymbol{\nabla} \times \vec{A}\right) = -ik\vec{E}. \tag{2.26}$$

Now, let's take one step back and consider the natural distance scales of most radiation systems we deal with. We have

1. d: diameter of the source;
2. r: distance to the observer;
3. λ: the wavelength of the radiation.

Now consider that we are very far away from the source (*far-field regime*), we have that $r \gg \lambda$ which implies $kr \gg 1$. In this regime, we have that:

$$\vec{E}, \vec{B} \sim \frac{1}{r}. \tag{2.27}$$

This regime is very often adopted as it helps to simplifies our lives in many scenarios. Furthermore, this regime is very common in many experiments and real life applications. In this regime, we have that:

$$R = \|\vec{x} - \vec{x}'\| \approx r - \frac{\vec{x} \cdot \vec{x}'}{r} \tag{2.28}$$

where $r = \|\vec{x}\|$. Calling $\vec{x}/r = \hat{n}$, we then have:

$$R \approx r - \hat{n} \cdot \vec{x}' \tag{2.29}$$

Plugging this into the integral definition of $\vec{A}(\vec{x})$ gives us:

$$\vec{A}(\vec{x}) \approx \frac{e^{ikr}}{cr} \int d^3x' \vec{J}(\vec{x}') e^{-ik\left(\hat{n} \cdot \vec{x}'\right)} \tag{2.30}$$

some things to note about this result, the factor in front of the integral is completely isotropic. What is important in the integral is the direction of the observer relative to the distribution of source points. From this, we can compute $\vec{B}$ by taking its gradient (acts on the e^{ikr}, $1/r$ and angular terms) as follows:

$$\vec{B} = \boldsymbol{\nabla} \times \vec{A}(\vec{x}) \tag{2.31}$$

in practice, we can often ignore the gradient acting on the $1/r$ term since we are in the far-field and can ignore terms in $1/r^2$. In addition to this, there is also a regime known as the *long wavelength limit* which is whereby $\lambda \gg d$ (common in atomic systems).

Note: For reference, a order of magnitude measure of wavelength for atoms is given be $d \sim \mathring{A}$, $\lambda \sim 1000\mathring{A}$.

In such a regime, we can then expand the exponent in the integral into a power series which grants us:

$$\begin{aligned}\vec{A}(\vec{x}) &= \frac{e^{ikr}}{cr}\sum_\ell \frac{1}{\ell!}\int d^3x'\left(-ik\hat{n}\cdot\vec{x}'\right)^\ell \vec{J}(\vec{x}')\\ &= \frac{e^{ikr}}{cr}\sum_\ell \frac{1}{\ell!}\int d^3x'\left(-ikr'\cos\theta\right)^\ell \vec{J}(\vec{x}')\end{aligned} \tag{2.32}$$

This is actually a *multipole expansion* of sorts, whereby each expansion term is approximately dependent on $(kd)^\ell$. We will now use this result to show that simple breathing monopole mode does not radiate!

Statement: *Monopole breathing modes do* ***not*** *produce radiation.*

Proof. First consider:

$$\Phi(\vec{x},t) = \int d^3x' \frac{\rho(\vec{x}',t-R/c)}{R} \tag{2.33}$$

Expanding $1/R$ in a Legendre polynomial series gives:

$$\frac{1}{R} = \frac{1}{r}\sum_\ell \left(\frac{r'}{r}\right)^\ell P_\ell(\cos\theta) \tag{2.34}$$

which tells us that the monopole $\ell = 0$ mode is:

$$\Phi_0(\vec{x},t') = \frac{1}{r}\int d^3x'\rho(\vec{x}',t-R/c) = \frac{Q}{r} \tag{2.35}$$

indeed showing that it is a constant in time and does not produce any radiation. □

Now consider the radiation from an electric dipole (electric dipole radiation). We have the dipole from the vector potential as:

$$\begin{aligned}\vec{A}_{\text{dipole}}(\vec{x}) &= \frac{e^{ikr}}{cr}\int d^3x\vec{J}(\vec{x})\\ &= \frac{e^{ikr}}{cr}(-i\omega)\int d^3x\rho(\vec{x})\vec{x}\end{aligned} \tag{2.36}$$

where the harmonic field result from earlier was adopted in the above derivation. Recalling that:

$$\int d^3x\rho(\vec{x})\vec{x} = \vec{p} \tag{2.37}$$

where $\vec{p}$ is the dipole moment, this grants us the result:

$$\boxed{\vec{A}(\vec{x}) = -\frac{ik}{r}\vec{p}e^{ikr}} \tag{2.38}$$

Lastly, there is a regime whereby $\lambda \gg r \gg d$ known as the *near zone regime*, which grants that $kr \sim 0$. This causes the exponent in the integral for $\vec{A}$ to evaluate to unity, which brings us back to a magnetostatics equation:

$$\vec{A}(\vec{x}) = \frac{1}{c}\int d^3x' \frac{\vec{J}(\vec{x}')}{\|\vec{x} - \vec{x}'\|} \tag{2.39}$$

This is less common but present in certain niche applications. Going back to the vector potential we had in the box above, the goal now is to get from this vector potential to the Poynting vector. First, we find the magnetic field as:

$$\begin{aligned}
\vec{B} &= \boldsymbol{\nabla} \times \vec{A} \\
&= \boldsymbol{\nabla} \times \left(-\frac{ik}{r}\vec{p}e^{ikr}\right) \\
&= ik\left(1 + \frac{i}{kr}\right)\frac{e^{ikr}}{r}\hat{n} \times (-ik\vec{p}) \\
&\approx \frac{k^2 e^{ikr}}{r}\hat{n} \times \vec{p}
\end{aligned} \tag{2.40}$$

where we adopted the long-wavelength limit and used the fact that:

$$\begin{aligned}
\boldsymbol{\nabla} r &= \sum_j \hat{j}\frac{x_j}{r} = \hat{n} \\
\Rightarrow \quad \boldsymbol{\nabla} r \cdot \boldsymbol{\nabla} r &= 1
\end{aligned} \tag{2.41}$$

As such, we can now compute the electric field as:

$$\begin{aligned}
\vec{E} &= -\hat{n} \times \vec{B} \\
&= -\frac{k^2 e^{ikr}}{r}\left[\hat{n} \times (\hat{n} \times \vec{p})\right] \\
&= -\frac{k^2 e^{ikr}}{r}\left[\hat{n}\,(\hat{n} \cdot \vec{p}) - \vec{p}\right]
\end{aligned} \tag{2.42}$$

From these fields, we can get the time-averaged Poynting vector:

$$\langle\vec{S}\rangle = \frac{c}{8\pi}\operatorname{Re}\left\{\vec{E} \times \vec{B}^*\right\} \tag{2.43}$$

when we are dealing with point source radiation, what we are concerned with is the power radiated per unit solid angle, which we can write as:

$$\begin{aligned}
\frac{dP}{d\Omega} &= r^2\hat{n} \cdot \langle\vec{S}\rangle \\
&= \frac{c}{8\pi}r^2\hat{n} \cdot \operatorname{Re}\left\{\vec{E} \times \vec{B}^*\right\} \\
&= \frac{c}{8\pi}r^2\operatorname{Re}\left\{\hat{n}\left[\hat{n}|E|^2 - \vec{E}^*\hat{n} \cdot \vec{E}\right]\right\} \\
&= \boxed{\frac{c}{8\pi}r^2|E|^2}
\end{aligned} \tag{2.44}$$

where it was noted that $\vec{E} \cdot \hat{n} = 0$ since the electric field would be transverse to the direction of propagation. The boxed formula above is the general form of the un-normalized antenna pattern.

Note: This formula seems to have no reason for us to consider the polarization of $\vec{E}$. However, this would not make sense since if we were to obscure the radiation with a polarizer, there would definitely be less power per unit solid angle. As such, we can amend the formula as follows:

$$\frac{dP_\varepsilon}{d\Omega} = \frac{c}{8\pi} r^2 \left| \hat{\varepsilon}^* \cdot \vec{E} \right|^2 \tag{2.45}$$

to account for polarization losses, where $\hat{\varepsilon}$ is the polarization axis.

As in the particular case of dipoles, we have that the final result for the **non-normalized antenna pattern of electric dipoles** being:

$$\boxed{\frac{dP}{d\Omega} = \frac{ck^4}{8\pi} \left[\|\vec{p}\|^2 - |\hat{n} \cdot \vec{p}|^2 \right]} \tag{2.46}$$

In the case where the **dipole alignment axis is fixed in space**. We can then simplify the formula by setting $\vec{p} \cdot \hat{n} = p \cos\theta$. This gives us that:

$$\frac{dP}{d\Omega} = \frac{ck^4}{8\pi} |p|^2 \sin^2\theta \tag{2.47}$$

$$\Rightarrow \quad \boxed{P = \frac{c}{3} |p|^2 k^4} \tag{2.48}$$

Let's now consider an example.

Example:

Consider a wire of length d aligned along the $\hat{z}$-axis where there are electric charges sloshing back and forth inside it ($I(t) = I(z)e^{-i\omega t}$). We also place the origin in the middle of the wire. Looking closer at the current, we have:

$$\begin{aligned} I &= \int \vec{J}\hat{n} dA \\ \text{s.t.} \quad I(z &= \pm d/2) = 0 \end{aligned} \tag{2.49}$$

Let's then model:

$$I(z) = I_0 \left[1 - \frac{2|z|}{d} \right] \tag{2.50}$$

$$\Rightarrow \quad \rho = \frac{1}{i\omega} \frac{dI}{dz} = \pm i \frac{2I_0}{\omega d} \tag{2.51}$$

Integrating this over the length of the wire to find the electric dipole moment gives us:

$$p_z = \int_{-d/2}^{d/2} z\rho(z) dz = i \frac{I_0 d}{2\omega} \tag{2.52}$$

As such, assume that we are far away enough from the antenna and adopt the long-wavelength limit ($\lambda \gg d$) so that we can compute the antenna pattern as:

$$\frac{dP}{d\Omega} = \frac{c}{8\pi}\left(\frac{I_0 d}{2\omega}\right)^2 k^4 \sin^2\theta = \left(\frac{I_0^2 k^2 d^2}{32\pi c}\right)\sin^2\theta \tag{2.53}$$

From this, we can compute the power as well and compare this to the "engineers" formula $P = I_0^2 R/2$ to get the *radiation resistance*:

$$R_{rad} = \frac{k^2 d^2}{6c} \tag{2.54}$$

§2.1.4 Higher Multipoles

If we go back to the equation for the vector potential:

$$\vec{A}(\vec{x}) = ik\frac{e^{ikr}}{cr}\int d^3x'\,(\hat{n}\cdot\vec{x}')\,\vec{J}(\vec{x}') \tag{2.55}$$

this equation is in fact hiding magnetic dipole and electric quadrupole moments.

Note: In this course, we will not be analyzing anything more than quadrupole moments since all higher order terms are generally unnecessary and painful to compute.

Considering the current in terms of its longitudinal and radial vector components $\vec{J} = \vec{J}_l + \vec{J}_\perp$, we have that:

$$\boldsymbol{\nabla}\cdot\vec{J}_l = 0; \quad \boldsymbol{\nabla}\times\vec{J}_\perp = 0 \tag{2.56}$$

Then to expand the vector potential in a way that is useful to us, we first rewrite it as:

$$\vec{A}(\vec{x}) = ik\frac{e^{ikr}}{cr}\int d^3x'\,(\hat{n}\cdot x'\hat{n}')\,\vec{J}(\vec{x}') \tag{2.57}$$

for which since we know the identity of double cross products to be:

$$\hat{n}\times(\hat{n}'\times\vec{J}) = \hat{n}'(\hat{n}\cdot\vec{J}) - (\hat{n}\cdot\hat{n}')\vec{J} \tag{2.58}$$

$$\Rightarrow \quad (\hat{n}\cdot\hat{n}')\vec{J} = \frac{1}{2}\left[(\hat{n}\cdot\hat{n}')\vec{J} + \hat{n}'(\hat{n}\cdot\vec{J})\right] - \frac{1}{2}\hat{n}\times(\hat{n}'\times\vec{J}), \tag{2.59}$$

where the first term in the second line above indeed corresponds to a quadrupole and the second term a monopole. As for the associated magnetic dipole of a charge distributions instead, we instead have magnetic moments given as:

$$\vec{m} = \int d^3x'\,\frac{1}{2c}\left[\vec{x}'\times\vec{J}(\vec{x}')\right] \tag{2.60}$$

With this, we then have that the **vector potential for magnetic dipoles** is given by:

$$\vec{A}(\vec{x}) = ik(\hat{n}\times\vec{m})\frac{e^{ikr}}{r} \tag{2.61}$$

$$\Rightarrow \boxed{\vec{B}(\vec{x}) = \boldsymbol{\nabla}\times\vec{A}(\vec{x}) = -\frac{k^2 e^{ikr}}{r}\left[\hat{n}\times(\hat{n}\times\vec{m})\right]} \tag{2.62}$$

To further analyze the vector potential, we can rewrite the vector potential as:

$$\begin{aligned}
A_j(\vec{x}) &= ik\frac{e^{ikr}}{cr}\int d^3x'\,(\hat{n}\cdot\vec{x}')\,J_j(\vec{x}') \\
&= ik\frac{e^{ikr}}{cr}\int d^3x'\,(\hat{n}\cdot\vec{x}')\left[\left(\boldsymbol{\nabla}'x_j\right)\cdot\vec{J}(\vec{x}')\right] \\
&= ik\frac{e^{ikr}}{cr}\int d^3x'\boldsymbol{\nabla}'\cdot\left[(\hat{n}\cdot\vec{x}')\,\vec{J}(\vec{x}')\right] \\
&= ik\frac{e^{ikr}}{2cr}\int d^3x'\left[\vec{x}'(\hat{n}\cdot\vec{x}')\boldsymbol{\nabla}'\cdot\vec{J}\right] = -\frac{k^2}{2}\frac{e^{ikr}}{r}\int\vec{x}'(\hat{n}\cdot\vec{x}')\rho(\vec{x}')d^3x'
\end{aligned} \tag{2.63}$$

where we used integration by parts to move the divergence around from line 2 to 3 (line 3 to 4 is a simplification where terms happen to cancel out) and the fact that $\boldsymbol{\nabla}'\cdot\vec{J} = i\omega\rho$ from the continuity equation. It works out that the final expression above is in fact related to quadrupole radiation, so we have that the magnetic field from an electric quadrupole is given by:

$$\boxed{\vec{B}_{\text{quad}} = -\frac{ik^3}{6}\frac{e^{ikr}}{r}\left[\hat{n}\times\vec{q}(\hat{n})\right]} \tag{2.64}$$

$$\text{where}\quad q_i(\hat{n}) = Q_{ij}n_j,\quad Q_{ij} = \int d^3x'\rho(\vec{x}')\left[3x_i'x_j' - \delta_{ij}|x'|^2\right] \tag{2.65}$$

where $\vec{q} = \vec{k}_0 - \vec{k}$. From this, we can compute the **un-normalized antenna pattern due to quadrupole radiation** as:

$$\boxed{\frac{dP}{d\Omega} = \frac{ck^6}{288\pi}|\hat{n}\times(\hat{n}\times\vec{q}(\hat{n}))|^2 = \frac{ck^6}{288\pi}\left[|q(\hat{n})|^2 - |\hat{n}\cdot\vec{q}(\hat{n})|^2\right]} \tag{2.66}$$

From this, it works out that the **total power radiated from a quadrupole** is given by:

$$\boxed{P = \frac{ck^6}{360}\sum_{i,j}|Q_{ij}|^2} \tag{2.67}$$

(Refer to Jackson pg 414 to 415 for the full derivation of this). The key trick to evaluating the integral to get the total power radiated is that the quadrupole tensor is traceless ($Q_{jj} = 0$). As an overview, we have the fields and power distributions due to electric/magnetic dipoles and quadrupoles summarized in the listing below:

Electric Dipoles

$$\vec{B} = \frac{k^2 e^{ikr}}{r}\left[\hat{n}\times\vec{p}\right]$$
$$\vec{E} = -\frac{k^2 e^{ikr}}{r}\left[\hat{n}\times(\hat{n}\times\vec{p})\right]$$
$$\frac{dP}{d\Omega} = \frac{ck^4}{8\pi}|\hat{n}\times(\hat{n}\times\vec{p})|^2$$
$$P = \frac{ck^4}{3}\|\vec{p}\|^2$$

Magnetic Dipoles

$$\vec{B} = -\frac{k^2 e^{ikr}}{r}\left[\hat{n}\times(\hat{n}\times\vec{m})\right]$$
$$\vec{E} = -\frac{k^2 e^{ikr}}{r}\left[\hat{n}\times\vec{m}\right]$$
$$\frac{dP}{d\Omega} = \frac{ck^4}{8\pi}|\hat{n}\times(\hat{n}\times\vec{m})|^2$$
$$P = \frac{ck^4}{3}\|\vec{m}\|^2$$

Electric Quadrupoles

$$\vec{B} = -\frac{ik^3 e^{ikr}}{6r}\left[\hat{n}\times\vec{q}(\hat{n})\right]$$
$$\vec{E} = \frac{ik^3 e^{ikr}}{6r}\left[\hat{n}\times(\hat{n}\times\vec{q}(\hat{n}))\right]$$
$$\frac{dP}{d\Omega} = \frac{ck^6}{288\pi}|\hat{n}\times(\hat{n}\times\vec{q}(\hat{n}))|^2$$
$$P = \frac{ck^6}{360}\sum_{i,j}|Q_{ij}|^2$$

(2.68)

Example:

Consider a system of 3 charges along a line, equally space by a distance d between them. The middle charge has charge $2qe^{-i\omega t}$ and the others have charge $qe^{-i\omega t}$ (figure 2.1).

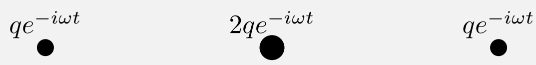

Figure 2.1: System of 3 colinear charges.

Let's try to compute the quadrupole tensor of this system. Recall that the angular distribution of the power and the total power emitted is given by:

$$\frac{dP}{d\Omega} = \frac{ck^6}{288\pi}\left[|q(\hat{n})|^2 - |\hat{n}\cdot\vec{q}(\hat{n})|^2\right] \tag{2.69}$$
$$P = \frac{ck^6}{360}\sum_{i,j}|Q_{ij}|^2 \tag{2.70}$$

When worked out, it turns out that the quadrupole tensor is diagonal with these entries being:

$$\begin{aligned} &Q_{xx} = Q_{yy} = -\frac{1}{2}Q_0, \quad Q_{zz} = Q_0 \\ \text{where} \quad &Q_0 = 4qd^2 \end{aligned} \tag{2.71}$$

As such, we have that the total power emitted is given by:

$$P = \frac{ck^6}{360}Q_0^2\left[1 + \left(-\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2\right] \tag{2.72}$$

Now to get the angular power distribution, we first need to compute the quadrupole vector $\vec{q}$ as follows:

$$\begin{aligned} q_i(\hat{n}) &= \hat{Q}_{ij}n_j \\ &= \hat{e}_i(\hat{n}\cdot\hat{e}_j)Q_{ij} = \hat{e}_i(\hat{n}\cdot\hat{e}_i)Q_{ii} \end{aligned} \tag{2.73}$$

where we had that $\hat{Q}_{ij} = \hat{e}_i \hat{e}_j Q_{ij}$ and we noted that Q_{ij} is diagonal. So we get that:

$$\begin{aligned} |q|^2 &= (\hat{n} \cdot \hat{e}_i)^2 Q_{ii}^2 \\ &= Q_0^2 \left[\cos^2\theta + \frac{1}{4}\sin^2\theta\left(\cos^2\phi + \sin^2\phi\right)\right] \end{aligned} \tag{2.74}$$

As for $\hat{n} \cdot \vec{q}$, we compute this as:

$$\begin{aligned} \hat{n} \cdot \vec{q} &= (\hat{n} \cdot \hat{e}_i) Q_{ii} \\ \Rightarrow \quad |\hat{n} \cdot \vec{q}|^2 &= Q_0^2 \frac{9}{16} \sin(2\theta) \end{aligned} \tag{2.75}$$

Plugging these back into the equation for the angular power distribution gives us:

$$\frac{dP}{d\Omega} = \left(\frac{ck^6}{288\pi}\right)\left(\frac{9Q_0^2}{6}\sin^2(2\theta)\right) \tag{2.76}$$

It turns out that there is class of problems with current densities that can be solved without the use of the multipole expansion. To see this, we first go back to the vector potential formula:

$$\vec{A}(\vec{x}) = \frac{1}{c}\frac{e^{ikr}}{r}\int d^3x e^{-ik\hat{n}\cdot\vec{x}'}\vec{J}(\vec{x}') \tag{2.77}$$

If we now consider 2 co-linear thin wires with a small gap between them, along with another thin wire orthogonal to the first 2 that has an end sitting right next to the gap (center-fed antenna). We can write the current as:

$$\vec{J}(\vec{x}') = \hat{z} I(z')\delta(x')\delta(y')\Theta\left(\frac{d}{2} - |z'|\right) \tag{2.78}$$

If we have the current as some sinusoidal function, say:

$$I(z) = I_0 \sin\left[k\left(\frac{d}{2} - |z'|\right)\right] \tag{2.79}$$

This gives us the vector potential as:

$$\begin{aligned} \vec{A}(\vec{x}) &= \frac{I_0}{c}\frac{e^{ikr}}{r}\tilde{A}_z\hat{z} \\ \Rightarrow \quad \tilde{A}_z(\vec{x}) &= \frac{I_0}{c}\int_{-d/2}^{d/2} dz' \sin\left(k\left(\frac{d}{2} - |z'|\right)\right) e^{-ikz'\cos\theta} \\ \Rightarrow \quad A_z(\vec{x}) &= \frac{2I_0}{c}\left[\frac{\cos\left(\frac{kd}{2}\cos\theta\right) - \cos\left(\frac{kd}{2}\right)}{\sin^2\theta}\right]\frac{e^{ikr}}{r} \end{aligned} \tag{2.80}$$

So we can compute the electric and magnetic fields from this to get the angular power distribution as:

$$\boxed{\frac{dP}{d\Omega} = ck^2\left|\tilde{A}_z\right|^2 \sin^2\theta} \tag{2.81}$$

which looks much like the radiation from an electric dipole! The point of what we have done, is to observe that the moment we have a vector potential which points in the z-direction, we are just back to the linear electric dipole but tacked on with $|A_z|^2$. First, consider the limit where $kd \ll 1$, we get:

$$\begin{aligned} A_z &\approx \frac{2I_0}{c}\left[\frac{1-\frac{1}{2}\left(\frac{kd}{2}\cos\theta\right)^2-\left(1-\frac{1}{2}\left(\frac{kd}{2}\right)^2\right)}{\sin^2\theta}\right] \\ &= \frac{2I_0}{c}\left[\frac{\frac{1}{2}\left(\frac{kd}{2}\right)^2\left(1-\cos^2\theta\right)}{\sin^2\theta}\right] \\ &= \frac{I_0}{c}\left(\frac{kd}{2}\right)^2 \end{aligned} \tag{2.82}$$

So all the angular dependence goes away and we are exactly back to the linear electric dipole. To see how well this approximation works, we can consider a "half-wavelength" antenna ($kd = \pi$), which would give us:

$$\frac{dP}{d\Omega} = \frac{I_0^2}{2\pi c}\frac{\cos^2\left(\frac{\pi}{2}\cos\theta\right)}{\sin^2\theta} \tag{2.83}$$

If we compare this to the dipole approximation, we had above, we get:

$$\frac{dP/d\Omega}{(dP/d\Omega)_{\text{dipole}}} \approx 0.66\left[\frac{\cos^2\left(\frac{\pi}{2}\cos\theta\right)}{\sin^2\theta}\right] \tag{2.84}$$

So we get that the dipole approximated angular power distribution has radiation resistance about 2/3 that of the actual angular power distribution, but when we just care about the antenna pattern, this factor is of no significance and the dipole approximation actually works really well.

Chapter 3

Scattering

When we talk about scattering, the starting point is to consider a plane wave incoming and interacting with a target at the origin, which then produces an outgoing spherical wave as illustrated in figure 3.1 below.

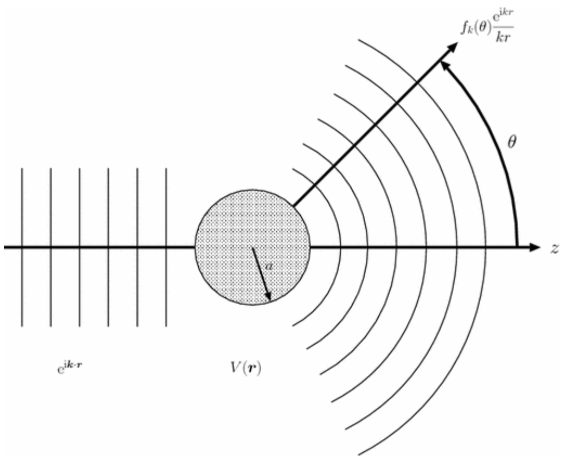

Figure 3.1: Scattered electromagnetic wave.

The scalar field ansatz for such a system can be written as:

$$\psi(\vec{x}) = E_0 e^{ikz} + \frac{e^{ikr}}{r} F(\hat{n}, \hat{n}_0) \tag{3.1}$$

where $\hat{n}_0$ is the direction of the incoming beam, and $\hat{n}$ is the direction of detector from the target. we also define

$$f(\hat{n}, \hat{n}_0) \equiv \frac{F(\hat{n}, \hat{n}_0)}{E_0} \tag{3.2}$$

$$\Rightarrow \quad \boxed{\psi(\vec{x}) = E_0 \left[e^{ikz} + \frac{e^{ikr}}{r} f(\hat{n}, \hat{n}_0) \right]} \tag{3.3}$$

where $f(\hat{n}, \hat{n}_0)$ is known as the *scattering amplitude* which gives a measure of the strength of the scattered wave. It turns out that we can think of the scattered wave is analogous to radiation from a source, where the radiation is due to the induced multipole moment from the incoming wave. This allows us to then write we have:

$$\frac{dP_{\text{scatt}}}{d\Omega} \sim r^2 \left| \frac{e^{ikr}}{r} f(\hat{n}, \hat{n}_0) \right|^2 \tag{3.4}$$

for which the angular power distribution is related to a quantity known as the *differential scattering cross-section*:

$$\frac{d\sigma}{d\Omega} = |f(\hat{n}, \hat{n}_0)|^2 \sim \frac{1}{E_0^2} \frac{dP}{d\Omega}. \tag{3.5}$$

The differential cross-section tells us "*the ratio of the scattered electric field flux of a scatterer into a differential solid to the incident electric field flux*". This will be presented mote formally in just a bit, but first we note that in general, there are 2 paths to getting solutions to scattering problems. These are:

1. Exact Solutions: Analytic results with direct solutions or exact parameterizations (e.g. optical theorem 3.0.1).
2. Approximate Treatments: Often involves perturbation theory where the scatterer is weak (e.g. Feynman diagrams), or some useful kinematic limit.

Theorem 3.0.1. Optical Theorem:

$$\sigma = \int \frac{d\sigma}{d\Omega} d\Omega = \frac{4\pi}{k} \operatorname{Im}\{f(\hat{n}, \hat{n}_0)\} \tag{3.6}$$

Exact methods are, as you would expect, in general difficult. So we differ to approximation schemes in which the results are more often than not, precise enough for all intends and purposes and allow us to extract useful physics.

§3.1 Long-Wavelength Scattering

Let's start by considering a scattering event whereby the incident electric field on a target has a large wavelength. That is, we have the condition that $\lambda \gg d$ with d being the effective diameter of the scatterer. In this regime, we have that the incident field is approximately **uniform across the scatter**, which grants that the problem is analogous to dipole radiation where the dipole moment is induced by the incident field. As such, we can immediately write down the scattered wave angular power distribution as the dipole power radiation formula:

$$\frac{dP_{\text{scatt}}}{d\Omega} = \frac{c}{8\pi} r^2 |E_{\text{out}}|^2 \sim |E_{\text{out}}|^2 \tag{3.7}$$

Then, we consider the electromagnetic power flux entering the system (incident on the target) which is in fact just the time-averaged Poynting vector:

$$\Phi_{\text{in}} = \frac{c}{8\pi} |E_{\text{in}}|^2, \tag{3.8}$$

from which we can define, from the qualitative description of the **differential cross-section** given in the previous section, the quantity:

$$\boxed{\frac{d\sigma}{d\Omega} = \frac{1}{\Phi_{\text{in}}}\frac{dP_{\text{scatt}}}{d\Omega}} \tag{3.9}$$

This is the analogous quantity from that used in classical hard sphere scattering processes. To state again, "*the differential scattering cross-section is the ratio of the scattered radiation flux off a scatterer/target into some differential solid angle $dP_{scatt}/d\Omega$, to the incident radiation flux on the scatterer/target Φ_{in}*". The 2 main ways we are going to compute this is via the following 2 methods.

1. Direct use of the induced dipole moment;
2. *Born's approximation.*

§3.1.1 Induced Dipole Scattering

Let's start with method 1. Clearly, this method is only useful when the induced dipole moment is a quantity that can be easily computed. This happens when the geometry of the scatterer is simple like a perfect sphere. As such, we consider a perfect dielectric sphere of radius a, for which radiation is incident on it with a wavelength $\lambda \gg a$. The incident fields can be modeled as plane waves which are written as:

$$\begin{aligned} \vec{E}_0(\vec{x},t) &= \hat{\varepsilon}_0 E_0 e^{i(k\hat{n}_0\cdot\vec{x}-\omega t)} \\ \vec{B}_0(\vec{x},t) &= \hat{n}_0 \times \vec{E}_0(\vec{x},t) \end{aligned} \tag{3.10}$$

Since $\lambda \gg a$, we treat the incident electric field as uniform across the dielectric sphere which is the basis of the long wavelength approximation. The induced electric dipole moment would then be given by:

$$\vec{p}(t) = \left[\frac{\epsilon-1}{\epsilon+2}\right] a^3 \vec{E}_0(t) \tag{3.11}$$

in Gaussian units. This result follows directly from the dipole moment generated by a dielectric sphere in a uniform electric field as done in Jackson section 4.4 (page 158). The scattered fields are then easily computed via the electric dipole radiation field formulas:

$$\vec{B}_s = \frac{k^2 e^{ikr}}{r}\left(\hat{n}\times\vec{p}\right), \quad \vec{E}_s = -\frac{k^2 e^{ikr}}{r}\left[\vec{p} - \hat{n}\left(\hat{n}\cdot\vec{p}\right)\right] \tag{3.12}$$

$$\Rightarrow \quad \frac{dP}{d\Omega} = \frac{cr^2}{8\pi}|E_s|^2 \tag{3.13}$$

In the above derivation, we realize that we did not consider the polarization of the electric field. However, this can conveniently be accounted for by writing the scattered polarization $\hat{\varepsilon}$ as a linear combination over polarization axes (inserting the resolution of the identity):

$$
\begin{aligned}
\vec{E}_s &= \sum_j \hat{\varepsilon}_j \left(\hat{\varepsilon}_j^* \cdot \vec{E}_s\right) \\
\Rightarrow \quad \frac{dP}{d\Omega}(\hat{\varepsilon}) &= \frac{cr^2}{8\pi}\left|\hat{\varepsilon}^* \cdot \vec{E}_s\right|^2 \\
&= \frac{ck^4}{8\pi}\|\vec{p}\|^2 |\hat{\varepsilon}^* \cdot \hat{\varepsilon}_0|^2 \\
&= \frac{ck^4a^6}{8\pi}\left(\frac{\epsilon-1}{\epsilon+2}\right)^2 |\hat{\varepsilon}^* \cdot \hat{\varepsilon}_0|^2 E_0^2
\end{aligned}
\tag{3.14}
$$

where we noted above that $\hat{\varepsilon}^* \cdot \hat{n} = 0$ since the polarization would always be orthogonal to the direction of propagation of the scattered wave (for the systems we are concerned with in this class at least). As such, we can compute the differential cross-section:

$$
\begin{aligned}
\frac{d\sigma}{d\Omega} &= \frac{1}{\Phi_{\text{in}}}\frac{dP}{d\Omega} \\
&= \frac{k^4|\hat{\varepsilon}^* \cdot \vec{p}|}{E_0^2} = \left|\frac{\epsilon-1}{\epsilon+2}\right|^2 k^4a^6|\hat{\varepsilon}^* \cdot \hat{\varepsilon}_0|^2.
\end{aligned}
\tag{3.15}
$$

The derivation above was pretty straight forward, for which the final expression above is particular for dipole scattering in this spherical target geometry, however the k^4 dependence of the differential cross-section is in fact a **universal** result and known as *Rayleigh's law*:

$$
\boxed{\frac{d\sigma}{d\Omega} \propto k^4}. \tag{3.16}
$$

In practice, we also want to know how to compute $|\hat{\varepsilon}^* \cdot \hat{\varepsilon}_0|^2$. Often, what we normally do is average over all polarizations since it is common that the incoming field unpolarized. To do this averaging, it is convenient to pick coordinates that make in the relative (incoming vs scattered) polarizations convenient for us to work with. The convenient coordinate system is one whereby a plane is defined such that we have a parallel and perpendicular component of the polarizations relative to this plane. A visualization of this is given in figure 3.2 below.

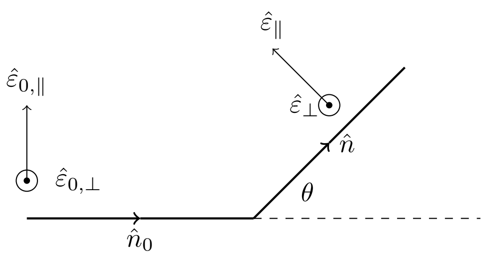

Figure 3.2: Visualization of a scattering process.

From figure 3.2, we can derive the following projection relations of the relative polarizations:

$$\begin{aligned} \hat{\varepsilon}_0^{(1)} \cdot \hat{\varepsilon}^{(1)} &= \cos\theta \\ \hat{\varepsilon}_0^{(2)} \cdot \hat{\varepsilon}^{(2)} &= 1 \\ \hat{\varepsilon}_0^{(i)} \cdot \hat{\varepsilon}^{(j)} &= 0, \quad i \neq j \end{aligned} \tag{3.17}$$

So these give us the differential cross-section of incoming radiation polarized in and out of the scattering plane as:

$$\begin{aligned} \text{in-plane}: \quad \frac{d\sigma_\parallel}{d\Omega} &= \frac{1}{2}\sigma_0 \cos^2\theta \\ \text{out-of-plane}: \quad \frac{d\sigma_\perp}{d\Omega} &= \frac{1}{2}\sigma_0 \end{aligned} \tag{3.18}$$

where $\sigma_0 \equiv \left|\frac{\epsilon-1}{\epsilon+2}\right|^2 k^4 a^6$.

§3.1.2 Scattering Off Perfect Conducting Spheres

Recall from electro and magnetostatics, we had that the with a perfectly conducting sphere of radius a, we had the fields:

$$\vec{E}_{in} = \frac{3}{\epsilon_r + 2}\vec{E}_{out} \tag{3.19}$$

$$\vec{B}_{in} = \frac{3\mu}{\mu + 2}\vec{B}_{out} \tag{3.20}$$

The dipole moments due to the induced currents on the conducting sphere are then:

$$\vec{p} = \left(\frac{\epsilon_r - 1}{\epsilon_r + 2}\right) a^3 \vec{E}_0, \quad \vec{m} = \left(\frac{\mu - 1}{\mu + 2}\right) a^3 \vec{B}_0 \tag{3.21}$$

So to get what a scattered wave would look like off this conducting sphere, we simply tack on the harmonic field time-dependence and take the sum of the 2 field emitted due to the electric and magnetic dipole moments as:

$$\begin{aligned} \vec{E}_{scatt} &= \vec{E}_E + \vec{E}_B \\ &= -k^2 \frac{e^{ikr}}{r}\hat{n} \times (\hat{n} \times \vec{p}) - k^2 \frac{e^{ikr}}{r}\hat{n} \times \vec{m} \\ &= -k^2 a^3 E_0 \frac{e^{ikr}}{r}\hat{n} \times (\hat{n} \times \hat{\varepsilon}_0) + \frac{k^2 a^3}{2} E_0 \frac{e^{ikr}}{r}\hat{n} \times (\hat{n} \times \hat{\varepsilon}_0) \end{aligned} \tag{3.22}$$

This then allows us to compute the differential cross-section as:

$$\frac{d\sigma}{d\Omega}(\hat{\varepsilon}, \hat{n}; \hat{\varepsilon}_0, \hat{n}_0) = k^4 a^6 \left| \hat{\varepsilon}^* \cdot \hat{\varepsilon}_0 - \frac{1}{2}(\hat{n} \times \hat{\varepsilon}^*) \cdot (\hat{n}_0 \times \hat{\varepsilon}_0) \right|^2 \tag{3.23}$$

Now averaging over $\hat{\varepsilon}_0$, we get:

$$\frac{d\sigma_\parallel}{d\Omega} = \frac{k^4a^6}{2}\left|\cos\theta - \frac{1}{2}\right|^2; \quad \frac{d\sigma_\perp}{d\Omega} = \frac{k^4a^6}{2}\left|1 - \frac{1}{2}\cos\theta\right|^2 \tag{3.24}$$

$$\Rightarrow \quad \frac{d\sigma}{d\Omega} = k^4a^6\left[\frac{5}{8}(1+\cos^2\theta) - \cos\theta\right] \tag{3.25}$$

We then defined the metric of back-scattering:

$$\boxed{\Pi(\theta) \equiv \frac{\frac{d\sigma_\perp}{d\Omega} - \frac{d\sigma_\parallel}{d\Omega}}{\frac{d\sigma_\perp}{d\Omega} + \frac{d\sigma_\parallel}{d\Omega}}} \tag{3.26}$$

For which in the case of perfect conducting sphere scattering, this metric evaluates to:

$$\Pi(\theta) = \frac{3\sin^2\theta}{5(1+\cos^2\theta) - 8\cos\theta} \tag{3.27}$$

More details on this are also given in Jackson section 10.1 (pages 456 - 460).

§3.2 The Born Approximation

We are now going to look at how to approach scattering problems with method 2, that is by using the *Born's approximation*. This method is a perturbative treatment of sorts, and assumes that the scattered field is small compared to the incident field on the scatterer. The way we can see anything in the world, is through the fact that the dielectric constant of objects is different from that in air. More specifically, the fluctuations in $\epsilon(\vec{x})$ or $\mu(\vec{x})$ induces scattering. To see this, we of course start from Maxwell's equations:

$$\begin{aligned} &\boldsymbol{\nabla}\cdot\vec{D} = 4\pi\rho, \quad \boldsymbol{\nabla}\cdot\vec{B} = 0 \\ &\boldsymbol{\nabla}\times\vec{E} + \frac{1}{c}\frac{\partial\vec{B}}{\partial t} = 0, \quad \boldsymbol{\nabla}\times\vec{H} - \frac{4\pi}{c}\vec{J} - \frac{1}{c}\frac{\partial\vec{D}}{\partial t} = 0, \end{aligned} \tag{3.28}$$

from which we derive that:

$$\begin{aligned} &\boldsymbol{\nabla}\times\Big(\boldsymbol{\nabla}\times(\vec{D}-\vec{E})\Big) = \boldsymbol{\nabla}\Big(\boldsymbol{\nabla}\cdot\vec{D}\Big) - \boldsymbol{\nabla}^2\vec{D} + \boldsymbol{\nabla}\left(\frac{1}{c}\vec{B}\frac{\partial\vec{B}}{\partial t}\right) \\ &\frac{\partial}{\partial t}\left(\frac{\partial\vec{D}}{\partial t}\right) = c\frac{\partial}{\partial t}\Big(\boldsymbol{\nabla}\times\vec{H}\Big) \\ \Rightarrow \quad &\left(\boldsymbol{\nabla}^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\vec{D} = -\boldsymbol{\nabla}\times\left[\boldsymbol{\nabla}\times\Big(\boldsymbol{\nabla}\times(\vec{D}-\vec{E})\Big) + \frac{1}{c}\frac{\partial}{\partial t}\boldsymbol{\nabla}\times(\vec{B}-\vec{H})\right] \end{aligned} \tag{3.29}$$

Then taking $D \sim e^{-i\omega t}$, we get:

$$\begin{aligned}
&\left(\boldsymbol{\nabla}^2 + k^2\right)\vec{D} = -\boldsymbol{\nabla}\times\left[\boldsymbol{\nabla}\times\left(\boldsymbol{\nabla}\times(\vec{D}-\vec{E})\right) + \frac{1}{c}\frac{\partial}{\partial t}\boldsymbol{\nabla}\times(\vec{B}-\vec{H})\right] \\
\Rightarrow \quad &\vec{D}(\vec{x}) = \vec{D}_0(\vec{x}) + \vec{D}_{scatt}(\vec{x}) \\
\text{where} \quad &\vec{D}_{scatt}(\vec{x}) = \int d^3x G_k(\vec{x},\vec{x}')\left(-\boldsymbol{\nabla}\times\left[\boldsymbol{\nabla}\times\left(\boldsymbol{\nabla}\times(\vec{D}-\vec{E})\right) + \frac{1}{c}\frac{\partial}{\partial t}\boldsymbol{\nabla}\times(\vec{B}-\vec{H})\right]\right)
\end{aligned} \tag{3.30}$$

The Green's function can be taken as:

$$G_k(\vec{x},\vec{x}') = -\frac{1}{4\pi}\frac{e^{ikR}}{R} \approx -\frac{1}{4\pi}\frac{e^{ikr}}{r}e^{-ik\hat{n}\cdot\vec{x}'} \tag{3.31}$$

which effectively allows us to simplify our expression for $\vec{D}$ scattered off the target as:

$$\begin{aligned}
&\vec{D}_{scatt} = \frac{e^{ikr}}{r}\hat{n}\times\left(\hat{n}\times\vec{f}_s(\hat{n})\right) \\
\text{where} \quad &\vec{f}_s(\hat{n}) = -\frac{k^2}{4\pi}\int d^3x' e^{-ik\hat{n}\cdot\vec{x}'}\left[\vec{D}(\vec{x}') - \vec{E}(\vec{x}')\right] \\
\Rightarrow \quad &\frac{d\sigma}{d\Omega}(\hat{\varepsilon},\hat{\varepsilon}_0) = \frac{\left|\hat{\varepsilon}^*\cdot\vec{f}_s(\hat{n})\right|^2}{|D_0|^2}
\end{aligned} \tag{3.32}$$

where the scattering amplitude is now a vector since we are now working with vector fields. Strangely, the scattering amplitude is dependent on the total D and E field, which if we had we wouldn't need to compute all these things in the first place! To resolve this, we will assume that the variation in the relative dielectric constant can be written as a perturbation from unity:

$$\epsilon(\vec{x}) = 1 + \delta\epsilon(\vec{x}) \tag{3.33}$$

where $\delta\epsilon(\vec{x}) \ll 1$. The *Born approximation* is then the approximation where we assume that we can approximate the total field as just the incident field and simply replace $\vec{D} \to \vec{D}_0$, $\vec{E} \to \vec{E}_0$ to give us:

$$\vec{E} = E_0\hat{\varepsilon}_0 e^{i\vec{k}\cdot\vec{x}}. \tag{3.34}$$

This allows the reduction:

$$\vec{D}(\vec{x}) - \vec{E}(\vec{x}) \quad \to \quad \vec{D}_0(\vec{x}) - \vec{E}_0(\vec{x}) = \left[\epsilon(\vec{x})E_0 - E_0\right]e^{i\vec{k}\cdot\vec{x}}\hat{\varepsilon}_0 \tag{3.35}$$

$$\Rightarrow \quad \boxed{\vec{F}_{Born} = -\frac{k^2}{4\pi}E_0\hat{\varepsilon}_0\int d^3x' e^{i(\vec{k}_0-\vec{k})\cdot\vec{x}'}\delta\epsilon(\vec{x}') = -\frac{k^2}{4\pi}E_0\hat{\varepsilon}_0\tilde{\delta\epsilon}(\vec{k})} \tag{3.36}$$

So we have that the Born approximation reduces the scattering amplitude to a Fourier transform of the dielectric variation in space into k-space. The differential cross-section also with a variation in μ ($\mu(\vec{x}) = 1 + \delta\mu(\vec{x})$) is then:

$$\boxed{\frac{d\sigma}{d\Omega} = \frac{k^4}{16\pi^2}\left|\hat{\varepsilon}^*\cdot\hat{\varepsilon}_0\int d^3x' e^{i(\vec{k}_0-\vec{k})\cdot\vec{x}'}\delta\epsilon(\vec{x}') + (\hat{n}\times\hat{\varepsilon})^*\cdot(\hat{n}_0\times\hat{\varepsilon}_0)\int d^3x' e^{i\vec{q}\cdot\vec{x}'}\delta\mu(\vec{x}')\right|^2} \tag{3.37}$$

Notice that in this approximation, we are keeping all the k-values and using the assertion that:

$$\frac{\delta\epsilon}{\epsilon}, \ \frac{\delta\mu}{\mu} \ll 1 \tag{3.38}$$

This was all for individual scatterers and individual cross-sections. What about if we had an ensemble of scatterers?

§3.2.1 Multi-Scatterer Systems

To first tackle this, let's first consider our coordinate system is such that the scatterer is **not** centered at the origin (which is what we have been assuming all this time). The electric field (an similarly also the magnetic field) would then be:

$$\vec{E}_{\text{in}}(\vec{x}) = \hat{\varepsilon}_0 E_0 e^{i(k\hat{n}_0 - \vec{x}_0)} \tag{3.39}$$

where $\vec{x}_0$ is the scattering center. We now define $\vec{x}' = \vec{x}_0 + \vec{x}''$ where $\vec{x}''$ is some convenient relative coordinate. The current density can then be written as:

$$\vec{J}(\vec{x}') = e^{ik\hat{n}_0\cdot\vec{x}_0}\vec{J}(\vec{x}'') \tag{3.40}$$

We can then use this to construct the vector potential in the far-field limit as:

$$\begin{aligned} \vec{A}(\vec{x}) &= \frac{e^{ikr}}{r}\int e^{ik(\hat{n}_0-\hat{n})\cdot\vec{x}_0} e^{-ik\hat{n}\cdot\vec{x}''}\vec{J}(\vec{x}'')d^3x'' \\ &= \frac{e^{ikr}}{r}\int e^{ik\vec{q}\cdot\vec{x}_0} e^{-ik\hat{n}\cdot\vec{x}''}\vec{J}(\vec{x}'')d^3x'' \end{aligned} \tag{3.41}$$

Notice that the phase $e^{ik\vec{q}\cdot\vec{x}_0}$ is not being integrated over, so it actually plays no role when we are concerned with differential cross-sections (normally left out in textbooks including Jackson). The resulting many-scatterer differential cross-section is then:

$$\boxed{\frac{d\sigma}{d\Omega} = \frac{k^4}{E_0^2}\left|\sum_j \hat{\varepsilon}^* \cdot \vec{p}_j e^{i\vec{q}\cdot\vec{x}_j}\right|^2} \tag{3.42}$$

where j denotes the scatterer index. So we infact have that:

$$\vec{F}(\hat{n}, \vec{x}_j) = \vec{F}(\hat{n}, 0)e^{i\vec{q}\cdot\vec{x}_j} \tag{3.43}$$

If we then assume **identical scatterers**, we then get:

$$\boxed{\frac{d\sigma}{d\Omega} = \frac{\left|\hat{\varepsilon}^* \cdot \vec{F}(\hat{n}, 0)\right|^2}{E_0^2}\left|\sum_j e^{i\vec{q}\cdot\vec{x}_j}\right|^2} \tag{3.44}$$

The absolute-squared sum term is called several things in different fields, such as the *structure factor* is atomic physics and *form factor* in particle physics. We can thus rewrite the many-scatterer differential cross-section as:

$$\begin{aligned} \frac{d\sigma}{d\Omega} &= \frac{d\sigma_0}{d\Omega}(\hat{\varepsilon})\, F(\vec{q}) \\ \text{where} \quad \frac{d\sigma_0}{d\Omega}(\hat{\varepsilon}) &\equiv \frac{\left|\hat{\varepsilon}^* \cdot \vec{F}(\hat{n}, 0)\right|^2}{E_0^2}, \quad F(\vec{q}) \equiv \left|\sum_j e^{i\vec{q}\cdot\vec{x}_j}\right|^2 \end{aligned} \tag{3.45}$$

2 common scenarios for many-scatterer systems are

1. <u>regular array of scatteres</u>: (leads to Bragg peaks);
2. <u>randomly-positioned scatterers</u>: The structure factor in such systems can also be:

$$F(\vec{q}) = \sum_{ij} e^{i\vec{q}\cdot(\vec{x}_i - \vec{x}_j)} \tag{3.46}$$

however in the sum above, all the terms where $i \neq j$ sum to zero due to their random positions. As such, we have only the diagonal ($i = j$) terms as non-trivial which grants:

$$\begin{aligned} F(\vec{q}) &= \sum_{j=1}^{N} 1 = N \\ \Rightarrow \quad & \boxed{\frac{d\sigma}{d\Omega} = N \frac{d\sigma_0}{d\Omega}} \end{aligned} \tag{3.47}$$

This result is known as *incoherent scattering*. However in the **forward direction** ($\vec{q} = \vec{k}_0 - \vec{k} \to 0$), we instead get:

$$\boxed{\frac{d\sigma}{d\Omega} = N^2 \frac{d\sigma_0}{d\Omega}} \tag{3.48}$$

This is known as *coherent scattering*, but have peaks that are extremely narrow.

Now consider an incoming electromagnetic wave entering some small volume of cross-sectional area A and width Δr. We now want to know how the scattering beam is attenuated by power losses through the this volume. In general, we have that the power loss through a volume is given by:

$$\begin{aligned} \text{power loss through volume} = &\left[\frac{\text{incident flux}}{\text{unit area}}\right] \\ &\times \left[\frac{\text{power loss per mole}}{\text{incident flux}} = \sigma\right] \\ &\times [\text{number of molecules}] \end{aligned} \tag{3.49}$$

How can we relate the structure factor, to the computation of such a loss? Well, we first write:

$$F(\vec{q}) = \left|\sum_j e^{i\vec{q}\cdot\vec{x}_j}\right|^2 = \left|\int d^3x \delta n(\vec{x}) e^{i\vec{q}\cdot\vec{x}}\right|^2 \tag{3.50}$$

where $\delta(\vec{x}) \equiv n(\vec{x}) - \overline{n}$ is the fluctuation in density of the medium the incoming field is traversing with:

$$\overline{n} = \frac{1}{V}\int d^3x n(\vec{x}) \tag{3.51}$$

Now, we can take the average over the molecular ensemble:

$$\begin{aligned} F(\vec{q}) &= \left\langle \left| \int d^3x \delta n(\vec{x}) e^{i\vec{q}\cdot\vec{x}} \right|^2 \right\rangle \\ &= \left\langle \int d^3x \delta n(\vec{x}) e^{i\vec{q}\cdot\vec{x}} \int d^3y \delta n(\vec{y}) e^{-i\vec{q}\cdot\vec{y}} \right\rangle \\ &= \int d^3x d^3y e^{i\vec{q}\cdot(\vec{x}-\vec{y})} \left\langle \delta n(\vec{x}) \delta n(\vec{y}) \right\rangle = V \int d^3r e^{i\vec{q}\cdot\vec{r}} \left\langle \delta n(\vec{r}) \delta n(\vec{0}) \right\rangle \end{aligned} \tag{3.52}$$

where the last simplification came from the appropriate change of variable as commonly done in statistical mechanics. So we see we get some kind of correlation function of the fluctuations between 2 different positions. So in fact see that scattering light off some volume of scatterers can tell us about the structure of the material. We note that:

$$\begin{aligned} \lim_{q\to 0} F(\vec{q}) &= V \int d^3r \left\langle \delta n(\vec{r}) \delta n(\vec{0}) \right\rangle \\ &= \langle N^2 \rangle - \langle N \rangle^2 \end{aligned} \tag{3.53}$$

which is just the variance of the number of particles in the volume.

Example:

Consider an experiment which tries to perform the magnetic scattering over a material's local spin density. We then have that:

$$\begin{aligned} F_{spin}(\vec{q}) &= \int d^3r \left\langle (\sigma(r) - \overline{\sigma}) (\sigma(0) - \overline{\sigma}) \right\rangle \\ \Rightarrow \quad \lim_{q\to 0} F_{spin}(\vec{q}) &= \langle \sigma^2 \rangle - \langle \sigma \rangle^2 \end{aligned} \tag{3.54}$$

This quantity $\langle \sigma^2 \rangle - \langle \sigma \rangle^2$ is in fact the magnetic susceptibility χ, which is also computed as:

$$\chi = \left.\frac{\partial M}{\partial h}\right|_{h=0} = \langle \sigma^2 \rangle - \langle \sigma \rangle^2 \tag{3.55}$$

Example:

Consider water. The P-T phase diagram of water has gaseous, liquid and solid phases, for which liquids and gases get indistinguishable (in terms of correlation functions) past the *critical point* ($P = 218$ atm, $T = 647$ K). At the critical point, we have that:

$$\begin{aligned} \frac{\partial P}{\partial V}\bigg|_T^{(\text{critical})} &= 0 \\ \Rightarrow \quad k_T = -\frac{1}{V}\frac{1}{\frac{\partial P}{\partial V}\big|_T} &\to \infty \end{aligned} \tag{3.56}$$

This is known as *critical opalescence*, which causes the differential scattering cross-section to be extremely large and allows a measurement of this phase via light scattering experiments.
What exactly is happening at these critical points then? Consider again the structure factor as a correlation function:

$$F(\vec{q}) \sim \int \langle \delta n(\vec{r})\delta n(0)\rangle e^{i\vec{q}\cdot\vec{r}} d^3r \tag{3.57}$$

The correlation function can be modelled by the *Orenstein-Zernike* function:

$$\langle \delta n(\vec{r})\delta n(0)\rangle = \frac{e^{-r/\xi}}{r} \tag{3.58}$$

where ξ is the *correlation length*. Using this model, this grants us:

$$F(q) \sim \frac{1}{q^2 + 1/\xi^2} \tag{3.59}$$

At these critical points, what happens is that the correlation length becomes extremely long (note that this is a correlation length of the fluctuations). When this happens, $F(q)$ goes like $1/q^2$ and so in the limit where $q \to 0$, the differential cross-section indeed blows up.

Example:

Now we can also ask a common question of why is the sky blue? Molecular polarizability is given by:

$$\begin{aligned} \vec{p}_j &= \gamma_{mol} E_0 \hat{\varepsilon}_0 \\ \Rightarrow \quad \sigma &= \frac{8\pi}{3} k^4 \gamma_{mol}^2 N = \sigma_0 N \end{aligned} \tag{3.60}$$

So because we have this k^4 dependence, light with higher frequencies get scattered more often which is indeed why we see blue skies. More information if given in Jackson 10.2c (pages 465 - 468).

§3.2.2 The Optical Theorem

Consider having a scattered electric field as:

$$\hat{\varepsilon}^* \cdot \vec{E}_{sc} = \frac{e^{ikr}}{r} \hat{\varepsilon}^* \cdot \vec{f}(\vec{k}_0, \vec{k}) E_0 \tag{3.61}$$

with the incident wave being:

$$\vec{E} = \hat{\varepsilon}_0 E_0 e^{i\vec{k}_0\cdot\vec{r}} \tag{3.62}$$

The cross-section formulas are also given by:

$$\frac{d\sigma}{d\Omega} = \left|\hat{\varepsilon}^* \cdot \vec{f}\right|^2, \quad \sigma = \int d\Omega \frac{d\sigma}{d\Omega} \tag{3.63}$$

The optical theorem is then given by:

$$\boxed{\sigma_{\text{total}} = \frac{4\pi}{k_0} \operatorname{Im}\left\{\hat{\varepsilon}_0 \cdot \vec{f}(\vec{k}_0, \vec{k} = \vec{k}_0\right\}} \tag{3.64}$$

The way the theorem is derived is by considering 2 spherical surfaces S_1 and S_2 enclosing the scatterer, and then σ_{total} decomposing the cross-section into an elastic and a absorption cross-section to give $\sigma_{\text{total}} = \sigma_{\text{elastic}} + \sigma_{\text{absorption}}$ (the $\sigma_{\text{absorption}}$ constitutes an inelastic scattering cross-section in quantum mechanics). The mathematical details are sorted out in Jackson. The result however, is very useful because it is an exact result (corresponding to probability conservation in quantum mechanics).

> **Note**: Approximation schemes to find scattering cross-sections often do not result in them obeying the optical theorem (e.g. in perturbative treatments).

§3.3 Diffraction

"*Diffraction is basically scattering from a hole*".

− T. DeGrand, 2020

We are going to start this off by talking about scalar fields (much like sound waves) rather than vector fields. This makes the analysis simpler for now, and we can just treat the scalar field as some component of a vector field $\vec{E}$ or $\vec{B}$. As a broad overview, the set-up of a diffraction problem can be described by 2 surfaces S_1 and S_2, in which a wave is first incident S_1, diffracted through it due to some geometrically specified "openings" (apertures), then observed by detectors on surface S_2.

In most cases we are concerned with, S_1 is a flat-plane with some rectangular or circular apertures and S_2 is some detector situated far away from S_1. Formally, we can take the incoming wave to be plane-waves (since the governing equation is linear which will allow solutions to be constructed by linear combination), resulting in us dealing with solutions to the Helmholtz equation:

$$\left(\boldsymbol{\nabla}^2 + k^2\right)\psi = 0 \tag{3.65}$$

$$\Rightarrow \quad \left(\boldsymbol{\nabla}^2 + k^2\right) G(\vec{x}, \vec{x}'; k) = -\delta^3(\vec{x} - \vec{x}') \tag{3.66}$$

We recall Green's theorem:

$$\begin{aligned} &\oint_S (\mathcal{G}(x,x')\frac{\partial\psi}{\partial n'} - \psi\frac{\partial\mathcal{G}}{\partial n'})dA' = \int_V \left(\mathcal{G}(x,x')\boldsymbol{\nabla}^2_{x'}\psi(x') - \psi(x')\boldsymbol{\nabla}_{x'}\mathcal{G}(x,x')\right) d^3x' \\ \Rightarrow \quad &\psi(\vec{x}) = \int_S \left[\mathcal{G}(x,x')\hat{n}' \cdot \boldsymbol{\nabla}_{x'}\psi(\vec{x}') - \psi(\vec{x}')\hat{n}' \cdot \boldsymbol{\nabla}_{x'}\mathcal{G}(x,x')\right] dA' \qquad \text{(Jackson 10.75)} \end{aligned} \tag{3.67}$$

Taking the Green's function as:

$$\mathcal{G}(x,x') = \frac{1}{4\pi}\frac{e^{i\vec{k}\cdot(\vec{x}-\vec{x}')}}{r} \tag{3.68}$$

$$\Rightarrow \quad \psi(\vec{x}) = \int_S \hat{n}'\cdot\left[\nabla_{x'}\psi(\vec{x}') - \psi(\vec{x}')\frac{\vec{R}}{R}\left(ik - \frac{1}{R}\right)\right]\frac{e^{i\vec{k}\cdot\vec{R}}}{R}dA' \tag{3.69}$$

where $\vec{R} \equiv \vec{x} - \vec{x}'$. Green's theorem relates the solution in a volume to that on the boundary of that volume, so we consider the physical argument where if we pick the boundary to lie on the diffraction screen S_2 and a great-dome that extends to infinity, we have terms that cancel at infinity. That is, in the region where r' is very large, we have a result known as the *Kirchhoff's integral formula*:

$$\boxed{\psi(x) = -\frac{1}{4\pi}\int dA'\frac{e^{ikR}}{R}\hat{n}'\cdot\left[\nabla_{x'}\psi(x') + \frac{\vec{R}}{R}\psi(x')\right]} \tag{3.70}$$

There are 3 things we need to keep in mind however, when we use the Kirchoff integral formula.

1. Mathematical consistencies:

   To be consistent mathematically, solutions to the Helmholtz equation only admit Dirichlet **or** Neumann boundary conditions (**not** both). However we somehow want Neumann conditions for the incoming wave and Dirichlet conditions for the outgoing wave. Let's just first think about how to satisfy both boundary conditions. To do this, we construct Green's functions that vanish adequately at the boundary (or its derivative) by the method of images, which gives:

$$\begin{aligned}
\text{Dirichlet}: \quad \mathcal{G}_D(x,x') &= \frac{1}{4\pi}\left[\frac{e^{ikR}}{R} - \frac{e^{ikR''}}{R''}\right] \\
\text{Neumann}: \quad \mathcal{G}_N(x,x') &= \frac{1}{4\pi}\left[\frac{e^{ikR}}{R} + \frac{e^{ikR''}}{R''}\right]
\end{aligned} \tag{3.71}$$

$$\begin{aligned}
\Rightarrow \quad \text{Dirichlet}: \quad \psi_D(x) &= \frac{1}{4\pi}\int_{S_1}\psi(x')2\left(\frac{\hat{z}\cdot\vec{R}}{R}\right)\frac{e^{ikR}}{R} \\
&= \frac{k}{2\pi i}\int_{S_1}dA'\left(\frac{e^{ikr'}}{r'}\right)\frac{e^{ikr}}{r}\cos\theta \\
\text{Neumann}: \quad \psi_N(x) &= \frac{1}{4\pi}\int_{S_1}\psi(x')\frac{e^{ikR}}{R}2\hat{n}'\cdot\nabla_{x'}\psi(x') \\
&= \frac{k}{2\pi i}\int_{S_1}dA'\left(\frac{e^{ikr'}}{r'}\right)\frac{e^{ikr}}{r}\cos\theta'
\end{aligned} \tag{3.72}$$

   where θ is the angle of the outgoing wave from the aperture normal and θ' the angle of the incoming wave.

2. Practical considerations:

Often times in optics, we have that $\psi_a(x') = t(x')\psi_{source}(x_0)$ where ψ_a is the wavefunction at the aperture and $t(x')$ is some transfer function. However, for most cases we care about, we can simply at $t(x') = 1$ which makes life easier.

A thing that is relevant to us is the uncertainty principle in waves ($\Delta k \Delta x \geq 1$). This is due to the fact that we are squeezing the incoming wave through a small aperture, which then cause Δx to be small. This causes a spreading over the wave-vectors and a spreading of the incoming wave onto its "*geometrical shadow*". We can then take the approximation with quadratures as:

$$\Delta x = \sqrt{a^2 + \left(\frac{r}{ak}\right)^2} \tag{3.73}$$

where a is the intrinsic width of the beam whereas r is the distance of the screen from the aperture. The regime where $r/k \ll a^2$ (or $\lambda r \ll a^2$) is known as the *Fresnel limit*, which grants that what we see on the detector is just the shape of the hole since $\Delta x \to a$. Alternatively, we have the *Fraunhofer limit* where $\lambda r \gg a^2$ (we can use the far-field approximations we did for antennas or scattering).

Both of the approximations above are in fact far field approximations and only differ in terms of the wavelength relative to the aperture size. This difference in application lies in how e^{ikR} is expanded (recalling $R = \|\vec{x} - \vec{x}'\|$). Let's consider an important expansion of kR for these approximations:

$$\boxed{kR = kr - k\hat{n} \cdot \vec{x}' + \frac{k^2}{2r}\left[(r')^2 - (\hat{n} \cdot \vec{x}')^2\right] + \dots} \tag{3.74}$$

where $r = |x|$ and $r' = |x|'$. Noting that $k \propto 1/\lambda$, have that in each approximation,

(a) Fraunhofer diffraction:

$$\begin{aligned} \frac{e^{ikR}}{R} &\approx \frac{e^{ikr}}{r} e^{-ik\hat{n}\cdot\vec{x}'} \\ \Rightarrow \quad & \boxed{\psi_D(r) = \frac{k}{2\pi i}\frac{e^{ikr}}{r} A\cos\theta \int_{\text{ap}} dA' e^{-iq\hat{n}\cdot\vec{x}'}} \end{aligned} \tag{3.75}$$

given that $\psi_0(\vec{x}') = Ae^{i\vec{k}_0\cdot\vec{x}'}$ where $\vec{q} = \vec{k}_0 - \vec{k}$ with the integral going over the area of the aperture. This tells us that the "*diffracted wave is just the* ***Fourier transform*** *of the incoming wave through the aperture*".

(b) Fresnel diffraction:

For this approximation, in the region close to the aperture, we can consider the diffraction that occurs off a knifes edge. That is, we consider diffraction just due to one edge of the aperture.

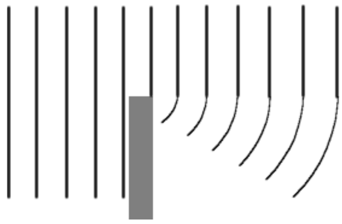

Figure 3.3: Diffraction of a knife's edge.

we can in fact a pick a coordinate system such that the linear term in the expansion in equation (3.74) goes away if we are considering diffraction off a knifes edge. We do this by picking the coordinate system to have $x = 0$ at the top edge of the lower aperture barrier, and $y = 0$ where the detector is (z is the optical axis). So we have:

$$\begin{aligned} R &= \sqrt{(x-x')^2 + (y')^2 + z^2} \\ &\approx z + \frac{1}{2z}\left[(x-x')^2 + (y')^2\right] \end{aligned} \tag{3.76}$$

where z is the large coordinate. We then also assert that $\hat{n} \cdot \hat{R} = 1$ (incident wave is normal to the barrier). So we can set up this problem as:

$$\begin{aligned} &\psi_0(\vec{x}') = \sqrt{I_0} e^{i\vec{k}_0 \cdot \vec{x}'} \\ \Rightarrow\ &\boxed{\psi_D(\vec{x}) = \frac{k}{2\pi i}\frac{e^{ikz}}{z}\sqrt{I_0}\int_0^\infty dx' \int_\infty^\infty dy' \exp\left(ik\left[\frac{(x-x')^2}{2z} + \frac{(y')^2}{2z}\right]\right)} \end{aligned} \tag{3.77}$$

3. Scalar to vector field modifications to the equation above:

   To now extend what we have been doing with scalar waves into vector fields, we use the *vector Smythe-Kirchhoff formula*, which is given as:

$$\boxed{\vec{E}(\vec{x}) = \frac{1}{2\pi}\boldsymbol{\nabla} \times \left[\int_{\text{ap}} \frac{e^{ikR}}{R}\hat{n} \times \vec{E}(\vec{x}')dA'\right]} \tag{3.78}$$

   Derivation of this is provided in Jackson section 10.7. This formula simplifies nicely in the Fraunhofer limit to:

$$\vec{E}_F(\vec{x}) = \frac{ik}{2\pi}E_0\frac{e^{ikr}}{r}\left[\hat{n} \times (\hat{z} \times \hat{\varepsilon}_0)\right]\int dA' \exp\left(i\left[\vec{k}_0 - k\hat{n}\right] \cdot \vec{x}'\right) \tag{3.79}$$

   which differs from the result we had for the scalar field mostly in polarization direction considerations. The magnetic field is another story because we have to consider the *complimentary surface* but this will usually not be of concern to us.

Note: The assumption of diffraction we will adopt here is that the medium that causes the diffraction is only modified by the geometry of the aperture and not the material itself.

Example:

Consider diffraction through a square aperture in the Fraunhofer limit. We have that:

$$\begin{aligned}\vec{k}_0 &= [0, 0, k]^T \\ \vec{k} &= [k_x, k_y, k_z]^T \\ \vec{x}' &= [x', y', 0]\end{aligned} \tag{3.80}$$

We then have that the integral over the incoming wave across the aperture is then given by:

$$\begin{aligned}\int_{\mathrm{ap}} e^{i\vec{k}_0\cdot\vec{x}'}\,dA' &= \int_{-L/2}^{L/2} e^{ik_y y'}\,dy' \int_{-L/2}^{L/2} e^{ik_x x'}\,dx' \\ &= \frac{L^2}{4}\left(\frac{\sin\left(\frac{k_x L}{2}\right)}{k_x L/2}\right)\left(\frac{\sin\left(\frac{k_y L}{2}\right)}{k_y L/2}\right)\end{aligned} \tag{3.81}$$

which are these nice sinc functions we've all seen before in diffraction experiments. We can also compute the intensity, which is found by:

$$\begin{aligned}I &= |\psi_d|^2 \\ &\sim \frac{\sin^2\left(\frac{kL\theta}{2}\right)}{\left(\frac{kL\theta}{2}\right)^2}\end{aligned} \tag{3.82}$$

where we have taken that $k_x = k\sin\theta \approx k\theta$. So we see that the intensity falls off like $1/\theta^2$ further away from the optical axis on the detector screen.

Note: All we have done so far is that there is no field emerging from the aperture barrier (not just the aperture/hole), but in real life this is **not** the case! In fact, derivation have mostly been done only for conducting surfaces. Most of what we will see is not having to worry about the barrier but only what comes through the hole.

§3.3.1 Circular Aperture Diffraction

Setting up the coordinate frame such that the x, y-plane lies on the surface of the barrier and the z-axis is in the direction of incident wave propagation, it can be worked out (as done in problem set 4) that the outgoing power distribution of a wave through a circular aperture of radius a is given by:

$$\frac{1}{P_i}\frac{dP}{d\Omega} = \frac{(ka)^2}{4\pi}\left[\frac{j_1(qa)}{qa}\right]^2 \times \begin{cases} 1 - \sin^2\theta\cos^2\phi, & \hat{\varepsilon}_0 = \hat{y} \\ 1 - \sin^2\theta\sin^2\phi, & \hat{\varepsilon}_0 = \hat{x}\end{cases} \tag{3.83}$$

where $j_1(qa)$ is the Bessel function of the first kind of index 1, $q = \left|\vec{k}_0 - \vec{k}\right|$ and the angular dependence arises from the polarization of the incident wave. Because of the Bessel function dependence, the largest angular resolution that can be resolved is given by:

$$\boxed{\Delta\theta \approx 1.22\frac{\lambda}{2d}} \tag{3.84}$$

This is common in astronomy (telescope lenses) and optics.

Example:

Consider a human eye that detects wavelengths on the order of $\lambda \sim 5.6 \times 10^{-3}$ cm. The diameter of the pupil is on the order of $d \sim 1.6$ to 6 mm, so we get the angular resolution of a human eye is bounded by:

$$1 \times 10^{-4} < \Delta\theta < 5 \times 10^{-4} \text{ rad} \tag{3.85}$$

§3.3.2 Babinet's Principle

We will be discussing this concept from the point of view of scalar diffraction theory. We start with the Kirchhoff's diffraction integral before any approximations:

$$\psi(x) = -\frac{1}{4\pi}\int dA'\frac{e^{ikR}}{R}\hat{n}'\cdot\left[\boldsymbol{\nabla}_{x'}\psi(x') + ik\left(1 - \frac{i}{kR}\right)\frac{\vec{R}}{R}\psi(x')\right] \tag{3.86}$$

Recall the way we set-up the surface relevant in a diffraction experiment, where we bound the volume enclosing the diffraction wave by the flat-plane barrier S_1 and a great hemisphere that bounds infinity S_2. Now Babinet further decomposed the problem in a smart way where we considered 2 regimes.

1. The entire S_1 region is the aperture except for a small opaque region. This causes some slight scattering off this small region that is negligable). This gives rise to scattered wave ψ_A.
2. The converse region from above, where the aperture is now only the small region and the rest of the S_1 is opaque. This gives rise to scattered wave ψ_B.

We then notice that if the whole region S_1 were to be empty, no diffraction would occur, so $\psi_s = 0$. So we get that by superposition:

$$\begin{aligned} &\psi_A + \psi_B = 0 \\ \Rightarrow\ &|\psi_A|^2 = |\psi_B|^2 \\ \Rightarrow\ &\boxed{I_a^{\text{diff}} = I_b^{\text{diff}}} \end{aligned} \tag{3.87}$$

This result is known as *Babinet's principle*. A result from this principle is a phenomena known as the *Arago/Fresnel spot*.

Chapter 4

Special Relativity

"*Einstein said it all.*"

– T. DeGrand, 2020

Here, we will be learning a lot of frame transformations which has roots in dynamics. Before relativity, classical mechanics and electricity and magnetism were separate. Einstein's contribution was to say that these had nothing distinct but related by spacetime. Linear operators and the appropriate symmetries cause E&M to fall out of the theory of special relativity naturally. Special relativity is a peculiar subject, and is often taught largely differently (depending on the audience). For instance even though special relativity is indeed the precursor of general relativity, but the emphasis of general relativity is largely different from literature focused on special relativity. In AMO, relativistic effects can often be treated perturbatively, but is still a essential piece for precise theories.

§4.1 Introduction

Consider 2 *inertial frames*. This means that we have 2 coordinate systems of interest and are moving relative to each other at constant velocity. The origin of observer in frame 2 as seen be the observer in frame 1 is then:

$$\vec{r} + \vec{v}(t - t_0) \tag{4.1}$$

where $\vec{r}$ and t are what the observers measure. The postulate of special relativity is stated as follows.

> **Postulate of Special Relativity**
>
> *The speed of light c is constant in* ***all*** *inertial reference frames.*

Now consider 2 events that are separated by $\Delta\vec{x}$ in space and Δt in time, that are seen by a relativistic observer in some inertial reference frame. A relativistic observer in another inertial frame would see these events occur at $\Delta\vec{x}'$ and $\Delta t'$ where:

$$\Delta\vec{x} \neq \Delta\vec{x}, \quad \Delta t \neq \Delta t' \tag{4.2}$$

Despite this, there is a quantity that will be invariant to both observers, which is the quantity:

$$\boxed{(\Delta s)^2 = c^2(\Delta t)^2 - (\Delta \vec{x})^2 = c^2(\Delta t')^2 - (\Delta \vec{x}')^2} \tag{4.3}$$

where Δs is known as the *spacetime interval.* Contrasting this to *Galilean invariance*, where if we have:

$$\begin{aligned} \vec{x}' &= \vec{x} + \vec{v}t \\ t' &= t \end{aligned} \tag{4.4}$$

then we have:

$$(\Delta \vec{x})^2 = (\Delta \vec{x}')^2, \quad (\Delta t)^2 = (\Delta t')^2 \tag{4.5}$$

So in Galilean physics, we have:

$$\vec{F} = m\frac{d^2x}{dt^2} = m\frac{d^2x'}{d(t')^2} \tag{4.6}$$

But this does **not** hold in a *Lorentz transformation* (relativistic frame transformation)! What is a Lorentz transformation? Well we can think of it in terms of a rotation. We know that rotations do not change the norm of a vector. Analogously, we want the spacetime interval to remain unchanged under this rotation as asserted by the postulate of special relativity. In fact, if we write $t = iw$ where i is the imaginary number, we exactly get that the rotation matrix allows *Lorentz invariance.* So first working in 1 spatial dimension, we have the transformations being written as:

$$\begin{aligned} x' &= x\cosh\theta - ct\sinh\theta \\ ct' &= -x\sinh\theta + ct\cosh\theta \end{aligned} \tag{4.7}$$

To find what θ is, we suppose that in the frame S', we have $x' = 0$. However, the origin of S' is moving away by $x = vt$ as seen from another observer in another frame S. So we get that:

$$\begin{aligned} & x' = 0 = x\cosh\theta - ct\sinh\theta \\ \Rightarrow \quad & \tanh\theta = \frac{x}{ct} = \frac{vt}{ct} = \frac{v}{c} \equiv \beta \\ \Rightarrow \quad & \boxed{\cosh\theta = \frac{1}{\sqrt{1-\beta^2}} \equiv \gamma, \quad \sinh\theta = \frac{\beta}{\sqrt{1-\beta^2}} \equiv \beta\gamma} \end{aligned} \tag{4.8}$$

So plugging this back into the frame-transformations, we get:

$$\begin{aligned} x' &= \gamma(x - vt) \\ t' &= \gamma\left(t - \frac{vx}{c^2}\right) \end{aligned} \tag{4.9}$$

Generally in relativity, we draw spacetime diagrams with the time-axis being vertical and the spatial axis horizontal. 45° lines that pass through the origin are known as *light cones*, and divide the spacetime diagram into regions that are accessible by physical particles (with velocities $v < c$, inside the light cone) and regions that are not. Trajectories of particles in spacetime diagrams are known as *world lines.*

§4.1.1 Time-Like and Space-Like Separation

Because spacetime intervals are invariant, we have lines:

$$x^2 - c^2t^2 = \text{constant} \tag{4.10}$$

which make up hyperbolas which define the positions of particles vary under Lorentz transformations. Stated again,

Under a Lorentz transformation, a spacetime point moves along the hyperbola $x^2 - c^2t^2 =$ constant.

For 2 events, we call the time interval between them occurring $\Delta t_0 = t_1 - t_2$ the *proper time*, which vanishes when these events happen on the surface of the light cone ($(\Delta s)^2 = 0$). Another way to define this, is the shortest time an observer in an arbitrary frame can measure between 2 events at the same spatial point.

Conversely, events that occur at the same time but at different spatial positions are known as *space-like separated*. In the spacetime diagram, these move along a hyperbola **outside** the light cone under a Lorentz transformation. This distinction leads us to worry about whether events occur inside or outside the light cone when we talk about dynamics.

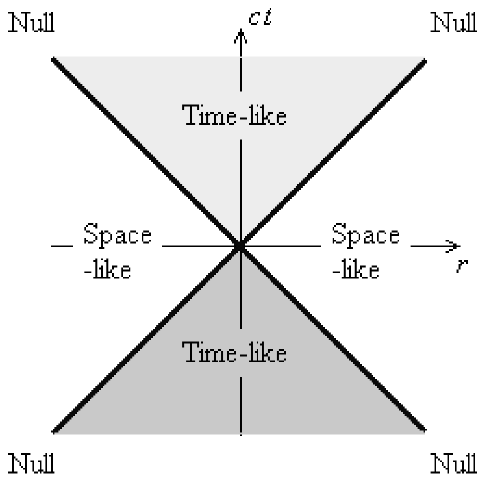

Figure 4.1: Space-like and Time-like Regions.

Example:

Consider a frame K which is the frame of some stationary observer that observes a particle moving with velocity $u(t)$. Also, there is another frame K' in which the particle is stationary and time ticks in this frame as τ (the *proper time*). We know that in both frames, we have the spacetime interval is constant:

$$\begin{aligned}(ds)^2 &= c^2 d\tau^2 \\ &= c^2(dt)^2 - |dx|^2 \\ &= c^2(dt)^2\left[1 - \frac{1}{c^2}\left(\frac{dx}{dt}\right)^2\right] \\ &= c^2(1-\beta^2)dt^2\end{aligned} \tag{4.11}$$

where $dx = udt$. So we can get:

$$\begin{aligned}d\tau &= \frac{dt}{\gamma} \\ \Rightarrow \quad \int_{\tau_1}^{\tau_2} \gamma(\tau)d\tau &= \int_{t_1}^{t_2} dt = \Delta t > \int_{\tau_1}^{\tau_2} d\tau\end{aligned} \tag{4.12}$$

§4.2 4-Vectors

In this class, we are going to define *4-vectors* as any quantity which transforms like x and ct under a Lorentz transformation. That is:

$$\begin{aligned}&A_\mu = [A_0, A_1, A_2, A_3]^T \\ \Rightarrow \quad &A'_0 = \gamma(A_0 - \vec{\beta}\cdot\vec{A}) \\ &A'_\parallel = \gamma(A_\parallel - \beta A_\perp) \\ &A'_\perp = A_\perp\end{aligned} \tag{4.13}$$

where overhead-arrows denote 3-vectors as per in Euclidean space. The scalar product of 2 4-vectors are then defined as:

$$\boxed{A_\mu B^\mu = A_0 B_0 - \vec{A}\cdot\vec{B}} \tag{4.14}$$

Now, consider a wave with frequency ω and wave-vector $\vec{k}$. We can then think about ω and $\vec{k}$ in 2 different frames K and K'. The way to think about this is considering a wave-train (pulse) with 3 wave peaks and 2 wave troughs. We note that in either frame, the number of these peaks and troughs does **not** change. Defining ϕ as the *number of wave crests*, we have that this quantity is invariant under Lorentz transformation. So we have:

$$e^{i\phi} e^{i(\omega t - \vec{k}\cdot\vec{x})} \tag{4.15}$$

So this tells us that $k_\mu x^\mu = \omega t - \vec{k}\cdot\vec{x}$ remains invariant, for which we can define:

$$k_\mu = [\omega/c, \vec{k}] \tag{4.16}$$

which is indeed a 4-vector since we know $[ct, \vec{x}]$ is one and will remain invariant under Lorentz transformation. So we get that k_μ transforms like:

$$\begin{aligned} k'_0 &= \gamma(k_0 - \vec{\beta} \cdot \vec{k}) \\ k'_\parallel &= \gamma(k_\parallel - \beta k_0) \\ k'_\perp &= k_\perp \end{aligned} \tag{4.17}$$

So we get that:

$$\boxed{\omega' = \frac{\omega(1-\beta)}{\sqrt{1-\beta^2}} = \omega\sqrt{\frac{1-\beta}{1+\beta}}} \tag{4.18}$$

which is the *relativistic Doppler shift* formula. We also get:

$$\begin{aligned} \frac{k'_\perp}{k'_\parallel} &= \tan\theta' \\ &= \frac{k\sin\theta}{\gamma\left(k\cos\theta - \beta k_0\right)} \\ &= \frac{\sin\theta}{\gamma(\cos\theta - \beta)} \end{aligned} \tag{4.19}$$

where θ is the angle of the $\vec{k}$ vector and the $\vec{x}$ axis in the spacetime diagram.

Side Note: There is another type of Doppler shift known as *gravitational red shift*. This occurs if we were to shoot a photon away from the surface of a massive body (usually taken as the Earth). This shift goes like:

$$\frac{\Delta\omega}{\omega} = -\frac{gh}{c^2} \sim 10^{-17} \tag{4.20}$$

where g is the acceleration due to gravity close to the surface of the massive body, h is the height and c is the speed of light. This alludes to the *principle of equivalence*.

§4.2.1 4-Velocity

How do we generalize the 3-velocity in Newtonian mechanics to relativistic 4-velocities? Well, we can consider a particle with velocity u' in the frame K', while K' moves with a velocity u with respect to another stationary frame K. We want to then compute the velocity of the particle as observed from K, which is done by boosting:

$$\begin{aligned} dx_0 &= \gamma_v(dx'_0 + \beta dx'_1) \\ dx_1 &= \gamma_v(dx'_1 + \beta dx'_0) \end{aligned} \tag{4.21}$$

We also have:

$$\begin{aligned} u'_j &= c\frac{dx'_j}{dx'_0}, \quad u_j = c\frac{dx_j}{dx_0} \\ \Rightarrow \quad u_1 &= c\frac{dx_1}{dx_0} = \frac{dx'_1 + \beta dx'_0}{dx'_0 + \beta dx'_1} \\ \Rightarrow \quad & \boxed{u_\parallel = \frac{u'_\parallel + v}{\sqrt{1 + \frac{\vec{v}\cdot\vec{u}'}{c^2}}}, \quad u_\perp = \frac{u'_\perp}{\gamma\left(1 + \frac{\vec{v}\cdot\vec{u}'}{c^2}\right)}} \end{aligned} \tag{4.22}$$

These are known as the *relativistic velocity addition* formulas. These makes things not so elegant, but we can get around this by defining the vector:

$$u_\mu \equiv \left[\frac{dt}{d\tau}, \gamma_u \vec{u}\right]^T \tag{4.23}$$

where we note that $dt/d\tau = c\gamma_u$. More concisely, we have:

$$\boxed{u_\mu = \frac{dx_\mu}{d\tau}} \tag{4.24}$$

In the rest frame, this 4-vector reduces to $[c, 0, 0, 0]^T$. The invariant quantity found by contracting this 4-vector with itself is:

$$u_\mu u^\mu = c^2 \tag{4.25}$$

With this definition, complicated velocity addition formula we had earlier is in fact encoded into this 4-velocity when it is Lorentz transformed (this working is left to the reader)!

§4.2.2 4-Acceleration

After the introduction of the 4-velocity, we can also define a *4-acceleration* as:

$$a_\mu = \frac{d}{d\tau} u_\mu \tag{4.26}$$

To get how this looks like explicitly, we compute:

$$\begin{aligned} \frac{d\gamma}{dt} &= \frac{\gamma^3}{c^2}\vec{u}\cdot\vec{a} \\ \Rightarrow \quad \frac{d}{dt}\left(\gamma\vec{u}\right) &= \gamma\vec{a} + \vec{u}\frac{\gamma^3}{c^2}\vec{u}\cdot\vec{a} \end{aligned} \tag{4.27}$$

Plugging this result into the 4-acceleration gives us:

$$\boxed{a_\mu = \left[\gamma^4 \frac{\vec{u}\cdot\vec{a}}{c}, \quad \gamma^4\left(\frac{\vec{u}\cdot\vec{a}}{c^2}\right)\vec{u} + \gamma^2\vec{a}\right]^T} \tag{4.28}$$

In the frame where $u = 0$ and $\gamma = 1$, we notice that we have:

$$\begin{aligned} &a_\mu = [0, \quad \vec{a}]^T, \quad u_\mu = [c, \quad \vec{0}]^T \\ \Rightarrow \quad &a_\mu u^\mu = 0 \end{aligned} \tag{4.29}$$

So this tells us that we can pick a frame that makes life easy and we will always have $a_\mu u^\mu = 0$ since this is a invariant quantity.

§4.2.3 4-Momentum

It seems that having the 4-velocity would make writing the *4-momentum* trivial, however although this works out to be true, the rationale is not so straight forward. In relativity, we construct 4-vectors such that they transform the same way under Lorentz transformations, but also so that they can be useful in doing physics. The way 4-momentum is constructed in relativity, is such that there is an intrinsic relation between momentum and energy.

To get some intuition of the momentum/energy relation, we look to quantum mechanics. Consider a spatial translation in quantum mechanics:

$$\psi(q + \delta q) = \psi(q) + \delta q \frac{\partial \psi(q)}{\partial q} \tag{4.30}$$

This leads to the derivation of the momentum operator with the commutation relation $[\hat{p}, \hat{q}] = i\hbar$, so we get that the momentum operator is the generator of the spatial translations. Similar, we get that the Hamiltonian is the generator of time-translations.

Now going back to relativity, we know that changes in frames interchange the translation of time and space, so from the notion of quantum generators, we see that $\hat{H}$ and $\hat{p}$ have to be related and therefore E and p. A good check of the derivation in relativity, is that when we move to the non-relativistic limit (e.g. $v \ll c$), we should retrieve classical mechanics. With these things in mind, we indeed find that we can define the 4-momentum to be $p_\mu = mu_\mu$. Explicitly, we have:

$$\boxed{p_\mu = \left[\frac{E(\vec{p})}{c}, \quad \vec{p}\right]^T = [\gamma mc, \quad \gamma m\vec{u}]^T} \tag{4.31}$$

The relation above tells gives us the famous Einstein relation:

$$\boxed{E = \gamma mc^2} \tag{4.32}$$

Or also written as $E^2 - (pc)^2 = \left(mc^2\right)^2$. This relation is very useful because c and m are spacetime invariant. This gives us 3 other useful relations:

$$\beta = \frac{cp}{E}, \quad \gamma = \frac{E}{mc^2} \tag{4.33}$$

Now checking the classical limit, we can Taylor expand the energy and 3-momentum terms:

$$\begin{aligned} E &= \gamma mc = mc^2 + \frac{1}{2}mu^2 + \dots \\ \vec{p} &= \gamma m\vec{u} = m\vec{u} + \dots \end{aligned} \tag{4.34}$$

which indeed gives us the classical kinetic energy and momentum. Now let's see what happens when we Lorentz transform this 4-momentum ($\Lambda^{\mu\nu} p_\nu$):

$$\begin{aligned} E' &= \gamma\left(E - \beta c p_\parallel\right) \\ p'_\parallel &= \gamma\left(p_\parallel - \frac{\beta}{c}E\right) \\ p'_\perp &= p_\perp \end{aligned} \tag{4.35}$$

Now if we consider a relativistic high energy process where we have incoming particles colliding and producing other particles, the sum of all the incoming and outgoing particles should have total energy and momentum add to 0 ($\sum_j p_j = \sum_j E_j = 0$). We see that even when we Lorentz transform these quantities, they still indeed do sum to 0, so energy and momentum are indeed conserved. When we deal with **massless** particles, we only write the 4-vector as:

$$\boxed{p_\mu = \begin{bmatrix} \frac{E}{c}, & \vec{p} \end{bmatrix}^T} \tag{4.36}$$

and not in the form where m is involved because it is **not** true that massless particles have no energy/momentum.

Another way to appeal to your senses about the construction of this 4-momentum, we look at a very crude derivation (that happens to be exact for non-interacting particles). Recall that:

$$k_\mu = \begin{bmatrix} \frac{\omega}{c}, & \vec{k} \end{bmatrix}^T \tag{4.37}$$

Nature is described by quantum fields that have fundamental modes as plane waves with frequency ω and wave-number k. Then quantum mechanics came along and asserted $E = \hbar\omega$ and $p = \hbar k$, so if we frivolously put things together, we get:

$$\hbar k_\mu = \begin{bmatrix} \frac{\hbar\omega}{c}, & \hbar\vec{k} \end{bmatrix}^T = \begin{bmatrix} \frac{E}{c}, & \vec{p} \end{bmatrix} \tag{4.38}$$

So we indeed retrieve the 4-momentum we had gotten to earlier.

§4.3 Relativistic Scattering

First, we consider elastic scattering. Consider an oncoming particle a that collides with another particle b both with the same mass m. We consider the rest-frame of particle b and pick the coordinate frame such that we have the 4-momenta of these particles being:

$$\begin{aligned} p_\mu^{(a)} &= [E, p, 0, 0]^T \\ p_\mu^{(b)} &= [m, 0, 0, 0]^T \end{aligned} \tag{4.39}$$

where we have now set c to 1. The cost-collision particles (now labeled by c and d) will have momentas

$$\begin{aligned} p_\mu^{(c)} &= [E_c, p_c\cos\theta_c, p_c\sin\theta_c, 0]^T \\ p_\mu^{(d)} &= [E_d, p_d\cos\theta_d, p_d\sin\theta_d, 0]^T \end{aligned} \tag{4.40}$$

where by conservation of momentum, we have $p_c \sin\theta_c = p_d \sin\theta_d$. This problem can actually be more conveniently formalized if we move to the center of mass frame, for which by the assertion of conservation of energy and momentum, we have:

$$\begin{aligned}
(p_\mu^{(a)})' &= [E', p', 0, 0]^T \\
(p_\mu^{(b)})' &= [E', -p', 0, 0]^T \\
(p_\mu^{(c)})' &= [E', p'\cos\theta, p'\sin\theta, 0]^T \\
(p_\mu^{(d)})' &= [E', -p'\cos\theta, -p'\sin\theta, 0]^T
\end{aligned} \tag{4.41}$$

where the primes denote the center of mass frame 4-momenta. Now, what we if wanted to move this back into the rest-frame of particle b, well we do a Lorentz transformation. However, we need to know γ and β to do that. Well, that's not too hard because we know things from Lorentz invariants which gives us:

$$\gamma = \frac{E'}{m}, \quad \beta = \frac{p}{E} = \sqrt{1 - \frac{m^2}{E^2}} \tag{4.42}$$

Also, we know that:

$$\begin{aligned}
& (p^{(a)} + p^{(b)})_\mu (p^{(a)} + p^{(b)})^\mu = (p^{(a)\prime} + p^{(b)\prime})_\mu (p^{(a)\prime} + p^{(b)\prime})^\mu \\
\Rightarrow \quad & 2m^2 + 2mE = 4(E')^2
\end{aligned} \tag{4.43}$$

where E is the b-frame energy and E' is half of the CoM-frame energy. Now what about the angles θ? Well, we can consider:

$$\begin{aligned}
& (p^{(a)} - p^{(c)})_\mu (p^{(a)} - p^{(c)})^\mu = (p^{(a)\prime} - p^{(c)\prime})_\mu (p^{(a)\prime} - p^{(c)\prime})^\mu \\
\Rightarrow \quad & (p^{(a)} - p^{(c)})_\mu (p^{(a)} - p^{(c)})^\mu = 2m^2 - 2\left[(E')^2 - (p')^2 \cos\theta\right]
\end{aligned} \tag{4.44}$$

§4.4 Relativistic Dynamics

To do dynamics in relativity, we want to work with simple building blocks so that it is easy to keep track of things when we move between frames. To start, we will be considering differential geometry that is relevant for special relativity. To describe events in terms of 4-vectors, we write:

$$x^\mu = [x_0, x_1, x_2, x_3] \tag{4.45}$$

where we imagine that there are some transformation rules such that:

$$\Lambda : x^\mu \mapsto x^{\mu\prime}(x^\mu) \tag{4.46}$$

indicating that the transformed 4-vector is a function of the initial 4-vector. We also want that observables such as positions, momenta, fields, etc. transform "simply". Math tells us that these observables have a mathematical description with things called *tensors*, which are objects of rank k. Some instances of tensors are

1. <u>rank-0</u>: scalars ($s = s'$);

2. rank-1: vectors "generalization of 4-vectors" ($A^{\mu\prime} = \Lambda^{\mu'}{}_{\nu} A^{\nu}$ or $A^{\mu\prime} = \frac{\partial x^{\mu\prime}}{x^{\nu}} A^{\nu}$);
3. rank-2: "vectors with more indices";

When we work with 4-vectors, some literature will write components of a vector in terms of unit vectors:

$$\begin{aligned} \tilde{V} &= V^{\mu} \hat{e}^{(\mu)} \\ &= \Lambda^{\mu}{}_{\nu} V^{\nu} \hat{e}_{(\mu)} \end{aligned} \tag{4.47}$$

where we also have that:

$$\hat{e}_{(\mu)} = \Lambda^{\nu}{}_{\mu} \hat{e}_{(\nu)} \tag{4.48}$$

Furthermore, we notice that if we act on 4-vectors with:

$$\begin{aligned} & V^{\mu} = \Lambda^{\mu'}{}_{\nu} V^{\nu'} = \Lambda^{\mu}{}_{\nu'} \Lambda^{\nu'}_{\rho} V^{\rho} \\ \Rightarrow \quad & \boxed{\Lambda^{\mu}{}_{\nu'} \Lambda^{\nu'}{}_{\rho} = \delta^{\mu}_{\rho}} \end{aligned} \tag{4.49}$$

So the reverse transformation of a Lorentz transformation is simply its inverse.

§4.4.1 Covariant 4-Vectors and Index Contraction

So far, we have been writing 4-vectors with their indices on the top, these are known as *contravariant 4-vectors*. However we also have 4-vectors with indices below, and these are known as *covariant/dual/one-form 4-vectors*. These are analogous to bras for kets in quantum mechanics. These objects transform like:

$$\boxed{B_{\mu'} = \frac{\partial x^{\nu}}{\partial x^{\mu'}} B_{\nu} = \Lambda^{\nu}{}_{\mu'} B_{\nu}} \tag{4.50}$$

An example of this object, would be a 4-gradient operators:

$$\begin{aligned} \partial_{\alpha'} &= \frac{\partial}{\partial x^{\alpha'}} \\ &= \frac{\partial x^{\beta}}{\partial x^{\alpha'}} \frac{\partial}{\partial x^{\beta}} = \begin{bmatrix} \frac{\partial}{\partial x^0} \\ \boldsymbol{\nabla} \end{bmatrix} \end{aligned} \tag{4.51}$$

The contravariant (upper index) version of this is written as:

$$\partial^{\alpha} = \frac{\partial}{\partial x_{\alpha}} = \begin{bmatrix} \frac{\partial}{\partial x^0} \\ -\boldsymbol{\nabla} \end{bmatrix}. \tag{4.52}$$

Contraction of the covariant and contravariant gives an operator known as the *d'Alembert operator*:

$$\Box = \partial_{\alpha} \partial^{\alpha} = \frac{\partial^2}{\partial (x^0)^2} - \boldsymbol{\nabla}^2 \tag{4.53}$$

which is the generalization of the Laplacian in 4-space. These lower-index objects are necessary to take inner-products which we refer to as *contracting the indices.* To see how this is done we have:

$$\begin{aligned}\tilde{V}\cdot\tilde{A} &= V_\mu A^\mu \\ &= \left(\frac{\partial x^\nu}{\partial x^\alpha}V_\nu\right)\left(\frac{\partial x^\alpha}{\partial x^\mu}A^\mu\right) \\ &= V_\nu\left(\frac{\partial x^\nu}{\partial x^\alpha}\frac{\partial x^\alpha}{\partial x^\mu}\right)A^\mu\end{aligned} \tag{4.54}$$

So we see that there is this 2-index object:

$$T^\nu{}_\mu = \frac{\partial x^\nu}{\partial x^\alpha}\frac{\partial x^\alpha}{\partial x^\mu} \tag{4.55}$$

which relates this inner product. In special relativity, if we again consider the invariant spacetime interval:

$$(ds)^2 = (dx^0)^2 - dx^j dx^j \tag{4.56}$$

where $j \in \{1,2,3\}$. We can also write this as:

$$\boxed{(ds)^2 = g_{\mu\nu}dx^\mu dx^\nu} \tag{4.57}$$

where $g_{\mu\nu}$ is known as the *metric tensor*. The metric tensor is a $(0,2)$-Cartesian tensor for which in special relativity (*Minkowski space*), is diagonal and has components:

$$g_{00} = 1, \quad g_{jj} = -1 \tag{4.58}$$

These objects allow us to take contravariant objects to their covariant counter parts (lowers upper-indices). The metric tensor has the property that:

$$\boxed{g_{\alpha\nu}g^{\nu\beta} = \delta^\beta{}_\alpha} \tag{4.59}$$

The matrix representation of the Minkowski metric tensor is written as:

$$g_{\mu\nu} = \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & -1 & 0\\ 0 & 0 & 0 & -1\end{bmatrix} \tag{4.60}$$

for which contraction in the matrix language will be written as:

$$\begin{aligned}A_\mu B^\mu &= A^\mu B^\nu g_{\mu\nu} \\ &= \mathbf{A}^T\mathbf{g}\mathbf{B}\end{aligned} \tag{4.61}$$

§4.4.2 Cartesian Tensors and Lorentz Transformations

As alluded in the previous section, we can generalize 4-vectors into objects with more indices:

$$T^{\mu_1 \ldots \mu_k}{}_{\nu_1 \ldots \nu_l} \tag{4.62}$$

which are known as rank-(k, l) *Cartesian tensors*. These transform as you would expect, that is Lorentz transformations are necessary for every index for instance:

$$T^{\mu_1 \ldots \mu_k} = \Lambda^{\mu_1}{}_{\nu_1'} \ldots \Lambda^{\mu_k}{}_{\nu_k'} T^{\nu_1' \ldots \nu_k'} \tag{4.63}$$

We can then ask, what is the most general transformation we can write such that we preserve the norm of the vector with respect to the metric tensor? One way of going about this is by constructing the group of all such *isometries* of the 4-vectors in Minkowski space. To do so, we consider some transformation:

$$\begin{aligned} & x^{\mu'} = \Lambda^{\mu'}{}_{\nu} x^{\nu} \\ \Rightarrow \quad & \mathbf{x}' = \mathbf{\Lambda x} \end{aligned} \tag{4.64}$$

and assert that:

$$\begin{aligned} & (\mathbf{x}', \mathbf{g x}') = (\mathbf{x}, \mathbf{g x}) \\ \Rightarrow \quad & \mathbf{x}^T \mathbf{\Lambda}^T \mathbf{g \Lambda x} = \mathbf{x}^T \mathbf{g x} \\ \Rightarrow \quad & \mathbf{\Lambda}^T \mathbf{g \Lambda} = \mathbf{g} \\ \Rightarrow \quad & g_{\mu\nu} \Lambda^{\mu}{}_{\rho} \Lambda^{\nu}{}_{\sigma} g_{\rho\sigma} \end{aligned} \tag{4.65}$$

where the parentheses above denote an inner product in 4-space. Now, we consider the determinant of these expressions:

$$\det\left(\mathbf{\Lambda}^T \mathbf{g \Lambda}\right) = \det(\mathbf{g}) \tag{4.66}$$

Since the determinant of a product is the product of the determinants, which implies that the $\det \Lambda = \pm 1$. It turns out that there are 2 types of possible Λ transformations:

1. Proper Transformations: **Can** be generated by continuous deformation of the identity element (e.g. rotations).
2. Improper Transformations: **Cannot** be generated by continuous deformation of the identity element (e.g. reflections).

The former constitutes a *Lie group*. We are now going to consider Λ being a proper transformation. We can thus consider the Lorentz transformation as a deviation ε away from the identity operator:

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} + \varepsilon^{\mu}{}_{\nu} \tag{4.67}$$

where $\delta^{\mu}{}_{\nu}$ is the identity. Then consider a multiplication of these matrices on the metric tensor up to first-order in ε:

$$\begin{aligned} & g_{\mu\nu} \left[\delta^{\mu}{}_{\rho} + \varepsilon^{\mu}{}_{\rho}\right] \left[\delta^{\nu}{}_{\sigma} + \varepsilon^{\nu}{}_{\sigma}\right] \approx g_{\rho\sigma} + \varepsilon_{\sigma\rho} + \varepsilon_{\rho\sigma} = g_{\rho\sigma} \\ \Rightarrow \quad & \boxed{\varepsilon_{\sigma\rho} = -\varepsilon_{\rho\sigma}} \end{aligned} \tag{4.68}$$

So we get that the additional variation on the identity **must** be antisymmetric. Plugging this result and applying the transformation on some 4-vector gives:

$$\begin{aligned} x^{\mu'} &= [\delta^{\mu}{}_{\nu} + \varepsilon^{\mu}{}_{\nu}]\, x^{\mu} \\ &= x^{\mu} + \varepsilon^{\mu}{}_{\nu} x^{\nu} \end{aligned} \tag{4.69}$$

We can then define:

$$\boxed{L_{\mu\nu} = i\,[x_{\mu}\partial_{\nu} - x_{\nu}\partial_{\mu}]} \tag{4.70}$$

$$\Rightarrow \quad \delta x^{\mu} \equiv \varepsilon^{\mu}{}_{\nu} x^{\nu} = \frac{i}{2}\varepsilon^{\rho\sigma} L_{\rho\sigma} x^{\mu} \tag{4.71}$$

where this antisymmetric $L_{\mu\nu}$ operator looks like a generalized angular momentum operator that generates rotations in 4-space (just as in quantum mechanics)! Note here that since imaginary numbers are introduced above, we are encroaching the realm of quantum mechanics where complex values are intrinsic to the theory. With this definition, we can further simplify:

$$\delta x^{\mu} = -\frac{1}{2}\,[\varepsilon^{\rho\mu} x_{\rho} - \varepsilon^{\mu\sigma} x_{\sigma}] = \varepsilon^{\mu\nu} x_{\nu} \tag{4.72}$$

Now we can ask, what is the commutation relations of these new generators we have constructed? Well, let's check:

$$[L_{\mu\nu}, L_{\rho\sigma}] = ig_{\nu\rho}L_{\nu\sigma} - ig_{\mu\rho}L_{\nu\sigma} - ig_{\nu\sigma}L_{\mu\rho} + ig_{\mu\sigma}L_{\nu\rho} \tag{4.73}$$

So we have that these objects L are generators of a Lie algebra:

$$\begin{aligned} &[L_i, L_j] = i\epsilon_{ijk} L_k \\ \text{where} \quad &L_i = \frac{1}{2}\epsilon_{ijk} L_{jk} \end{aligned} \tag{4.74}$$

with $i, j, k \in \{1, 2, 3\}$. However, we can consider objects that have intrinsic properties (e.g. spin in quantum mechanics) which follow the same commutation relations as in equation (4.73). Calling these other operators $S_{\mu,\nu}$, this will allow us to write a more generic generator:

$$\begin{aligned} &M_{\mu\nu} = L_{\mu\nu} + S_{\mu\nu} \\ \text{where} \quad &[L_{\mu\nu}, S_{\mu\nu}] = 0, \end{aligned} \tag{4.75}$$

which grants the most general form of an infinitesimal rotation as:

$$\mathscr{D}(\varepsilon) = 1 + \frac{i}{2}\varepsilon^{\mu\nu} M_{\mu\nu}. \tag{4.76}$$

We then further define:

$$\boxed{J_i \equiv \frac{1}{2}\epsilon_{ijk} M_{jk}, \quad K_i \equiv M_{0i}} \tag{4.77}$$

which have the commutation relations:

$$[J_i, J_j] = i\epsilon_{ijk} J_k \tag{4.78}$$
$$[J_i, K_j] = i\epsilon_{ijk} K_k \tag{4.79}$$
$$[K_i, K_j] = -i\epsilon_{ijk} J_k. \tag{4.80}$$

It turns out that these J **operators are the generators of rotations**, and K **the generators of Lorentz boosts**! With this in mind, it is also good to note the following commutation relations and their implications:

1. $[J_i, J_j] = i\epsilon_{ijk} J_k$: rotations do **not** commute;
2. $[J_i, K_j] = i\epsilon_{ijk} K_k$: rotations and boosts do **not** commute;
3. $[K_i, K_j] = -i\epsilon_{ijk} J_k$: Non-colinear boosts do **not** commute and in fact result in a rotation.

In 4-space, these generators have a well-defined matrix representation given by:

$$J_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad J_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \quad J_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{4.81}$$

$$K_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad K_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad K_3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \tag{4.82}$$

Notice that the square of these matrices are diagonal, so all squared generators commute. Additionally, we can take linear combinations of these generators to get:

$$\boxed{A_i \equiv \frac{1}{2}\left[J_i + iK_i\right], \quad B_i \equiv \frac{1}{2}\left[J_i - iK_i\right]}. \tag{4.83}$$

It then works out that these have:

$$[A_i, B_j] = 0 \tag{4.84}$$
$$[A_i, A_j] = i\epsilon_{ijk} A_k \tag{4.85}$$
$$[B_i, B_j] = i\epsilon_{ijk} B_k \tag{4.86}$$

which tells us that these A and B objects have the Lie algebra of $su(2)$ (spin operators). Hearkening back to non-relativistic quantum mechanics, we recall that rotations are generated by angular momentum operators for which the angular momentum eigenstates are labeled by the quantum number j (denoting total angular momentum) and $m \in [-j, j]$ (denoting the z projection of angular momentum leading to $2j+1$, m-states). As such, we get that each of these A and B operators have a total angular momentum-like quantity associated to them $\{a, b\}$, for which we can write:

$$A^2 \left|\psi_{AB}\right\rangle = a(a+1) \left|\psi_{AB}\right\rangle \tag{4.87}$$
$$B^2 \left|\psi_{AB}\right\rangle = b(b+1) \left|\psi_{AB}\right\rangle \tag{4.88}$$

where $|\psi_{AB}\rangle$ is indeed a quantum state. This is in fact where intrinsic spin of particles pop out from in relativistic quantum field theories! Let's now look at some properties.

It turns out that the eigenstates of the irreducible representations of the generators of the Lorentz group are **not** necessarily parity eigenstates under exchange of A and B (only parity eigenstates if $a = b$). Specifically, under the parity operator P^{AB}_{ij} we have:

$$P^{AB}_{ij} J_j = J_i \tag{4.89}$$
$$P^{AB}_{ij} K_j = -K_i \tag{4.90}$$

implying that while K_i transforms as an *axial vector* (*pseudovector*), J_i transforms as a vector. In general, we have that the eigenstate of the matrix representation of the operator $J = A + B$ (analogous to the addition of angular momentum) to be a $[(2a+1)+(2b+1)]$-dimesional vector. To see this, we consider some examples.

Examples:

The left-handed (chirality) neutrino is a fermionic particle (spin-1/2) that interacts with only the weak force and gravity. It turns out that these particles have the property of either $(a = 1/2, b = 0)$ or $(a = 0, b = 1/2)$. In the 2-dimensional spinor representation of spin-1/2 particles, these would clearly not be parity symmetric.

However, we see that if we have the Direct product $(a, b) = (0, 1/2) + (2/1, 0)$ as proposed by Dirac to construct *Dirac fermions* which have a 4-component spinor representation, these are indeed parity symmetric.

Having the structure of the generators, we can now ask how we can actually generate elements of the Lorentz group. That is asking, how do states actually transform under general Lorentz transformations? Well, as per in quantum mechanics, we can apply the exponential mapping to these generators to get:

$$\begin{aligned}\mathscr{D}(\varepsilon) &= 1 + \frac{i}{2}\varepsilon^{\mu\nu}M_{\mu\nu} \\ &= 1 + i\vec{\theta}_A\cdot\vec{A} + i\vec{\theta}_B\cdot\vec{B} \\ \Rightarrow \quad \mathscr{D}(\vec{\theta}_A,\vec{\theta}_B) &= \exp\left(i\vec{\theta}_A\cdot\vec{A}\right)\exp\left(i\vec{\theta}_B\cdot\vec{B}\right)\end{aligned} \tag{4.91}$$

where the factorization of the exponential maps came from the fact that $[A_i, B_i] = 0$. The vectors A and B in the expression above denote vector-operators, whereas $\vec{\theta}_A$ and $\vec{\theta}_B$ are the associated unit-vectors just like in quantum angular momentum. Alternatively, we can decompose this into the J and K operators as follows:

$$\boxed{\mathscr{D}(\vec{\omega},\vec{\zeta}) = 1 + i\vec{\omega}\cdot\vec{J} - \vec{\zeta}\cdot\vec{K}} \tag{4.92}$$

where $\vec{\omega} = (\vec{\theta}_A + \vec{\theta}_B)/2$ and $\vec{\zeta} = (\vec{\theta}_A - \vec{\theta}_B)/2$. Most of the concerns in SR are for Lorentz boosts, so let us consider the case where $\vec{\omega} = \vec{0}$ (recalling J's are associated to rotations and K's to boosts). Furthermore, let's say that the boost is purely along the x direction. This gives us that:

$$\mathscr{D}(\zeta) = 1 + \zeta\sigma_x + \dots \tag{4.93}$$

where we can represent K_x as the Pauli-x matrix since it follow the same Lie algebra. This gives the matrix representation of this boost as:

$$\begin{aligned} \mathscr{D}(\zeta) &= \exp\left(\zeta\sigma_x\right) \\ &= \begin{bmatrix} \cosh(\zeta) & \sinh(\zeta) \\ \sinh(\zeta) & \cosh(\zeta) \end{bmatrix} \\ \Rightarrow \quad V' &= \begin{bmatrix} \cosh(\zeta) & \sinh(\zeta) & 0 & 0 \\ \sinh(\zeta) & \cosh(\zeta) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} V \end{aligned} \tag{4.94}$$

The hyperbolic functions are thus indicative of the hyperbolic geometry of spacetime which we implicitly assert when we impose the Minkowski metric tensor $g_{\mu\nu}$.

§4.4.3 Thomas Precession

The property stated in the previous section that non-colinear boosts do not have a commutative structure gives rise to an interesting phenomenon in relativistic electronic orbits known as *Thomas precession*. That is, the commutation relations imply 2 successive Lorentz boosts are equivalent to one Lorentz boost and a proper rotation:

$$(\text{Lorentz Boost})_1 + (\text{Lorentz Boost})_2 = (\text{Lorentz Boost})_3 + (\text{3D Rotation}) \tag{4.95}$$

Thomas precession is a relativistic kinematic correction to the measured g factor of an electron, for which it accounts for the electron following a curvilinear orbit. This was explained by Llewellyn Thomas in 1927 after knowledge of the *Zeeman effect*. To understand this, we first recall that in classical mechanics, a transformation into a rotating frame results in the emergence of *fictitious forces* (centrifugal force, Coriolis force, Euler force). Furthermore, the time derivative of any vector function $\vec{f}(t)$ in the lab-frame is related to its rotating frame correspondent by:

$$\left.\frac{d\vec{f}(t)}{dt}\right|_{\text{lab}} = \left.\frac{d\vec{f}(t)}{dt}\right|_{\text{rot}} + \vec{\omega} \times \vec{f}(t) \tag{4.96}$$

where the subscript "lab" indicates the quantity as observed in the lab-frame, and "rot" as observed in the rotating frame. $\omega = \dot{\theta}(t)$. Knowing this, we now consider an electron orbiting a charged nucleus. We know from quantum theory that the magnetic moment $\vec{\mu}$ of electron is generated from its intrinsic spin-1/2 character as:

$$\vec{\mu} = \frac{ge}{2m_e c}\vec{S} \tag{4.97}$$

wehere $g \approx 2$ (as derived from the Zeeman effect experiments). As the electron traverses along its orbit, we have that its velocity would vary due to curvilinear trajectories. However, we can consider an infinitesimal interval of time δt such that within any δt instant, the electron travels will a fixed velocity $\vec{v}$. We will show in section **??** that a relativistic particle moving in external electric and magnetic fields will experience a magnetic field in its rest-frame given by:

$$\vec{B}\,' \approx \vec{B} - \vec{\beta} \times \vec{E}, \tag{4.98}$$

where $\vec{\beta} = \vec{v}/c$. So for an electron orbiting a nucleus which emits an electric field $\vec{E}$ (due to its charge distribution) and a magnetic field $\vec{B}$ (due to its spin), we have that its spin equation of motion in its rest-frame will be governed by:

$$\begin{aligned} \left.\frac{d\vec{S}}{dt}\right|_{\text{rest}} &= \vec{\mu} \times \vec{B}\,' \\ &= \vec{\mu} \times \left[\vec{B} - \vec{\beta} \times \vec{E}\right] \end{aligned} \tag{4.99}$$

Now, if we move out of the momentarily-comoving frame (rotating-frame) of the electron into the lab-frame, the associated change in the electron's velocity within each δt interval would be $\beta \to \beta + \delta\beta$. The electron has a spin equation of motion would have to be corrected as per equation (4.96) to:

$$\left.\frac{d\vec{S}}{dt}\right|_{\text{lab}} = \left.\frac{d\vec{S}}{dt}\right|_{\text{rest}} + \vec{\omega} \times \vec{S} \tag{4.100}$$

$$= \vec{S} \times \left[\frac{ge}{2m_e c}\left(\vec{B} - \vec{\beta} \times \vec{E}\right) - \vec{\omega}_T\right] \tag{4.101}$$

where $\vec{\omega}_T$ is known as the *Thomas frequency* derived of course, by Thomas. The derivation involves the commutative structure of J_i and K_i considering these infinitesimal velocity changes and boosts to the lab-frame (Jackson section 11.8 pages 550 - 552). Fast forward to the result, the Thomas frequency works out to be:

$$\vec{\omega}_T = -\lim_{\delta t \to 0} \frac{\Delta\Omega}{\delta t} = \frac{\gamma^2}{1+\gamma}\frac{\vec{a} \times \vec{v}}{c^2} \tag{4.102}$$

where $\vec{a}$ is the acceleration of the electron in the lab-frame. For the electron in particular (in a "screened" Coulomb field), we have:

$$\vec{\omega}_T = -\frac{\vec{L}}{2m_e^2 c^2}\frac{1}{r}\frac{dV}{dr} \tag{4.103}$$

where $\vec{L}$ is the angular momentum of the electron. As such, we have that the Thomas corrected interaction energy is given as:

$$\begin{aligned} U' &= -\vec{S} \cdot \left[\frac{ge\vec{B}'}{2m_e c} - \vec{\omega}_T\right] \\ \Rightarrow \quad & \boxed{U' = -\frac{ge}{2mc}\vec{S} \cdot \vec{B} + \frac{g-1}{2m_e^2 c^2}\left(\vec{S} \cdot \vec{L}\right)\frac{1}{r}\frac{\partial V}{\partial r}} \end{aligned} \tag{4.104}$$

which is the sum of the interaction energy due to the rest coupling of the electron spin and nuclear magnetic fields, and the interaction energy due to relativistic kinematics. We see that in the kinematic interaction energy term, the g-factor is suppressed $g \to g-1$ which was indeed measured by Uhlenbeck and Goudsmit in 1926. This result can also be pulled out of the Dirac equation when expanded in power of v/c. This result ignores the details about the value of g in higher orders of α (as done in relativistic quantum field theories):

$$\frac{g}{2} = 1 + c_1 \left(\frac{\alpha}{\pi}\right) + c_2 \left(\frac{\alpha}{\pi}\right)^2 + \dots \tag{4.105}$$

where α is the *fine-structure constant*. Much work (both experimental and theoretical) is still ongoing for this, for which experiments (like the eEDM experiment performed by the *Cornell group*) are trying to measure the electron g-factor so as bound the allowed quantum field theories.

"*If you can't be the giant you want to be the giant killer.*"

– T. DeGrand, 2020.

Chapter 5

The Covariance of Electrodynamics

We are now going to study about how electrodynamics looks different in different inertial frames. The hope is that physics remains the same in all inertial reference frames. We will then be looking at electrodynamics from a field theory perspective which will elucidate important notions such as gauge invariance a little more clearly.

To start, we recall the continuity equation:

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \vec{J} = 0 \tag{5.1}$$

which we know is true in classical electrodynamics. We can extrapolate this into 4-space and postulate that these quantities ρ and $\vec{J}$ form a 4-vector:

$$\begin{aligned} & J^\mu = \left[c\rho, \;\; \vec{J}\right] \\ \Rightarrow \quad & \partial_\mu J^\mu = 0 \end{aligned} \tag{5.2}$$

Let's push on this a little. We recall that charge is defined as:

$$Q = \int d^3x \rho(\vec{x}) \tag{5.3}$$

This should intuitively be an invariant quantity, however the volume contracts when an observer is boosted to another inertial frame. To possibly fix this, we can consider the integral over spacetime volume:

$$d^4x = dx^0 d^3x \tag{5.4}$$

for which this quantity is indeed Lorentz invariant. This immediately then tells us that if Q is invariant, ρ must transform like x^0 which indeed tells us our postulate checks out. Now we also recall the Lorentz gauge condition:

$$\begin{aligned} &\frac{1}{c}\frac{\partial \Phi}{\partial t} + \boldsymbol{\nabla}\cdot\vec{A} = 0 \\ \Rightarrow \quad &\left[\frac{1}{c}\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2\right]\begin{bmatrix}\Phi\\ \vec{A}\end{bmatrix} = \frac{4\pi}{c}\begin{bmatrix}c\rho\\ \vec{J}\end{bmatrix} \end{aligned} \tag{5.5}$$

which this, we can also then postulate that:

$$A^\mu = \left[\Phi,\ \vec{A}\right] \tag{5.6}$$

is another 4-vector. However, Lorentz gauge is completely arbitrary and what we then choose another gauge? Well, let's take a step back. We are going to start with a classical Lagrangian and encode/impose the appropriate symmetries on it. From this, equations of motion will then naturally pop out of this Lagrangian which intrinsically encode these symmetries. We can then ask questions like if we retrieve Maxwell's equations from this, are these unique? Can we quantize this theory to get a quantum theory (get a Hamiltonian)?

§5.1 Classical Field Theory and Noether's Theorem

Having these newly defined fields in mind, we are going to consider electric and magnetic fields as classical fields instead of things emergent from point particles. This also gives the foundational formalism to bring us to quantum field theories in the future (not in this course). To get started, we make some remarks on classical mechanics. In classical mechanics, we have a set of coordinates $\{q_j, \dot{q}_j\}$, a Lagrangian $\mathcal{L}(q_j, \dot{q}_j)$ and a set of equations of motion:

$$\sum_j \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \frac{\partial \mathcal{L}}{\partial q_j} = 0 \tag{5.7}$$

from the *least-action principle.* Classical field theories are very similar, but instead of the conjugate variables above, the coordinates we will be considering are the fields themselves ($q_j \to \phi_j(x)$ and $p_j \to \partial_\mu \phi_j$). Furthermore, we deal with a quantity known as the *Lagrangian density* $\mathscr{L}$ instead of the Lagrangian, which is defined as:

$$\mathcal{L} = \int d^3x \mathscr{L} \tag{5.8}$$

for which the action is then defined for the Lgrangian density as:

$$S = \int d^4x \mathscr{L}(\phi_j, \partial_\mu \phi_j). \tag{5.9}$$

Minimizing the action via the calculus of variations give the result:

$$\boxed{\partial_\mu\left[\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi_j)}\right] - \frac{\partial \mathscr{L}}{\partial \phi_j} = 0} \tag{5.10}$$

which is analogous to the Euler-Lagrange equations of motion in traditional analytical mechanics. To analyze this, we consider a small variation of the field variable:

$$\phi_j \quad \rightarrow \quad \phi_j + \delta\phi_j \tag{5.11}$$
$$\Rightarrow \quad \partial_\mu\phi_j \quad \rightarrow \quad \partial_\mu\phi_j + \delta\left(\partial_\mu\phi_j\right) \tag{5.12}$$

Plugging this into the Lagrangian density, we get that the variation in the Lagrangian density is given by:

$$\begin{aligned}
\delta\mathscr{L} &= \sum_j \left[\frac{\partial\mathscr{L}}{\partial\phi_j}\delta\phi_j + \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_j)}\delta\left(\partial_\mu\phi_j\right)\right] \\
&= \sum_j \left[\partial_\mu\left(\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_j)}\right)\delta\phi_j + \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_j)}\delta\left(\partial_\mu\phi_j\right)\right] \\
&= \partial_\mu\left[\sum_j \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_j)}\delta\phi_j\right].
\end{aligned} \tag{5.13}$$

We now define the term in the square-brackets above as a form of generalized contravariant current:

$$\boxed{J^\mu \equiv \sum_j \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_j)}\delta\phi_j}, \tag{5.14}$$

from which we see that if $\delta\mathscr{L} = 0$ under some continuous transformation of ϕ_j, this implies $\partial_\mu J^\mu = 0$, telling us that J^μ is conserved! So we have the statement of *Noether's theorem*:

> "*Given any continuous symmetry (transformation of the field variables, ϕ_j) which leaves the Lagrangian density invariant, there will be an associated* ***Noether current****, J^μ which remains conserved under this symmetry transformation*".

This is the statement that symmetries lead to conserved quantities, which is one of the most profound results in physics! The Noether current will also have an associated *Noether charge* defined as:

$$Q = \int d^3x J^0. \tag{5.15}$$

These Noether charges in ϕ can either be *internal Neother charges* or *external Neother charges*. The former (a.k.a *internal symmetries*) pertain to a transformation acting only on the field variables, therefore do not transform spacetime points, and leave the Lagrangian density or physics invariant (e.g. $\phi(t,\vec{x}) \rightarrow e^{i\theta(t,\vec{x})}\phi(t,\vec{x})$). The latter however (a.k.a *external symmetries*), pertains to transformations of the spacetime coordinates, leading to results like the conservation of momentum and energy. Let us first consider internal symmetries since these will prove rather interesting. Let's an example.

Example:

Consider the Lagrangian density:

$$\mathscr{L} = \frac{1}{2}\left[(\partial_\mu\phi_1)^2 + (\partial_\mu\phi_2)^2\right] - V(\phi_1, \phi_2) \tag{5.16}$$

We postulate that there is a symmetry in the system under linear combination of the 2 field variables $\{\phi_1, \phi_2\}$:

$$\begin{bmatrix}\phi_1'\\ \phi_2'\end{bmatrix} = \begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}\phi_1\\ \phi_2\end{bmatrix} \approx \begin{bmatrix}1 & \theta\\ -\theta & 1\end{bmatrix}\begin{bmatrix}\phi_1\\ \phi_2\end{bmatrix}. \tag{5.17}$$

This would result in the variation of the Lagrangian density as:

$$\delta\mathscr{L} = \sum_{j=1}^{2}\left[\frac{\partial\mathscr{L}}{\partial\phi_j}\delta\phi_j + \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi_j)}\delta\left(\partial_\mu\phi_j\right)\right] \tag{5.18}$$

for which we notice that the terms $\frac{\partial\mathscr{L}}{\partial\phi_j}\delta\phi_j$ would vanish if the potential is some function of $\phi_1^2 + \phi_2^2 = \left\|\vec{\phi}\right\|^2$. Asserting this, we then have that:

$$\delta\mathscr{L} = (\partial_\mu\phi_1)(\theta\partial_\mu\phi_2) + (\partial_\mu\phi_2)(-\theta\partial_\mu\phi_1) = 0 \tag{5.19}$$

So we get the transformation we write down above is indeed a symmetry! For which we get that the associated Noether current as:

$$J^\mu = (\partial_\mu\phi_1)\,\theta\phi_2 - (\partial_\mu\phi_2)\,\theta\phi_1 \tag{5.20}$$

Alternatively, we had consider the field variable as a single complex valued field $\phi = \phi_1 + i\phi_2$, we would retrieve the equivalent derivation above with the Lagrangian:

$$\mathscr{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi^* - V(\phi, \phi^*) \tag{5.21}$$

where the transformation is equivalent to tacking on a coordinate dependent phase to the field variable:

$$\delta\phi = i\theta\phi. \tag{5.22}$$

However, what happens if the phase is now coordinate dependent? That is $\theta = q\epsilon(\vec{x}, t)$. This implies that:

$$\begin{aligned} &\delta\phi(\vec{x}, t) = iq\epsilon(\vec{x}, t)\phi(\vec{x}, t)\\ \Rightarrow\quad &\phi'(\vec{x}, t) = \exp\left[iq\epsilon(\vec{x}, t))\right]\phi(\vec{x}, t) \end{aligned} \tag{5.23}$$

To get the variation in $\mathscr{L}$, we consider then:

$$\begin{aligned} &\delta\left(\partial_\mu\phi\right) = iq\epsilon(\vec{x}, t)\partial_\mu\phi + iq\phi\partial_\mu\epsilon(\vec{x}, t)\\ \Rightarrow\quad &\delta\mathscr{L} = \epsilon(\vec{x}, t)\partial_\mu J^\mu + \left[\frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi)}iq\phi + \frac{\partial\mathscr{L}}{\partial(\partial_\mu\phi^*)}(-iq\phi^*)\right]\partial_\mu\epsilon(\vec{x}, t) \end{aligned} \tag{5.24}$$

We already know that $\partial_\mu J^\mu = 0$, and we also notice that term in the square bracket is in fact the Noether current which gives:

$$\delta\mathscr{L} = \epsilon(\vec{x},t)\partial_\mu J^\mu(\vec{x},t) + J^\mu(\vec{x},t)\partial_\mu\epsilon(\vec{x},t) \tag{5.25}$$

which gives that the variation on the Lagrangian is in fact **not** zero! What then do we do? Well, let's perhaps construct another Lagrangian which has another degree of freedom such that we get $\delta\mathscr{L}$ under this transformation. Consider:

$$\delta\phi(\vec{x},t) = iq\epsilon(\vec{x},t)\phi(\vec{x},t) \tag{5.26}$$

$$\begin{aligned} &\delta A_\mu(\vec{x},t) = \partial_\mu\phi(\vec{x},t) \\ \Rightarrow\quad &\vec{A}\,' = \vec{A} + \boldsymbol{\nabla}\epsilon(\vec{x},t), \quad A^{0'} = A^0 + \frac{\partial\epsilon}{\partial t} \end{aligned} \tag{5.27}$$

So the variation of the Lagrangian density now is the same as before but with an additional term:

$$\begin{aligned} &\frac{\partial\mathscr{L}}{\partial A_\lambda}\delta A_\lambda + \frac{\partial\mathscr{L}}{\partial(\partial_\mu A_\lambda)}\delta(\partial_\mu A_\lambda) \\ \Rightarrow\quad &\delta\mathscr{L} = (\partial_\lambda\epsilon)\,J^\lambda + \frac{\partial\mathscr{L}}{\partial A^\lambda}\partial_\lambda\epsilon + \frac{\partial\mathscr{L}}{\partial(\partial_\nu A_\lambda)}\partial_\nu\partial_\lambda\epsilon = 0 \end{aligned} \tag{5.28}$$

where we asserted the above variation is zero so as to solve for our new invariant Lagrangian density. The above result tells us that the variation of the Lagrangian vanishes if:

$$\frac{\partial\mathscr{L}}{\partial A_\lambda} = -J^\lambda \tag{5.29}$$

$$\frac{\partial\mathscr{L}}{\partial(\partial_\nu A_\lambda)} = -\frac{\partial\mathscr{L}}{\partial(\partial_\lambda A_\nu)} \tag{5.30}$$

As such, we get that a new invariant Lagrangian density term and a necessary tensor dependence in the total Lagrangian density:

$$\mathscr{L}_I = -J_\mu A^\mu \tag{5.31}$$

$$\mathscr{L} = \mathscr{L}(\partial_\mu A_\nu - \partial_\nu A_\mu) \tag{5.32}$$

The example above might have been quite confusing, but we can summarize the take-aways. In summary, we have that if we have a complex "matter field" $\phi(\vec{x},t)$, the Lagrangian density $\mathscr{L}$ is invariant under the global symmetry:

$$\delta\phi(\vec{x},t) = i\epsilon q\phi(\vec{x},t) \tag{5.33}$$

However, if we replace the global transformation by a local coordinate dependent one (known as a *local gauge transformation*):

$$\delta\phi(\vec{x},t) = i\epsilon(\vec{x},t)q\phi(\vec{x},t) \tag{5.34}$$

then we get conservation if the following 3 properties are satisfied:

1. An additional gauge field $A_\mu(\vec{x},t)$ needs to be added such that $\delta A_\mu(\vec{x},t) = \partial_\mu \epsilon(\vec{x},t)$;
2. The gauge field must couple to the global conserved Noether current J^μ under the global symmetry;
3. The terms in the Lagrangian density that determines the dynamics of the gauge field must be a function of an antisymmetric tensor:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{5.35}$$

which gives the total Lagrangian density of the system as:

$$\mathscr{L} = \mathscr{L}(\phi, \partial_\mu \phi) - J_\mu A^\mu + \mathscr{L}(F_{\mu\nu}) \tag{5.36}$$

Theories which are invariant under these local gauge transformations are known as *gauge theories*".

"*Everything that we know about that is fundamental is a gauge theory.*"

− T. DeGrand, 2020.

It turns out that there is another way to build in local gauge invariances. To do this, we once again make the Lagrangian density a function of the anti-symmetric tensor, but also we convert the one of the cojugate field variables to one with a *covariant derivative*:

$$\partial_\mu \phi \quad \to \quad \boxed{D_\mu \phi \equiv \partial_\mu \phi - iqA_\mu \phi} \tag{5.37}$$

With this modification, we have that this new conjugate variable $D_\mu \phi$ transforms like ϕ under gauge transformations. As such, we have:

$$\mathscr{L}\left(\phi, D_\mu \phi, F_{\mu\nu}\right) = \mathscr{L}\left(\phi', D'_\mu \phi', F'_{\mu\nu}\right) \tag{5.38}$$

To check this, we consider:

$$\begin{aligned} D'_\mu \phi' &= \partial_\mu \phi' - iqA'_\mu \phi' \\ &= \partial_\mu \left[e^{iq\epsilon}\phi\right] - iq\left[A_\mu - \partial_\mu \epsilon\right] e^{iq\epsilon}\phi \\ &= e^{iq\epsilon}\left[\partial_\mu \phi - iqA_\mu \phi\right] \\ &= e^{iq\epsilon} D_\mu \phi \end{aligned} \tag{5.39}$$

indeed showing that $D_\mu \phi$ transforms just like ϕ. Now we ask, what then is a good candidate Lagrangian? Well, we know that $\mathscr{L}(F_{\mu\nu})$ is a scalar, so we want to write something that is a function of $F_{\mu\nu}$ but with all indices contracted:

$$\mathscr{L}(F_{\mu\nu}) = c_1 F_{\mu\nu}F^{\mu\nu} + c_2(F_{\mu\nu}F^{\mu\nu})^2 + \dots \tag{5.40}$$

§5.2 The Faraday Tensor

It turns out that the Lagrangian in this form that produces the Maxwell equations, is written as:

$$\boxed{\mathscr{L}_{EM} = -\frac{1}{16\pi}F_{\mu\nu}F^{\mu\nu} - \frac{1}{c}J_\mu A^\mu}. \tag{5.41}$$

This is **the Lagrangian for classical electrodynamics**. Let's now see how Maxwell's equations fall out of this. Recall that the equations of motion for such Lagrangian densities are given by:

$$\partial^\beta \frac{\partial \mathscr{L}}{\partial(\partial^\beta A^\alpha)} - \frac{\partial \mathscr{L}}{\partial A^\alpha} = 0, \tag{5.42}$$

so plugging in the Lagrangian density that we have term by term gives:

$$\begin{aligned}
&\frac{\partial \mathscr{L}}{\partial(\partial^\beta A^\alpha)} = -\frac{1}{16\pi} g_{\lambda\mu} g_{\nu\sigma} \left[\delta^\mu_\beta \delta^\sigma_\alpha F^{\lambda\nu} - \delta^\sigma_\beta \delta^\mu_\alpha F^{\lambda\nu} + \delta^\lambda_\beta \delta^\nu_\alpha F^{\mu\sigma} - \delta^\nu_\beta \delta^\lambda_\alpha F^{\mu\sigma}\right] \\
\Rightarrow \quad &\partial^\beta \frac{\partial \mathscr{L}}{\partial(\partial^\beta A^\alpha)} - \frac{\partial \mathscr{L}}{\partial A^\alpha} = -\frac{4}{16\pi} \partial^\beta F_{\beta\alpha} + \frac{1}{c} J_\alpha = 0 \\
\Rightarrow \quad &\boxed{\partial^\beta F_{\beta\alpha} = \frac{4\pi}{c} J_\alpha}
\end{aligned} \tag{5.43}$$

where we used the fact that $F^{\mu\nu}$ is antisymmetric above. As such, we can derive the explicit matrix form of $F^{\alpha\beta}$ as:

$$\begin{aligned}
&\vec{B} = \boldsymbol{\nabla} \times \vec{A}, \quad \vec{E} = -\boldsymbol{\nabla}\Phi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t} \\
\Rightarrow \quad &\boxed{F^{\alpha\beta} = \begin{bmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{bmatrix}}.
\end{aligned} \tag{5.44}$$

The antisymmetric tensor above is known as the *Faraday tensor*. Taking the 4-dimensional divergence of J^α then gives:

$$\begin{aligned}
&\partial_\alpha J^\alpha = 0 \\
\Rightarrow \quad &\boldsymbol{\nabla} \cdot \vec{E} = 4\pi\rho, \quad \boldsymbol{\nabla} \times \vec{B} - \frac{1}{c}\frac{\partial \vec{E}}{\partial t} = \frac{4\pi}{c}\vec{J}
\end{aligned} \tag{5.45}$$

The lack of magnetic monopoles $\boldsymbol{\nabla} \cdot \vec{B} = 0$ then follows naturally by construction and **not** by the assertion of no experimental evidence of magnetic monopoles. We can then construct a *dual field strength tensor* via the 4-dimensional Levi-Civita tensor:

$$\mathcal{F}^{\alpha\beta} \equiv \frac{1}{2}\epsilon^{\alpha\beta\mu\nu} F_{\mu\nu} \tag{5.46}$$

which also allows us to pop-out the Maxwell's equations via:

$$\begin{aligned}
\partial_\alpha \mathcal{F}^{\alpha\beta} &= \partial_\alpha \left(\frac{1}{2}\epsilon^{\alpha\beta\mu\nu} F_{\mu\nu}\right) \\
&= \frac{1}{2}\epsilon^{\alpha\beta\mu\nu} \partial_\alpha \partial_\mu A_\nu = 0
\end{aligned} \tag{5.47}$$

which vanishes since $\epsilon^{\alpha\beta\mu\nu}$ is completely anti-symmetric whereas $\partial_\alpha \partial_\mu$ is symmetric, causing the contraction of the indices α and μ to sum to zero. Now, considering the transformation properties of the Faraday tensor, it is a 2-index object so we need 2 Lorentz transformation to move it into another frame as follows:

$$F^{\mu'\nu'} = \Lambda^{\mu'}{}_{\mu}\Lambda^{\nu'}{}_{\nu}F^{\mu\nu} \quad \text{or} \quad \mathbf{F}' = \mathbf{\Lambda F \Lambda}^T \tag{5.48}$$

If we take Λ as a Lorentz boost along the x-direction (x^1-direction), we have the fields become:

$$\begin{aligned} &E_1' = E_1, \quad E_1' = B_1 \\ &E_2' = \gamma(E_2 - \beta B_3), \quad B_2' = \gamma(B_2 + \beta E_3) \\ &E_3' = \gamma(E_3 + \beta B_2), \quad B_3' = \gamma(B_3 - \beta E_2) \end{aligned} \tag{5.49}$$

or more generally (and succinctly):

$$\boxed{\vec{E}' = \gamma\left(\vec{E} + \vec{\beta}\times\vec{B}\right) - \frac{\gamma^2}{\gamma+1}\vec{\beta}(\vec{\beta}\cdot\vec{E})} \tag{5.50}$$

$$\boxed{\vec{B}' = \gamma(\vec{B} - \vec{\beta}\times\vec{E}) - \frac{\gamma^2}{\gamma+1}\vec{\beta}(\vec{\beta}\cdot\vec{B})} \tag{5.51}$$

which tells us that magnetic and electric fields intrinsically couple when we Lorentz transform between relativistic inertial frames! Let us consider an example to see this occurring explicitly.

Example:

Consider a point charge q in an inertial reference S' moving with velocity v in the x-direction relative to an observer in a stationary lab-frame S. we further assert that the observer in S is at position $(0, b, 0)$ where $b > 0$. In the frame S', we have that the electric field from the point charge is given by:

$$\vec{E}' = \frac{q}{(r')^2}\hat{n}', \quad \vec{B}' = 0 \tag{5.52}$$

Now to consider what the observer in S sees, we note that we need to convert **both** the fields **and** the coordinates! We write the components explicitly in the S' frame of the field with coordinates in the S frame as:

$$t' = \gamma t \tag{5.53}$$

$$E_1' = \frac{q(-\gamma vt)}{[b^2 + (\gamma bt)^2]^{3/2}}, \quad E_2' = \frac{qb}{[b^2 + (\gamma bt)^2]^{3/2}} \tag{5.54}$$

Then further boosting the fields into the S frame by Lorentz transformation gives:

$$E_1 = E_1' = \frac{q(-\gamma vt)}{[b^2 + (\gamma vt)^2]^{3/2}}, \quad E_2 = \gamma E_2' = \frac{\gamma qb}{[b^2 + (\gamma vt)^2]^{3/2}} \tag{5.55}$$

$$B_3 = \beta\gamma E_2' = \beta E_2 \tag{5.56}$$

So we get that in the frame S, the observer sees that the charge also produces a non-trivial magnetic field! Consider the plot of E_2 vs vt, we see that this is in fact a Lorentzian curve with the peak at $qb\gamma/b^3$ and width $\sim b/\gamma$. On the other hand, E_1 vs vt is an odd function, implying that the the radially symmetric field in the rest-frame will become squeeze along one-axis and stretched in the other when we boost to another frame.

Let's go back and ask the equation, why is the Lagrangian density for electromagnetism given as:

$$\mathscr{L}_{EM} = -\frac{1}{16\pi}F_{\mu\nu}F^{\mu\nu} \tag{5.57}$$

and does not include higher order terms? Well, it turns out that there is actually more but these are "small corrections" to what is written above. To see this, we first set $\hbar = c = 1$. In these units, $\hbar c$ gives a quantity with units of energy $\times$ length. The Lagrangian density gives us a quantity of energy density ($\sim E/L^3$), for which in these units are $\sim 1/L^4$. We further note that:

$$\begin{aligned} & F_{\mu\nu} \sim \text{electric field} \sim \frac{q}{L^2} \\ \Rightarrow \quad & F_{\mu\nu} \sim \frac{1}{L^2} \\ \Rightarrow \quad & F_{\mu\nu}F^{\mu\nu} \sim \frac{1}{L^4} \end{aligned} \tag{5.58}$$

So if we want a Lagrangian density in higher powers in $F_{\mu\nu}F^{\mu\nu}$, the coefficients that multiply these higher powers of $F_{\mu\nu}F^{\mu\nu}$ must be dimensionful quantities so that we retrieve $\mathscr{L} \sim 1/L^4$. So the most general Lagrangian density we can write down is given as:

$$\mathscr{L} = c_1 F_{\mu\nu}F^{\mu\nu} + \frac{c_2}{\Lambda^4}(F_{\mu\nu}F^{\mu\nu})^2 + \ldots \tag{5.59}$$

where Λ^4 is some energy scale. We can ask then what is Λ? Well, it could be a scale for new physics! One theory (quantum field theory) is that Λ gives a scale for virtual electron-positron pairs which will give us $\Lambda \sim m_e \sim 0.5$ MeV. The point being made here is that when we are dealing with energy scales in which classical electrodynamic scales dominate, so we really can throw away these higher order "corrections", but not so much at lower energy quantum electrodynamic scales.

§5.3 The "Mass of the Photon"

There is discussion of the "mass of a photon" in modern literature on relativistic physics, which is really asking the question, '*what is the dispersion relation for classical fields?*". Photons are actually a particular type of particle which falls into a broader category of particles known as *gauge bosons*. Gauge bosons in particle physics, are force carriers of the fundamental interactions in nature, photons of course, being the carriers of electromagnetic interactions. It turns out that even though the photon is massless (due to gauge invariance), gauge bosons (which are things analogous to the photon) can in fact have mass. To comprehend this, it would be good to diverge a little from pure electrodynamics and discuss 2 related topics:

1. Goldstone's theorem (Goldstone bosons);
2. the Higgs effect.

The Higgs effect is what actually causes these gauge bosons to pick up a mass. Let us first consider a general classical field with field variables $\{\phi_j, \partial_\mu \phi_j\}$ for which j runs from $[1, n]$. We can write the Lagrangian density of this field to be:

$$\mathscr{L} = \frac{1}{2} \sum_{j=1}^{n} (\partial_\nu \phi_j)(\partial^\nu \phi_j) - V(\phi_j) \tag{5.60}$$

$$\text{where} \quad V(\phi_j) = \frac{\mu_0^2}{2} \left(\sum_{j=1}^{n} \phi_j^2 \right) + \frac{\lambda}{4!} \left(\sum_{j=1}^{n} \phi_j^2 \right)^2 \tag{5.61}$$

with μ_0 and λ being interaction parameters. This relatively simple model is adopted in a variety of contexts, one of which arises in condensed matter physics when describing the spin of a "patch" of a system of atoms (for which this Lagrangian is referred to as the *Ginzburg-Landau model*). In this context, the $(\partial_\nu \phi_j)(\partial^\nu \phi_j)$ measures the interaction between spin-patches, and $V(\phi_j)$ measures the self-interactions within spin-patches. This is also used in mesoscopic systems which exhibit quantum behaviour such as a *Bose-Einstein condensate* (BEC) or superconducting liquid Helium using complex valued scalar fields.

The context we are interested in and will get a little into the meat of, is that for fundamental particle physics. Here, we take μ as the particle mass so we have a "force":

$$-\frac{\partial V}{\partial \phi} = \frac{\partial \mathscr{L}}{\partial \phi}, \tag{5.62}$$

arising from the potential which attempts to drive the field variables ϕ toward the minima of V. This is known as the *Higgs effect.* Clearly, this effect will only manifest if $V(\phi)$ has at least one minimum in ϕ, so we explore such a case. The easiest way to approach this is to first linearize the equations of motion about this minimum (defined to be at ϕ_0):

$$V(\phi) \approx V(\phi)|_{\phi=\phi_0} + \frac{1}{2} V''(\phi)|_{\phi=\phi_0} (\phi - \phi_0)^2 \tag{5.63}$$

Using this expansion, we get that the equation of motion goes like:

$$\left(\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2 \right) (\phi - \phi_0) = -V''(\phi_0)(\phi - \phi_0). \tag{5.64}$$

Recall that for a massless vector field, we had:

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2 \right) \vec{A} = 0, \tag{5.65}$$

for which if we have a plane-wave solution $\vec{A} \sim e^{i(\vec{k}\cdot\vec{x} - \omega t)}$, we get:

$$\left(-\frac{\omega^2}{c^2} + k^2 \right) \vec{A} = 0. \tag{5.66}$$

It also works out that for a massive scalar relativistic field, the equations of motion is:

$$\left(\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2 + m^2 \right) \Phi = 0 \tag{5.67}$$

known as the *Klein-Gordon equation* (setting $c = 1$), for which static solutions ($\partial^2\Phi/\partial t^2 = 0$) gives us solutions where $\Phi \sim e^{-mr}/r$, which is a short-ranged potential. So particles that are massive tell us that masses cause the potential to be "*screened*". The Klein-Gordon equation comes from asserting special relativity, in which its most general formulation comes from quantizing the mass-energy relativistic relation (not done in this class). Comparing this with the Klein-Gordon equation of motion, this tells us that the mass is related to the second derivative of the potential at its minimum:

$$\mu^2 = V''(\phi_0) \tag{5.68}$$

$$\Rightarrow \quad \left[\omega^2 - k^2 - V''(\phi_0)\right]\phi = 0 \tag{5.69}$$

where we adopt the notation μ for mass now. Now, since μ_0 in general is just a model parameter, we can consider 2 scenarios for the potential:

$$V(\phi;\mu_0) = \frac{1}{2}\mu_0^2\phi^2 + \frac{\lambda}{4!}\phi^4. \tag{5.70}$$

1. $\underline{\mu_0^2 > 0}$:

   In this parameter regime, we have that the potential curve looks like that shown in figure 5.1 below.

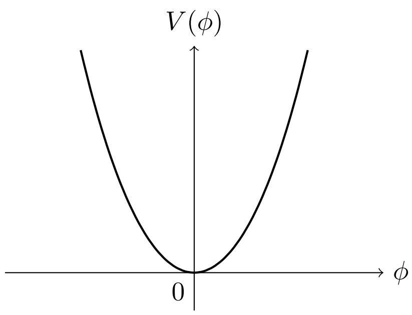

Figure 5.1: Plots of $V(\phi)$ for $\mu_0^2 > 0$.

   From the plot above, we see that the minimum here occurs at $\phi_0 = 0$, granting the linearization result that $\mu^2 = \mu_0^2$.

2. $\underline{\mu_0^2 < 0}$:

   This parameter regime is a little more interesting, as we have that the potential curve looks like that shown in figure 5.2 below.

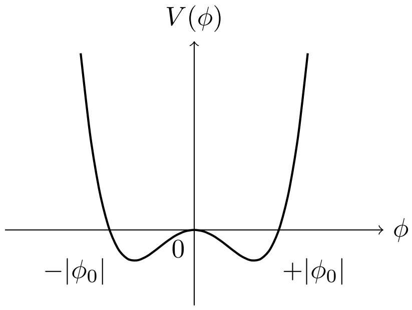

Figure 5.2: Plots of $V(\phi)$ for $\mu_0^2 < 0$.

From the plot above, we see that there are now 2 different minima which occur at:

$$\begin{aligned} & V'(\phi) = \mu^2\phi + \frac{\lambda}{6}\phi^3 \\ \Rightarrow \quad & \phi_0 = \pm\sqrt{-\frac{6\mu_0^2}{\lambda}} \end{aligned} \tag{5.71}$$

With this result, we get that the second derivative of the potential with respect to ϕ evaluated at ϕ_0 is:

$$\left.\frac{d^2V(\phi)}{d\phi^2}\right|_{\phi=\phi_0} = \mu_0^2 + \frac{\lambda}{2}\phi_0^2 = -2\mu_0^2 \tag{5.72}$$

$$\Rightarrow \quad \mu^2 = -2\mu_0^2 \tag{5.73}$$

which tells us that μ^2 measures the “rocking” of the system (much like the frequency of a harmonic oscillator), keeping in mind that $\mu_0^2 < 0$. We note that since the initial Lagrangian density consists of even powers of the conjugate field variables, granting it a discrete symmetry in the transformation:

$$\phi(x,t) \quad \rightarrow \quad -\phi(x,t). \tag{5.74}$$

However if the system then “chooses” to preferentially minimize its potential by sitting in any one of the 2 minima, one refers to this as ***spontaneous symmetry breaking***. The symmetry being referenced here is a global one and so, if the system is within one of these 2 minima, it is very hard from a local picture to to see the discrete symmetry mentioned above. To see this, we consider a small perturbation away from one of these minima:

$$\begin{aligned} & \phi(x,t) = \phi_0 + \chi(x,t) \\ \Rightarrow \quad & V(\phi_0 + \chi(x,t)) = \frac{1}{2}\mu_0^2\left(\phi_0 + \chi(x,t)\right)^2 + \frac{\lambda}{4!}\left(\phi_0 + \chi(x,t)\right)^4 \\ & \qquad = \left[\frac{1}{2}\mu_0^2 + \frac{\lambda}{4}\phi_0^2\right]\chi^2 + \mathcal{O}(\chi^3) \end{aligned} \tag{5.75}$$

Showing that it is very difficult now to still tell that $\phi \rightarrow -\phi$ (i.e. $\chi \rightarrow -\phi_0 - (\chi + \phi_0)$) is still a symmetry.

"An ant living in a magnetized ferromagnet, has a hard time realizing that the underlying system is rotationally invariant."

– T. Degrand, 2020.

§5.3.1 Goldstone Bosons

Now we are going to consider another example where by we can write the Lagrangian as:

$$\mathscr{L} = \frac{1}{2}\left[(\partial_\mu \phi_1)(\partial^\mu \phi_1) + (\partial_\mu \phi_2)(\partial^\mu \phi_2)\right] + \frac{1}{2}\mu_0^2\left[\phi_1^2 + \phi_2^2\right] - \frac{\lambda}{4!}\left[\phi_1^2 + \phi_2^2\right]^2 \tag{5.76}$$

which gives us a **continuous** symmetry (as opposed to a discrete one seen in the previous example) under the transformation:

$$\begin{bmatrix}\phi_1' \\ \phi_2'\end{bmatrix} = \begin{bmatrix}\cos\Omega & \sin\Omega \\ -\sin\Omega & \cos\Omega\end{bmatrix}\begin{bmatrix}\phi_1 \\ \phi_2\end{bmatrix}. \tag{5.77}$$

A continuous symmetry implies a conserved Noether current associated to this symmetry. If we try and draw the potential in a 3D-space with the axes being $\{\phi_1, \phi_2, V(\phi_1, \phi_2)\}$ (just like we did above), the case where $\mu_0^2 > 0$ is not very interesting once again as it will just be some cylindrically symmetric bowl like potential. However, the case where $\mu_0^2 < 0$ will give us a potential that looks like that shown in figure 5.3 below.

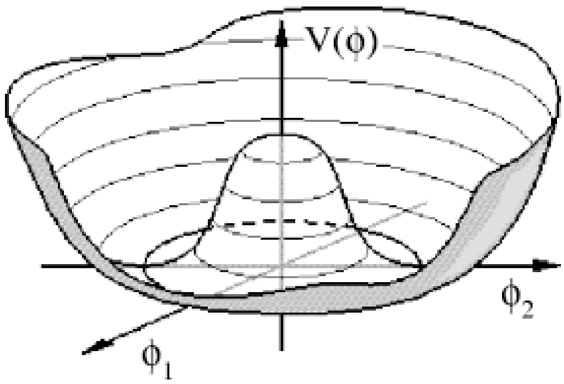

Figure 5.3: Potential function for $\mu_0^2 < 0$.

This is known as the *Mexican hat potential.* A system which wants to minimize its energy then wants to sit in the trough of this potential which is explicitly given as:

$$\phi_1^2 + \phi_2^2 = -\frac{6\mu_0^2}{\lambda} \equiv \phi_0^2 \tag{5.78}$$

Suppose now that the symmetry is now broken by setting:

$$\begin{aligned}\phi_1^2 &= \phi_0^2 \\ \phi_2 &= 0.\end{aligned} \tag{5.79}$$

A perturbation applied to this system can then be written as:

$$\begin{aligned}\phi_1(x,t) &= \phi_0 + \chi_1(x,t)\\ \phi_2(x,t) &= \chi_2(x,t)\end{aligned} \tag{5.80}$$

Plugging this into the potential and then looking at the terms up to quadratic powers in χ_1 and χ_2 gives us:

$$\begin{aligned}V(\phi_1,\phi_2) &\approx \chi_1\left[\phi_0\left(\mu_0^2+\frac{\lambda}{6}\right)+\phi_0^2\right]+\chi_1^2\left[\frac{\mu_0^2}{2}+\frac{\lambda}{4}\phi_0^2\right]+\chi_2^2\left[\frac{\mu_0^2}{2}+\frac{\lambda}{12}\phi_0^2\right]\\ &= \chi_1^2\left[-\frac{\mu_0^2}{2}\right]\end{aligned} \tag{5.81}$$

where we plugged in the solution to ϕ_0 into the expression above to get the simplification. This tells us that the χ_1 field has a mass, but the χ_1 field is massless! This result is in fact rather generic (universal), and it an example of something called *Goldstone's theorem*, for which the general statement of this theorem is as follows.

Theorem 5.3.1. Goldstone's Theorem: *When a global continuous symmetry ($U(1)$ symmetry) is spontaneously broken, there is an accompanying massless mode known as the* ***Goldstone boson****.*

Massless modes imply that arbitrarily long wavelength excitations can be supported. An example of a Goldstone boson is a spin-wave in a magnet (behaves like soundwaves).

§5.3.2 The Higgs Effect

Thus far we have been talking about global gauge symmetries and associated spontaneously breaking of these symmetries. What happens if these occurred on local gauge symmetries? Well let's look back at classical electrodynamics for an answer. Recall that we derived Maxwell's equations by imposing local gauge invariances through the use of covariant derivatives. That is, the Lagrangian remains invariant under the transformation:

$$\phi(\vec{x},t) \quad \to \quad e^{i\Lambda(\vec{x},t)}\phi(\vec{x},t). \tag{5.82}$$

So let's try to do the same for the Lagrangian we had for Goldstone bosons but with the appropriate function of the Faraday tensor inserted. This gives us:

$$\mathscr{L} = \left[\left(\partial_\mu - iqA_\mu\right)\phi\right]\left[\left(\partial_\mu + iqA_\mu\right)\phi^*\right] - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \mu_0^2|\phi|^2 + \lambda|\phi|^4 \tag{5.83}$$

where the fields now are asserted to be complex valued (sines and cosines replaced with complex phases). The sign convention for the potential is also chosen such that we have the Mexican hat potential with minima at:

$$|\phi|^2 = -\frac{\mu_0^2}{\lambda} \equiv \frac{a^2}{2} \tag{5.84}$$

where a is defined such that it is exactly the radius of the minima from the origin. As per before, let's consider a small perturbation away from the minimum just as before such that:

$$\phi(\vec{x},t) = \frac{a + \chi_1(\vec{x},t) + i\chi_2(\vec{x},t)}{\sqrt{2}}. \tag{5.85}$$

Plugging this into the potential and simplifying gives:

$$V(\phi) = \frac{\lambda a^2}{2}\chi_1^2, \tag{5.86}$$

showing that the χ_1^2 is massive and the χ_2 field is massless. As for the derivative terms in the Lagrangian, we plug in this perturbative solution to get:

$$\begin{aligned} D_\mu\phi &= \frac{1}{\sqrt{2}}\left[\partial_\mu\chi_1 - qA_\mu\chi_2 + i\left(\partial_\mu\chi_2 - qA_\mu\chi_1 - qA_\mu a\right)\right] \\ \Rightarrow \quad (D_\mu\phi)^*(D_\mu\phi) &= \frac{1}{2}\left[\left(\partial_\mu\chi_1 - qA_\mu\chi_2\right)^2 + \left(\partial_\mu\chi_2 - qA_\mu\chi_1 - qA_\mu a\right)^2\right], \end{aligned} \tag{5.87}$$

for which if we only keep the quadratic term (because that's all we're interested in), we get:

$$(D_\mu\phi)^*(D_\mu\phi) \approx \frac{1}{2}\left[\partial_\mu\chi_1\partial^\mu\chi_1 + \partial_\mu\chi_2\partial^\mu\chi_2\right] + \frac{q^2a^2}{2}A_\mu A^\mu + \frac{qa}{\sqrt{2}}A_\mu\partial^\mu\chi_2 \tag{5.88}$$

$$\Rightarrow \mathscr{L} \approx \frac{1}{2}\left[\partial_\mu\chi_1\partial^\mu\chi_1 + \partial_\mu\chi_2\partial^\mu\chi_2\right] + \frac{q^2a^2}{2}A_\mu A^\mu + \frac{qa}{\sqrt{2}}A_\mu\partial^\mu\chi_2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{\lambda a^2}{2}\chi_1^2. \tag{5.89}$$

Looking at this equation, we see that we have a term that is quadratic in the vector potential! This seems to indicate that the photon has acquired a mass. However, the $A_\mu\partial^\mu\chi_2$ term presents a confusion as it is an object we have not yet encountered. To tackle this puzzle, we first go back and consider the local gauge transformation since that is where all this fell out of. Denoting the gauge transformed field with a prime, we have:

$$\phi'(\vec{x},t) = e^{i\Lambda(\vec{x},t)}\phi(\vec{x},t) \approx \left[1 + i\Lambda(\vec{x},t)\right]\phi(\vec{x},t), \tag{5.90}$$

for which plugging in the perturbative ansatz once more gives:

$$\phi'(\vec{x},t) \approx \frac{(a + \chi_1 - \Lambda\chi_2) + i\left(\chi_2 + \Lambda\chi_1 + a\Lambda\right)}{\sqrt{2}}. \tag{5.91}$$

However, since gauge transformations do not change the physical system, we must have that ϕ' still lives in the circular trough of the Mexican hat potential. So we can write:

$$\phi'(\vec{x},t) = \frac{a + \chi_1' + i\chi_2'}{\sqrt{2}}. \tag{5.92}$$

Knowing this, we can now solve the puzzle we had with that strange new term by choosing a gauge such that $\chi_2' = 0$, which is done by:

$$\begin{aligned} &\chi_2 + \Lambda\chi_1 + a\Lambda = 0 \\ \Rightarrow \quad &\Lambda = -\frac{\chi_2}{\chi_1 + a}. \end{aligned} \tag{5.93}$$

With this change of gauge, we have that the Lagrangian is now:

$$\mathscr{L} \approx \frac{1}{2}\left[\partial_\mu \chi_1' \partial^\mu \chi_1' + \partial_\mu \chi_2' \partial^\mu \chi_2'\right] + \frac{q^2 a^2}{2}(A')_\mu (A')^\mu - \frac{1}{4}(F')_{\mu\nu}(F')^{\mu\nu} + \frac{\lambda a^2}{2}(\chi_1')^2, \tag{5.94}$$

which gives rise to a massive gauge boson from the field A_μ of mass:

$$\mu_A^2 = q^2 a^2 = \frac{q^2 \mu_0^2}{\lambda}, \tag{5.95}$$

and a massive scalar field χ_1 which is in fact the Higgs field with mass:

$$\mu_H^2 = \lambda a^2. \tag{5.96}$$

This process in which gauge bosons which are “supposed” to be massless pick up a mass because there is spontaneous symmetry breaking in the scalar field (of a local gauge symmetry) is known as the *Higgs effect*.

Having seen both local and global symmetry breaking, let us compare and contrast these a little by constructing a table 5.1.

	Unbroken Symmetry	**Broken Symmetry**
Global	2 massive scalar fields	1 massive scalar field + Goldstone boson
Local	2 massive scalar fields + massless photon	Higgs Effect

Table 5.1: Local vs global symmetry breaking.

§5.3.3 The Meissner Effect and Superconductivity

We are now going to look at a phenomenon that is the basis for *superconductivity* (the property of a material an extremely low temperature regime such that its electrical resistance vanishes and a magnetic flux is expelled). A means of observing superconductivity is through the exponential suppression of electromagnetic fields inside the superconducting material. The *Meissner effect* is the screening of magnetic fields in a superconductor that effectively arises from the same mechanism as what we saw for gauge boson mass generation. What happens physically is that at some critical temperature, the electron-phonon interactions in a material cause electron-electron interactions to become attractive, allowing the system to lower its energy from the usual free-electron gas state and condense (forming a fermionic condensation) via the formation of *cooper-pairs* (a pair of bound fermions). Formally, we can write down a wavefunction for the superconducting state $\Psi_s(r)$ such that its norm-square gives the number density of cooper-pairs $n_{cp}(r)$:

$$|\Psi_s(r)|^2 = n_{cp}(r) = \frac{n_e(r)}{2} \tag{5.97}$$

where $n_e(r)$ is the number density of electrons in the superconducting state. The goal is then to write down a free-energy function (portion of energy that is able to perform thermodynamic work at constant temperature) in term of this wavefunction and its derivatives then try to derive the equations of motion from it. Firstly, the cooper-pair state will have the intrinsic charge and mass parameters, $q^* = 2e$ and $m^* \approx 2m_e$ respectively. Because these things are charged and interact with the electromagnetic field, all derivatives will have to be replaced by covariant derivatives:

$$\frac{\hbar}{i}\boldsymbol{\nabla} \quad \rightarrow \quad \frac{\hbar}{i}\boldsymbol{\nabla} - \frac{q^*}{c}\vec{A}. \tag{5.98}$$

The free-energy for the superconductor is then written as the free-energy for a free-electron gas $F_0(T_0)$ at some temperature T_0 plus other terms associated to the expelled magnetic field and the cooper-pairs:

$$\begin{aligned} F(\Psi_s) = F_0(T_0) + \frac{1}{8\pi}\int d^3r \left\|\boldsymbol{\nabla}\times\vec{A}\right\| + \int d^3r \left[a|\Psi_s|^2 + \frac{b}{2}|\Psi_s|^4 + \ldots\right] \\ + \int d^3r \left|\frac{1}{2m^*}\left(\frac{\hbar}{i}\boldsymbol{\nabla} - \frac{q^*}{c}\vec{A}\right)\Psi_s\right|^2. \end{aligned} \tag{5.99}$$

The minimum of this free-energy function occurs when we solve for the equations of motion from this energy function which will give us:

$$\boldsymbol{\nabla}\times\vec{B} = \frac{4\pi}{c}\vec{J}_s \tag{5.100}$$

where $\vec{J}_s$ is the *superconducting current*. The superconducting current comes from taking the derivative of the free-energy with respective to the vector-potential, which gives:

$$\vec{J}_s(\vec{r}) = i\frac{q^*\hbar}{2m^*}\left[\Psi_s^*\boldsymbol{\nabla}\Psi_s - \Psi_s(\boldsymbol{\nabla}\Psi_s)^*\right] - \frac{(q^*)^2}{m^*c}|\Psi_s|^2\vec{A}(\vec{r}). \tag{5.101}$$

The equation of motion for the Ψ_s itself is also then given as:

$$\left[\frac{1}{2m^*}\left(\frac{\hbar}{i}\boldsymbol{\nabla} - \frac{q^*\vec{A}}{c}\right)^2 + b|\Psi_2|^2\right]\Psi_s = -a\Psi_s \tag{5.102}$$

which is a nonlinear equation in Ψ_s. We see that the nonlinear term associated to the constant b acts as repulsive assuming $b > 0$, which tells us that the superconducting wavefunction seemingly wants to the spread out over the entire volume. To be consistent with equation (5.100), we require that the magnetic field inside the superconductor is zero and as a simplification of our model, we assert that the density of superconducting electrons in the material is uniform (derivatives of Ψ_s vanish). This grants us that the difference between the free-energy of the superconducting state and the free-electron state is given as:

$$\begin{aligned} \Delta F &= F_s - F_0 \\ &= \int d^3r \left[a|\Psi_s|^2 + \frac{b}{2}|\Psi_s|^4 + \ldots\right] \\ &\approx V\left[a|\Psi_s|^2 + \frac{b}{2}|\Psi_s|^4\right] \end{aligned} \tag{5.103}$$

where $V = \int d^3r$. Furthermore, we have that the equation of motion for the Ψ_s field reduces to:

$$\left[a + b|\Psi_2|^2\right]\Psi_s = 0, \tag{5.104}$$

which grants us 2 possible regimes in a (assuming b is positive as we already did earlier).

1. $\underline{a > 0}$: $\Psi_s = 0$;
2. $\underline{a < 0}$: $|\Psi_s|^2 = -a/b > 0$, which will give:
$$\frac{\Delta F}{V} = -\frac{a^2}{2b} < 0, \tag{5.105}$$
which says that the free-energy favors the forming of the superconducting state.

Now looking back at the equation for the superconducting current, we have for the uniform density superconductor that:

$$\begin{aligned} \vec{J}_s(\vec{r}) &= -\frac{(q^*)^2}{m^*c}|\Psi_s|^2\vec{A}(\vec{r}) \\ \Rightarrow \quad \frac{c}{4\pi}(\boldsymbol{\nabla}\times\vec{B}) &= -\frac{(q^*)^2}{m^*c}|\Psi_s|^2\vec{A} \\ \Rightarrow \quad \boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\vec{B}) &= -\frac{4\pi(q^*)^2}{m^*c^2}|\Psi_s|^2\left(\boldsymbol{\nabla}\times\vec{A}\right) \\ \Rightarrow \quad & \boxed{\boldsymbol{\nabla}^2\vec{B} = \frac{4\pi(q^*)^2}{m^*c^2}|\Psi_s|^2\vec{B}} \end{aligned} \tag{5.106}$$

where we used that fact that $\boldsymbol{\nabla}\cdot\vec{B} = 0$ above. This result is known as the *Meissner effect* (*Meissner equation*). We then define a length parameter:

$$\begin{aligned} \lambda_L &\equiv \sqrt{\frac{m^*c^2}{4\pi(q^*)^2 n_s}} \\ \Rightarrow \quad \boldsymbol{\nabla}^2\vec{B} &= \frac{1}{\lambda_L^2}\vec{B} \end{aligned} \tag{5.107}$$

where $n_s = |\Psi_s|^2$. Solving the equation above for $\vec{B}$ will indeed give us that magnetic field is “screened” and to quickly see this, we consider a 1D system:

$$\begin{aligned} \frac{d^2B(z)}{dz^2} &= \frac{1}{\lambda_L^2}B(z) \\ \Rightarrow \quad B(z) &= B_0 e^{-z/\lambda_L}, \end{aligned} \tag{5.108}$$

showing that the magnetic field dies away exponentially. This equation for $\vec{B}$ is the just a Helmholtz equation that follows from the *London equation* (will be discussed further down), for which λ_L is known as the *London penetration depth*. Earlier, we mentioned that this mechanism is equivalent to that for photon mass generation. This can be seen by considering the Green’s function solution to the Klein-Gordon equation used when we studied the Higgs effect:

$$\left(\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2 - m^2\right)G(t,\vec{r};t',\vec{r}') = \delta(t-t')\delta^3(\vec{r}-\vec{r}'). \tag{5.109}$$

Considering just the static behavior, the above expression simplifies to:

$$\left(\boldsymbol{\nabla}^2 + m^2\right)G(\vec{r};\vec{r}') = -\delta^3(\vec{r}-\vec{r}') \tag{5.110}$$

$$\Rightarrow \quad G(\vec{r},\vec{r}') \sim \frac{e^{-m\|\vec{r}-\vec{r}'\|}}{\|\vec{r}-\vec{r}'\|}, \tag{5.111}$$

which we see it matches up exactly to the solution for the screened magnetic field! Going back to superconductivity, what does it mean to have no resistance (infinite conductivity)? We know for an ordinary conductor, we have:

$$\vec{J} = \sigma\vec{E} = en\vec{v} \tag{5.112}$$

where e is the electronic charge and n is the electron number density. In a superconductor however, it experimentally turns out that Ohm's law becomes a ballistic transport equation:

$$\begin{aligned} m^*\frac{d\vec{v}}{dt} &= q^*\vec{E} \\ \Rightarrow\quad \Aboxed{\frac{d}{dt}\vec{J}_s &= \frac{n_s(q^*)^2}{m^*}\vec{E}} \end{aligned} \tag{5.113}$$

known as the *London equation* that governs superconductivity. If we now go back to the Meissner equation, we can in fact see that:

$$\begin{aligned} \vec{J}_s(\vec{r}) &= -\frac{(q^*)^2}{m^*c}|\Psi_s|^2\vec{A}(\vec{r}) \\ \Rightarrow\quad \nabla\times\vec{J}_s &= -\frac{(q^*)^2}{m^*c}|\Psi_s|^2\left(\nabla\times\vec{A}\right) = -\frac{(q^*)^2}{m^*c}|\Psi_s|^2\vec{B}, \end{aligned} \tag{5.114}$$

and from Maxwell's equations:

$$\begin{aligned} \nabla\times\vec{E} &= -\frac{1}{c}\frac{\partial\vec{B}}{\partial t} \\ \Rightarrow\quad \nabla\times\vec{E} &= -\frac{1}{c}\frac{\partial}{\partial t}\left[-\frac{m^*c}{(q^*)^2|\Psi_s|^2}\left(\nabla\times\vec{J}_s\right)\right] \\ \Rightarrow\quad \vec{E} &= \frac{m^*}{(q^*)^2n_s}\frac{\partial\vec{J}_s}{\partial t}, \end{aligned} \tag{5.115}$$

bringing us back to the London equation. So we see that the London equation and thus superconductivity, falls out of the Meissner effect, implying that superconductivity is a result of photons picking up a mass.

Chapter 6

Relativistic Particle Dynamics

Up to now, we have gotten away with dealing with relativistic electromagnetism without ever touching on the Lorentz force law. To have a complete theory of electrodynamics however, we need to know how this fits into the relativistic formalism we have been working with. Thus far, we have seen electromagnetism fall out of the Lagrangian formalism in which symmetries are encoded. One of these symmetries is translation invariance, and is what will lead us to forces in relativistic electrodynamics.

§6.1 The Stress-Energy Tensor

To start off, we consider an infinitesimal translation and see what this leads us to.

$$\begin{aligned} & x^{\mu'} = x^\mu + \epsilon^\mu \\ \Rightarrow \quad & \delta\mathcal{L} = \frac{\partial \mathcal{L}}{\partial x_\mu}\epsilon_\mu = \epsilon_\mu \partial^\mu \mathcal{L}. \end{aligned} \tag{6.1}$$

However, we also know that:

$$\begin{aligned} & \delta\mathcal{L} = \sum_j \frac{\partial \mathcal{L}}{\partial \phi_i}\delta\phi_j + \frac{\partial \mathcal{L}}{\partial\left(\partial^\mu \phi_i\right)}\delta\left(\partial^\mu \phi_i\right) \\ \Rightarrow \quad & \epsilon_\mu \partial^\mu \mathcal{L} = \partial^\mu\left[\sum_j \frac{\partial \mathcal{L}}{\partial\left(\partial^\mu \phi_i\right)}\epsilon_\nu \partial^\nu \phi_j\right]. \end{aligned} \tag{6.2}$$

Noting that the contracted index of the variation in coordinates by ϵ_μ is arbitrary, we can now define the conserved quantity:

$$\boxed{T^{\mu\nu} \equiv -g^{\mu\nu}\mathcal{L} + \frac{\partial \mathcal{L}}{\partial\left(\partial^\mu \phi_i\right)}\partial^\nu \phi_j} \tag{6.3}$$

known as the *stress-energy tensor* (a.k.a. *energy-momentum tensor*). We can see that this is a **symmetric** tensor. To get an understanding of what this is, we look at an explicit example.

§6.1.1 Understanding the Stress-Energy Tensor

Consider a finite box of N particles with a rest-number density $n_0 = N/V^*$. If we go into a frame where the box is moving with velocity v, the number of particles in invariant but lengths of the box will be contracted so we have:

$$n_0 \quad \rightarrow \quad n = \gamma n_0. \tag{6.4}$$

We can then construct a Noether current associated to the number of particle which we call the *number current density*:

$$N^\mu = nv^\mu, \tag{6.5}$$

where v^μ is the 4-velocity, so we see that N^μ is conserved, $\partial_\mu N^\mu = 0$. Now further consider the flow of particle through some differential area dA of the box. Then the number of particles exiting the box through this area in a differential time interval dt would be:

$$\frac{dN}{dt} = \vec{N} \cdot \hat{n} dA, \tag{6.6}$$

where $\hat{n}$ is the normal vector to dA. This is also known as the number-flux of particles through dA. Before proceeding, we note that a density is a scalar quantity associated with a 3-volume. Defining a 3-volume in 4-space requires us to define a vector of volumetric orientation n_α, which allows us to say that the number of particles in a 3-volume ΔV is given by:

$$\Delta n = N^\alpha n_\alpha \Delta V. \tag{6.7}$$

With this, we can now turn our attention to the momentum of the particles in this volume. To do so, we in fact need something with 2-indices $T^{\mu\nu}$ (with the right units), since we need to contract one of those indices with the orientation vector associated to the 3-volume and get a 1-index momentum object:

$$\Delta p^\mu = T^{\mu\nu} n_\nu \Delta V. \tag{6.8}$$

To see what $T^{\mu\nu}$ would constitute of, we consider the case of the box at rest. The intuition we gain from this will then generalize to all inertial reference frames. IN the rest-fram of the box, we just have that $n^\mu = (1, 0, 0, 0)$, which gives us that:

$$\Delta p^\mu = T^{\mu 0} \Delta V \tag{6.9}$$

$$\Rightarrow \quad T^{00} = \mathcal{E}, \quad T^{i0} = \Pi_i \tag{6.10}$$

where $\mathcal{E}$ is an *energy-density* and $\Pi_i \equiv \Delta p^i/\Delta V$ a *momentum-density.* Furthermore, we also have:

$$T^{i1} = \frac{\Delta p^i/\Delta t}{\Delta y \Delta z} = \frac{\text{force}}{\text{area}} \tag{6.11}$$

which tells us that the spatial part of this tensor T is a pressure-like quantity, while:

$$T^{01} = \frac{\Delta p^0/\Delta t}{\Delta y \Delta z} = \frac{\text{power}}{\text{area}}, \tag{6.12}$$

telling us that the temporal part of T is an energy-flux like quantity. So we see that all components of the stress-energy tensor would have well-defined physical quantities in all reference frames. Harkening back to Noether's theorem, it works out that by construction, integrals of T are conserved quantities. The one in particular we are most concerned with would be:

$$P^0 = \int d^3x T^{00} = \int d^3x \left[\sum_j \Pi_j \dot{\phi}_j - \mathcal{L} \right] = \text{total energy.} \tag{6.13}$$

At this point, it would be instructive to write the stress-energy tensor explicitly for actual physical systems. We will look at 2 in particular.

1. Dust:

   This is in fact the system we have already seen above, which is defined to be a system of non-interacting particles. As we saw above, within the rest-frame, the dust stress-energy tensor just has one non-trivial element:

   $$T^{00} = \rho_E \tag{6.14}$$

   where ρ_E is the energy density. Generalizing this to any inertial reference frame is given by:

   $$T^{\mu\nu} = \rho_E u^\mu u^\nu, \tag{6.15}$$

   where u^μ is the 4-velocity.

2. Perfect Fluid:

   A perfect fluid is system that only requires 2 quantities to characterize it. These are the pressure P and energy density ρ. In its rest-frame, we have:

   $$T^{\mu\nu} = \begin{bmatrix} \rho_E & 0 & 0 & 0 \\ 0 & P & 0 & 0 \\ 0 & 0 & P & 0 \\ 0 & 0 & 0 & P \end{bmatrix}, \tag{6.16}$$

   whereas in a generic inertial frame, we have:

   $$T^{\mu\nu} = (\rho_E + P)\, u^\mu u^\nu + P g^{\mu\nu} \tag{6.17}$$

This is all well and good, but let's remember that this is a class on electrodynamics, so we can ask how this in fact comes into play for electrodynamic systems. In electrodynamics, the stress-energy tensor for the Lagrangian in free-space (no currents J^μ) is given by:

$$\mathcal{L} = -\frac{1}{16\pi} F^{\mu 0} F_{\mu 0}. \tag{6.18}$$

So using equation (6.3), the stress-energy tensor works out to be:

$$T^{\mu\nu} = -g^{\mu\nu}\mathcal{L} - \frac{1}{4} g^{\mu\alpha} F_{\alpha\beta} \partial^\nu A^\beta. \tag{6.19}$$

This can be massaged into something gauge invariant (by dropping terms that do not lead to energy-momentum conservation):

$$\boxed{\Theta^{\mu\nu} = \frac{1}{4\pi}\left[g^{\mu\alpha}F_{\alpha\beta}F^{\beta\nu} + \frac{1}{4}g^{\mu\nu}F_{\alpha\beta}F^{\alpha\beta}\right]}, \tag{6.20}$$

such that $\partial_\mu \Theta^{\mu\nu} = 0$ with components:

$$\Theta^{00} = \frac{1}{8\pi}\left[E^2 + B^2\right] \equiv u_E \tag{6.21}$$

$$\Theta^{0i} = \Theta^{i0} = \frac{1}{4\pi}\left(\vec{E}\times\vec{B}\right)_i \equiv c\vec{g} \tag{6.22}$$

$$\Theta^{ij} = -\frac{1}{4\pi}\left[E_iE_j + B_iB_j - \frac{1}{2}\delta_{ij}\left(E^2+B^2\right)\right] = -T^{ij}. \tag{6.23}$$

We see that the Θ^{00} entry resembles an energy density while the Θ^{0i} entries a Poynting-vector (energy-flux) like quantity. Altogether, we can write this as:

$$\boxed{\Theta^{\mu\nu} = \begin{bmatrix} u_E & c\vec{g} \\ c\vec{g} & -T^{ij}\end{bmatrix}}. \tag{6.24}$$

If we now want to re-introduce currents (go from free-space to a charge-rich environment), we can do some algebra to get:

$$\boxed{\partial_\mu\Theta^{\mu\nu} = -\frac{1}{c}F^{\alpha\beta}J_\beta}. \tag{6.25}$$

More details of this is given in Jackson section 12.10. Recall that at the start of the section, our motivation was to look for an object in 4-space that encapsulates the Lorentz force law. Well, the object we have just constructed, $\Theta^{\mu\nu}$ is indeed that object where the time-like components produce energy conservation and space-like components the Lorentz force law (in terms of densities):

$$\left[\partial_\alpha\Theta^{\alpha\beta}\right]_{\beta=0}: \quad \frac{\partial u_E}{\partial t} + \nabla\cdot\vec{S} = -\vec{J}\cdot\vec{E} \tag{6.26}$$

$$\left[\partial_\alpha\Theta^{\alpha\beta}\right]_{\beta=i}: \quad \frac{\partial g_i}{\partial t} - \partial_j T_{ij} = -\left[\rho_E E_i + \frac{1}{c}\left(\vec{J}\times\vec{B}\right)_i\right]. \tag{6.27}$$

§6.2 Point Particles in External Fields

Knowing how forces fit into the relativistic picture of electrodynamics, we can now look at the motion of charged, relativistic particles in external electric and magnetic fields. To start off, we consider a static and uniform magnetic field which is pretty much the simplest example we can do.

§6.2.1 Particles in Uniform Magnetic Fields

Since there is no electric field, we have that the energy is conserved (magnetic fields do not do work on charged particles):

$$\begin{aligned} &\frac{dE}{dt} = 0 && (6.28)\\ \Rightarrow\quad &|v| = \text{constant} && (6.29)\\ \Rightarrow\quad &\gamma = \text{constant}. && (6.30) \end{aligned}$$

A quantity of interest retrieved from these constants of motion is known as the *cyclotron frequency* defined as:

$$\boxed{\vec{\omega}_B \equiv \frac{q}{\gamma mc}\vec{B}}, \tag{6.31}$$

where q is the charge of the particle (this may sometimes be swapped out with e, denoting the electronic charge). This frequency will grants us a rate of change of the velocity of the charged particle in this uniform magnetic field, starting with the Lorentz force law:

$$\begin{aligned} &\frac{d}{dt}\vec{p} = \frac{q}{c}\left(\vec{v}\times\vec{B}\right)\\ \Rightarrow\quad &\frac{d\vec{v}}{dt} = \vec{v}\times\vec{\omega}_B. \end{aligned} \tag{6.32}$$

If we now take that $\vec{\omega}_B = \omega_B\hat{z}$, which gives us that:

$$\begin{aligned} &\vec{v}(t) = \hat{z}v_z + \omega_B r_0\left(\hat{x}-i\hat{y}\right)e^{-i\omega_B t}\\ \Rightarrow\quad &\vec{r}(t) = \vec{R} + v_z t\hat{z} + ir_0\left(\hat{x}-i\hat{y}\right)e^{-i\omega_B t}\\ \text{and}\quad &\frac{d\vec{v}}{dt} = -i\omega_B^2 r_0\left(\hat{x}-i\hat{y}\right)e^{-i\omega_B t}, \end{aligned} \tag{6.33}$$

where $\vec{R}$ is a constant vector and r_0 is known as the *gyration radius* set by initial conditions. Of course, the actual physical quantity would be the real part of what we got above so we have:

$$\frac{d\vec{v}}{dt} = \hat{x}r_0\sin(\omega_B t) + \hat{y}r_0\cos(\omega_B t) \tag{6.34}$$

which tells us that the motion of the positively charged particle in a uniform magnetic field would move in a circle of radius r_0. Not too surprising!

> **Note**: If we track the direction of circular motion, we see that this has a minus sign with respect to the right-hand rule. So, we get the right-hand rule for electrons (with negative charge) and the opposite for positively charged particles.

We can also consider if we initiate the particle motion with a tilt angle α from the x, y-plane such that it then follows a helix around the z-axis. In this scenario, we have now have transverse and axial components of the momentum. Often times, r_0 here is known as the *bend radius* which can directly give us the transverse component of the momentum $p_\perp$ as follows:

$$\begin{aligned} p_\perp = \gamma m v_\perp = \gamma m \omega_B r_0 = \frac{eBr_0}{c} \\ \Rightarrow \quad \boxed{cp_\perp = eBr_0}. \end{aligned} \tag{6.35}$$

That is, knowing the bend radius and strength of the magnetic field immediately allows us to compute the transverse momentum of the particle.

§6.2.2 Particles in Cross-Fields

Another interesting set-up of uniform fields is uniform $\vec{E}$ and $\vec{B}$ fields that are orthogonal to each other ($\vec{E}\cdot\vec{B} = 0$). In this scenario, there is a non-trivial electric field so energy is not constant in time. However, we note that we can always boost to a frame where either $\vec{E}$ or $\vec{B}$ vanishes. To derive this boost, we consider:

$$\vec{E}\,' = \gamma\left(\vec{E} + \vec{\beta}\times\vec{B}\right) - \frac{\gamma^2}{\gamma+1}\vec{\beta}\left(\vec{\beta}\cdot\vec{E}\right) \tag{6.36}$$

$$\vec{B}\,' = \gamma\left(\vec{B} - \vec{\beta}\times\vec{E}\right) - \frac{\gamma^2}{\gamma+1}\vec{\beta}\left(\vec{\beta}\cdot\vec{B}\right). \tag{6.37}$$

There are a few conditions that have to be met if we want to kill off anyone of the fields in this primed frame. Let's consider the 2 cases separately.

1. $\underline{\vec{E}\,' = 0}$:

   This can only be achieved if $|E| < |B|$, for which it works out (with some algebra) that this is acheived by setting the moving frame-velocity to:

   $$\frac{\vec{u}\,'}{c} = \frac{\vec{B}\times\vec{E}}{B^2} \tag{6.38}$$

   $$\Rightarrow \quad \begin{cases} \vec{E}\,' = 0; \\ \vec{B}\,' = \gamma\vec{B}\left[1 - \left|\frac{E}{B}\right|^2\right] = \frac{\vec{B}}{\gamma}. \end{cases} \tag{6.39}$$

   In the primed frame, the motion of the particle would thus just be a circle. However in the unprimed frame, the motion of the particle becomes a cycloid along the direction of $\vec{E}\times\vec{B}$. This is commonly referred to as "$\vec{E}\times\vec{B}$ *drift*".

2. $\underline{\vec{B}\,' = 0}$:

   This can only be achieved if $|B| < |E|$, and is achieved by setting:

   $$\frac{\vec{u}\,'}{c} = \frac{\vec{E}\times\vec{B}}{|E|^2} \tag{6.40}$$

   $$\Rightarrow \quad \begin{cases} \vec{B}\,' = 0; \\ \vec{E}_\parallel' = 0, \quad \vec{E}_\perp' = \frac{\vec{E}}{\gamma}. \end{cases} \tag{6.41}$$

> In this frame, only an electric field is present in the direction transverse to the particles initial velocity. This results in the particle moving in a hyperbolic trajectory with ever increasing velocity.

This is also found in Jackson section 12.3, with some elaboration on the use of cross-fields to construct a *velocity selector*.

Thus, far We have seen particle motion in uniform and static fields. What happens when fields are non-uniform? One way to treat this is via approximation schemes as we will do in the proceeding sections.

§6.2.3 Approximate Methods for Particle Dynamics

The first non-uniform field system we are going to consider is a non-uniform but static magnetic field. The first thing to do here is a Taylor expansion of the magnetic field around some position in which the magnetic field looks locally uniform:

$$\vec{B}(\vec{x}) = \vec{B}(\vec{x}_0) + \delta\vec{x} \cdot \boldsymbol{\nabla}\vec{B}(\vec{x}_0) + \dots . \tag{6.42}$$

We also assume that all inhomogeneous variations in B are transverse to B_z, so we can treat the motion as roughly circular about the field lines (i.e. magnetic field is approximately uniform along z with some "wiggles" in the transverse directions). Defining then that $\boldsymbol{\nabla}_\perp B_z$ points in the direction $\hat{n}$ (such that $\hat{n} \cdot \vec{B} = 0$) and a coordinate ξ being the radius away from the z-axis, we then have:

$$\boldsymbol{\nabla}_\perp B_z = \frac{\partial B}{\partial \xi}\hat{n}. \tag{6.43}$$

This allows us to write a coordinate dependent cyclotron frequency as:

$$\vec{\omega}_B(\vec{x}) = \vec{\omega}_0 \left[1 + \frac{1}{B_0} \left.\frac{\partial B}{\partial \xi}\right|_0 \hat{n} \cdot \vec{x}\right] \tag{6.44}$$

where ω_0 is the cyclotron frequency that would arise from a uniform field B_0. This is also saying that the deviation from a uniform magnetic field we are considering is some gradient in the transverse direction. Taking that the deviation from this uniform field cyclotron frequency is small, we define:

$$\begin{aligned} \delta &\equiv \frac{1}{B_0} \left.\frac{\partial B}{\partial \xi}\right|_0 \hat{n} \cdot \vec{x} \\ \Rightarrow \quad \vec{\omega}_B(\vec{x}) &= \vec{\omega}_0 \left[1 + \delta\right]. \end{aligned} \tag{6.45}$$

Then taking that we can also expand $\vec{v}_\perp = \vec{v}_0 + \vec{v}_1 + \dots$ (expansion terms for velocities away from $\vec{v}_0$ which is the velocity that would be due to a uniform magnetic field), we get:

$$\begin{aligned} \frac{d\vec{v}_\perp}{dt} &= [\vec{v}_0 + \vec{v}_1] \times \vec{\omega}_B \\ &\approx \vec{v}_0 \times \vec{\omega}_0 + (\vec{v}_0\delta + \vec{v}_1) \times \vec{\omega}_0, \end{aligned} \tag{6.46}$$

where we dropped the $\delta\vec{v}_1$ term. Now considering the terms separately, we get that the zeroth-order term grants us:

$$\vec{x}_0(t) = \vec{\chi} + \frac{\vec{v}_0 \times \vec{\omega}_0}{\omega_0^2} \tag{6.47}$$

which is just the circular motion we had from a uniform magnetic field with $\vec{\chi}$ being a constant. To isolate what is happening with the inhomogeneity, we take a time-average. Firstly, we note that the time-average of periodic circular motion is zero, so the zeroth order motion falls away and we are left with:

$$\begin{aligned}
&\left\langle \frac{d\vec{v}_\perp}{dt} \right\rangle = \langle \vec{v}_0 + \delta\vec{v}_1 \rangle \times \vec{\omega}_B = 0 \\
\Rightarrow\quad &\langle \vec{v}_\perp \rangle = \frac{1}{B_0} \left.\frac{\partial B}{\partial \xi}\right|_0 \langle (\hat{n}\cdot\vec{x}_0)(\vec{\omega}_0 \times \vec{x}_0) \rangle \\
\Rightarrow\quad &\vec{v}_1 = \frac{1}{B_0} \left.\frac{\partial B}{\partial \xi}\right|_0 \frac{a^2}{2} (\hat{\omega}_0 \times \hat{n}),
\end{aligned} \tag{6.48}$$

where a is the peak amplitude of $\vec{x}_0$ and we have now defined $\omega_B = eB_0/(2mc)$. In summary, we get:

$$\boxed{\vec{v}_1 = \frac{\omega_B^2 a^2}{2B_0} \left(\vec{B} \times \boldsymbol{\nabla}_\perp B \right)}. \tag{6.49}$$

This tells us that for a magnetic field that is close to uniform but with a slight inhomogeneity, we get cyclotron motion around the magnetic field but with an added component in the radial direction which plasma physicist refer to as *radiant drift*. Let's now consider a specific example.

Example:

Consider an approximately uniform linear external field $\vec{B}_0$ but with a slight curvature of the field lines (taken to be the arc of some large circle of radius R). As the particle moves through the field, it is going to experience some form of centrifugal acceleration given by $v_\parallel^2/R$, which gives us the electric field:

$$\vec{E} = \frac{\vec{F}}{q} = \frac{\gamma m v_\parallel^2}{q} \frac{\vec{R}}{R^2}. \tag{6.50}$$

This results in a sort of $\vec{E} \times \vec{B}$ drift velocity known as *curvature field velocity*:

$$\frac{\vec{v}_c}{c} = \frac{v_\parallel^2}{\omega_B R} \left(\frac{\vec{R} \times \vec{B}_0}{R B_0} \right). \tag{6.51}$$

More information on this is given in Jackson section 12.4.

§6.2.4 Adiabatic Invariants

Thus far, we have been dealing with inhomogeneities in external fields by asserting that they are static. What we will now look at is if these fields were to be time-varying. The first type of time-dependence we consider is if the field vary extremely slowly in time (i.e. adiabatically). In such contexts, the use of a quantity known as the *adiabatic invariant* is very handy.

Consider working with the Hamiltonian formalism of an electromagnetic system such that we have conjugate variables q and p that undergo periodic motion. In Hamilton-Jacobi theory, we can define an "action" (not the actual action from Lagrangian mechanics) quantity:

$$\boxed{J \equiv \oint pdq}, \tag{6.52}$$

which is in fact a constant of motion of mechanical systems even if we take our system and change its parameters, given that this change occurs adiabatically. This is the adiabatic invariant. The invariance of this quantity is best understood by considering a simple example. Let's take this guy and apply it to a system where we are increasing the strength of a uniform magnetic field with a charged particle moving through it. We are going to pick the closed integral to traverse over one cycle of the circular trajectory S of radius a and we recall that we have:

$$\begin{aligned} \vec{p} &= \gamma m\vec{v} + \frac{e\vec{A}}{c} \\ \Rightarrow \quad J &= \oint_S \left[\gamma m\vec{v} + \frac{e\vec{A}}{c}\right] \cdot d\vec{l} \\ &= 2\pi\gamma m\omega_B a^2 - \frac{e}{c}\Phi_B \\ &= \gamma m\omega_B \pi a^2, \end{aligned} \tag{6.53}$$

where $\Phi_B = B\pi a^2$ is the magnetic flux through the area of the loop S for which we applied Stoke's theorem to achieve this result. Because this is invariant, there are several invariants we can consider from this result such as the flux of the orbit Ba^2, or the orbital magnetic moment $\mu = e\omega_B a^2/(2c)$, or the the transverse momentum squared over the field $p_\perp^2/B = (\gamma m\omega_B a)^2/B$. The invariance of the orbital flux Ba^2, tells us that a particle would have an orbit that decreases in radius as B increases over time. However, $v^2 = v_\perp^2 + v_\parallel^2$ would remain constant since a magnetic field does not do work and we have energy conservation. So along with the invariant $p_\perp^2/B$, we get:

$$\boxed{\frac{v_\perp^2(z)}{B(z)} = \text{constant}}. \tag{6.54}$$

This implies that if we adiabatically increase the strength of a magnetic field ($B \to \infty$), we magnetically confine a charged particle ($\vec{v}_\perp \to 0$) along the z-direction. This is analogous to modifying the potential such that the particle would bounce off a potential barrier given that the magnetic field is sufficiently large by the end of the adiabatic process (More details in Jackson 12.5).

§6.3 Relativistic Radiation of Moving Charges

We are now going to be looking at radiation that is emitted by accelerated, relativistic charged particle. This phenomenon was used by Bohr to argue that the motion of electrons around a nucleus had to be quantized, otherwise they would radiate in circular motion and spiral inward, causing the atom to collapse. To approach this topic we start from the covariant formulation of Maxwell's equations in the presence of external sources:

$$\partial_\mu F^{\mu\nu} = \frac{4\pi}{c} J^\nu(x). \tag{6.55}$$

Using the definition of the Faraday tensor in terms of 4-potentials and working in the Lorenz gauge ($\partial_\mu A^\mu = 0$), get:

$$\Box A^\nu = \frac{4\pi}{c} J^\nu(x), \tag{6.56}$$

where $\Box$ is once again the d'Alembert operator (4-Laplacian). This is essentially a 4-space version of the wave equation, for which solutions can be found through the use of Green's functions, giving us:

$$A^\mu(x) = \frac{4\pi}{c} \int d^4x' D(x - x') J^\mu(x'). \tag{6.57}$$

where $D(x - x')$ is known as the *retarded/causal Green's function* because it ensures that the observation time x^0 is always after the source emission time $x^{0'}$. To solve this expression, we need to know what the retarded Green's function and the currents are. To start with, we know the Green's function in the Lorenz gauge (as derived in Jackson 12.11) is written as:

$$D(x - x') = \frac{\Theta(t - t')}{4\pi R} \delta(t - t' - R/c) \tag{6.58}$$

where $\vec{R} = \vec{x} - \vec{x}\,'$. We are now going to define $z_0 \equiv t - t'$ such that $z^\mu = (z_0, \vec{R})$, for which can then do a change of variable within the δ-function using its properties to get:

$$\begin{aligned} \delta(z_\mu z^\mu) &= \delta(z_0^2 - R^2) \\ &= \frac{\delta(z_0 - R) + \delta(z_0 + R)}{2R} \\ \Rightarrow \quad D(x - x') &= \Theta(z_0) \frac{\delta(z_\mu z^\mu)}{2\pi}. \end{aligned} \tag{6.59}$$

We can then also define $\eta^\mu = (1, 0, 0, 0)$ which allows us to write:

$$D(x - x') = \Theta(z_\mu \eta^\mu) \frac{\delta(z_\mu z^\mu)}{2\pi}, \tag{6.60}$$

which is intuitively a completely covariant object. Since we are dealing with point charges moving in external fields, we take the charge density as:

$$\begin{aligned}\rho(x) &= q\delta^3(\vec{x}-\vec{w}(t)) \\ &= q\int dt'\delta(t-t')\delta^3(\vec{x}-\vec{w}(t')) \\ &= \frac{q}{c}\int dt'\delta^4(x-w(\tau)) \\ &= \frac{q}{c}\int d\tau c\gamma(t)\delta^4(x-w(\tau)),\end{aligned} \tag{6.61}$$

where w denotes the coordinates of the charge. From this, we can get the current:

$$J^\mu(x) = qc\int d\tau u^\mu(\tau)\delta^4(x-w(\tau)), \tag{6.62}$$

and plug this back into the 4-vector potential to get:

$$A^\mu(x) = 2q\int d\tau u^\mu(\tau)\Theta(x_0-w_0(\tau))\delta\left[(x-w(\tau))_\mu(x-w(\tau))^\mu\right]. \tag{6.63}$$

This formula tells us that given some radiative phenomenon occurring at some earlier time, it would reach us, the observer at some later time. Evaluating the δ-function in the integral, We can simplify the 4-potential to:

$$\boxed{A^\mu(x) = \left.\frac{qu^\mu(\tau)}{(x-w(\tau))_\mu u^\mu(\tau)}\right|_{\tau=\tau_0}}. \tag{6.64}$$

This is known as the *Lienard-Wiechart potential.* Explicitly, we can write the retarded 4-potential entries as:

$$\boxed{\Phi(\vec{r},t) = \frac{q}{(1-\vec{\beta}\cdot\hat{n})R}, \quad \vec{A}(\vec{r},t) = \frac{q\vec{\beta}}{(1-\vec{\beta}\cdot\hat{n})R}}, \tag{6.65}$$

where $\hat{n}$ is a unit vector in the direction of $\vec{x}-\vec{w}(\tau)$ and $\vec{\beta} = \vec{v}(\tau)/c$. More details are found in Jackson 14.1.

§6.3.1 The Relativistic Larmor's Formula

We are now going to look at the relativistic extension of a classical expression for radiation of an accelerating charge. This is the *Larmor formula* and is used to calculate the total power radiated from accelerated non-relativistic charges:

$$\boxed{P = \frac{2}{3}\frac{q^2a^2}{c^3}}, \tag{6.66}$$

where q is the charge and a is the proper acceleration. To derive this, we use the Lienard-Wiechart potential derived in the previous section (equation 6.64) and take the limit where $v \ll c$ to get the gradient of the scalar potential:

$$-\boldsymbol{\nabla}\Phi = \frac{q}{rc}\left(\hat{n}\cdot\dot{\vec{\beta}}\right)\hat{n}, \tag{6.67}$$

and the time-derivative of the magnetic vector potential (using the multipole expansion and only keeping the dipole term):

$$\begin{aligned} \vec{A} &= \frac{1}{r}\dot{\vec{p}}(t_0) \\ \Rightarrow \quad \frac{1}{c}\frac{\partial\vec{A}}{\partial t} &= \frac{1}{rc}\ddot{\vec{p}}(t_0). \end{aligned} \tag{6.68}$$

Putting these together, we get:

$$\begin{aligned} \vec{E} &= -\boldsymbol{\nabla}\Phi - \frac{1}{c}\frac{\partial\vec{A}}{\partial t} \\ \Rightarrow \quad & \boxed{\vec{E} = \frac{q}{cr}\left[\hat{n}\left(\hat{n}\cdot\dot{\vec{\beta}}\right) - \dot{\vec{\beta}}\right] = \frac{q}{rc}\left[\hat{n}\times\left(\hat{n}\times\dot{\vec{\beta}}\right)\right]}. \end{aligned} \tag{6.69}$$

This results in fact looks rather similar to the radiation from a classical electric dipole which takes the form:

$$\vec{E}_{\text{rad}} = -\frac{k^2 e^{ikr}}{r}\left[\hat{n}\times\left(\hat{n}\times\dot{\vec{\beta}}\right)\right]. \tag{6.70}$$

So it turns out that these 2 systems will in fact have very similar antenna patterns, except for the following. In the case of a moving accelerating charge in an external field, the radiated field computed in equation (6.69) is associated to an instantaneous energy flux, implying $\vec{E}$ and $\vec{B}$ are fully real. On the other hand, the dipole radiation in equation (6.70) occurs due to a complex $e^{i\omega t}$ time-dependence which we later take the time-average of. So we see that when we compute the angular power distribution from the Lienard-Wiechart potentials, the moving charge produces:

$$\begin{aligned} \frac{dP}{d\Omega} &= R^2\hat{n}\cdot\vec{S} \\ \Rightarrow \quad \frac{dP}{d\Omega} &= \frac{q^2}{4\pi c}\left[\hat{n}\times\left(\hat{n}\times\dot{\vec{\beta}}\right)\right]^2, \end{aligned} \tag{6.71}$$

where the angular power distribution now can be an explicit function of time, unlike how we were doing these time-averages before. In the simple case where we have linear motion of the charged particle with acceleration $\vec{a} = c\dot{\vec{\beta}}$ and $\hat{R}\cdot\hat{a} = \cos\theta$, we get:

$$\frac{dP}{d\Omega} = \frac{q^2a^2}{4\pi c}\sin^2\theta \tag{6.72}$$

$$\Rightarrow \quad P = \frac{2}{3}\frac{q^2a^2}{c^3} \tag{6.73}$$

which is indeed Larmor formula for power radiated. Now what about relativity? Well we can generalize the result above to account for relativity by first rewriting the power from the Larmor formula as:

$$P = \frac{2q^2}{3m^2c^3}\left(\frac{d\vec{p}}{dt}\cdot\frac{d\vec{p}}{dt}\right), \tag{6.74}$$

where $\vec{p}$ above is the momentum (not dipole moment). So to generalize this to relativity, we instead write the momenta as 4-momenta:

$$P = -\frac{2q^2}{3m^2c^3}\left(\frac{dp^\mu}{d\tau}\frac{dp_\mu}{d\tau}\right). \tag{6.75}$$

Recall that the 4-acceleration is defined as $a^\mu = du^\mu/d\tau$, which works out explicitly to be:

$$a^\mu = c\gamma^2\begin{bmatrix}\gamma^2\left(\vec{\beta}\cdot\dot{\vec{\beta}}\right)\\ \dot{\vec{\beta}}+\gamma^2\vec{\beta}\left(\vec{\beta}\cdot\dot{\vec{\beta}}\right)\end{bmatrix}. \tag{6.76}$$

Plugging this into the genrealized Larmor formula gives:

$$\begin{aligned}P &= \frac{2}{3}\frac{q^2}{c}\gamma^4\left[\dot{\beta}^2+\gamma^2\left(\vec{\beta}\cdot\dot{\vec{\beta}}\right)^2\right]\\ &= \frac{2}{3}\frac{q^2}{c}\gamma^4\left[(1+\gamma^2\beta^2)\dot{\beta}^2-\gamma^2\left(\vec{\beta}\times\dot{\vec{\beta}}\right)^2\right]\\ \Rightarrow\quad &\boxed{P = \frac{2}{3}\frac{q^2}{c}\gamma^6\left[\dot{\vec{\beta}}^2-\left(\vec{\beta}\times\dot{\vec{\beta}}\right)^2\right]}.\end{aligned} \tag{6.77}$$

It turns that that this was already derived in 1898 by Lienard before the theory of relativity! What is interesting about this, is that there is a factor of γ to the sixth power, so it is thoroughly affected by relativistic effects. To see this in application, we consider 2 examples, the first being an example where relativistic radiation is negligible, while the second the converse.

1. Linear Accelerated Motion:

   Let's consider we have some particle accelerating along its initial direction of motion. This causes the cross-product term to vanish in the power formula, leaving us with:

   $$P = \frac{2}{3}\frac{q^2}{c}\gamma^6\dot{\vec{\beta}}^2. \tag{6.78}$$

   If we go back to the original generalization of the radiated power, we also note that we can recast the expression into the following:

   $$\begin{aligned}P &= -\frac{2q^2}{3m^2c^3}\left(\frac{dp^\mu}{d\tau}\frac{dp_\mu}{d\tau}\right)\\ &= -\frac{2q^2}{3m^2c^3}\gamma^2\left[\frac{1}{c^2}\left(\frac{dE}{d\tau}\right)^2-\left(\frac{d\vec{p}}{d\tau}\right)^2\right]\\ &= \frac{2q^2}{3m^2c^3}\left(\frac{d\vec{p}}{dt}\right)^2 = \frac{2q^2}{3m^2c^3}\left(\frac{\partial E}{\partial x}\right)^2.\end{aligned} \tag{6.79}$$

Let's then compare this to the input power, which we take as $P_{in} = (\partial E/\partial x)v$. So we have:

$$\boxed{\frac{P_{rad}}{P_{in}} = \frac{2q^2}{3m^2c^3}\frac{1}{v}\frac{\partial E}{\partial x}}. \tag{6.80}$$

This number turns out to be extremely small, so the power loss from radiation from a linear accelerating charge is negligible.

2. Circular Motion:

   For circular motion, we consider the regime whereby in each orbit, the charge is accelerated such that the change in energy is small but the change in linear momentum is large. In this case, we have:

$$\left|\frac{d\vec{p}}{d\tau}\right| = \gamma\left|\frac{d\vec{p}}{dt}\right| \approx \gamma\omega p, \tag{6.81}$$

   if we take $p \sim e^{i\omega t}$. Plugging this back into the radiated power formula gives us:

$$P = \frac{2}{3}\frac{q^2}{m^2c^3}\gamma^2\omega^2p^2. \tag{6.82}$$

   Then taking $p = \gamma m\omega r_0$ where r_0 is the gyration/bend radius and $\omega = c\beta/r_0$, we have:

$$P = \frac{2}{3}\frac{q^2}{c}\frac{(\gamma\beta)^4}{r_0^2}. \tag{6.83}$$

   Let's now ask, how much energy so we have to put in to compensate for the radiation loss so as to keep the particle in the same orbit? Well, for this we consider:

$$\delta E_{rad} = PT, \tag{6.84}$$

   where $T = 2\pi/\omega$. Plugging all the quantities in gives:

$$\boxed{\delta E_{rad} = \frac{4\pi}{3}\frac{q^2}{r_0}\beta^3\gamma^4}. \tag{6.85}$$

   This is typically a significant number especially compared to the radiative losses from linear accelerating charges!

   > **Remark**: These losses caused technological issues for particle physicists back in the 90s when they were trying to maintain the particle orbits in LEP (large electron-positron collider).

The down side of the Larmor formula, is that it doesn't give us the antenna pattern and frequency distribution of the radiation. To get these quantities, we pretty much need to go back to the drawing board and rederive these things in the relativistic picture (Jackson 14.3). From the retarded Green's function, we can derive the radiated electric field as:

$$\vec{E} = \frac{q}{c}\left[\frac{\hat{n}\times\left[\left(\hat{n}-\vec{\beta}\right)\times\dot{\vec{\beta}}\right]}{R(1-\vec{\beta}\cdot\hat{n})^3}\right]_{ret}, \tag{6.86}$$

where we recall that R is the distance between the observer and the retarded position. If we just take this to compute $dP/d\Omega$ via the relation:

$$\frac{dP}{d\Omega} = R(\tau_0)^2\hat{n}\cdot\vec{S}, \tag{6.87}$$

this observed energy flux will **not** be the same as the rate of energy emitted by the particle! This is due to the retardation of the radiation observed compared to that emitted by the particle for a given time interval. It works out that if we define the power radiated as that observed in the charge's rest frame, this becomes:

$$\frac{dP}{d\Omega} = R^2(t)\hat{n}\cdot\vec{S}\,\frac{dt_{obs}}{dt_r}, \tag{6.88}$$

where $t_{obs} = t_r + R/c$ (derivation in Jackson 14.3). Computing this Jacobian term gives:

$$\begin{aligned}\frac{dt_{obs}}{dt_r} &= \frac{d}{dt}\left[t_r + \frac{R}{c}\right]\\ &= 1 + \frac{1}{2Rc}\frac{d}{dt}\left[\vec{x}-\vec{w}(t)\right]^2\\ &= 1 - \frac{\vec{v}}{c}\cdot\frac{\vec{R}(t)}{R}\\ &= 1-\vec{\beta}\cdot\hat{n}.\end{aligned} \tag{6.89}$$

Plugging this into the angular power distribution gives:

$$\boxed{\frac{dP}{d\Omega} = \frac{q^2}{4\pi c}\frac{\left[\hat{n}\times\left[\left(\hat{n}-\vec{\beta}\right)\times\dot{\vec{\beta}}\right]\right]^2}{\left(1-\vec{\beta}\cdot\hat{n}\right)^5}}. \tag{6.90}$$

Great, now we have an angular power distribution per unit solid angle which we sought to find. Let's consider this in 2 specific cases.

1. <u>Linear Motion</u>:

   The first and simplest case is when the direction of $\vec{\beta}$ and $\dot{\vec{\beta}}$ are colinear, i.e. the particle undergoes linear acceleration. In this scenario, the angular radiated power distribution is given by:

$$\frac{dP}{d\Omega} = \frac{q^2a^2}{4\pi c^3}\frac{\sin^2\theta}{(1-\beta\cos\theta)^5}, \tag{6.91}$$

where θ is the angle between the axis of $\vec{\beta}$ or $\dot{\vec{\beta}}$ and the observer, and a is the proper acceleration. This is very similar to the angular distribution from a linear nonrelativistic dipole, but modified by the factor in the denominator. In the highly relativistic limit, the angle θ becomes small due to the deformation of the antenna pattern. This grants the approximation:

$$\frac{dP}{d\Omega} \approx \frac{q^2 a^2}{4\pi c^3} \frac{\gamma^8 (\gamma\theta)^2}{(1+\gamma^2\theta^2)^5}, \tag{6.92}$$

which tells us we get a lot more radiation if we are observing along the direction of the charge. This can be seen from a plot of the angular radiated power distribution, comparing relativistic and non-relativistic particles experiencing linear acceleration as shown in figure 6.1 below.

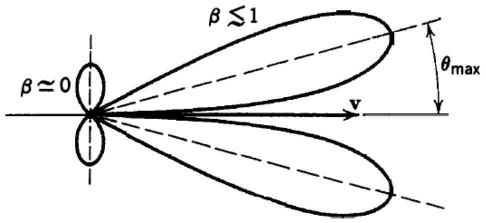

Figure 6.1: $dP/d\Omega$ for relativistic and non-relativistic linear acceleration.

Circular Motion:

The other case we can consider is when the directions of $\vec{\beta}$ and $\dot{\vec{\beta}}$ are orthogonal. This constitutes circular motion, where acceleration is directed to the center of the circular path. If we consider an instantaneous slice of time and define our coordinate frame such that $\vec{\beta}$ lies along $\hat{z}$ and $\dot{\vec{\beta}}$ lies along $\hat{x}$, we get:

$$\frac{dP}{d\Omega} = \frac{q^2 a^2}{4\pi c^3 (1-\beta\cos\theta)^3} \left[1 - \frac{\sin^2\theta\cos^2\phi}{\gamma^2(1-\beta\cos\theta)^2}\right], \tag{6.93}$$

for which in the highly relativistic limit ($\gamma \gg 1$), we get:

$$\frac{dP}{d\Omega} \approx \frac{2e^2 a^2}{\pi c^3} \frac{\gamma^6}{(1+\gamma^2\theta^2)^3} \left[1 - \frac{4\gamma^2\theta^2\cos^2\phi}{(1+\gamma^2\theta^2)^2}\right]. \tag{6.94}$$

This once again implies that we get peak power at forward angles, similar to the in the linear motion case.

More details on these derivations can be found in Jackson section 14.4.

§6.3.2 Frequency Distribution of Radiation

More notes to be added soon (reference: ***Jackson section 14.5****).*

$$\boxed{\frac{d^2 I}{d\omega d\Omega} = \frac{e^2}{4\pi^2 c}\left|\int_{-\infty}^{\infty} \frac{\hat{n}\times\left[\left(\hat{n}-\vec{\beta}\right)\times\dot{\vec{\beta}}\right]}{\left(1-\vec{\beta}\cdot\hat{n}\right)^2}\exp\left(i\omega\left[t-\frac{\hat{n}\cdot\vec{w}(t)}{c}\right]\right)dt\right|^2} \tag{6.95}$$

§6.3.3 Bremsstrahlung Radiation

Bremsstrahlung is the German word for "breaking radiation" and pertains to the following. If we have a collision of 2 charged particles, the collision causes a deceleration of the particles and in turn, also radiation. In most situations, we have a low-mass particle colliding with a highly massive particle. In such scenarios, the low-mass particle would experience most of the acceleration, so radiation from this process is mostly due to the low-mass particle. Furthermore, we assume that the collision occurs within the time interval $[0, \tau]$, and the frequency ω is small but $\omega\tau \ll 1$.

This results in the energy radiated per unit solid angle for unit frequency interval:

$$\frac{d^2 I(\hat{\epsilon})}{d\omega d\Omega} = \frac{e^2}{4\pi^2 c}\left|\frac{\hat{\epsilon}^*\cdot\vec{\beta}_f}{1-\vec{\beta}_f\cdot\hat{n}} - \frac{\hat{\epsilon}^*\cdot\vec{\beta}_i}{1-\vec{\beta}_i\cdot\hat{n}}\right|^2. \tag{6.96}$$

Amazingly, even though this is a classical result, the quantum mechanical process of photon emission from a collision process reduces to this result if $\hbar\omega \ll$ any other energy scale of the system.

More notes to be added soon (classical Bremsstrahlung, infrared divergence. reference: ***Jackson chapter 15****).*

For relativistic Bremsstrahlung radiation, there is a long way and a short way to do this. The long way is by through considering a Lorentz boost and churning through all the math. A faster way is to think about the system in its kinematic picture. Before we get into this, we first note several useful facts.

1. Quanta of photons can be counted and leads to the relation $dI(\omega) = \hbar\omega\, dN(\omega)$.
2. The differential segment of the wave-vector is invariant, and can be rewritten as:

$$\frac{d^3 k}{2k_0} = d^3 k dk_0 \delta(k^2 - k_0^2) = d^4 k \delta(k_\mu k^\mu), \tag{6.97}$$

   where k_0 is the zeroth component of the 4-momentum. This leads to the following invariants:

$$\begin{aligned} &\frac{dN}{d^3k/k_0} \propto \frac{dN}{\omega d\omega d\Omega_\gamma} \propto \frac{d^2 I}{\omega^2 d\omega d\Omega_\gamma} \\ \Rightarrow\quad &\frac{d^2 I}{d\omega d\Omega_\gamma} = \left(\frac{\omega}{\omega'}\right)^2 \frac{d^2 I'}{d\omega' d\Omega'_\gamma}. \end{aligned} \tag{6.98}$$

   This tells us that energies radiated per unit solid angle/frequency are related in different relativistic frames by relativistic Doppler shifts. Geometrically, we can work out the

Doppler shift as:

$$\frac{\omega}{\omega'} = \frac{1}{\gamma(1 - \beta\cos\theta)}. \tag{6.99}$$

Chapter 7

Quantum Electrodynamics

So far, we have been studying much about classical electrodynamics, however the world is fundamentally quantum mechanical, so it is appropriate to end the course drifting away from classical mechanics and asking how quantum mechanics interacts with classical E&M fields? That is we ask, what is the photon? What is normally done in introductory quantum mechanics is known as semiclassical radiation theory, where matter is treated as quantum mechanical while the electromagetic field is treated classically. A more comprehensive and accurate theory is one that quantizes the electromagnetic field itself. This is known as ***quantum electrodynamics (QED)****. We will see that to do this, working in the Coulomb gauge* ($\boldsymbol{\nabla} \cdot \vec{A} = 0$) *is most convenient.*

§7.1 Quantization

In classical mechanics, we start with a Lagrangian $\mathcal{L}(q, \dot{q})$, with conjugate variables $\{q, \dot{q}\}$, for which we derive equations of motion via the Euler-Lagrange equations. Furthermore, we can derive a Hamiltonian from the Lagrangian formalism:

$$H = \sum_k p_k q_k - \mathcal{L} \tag{7.1}$$

where $p_k = \partial \mathcal{L}/\partial \dot{q}_k$. What we do now to quantize such a system is by promoting the conjugate variables to *quantum operators*, such that their Poisson bracket relations are now *commutator relations*:

$$[\hat{q}_k, \hat{p}_l] = i\hbar \delta_{k,l}. \tag{7.2}$$

This procedure can also be done for the electromagnetic field, where we are now working with a field, so we replace the Lagrangian with a Lagrangian density as in classical field theory. The Lagrangian density we recall, is:

$$\mathcal{L}(A_{\mu, \partial_\mu A_\mu}) = -\frac{1}{16\pi} F_{\mu\nu} F^{\mu\nu} = \frac{1}{8\pi}(E^2 - B^2). \tag{7.3}$$

We can immediately see that the canonical coordinate variable would be A_μ, but what is the canonical momentum? Well, we can try (as motivated from Hamiltonian mechanics):

$$\Pi^k = \frac{\partial \mathcal{L}}{\partial \dot{A}_k} = -\frac{1}{8\pi}\left[\frac{\partial A_k}{\partial x^0} - \frac{\partial A_0}{\partial x^k}\right] = \frac{E_k}{4\pi}, \tag{7.4}$$

where $k = 1, 2, 3$. As for the zeroth component, it turns out that there is no momentum conjugate to A_0. So we have that $A_0 = \Phi$ is **not** a dynamical variable! As such, we get a Hamiltonian:

$$\mathcal{H} = \sum_{k=1}^{3} \Pi_k \dot{A}_k - \mathcal{L} = \frac{1}{8\pi}\left[E^2 + \left(\boldsymbol{\nabla} \times \vec{A}\right)^2\right]. \tag{7.5}$$

Now we stop and think. In free-space, we have that $\boldsymbol{\nabla} \cdot \vec{E} = 0$, which implies that the 3-components of the electric field which is a dynamical variable, are not independent (only 2 of them are). Counting degrees of freedom then becomes a mess. To avoid getting confused here, let's make a gauge choice such that the number of degrees of freedom in $\vec{A}$ is the same as those in $\vec{E}$. This is of course, the Coulomb gauge which asserts $\boldsymbol{\nabla} \cdot \vec{A} = 0$. We note again that Φ does not participate in any of this story (i.e. Φ has nothing to do with photons), so we only work with 4 dynamical variables (2 A's and 2 E's).

§7.2 Semiclassical Radiation Theory

In semiclassical radiation theory, as mentioned, is where we treat the electromagnetic fields as classical while everything else is quantum. As such, we take the vector potential as a classical plane wave:

$$\vec{A}(\vec{x}, t) = C\hat{\epsilon}\left[e^{i(\vec{k}\cdot\vec{x} - \omega t)} + e^{-i(\vec{k}\cdot\vec{x} - \omega t)}\right]. \tag{7.6}$$

We then plug this guy into the Hamiltonian for a charged particle in an electric field, giving us:

$$H = \frac{1}{2m}\left[\vec{p} - e\frac{\vec{A}}{c}\right]^2 + V(\vec{x}). \tag{7.7}$$

What we'll do now is expand and collect terms:

$$\begin{aligned} H &= \frac{p^2}{2m} + V(\vec{x}) - \frac{e}{mc}\vec{A}\cdot\vec{p} + \frac{e^2A^2}{2mc^2} \\ &= H_0 + H_1 + H_3 \\ \text{where} \quad H_0 &\equiv \frac{p^2}{2m} + V(\vec{x}) \\ H_1 &\equiv -\frac{e}{mc}\vec{A}\cdot\vec{p} \\ H_2 &\equiv \frac{e^2A^2}{2mc^2}. \end{aligned} \tag{7.8}$$

From this, we can promote $\vec{A}$ and $\vec{p}$ to operators (noting that they commute), and look for radiation via computing transition probabilities using *Fermi's golden rule*. To apply the golden rule, we need to treat H_1 as a harmonic perturbation:

$$\hat{H}_1 = -\frac{e}{mc} C \left[e^{i(\vec{k}\cdot\vec{x}-\omega t)} + e^{-i(\vec{k}\cdot\vec{x}-\omega t)} \right] \hat{\epsilon}\cdot\hat{\vec{p}}, \tag{7.9}$$

and drop the H_2 term. This gives us:

$$d\Gamma_{i\to f} = \frac{2\pi}{\hbar} \left| \langle f| \hat{H}_1 |i\rangle \right|^2 \delta(E_f - E_i \pm \hbar\omega) \left[\frac{V d^3k}{(2\pi)^3} \right], \tag{7.10}$$

The term in square-brackets above is the phase-space factor to count classical modes in a volume V. We now need to get into some details. First of, what is the factor C that normalizes the vector potential. There are (at least) 2 choices (which turn out to be related in the end but not in the semiclassical approximation). These 2 choices depend on the process we are studying. These processes can be:

1. *absorption* or *stimulated emission* which gives us a transition rate that is proportional to the intensity of the radiation source. In this case, we can define an intensity within some frequency interval:

$$\begin{aligned} I(\omega)d\omega &= \frac{1}{8\pi}(E^2+B^2)d\omega \\ &= \frac{1}{8\pi}\left[\frac{1}{c^2}\left|\frac{\partial \vec{A}}{\partial t}\right|^2 + (\boldsymbol{\nabla}\times\vec{A})^2 \right] d\omega \\ &= \frac{1}{2\pi}|C|^2 \sin^2(\vec{k}\cdot\vec{x}-\omega t)\left[\frac{\omega^2}{c^2} + (\vec{k}\times\hat{\epsilon})^2 \right]. \end{aligned} \tag{7.11}$$

2. using the idea of "*normalizing the energy of a photon*". The electromagnetic field for a photon has energy density approximated as:

$$\begin{aligned} \frac{\hbar\omega}{V} &= \frac{\langle E^2+B^2\rangle}{8\pi} = \frac{|C|^2\omega^2}{2\pi c^2} \\ \Rightarrow \quad |C|^2 &= \frac{2\pi\hbar c}{\omega V}. \end{aligned} \tag{7.12}$$

   So this gives us:

$$\vec{A}(\vec{x},t) = \sqrt{\frac{2\pi\hbar c}{\omega V}}\, \hat{\epsilon} \left[e^{i(\vec{k}\cdot\vec{x}-\omega t)} + e^{-i(\vec{k}\cdot\vec{x}-\omega t)} \right]. \tag{7.13}$$

Another little detail, is that the perturbative Hamiltonian $\hat{H}_1$ gives a matrix element:

$$C \langle f| \frac{e}{mc} \left[e^{i(\vec{k}\cdot\vec{x}-\omega t)} + e^{-i(\vec{k}\cdot\vec{x}-\omega t)} \right] \hat{\epsilon}\cdot\hat{\vec{p}}|i\rangle, \tag{7.14}$$

which is in general nasty to compute. However, what we can do is do a multipole expansion, which turns out to the be easier to compute than in the classical case! In quantum mechanics, the dipole approximation simply gives us:

$$e^{\pm i\vec{k}\cdot\vec{x}} \approx 1. \tag{7.15}$$

The matrix element is then just:

$$\langle f|\,\hat{H}_1\,|i\rangle \approx \left(\frac{Ce}{mc}\right)\hat{\epsilon}\cdot\langle f|\,\hat{\vec{p}}\,|i\rangle\,. \tag{7.16}$$

This can in fact be further simplified by recalling the Heisenberg equation of motion:

$$i\hbar\frac{\partial\hat{x}}{\partial t} = \left[\hat{x},\hat{H}\right], \tag{7.17}$$

along with the Ehrenfest's theorem which gives:

$$\begin{aligned}\langle f|\,\hat{\vec{p}}\,|i\rangle &= -\frac{im}{\hbar}\langle f|\left[\hat{x},\hat{H}\right]|i\rangle\\ &= -im\omega_{f,i}\,\langle f|\,\hat{\vec{x}}\,|i\rangle\\ \Rightarrow\quad \langle f|\,\hat{H}_1\,|i\rangle &= -iC\frac{e\omega_{f,i}}{c}\hat{\epsilon}\cdot\langle f|\,\hat{\vec{x}}\,|i\rangle\,.\end{aligned} \tag{7.18}$$

So this gives us a transition rate:

$$\frac{d\mathbb{P}}{dt} = \frac{2\pi e^2\omega_{f,i}^2}{\hbar c^2}\left|\langle f|\,\hat{H}_1\,|i\rangle\right|^2\left(\frac{2\pi c^2}{\omega^2}\right)I(\omega)\Delta\omega\;\delta(\Delta E-\hbar\omega). \tag{7.19}$$

We can then ask, how do we interpret this? Well, we can think of this as an absorption rate Γ_{abs} by integrating over the frequency intervel $\Delta\omega$:

$$\begin{aligned}\Gamma_{\text{abs}} &= \frac{2\pi e^2\omega_{f,i}^2}{\hbar^2c^2}\int d\omega\left|\langle f|\,\hat{H}_1\,|i\rangle\right|^2\left(\frac{2\pi c^2}{\omega^2}\right)I(\omega)\;\delta(\Delta E-\hbar\omega)\\ &= \frac{4\pi e^2}{\hbar^2}\left|\hat{\epsilon}\cdot\langle f|\,\hat{\vec{x}}\,|i\rangle\right|^2 I(\Delta E/\hbar=\omega_{f,i}).\end{aligned} \tag{7.20}$$

The prefactor, in the result above is known as the *Einstein coefficient*:

$$\boxed{B_{f,i} = \frac{4\pi e^2}{\hbar^2}\left|\hat{\epsilon}\cdot\langle f|\,\hat{\vec{x}}\,|i\rangle\right|^2}. \tag{7.21}$$

These Einstein coefficients in fact give rise to *selection rules* for the transition. There are several types of such transitions.

1. Forbidden transitions: $\langle f|\,\hat{\vec{x}}\,|i\rangle = 0$.

   Despite its name, these might not actually be completely forbidden, but just largely suppressed compared to the electric dipole transition, which is the approximation we are working with. Note that $\hat{\vec{x}}$ is a vector operator and transforms under rotations like an $l = 1$ object. So all these matrix elements are proportional to Clebsch-Gordan coefficients.

Let's now change gears a little and consider the spontaneous emission of a photon in the process $i \to f + \gamma$. In this case, the decay probability rate can then be thought of by considering the direction of the outgoing photon and its wave number. This gives us:

$$\begin{aligned} d\Gamma_{\text{se}} &= \frac{2\pi e^2\omega_{f,i}^2}{\hbar c^2}|\hat{\epsilon}\cdot\langle f|\,\hat{x}\,|i\rangle|^2\left(\frac{2\pi c^2}{\omega^2}\right)\delta(\Delta E-\hbar\omega)\left(\frac{2\pi\hbar c}{\omega V}\right)\left(\frac{Vd^3k}{(2\pi)^2}\right) \\ \Rightarrow\quad & \boxed{\frac{d\Gamma_{\text{se}}}{d\Omega_\gamma} = \frac{e^2\omega_{f,i}^3}{2\pi\hbar c^2}|\hat{\epsilon}\cdot\langle f|\,\hat{x}\,|i\rangle|^2}, \end{aligned} \tag{7.22}$$

where we integrated over $\omega^2 d\omega$, with the differential phase-space volume element being $d^3k = \omega^2 d\omega d\Omega_\gamma/c^3$. What we can do now is sum over all the polarizations, then integrate over all angular space to get:

$$\begin{aligned} \left.\frac{d\Gamma_{\text{se}}}{d\Omega_\gamma}\right|_{\text{sum pol}} &= \frac{e^2\omega_{f,i}^3}{2\pi\hbar c^2}|\langle f|\,\hat{x}\,|i\rangle|^2\sin^2\theta \\ \Rightarrow\quad & \boxed{\Gamma_{\text{se}} = \frac{4}{3}\frac{e^2\omega_{f,i}^3}{\hbar c^3}|\langle f|\,\hat{x}\,|i\rangle|^2}. \end{aligned} \tag{7.23}$$

§7.3 Quantizing the Electromagnetic Field

The quantum mechanics of an EM field is in fact just the quantum mechanics of a simple harmonic oscillator but with lots of indices. What characterizes a harmonic oscillator is that the conjugate variables in the Hamiltonian are quadratic:

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2. \tag{7.24}$$

This classical Hamiltonian will give rise to the Hamilton's equations of motion, for which by taking the appropriate linear combination, we get:

$$\frac{d}{dt}\left[q \pm \frac{ip}{m\omega}\right] = \mp i\omega\left[q \pm \frac{ip}{m\omega}\right]. \tag{7.25}$$

Then defining:

$$\begin{aligned} & A \equiv q \pm \frac{ip}{m\omega} \\ \Rightarrow\quad & \frac{dA}{dt} = -i\omega A, \qquad \frac{dA^*}{dt} = i\omega A^* \\ \Rightarrow\quad & \boxed{H = \frac{m\omega^2}{4}\left(A^*A + AA^*\right)}. \end{aligned} \tag{7.26}$$

In quantum mechanics on the other hand, this system is usually solved by constructing the creation and annihilation operators:

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\hat{q} + \frac{i\hat{p}}{\sqrt{2m\omega\hbar^2}} \tag{7.27}$$

$$\hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} - \frac{i\hat{p}}{\sqrt{2m\omega\hbar}} \tag{7.28}$$

$$\Rightarrow\quad \hat{H} = \hbar\omega\left[\hat{a}^\dagger\hat{a} + \frac{1}{2}\right] = \frac{\hbar\omega}{2}\left[\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}\right], \tag{7.29}$$

for which in the Heisenberg picutre, we indeed get analogous equations of motion for these operators to these A objects in classical mechanics:

$$\frac{d}{dt}\hat{a}(t) = -i\omega\hat{a}(t) \tag{7.30}$$

$$\frac{d}{dt}\hat{a}^\dagger(t) = i\omega\hat{a}^\dagger(t). \tag{7.31}$$

These operators then act on the eigenstates as:

$$\hat{a}\left|n\right\rangle = \sqrt{n}\left|n-1\right\rangle, \quad \hat{a}^\dagger\left|n\right\rangle = \sqrt{n+1}\left|n+1\right\rangle \tag{7.32}$$

$$\hat{H}\left|n\right\rangle = \hbar\omega\left(n+\frac{1}{2}\right)\left|n\right\rangle. \tag{7.33}$$

Now in the case of multiple oscillators, we can then instead have:

$$\begin{aligned}\hat{H} &= \sum_k\left(A_k\hat{p}_k^2 + B_k\hat{q}_k^2\right)\\ &= \sum_k\frac{\hbar\omega_k}{2}\left[\hat{a}_k^\dagger\hat{a}_k + \hat{a}_k\hat{a}_k\hat{\dagger}\right]\\ &= \sum_k\hbar\omega_k\left[\hat{a}_k^\dagger\hat{a}_k + \frac{1}{2}\right],\end{aligned} \tag{7.34}$$

where we have the commutation relations:

$$\left[\hat{a}_i, \hat{a}_j^\dagger\right] = \delta_{ij}. \tag{7.35}$$

If the oscillators are uncoupled, we then have a completely separable system in terms of tensor product states so we have:

$$\begin{aligned}&\left|\psi\right\rangle = \left|\{n_k\}\right\rangle\\ \Rightarrow\quad &\hat{H}\left|\{n_k\}\right\rangle = \sum_k\hbar\omega_k\left(n_k+\frac{1}{2}\right)\left|\{n_k\}\right\rangle\\ \Rightarrow\quad &E_n = \hbar\omega_k\left(n_k+\frac{1}{2}\right).\end{aligned} \tag{7.36}$$

Let's go back to classical E&M for awhile. In classical electrodynamics, the extension to a quantum theory of electrodynamics starts by first writing the magnetic vector potential as a sum over Fourier modes:

$$\vec{A}(\vec{x},t) = \sum_{\vec{k}}\sum_{\sigma=1}^{2}\left(\frac{2\pi\hbar c}{V\omega_k}\right)^{1/2}\hat{\epsilon}_{k,\sigma}\left[a_{k,\sigma}(t)e^{i\vec{k}\cdot\vec{x}} + a_{k,\sigma}^*(t)e^{-i\vec{k}\cdot\vec{x}}\right] \tag{7.37}$$

where σ denotes the polarization index, $a_{k,\sigma}(t)$ is the classical Fourier coefficients and we have $\epsilon_{\hat{k,\sigma}}\cdot\vec{k} = 0$. Also we put in a *photon normalization* factor with $\hbar$ in it, which will make the quantization of this be convenient later. Plugging this into the classical wave equation in free-space, we get:

$$\begin{aligned} &\nabla^2 \vec{A} - \frac{1}{c^2}\frac{\partial^2 \vec{A}}{\partial t^2} = 0 \\ \Rightarrow \quad &\sum_{k,\sigma} \left(\frac{2\pi\hbar c}{V\omega_k}\right)^{1/2} \hat{\epsilon}_{k,\sigma} \left[-\omega_k^2 + \frac{\partial^2}{\partial t^2}\right] a_{k,\sigma}(t) = 0 \\ \Rightarrow \quad &\frac{d^2 a_{k,\sigma}(t)}{dt^2} + \omega_k^2 a_{k,\sigma}(t) = 0 \\ \Rightarrow \quad &a_{k,\sigma}(t) = a^{(1)}_{k,\sigma} e^{-i\omega_k t} + a^{(1)}_{k,\sigma} e^{i\omega_k t}. \end{aligned} \tag{7.38}$$

It turns out that one of the terms in the solution can be dismissed because of the construction of the Fourier expansion, which gives us:

$$\boxed{\frac{d}{dt} a_{k,\sigma}(t) = -i\omega_k a_{k,\sigma}(t)}, \tag{7.39}$$

which tells us that we can formulate classical electromagnetism in terms of variables that effectively follow the harmonic oscillator equations of motion! Now considering the energy density:

$$u = \frac{1}{8\pi}\int d^3x \left(E^2 + B^2\right) = \frac{1}{8\pi}\int d^3x \left[\left(\frac{\partial \vec{A}}{\partial t}\right)^2 + \left(\nabla \times \vec{A}\right)^2\right], \tag{7.40}$$

for which plugging in the Fourier series of the vector potential gives:

$$\boxed{u = \frac{1}{2}\sum_{k,\sigma} \hbar\omega_k \left[a^*_{k,\sigma} a_{k,\sigma} + a_{k,\sigma} a^*_{k,\sigma}\right]}. \tag{7.41}$$

So what this says, is that classical electromagnetism in free-space, is a set of uncoupled harmonic oscillators! Now to quantize this, we simply promote these Fourier coefficients to operators. In quantum electrodynamics, the conjugate variables in our Hamiltonian will be $\vec{A}$ and $\vec{E}$:

$$\left[\hat{\vec{A}}_k, \hat{\vec{E}}_{k'}\right] = i\hbar\delta_{k,k'}, \tag{7.42}$$

and we also have these quantized Fourier coefficient operators following:

$$\left[\hat{a}_{k,\sigma}, \hat{a}^\dagger_{k',\sigma'}\right] = \delta_{k,k'}\delta_{\sigma,\sigma'}. \tag{7.43}$$

From this, we then write our Hamiltonian as:

$$\boxed{\hat{H} = \sum_{k,\sigma} \hbar\omega_k \left[\hat{a}^\dagger_{k,\sigma}\hat{a}_{k,\sigma} + \frac{1}{2}\right]}, \tag{7.44}$$

and we have the eigenstates as $|\psi\rangle = |\{n_{k,\sigma}\}\rangle$ where $n_{k,\sigma} = 0, 1, 2, \ldots$, denotes a quantized mode excitation of the field with wave-vector k and polarization σ.

We can now ask, what actually is the quantum mechanical vector potential? This is actually analogous to asking what is $\hat{x}$ in the standard simple harmonic oscillator system since it is the conjugate position in the Hamiltonian. The most general way to see what this is is by writing this conjugate position operator in terms of the creation and annihilation operators ($\hat{x}$ is a linear combination of $\hat{a}$ and $\hat{a}^\dagger$). This then tells us that $\hat{A}$ in fact creates and annihilates photons in a pair-like fashion.

Now let us consider a system in which we have an atom irradiated by some photons. This system would have a Hamiltonian:

$$\begin{aligned} H &= H_{\text{atom}} + H_{\text{rad}} + H_{\text{interaction}} \\ &= \left[\frac{p^2}{2m} + V\right] + \left[\sum_k \hbar\omega_k n_{k,\sigma} + \frac{e^2}{c^2}A^2\right] - \left[\frac{e}{mc}\vec{A}\cdot\vec{p}\right] \end{aligned} \tag{7.45}$$

where we dropped the groundstate energy in the photon excitation Hamiltonian and hats for operators. What we can now do, is consider the interaction term as a perturbation, and drop all constant terms in the Hamiltonian:

$$\begin{aligned} \hat{H} &= \hat{H}_0 + \delta\hat{H} \\ &= \left[\frac{p^2}{2m} + V + \sum_k \hbar\omega_k n_{k,\sigma}\right] - \left[\frac{e}{mc}\vec{A}\cdot\vec{p}\right], \end{aligned} \tag{7.46}$$

for which the zeroth-order Hamiltonian $\hat{H}_0$ will have separable eigenstates $|a\rangle \otimes |n_{k,\sigma}\rangle$. Having these eigenstates, we can then compute transitions in the system with usual quantum perturbation theory by computing the matrix elements:

$$\begin{aligned} \langle f|\,\delta\hat{H}\,|i\rangle &= \langle f|\,\frac{e}{mc}\vec{A}\cdot\vec{p}\,|i\rangle \\ &= -\frac{e}{mc}\sum_{k',\sigma'}\sqrt{\frac{2\pi\hbar c^2}{\omega_k V}}\ \langle n_{k,\sigma}+1|\,a^\dagger_{k',\sigma'}\,|n_{k,\sigma}\rangle\,\langle a'|\,\hat{\epsilon}_{k',\sigma'}\cdot\vec{p}e^{i\vec{k}\,'\cdot\vec{x}}\,|a\rangle \\ &= -\frac{e}{mc}\sqrt{\frac{2\pi\hbar c^2}{\omega_k V}}\ \sqrt{n_{k,\sigma}+1}\ \langle a'|\,\hat{\epsilon}_{k,\sigma}\cdot\vec{p}e^{i\vec{k}\cdot\vec{x}}\,|a\rangle\,. \end{aligned} \tag{7.47}$$

We can then plug this into Fermi's golden rule to get the transition rate:

$$\begin{aligned} d\Gamma &= \frac{2\pi}{\hbar}\left(\frac{2\pi\hbar c^2}{\omega_k V}\right)(n_{k,\sigma}+1)\left|\langle a'|\,\hat{\epsilon}_{k,\sigma}\cdot\hat{\vec{p}}e^{i\vec{k}\cdot\vec{x}}\,|a\rangle\right|^2\delta(E_{a'}+\hbar\omega_k-E_a)\frac{Vd^3k}{(2\pi)^3} \\ \Rightarrow\quad \Gamma_{\text{se}} &= \frac{4}{3}\frac{e^2\omega_{f,i}^2}{2\pi\hbar c^3}\left|\langle a'|\,x\,|a\rangle\right|^2, \end{aligned} \tag{7.48}$$

which is the same result we got from the semi-classical approximation! What is new in the formalism however, is that we can consider the phenomena of stimulated emission and absorption.

§7.4 The End

In this second semester of graduate E&M at Boulder, there has been one unifying thing, that being the physics of radiation. Classically, this really only had one formula:

$$\vec{A}(\vec{x},t) = \frac{4\pi}{c}\int d^3x' dt' \vec{J}(\vec{x}\,',t) D(\vec{x}-\vec{x}\,', t-t'). \tag{7.49}$$

Getting to this of course had a long backstory, such as multipole antennas, dipole scattering, diffraction, all of which had the assumption that $J \sim e^{i\omega t}$ (harmonic). This all had no reference to special relativity, and thus was rather incomplete in terms of its connection to the rest of physics. So indeed, after this came the story about special relativity. This taught us the use of 4-vector, tensors and the rigid structures that bake in special relativity to the theory. The crux of all this was the formalism of Lagrangians for E&M, from which we saw how symmetries constrained Lagrangians. But aside from physics...

The real point of a graduate physics class is often not the physics, but the process to which we learn to function as physicists. We learned many methods of approximations, for which none of that stuff is useful (practical) unless one or two expansion terms are important. This story of approximations is especially true this semester, with things like partial wave expansions. Making approximations is what we do as physicist, and the job is to ask where these approximations come from. This is hard to teach, and perhaps comes only from some sort of experience/intuition. Nonetheless, this is what a class on Jackson may help us to uncover.

"*Learning how to do approximations is what Jackson is all about.*"

Appendices