An Undergraduate View of a Second Year Course in Physics

Josuan Calderon, Sruthi Narayanan, Farid Salazar, Pedro Soto

November 1, 2015
# Contents

**Foreword**

**Preface**

1 Electromagnetism and the Wave Equation for Light

1.1 Integral Form of Maxwell’s Laws .................................................. 9
1.2 Differential Form of Maxwell Equations .......................................... 11
1.3 Solving the Homogeneous Equations ............................................... 14
1.4 Non-uniqueness of $\vec{A}$ .......................................................... 16
1.5 Solving the Inhomogeneous Equations ............................................. 17
1.5.1 Conservation of Charge .......................................................... 19
1.6 Existence of Electromagnetic Waves in Vacuum .................................. 19

2 Coordinate Transformations

2.1 Definitions ................................................................. 23
2.2 Translations .............................................................. 24
  2.2.1 Physical Constants ...................................................... 24
  2.2.2 Scalar Functions ....................................................... 24
  2.2.3 Vector Valued Functions ............................................... 24
  2.2.4 The Gradient .......................................................... 25
  2.2.5 Divergence ............................................................. 26
  2.2.6 Second Order Derivative Operator ..................................... 28
2.3 Rotations ............................................................... 29
  2.3.1 Rotations in two dimensions ............................................ 29
  2.3.2 Sequence of Rotations .................................................. 30
  2.3.3 Physical Constants ...................................................... 31
  2.3.4 Scalar Functions ........................................................ 32
  2.3.5 Vector Valued Functions ............................................... 32
  2.3.6 The Gradient .......................................................... 33
  2.3.7 Divergence ............................................................. 34
  2.3.8 Second Order Derivative Operator ..................................... 35
2.4 Relativistic Invariance of the Wave Equation .................................. 37
### 3 Physical Consequences of Lorentz Transformations

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Effect on Wavelength: Red and Blue Shifts</td>
</tr>
<tr>
<td>3.2</td>
<td>Effect on Wavelength: General Direction of Motion</td>
</tr>
<tr>
<td>3.3</td>
<td>Effect on Distance: Length Contraction</td>
</tr>
<tr>
<td>3.4</td>
<td>Effect on Time: Time Dilation</td>
</tr>
<tr>
<td>3.5</td>
<td>Effect on Movement: Velocity Transformations</td>
</tr>
</tbody>
</table>

### 4 Relativistic Mass and Momentum of a Particle

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Momentum conservation with respect to a static reference frame</td>
</tr>
<tr>
<td>4.2</td>
<td>Momentum conservation with respect to a moving reference frame</td>
</tr>
<tr>
<td>4.3</td>
<td>Mass as a function of speed</td>
</tr>
<tr>
<td>4.4</td>
<td>Energy of a Particle</td>
</tr>
<tr>
<td>4.5</td>
<td>Energy of a particle as a function of momentum and rest energy</td>
</tr>
<tr>
<td>4.6</td>
<td>Energy and momentum transformations</td>
</tr>
<tr>
<td>4.7</td>
<td>Transformation of Forces</td>
</tr>
</tbody>
</table>

### 5 Coulomb’s Law and Special Relativity

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Interactions of moving charges</td>
</tr>
</tbody>
</table>

### 6 Black-Body Radiation - The Need for Quantization

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Energy radiated by a Black-body at a given temperature</td>
</tr>
<tr>
<td>6.2</td>
<td>Classical Approach</td>
</tr>
<tr>
<td>6.3</td>
<td>Quantum Approach</td>
</tr>
</tbody>
</table>

### 7 Setting Up Wave Mechanics

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1</td>
<td>Interference of Light Waves</td>
</tr>
<tr>
<td>7.2</td>
<td>Wave Mechanics - A Motivation</td>
</tr>
</tbody>
</table>

### 8 Quantum Mechanics of a Free Particle at Rest

### 9 Attractive Square Well Potential in One Dimension: Bound States

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.1</td>
<td>Classical Approach</td>
</tr>
<tr>
<td>9.2</td>
<td>Wave Mechanical Approach</td>
</tr>
<tr>
<td>9.3</td>
<td>Outside the well: Regions I and III ($</td>
</tr>
<tr>
<td>9.4</td>
<td>Inside the well: Region II ($</td>
</tr>
<tr>
<td>9.5</td>
<td>Matching Conditions and Boundary Conditions</td>
</tr>
<tr>
<td>9.6</td>
<td>Finding $\alpha$ and $\beta$</td>
</tr>
</tbody>
</table>

### 10 Harmonic Oscillator in One Dimension: Bound States

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.1</td>
<td>The Classical Approach</td>
</tr>
<tr>
<td>10.2</td>
<td>The Quantum Mechanical Approach</td>
</tr>
</tbody>
</table>

### 11 Single Delta Function Potential

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.1</td>
<td>Dirac Delta Function</td>
</tr>
<tr>
<td>11.2</td>
<td>Schrödinger Equation of the Attractive Delta Function Potential</td>
</tr>
<tr>
<td>12 Multiple Delta Function Potentials: Using the Concept of Transfer Matrices</td>
<td>123</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>12.1 Many Delta Function Potentials</td>
<td>123</td>
</tr>
<tr>
<td>12.2 Transfer Matrix</td>
<td>125</td>
</tr>
<tr>
<td>12.3 Analysis for various values of ( N )</td>
<td>126</td>
</tr>
<tr>
<td>12.4 Numerical Computations of Eigenvalues and Probability Distributions for large ( N ): Band Structures</td>
<td>128</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>13 Solid State Physics Using the Delta Function Potential</th>
<th>134</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.1 Energy levels for a free particle in a one dimensional periodic box</td>
<td>134</td>
</tr>
<tr>
<td>13.2 Derivation of Bloch’s Theorem</td>
<td>135</td>
</tr>
<tr>
<td>13.3 Bound and unbound energy levels in the presence of a periodic delta function potential</td>
<td>137</td>
</tr>
<tr>
<td>13.4 Numerical Calculations of Energy Eigenvalues for Periodic Delta Potentials</td>
<td>139</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>14 Semiconductors: A Problem in Tunable Band Gaps</th>
<th>143</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.1 Model Definitions</td>
<td>143</td>
</tr>
<tr>
<td>14.2 Outline of method</td>
<td>144</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>15 Quantum Mechanics in Two Dimensions</th>
<th>147</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.1 Lagrangian Mechanics</td>
<td>147</td>
</tr>
<tr>
<td>15.2 The Schrödinger Equation in Two Dimensions</td>
<td>148</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>16 Quantum Mechanics in Three Dimensions</th>
<th>153</th>
</tr>
</thead>
<tbody>
<tr>
<td>16.1 Schrödinger Equation in Spherical Coordinates</td>
<td>153</td>
</tr>
<tr>
<td>16.2 Angular Momentum Operator</td>
<td>159</td>
</tr>
<tr>
<td>16.3 Algebra of Angular Momentum Operator</td>
<td>160</td>
</tr>
<tr>
<td>16.4 Generalization of the Angular Momentum Operator</td>
<td>162</td>
</tr>
<tr>
<td>16.5 Representation of the Algebra for ( s = 1 )</td>
<td>165</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>17 Statistical Mechanics</th>
<th>167</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.1 Particle Exchange</td>
<td>167</td>
</tr>
<tr>
<td>17.2 Statistical Mechanics: Partition Function</td>
<td>170</td>
</tr>
<tr>
<td>17.3 Expectation Values</td>
<td>172</td>
</tr>
<tr>
<td>17.4 Free Fermions in a Periodic Box</td>
<td>173</td>
</tr>
<tr>
<td>17.4.1 Free Fermions in a Three Dimensional Periodic Box</td>
<td>173</td>
</tr>
<tr>
<td>17.4.2 Free fermions in Two Dimensional Periodic Box</td>
<td>179</td>
</tr>
<tr>
<td>17.4.3 Free fermions in One Dimensional Periodic Box</td>
<td>182</td>
</tr>
<tr>
<td>17.5 Free Bosons in a Periodic Box</td>
<td>184</td>
</tr>
<tr>
<td>17.5.1 Free Bosons in a Three Dimensional Periodic Box</td>
<td>185</td>
</tr>
<tr>
<td>17.5.2 Free Bosons in a Two Dimensional Periodic Box</td>
<td>191</td>
</tr>
<tr>
<td>17.5.3 Free Bosons in a One Dimensional Box</td>
<td>192</td>
</tr>
</tbody>
</table>
Foreword

This book, written solely by four undergraduate students in physics, is an outcome of a one year course I taught for second year undergraduates in physics from 2005 to 2008. The students for my course at that time came with a background in basic physics (typically from a text book like Sears and Zemansky's University Physics) and basic calculus. The aim of the second year course was to provide a semi-rigorous foundation of relativity, quantum mechanics, solid state physics and statistical mechanics. I did not find one text-book that fit the syllabus and I wanted to experiment teaching a course where the students had to study the material from several books. Some of the books I had recommended were:

1. The Feynman Lectures on Physics by Richard P. Feynman, Robert B. Leighton and Matthew Sands.

2. Electricity and Magnetism by Edward M. Purcell.


I prepared lecture notes and the logical flow of my lectures is shown the accompanying fchart. The students were encouraged to prepare detailed lecture notes on their own based on my lectures and by referring to the above books and others. The idea was to get students used to a teaching style where there is no required text book. Students were graded based on their note book and one or two oral examinations where the student had to explain a concept or two with the aid of her/his note book. Like all experiments in teaching style, this has its advantages and disadvantages. One thing was clear and this came as no surprise it worked well for a motivated student.

One of the authors of this book, Farid, approached me in Spring 2013 with the idea of working on an Independent Study Course with me for one year. During the one year period, he wanted to work through the contents of my second year course for undergraduates. Two other students, Josuan and Pedro were interested in joining him and the three of them met with me on a weekly
basis for one year and we went through the contents of the entire course. This book is their joint
version of the note book for this course. The fourth author, Sruthi, did not participate in the
weekly meetings. Instead, she remotely participated by reworking the notes and adding some of
her own. I did not play a role in the style of presentation nor did I edit the contents of this note
book. I hope the release of this note book to undergraduate students and faculty teaching physics
undergraduates will serve two purposes: present a view of basic physics from the view point of
motivated undergraduates; provide a medium for discussion of basic physics among beginning un-
dergraduates.

Rajamani Narayanan
Department of Physics
Florida International University
Miami, FL 33199
November 2014.
Preface

The aim of this book is to introduce sophomore undergraduate students in Physics to two important branches: Special Theory of Relativity and Quantum Mechanics by providing a smooth transition between electromagnetism and these topics. The reader is assumed to understand basics of general physics, calculus, linear algebra and some differential equations.

The first section starts with Electromagnetism and the derivation of the wave operator by expressing Maxwell’s equations on their differential form and solving them employing the gauge fixing condition. It also shows the existence of homogeneous solutions to the wave equation which imply that there exist electromagnetic waves on vacuum. The second section discusses coordinate transformations and the role they play to relate physical observations performed by two different observers on different frames. The familiar examples of translations and rotations are used to illustrate that physics should be the same regardless of position and orientation; particularly, the wave operator is put into perspective. Then we derive Lorentz Boosts as the transformation of space and time that preserves the physics of Electromagnetism. The interpretation and geometry of this transformation is briefly explained.

The third section explores the implications of this transformation such as red and blue shifts, length contraction, time dilation and velocity transformations. The fourth section addresses the results of conservation of momentum within the framework of Lorentz Transformations which lead to mass dependence on velocity, and the proper transformation of momentum, energy and forces. It also discusses Newton’s Laws. The fifth proves that the magnetic force is a consequence of Lorentz Transformations applied to the electric force by using the correct form of Coulomb’s Law. This last section in Special Relativity shows the consistency, symmetry and power of Special Theory of Relativity.

Section six discusses the need for quantization of light based on the analysis of the radiation emitted by a black body, it also introduces Planck’s constant as the width partition of light quanta. Section seven develops wave mechanics in order to describe and to explain experiments such as electron diffraction and interference. It also introduces the operators associated with energy and momentum, and naturally leads to the Schrödinger equation. Moreover, it talks about the interpretation of the solutions to this equation (wavefunctions) and its relationship to probabilities. The next several sections explore different examples of Schrödinger equations. Section eight solves the problem of a free particle, and it discusses the conditions under which wave mechanics is relevant. Sections nine solves the case of a particle in a finite square well potential, and concludes that the presence of a potential constraints the possible energy states of a particle; thereby, bringing up the concept of quantization. Section ten solves the harmonic oscillator problem and it proposes that it
is a good model for light.

Sections eleven to fourteen study models of solids starting at the single delta potential and the existence of a bound state. Then it generalizes the ideas to an $N$ number of delta potentials using the Transfer Matrix, and for the first time energy band gaps arise. Furthermore, Bloch’s theorem for periodic solids is proven and applied to periodic delta potentials. Also, a simple model of a semiconductor with a tunable parameter is analyzed.

Section fifteen and sixteen address wave mechanics in two and three dimensions. It emphasizes the importance of symmetry and their relation to eigenvalues of certain special operators. It also talks about degeneracy and commutators, and the algebraic aspects of wave mechanics operators to solve and simplify problems.

The last section discusses new ideas of the wave mechanics of several particle. First, it discusses particle exchange and classify particle accordingly their symmetric or antisymmetric exchange as bosons and fermions. Then, it introduces the basic tools of Thermodynamics: the Partition Function and its relation to expectation values. Finally, the core of this section is to study the physics of a free gas of fermions and Bose Einstein condensate in three, two and one dimension.
Chapter 1

Electromagnetism and the Wave Equation for Light

1.1 Integral Form of Maxwell’s Laws

A first year course in Electricity and Magnetism ends with the familiar integral forms of Maxwell’s equations.

1. We know that the Electric Field at a certain position due to a charged particle is defined as the electric force due to that particle on a positive one coulomb charge at that position. Any closed surface around charges will have electric field lines going in and out of it (the electric flux)\(^\text{\[1\]}\) Gauss’s Law of Electric Fields, the statement that the total electric flux through a closed surface is proportional to the net charge enclosed by the surface, is

\[ \oint_S \vec{E} \cdot d\vec{A} = \frac{Q}{\varepsilon_0}. \]  

(1.1)

2. We can consider a source of a magnetic field such as a current loop. We know that moving charges are sources of magnetic field. In addition we can think of having magnetic monopoles as sources for magnetic field. Magnetic monopoles are not known to exist although there are theories as to their proposed existence. Thus, when talking about the magnetic flux through a surface enclosing the source, all field lines that go out of the surface will come back into the surface resulting in zero net magnetic flux. We then obtain Gauss’s Law for Magnetic Fields

\[ \oint_S \vec{B} \cdot d\vec{A} = 0. \]  

(1.2)

3. Suppose we have a closed loop with a changing magnetic field through the loop. This changing magnetic field will result in a changing magnetic flux. In order to oppose this changing magnetic flux, there will be an induced electromotive force (EMF) around the loop. Thus,

\(^1\)The integrals with a loop are closed integrals.
Figure 1.1: Capacitor in a closed circuit

we obtain Faraday’s Law of Induction

\[
\oint \vec{E} \cdot d\vec{l} = -\frac{d}{dt} \iint_S \vec{B} \cdot d\vec{A}. \tag{1.3}
\]

4. We know that moving charges result in a magnetic field. A current is defined as moving charges and thus results in a magnetic field. Similar to Gauss’s law of electricity we have a relation between the magnetic field on a loop and the current that pierces the loop. That is the simpler version of Ampere’s Law. In addition, we require a correction known as displacement current. This is similar to Faraday’s Law as it comes from the fact that a changing electric flux results in an induced magnetic field. Putting these two together, Ampere’s Law with Maxwell’s Displacement Current is stated as

\[
\oint_C \vec{B} \cdot d\vec{l} = \mu_0 \left( I_c + \epsilon_0 \frac{d}{dt} \iint_S \vec{E} \cdot d\vec{A} \right). \tag{1.4}
\]

**Why Displacement Current?** The notion of displacement current is not necessarily intuitive so let us understand why it is needed. Consider a capacitor in a closed circuit, and choose a loop around a piece of wire as shown in Fig. 1.1. According to Ampere’s Law without the displacement current (through surface 1)

\[
\oint_C \vec{B} \cdot d\vec{l} = \mu_0 I_c \tag{1.5}
\]

\(^2I_c\) is called conducting current (it represents a physical flux of charge), and \(\epsilon_0 \frac{d}{dt} \iint_S \vec{E} \cdot d\vec{A}\) is known as Maxwell’s displacement current.
and, (through surface 2)

\[
\oint_C \vec{B} \cdot d\vec{l} = 0.
\] (1.6)

Hence, there is a disagreement in the choice of surface. This disagreement must be reconciled by an added contribution to the magnetic flux through surface 2 that will make it match that through surface 1. This can be fixed if we realize that there is a changing electric field between the capacitor, and that this is related to the current through the wire

\[
I_c = \frac{dq}{dt} = C \frac{dV}{dt}
\] (1.7)

where \( C \) is the capacitance. The capacitance for such an arrangement is given by

\[
C = \frac{\epsilon_0 A}{d}
\] (1.8)

where \( A \) is the area of the capacitor plate and \( d \) is the distance between the plates. Also, since the Electric Field is nearly constant we can state simply

\[
V = E \cdot d.
\] (1.9)

Then, combining 1.8, 1.9 and 1.7 we obtain

\[
I_c = \epsilon_0 \frac{d(EA)}{dt} = \epsilon_0 \frac{d}{dt} \oint_S \vec{E} \cdot d\vec{A},
\] (1.10)

Maxwell’s displacement current.

### 1.2 Differential Form of Maxwell Equations

We now have a set of integral equations, which typically are extremely hard to solve analytically. Instead, a system of partial differential equations would be easier to handle. So let us change each equation to its differential form to obtain a system of PDEs.

1. Gauss’s Law of Electric Fields

\[
\nabla \cdot \vec{E} = \frac{\rho}{\epsilon_0},
\] (1.11)

2. Gauss’s Law of Magnetic Fields

\[
\nabla \cdot \vec{B} = 0.
\] (1.12)

3. Faraday’s Law of Induction

\[
\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}.
\] (1.13)

4. Ampere’s Law

\[
\frac{1}{\epsilon_0 \mu_0} \nabla \times \vec{B} = \frac{j}{\epsilon_0} + \frac{\partial \vec{E}}{\partial t}.
\] (1.14)

\[3\]The 0 on the right hand side is due to the absence of magnetic monopoles in nature. This type of equation is called homogeneous.
Below, we rigorously derive each of these equations from the integral forms we stated before.

**Deriving the System of PDEs:**

(1) Consider an infinitesimal cube of dimensions $dx, dy, dz$. Given that the charge density is $\rho(x, y, z)$ we can say that

$$Q_{encl} = \int \int \int \rho(x, y, z) dx dy dz$$

and since the dimensions of the space are infinitesimal this Taylor Expands to $\rho(x, y, z) dx dy dz$. Therefore, the right hand side of Gauss’s Law of Electric Fields is

$$\frac{\rho(x, y, z) dx dy dz}{\epsilon_0}.$$  

For the other side of the equation, first consider the bottom and top surfaces of the cube. The component of the Electric Field that contributes to the left hand side integral is just the $z$ component so,

$$E_z(x, y, z + \frac{dz}{2}) dx dy - E_z(x, y, z - \frac{dz}{2}) dx dy = \frac{E_z(x, y, z + \frac{dz}{2}) dx dy dz - E_z(x, y, z - \frac{dz}{2}) dx dy dz}{dz}$$

which by the definition of the derivative is just

$$\frac{\partial E_z}{\partial z} dx dy dz.$$

Similarly, the same logic can be used for the other two directions. Thus, the left hand side becomes

$$\left(\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z}\right) dx dy dz$$

and therefore

$$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = \frac{\rho(x, y, z)}{\epsilon_0} \vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0}.$$  \hspace{1cm} (1.15)

(2) In the case of the magnetic field, the right hand side of the expression is clearly 0. The left hand side again uses the infinitesimal cube. Consider the top and bottom components again. Only the $z$ component of the magnetic field is considered so

$$B_z(x, y, z + \frac{dz}{2}) dx dy - B_z(x, y, z - \frac{dz}{2}) dx dy = \frac{B_z(x, y, z + \frac{dz}{2}) dx dy dz - B_z(x, y, z - \frac{dz}{2}) dx dy dz}{dz}.$$  

By the definition of the derivative this is just

$$\frac{\partial B_z}{\partial z} dx dy dz.$$

Similarly, this can be done to the other two component directions so we can conclude

$$\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} = 0$$

\footnote{Although it is common to see these relations easily derived using the Stokes and Divergence Theorems, we choose not to make the assumptions necessary in order to use them. Instead we consider $\mathbb{R}^3$ space and assume solely the existence of the first (and in certain cases the second) derivatives.}
\( \nabla \cdot \vec{B} = 0. \)  \hspace{1cm} (1.16)

(3) Now we consider a closed square since we want the integral over a closed loop. Consider first the closed square in the \( x-y \) plane. Suppose the current is in the \( \hat{z} \) direction and current density is \( \vec{j} = \langle j_x, j_y, j_z \rangle \). Then we can write

\[
\int_C \vec{B} \cdot d\vec{l} = B_y(x + \frac{dx}{2}, y) dy - B_z(x, y + \frac{dy}{2}) dx - B_y(x - \frac{dx}{2}, y) dy + B_z(x, y - \frac{dy}{2}) dx.
\]

By Taylor Expansion this is equal to

\[
\frac{\partial B_y}{\partial x} dx dy - \frac{\partial B_z}{\partial y} dx dy.
\]

Similarly this can be done for the \( x-z \) plane to obtain

\[
\int_C \vec{B} \cdot d\vec{l} = -B_z(x + \frac{dx}{2}, z) dz - B_x(x, z - \frac{dz}{2}) dx + B_z(x, z + \frac{dz}{2}) dz + B_x(x, z) dx.
\]

Similarly this can be done for the \( y-z \) plane to obtain

\[
\int_C \vec{B} \cdot d\vec{l} = -B_z(y - \frac{dy}{2}, z) dz + B_y(y, z - \frac{dz}{2}) dy + B_z(y, z + \frac{dz}{2}) dz - B_y(y, z) dy.
\]

Now consider the right hand side of the equation for all three planes.

- For the \( x-y \) plane:
  \[ \mu_0 j_z dx dy + \epsilon_0 \mu_0 \frac{\partial E_z}{\partial t} dx dy \]

- For the \( x-z \) plane:
  \[ \mu_0 j_y dx dz + \epsilon_0 \mu_0 \frac{\partial E_y}{\partial t} dx dz \]

- For the \( y-z \) plane:
  \[ \mu_0 j_x dy dz + \epsilon_0 \mu_0 \frac{\partial E_x}{\partial t} dy dz \]

Equating each left hand side component with each right hand side equation we obtain

\[
\begin{align*}
\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y} &= \mu_0 j_z + \epsilon_0 \mu_0 \frac{\partial E_z}{\partial t} \\
\frac{\partial B_z}{\partial x} - \frac{\partial B_x}{\partial z} &= \mu_0 j_y + \epsilon_0 \mu_0 \frac{\partial E_y}{\partial t} \\
\frac{\partial B_z}{\partial y} - \frac{\partial B_y}{\partial z} &= \mu_0 j_x + \epsilon_0 \mu_0 \frac{\partial E_x}{\partial t} .
\end{align*}
\]

(1.19)

Recall that

\[
\nabla \times \vec{B} = (\frac{\partial B_z}{\partial y} - \frac{\partial B_y}{\partial z}), (\frac{\partial B_x}{\partial z} - \frac{\partial B_z}{\partial x}), (\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y})
\]

so we can conclude

\[
\frac{1}{\epsilon_0 \mu_0} \nabla \times \vec{B} = \frac{1}{\epsilon_0} \vec{j} + \frac{\partial \vec{E}}{\partial t} .
\]

(1.20)

(4) In a similar fashion, expanding the left hand side of Ampere’s Law results in
• For the $x$-$y$ plane: \[
\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}
\]
• For the $x$-$z$ plane: \[
\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}
\]
• For the $y$-$z$ plane: \[
\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}
\]

Expanding the right hand side results in

• For the $x$-$y$ plane: \[-\frac{\partial B_z}{\partial t}\]
• For the $x$-$z$ plane: \[-\frac{\partial B_y}{\partial t}\]
• For the $y$-$z$ plane: \[-\frac{\partial B_x}{\partial t}\]

Thus we can conclude that
\[
\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}.
\] (1.21)

1.3 Solving the Homogeneous Equations

Let us return to Maxwell’s Equations in their differential form. Observe that the first two equations are homogeneous. We can find a solution to these homogeneous equations by introducing $\vec{A}$, the magnetic potential\footnote{It might be arbitrary to call it Magnetic Potential at this point. This will make more sense after taking a Mathematical Physics course; however, it is important to mention that $\vec{A}$ is related to energy stored in the field.}
\[
\vec{B} = \nabla \times \vec{A}.
\] (1.22)

In order to prove why this is in fact a solution for Gauss’s Law of Magnetic Fields, we will need a rather simple vector calculus identity
\[
\nabla \cdot (\nabla \times \vec{A}) = 0.
\] (1.23)

Mathematical Identity 1: Let $\vec{A}$ be any vector valued function, then $\nabla \cdot (\nabla \times \vec{A}) = 0$. 
Proof. Let $\vec{A}$ be any continuous vector valued function in $\mathbb{R}^3$. By the definition of curl,
\[ \vec{\nabla} \times \vec{A} = \left( \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \hat{i} + \left( \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \hat{j} + \left( \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \hat{k}. \] (1.24)

Taking the gradient,
\[ \vec{\nabla} \cdot (\vec{\nabla} \times \vec{A}) = \frac{\partial^2 A_z}{\partial x \partial y} - \frac{\partial^2 A_y}{\partial x \partial z} + \frac{\partial^2 A_x}{\partial y \partial z} - \frac{\partial^2 A_z}{\partial y \partial x} + \frac{\partial^2 A_y}{\partial z \partial x} - \frac{\partial^2 A_x}{\partial z \partial y}. \] (1.25)

The mixed partials here are all equal thus, we obtain \[ \text{(1.23)} \]

For $\vec{B}$ as defined in \[ \text{(1.22)} \] and Mathematical Identity 1, it is clear that $\vec{\nabla} \cdot \vec{B} = 0$. Thus, Gauss’s Law of Magnetic Fields is satisfied.

Next, we will find a solution for $\vec{E}$ using Faraday’s Law. Substituting \[ \text{(1.22)} \] into the right hand side of Faraday’s Law, we obtain
\[ \vec{\nabla} \times \vec{E} = -\frac{\partial}{\partial t} \left( \vec{\nabla} \times \vec{A} \right). \] (1.26)

Since the $\vec{\nabla}$ operator is independent of time we are allowed to interchange the position of the curl and the derivative with respect to time to obtain
\[ \vec{\nabla} \times \vec{E} = -\vec{\nabla} \times \frac{\partial \vec{A}}{\partial t}, \] (1.27)

which by simple algebra and an application of the distributive law can be written as,
\[ \vec{\nabla} \times \left( \vec{E} + \frac{\partial \vec{A}}{\partial t} \right) = 0. \] (1.28)

In order to solve this further we will need another vector calculus identity
\[ \vec{\nabla} \times (\vec{\nabla} \phi) = 0. \] (1.29)

**Mathematical Identity 2:** Let $\phi$ be any scalar function, then $\vec{\nabla} \times (\vec{\nabla} \phi) = 0$.

Proof. The gradient of $\phi$ is defined as
\[ \vec{\nabla} \phi = \frac{\partial \phi}{\partial x} \hat{i} + \frac{\partial \phi}{\partial y} \hat{j} + \frac{\partial \phi}{\partial z} \hat{k}. \] (1.30)

Taking the Curl,
\[ \vec{\nabla} \times (\vec{\nabla} \phi) = \left( \frac{\partial^2 \phi}{\partial y \partial z} - \frac{\partial^2 \phi}{\partial z \partial y} \right) \hat{i} + \left( \frac{\partial^2 \phi}{\partial z \partial x} - \frac{\partial^2 \phi}{\partial x \partial z} \right) \hat{j} + \left( \frac{\partial^2 \phi}{\partial x \partial y} - \frac{\partial^2 \phi}{\partial y \partial x} \right) \hat{k}. \] (1.31)

Since continuity of the function is assumed, the mixed partials of (1.31) are identical, and each term on the right hand side is equal to 0, (1.29) easily follows. \qed
Using Mathematical Identity 2, we obtain
\[ \vec{E} + \frac{\partial \vec{A}}{\partial t} = -\vec{\nabla} \phi \]  
(1.32)
where \( \phi \) is known as the Electrostatic Potential\(^6\).

Then,
\[ \vec{E} = -\vec{\nabla} \phi - \frac{\partial \vec{A}}{\partial t}. \]  
(1.33)

Let us briefly illustrate why \( \phi \) is called the Electrostatic Potential. Consider \( \vec{A} = 0 \), then \( \vec{B} = 0 \) and \( \vec{E} = -\vec{\nabla} \phi \). This is the familiar electrostatic case where \( \phi \) is more familiarly known as the potential difference. If \( \vec{A} \neq 0 \), then \(- \frac{\partial \vec{A}}{\partial t}\) resembles the induction term that gives rise to a non-conservative potential. Both \( \phi \) and \( \vec{A} \) together define \( \vec{E} \) as in (1.33) and \( \vec{B} \) as in (1.22) to form the so-called Electromagnetic Potential.

### 1.4 Non-uniqueness of \( \vec{A} \)

As previously shown, a pair \((\phi, \vec{A})\) defines a pair \((\vec{E}, \vec{B})\). In fact, due to the homogeneity, there are infinitely many pairs \((\phi, \vec{A})\) that define the same pair \((\vec{E}, \vec{B})\). Therefore, if you are given a pair \((\vec{E}, \vec{B})\), you cannot uniquely find a pair \((\phi, \vec{A})\). In order to show that this is true, we will suppose that we have a pair \((\phi, \vec{A})\) which defines some pair \((\vec{E}, \vec{B})\). We can then apply a gauge transformation to \((\phi, \vec{A})\) using some arbitrary function \(\psi\) as follows

\[ \phi' = \phi - \frac{\partial \psi}{\partial t} \]  
(1.34)
\[ \vec{A}' = \vec{A} + \vec{\nabla} \psi. \]  
(1.35)

Then, \( \vec{B} \) as in (1.22) becomes
\[ \vec{B}' = \vec{\nabla} \times (\vec{A} + \vec{\nabla} \phi) \]  
(1.36)
which simplifies to
\[ \vec{B}' = \vec{\nabla} \times \vec{A} + \vec{\nabla} \times \vec{\nabla} \phi. \]  
(1.37)

By Mathematical Identity 2 our original equation simplifies to
\[ \vec{B}' = \vec{B}. \]  
(1.38)

This shows us that the pairs \((\phi', \vec{A}')\) and \((\phi, \vec{A})\) both define the same \( \vec{B} \). Similarly \( \vec{E} \) as in (1.33) becomes
\[ \vec{E}' = -\vec{\nabla} \left( \phi - \frac{\partial \psi}{\partial t} \right) - \frac{\partial}{\partial t} \left( \vec{A} + \vec{\nabla} \psi \right). \]  
(1.39)

---

\(^6\)The gradient could have been positive or negative; however, it is more convenient to choose the negative sign.
Applying the commutative and distributive laws for the vector operators we obtain

\[ \vec{E}' = -\nabla \phi + \nabla \left( \frac{\partial \psi}{\partial t} \right) - \frac{\partial \vec{A}}{\partial t} - \nabla \left( \frac{\partial \psi}{\partial t} \right). \]  

(1.40)

Finally, we observe that our definition of \( \vec{E} \) is also invariant under this transformation

\[ \vec{E}' = \vec{E}. \]  

(1.41)

It is important to point out that our choice of \( \psi \) was arbitrary and so the pair \((\phi, \vec{A})\) is not unique for a given \((\vec{E}, \vec{B})\). We can now consider what we will call the gauge fixing condition

\[ \frac{1}{\epsilon_0 \mu_0} \nabla \cdot \vec{A} + \frac{\partial \phi}{\partial t}. \]  

(1.42)

This condition is not invariant under the transformations above and so this can be used to find a unique solution. The easiest case of the above equation to consider is

\[ \frac{1}{\epsilon_0 \mu_0} \nabla \cdot \vec{A} + \frac{\partial \phi}{\partial t} = 0. \]  

(1.43)

It can be argued that even though our choice of \( \psi \) is arbitrary, the condition \(1.43\) may not necessarily hold. Let’s show that it actually can be found by applying the previous transformation,

\[ \frac{1}{\epsilon_0 \mu_0} \nabla \cdot (\vec{A} + \nabla \psi) + \frac{\partial}{\partial t} \left( \phi - \frac{\partial \psi}{\partial t} \right) = 0 \]  

(1.44)

which can be re-written as

\[ \epsilon_0 \mu_0 \frac{\partial^2 \psi}{\partial t^2} - (\nabla \cdot \nabla) \psi = \nabla \cdot \vec{A} + \epsilon_0 \mu_0 \frac{\partial \phi}{\partial t}. \]  

(1.45)

This is a non-homogeneous wave equation that can be solved for analytic functions \( \vec{A} \) and \( \phi \), which is usually the case for an electromagnetic potential.

### 1.5 Solving the Inhomogeneous Equations

Now let us solve Gauss’s Law of Electric Fields. Combining \(1.33\) and \(1.11\) we obtain

\[ \nabla \cdot \left( -\nabla \phi - \frac{\partial \vec{A}}{\partial t} \right) = \frac{\rho}{\epsilon_0}. \]  

(1.46)

Expanding and combining terms this becomes

\[ - (\nabla \cdot \nabla) \phi - \frac{\partial}{\partial t} (\nabla \cdot \vec{A}) = \frac{\rho}{\epsilon_0}. \]  

(1.47)

Now let us use Ampere’s Law. Combining \(1.22\), \(1.33\) and \(1.13\) we obtain

\[ \frac{1}{\epsilon_0 \mu_0} \nabla \times (\nabla \times \vec{A}) = \frac{\vec{j}}{\epsilon_0} + \frac{\partial}{\partial t} \left( -\nabla \phi - \frac{\partial \vec{A}}{\partial t} \right). \]  

(1.48)

In order to proceed we need one more vector calculus identity

\[ \nabla \times (\nabla \times \vec{A}) = \nabla (\nabla \cdot \vec{A}) - (\nabla \cdot \nabla) \vec{A}. \]  

(1.49)
Mathematical Identity 3: Let \( \vec{A} \) be any vector valued function, then \( \vec{\nabla} \times (\vec{\nabla} \times \vec{A}) = \vec{\nabla}(\vec{\nabla} \cdot \vec{A}) - (\vec{\nabla} \cdot \vec{\nabla}) \vec{A} \).

Proof. Taking the curl of (1.24),
\[
\vec{\nabla} \times (\vec{\nabla} \times \vec{A}) = \left( \frac{\partial^2 A_y}{\partial y \partial x} - \frac{\partial^2 A_x}{\partial y^2} + \frac{\partial^2 A_z}{\partial z \partial x} \right) \hat{i} + \left( \frac{\partial^2 A_z}{\partial z \partial y} - \frac{\partial^2 A_y}{\partial z^2} + \frac{\partial^2 A_x}{\partial x \partial y} \right) \hat{j} + \left( \frac{\partial^2 A_x}{\partial x \partial z} - \frac{\partial^2 A_y}{\partial y \partial z} + \frac{\partial^2 A_z}{\partial z^2} \right) \hat{k} \tag{1.50}
\]
The divergence of \( \vec{A} \) is
\[
\vec{\nabla} \cdot \vec{A} = \frac{\partial A_z}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}. \tag{1.51}
\]
Taking the gradient of (1.51),
\[
\vec{\nabla} (\vec{\nabla} \cdot \vec{A}) = \left( \frac{\partial^2 A_x}{\partial x^2} + \frac{\partial^2 A_y}{\partial y^2} + \frac{\partial^2 A_z}{\partial z^2} \right) \hat{i} + \left( \frac{\partial^2 A_z}{\partial z \partial x} + \frac{\partial^2 A_y}{\partial z \partial y} + \frac{\partial^2 A_x}{\partial z \partial z} \right) \hat{j} + \left( \frac{\partial^2 A_x}{\partial x \partial z} + \frac{\partial^2 A_y}{\partial y \partial z} + \frac{\partial^2 A_z}{\partial z^2} \right) \hat{k} \tag{1.52}
\]
Then, subtracting (1.50) from (1.51),
\[
\vec{\nabla} (\vec{\nabla} \cdot \vec{A}) - \vec{\nabla} \times (\vec{\nabla} \times \vec{A}) = \left( \frac{\partial A_z}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \right) \hat{i} + \left( \frac{\partial A_z}{\partial z \partial x} + \frac{\partial A_y}{\partial z \partial y} + \frac{\partial A_x}{\partial z \partial z} \right) \hat{j} + \left( \frac{\partial A_x}{\partial x \partial z} + \frac{\partial A_y}{\partial y \partial z} + \frac{\partial A_z}{\partial z^2} \right) \hat{k} \tag{1.53}
\]
This can be written as
\[
\vec{\nabla} (\vec{\nabla} \cdot \vec{A}) - \vec{\nabla} \times (\vec{\nabla} \times \vec{A}) = (\vec{\nabla} \cdot \vec{\nabla}) \vec{A}. \tag{1.54}
\]
Finally, rearranging terms, we obtain (1.49).

Continuing with Ampere’s Law, we can apply Mathematical Identity 3 and use the distributive law to obtain
\[
\frac{1}{\epsilon_0 \mu_0} (\vec{\nabla} (\vec{\nabla} \cdot \vec{A}) - (\vec{\nabla} \cdot \vec{\nabla}) \vec{A}) = \frac{\vec{j}}{\epsilon_0} - \vec{\nabla} \left( \frac{\partial \phi}{\partial t} \right) - \frac{\partial^2 \vec{A}}{\partial t^2}. \tag{1.55}
\]
Then, rearranging the terms and factoring out the gradient operator
\[
\frac{\partial^2 \vec{A}}{\partial t^2} - \frac{1}{\epsilon_0 \mu_0} (\vec{\nabla} \cdot \vec{\nabla}) \vec{A} = \frac{\vec{j}}{\epsilon_0} - \vec{\nabla} \left( \frac{1}{\epsilon_0 \mu_0} \vec{\nabla} \cdot \vec{A} + \frac{\partial \phi}{\partial t} \right). \tag{1.56}
\]
This is where the Lorentz Gauge choice (our specific choice of the gauge fixing condition) comes in handy since it lets the second term on the right hand side of (1.56) reduce to 0 and
\[
\epsilon_0 \mu_0 \frac{\partial^2 \vec{A}}{\partial t^2} - (\vec{\nabla} \cdot \vec{\nabla}) \vec{A} = \mu_0 \vec{j}. \tag{1.57}
\]
Additionally this choice lets (1.47) reduce to
\[
\epsilon_0 \mu_0 \frac{\partial^2 \phi}{\partial t^2} - (\vec{\nabla} \cdot \vec{\nabla}) \phi = \frac{\rho}{\epsilon_0}. \tag{1.58}
\]
Finally, both (1.57) and (1.58) reduce to non-homogeneous, independent wave equations.

Setting \( \epsilon_0 \mu_0 = \frac{1}{c^2} \), both equations can be rewritten in short hand notation as

\[
\Box \left[ \frac{\phi}{c^2} \right] = \frac{1}{\epsilon_0} \left[ \frac{\rho}{c^2} \right]
\]

where, \( \Box = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla \cdot \nabla \) and is known as the Wave or D'Alembert Operator.

The reason we multiply \( \vec{A} \) by \( c \) is so that the arguments of the column vectors of (1.59) are dimensionally equivalent; in other words, they have the same units. This shows the symmetry between \( \phi \) and \( c\vec{A} \) and between \( \rho \) and \( \vec{j}c \) respectively.

### 1.5.1 Conservation of Charge

In this small subsection, we will show that conservation of charge is a direct consequence of the symmetry of Maxwell's Equations. Let us take the divergence of Ampere's Law (1.14) and apply commutativity for vector operators to obtain

\[
\frac{1}{\epsilon_0 \mu_0} \nabla \cdot (\nabla \times \vec{B}) = \frac{1}{\epsilon_0} \nabla \cdot \vec{j} + \frac{\partial}{\partial t} (\nabla \cdot \vec{E}).
\]

We observe that the left hand side reduces to 0 by Mathematical Identity 1 and that the term inside the time derivative can be replaced by \( \frac{\vec{E}}{\epsilon_0} \) in accordance with Gauss's Law (1.11). Hence

\[
\nabla \cdot \vec{j} + \frac{\partial \rho}{\partial t} = 0.
\]

This result can be understood as follows

- If \( \nabla \cdot \vec{j} > 0 \), then \( \frac{\partial \rho}{\partial t} < 0 \). An outward flux of current density at a point implies a proportional decrease of the current density at the same point.
- If \( \nabla \cdot \vec{j} < 0 \), then \( \frac{\partial \rho}{\partial t} > 0 \). An inward flux of current density at a point implies a proportional increase of the current density at the same point.

Indeed, this is conservation of charge: Charge is neither created nor destroyed This is illustrated in Fig. 1.2.

### 1.6 Existence of Electromagnetic Waves in Vacuum

Let us consider the special case where \( \rho = 0 \) and \( \vec{j} = 0 \). This is the situation where we are in a vacuum. We can show the existence of a non-trivial solution for the following case

\[
\Box \left[ \frac{\phi}{c^2} \right] = 0
\]

which is the homogeneous wave equation.
The figure on the left represents an outward flux of current and a decrease in charge density, while the Fig. on the right represents an inward flux of current and an increase in charge density.

For simplicity, let $\phi = 0$ and let $\vec{A}$ have components in the $\hat{y}, \hat{z}$ directions. Then (1.62) reduces to

$$1 \frac{\partial^2 A_y}{c^2 \partial t^2} - \frac{\partial^2 A_y}{\partial x^2} = 0$$
$$1 \frac{\partial^2 A_z}{c^2 \partial t^2} - \frac{\partial^2 A_z}{\partial x^2} = 0. \quad (1.63)$$

These equations have the following solution\(^8\)

$$A_y(x, t) = A_0 \cos(kx - \omega t)$$
$$A_z(x, t) = A_0 \sin(kx - \omega t). \quad (1.64)$$

We can show that (1.64) is actually a solution for (1.63). Taking the first and second partial derivatives with respect to $t$ and $x$

$$\frac{\partial A_y}{\partial t} = A_0 \omega \sin(kx - \omega t), \quad \frac{\partial^2 A_y}{\partial t^2} = -A_0 \omega^2 \cos(kx - \omega t) \quad (1.65)$$
$$\frac{\partial A_y}{\partial x} = -A_0 k \sin(kx - \omega t), \quad \frac{\partial^2 A_y}{\partial x^2} = -A_0 k^2 \cos(kx - \omega t). \quad (1.66)$$

(Similarly done for $A_z$.) Substituting the second derivatives into the homogeneous wave equation given in (1.63),

$$\left(\frac{\omega^2}{c^2} - k^2\right) \cos(kx - \omega t) = 0, \quad \forall x, t. \quad (1.67)$$

---

\(^7\)It is possible to take $\phi = 0$, and still have both an electric and a magnetic field. This is not the case when $A = 0$.

\(^8\)Any function of the form $f(x - ct)$ can solve the homogeneous wave equation, this is known as D’Alembert Method.
Therefore,
\[ \omega^2 = c^2 k^2 \]  \hspace{2cm} (1.68)

or equivalently
\[ |\omega| = ck. \]  \hspace{2cm} (1.69)

Now, let us see what the corresponding electric and magnetic fields are. From (1.33) the electric field is given by
\[
\vec{E} = -\frac{\partial}{\partial t} (A_0 \cos(kx - \omega t) \hat{y} + A_0 \sin(kx - \omega t) \hat{z}) \\
= -\omega A_0 \sin(kx - \omega t) \hat{y} + \omega A_0 \cos(kx - \omega t) \hat{z}. \]  \hspace{2cm} (1.70)

Similarly, the magnetic field as in (1.22) is given by
\[
\vec{B} = \nabla \times (A_0 \cos(kx - \omega t) \hat{y} + A_0 \sin(kx - \omega t) \hat{z}). \]  \hspace{2cm} (1.71)

Evaluating the curl we obtain
\[
\vec{B} = \frac{\partial}{\partial x} (A_0 \cos(kx - \omega t)) \hat{z} - \frac{\partial}{\partial x} (A_0 \sin(kx - \omega t)) \hat{y} \\
= -k A_0 \sin(kx - \omega t) \hat{y} - k A_0 \cos(kx - \omega t) \hat{z}. \]  \hspace{2cm} (1.72)

Furthermore, recognizing the relationship between \( \omega \) and \( k \), and setting \( E_0 = -\omega A_0 \), (1.70) and (1.72) can be rewritten as
\[
\vec{E} = E_0 (\sin(k(x - ct)) \hat{y} - \cos(k(x - ct)) \hat{z}) \]  \hspace{2cm} (1.73)
\[
\vec{B} = \frac{E_0}{c} (\sin(k(x - ct)) \hat{z} + \cos(k(x - ct)) \hat{y}) \]  \hspace{2cm} (1.74)

where \( E_0 \) and \( \frac{E_0}{c} \) are the amplitudes of the electric field and magnetic fields respectively. Interestingly enough this shows that electric and magnetic fields can exist in vacuum. Additionally both the electric and magnetic fields are circularly polarized which means that their amplitudes do not change with time. By changing the sign of one of the terms, we can also switch between being right circularly polarized and left circularly polarized, both are viable options. This will be useful later on when we introduce the relation between the energy and the magnitude of a wave. Also, both are in phase with each other, and their magnitudes are proportional
\[ ||E|| = c||B||. \]  \hspace{2cm} (1.75)

The fact that both equations have the term \((x - ct)\) implies that this combination of electric and magnetic field is a disturbance that is moving in space with velocity \( c \). This disturbance is called an electromagnetic wave: light! A plane electromagnetic wave is usually described by the single equation
\[ \sin(k(x - ct)) \]  \hspace{2cm} (1.76)

or more generally as
\[ \sin(\vec{k} \cdot \vec{r} - \omega t). \]  \hspace{2cm} (1.77)

This is a plane wave propagating in the \( \vec{k} \) direction, where \( \vec{r} \) is the vector position of a point in the wave. It is easy to show that this is a solution to the wave equation in three dimensions if
\[ ||\vec{k}|| = \frac{\omega}{c}. \]  \hspace{2cm} (1.78)
Unlike mechanical waves, it has been shown that light does not need a medium to propagate. It is the changing electric and magnetic fields that sustain each other. Moreover, the speed of the wave, \( c \), is a universal constant. Thus, light moves with the same speed for any observer in any inertial frame of reference. This is one of the most remarkable consequences of Maxwell’s Equations as well as one of the postulates of Einstein’s theory of Special Relativity.
Chapter 2

Coordinate Transformations

We are familiar with electric charges and currents as well as electric and magnetic potentials. We can imagine that we measure charge and currents and then wish to Fig. out the corresponding potentials or vice-a-versa. Furthermore, we can think of the situation where two people measure the same charge and current and find the corresponding potentials but, they use two different coordinate systems or labels to make their measurements. Suppose they want to know if they are consistent with one another. In order to do this, they would need to relate their two coordinate systems somehow and show consistency in their measurements. This is where coordinate transformations would be handy. In this section we will also show why it is interesting to show invariance of certain quantities under these coordinate transformations. Common examples of coordinate transformations are: rectangular-to-polar, spherical-to-cylindrical, translations, rotations, one frame moving with respect to another, etc.

2.1 Definitions

Before going into the actual transformations, we should first state some important definitions.

- **Physical constants** are the universal constants we often consider such as \( \mu_0 \) and \( \epsilon_0 \). Additionally, the Gravitational constant \( (G) \) and Coulomb’s constant \( (k) \) are examples. These are quantities that are invariant with respect to any frame of reference; in other words, they never change under any circumstances.

Now let us briefly talk about functions. In this discussion we will put into context the idea of the scalar and vector potentials that we mentioned earlier when talking about the wave equation. It is useful to keep in mind that in general the solutions for the two potentials are found given both a charge and current density.

- **Scalar Functions** are functions that map a point in an \( n \) dimensional space (for instance, Euclidean Space \( \mathbb{R}^3 \)) to a value in \( \mathbb{R} \). The scalar potential \( \phi \) is an example of a scalar function. In particular we have that under transformations from an unprimed system to a primed system
  \[
  \phi(x, y, z, t) = \phi'(x', y', z', t').
  \]

- **Vector-Valued Functions** are functions that map a point in an \( n \) dimensional space to a vector in an \( m \) dimensional space. Examples of this are the magnetic potential \( \vec{A} \) as well as
the electric and magnetic fields: $\vec{E}$ and $\vec{B}$. In this case, under transformations, we should have the relations

$$A(x, y, z, t) = A'(x', y', z', t'), \quad E(x, y, z, t) = E'(x', y', z', t'), \quad B(x, y, z, t) = B'(x', y', z', t').$$

We will first discuss some basic transformations like translations and rotations, under which all the above relations should hold true.

### 2.2 Translations

In this section we will discuss one type of transformation: translations, which are **shifts of one or more of the coordinates**, and their consequences on physical constants, scalar variables and vector variables.

#### 2.2.1 Physical Constants

If we consider what happens when we translate our reference frame, the one thing that is guaranteed not to change is a physical constant.

#### 2.2.2 Scalar Functions

What about the value of a function? If we choose the same coordinates and plug their values into two different functions, we would expect to get two different values. However, it is important to point out that the value of two functions describing some property at a physical point should be the same, even though they might have different coordinates (labels) for the same physical point. Therefore,

$$f(x, y, z, t) = f'(x', y', z', t') \quad (2.1)$$

where $f(x, y, z, t)$ and $f'(x', y', z', t')$ represent scalar functions in two different frames. As shown in Fig. 2.1, the relationship between the coordinates in each frame can be given by

$$
x' = x + x_0 \\
y' = y \\
z' = z \\
t' = t.
$$

Combining (2.2) with (2.1), we obtain

$$f(x, y, z, t) = f'(x + x_0, y, z, t). \quad (2.3)$$

This is the proper relationship of how scalar functions change under a translation.

#### 2.2.3 Vector Valued Functions

Now, let us explain how vector valued functions change under translations. By following the previous arguments that the values at a physical point must be the same, we should have

$$F(x, y, z, t) = F'(x', y', z', t') \quad (2.4)$$

24
The point $P$ can be described by different coordinates according to each reference frame. The picture is only in two dimensions but most generally a point can be described by coordinates $(x, y, z, t)$.

where $\vec{F}(x, y, z, t)$, $\vec{F}'(x', y', z', t')$ are vector valued functions in two different reference frames evaluated at the same physical point. This can be written in component notation as

$$f(x, y, z, t)\hat{i} + g(x, y, z, t)\hat{j} = f'(x', y', z', t')\hat{i}' + g'(x', y', z', t')\hat{j}' \quad (2.5)$$

where $f(x, y, z, t)$ and $g(x, y, z, t)$, and $f'(x', y', z', t')$ and $g'(x', y', z', t')$ are scalar functions that represent the coordinates of the vector valued functions $\vec{F}(x, y, z, t)$ and $\vec{F}'(x', y', z', t')$ respectively. From Fig. 2.1, it is easy to see that the basis vectors $(\hat{i}, \hat{j})$ and $(\hat{i}', \hat{j}')$ are the same. Hence (2.5) can be simplified to

$$f(x, y, z, t) = f'(x', y', z', t')$$
$$g(x, y, z, t) = g'(x', y', z', t') \quad (2.6)$$

Combining (2.2) with (2.6) we obtain

$$f(x, y, z, t) = f'(x + x_0, y, z, t)$$
$$g(x, y, z, t) = g'(x + x_0, y, z, t) \quad (2.7)$$

The above equations show how to relate the components of a vector valued function under a translation.

### 2.2.4 The Gradient

We will now discuss how the gradient of a scalar function changes under translations. In order to do this we have to consider the partial derivatives of a scalar function under a translation. By
definition this is given by
\[
\frac{\partial f'}{\partial x'}(x', y', z', t') = \lim_{\Delta x' \to 0} \frac{f'(x' + \Delta x', y', z', t') - f'(x', y', z', t')}{\Delta x'}.
\] (2.8)

It is clear that \(\Delta x = \Delta x'\), then (2.8) is
\[
\frac{\partial f'}{\partial x'}(x', y', z', t') = \lim_{\Delta x \to 0} \frac{f'(x + \Delta x, y, z, t) - f'(x, y, z, t)}{\Delta x}.
\] (2.9)

Next, using the translation given by (2.9) we obtain
\[
\frac{\partial f'}{\partial x'}(x', y', z', t') = \lim_{\Delta x \to 0} \frac{f'(x + x_0 + \Delta x, y, z, t) - f'(x + x_0, y, z, t)}{\Delta x}.
\] (2.10)

Applying the relations derived in (2.7) into the previous equation,
\[
\frac{\partial f'}{\partial x'}(x', y', z', t') = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y, z, t) - f(x, y, z, t)}{\Delta x}.
\] (2.11)

Finally, by definition of the partial derivative, the right hand side becomes
\[
\frac{\partial f'}{\partial x'}(x', y', z', t') = \frac{\partial f}{\partial x}(x, y, z, t).
\] (2.12)

The last equation states that the partial derivatives of the scalar functions at the same physical point are identical. It is useful to note that alternatively we could have arrived at this result using the chain rule. It is then fairly easy to show that the gradient is related as follows,
\[
\vec{\nabla} f(x, y, z, t) = \vec{\nabla}' f'(x', y', z', t')
\] (2.13)

where \(\vec{\nabla} = \frac{\partial}{\partial x}\hat{i} + \frac{\partial}{\partial y}\hat{j} + \frac{\partial}{\partial z}\hat{k}\). This follows directly from the above relations, so we choose not to write out this reasoning.

### 2.2.5 Divergence

Another interesting quantity to look at under translations is the divergence of a vector valued function. In order to do that we will consider how the partial derivatives of a vector valued function change under a translation where \(\vec{F}\) in \(\mathbb{R}^3\) is given by
\[
\vec{F} = f(x, y, z, t)\hat{i} + g(x, y, z, t)\hat{j}
\] (2.14)

whose partial derivative with respect to \(x\) is
\[
\frac{\partial \vec{F}}{\partial x}(x, y, z, t) = \frac{\partial f}{\partial x}(x, y, z, t)\hat{i} + \frac{\partial g}{\partial x}(x, y, z, t)\hat{j}.
\] (2.15)

In the translated frame, the partial derivative is
\[
\frac{\partial \vec{F'}}{\partial x'}(x', y', z', t') = \frac{\partial f'}{\partial x'}(x', y', z', t')\hat{i} + \frac{\partial g'}{\partial x'}(x', y', z', t')\hat{j}.
\] (2.16)
Since the basis vectors are the same, using a generalization of (2.12), we obtain

\[
\frac{\partial \mathbf{F}'(x', y', z', t')}{\partial x'}(x', y', z', t') = \frac{\partial f}{\partial x}(x, y, z, t)\hat{i} + \frac{\partial g}{\partial x}(x, y, z, t)\hat{j}
\]  

(2.17)

which, using (2.15), can be written as,

\[
\frac{\partial \mathbf{F}'(x', y', z', t')}{\partial x'}(x', y', z', t') = \frac{\partial \mathbf{F}}{\partial x}(x, y, z, t).
\]  

(2.18)

This implies that the partial derivatives of a vector valued functions evaluated at the same physical point are identical under translations.

**Useful Partial Derivative Relations:** We will start by defining a vector valued function \( \vec{A}(x, y, z, t) = A_x(x, y, z, t)\hat{i} + A_y(x, y, z, t)\hat{j} + A_z(x, y, z, t)\hat{k} \). Then using the transformations given by the rotation matrices we obtain the following nine relations

\[
\begin{align*}
\frac{\partial}{\partial x} A_x(x, y, z, t) &= \frac{\partial}{\partial x'} A_x(x, y, z, t) \\
\frac{\partial}{\partial y} A_x(x, y, z, t) &= \frac{\partial}{\partial y'} A_x(x, y, z, t) \\
\frac{\partial}{\partial z} A_x(x, y, z, t) &= \frac{\partial}{\partial z'} A_x(x, y, z, t) \\
\frac{\partial}{\partial x} A_y(x, y, z, t) &= \frac{\partial}{\partial x'} A_y(x, y, z, t) \\
\frac{\partial}{\partial y} A_y(x, y, z, t) &= \frac{\partial}{\partial y'} A_y(x, y, z, t) \\
\frac{\partial}{\partial z} A_y(x, y, z, t) &= \frac{\partial}{\partial z'} A_y(x, y, z, t) \\
\frac{\partial}{\partial x} A_z(x, y, z, t) &= \frac{\partial}{\partial x'} A_z(x, y, z, t) \\
\frac{\partial}{\partial y} A_z(x, y, z, t) &= \frac{\partial}{\partial y'} A_z(x, y, z, t) \\
\frac{\partial}{\partial z} A_z(x, y, z, t) &= \frac{\partial}{\partial z'} A_z(x, y, z, t).
\end{align*}
\]

(2.19)

Let us now look at the divergence of \( \vec{A} \), a vector valued function. We know the divergence is defined as

\[
\nabla \cdot \vec{A}(x, y, z, t) = \frac{\partial A_x}{\partial x}(x, y, z, t) + \frac{\partial A_y}{\partial y}(x, y, z, t) + \frac{\partial A_z}{\partial z}(x, y, z, t).
\]

By the relations we derived for the partial derivatives of vector valued functions under translations we can easily conclude

\[
\nabla \cdot \vec{A} = \nabla' \cdot \vec{A}'.
\]

(2.20)
2.2.6 Second Order Derivative Operator

Finally, it is useful to look at how second order derivatives change under translations. We can do this using what we have already found about first order derivatives. If we consider \( \frac{\partial^2}{\partial x'^2} \) then using (2.3) we can write a scalar function \( \phi(x, y, z, t) \) as

\[
\frac{\partial^2}{\partial x'^2} \phi'(x', y', z', t') = \frac{\partial^2}{\partial x'^2} \phi(x, y, z, t).
\] (2.21)

We can re-write this by separating the two derivatives and using the fact that for translations \( \frac{\partial x}{\partial x'} = 1 \)

\[
\frac{\partial^2}{\partial x'^2} \phi(x, y, z, t) = \frac{\partial}{\partial x'} \left[ \frac{\partial x}{\partial x'} \frac{\partial}{\partial x} \phi(x, y, z, t) \right] = \frac{\partial}{\partial x'} \left[ \frac{\partial}{\partial x} \phi(x, y, z, t) \right] = \frac{\partial^2}{\partial x'^2} \phi(x, y, z, t).
\]

This is true for the partial derivatives with respect to \( y, z, t \) as well, since they transform similarly. Thus we know how translations affect the second order derivative operator in terms of scalar functions. For vector-valued functions we have

\[
\frac{\partial^2}{\partial x'^2} \vec{A}'(x', y', z', t') = \frac{\partial^2}{\partial x'^2} \vec{A}(x, y, z, t).
\]

Using the same fact that for transformations we have \( \frac{\partial x}{\partial x'} = 1 \) we can come to the conclusion

\[
\frac{\partial^2}{\partial x'^2} \vec{A}(x, y, z, t) = \frac{\partial^2}{\partial x'^2} \vec{A}(x, y, z, t).
\] (2.22)

Again, this is true for the partial derivatives with respect to \( y, z, t \) as well since they transform similarly.

**Putting it into perspective: Translations.** We have considered the general idea of translations but now we should understand how that relates to the wave equation we obtained earlier. We know that given a charge and current density we should be able to find a vector and scalar potential as we previously outlined. The important thing to understand is that the wave operator is actually invariant under translations. It is quite simple to show this. Recall

\[
\left( \frac{1}{c^2} \frac{\partial^2}{\partial t'^2} - \vec{\nabla}' \cdot \vec{\nabla}' \right) \left[ \phi c \vec{A} \right] = \frac{1}{\epsilon_0} \left[ \frac{\rho}{c} \right].
\]

In the primed system this becomes

\[
\left( \frac{1}{c^2} \frac{\partial^2}{\partial t'^2} - \vec{\nabla}' \cdot \vec{\nabla}' \right) \left[ \phi' c \vec{A}' \right] = \frac{1}{\epsilon_0} \left[ \frac{\rho'}{c} \right].
\]

Then using the relations we obtained about translations

\[
\left( \frac{1}{c^2} \frac{\partial^2}{\partial t'^2} - \vec{\nabla}' \cdot \vec{\nabla}' \right) \left[ \phi c \vec{A} \right] = \frac{1}{\epsilon_0} \left[ \frac{\rho}{c} \right].
\]

We showed above that translations do not affect any of the second derivatives, thus we conclude that the wave operator is invariant under translations. 

\[^1\]Time does not change under translations, i.e. \( t' = t \).
2.3 Rotations

Another important set of transformations is the set of rotations. Rotations describe a change in coordinates with respect to a fixed point or axis, where a plane is rotated. It is usually more intuitive to explain what a rotation is by using polar coordinates. It immediately follows that a rotation in two dimensions only changes the angular component, but not the radial one. This is an important property of rotations, the distance between two points does not change. Rotations are very important in physics because objects subject to measurements can be observed at an angle, and it is sometimes desirable to apply a rotation and have the object at a certain axis. In fact, we will use rotations in a later section.

2.3.1 Rotations in two dimensions

First, let’s focus on rotations in two dimensions and Fig. out how to relate a frame \((x, y, z, t)\) to a frame \((x', y', z', t')\) rotated counterclockwise by an angle \(\theta\) as in Fig. 2.2. If the basis in the unprimed frame is given by \((\hat{i}, \hat{j}, \hat{k})\), it is easy to see that the basis in the new frame is \((\hat{i}', \hat{j}', \hat{k}')\) which can be expanded to

\[
\hat{i}' = \hat{i} \cos \theta + \hat{j} \sin \theta \\
\hat{j}' = -\hat{i} \sin \theta + \hat{j} \cos \theta \\
\hat{k}' = \hat{k}.
\]  

(2.23)

Point \(P\) can now be described either by \(x \hat{i} + y \hat{j} + z \hat{k}\) or \(x' \hat{i}' + y' \hat{j}' + z' \hat{k}'\). Then it follows that

\[
x \hat{i} + y \hat{j} = x' \hat{i}' + y' \hat{j}'.
\]

(2.24)

After some simple vector algebra we obtain

\[
x \hat{i} + y \hat{j} = x' (\hat{i} \cos \theta + \hat{j} \sin \theta) + y' (-\hat{i} \sin \theta + \hat{j} \cos \theta).
\]

(2.25)
Rearranging the terms on the right hand side,
\[ x\hat{i} + y\hat{j} = (x' \cos \theta - y' \sin \theta)\hat{i} + (x' \sin \theta + y' \cos \theta)\hat{j}. \tag{2.26} \]

Finally the transformation is given by
\[
\begin{align*}
x &= x' \cos \theta - y' \sin \theta \\
y &= x' \sin \theta + y' \cos \theta \\
z &= z' \\
t &= t'.
\end{align*}
\tag{2.27}
\]

The rotation, which only combines \(x\) and \(y\), can be written in matrix notation as
\[
\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}. \tag{2.28}
\]

If we wish to find \((x', y')\) in terms of \((x, y)\), we can take the inverse of the matrix and we obtain
\[
\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \tag{2.29}
\]

This rotation matrix can be denoted by
\[
R(\theta) = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}. \tag{2.30}
\]

Then (2.29) can be written as
\[
\begin{bmatrix} x' \\ y' \end{bmatrix} = R(\theta) \begin{bmatrix} x \\ y \end{bmatrix}. \tag{2.31}
\]

We have found a way to relate the rotated coordinates \((x', y')\) to the coordinates \((x, y)\) through the use of a rotation matrix \(R(\theta)\).

### 2.3.2 Sequence of Rotations

The next step is to show what happens when we perform a sequence of rotations. We will demonstrate what happens by performing two subsequent rotations and calculating the result. Let \((x'', y'')\) be the coordinates of a point in a frame (double primed frame) that has been rotated counterclockwise by an angle \(\phi\) with respect to the \(x'y'\)-frame (primed frame). The primed frame has been rotated counterclockwise by an angle \(\theta\) with respect to the unprimed frame. Then,
\[
\begin{bmatrix} x'' \\ y'' \end{bmatrix} = \begin{bmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}, \tag{2.32}
\]

and as previously done
\[
\begin{bmatrix} x'' \\ y'' \end{bmatrix} = \begin{bmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \tag{2.33}
\]
Intuitively, if we perform two rotations first by an angle $\theta$ and then by $\phi$, we should obtain a rotation matrix equivalent to a single rotation matrix by an angle $\theta + \phi$. Let us now verify that this is true,

$$R(\theta)R(\phi) = \begin{bmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} = \begin{bmatrix} \cos \phi \cos \theta - \sin \phi \sin \theta & \cos \phi \sin \theta + \sin \phi \cos \theta \\ -\sin \phi \cos \theta - \cos \phi \sin \theta & -\sin \phi \sin \theta + \cos \phi \cos \theta \end{bmatrix}. \quad (2.34)$$

Using the trigonometric identities for the sums and differences of angles this matrix transforms to

$$R(\theta)R(\phi) = \begin{bmatrix} \cos(\theta + \phi) & \sin(\theta + \phi) \\ -\sin(\theta + \phi) & \cos(\theta + \phi) \end{bmatrix} = R(\theta + \phi). \quad (2.35)$$

This is clearly what we expected, a sequence of two rotations is still a rotation\(^2\)

$$R(\phi)R(\theta) = R(\phi + \theta). \quad (2.36)$$

By induction, it is easy to show that a sequence of $n$ rotation matrices is still a rotation matrix.

---

**Mathematical Identity 4**: The composition of $n$ rotation matrices is equal to one rotation matrix for the sum of all the angles i.e. $R(\theta_1)R(\theta_2)\ldots R(\theta_n) = R(\theta_1 + \theta_2 + \ldots + \theta_n)$.

**Proof.** We will prove this by induction. Suppose that $R(\theta_1)R(\theta_2)\ldots R(\theta_n) = R(\theta_1 + \theta_2 + \ldots + \theta_n)$ and for simplicity that $\theta_1 + \theta_2 + \ldots + \theta_n = \phi_n$. Then

$$R(\theta_1)R(\theta_2)\ldots R(\theta_n)R(\theta_{n+1}) = R(\phi_n)R(\theta_{n+1}).$$

But, we already know that for two angles $\theta$ and $\phi$ that $R(\theta)R(\phi) = R(\theta + \phi)$ so we have that

$$R(\theta_1)R(\theta_2)\ldots R(\theta_n)R(\theta_{n+1}) = R(\phi_n + \theta_{n+1}) = R(\theta_1 + \theta_2 + \ldots + \theta_n + \theta_{n+1}).$$

Since we know that $R(\theta_1)R(\theta_2) = R(\theta_1 + \theta_2)$, by induction our claim is also true. \qed

---

### 2.3.3 Physical Constants

As we previously discussed for translations we will discuss how physical constants, scalar functions and vector valued functions change under rotations. By definition, physical constants must remain the same under rotations.

\(^2\)Rotation matrices form a group under composition since it is closed under this operation and each rotation matrix possesses an inverse.
2.3.4 Scalar Functions

Scalar functions have the property that their value at a particular physical point in space should be the same no matter what the coordinates. In other words, the physics must be the same under rotations. Then,

$$\phi(x, y, z, t) = \phi'(x', y', z', t').$$ (2.37)

Combining (2.29) and (2.37),

$$\phi(x, y, z, t) = \phi'(x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta, z, t)$$ (2.38)

where \(\phi\) represents a scalar function which is a mapping \((x, y, z, t) \rightarrow \mathbb{R}\). This is the relationship between two scalar functions under a counterclockwise rotation by an angle \(\theta\).

2.3.5 Vector Valued Functions

Vector valued functions will also change under rotation. In fact, the correct statement of how vector valued functions change involves slightly more work than the one for scalar functions since not only the coordinates change but the basis vectors do as well. Similar to scalar functions, vector valued functions have the property

$$\vec{F}(x, y, z, t) = \vec{F}'(x', y', z', t')$$ (2.39)

or more explicitly,

$$f(x, y, z, t)\hat{i} + g(x, y, z, t)\hat{j} = f'(x', y', z', t')\hat{i}' + g'(x', y', z', t')\hat{j}'$$ (2.40)

where \(f(x, y, z, t), g(x, y, z, t)\) and \(f'(x', y', z', t'), g'(x', y', z', t')\) are scalar functions that represent the components of the vector valued functions \(\vec{F}(x, y, z, t)\) and \(\vec{F}'(x', y', z', t')\) respectively.

The change in basis was determined earlier and the change in the scalar function components is given by (2.38). Applying these to the right hand side of (2.40), we obtain

$$f(x, y, z, t)\hat{i} + g(x, y, z, t)\hat{j} = f'(x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta, z, t)(\hat{i} \cos \theta + \hat{j} \sin \theta)$$

$$+ \ g'(x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta, z, t)(-\hat{i} \sin \theta + \hat{j} \cos \theta)$$ (2.41)

Rearranging terms, we obtain,

$$f(x, y, z, t)\hat{i} + g(x, y, z, t)\hat{j} = \begin{bmatrix} f'(x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta, z, t) \cos \theta \\ - g'(x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta, z, t) \sin \theta \end{bmatrix} \hat{i}$$

$$+ \begin{bmatrix} g'(x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta, z, t) \cos \theta \\ f'(x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta, z, t) \sin \theta \end{bmatrix} \hat{j}$$ (2.42)

which can be written in vector notation as

$$\begin{bmatrix} f(x, y, z, t) \\ g(x, y, z, t) \end{bmatrix} = R(\theta) \begin{bmatrix} f'(x', y', z', t') \\ g'(x', y', z', t') \end{bmatrix}.$$ (2.43)

This shows how the components of a vector valued function from two different frames (under rotation) are related.
2.3.6 The Gradient

It is also important to know how the gradient for scalar functions like \( \phi(x, y, z, t) \) changes under rotations. We will need partial derivatives but unlike the previous section, we will not use the limit definitions of partial derivatives since the computations will get extremely tedious and long. Instead, we will use the chain rule which should be familiar from multivariable calculus. If \( \phi(x, y, z, t) \) and \( \phi'(x', y', z', t') \) are equal and continuous in all space then the partial derivatives with respect to the same variable have to also be the same (2.37). For instance, the partial derivative with respect to \( x' \) is

\[
\frac{\partial \phi}{\partial x'} = \frac{\partial \phi'}{\partial x'}.
\] (2.44)

Applying the chain rule to the right hand side we conclude

\[
\frac{\partial \phi}{\partial x} = \frac{\partial \phi}{\partial x'} \frac{\partial x'}{\partial x} + \frac{\partial \phi}{\partial y'} \frac{\partial y'}{\partial x} + \frac{\partial \phi}{\partial z'} \frac{\partial z'}{\partial x} + \frac{\partial \phi}{\partial t'} \frac{\partial t'}{\partial x}.
\] (2.45)

Taking into consideration the relations from (2.29) we obtain the following results

\[
\frac{\partial x'}{\partial x} = \cos \theta, \quad \frac{\partial y'}{\partial x} = -\sin \theta, \quad \frac{\partial z'}{\partial x} = 0, \quad \frac{\partial t'}{\partial x} = 0.
\] (2.46)

Then, substituting these partials into (2.45) by the previous equations,

\[
\frac{\partial \phi}{\partial x} = \frac{\partial \phi}{\partial x'} \cos \theta - \frac{\partial \phi}{\partial y'} \sin \theta.
\] (2.47)

Next, using a generalization of (2.44),

\[
\frac{\partial \phi}{\partial x} = \frac{\partial \phi'}{\partial x'} \cos \theta - \frac{\partial \phi'}{\partial y'} \sin \theta.
\] (2.48)

When the same procedure is applied for the partial derivative with respect to \( y \) we obtain

\[
\frac{\partial \phi}{\partial y} = \frac{\partial \phi}{\partial x'} \frac{\partial x'}{\partial y} + \frac{\partial \phi}{\partial y'} \frac{\partial y'}{\partial y} + \frac{\partial \phi}{\partial z'} \frac{\partial z'}{\partial y} + \frac{\partial \phi}{\partial t'} \frac{\partial t'}{\partial y}
\] (2.49)

where

\[
\frac{\partial x'}{\partial y} = \sin \theta, \quad \frac{\partial y'}{\partial y} = \cos \theta, \quad \frac{\partial z'}{\partial y} = 0, \quad \frac{\partial t'}{\partial y} = 0.
\] (2.50)

It follows that

\[
\frac{\partial \phi}{\partial y} = \frac{\partial \phi'}{\partial y'} \sin \theta + \frac{\partial \phi'}{\partial y'} \cos \theta.
\] (2.51)

The partial derivatives with respect to \( t, z \) are obviously equal under this transformation. Finally from (2.48) and (2.51), the relationship between the partial derivatives can be written in matrix notation as

\[
\begin{bmatrix}
\frac{\partial \phi}{\partial x} \\
\frac{\partial \phi}{\partial y}
\end{bmatrix} = \begin{bmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{bmatrix}
\begin{bmatrix}
\frac{\partial \phi'}{\partial x'} \\
\frac{\partial \phi'}{\partial y'}
\end{bmatrix}.
\] (2.52)

Furthermore, taking the inverse matrix, we obtain

\[
\begin{bmatrix}
\frac{\partial \phi'}{\partial x'} \\
\frac{\partial \phi'}{\partial y'}
\end{bmatrix} = \begin{bmatrix}
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta
\end{bmatrix}
\begin{bmatrix}
\frac{\partial \phi}{\partial x} \\
\frac{\partial \phi}{\partial y}
\end{bmatrix}.
\] (2.53)
which can also be written as
\[
\begin{bmatrix}
\frac{\partial \phi'}{\partial x'} \\
\frac{\partial \phi'}{\partial y'} \\
\end{bmatrix}
= R(\theta)
\begin{bmatrix}
\frac{\partial \phi}{\partial x} \\
\frac{\partial \phi}{\partial y} \\
\end{bmatrix}
\] (2.54)
where \(R(\theta)\) is given by (2.30). Now we can discuss the relationship between the gradient operator in two reference frames. We start with the definition of the gradient of a scalar function in the rotated frame
\[
\vec{\nabla}' \phi' = \frac{\partial \phi'}{\partial x'} \hat{i} + \frac{\partial \phi'}{\partial y'} \hat{j} + \frac{\partial \phi'}{\partial z'} \hat{k}'.
\] (2.55)

Combining the change of basis relations with (2.53) and using the fact that \(\hat{k}' = \hat{k}\) and the partial derivatives with respect to \(z\) and \(z'\) at the same physical point are the same, we obtain
\[
\vec{\nabla}' \phi' = \left[ \frac{\partial \phi}{\partial x} \cos \theta + \frac{\partial \phi}{\partial y} \sin \theta \right] \hat{i} + \left[ -\frac{\partial \phi}{\partial x} \sin \theta + \frac{\partial \phi}{\partial y} \cos \theta \right] \hat{j} + \frac{\partial \phi}{\partial z} \hat{k}.
\] (2.56)

Expanding and rearranging the terms,
\[
\vec{\nabla}' \phi' = \left( \frac{\partial \phi}{\partial x} \cos^2 \theta + \frac{\partial \phi}{\partial y} \sin \theta \cos \theta + \frac{\partial \phi}{\partial z} \sin^2 \theta - \frac{\partial \phi}{\partial y} \sin \theta \cos \theta \right) \hat{i} + \left( \frac{\partial \phi}{\partial x} \cos \theta \sin \theta + \frac{\partial \phi}{\partial y} \sin^2 \theta - \frac{\partial \phi}{\partial x} \sin \theta \cos \theta + \frac{\partial \phi}{\partial y} \cos^2 \theta \right) \hat{j} + \frac{\partial \phi}{\partial z} \hat{k}.
\] (2.57)

Simplifying we have
\[
\vec{\nabla}' \phi' = \frac{\partial \phi}{\partial x} \hat{i} + \frac{\partial \phi}{\partial y} \hat{j} + \frac{\partial \phi}{\partial z} \hat{k}.
\] (2.58)
which by the definition of the gradient of a scalar function in the unrotated frame is
\[
\vec{\nabla}' \phi'(x', y', z') = \vec{\nabla} \phi(x, y, z).
\] (2.59)

If you carefully think about it, this is a not an obvious result, but this is how the gradient changes under rotation. In other words, the gradient behaves like a vector valued function.

### 2.3.7 Divergence

In order to discuss the divergence of a vector valued function and how it behaves under rotations, we will first derive several simple relations for partial derivatives of vector valued function components.

---

**Useful Partial Derivative Relations:** We will start by defining a vector valued function \(\vec{A}(x, y, z, t) = A_x(x, y, z, t) \hat{i} + A_y(x, y, z, t) \hat{j} + A_z(x, y, z, t) \hat{k}\). Then using the transformations given by the rotation matrices we obtain the following nine relations
\[
\begin{align*}
\frac{\partial}{\partial x'} A_x(x, y, z, t) &= \left[ \frac{\partial x}{\partial x'} \frac{\partial}{\partial x} + \frac{\partial y}{\partial x'} \frac{\partial}{\partial y} + \frac{\partial z}{\partial x'} \frac{\partial}{\partial z} + \frac{\partial t}{\partial x'} \frac{\partial}{\partial t} \right] A_x(x, y, z, t) = \left[ \cos \theta \frac{\partial}{\partial x} + \sin \theta \frac{\partial}{\partial y} \right] A_x(x, y, z, t) \\
\frac{\partial}{\partial y'} A_x(x, y, z, t) &= \left[ \frac{\partial x}{\partial y'} \frac{\partial}{\partial x} + \frac{\partial y}{\partial y'} \frac{\partial}{\partial y} + \frac{\partial z}{\partial y'} \frac{\partial}{\partial z} + \frac{\partial t}{\partial y'} \frac{\partial}{\partial t} \right] A_x(x, y, z, t) = \left[ -\sin \theta \frac{\partial}{\partial x} + \cos \theta \frac{\partial}{\partial y} \right] A_x(x, y, z, t)
\end{align*}
\]

\(^3\)You should remember that the partial derivatives must be evaluated at the same physical point, not just at the same coordinates.
Similarly we have rotations. Let

\[ \frac{\partial}{\partial x} A_x(x, y, z, t) = \frac{\partial}{\partial y} A_y(x, y, z, t) = \frac{\partial}{\partial z} A_z(x, y, z, t) = \frac{\partial}{\partial t} A_t(x, y, z, t) \]

\[ \frac{\partial}{\partial x'} A'_x(x', y', z', t') = \frac{\partial}{\partial y'} A'_y(x', y', z', t') = \frac{\partial}{\partial z'} A'_z(x', y', z', t') = \frac{\partial}{\partial t'} A'_t(x', y', z', t') \]

Finally, we will discuss the second order derivatives and how they change under rotations. We

From these relations it is easy to see that the divergence transforms under rotations as follows

\[
\nabla' \cdot \vec{A}(x', y', z', t') = \frac{\partial}{\partial x'} A'_x(x', y', z', t') + \frac{\partial}{\partial y'} A'_y(x', y', z', t') + \frac{\partial}{\partial z'} A'_z(x', y', z', t')
\]

\[
= \cos^2 \theta \frac{\partial A_x}{\partial x} + \cos \theta \sin \theta \frac{\partial A_x}{\partial y} + \sin \theta \cos \theta \frac{\partial A_x}{\partial z} + \sin^2 \theta \frac{\partial A_y}{\partial z} - \cos \theta \sin \theta \frac{\partial A_y}{\partial x} + \cos \theta \sin \theta \frac{\partial A_y}{\partial z} + \frac{\partial A_z}{\partial z}
\]

\[
= \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} = \nabla \cdot \vec{A}(x, y, z, t).
\]

(2.61)

2.3.8 Second Order Derivative Operator

Finally, we will discuss the second order derivatives and how they change under rotations. We

will again first address the second derivatives for scalar functions and how they transform under

rotations. Let \( \phi \) be the scalar function and consider the quantity \( \frac{\partial^2 \phi}{\partial x^2} \). Then

\[
\frac{\partial^2}{\partial x^2} \phi(x', y', z', t') = \frac{\partial^2}{\partial x'^2} \phi(x, y, z, t).
\]

Now, using the transformation equations, we know that \( \frac{\partial x}{\partial x'} = \cos \theta, \ \frac{\partial y}{\partial y'} = \sin \theta, \ \frac{\partial z}{\partial z'} = \frac{\partial t}{\partial t'} = 0 \)

so we can then state

\[
\frac{\partial^2}{\partial x^2} \phi(x, y, z, t) = \frac{\partial}{\partial x'} \left[ \frac{\partial x}{\partial x'} \frac{\partial}{\partial x} + \frac{\partial y}{\partial y'} \frac{\partial}{\partial y} + \frac{\partial z}{\partial z'} \frac{\partial}{\partial z} + \frac{\partial t}{\partial t'} \right] \phi(x, y, z, t)
\]

\[
= \frac{\partial}{\partial x'} \left[ \cos \theta \frac{\partial}{\partial x} + \sin \theta \frac{\partial}{\partial y} \right] \phi(x, y, z, t)
\]

\[
= \left[ \cos^2 \theta \frac{\partial^2}{\partial x^2} + 2 \cos \theta \sin \theta \frac{\partial^2}{\partial x \partial y} + \sin^2 \theta \frac{\partial^2}{\partial y^2} \right] \phi(x, y, z, t).
\]

(2.62)

Similarly we have

\[
\frac{\partial^2}{\partial y^2} \phi(x, y, z, t) = \left[ \cos^2 \theta \frac{\partial^2}{\partial y^2} - 2 \cos \theta \sin \theta \frac{\partial^2}{\partial x \partial y} + \sin^2 \theta \frac{\partial^2}{\partial x^2} \right] \phi(x, y, z, t).
\]

35
From the transformation equations it is evident that $z, t$ are invariant we have that $\frac{\partial^2}{\partial z'^2} = \frac{\partial^2}{\partial z^2}$ and $\frac{\partial^2}{\partial t'^2} = \frac{\partial^2}{\partial t^2}$. We can finally conclude

\[
(\nabla' \cdot \nabla') \phi(x, y, z, t) = \frac{\partial^2}{\partial x'^2} + \frac{\partial^2}{\partial y'^2} + \frac{\partial^2}{\partial z'^2} \phi(x, y, z, t)
\]

(2.63)

**Putting it into perspective: Rotations.** We have considered the general idea of rotations but now we should understand how that relates to the wave equation we obtained earlier. We know that given a charge and current density we should be able to find a vector and scalar potential as we previously outlined. The important thing to understand is that the wave operator is actually invariant under rotations. It is quite simple to show this. Recall

\[
\left( \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla \cdot \nabla \right) \frac{\phi}{c^2} = \frac{1}{\epsilon_0} \left[ \frac{\rho}{c^2} \right].
\]

In the primed system the wave operator becomes

\[
\frac{1}{c^2} \frac{\partial^2}{\partial t'^2} - \nabla' \cdot \nabla'.
\]

We observed that since $t$ doesn’t vary with the rotation, all derivatives of $t$ are invariant. We also showed that the second derivative operator is invariant under the rotation. Thus, we can conclude that this wave operator is invariant under our rotation. Additionally $\phi$ is related to $\phi'$ and $A$ is related to $A'$ by the same rotation matrix as $\rho$ is related to $\rho'$ and $J$ is related to $J'$. Thus, under the transformation this rotation matrix can be pulled out on both sides. It is also important to note that even if the rotation were not only the in the $x$ and $y$ directions as it was here, this invariance would still be the case.

**Note:** We discussed how the divergence transforms for rotations and translations but we did not discuss the curl. The curl can be worked out, but the algebra is quite tedious. The relations required are the same as the relations required for showing divergence.

**Note:** After discussing rotations and translations there is an important result to realize. We have shown that under both rotations and translations the wave operator itself is invariant. This means that in the example where we have two people in two places making measurements and using different coordinate systems they will still be using the same operator!
2.4 Relativistic Invariance of the Wave Equation

We just finished discussing transformations including translations, rotations and even sequences of rotations. It seems natural to question if there are other transformations where the invariance of the wave operator is evident. Previously we looked at rotations which were physically motivated and observed the invariance of the wave operator under such transformations. We could have, in principle, gone backwards and started by wanting the wave operator to be invariant and found the transformation necessary. We would have, obviously, ended up with the answer we had. If we mix up \(x\) and \(t\) it is not so easy to think of what the transformation would be so we will do it backwards this time and ask if such a transformation can exist to make the wave operator invariant. Consider the following coordinate transformation, which we will call the Lorentz Boost, that involves space and time

\[
x = a_{11} x' + a_{14} c t'
\]
\[
ct = a_{41} x' + a_{44} c t'.
\]

The transformation is a linear combination of \(x'\) and \(ct'\). Thus, it maps lines on the \(x'\)-\(ct'\) plane to lines on the \(x\)-\(ct\) plane. The reason \(ct\) is chosen in place of \(t\) is to make both equations dimensionally consistent (have the same units.) Let us now determine which values of \(a_{11}, a_{14}, a_{41}, a_{44}\) we need to make the wave operator invariant.

Using the chain rule and differentiating (2.64) and (2.65), we obtain

\[
\frac{\partial}{\partial x'} = \left( \frac{\partial x}{\partial x'} \right) \frac{\partial}{\partial x} + \left( \frac{\partial c t}{\partial x'} \right) \frac{\partial}{\partial c t} = a_{11} \frac{\partial}{\partial x} + a_{41} \frac{\partial}{\partial c t}. \tag{2.66}
\]

For the partial derivative with respect to \(t\)

\[
\frac{\partial}{\partial c t'} = \left( \frac{\partial x}{\partial c t'} \right) \frac{\partial}{\partial x} + \left( \frac{\partial c t}{\partial c t'} \right) \frac{\partial}{\partial c t} = a_{14} \frac{\partial}{\partial x} + a_{44} \frac{\partial}{\partial c t}. \tag{2.67}
\]

Recall that we needed to show that

\[
-\frac{\partial^2}{\partial x'^2} + \frac{\partial^2}{\partial c t'^2} = -a_{11} \frac{\partial^2}{\partial x^2} - a_{44} \frac{\partial^2}{\partial c t^2} - a_{11} a_{41} \frac{\partial^2}{\partial x \partial c t} + a_{14} \frac{\partial^2}{\partial x^2} + a_{44} \frac{\partial^2}{\partial c t^2} + a_{14} a_{44} \frac{\partial^2}{\partial x \partial c t}. \tag{2.68}
\]

which for simplicity is the wave equation in one spatial dimension.

Using the relations (2.66), (2.67) and (2.68) becomes

\[
-\frac{\partial^2}{\partial x'^2} + \frac{\partial^2}{\partial c t'^2} = -a_{11} \frac{\partial^2}{\partial x^2} - a_{41} \frac{\partial^2}{\partial c t^2} - a_{11} a_{41} \frac{\partial^2}{\partial x \partial c t} + a_{14} \frac{\partial^2}{\partial x^2} + a_{44} \frac{\partial^2}{\partial c t^2} + a_{14} a_{44} \frac{\partial^2}{\partial x \partial c t}. \tag{2.69}
\]

This simplifies to

\[
-\frac{\partial^2}{\partial x'^2} + \frac{\partial^2}{\partial c t'^2} = -a_{11} \frac{\partial^2}{\partial x^2} - a_{14} \frac{\partial^2}{\partial c t^2} - a_{11} a_{41} \frac{\partial^2}{\partial x \partial c t} + a_{14} \frac{\partial^2}{\partial x^2} + a_{44} \frac{\partial^2}{\partial c t^2} + a_{14} a_{44} \frac{\partial^2}{\partial x \partial c t}
\]

\[
= - a_{11} \frac{\partial^2}{\partial x^2} + a_{14} \frac{\partial^2}{\partial c t^2} + a_{11} a_{41} \frac{\partial^2}{\partial x \partial c t} + a_{14} a_{44} \frac{\partial^2}{\partial x \partial c t}. \tag{2.70}
\]

In order to have equality between the left and right hand side we must have the following relationships

\[
a_{11} a_{41} = a_{14} a_{44}, \quad a_{11}^2 - a_{14}^2 = 1, \quad a_{44}^2 - a_{41}^2 = 1. \tag{2.71}
\]
If we parametrize $a_{11}$ and $a_{44}$ in terms of some variable $\theta$ such that

$$a_{11} = a_{44} = \pm \cosh(\theta)$$

(2.72)

and

$$a_{14} = a_{41} = \pm \sinh(\theta)$$

(2.73)

it obviously follows that

$$\cosh^2(\theta) - \sinh^2(\theta) = 1.$$  

(2.74)

It is important to note two things about the signs chosen. As the hyperbolic sine function can be both negative or positive depending on theta, it doesn’t matter whether the positive or negative sign is chosen. In the case of the hyperbolic cosine, changing the sign results in a a reflection, or what is generally known as a change in parity and time reversal. Then, this parametrization satisfies all three of our conditions given in (2.71). Hence, it satisfies the invariance of the wave equation. This implies that if we wish to transform between coordinate frames the proper transformation to use can be given by

$$x = \cosh(\theta)x' + \sinh(\theta)ct'$$

(2.75)

$$ct = \sinh(\theta)x' + \cosh(\theta)ct'.$$

(2.76)

In matrix notation we have

$$\begin{bmatrix} x \\ ct \end{bmatrix} = \begin{bmatrix} \cosh(\theta) & \sinh(\theta) \\ \sinh(\theta) & \cosh(\theta) \end{bmatrix} \begin{bmatrix} x' \\ ct' \end{bmatrix}.$$  

(2.77)

In order to solve for $(x', ct')$, the determinant of the matrix should be non-zero. Evaluating the determinant,

$$\det \begin{bmatrix} \cosh(\theta) & \sinh(\theta) \\ \sinh(\theta) & \cosh(\theta) \end{bmatrix} = 1.$$  

(2.78)

We now know the matrix is invertible, and solving for $(x', ct')$, we obtain the following result

$$x' = \cosh(\theta)x - \sinh(\theta)ct$$

(2.79)

$$ct' = -\sinh(\theta)x + \cosh(\theta)ct$$

(2.80)

or in matrix notation

$$\begin{bmatrix} x' \\ ct' \end{bmatrix} = \begin{bmatrix} \cosh(\theta) & -\sinh(\theta) \\ -\sinh(\theta) & \cosh(\theta) \end{bmatrix} \begin{bmatrix} x \\ ct \end{bmatrix}.$$  

(2.81)

Since this transformation (which is called a hyperbolic rotation) can be represented as a matrix, it is a linear transformation. We shall denote the rotation matrix as

$$R(\theta) = \begin{bmatrix} \cosh(\theta) & -\sinh(\theta) \\ -\sinh(\theta) & \cosh(\theta) \end{bmatrix}.$$  

(2.82)

and its inverse as

$$R^{-1}(\theta) = \begin{bmatrix} \cosh(\theta) & \sinh(\theta) \\ \sinh(\theta) & \cosh(\theta) \end{bmatrix}.$$  

(2.83)
Both matrix transformations $R(\theta)$ and $R^{-1}(\theta)$ allow us to relate two frames of reference whose rotational transformation is parametrized by $\theta$. These transition matrices preserve the wave operator ($\Box = \Box'$) and are called the Lorentz and Inverse Lorentz Transformations respectively. Now we will define a parameter

$$\beta = \frac{a_{14}}{a_{11}} \quad (2.84)$$

which can also be written as

$$\beta = \tanh(\theta) = \frac{e^\theta - e^{-\theta}}{e^\theta + e^{-\theta}} \quad (2.85)$$

using the exponential definition of the hyperbolic tangent. Since the denominator is never equal to zero and the exponential function takes on all values, $\beta$ is defined for all $\theta$. Furthermore, because

$$\frac{d\tanh(\theta)}{d\theta} = \text{sech}^2(\theta) > 0, \quad (2.86)$$

$\beta$ is also an increasing function.

We can evaluate the limit as $\theta$ approaches $-\infty$ and $\infty$

$$\lim_{\theta \to -\infty} \tanh(\theta) = -1 \quad (2.87)$$
$$\lim_{\theta \to \infty} \tanh(\theta) = 1. \quad (2.88)$$

This implies that $-1 < \beta < 1$. You may ask about the significance or interpretation of $\beta$. First, let us describe the motion of a particle in the $(x,ct)$ frame that is at rest in the $(x',ct')$ frame as shown in Fig. 2.3. The physical point $A(x'_1,ct'_1)$ maps to $A(x_1,ct_1)$ as follows

$$x_1 = x'_1 \cosh(\theta) + ct'_1 \sinh(\theta), \quad ct_1 = x'_1 \sinh(\theta) + ct'_1 \cosh(\theta). \quad (2.89)$$

Similarly, the physical point $B(x'_2,ct'_2)$ maps to $B(x_2,ct_2)$ as follows

$$x_2 = x'_2 \cosh(\theta) + ct'_2 \sinh(\theta), \quad ct_2 = x'_2 \sinh(\theta) + ct'_2 \cosh(\theta). \quad (2.90)$$

It can be observed from the diagram that the two points A and B in the unprimed frame do not lie on the same vertical line which represents a change in position as in Fig. 2.3.

The change in position can be written as

$$\Delta x = x'_2 \cosh(\theta) + ct'_2 \sinh(\theta) - (x'_1 \cosh(\theta) + ct'_1 \sinh(\theta)) \quad (2.91)$$

and simplified to

$$\Delta x = \cosh(\theta)(x'_2 - x'_1) + \sinh(\theta)(ct'_2 - ct'_1). \quad (2.92)$$

However, $x'_1$ and $x'_2$ are the same (since the particle is at rest in the primed frame) so,

$$\Delta x' = \sinh(\theta)(ct'_2 - ct'_1). \quad (2.93)$$

\footnote{This is not the unique transformation that preserves the form of the wave equation However, it is one that maps lines to lines. A property that is related to the fact that if there are no forces acting in one frame, then there will be no forces acting on the other frame. This will be explained shortly.}
Likewise, the change in time is

$$\Delta t = \cosh(\theta)(t_2' - t_1').$$  \hfill (2.94)

Dividing (2.93) by (2.94),

$$\frac{\Delta x}{\Delta t} = c \tanh(\theta)$$  \hfill (2.95)

or, in terms of $\beta$,

$$\frac{\Delta x}{\Delta t} = \beta c.$$  \hfill (2.96)

Recall from introductory physics that $\frac{\Delta x}{\Delta t}$ represents the velocity ($v$) of the particle in the unprimed frame, so

$$v = \beta c.$$  \hfill (2.97)

Therefore, we can conclude

$$\beta = \frac{v}{c}.$$  \hfill (2.98)

The quantity $\beta$ has now acquired a physical interpretation: it is the velocity of the particle in the unprimed frame that was at rest in the primed frame, divided by the speed of light. In particular the particle could have been chosen to lie at the origin of the $(x', ct')$ frame so $\beta$ can then be interpreted as the velocity of the primed frame with respect to the unprimed frame divided by $c$. It is easy to show that if we had chosen the other parametrization

$$\begin{bmatrix} x' \\ ct' \end{bmatrix} = \begin{bmatrix} \cosh(\theta) & \sinh(\theta) \\ \sinh(\theta) & \cosh(\theta) \end{bmatrix} \begin{bmatrix} x \\ ct \end{bmatrix}$$  \hfill (2.99)

then, the only thing that changes is (2.93), it acquires a negative sign. Hence, we obtain a reversal in direction.
Since we found $\beta$ to have a maximum magnitude of 1, $v$ is bounded to be strictly less than the speed of light $-c < v < c$. We can now ask what happens to the Lorentz Boost Transformation when $\beta$ tends to 0.

In order to answer this question, we will describe $\cosh \theta$ and $\sinh \theta$ in terms of $\beta$ by performing some simple algebraic manipulation and use of hyperbolic trigonometric identities

$$\text{sech} \theta = 1 - \beta^2 \quad \rightarrow \quad \cosh \theta = \frac{1}{\sqrt{1 - \beta^2}}, \quad \sinh \theta = \frac{\beta}{\sqrt{1 - \beta^2}}. \quad (2.100)$$

Hence, the Lorentz Boost Transformation$^6$ can be written as

$$x' = \frac{1}{\sqrt{1 - \beta^2}} (x - \beta ct) \quad (2.101)$$
$$ct' = \frac{1}{\sqrt{1 - \beta^2}} (-\beta x + ct) \quad (2.102)$$

or

$$x' = \gamma (x - vt) \quad (2.103)$$
$$ct' = \gamma (\frac{xv}{c} + ct) \quad (2.104)$$

where $\gamma = \frac{1}{\sqrt{1 - \beta^2}}$. Also, $y' = y$ and $z' = z$ since neither $y$ or $z$ are involved in the transformation.

Similarly, the Inverse Lorentz Boost is

$$x = \gamma (x' + vt') \quad (2.105)$$
$$ct = \gamma (\frac{x'v}{c} + ct') \quad (2.106)$$

The Taylor expansion of $\frac{1}{\sqrt{1 - \beta^2}}$ is given by

$$\frac{1}{\sqrt{1 - \beta^2}} = 1 + \frac{1}{2} \beta^2 + \frac{3}{8} \beta^4 + \ldots \quad (2.107)$$

So, for $\beta << 1$

$$\frac{1}{\sqrt{1 - \beta^2}} \approx 1. \quad (2.108)$$

Hence, the Lorentz Boost Transformation for $\beta << 1$ is

$$x' = x - \beta ct \quad (2.109)$$
$$ct' = -\beta x + ct \quad (2.110)$$

or (since $\frac{\beta}{c} \approx 0$)

$$x' = x - vt \quad (2.111)$$

$^6$Keep in mind that this transformation can be interpreted as the relation between two frame of reference. A primed frame that is moving with constant velocity with respect to the unprimed frame. These frames are called inertial frames.
\[ t' = t \] (2.112)

which is the simple, well-known Galilean Transformation.

Thus, for small values of \( \beta \), the Lorentz Boost Transformation tends to the Galilean Transformation as was known by Newton. It is interesting to note that the symmetry of space and time is broken at small values of \( \beta \), and that time is the same in both reference frames. This agrees with our everyday experiences.

---

**Space-time Interval Invariance:** We have looked at many transformations so far. In particular we just discussed how \( x \) and \( t \) transform. Since they transform, it is clear that the classical definitions of space and time are not invariant. Instead we can look at another quantity which we will call \( s^2 \), distance, which will be invariant. First we define the following quantities,

\[
s'^2 = (ct')^2 - x'^2 \tag{2.113}
\]

and

\[
s^2 = (ct)^2 - x^2. \tag{2.114}
\]

Combining (2.99) and (2.113)

\[
s^2 = (-x \sinh \theta + ct \cosh \theta)^2 + (x \cosh \theta - ct \sinh \theta)^2. \tag{2.115}
\]

Expanding and simplifying,

\[
s'^2 = (ct')^2 - x'^2. \tag{2.116}
\]

Then using (2.114)

\[
s'^2 = s^2. \tag{2.117}
\]

This means that the quantity \( s^2 \) is invariant under a Lorentz Transformation. This is also known as the space-time interval, and it can be interpreted as the norm of a vector in space-time. It can be shown that the more general invariant quantity is

\[
(ct')^2 - x'^2 - y'^2 - z'^2 = (ct)^2 - x^2 - y^2 - z^2. \tag{2.118}
\]

Even though the coordinates of \( ct \) and \( x \) might change under a Lorentz Transformation, they follow a relation that makes the difference of their squares invariant. This invariance of space and time will be important as it is known as the Minkowski metric.

---

We have concluded that we can mix up \( x \) and \( t \) and there is a transformation where the wave operator is invariant, these are the hyperbolic rotations. Previously we found that we could make transformations with \( x, y, \) and \( z \) that would make the wave operator invariant as well. We can combine these together using a sequence of rotations and the wave operator should still remain invariant. The most general transformation where we mix \( x, y, z, t \) is known as the Lorentz Transformation rather than just the Lorentz Boost in one particular direction. We will discuss this in more detail as we proceed.
Chapter 3

Physical Consequences of Lorentz Transformations

We previously outlined the most general case of the Lorentz Transformation. We will discuss multiple variations of this transformation as well as the various physical consequences.

3.1 Effect on Wavelength: Red and Blue Shifts

First we will examine the dependence of the wavelength of light on the chosen frame of reference. Let $\lambda$ be the wavelength of light coming from a star that is moving with constant velocity along the $x$-axis detected by an observer located on earth. We will use the coordinates as shown in Fig. 3.1. In this frame, light is propagating in the negative $\hat{x}$ direction so that it can be detected by the observer. We will consider the case where $\phi = 0$ and $\vec{A}$ is only in the $\hat{z}$ direction. We will suppose that in this case the wave is linearly polarized in the $\hat{z}$ direction because it will make the solution to the wave equation much easier to work with in both this section and the next. Recall that $\lambda = \frac{2\pi}{k}$. We now shift to the frame of the star $(x', y', z', ct')$ parametrized by $\beta$ as in (2.103) and (2.104). The equation describing light in this frame is then given by

$$A'_z(x', t') = A_0 \cos(k'(x' + ct'))$$ (3.1)

where we know from before that $A'_z(x', t') = A_z(x, t)$. Combining (2.103) and (2.104),

$$A'_z(x', t') = A_0 \cos \left(k' \left(\gamma(x - vt) + \gamma \left( ct - \frac{xv}{c} \right) \right) \right)$$ (3.2)

which can be rewritten as

$$A'_z(x', t') = A_0 \cos \left(k' \gamma \left( 1 - \frac{v}{c} \right) (x + ct) \right) = A_z(x, t)$$ (3.3)

where $v$ is the velocity of the star in the earth’s frame. This makes it clear that if an object has a wavelength in one frame, it has a different wavelength in another.

Comparing (3.1) and (3.3) we find that we can set them equal. We obtain the relation

$$k = k' \gamma \left( 1 - \frac{v}{c} \right)$$ (3.4)
or

\[ k = k' \frac{1 - \frac{v}{c}}{\sqrt{1 - \frac{v^2}{c^2}}} = k' \sqrt{\frac{1 - \frac{v}{c}}{1 + \frac{v}{c}}}. \] (3.5)

In terms of wavelength,

\[ \lambda = \lambda' \sqrt{\frac{1 + \frac{v}{c}}{1 - \frac{v}{c}}}. \] (3.6)

where \( \lambda' \) is the wavelength of light propagating in the primed frame. This is an amazing result that states that if \( v > 0 \) (receding star) then the wavelength observed in the unprimed frame is larger than in the primed frame. This is known as red shift, since the wavelength tends to the red side of the electromagnetic spectrum. Similarly, if \( v < 0 \) (approaching star) the wavelength observed in the unprimed frame is smaller than that one in the primed frame. This is known as blue shift. Both effects play an important role in observational astronomy where the light of distant stars is analyzed (spectroscopy) and different properties of the stars are recorded. This effect can also be known as the Doppler Effect for Electromagnetic Waves as it is similar to the phenomenon that sound waves undergo for moving sources or moving observers.

### 3.2 Effect on Wavelength: General Direction of Motion

We previously considered a rather simple case to observe the phenomena of red and blue shifts. Now we can think of some more arbitrary motion to get a more generalized relation for the change in wavelength. Let us consider the case where the light propagates at an angle \( \alpha \) with respect to the \( x \)-axis in the unprimed frame. Light observed in the unprimed frame is given by

\[ A_z(x, y, t) = A_0 \cos \left[ k(x \cos \alpha + y \sin \alpha + ct) \right]. \] (3.7)
Now let us go to the frame of the star (primed frame.) We will proceed differently from the previous calculations. We use the Lorentz Transformation given in (2.83), then

$$A'_x(x', y', t') = A'_0 \cos \left[ k \left[ (x' \cosh \theta + ct' \sinh \theta) \cos \alpha + y' \sin \alpha + (x' \sinh \theta + ct' \cosh \theta) \right] \right]$$  \hspace{1cm} (3.8)

and grouping like terms we obtain

$$A'_x(x', y', t') = A'_0 \cos \left[ k \left[ x'( \cosh \theta \cos \alpha + \sinh \theta) + y' \sin \alpha + ct'( \sinh \theta \cos \alpha + \cosh \theta) \right] \right].$$  \hspace{1cm} (3.9)

Recall that for a general wave in three dimensions given by

$$\vec{E}(x, y, z, t) = E_0 \cos(kx + ky + kz + \omega t) \hat{k}$$  \hspace{1cm} (3.10)

the wave number is given by

$$k = \sqrt{k_x^2 + k_y^2 + k_z^2}$$  \hspace{1cm} (3.11)

which can be proven using the wave equation. From this logic, the new wave number in (3.9) is

$$k' = k \sqrt{(\cosh \theta \cos \alpha + \sinh \theta)^2 + \sin^2 \alpha} = k \sqrt{\cosh^2 \theta \cos^2 \alpha + 2 \cosh \theta \cos \alpha \sinh \theta + \sinh^2 \theta + \sin^2 \alpha}. $$  \hspace{1cm} (3.12)

Using the identities $\cosh^2 \theta = 1 + \sinh^2 \theta$ and $\sin^2 \alpha + \cos^2 \alpha = 1$ we obtain

$$k' = k \sqrt{\sinh^2 \theta \cos^2 \alpha + 2 \cosh \theta \cos \alpha \sinh \theta + \cosh^2 \theta} = k \sqrt{(\sinh \theta \cosh \alpha + \cosh \theta)^2} = k(\sinh \theta \cos \alpha + \cosh \theta).$$  \hspace{1cm} (3.13)

It follows that (3.9) can be written as

$$A'_x(x', y', t') = A'_0 \cos \left[ k' \left[ x' \frac{\cosh \theta \cos \alpha + \sinh \theta}{\sinh \theta \cos \alpha + \cosh \theta} + y' \frac{\sin \alpha}{\sinh \theta \cos \alpha + \cosh \theta} + ct' \right] \right].$$  \hspace{1cm} (3.14)

There are two interesting pieces of information that can be deduced from this equation. First, notice how nicely $ct'$ came out after algebraic manipulation. This is naturally consequential from the fact that the wave equation is preserved under the Lorentz Transformation (i.e. the invariance we previously showed.) Secondly it is important to observe that the angle of propagation has changed. Let

$$\cos \alpha' = \frac{\cosh \theta \cos \alpha + \sinh \theta}{\sinh \theta \cos \alpha + \cosh \theta}, \quad \sin \alpha' = \frac{\sin \alpha}{\sinh \theta \cos \alpha + \cosh \theta}. $$  \hspace{1cm} (3.15)

The new angle of propagation is then given by,

$$\tan \alpha' = \frac{\sin \alpha}{\cosh \theta \cos \alpha + \sinh \theta} = \frac{\tan \alpha}{\cosh \theta + \frac{\sin \theta}{\cos \alpha}}.$$  \hspace{1cm} (3.16)

Relating this equation to the velocity of the star using (2.100) and the relation $\beta = \frac{v}{c}$,

$$\tan \alpha' = \frac{\sqrt{1 - \frac{v^2}{c^2}}}{1 + \frac{v}{c \cos \alpha}} \tan \alpha.$$  \hspace{1cm} (3.17)

Both equations are consistent with each other, since $\cos^2 \alpha' + \sin^2 \alpha' = 1$. This is not a coincidence because it comes from how we manipulate the equation.
From (3.13) we obtain

$$k' = k \frac{1 + \frac{v \cos \alpha}{c}}{\sqrt{1 - \frac{v^2}{c^2}}}$$

(3.18)

or in terms of wavelength

$$\lambda = \lambda' \frac{1 + \frac{v \cos \alpha}{c}}{\sqrt{1 - \frac{v^2}{c^2}}}.$$  

(3.19)

These are more general expressions for the change in wavelength and angle of propagation. It is easy to see that they reduce to the the formulas for the unrotated case ($\alpha = 0$.)

### 3.3 Effect on Distance: Length Contraction

Another consequence of the Lorentz Transformation is Length Contraction. Before starting to discuss the mathematics it is important to define what we mean by length. **Length is defined as the distance between two points that are measured simultaneously in one frame.** It is important that we go through this carefully because otherwise it can become very confusing very quickly. Let us consider two points that are stationary in the $(x', ct')$ frame and see how they look in the $(x, ct)$ frame. As discussed before, both the spatial and time coordinates will change. Hence, the coordinates $(x'_1, ct'_1)$ will map to a different pair $(x_1, ct_1)$ and the same is true for $(x'_2, ct'_2)$. It is important to remark that the simultaneity of events (measuring the distance) in the $(x', ct')$ frame is not true for the $(x, ct)$ frame. In other words, the events are not simultaneous in the $(x', ct')$ frame! Therefore, there is no such thing as length for these two physical points in the $(x, ct)$ frame. Instead what we can do is to find two points that have the same time component in the $(x, ct)$ frame and measure the spatial distance between them as in Fig. 3.2.

First we can find the change in position and the change in time between these two stationary physical points in the primed frame. Applying (2.105) and (2.106) to the two physical points and taking the difference

$$\Delta x = \gamma \Delta x'$$

$$\Delta ct = \gamma \beta \Delta x'.$$

(3.20)

As a consequence we can conclude that point B has moved an extra distance during the interval $\Delta t'$

$$\Delta_Bx = v \Delta t = \gamma \beta^2 \Delta x'.$$

(3.21)

Correcting for this, the length in the $(x', ct')$ frame is given by

$$L = \Delta x - \Delta_Bx = \gamma (1 - \beta^2) \Delta x'$$

(3.22)

where $L$ is the apparent length between point A and B in the primed frame. After some simple algebra,

$$L = \frac{\Delta x'}{\gamma}.$$

(3.23)

Since $\gamma > 1$ for all $\beta$, $L > \Delta x'$ which implies that the apparent length of an object in a frame that is not at rest with respect to it is smaller. This is the phenomena known as **length contraction**

---

2Keep in mind that now the measurements are being applied to two different physical points in space-time.
It is important to point out that this is a result of the fact that the time intervals in both frames are not the same. In other words, events are not simultaneous.

### 3.4 Effect on Time: Time Dilation

We will not discuss another phenomenon similar to length contraction. Consider the lifetime of a particle in the unprimed reference frame which moves with respect to the particle’s frame and compare it to a frame on earth $(x',ct')$. Since the particle is at rest with respect to itself it is obvious that $x_1 = x_2$. The time coordinates $t_1$ and $t_2$ map to

$$t'_1 = \gamma \left( t_1 - \frac{vx_1}{c^2} \right)$$

$$t'_2 = \gamma \left( t_2 - \frac{vx_2}{c^2} \right).$$

Taking the difference between them

$$\Delta t' = \gamma \Delta t$$

which shows that the lifetime measured on the earth is longer than the lifetime of the particle in its reference frame. This is known as time dilation, and has been experimentally tested for muons which have different lifetimes depending on the velocity. This is an important property because there are particles that are hard to analyze in their rest frame since they have extremely small lifetimes. However, if we let the particle have a velocity close to $c$ then its lifetime would increase considerably and become easier to analyze. In order to understand more physical consequences we will need to discuss some math concepts in more detail.

---

3 This can be done applying a constant magnetic field which will then cause the particle to travel in a circle as long as it is a charged particle.
3.5 Effect on Movement: Velocity Transformations

If we consider the case of length and time, the velocity measured by different observers in different frames also differs.

Consider a particle moving with velocity \( \vec{u} = u_x \hat{i} + u_y \hat{j} + u_z \hat{k} \) in the \((x, y, z, ct)\) frame. The velocity \( \vec{u}' = u'_x \hat{i} + u'_y \hat{j} + u'_z \hat{k} \) in the \((x', y', z', ct')\) frame parametrized by \( \beta \) according to (2.103) and (2.104) is found below. First we obtain the differentials in \( x' \) and \( t' \) from (2.103) and (2.104),

\[
dx' = \gamma(dx - vdt), \quad dt' = \gamma\left(dt - \frac{vdx}{c^2}\right).
\]

From this we see

\[
\frac{dx'}{dt'} = \frac{\gamma(dx - vdt)}{\gamma\left(dt - \frac{vdx}{c^2}\right)}
\]

which can be simplified to

\[
\frac{dx'}{dt'} = \frac{dx}{dt} - \frac{v}{1 - \frac{u_x v}{c^2}}.
\]

We know from aforementioned definitions that \( \frac{dx'}{dt'} = u'_x \) and \( \frac{dx}{dt} = u_x \). Thus, the transformation for velocity in the \( \hat{x} \)-direction becomes

\[
u'_x = \frac{u_x - v}{1 - \frac{u_x v}{c^2}}.
\]

Since \( y' = y \) and \( z' = z \)

\[
dy' = dy, \quad dz' = dz
\]

and so

\[
\frac{dy'}{dt'} = \frac{dy}{dt} \gamma\left(1 - \frac{dx}{dt} \frac{v}{c^2}\right), \quad \frac{dz'}{dt'} = \frac{dz}{dt} \gamma\left(1 - \frac{dx}{dt} \frac{v}{c^2}\right).
\]

We can conclude that the transformation of velocities in the \( \hat{y} \) and \( \hat{z} \) directions are given by

\[
u'_y = \frac{u_y}{\gamma\left(1 - \frac{u_x v}{c^2}\right)} \quad (3.32)
\]

\[
u'_z = \frac{u_z}{\gamma\left(1 - \frac{u_x v}{c^2}\right)} \quad (3.33)
\]

It immediately follows that if \( u_y = u_z = 0 \) then, \( u'_y = 0 \) and \( u'_z = 0 \). Additionally, components of the velocity depend on the speed of the moving frame. One interesting result is that if we let \( u_x = c \) in (3.30) then

\[
u'_x = \frac{c - v}{1 - \frac{cv}{c^2}} = \frac{c - v}{c} = c.
\]

This is a nice consequence since it is consistent with our assumption that the speed of light is invariant regardless of the reference frame chosen. In order to properly see that this result is

\[4\text{Remember that the physical interpretation of this parametrization is that the (x', y', z', ct') is moving with a velocity } \beta \text{ with respect to (x, y, z, ct).}\]
true, consider the simple case where light has velocity $c$ in only the $y$ direction and 0 in the other two directions (i.e. $\vec{u} = c\hat{y}$.) Then in the primed system we get, after some simple algebra, $
olinebreak \vec{u}' = -v\hat{x}' + \sqrt{c^2 - v^2}\hat{y}'$. It follows that $|u| = c$, thus the direction of the light changes but the speed of light does not change.
Chapter 4

Relativistic Mass and Momentum of a Particle

Having discussed in detail the transformations of velocity we can now talk about momentum. We begin by recalling the principle of momentum conservation. For example, in the case of a collision we know that if there are no external forces then the momentum of the system must be conserved. If there is no external force in one frame then it would seem intuitive to conclude that there is no external force in another frame. In lieu of this, it makes sense to conclude that if momentum is conserved in one frame then it is also conserved in the other frame. It wouldn’t really make sense to say that momentum was only conserved in one particular frame because that would be too frame-dependent and too subjective on the method of measurement. So, we can then ask whether mass is conserved or whether it changes from frame to frame. To answer this we are going to use a clever experiment devised by Richard Feynman to teach relativity to his first year physics students.

Consider two particles of mass \( m \) that approach one another with equal and opposite velocities in the xy-plane of the \( (x, y, z, t) \) frame. Particle 1 is moving in the second quadrant and particle 2 is moving in the fourth quadrant and the motion of both particles makes an angle \( \theta \) with respect to the x-axis. After the collision (which will be considered to be elastic), they head away from each other with equal and opposite velocities; particle 1 goes off in the first quadrant and particle 2 moves off in the third quadrant. Let us assume that the motion of both particles makes an angle \( \theta \) with respect to the x-axis after the collision.

We must first determine whether or not momentum is conserved. If we take the definition from introductory physics that the change in momentum is the force i.e.

\[
\vec{F} = \frac{d\vec{p}}{dt}
\]  

then considering that there are no external forces in our ideal collision, momentum must be conserved. It could also be deduced that the momentum of these two particles is conserved in any other reference frame because the laws of physics must be the same everywhere. Let us begin by showing how the momentum is conserved in the static frame.
Both particles have the same speed $u$, and their masses $m$ are identical. The different $p$’s denote the momenta of the particles.

4.1 Momentum conservation with respect to a static reference frame

From introductory physics we know that the momenta of the particles before collision are

$$\vec{p}_{1I} = m(u \cos \theta, -u \sin \theta), \quad \vec{p}_{1I} = m(-u \cos \theta, u \sin \theta).$$

The momenta of the particles after collision are

$$\vec{p}_{1F} = m(u \cos \theta, u \sin \theta), \quad \vec{p}_{1F} = m(-u \cos \theta, -u \sin \theta).$$

The total momentum of the two particles before collision is

$$\sum \vec{p}_i = m(u \cos \theta - u \cos \theta, -u \sin \theta + u \sin \theta) = (0, 0).$$

The total momentum of the two particles after collision is

$$\sum \vec{p}_f = m(u \cos \theta - u \cos \theta, u \sin \theta - u \sin \theta) = (0, 0).$$

Therefore

$$\sum \vec{p}_i = \sum \vec{p}_f$$

which confirms that momentum is indeed conserved.

4.2 Momentum conservation with respect to a moving reference frame

Let us go to a different frame $(x', y', z', ct')$ parametrized by $\beta$ such that,

$$\beta = \frac{v \cos \theta}{c}.$$
Both the magnitude and direction of the velocity changed in this moving frame. The different $p'$'s denote the momenta of the particles.

The initial and final velocity components of both particles will change according to (3.30) and (3.32). The momenta of the particles after the transformation are

$$
\vec{p}'_{1I} = m \left( 0, \frac{-u \sin \theta \sqrt{1 - \frac{u^2 \cos^2 \theta}{c^2}}}{1 - \frac{u^2 \cos^2 \theta}{c^2}} \right), \quad \vec{p}'_{2I} = m \left( \frac{-2u \cos \theta}{1 + \frac{u^2 \cos^2 \theta}{c^2}}, \frac{u \sin \theta \sqrt{1 - \frac{u^2 \cos^2 \theta}{c^2}}}{1 + \frac{u^2 \cos^2 \theta}{c^2}} \right)
$$

(4.8)

$$
\vec{p}'_{1F} = m \left( 0, \frac{u \sin \theta \sqrt{1 - \frac{u^2 \cos^2 \theta}{c^2}}}{1 - \frac{u^2 \cos^2 \theta}{c^2}} \right), \quad \vec{p}'_{2F} = m \left( \frac{-2u \cos \theta}{1 + \frac{u^2 \cos^2 \theta}{c^2}}, \frac{-u \sin \theta \sqrt{1 - \frac{u^2 \cos^2 \theta}{c^2}}}{1 + \frac{u^2 \cos^2 \theta}{c^2}} \right).
$$

(4.9)

Let us see if momentum is still conserved in this frame. According to our current premonitions it should be. The sum of the momentum before the collision is

$$
\sum \vec{p}'_{I} = m \left( \frac{-2u \cos \theta}{1 + \frac{u^2 \cos^2 \theta}{c^2}}, u \sin \theta \sqrt{1 - \frac{u^2 \cos^2 \theta}{c^2}} \frac{1}{1 + \frac{u^2 \cos^2 \theta}{c^2}} - \frac{1}{1 - \frac{u^2 \cos^2 \theta}{c^2}} \right).
$$

(4.10)

The sum of the momentum after the collision is

$$
\sum \vec{p}'_{F} = m \left( \frac{-2u \cos \theta}{1 + \frac{u^2 \cos^2 \theta}{c^2}}, u \sin \theta \sqrt{1 - \frac{u^2 \cos^2 \theta}{c^2}} \frac{1}{1 + \frac{u^2 \cos^2 \theta}{c^2}} + \frac{1}{1 - \frac{u^2 \cos^2 \theta}{c^2}} \right).
$$

(4.11)

It is clear that the sum of the $y$-components before and after are not the same; hence, momentum appears to not be conserved! This is a very unfortunate result. There is no way to explain why momentum is not conserved since there are no external forces acting on the particles in this frame. Therefore, something must be wrong in our calculations and we need to look closer at the formula for momentum and the assumptions we have made so far. Momentum depends on both velocity
and mass. So far we have satisfactorily found a velocity transformation, but we have assumed in these calculations that the masses of the particles remained identical. We can conclude that since our results here don’t make sense, the mass must transform in some way.

4.3 Mass as a function of speed

As has been pointed out previously, the masses of the particles must be different in order to comply with the law of conservation of momentum.

Questions to ask:

- What is the dependence of mass?
- Does it depend on its velocity? Speed? Neither?

Assuming that mass is invariant under rotation mass should not depend on velocity but it should depend on its magnitude i.e. on speed. In order to find the explicit relation between mass and speed we will assume momentum is conserved and find the conditions that follow

\[ \sum \vec{p}'_I = \sum \vec{p}'_F. \]  

(4.12)

Now it becomes necessary to state that the masses of the particles do not change in the collision since particle 1 has the same speed before and after the collision, and the same is true for particle 2. However, both particles have different speeds. This is easy to see from the Fig. 4.2. Therefore, (4.12) implies the following condition

\[
\sqrt{1 - \frac{u^2 \cos^2 \theta}{c^2}} \left( \frac{-2m_2 u \cos \theta}{1 + \frac{u^2 \cos^2 \theta}{c^2}} + \frac{m_1}{1 - \frac{u^2 \cos^2 \theta}{c^2}} \right) = \sqrt{1 - \frac{u^2 \cos^2 \theta}{c^2}} \left( \frac{m_1}{1 - \frac{u^2 \cos^2 \theta}{c^2}} - \frac{m_2}{1 + \frac{u^2 \cos^2 \theta}{c^2}} \right)
\]

(4.13)

where \( m_1 \) and \( m_2 \) are the masses of particle 1 and particle 2 respectively. The first condition is trivially satisfied, while the second one reduces to

\[
- \frac{m_1}{1 - \frac{u^2 \cos^2 \theta}{c^2}} + \frac{m_2}{1 + \frac{u^2 \cos^2 \theta}{c^2}} = \frac{m_1}{1 - \frac{u^2 \cos^2 \theta}{c^2}} - \frac{m_2}{1 + \frac{u^2 \cos^2 \theta}{c^2}}
\]

(4.14)

which can be further reduced to the following condition

\[
\frac{m_1}{m_2} = \frac{1 - \frac{u^2 \cos^2 \theta}{c^2}}{1 + \frac{u^2 \cos^2 \theta}{c^2}}.
\]

(4.15)

This relation must be expressed in terms of speed. Therefore, we will use the following identity,

\[
\sqrt{1 - \frac{w^2}{c^2}} = \frac{1 + \frac{u^2 \cos^2 \theta}{c^2}}{1 - \frac{u^2 \cos^2 \theta}{c^2}}
\]

(4.16)

where \( w_1 \) and \( w_2 \) are the speeds of particle 1 and particle 2 respectively.

\(^1\)Velocity is a vector and hence it is dependent on the set of coordinates chosen. This can also be understood by symmetry arguments.
Mathematical Identity 5: We assert that
\[
\frac{\sqrt{1-\frac{w_1^2}{c^2}}}{\sqrt{1-\frac{w_2^2}{c^2}}} = \frac{1+u^2 \cos^2 \theta}{1-\frac{u^2 \cos^2 \theta}{c^2}}
\]
is true.

**Proof.** The speeds of particle 1 and 2 are
\[
w_1^2 = \frac{u^2 \sin^2 \theta}{(1-u^2 \cos^2 \theta)}
\]
\[
w_2^2 = \frac{4u^2 \cos^2 \theta + u^2 \sin^2 \theta (1-u^2 \cos^2 \theta)}{(1+u^2 \cos^2 \theta)^2}.
\]
(4.17)

Hence,
\[
1-\frac{w_1^2}{c^2} = 1-\frac{u^2 \sin^2 \theta}{c^2(1-u^2 \cos^2 \theta)}
\]
\[
1-\frac{w_2^2}{c^2} = 1-\frac{4u^2 \cos^2 \theta + u^2 \sin^2 \theta (1-u^2 \cos^2 \theta)}{c^2(1+u^2 \cos^2 \theta)^2}.
\]
(4.18)

Creating a common denominator in both numerator and denominator and performing simplifications leads to
\[
1-\frac{w_1^2}{c^2} = \frac{1-\frac{u^2 \sin^2 \theta}{c^2(1-u^2 \cos^2 \theta)}}{1-\frac{4u^2 \cos^2 \theta + u^2 \sin^2 \theta (1-u^2 \cos^2 \theta)}{c^2(1+u^2 \cos^2 \theta)^2}}.
\]
(4.19)

Simplifying the numerator of the denominator,
\[
1-\frac{w_1^2}{c^2} = \frac{1-\frac{u^2 \sin^2 \theta}{c^2(1-u^2 \cos^2 \theta)}}{1-\frac{4u^2 \cos^2 \theta + u^2 \sin^2 \theta (1-u^2 \cos^2 \theta)}{c^2(1+u^2 \cos^2 \theta)^2}}.
\]
(4.20)

Factoring common terms in the numerator of the denominator,
\[
1-\frac{w_1^2}{c^2} = \frac{1-\frac{u^2 \sin^2 \theta}{c^2(1-u^2 \cos^2 \theta)}}{(1-\frac{4u^2 \cos^2 \theta + u^2 \sin^2 \theta (1-u^2 \cos^2 \theta)}{c^2(1+u^2 \cos^2 \theta)^2}}.
\]
(4.21)

and cancelling factors
\[
1-\frac{w_1^2}{c^2} = \left(1 + \frac{u^2 \cos^2 \theta}{c^2}\right)^2
\]
\[
1-\frac{w_2^2}{c^2} = \left(1 - \frac{u^2 \cos^2 \theta}{c^2}\right)^2.
\]
(4.22)

which finally reduces to (4.16) by taking the square root of both sides.

Combining (4.16) and (4.15) we obtain
\[
\frac{m_1}{m_2} = \frac{\sqrt{1-\frac{w_1^2}{c^2}}}{\sqrt{1-\frac{w_2^2}{c^2}}} \rightarrow m_1 \sqrt{1-\frac{w_1^2}{c^2}} = m_2 \sqrt{1-\frac{w_2^2}{c^2}}.
\]
(4.23)
Since this equation is true for arbitrary values of \( w_1 \) and \( w_2 \), then \( m \sqrt{1 - \frac{w^2}{c^2}} = \text{constant} \). If we call the constant \( m_0 \), then
\[
m(w) = \frac{m_0}{\sqrt{1 - \frac{w^2}{c^2}}} \quad \text{(4.24)}
\]

where \( w \) is the speed of the particle, \( m_0 \) is called the rest mass since it is the value of \( m \) at \( w = 0 \). This extraordinary result states that the mass of a particle depends upon its speed and this comes from the law of conservation of momentum. This completely challenges the classical notion of mass being a constant parameter because now it is a dynamic variable. However, the effects of the equation will not be noticeable unless the speed of the particle is several thousands miles per second (or as close to \( c \) as it can get.)

Another consequence of the dependence of mass on speed is that momentum will be written as
\[
\vec{p} = \frac{m_0 \vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}} = m(v)\vec{v}. \quad \text{(4.25)}
\]

Thus, force as defined by (4.1) will be
\[
\vec{F} = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}} \frac{d\vec{v}}{dt} + \frac{\left(\vec{v} \cdot \frac{d\vec{v}}{dt}\right)}{c^2} \frac{m_0 \vec{v}}{\left(1 - \frac{v^2}{c^2}\right)^{3/2}}. \quad \text{(4.26)}
\]

This implies that the acceleration of a particle is not dependent on the direction of the force applied to it. In other words, relativistically Newton’s second law does not hold. This seems rather counterintuitive in the classical way of thinking.

### 4.4 Energy of a Particle

As we discussed earlier, momentum has to be conserved everywhere. Consequently, the mass of each particle depends on its speed. Therefore, we have to modify the classical equations of energy and work in order to be compatible with these ideas. We assume motion along the \( x \)-axis (although the same analysis can be extended to the other coordinates.)

First, we will start with the definition of work
\[
W = \int_{x_1}^{x_2} F_x \, dx. \quad \text{(4.27)}
\]

Recalling that force and momentum are related using the one dimensional version of (4.26), we obtain
\[
W = \int_{t_1}^{t_2} m_0 \frac{dw}{dt} \frac{w}{\left(1 - \left(\frac{w^2}{c^2}\right)\right)^{3/2}} v \, dt. \quad \text{(4.28)}
\]

\(^3\)Note that \( \theta \) and the initial velocities can be adjusted properly to give any value for \( w_1 \) and \( w_2 \).

\(^4\)As usual, the relativistic effects are not observed for ordinary speeds.
Using the condition that the particle starts at rest at \( x_1 \),

\[
W = m_0 \int_0^v \frac{vdv}{\left(1 - \left(\frac{v^2}{c^2}\right)\right)^{\frac{3}{2}}}
\]  
(4.29)

and solving the integral we obtain

\[
W = \frac{mc^2}{\sqrt{1 - \left(\frac{v^2}{c^2}\right)}} - mc^2.
\]  
(4.30)

We know that \( W = K \) where \( K \) is the kinetic energy, so

\[
K = \frac{m_0 c^2}{\sqrt{1 - \left(\frac{v^2}{c^2}\right)}} - m_0 c^2.
\]  
(4.31)

It is important to note that the term \( m_0 c^2 \) does not depend on the particle’s speed. The term \( E_0 = m_0 c^2 \) is known as the rest energy of the particle. This rest mass term is an inherent property of the particle. In other words it is an invariant quantity preserved under any transformation. Therefore, the total energy can be represented as the sum of the kinetic energy and the rest energy,

\[
E = E_0 + K
\]  
(4.32)

or,

\[
E = \frac{m_0 c^2}{\sqrt{1 - \left(\frac{v^2}{c^2}\right)}}
\]  
(4.33)

Using (4.24), the following relation is obtained

\[
E = m(v)c^2.
\]  
(4.34)

This is known as the **Energy-Mass Equivalence Relation**.

### 4.5 Energy of a particle as a function of momentum and rest energy

With what we have derived so far, we can express energy as a function of momentum and rest energy, using the relations in (4.34) and (4.25)

\[
E^2 - (pc)^2 = (m_0)^2(c^4 - v^2c^2).
\]  
(4.35)

After some calculations,

\[
E^2 - (pc)^2 = (m_0 c^2)^2.
\]  
(4.36)

Solving for \( E^2 \),

\[
E^2 = (pc)^2 + (m_0 c^2)^2.
\]  
(4.37)

This is a powerful result that describes a quantity that stays invariant under Lorentz Transformations since the rest energy is preserved. Note that we expressed energy and momentum as functions
of particle speed, therefore these quantities cannot remain invariant in all moving reference frames, but their combination as in the previous equation remains the same. For example, if we consider the Feynman collision problem that we analyzed earlier we can see that when the reference frame is moving, the total momentum is different from the one we calculated at rest. The same result can be obtained if we consider the energy. Interesting results occur when we set the momentum or the mass of the particle equal to zero. For instance, if \( p = 0 \) then the energy can only be expressed as function of its rest mass \( (E = m_0c^2) \), this is why it is known as rest energy. On the other hand, for the case \( m_0 = 0 \), the energy of the particle is proportional to its momentum as is the case for photons.\(^4\)

\[
E^2 = p^2c^2. \tag{4.38}
\]

### 4.6 Energy and momentum transformations

We can now address the transformation of energy and momentum. However, before we go any further it is important to prove the following identity that is going to help us derive expressions for \((E', p')\) in the \((x', y, z, ct')\) frame in terms of \((E, p)\) in the \((x, y, z, ct)\) frame and \(v\),

\[
\sqrt{1 - \frac{u^2}{c^2}} \left( 1 + \frac{vu_x'}{c^2} \right) = \sqrt{1 - \frac{v^2}{c^2}} \sqrt{1 - \frac{u'^2}{c^2}}. \tag{4.39}
\]

**Mathematical Identity 6:** We assert that \(\sqrt{1 - \frac{u^2}{c^2}} \left( 1 + \frac{vu_x'}{c^2} \right) = \sqrt{1 - \frac{v^2}{c^2}} \sqrt{1 - \frac{u'^2}{c^2}}\) is true.

**Proof.** First, let us start with the following equation,

\[
1 - \frac{u^2}{c^2} = 1 - \frac{u^2_x + u^2_y + u^2_z}{c^2}. \tag{4.40}
\]

Using the velocity transformations in \((3.30), (3.32)\) and \((3.33)\), it follows that

\[
1 - \frac{u^2}{c^2} = 1 - \frac{(u'_{x'} + v)^2 + (u'_{y'} \sqrt{1 - \frac{v^2}{c^2}})^2 + (u'_{z'} \sqrt{1 - \frac{v^2}{c^2}})^2}{c^2 \left( \frac{vu'_{x'}}{c^2} + 1 \right)^2}. \tag{4.41}
\]

and multiplying the previous equation by \(c^2 \left( \frac{vu'_{x'}}{c^2} + 1 \right)\) we have

\[
c^2 \left( 1 - \frac{u^2}{c^2} \right) \left( \frac{vu'_{x'}}{c^2} + 1 \right)^2 = c^2 + 2u'_{x'} v + \frac{u^2_{x'} v^2}{c^2} - \left( u'^2_{x'} + 2u'_{x'} v + v^2 + u'^2_{y'} - u'^2_{y'} \frac{v^2}{c^2} + u'^2_{z'} - u'^2_{z'} \frac{v^2}{c^2} \right). \tag{4.42}
\]

Grouping terms and simplifying,

\[
c^2 \left( 1 - \frac{u^2}{c^2} \right) \left( \frac{vu'_{x'}}{c^2} + 1 \right)^2 = c^2 + \frac{v^2 u^2}{c^2} \left( u'^2_{x'} + u'^2_{y'} + u'^2_{z'} \right) - \left( u'^2_{x'} + u'^2_{y'} + u'^2_{z'} \right) - v^2 \tag{4.43}
\]

which is the same as

\[
c^2 \left( 1 - \frac{u^2}{c^2} \right) \left( 1 + \frac{vu'_{x'}}{c^2} \right)^2 = c^2 + \frac{v^2 u^2}{c^2} - u'^2 - v^2. \tag{4.44}
\]

\(^4\)Photons are quanta of electromagnetic waves and they will be discussed later on.
Grouping the terms on the right hand side,
\[ c^2 \left(1 - \frac{u'^2}{c^2}\right) \left(1 + \frac{vu'}{c^2}\right)^2 = c^2 \left(1 - \frac{v^2}{c^2}\right) \left(1 - \frac{u'^2}{c^2}\right). \]  
(4.45)

Finally, dividing both sides by \( c^2 \) and taking a square root, yields (4.39).

Using this identity we are set up to find \( E' \) and \( p' \). Multiplying and dividing (4.34) by \( 1 + \frac{vu'}{c^2} \)
\[ E = \frac{m_0 c^2}{\sqrt{1 - \frac{u'^2}{c^2}}} \left(1 + \frac{vu'}{c^2}\right), \]  
(4.46)

Then, substituting Mathematical Identity 6 into the denominator,
\[ E = \frac{m_0 c^2}{\sqrt{1 - \frac{v^2}{c^2}}} \frac{\left(1 + \frac{vu'}{c^2}\right)}{\sqrt{1 - \frac{u'^2}{c^2}}}, \]  
(4.47)

which can also be written as
\[ E = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \left(\frac{m_0 c^2}{\sqrt{1 - \frac{u'^2}{c^2}}} + \frac{m_0 u' c}{\sqrt{1 - \frac{v^2}{c^2}}}\right) = \frac{E' + \frac{p'}{c} v}{\sqrt{1 - \frac{v^2}{c^2}}}, \]  
(4.48)

Similarly, the relation for momentum can be found using (4.25) and (3.30),
\[ p_x = \frac{m_0 \left(\frac{u_x'}{c} + v\right)}{\sqrt{1 - \frac{u'^2}{c^2}} \sqrt{1 - \frac{v^2}{c^2}}}, \]  
(4.49)

Using Mathematical Identity 6 once again,
\[ p_x = \frac{m_0}{\sqrt{1 - \frac{u'^2}{c^2}}} \frac{u_x' + v}{\sqrt{1 - \frac{u'^2}{c^2}}} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \left(\frac{m_0 u_x'}{\sqrt{1 - \frac{u'^2}{c^2}}} + \frac{vm_0 c^2}{c^2 \sqrt{1 - \frac{u'^2}{c^2}}}\right) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \left(p_x' + \frac{v E'}{c^2}\right). \]  
(4.50)

This can be rearranged to obtain
\[ \frac{p_x'}{\sqrt{1 - \frac{v^2}{c^2}}} = p_x - \frac{E'}{c^2} \frac{v}{\sqrt{1 - \frac{v^2}{c^2}}} \]  
(4.51)

or
\[ \frac{E'}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{c^2}{v} \left(p_x - \frac{p_x'}{\sqrt{1 - \frac{v^2}{c^2}}}\right). \]  
(4.52)
Substituting this equation into (4.48),

\[
E = \frac{E'}{\sqrt{1 - \frac{v^2}{c^2}}} + v(p_x - \frac{E' v}{\sqrt{1 - \frac{v^2}{c^2}}}).
\]  

(4.53)

After some algebraic manipulations we finally obtain

\[
E' = \frac{E - vp_x}{\sqrt{1 - \frac{v^2}{c^2}}}. 
\]  

(4.54)

This is the correct transformation for energy. As can be seen, it depends on both energy and momentum (as it is expected.)

---

**Momentum Transformation:** In order to find the transformation for momentum, we can substitute (4.52) into (4.48), then

\[
E = \frac{c^2}{v} p_x - \frac{p'_{x'}}{\sqrt{1 - \frac{v^2}{c^2}}} \left( \frac{c^2}{v} - v \right).
\]  

(4.55)

Expanding the first term on the left hand side yields

\[
E = \frac{c^2}{v} p_x - \frac{p'_{x'}}{\sqrt{1 - \frac{v^2}{c^2}}} \left( \frac{c^2}{v} - v \right).
\]  

(4.56)

Pulling out a \(\frac{c^2}{v}\) from the second term on the left hand side, simplifying and grouping terms

\[
E = \frac{c^2}{v} \left( p_x - \frac{p'_{x'}}{\sqrt{1 - \frac{v^2}{c^2}}} \sqrt{1 - \frac{v^2}{c^2}} \right).
\]  

(4.57)

Finally, solving for \(p'_{x'}\),

\[
p'_{x'} = \frac{p_x - \frac{E}{c^2} v}{\sqrt{1 - \frac{v^2}{c^2}}}.
\]  

(4.58)

This is the proper transformation for momentum. Likewise, for \(p_y\) we obtain

\[
p_y = \frac{m_0 u_y}{\sqrt{1 - \frac{u^2}{c^2}}},
\]  

(4.59)

Using the velocity transformation in (3.32),

\[
p_y = \frac{m_0}{\sqrt{1 - \frac{u^2}{c^2}}} \left( \frac{u'_y \sqrt{1 - \frac{v^2}{c^2}}}{c^2 u'_x + 1} \right).
\]  

(4.60)

Once again, using Mathematical Identity 6 and simplifying

\[
p_y = \frac{m_0 u'_y}{\sqrt{1 - \frac{u^2}{c^2}}} = p'_y.
\]  

(4.61)

Similar results hold for \(p'_{z'}\), so

\[
p'_z = p_z.
\]  

(4.62)
Let us rewrite (4.54) and (4.58) in a slightly different way,
\[
\frac{E'}{c} = \gamma \left( \frac{E}{c} - \beta p_x \right)
\]
\[
p'_{x'} = \gamma \left( p_x - \beta \frac{E}{c} \right).
\]
These equations reveal the strong symmetry between the space-time and momentum-energy relations.\(^5\) Let us consider an interesting example: the case of massless particles \((E = pc)\). In this case (4.54) reduces to
\[
p'c = p(c - v) \sqrt{\frac{1}{1 - \frac{v^2}{c^2}}}
\]
which can be written as
\[
p'c = pc \sqrt{\frac{1 + \frac{v}{c}}{1 - \frac{v}{c}}} \quad (4.65)
\]
or
\[
E' = E \sqrt{\frac{1 + \frac{v}{c}}{1 - \frac{v}{c}}}.
\]
What is particularly interesting about this expression is its similarity to the Doppler shift effect. Therefore we can conclude that energy also gets “Doppler shifted”. Thus it can be inferred that, for massless particles, the energy and momentum are proportional to their frequencies.\(^6\)
\[
E \propto f, \quad p \propto f.
\]
\[\]  
**4.7 Transformation of Forces**

Now that we have discussed the transformation of energy and momentum we can derive expressions for \(F'\) in the \((x', y, z, ct')\) frame in terms of \(F\), \(u\), and \(v\). Using the one dimensional definition of force given in (4.1)
\[
F'_{x'} = \frac{dp'_{x'}}{dt'}.
\]
Combining this with (4.63)
\[
F'_{x'} = \frac{d}{dt'} \left( \frac{p_x - \frac{v}{c^2} E}{\sqrt{1 - \frac{v^2}{c^2}}} \right).
\]
\[\]  
\(^5\)Recall that in this case the associated invariant quantity is \(m_0c^2\), while for space-time the invariant quantity was \(-s^2\).

\(^6\)At this point, it might be confusing to talk about the frequency or wavelength of a massless particle, but keep in mind that photons (massless) are light so they are related to electromagnetic waves.
Using the chain rule,
\[
F'_{x'} = \frac{d}{dt} \left( p_x - \frac{u}{c^2} E \right) \frac{dt}{dt'} = \frac{dp_x}{dt'} - \frac{v}{c^2} \frac{dE}{dt'} \sqrt{1 - \frac{v^2}{c^2}} \frac{1 - \frac{u^2}{c^2}}{\frac{v}{c^2}}
\]
(4.70)
which simplifies to
\[
F'_{x'} = \frac{F_x - \frac{v}{c^2} \frac{dE}{dt'}}{1 - \frac{u^2}{c^2}}.
\]
(4.71)
In order to simplify the problem further, we will obtain another relation for energy that comes from
\[
E = \sqrt{(p\cdot\bar{p})c^2 + E_0^2}.
\]
(4.72)
Taking the derivative with respect to time,
\[
dE \frac{dt}{dt'} = c^2 \bar{F} \cdot \bar{p}.
\]
(4.73)
Multiplying both sides by \(E\),
\[
E \frac{dE}{dt} = c^2 \bar{F} \cdot \bar{p}.
\]
(4.74)
Applying the relation of energy and mass given in (4.34), and solving for \(\frac{dE}{dt}\)
\[
\frac{dE}{dt} = \bar{F} \cdot \bar{u}.
\]
(4.75)
Combining these equations with (4.71)
\[
F'_{x'} = \frac{F_x - \frac{v}{c^2} \bar{F} \cdot \bar{u}}{1 - \frac{u^2}{c^2}}.
\]
(4.76)
This equation shows the transformation of forces for different frames of reference. For the \(y\)-component of the force we have
\[
F'_{y'} = \frac{dp_y'}{dt'}
\]
(4.77)
using the momentum transformation for \(y\) and the chain rule
\[
F'_{y'} = \frac{dp_y}{dt} \frac{dt}{dt'}
\]
(4.78)
Thus, we obtain
\[
F'_{y'} = F_y \sqrt{1 - \frac{v^2}{c^2}} \frac{1 - \frac{u^2}{c^2}}{1 - \frac{u^2}{c^2}u_x}.
\]
(4.79)
Similar results hold for the \(z\)-component
\[
F'_{z'} = F_z \sqrt{1 - \frac{v^2}{c^2}} \frac{1 - \frac{u^2}{c^2}}{1 - \frac{u^2}{c^2}u_x}.
\]
(4.80)
In this section we derived and discussed further consequences of the Lorentz Transformation. Namely, the dependence of mass on velocity, relativistic energy, momentum and force, and their proper transformation to other frames. In addition we briefly discussed invariant quantities and their symmetry with their space-time counter parts. In particular, the force transformations will play an important role in the next section which is dedicated to the consequences of relativity on electricity.
Chapter 5

Coulomb’s Law and Special Relativity

Until now we have seen that space, time, mass, energy, momentum and forces transform under the Lorentz Transformation. These ideas will help us to derive the correct relation between electric and magnetic forces of stationary and moving charges. In other words, we will derive magnetism from electricity using special relativity. We will start with the correct form of Coulomb’s Law. Experiments have shown that Coulomb’s law gives the correct force for a test charge with any velocity as long as the source charge is stationary. Fig. 5.1 illustrates this interaction of charges and Coulomb’s law is given by

\[ \vec{F} = \frac{kq_s q_t}{r^2} \hat{r} \]  

where \( q_s \) is the source charge and \( q_t \) is the test charge.

However, Coulomb’s Law does not address the case when the source is moving with constant velocity. This can be understood from the fact that an electric charge causes an electric field everywhere in space. If the electric charge is at rest, the electric field will be static, meaning it will not evolve over time. However, if the charge is moving the electric field in space will change but the speed of propagation of the disturbance is finite so it travels at the speed of light. Thus, Coulomb’s law will not give the correct instantaneous force at a point in space. Fortunately, the story is not over because we can use our arsenal of transformations to find the correct electric force for a moving source and moving charges. We will consider different cases but the current underlying mechanics is that we will go to the inertial frame where the source charge is at rest and calculate the corresponding force in this frame. Then, we will use force transformations to return to our original frame.

5.1 Interactions of moving charges

Let us consider a source charge that is moving with constant velocity \( \vec{v} \) (with its position initially at the origin) in the \( \hat{x} \) direction and a test charge is at rest at some distance \( x \) from the source charge (refer to Fig. 5.2). In a classical introductory physics course we learned to consider Biot-Savart’s Law along with the Lorentz force because the source charge is moving. Before we consider that, we must ask ourselves if such a distinction makes any sense. The fact that one charge is moving with a certain velocity is completely arbitrary. For any charge there is always a reference frame

\[ \text{This could be tested with a Millikan apparatus at relativistic speeds.} \]
Figure 5.1: Interaction of charges.

\[ F = k \frac{q_1 q_2}{r^3} \]

Figure 5.2: Interaction of charges.
where it is moving and another where it is at rest. The relativity postulate tells us that all inertial reference frames are the same and therefore considering Coulomb’s Law in a reference frame that is inertial with the source charge should give us the correct force. Let us consider what it looks like in the frame that is at rest with source charge, where the test charge is moving with a speed $v$ in the negative direction as shown in Fig. 5.3. Before we continue we must consider what happens when we transform the coordinates into the other reference frame. The coordinates $(x,0,0,0)$ and $(0,0,0,0)$ will not get transformed to $(x',0,0,0)$ and $(0,0,0,0)$ but rather to $(x',0,0,t')$ and $(0,0,0,0)$. However we only need to know the force on the point $(x',0,0,t')$ because that is the coordinate of the test charge in the new reference frame (remember we have done away with Newton’s third law.) In our new reference frame the source charge stays at the origin at all times (i.e. the source charge is at $(0,0,0)$ at time $t'$ or we could say that the world line of the source charge passes through the point $(0,0,0,t')$.) We can then still calculate the force on the test charge using Coulomb’s Law and we get

$$F'_{x'} = \frac{kq_1q_s}{x'^2}$$

then using our appropriate Lorentz transformation we get that the force in our original frame is

$$F_x = F'_{x'} = \frac{kq_1q_s}{x'^2} = \frac{kq_1q_s}{\gamma^2 x'^2}.$$  

Now because

$$\frac{kq_1q_s}{(1 - \frac{v^2}{c^2}) x^2} = \frac{kq_1q_s}{x^2} \left(1 + \frac{1}{2}\frac{v^2}{c^2} + ...\right)$$

we know that if $v << c$ then the relativistic force is approximately the same as the force predicted by classical mechanics. In general we still have that instantaneous action at a distance contradicts Einstein’s theory of relativity. This is because there is no such thing as simultaneity. In addition, Newton’s third law also contradicts the relativity postulate. We can conclude that there does not exist an equal and opposite force. There is no question as to why Einstein wrote a paper on moving charges, it really highlights the faults involved with attempting to understand electromagnetism from the standpoint of classical mechanics.
We will now consider the situation where the test charge is at rest on the \( y \)-axis and the source charge is at the origin moving with a velocity \( v \) in the \( \hat{x} \)-direction as modeled in Fig. 5.4. If we use the Lorentz transformation to transform to the reference frame where the source charge is at rest at the origin and the test charge is at some some distance \( y \) on the \( y \)-axis from the source charge moving with a velocity \( v \) in the negative \( \hat{x} \)-direction as shown in Fig. 5.5. We have that

\[
F_{y'} = \frac{kq_1q_s}{y'^2}. \tag{5.5}
\]

If we transform back to the reference frame that is at rest with respect to the test charge then we have, by applying the appropriate Lorentz transformation, that the force in our original reference frame is

\[
F_y = \frac{F_{y'}}{\gamma(1 + \frac{vu}{c^2})} = \frac{F_{y'}}{\gamma(1 - \frac{v^2}{c^2})} = \gamma \frac{kq_1q_s}{y^2}, \tag{5.6}
\]

where \( u_{t'} \) is the velocity of the test charge in the other reference frame.
We will now consider an analogous situation: the test charge is at rest on the $z$-axis and the source charge is at the origin moving with a velocity $v$ in the $\hat{x}$-direction

$$F_z = \frac{F'_z}{\gamma(1 + \frac{vu_t}{c^2})} = \frac{F'_z}{\gamma(1 - \frac{v^2}{c^2})} = \gamma \frac{kq_t q_s}{z'^2}. \quad (5.7)$$

Now we are ready to consider the situation where we have the source charge at the origin moving in the $x$-direction with a velocity $v$ and the test charge is at rest at an arbitrary position. Once again we transform to a reference frame where the source charge is at rest as shown in Fig. 5.6. Then we have

$$\vec{F}'_x = \frac{kq_t q_s}{r'^3} \vec{x}'$$
$$\vec{F}'_y = \frac{kq_t q_s}{y'^3} \vec{y}'$$
$$\vec{F}'_z = \frac{kq_t q_s}{z'^3} \vec{z}'. \quad (5.8)$$

Moreover, we know from the transformation of forces given in (4.76), (4.79) and (4.80) that in this situation we have

$$F_x = F'_x$$
$$F_y = \gamma F'_y$$
$$F_z = \gamma F'_z. \quad (5.9)$$

Thus, if we transform back into our original reference frame we obtain

$$\vec{F}_x = \gamma \frac{kq_t q_s}{r'^3} \vec{x}$$
$$\vec{F}_y = \gamma \frac{kq_t q_s}{y'^3} \vec{y}$$
$$\vec{F}_z = \gamma \frac{kq_t q_s}{z'^3} \vec{z}. \quad (5.10)$$

From previous relations we have

$$x' = \gamma x$$
\[ y' = y, \quad z' = z. \] (5.11)

Then we obtain
\[ r' = (x'^2 + y'^2 + z'^2)^{\frac{1}{2}} = (\gamma^2 x^2 + y^2 + z^2)^{\frac{1}{2}}. \] (5.12)

These results together imply that
\[ \vec{F} = \frac{kq_t q_s}{((\gamma^2 x^2) + y^2 + z^2)^{\frac{3}{2}}} \gamma \vec{r}. \] (5.13)

which allows us to calculate the force that a moving source charge (with any arbitrary velocity) exerts on a stationary charge, since we have made the assumption that the laws of physics are invariant under rotations. Therefore we can always choose our axis such that the velocity of the source charge is parallel to the \( x \)-axis. We can also derive the correct formula for the electric field. Since we considered the force on an arbitrary test charge at rest, it is natural to define the electric field as
\[ \vec{E} = \frac{kq_s}{((\gamma^2 x^2) + y^2 + z^2)^{\frac{3}{2}}} \gamma \vec{r}. \] (5.14)

Notice that this equation implies that the electric field is no longer spherically symmetric. However, at each instant it is radially outward. It should also be noted that the reason why the original expression for the electric field did not make sense is because it assumed that there existed action at a distance. The relativity postulate states that nothing (mass, energy, or a force) can travel through space faster than the speed of light. This implies that any signal must travel slower than or at the speed of light. We can therefore conclude that the signal must travel with speed \( c \) from its position at time \( t_0 \) where it takes time \( \frac{r}{c} \) to reach a particle at an arbitrary distance \( r \) away from the particle at the origin. Therefore the particle will feel the force that it would feel from a source charge at the position the source charge was in at time \( t_0 \) as opposed to the force it would feel from the source charge when it is at the position it occupies during the time \( t_0 + \frac{r}{c} \). This seemingly simple observation is very surprising! It is as if nature is constantly keeping track of where all the source charges were during some prior time and calculates the force the test charge should have felt at that prior time.

Consider the case where both of the charges are moving with a velocity \( v \) in the positive \( x \)-direction, the source charge is at the origin and the test charge is at a distance \( y \) from the source charge on the \( y \)-axis as shown in Fig. 5.7. Transforming to the frame where they are both at rest
\[ F'_y = F_y = \frac{kq_t q_s}{\gamma y^2}. \] (5.15)

To avoid any confusion, recall that if the test charge were at rest then we would be multiplying by \( \gamma \) instead of dividing by \( \gamma \).

Then, what is the real magnetic force? We will define this as the difference between the force when the test charge is moving and the force when it is stationary. This can be calculated using
\[ F_m = k \frac{q_t q_s}{y^2} \left( \frac{1}{\gamma} - \gamma \right) = \gamma k \frac{q_t q_s}{y^2} \left( \frac{1}{\gamma^2} - \gamma \right) = \gamma k \frac{q_t q_s}{y^2} \left( \frac{-v^2}{c^2} \right). \] (5.16)
This implies that
\[ F_m = F_e \left( \frac{-v^2}{c^2} \right). \tag{5.17} \]
where \( F_e = \gamma k \frac{q_s q_t}{y^2} \).

This is the correct expression for what is referred to as **Magnetic Force**.

Suppose we consider the same situation as before but now the test charge has a velocity \( u_t \) and the source has a velocity \( v \) (refer to Fig. 5.8). If we use the equation for the force transformation this time we get a factor of \( 1 - \frac{vu_t}{c^2} \). That is
\[ F_y = \gamma k \frac{q_s q_t}{y^2} \left( 1 - \frac{vu_t}{c^2} \right). \tag{5.18} \]

In the case where the velocity \( u_t \) is in the positive \( \hat{y} \) direction as shown in Fig. ??, if we transform
to the reference frame where the source charge is at rest we would get (as expected) that

\[ F'_{y'} = F_y = \frac{k q_t q_s}{\gamma y^2} \]  (5.19)

but this time the test charge has a velocity in the \( \hat{x}' \) and the \( \hat{y}' \) direction. In particular we have

\[
\begin{align*}
    u'_x &= -v \\
    u'_y &= \frac{u_y}{\gamma} \\
    u'_z &= 0
\end{align*}
\]  (5.20)

which together gives us the following

\[
F_m = F_x = \gamma k \left( \frac{vu_t}{c^2} \right) \frac{q_s q_t}{y^2} 
\]  (5.21)

\[
F_e = F_y = \gamma k \frac{q_s q_t}{y^2}. 
\]  (5.22)

Finally, we consider the case where the source charge is at the origin with a velocity parallel to the \( x \)-axis, the test charge has an arbitrary position and a velocity \( u_t \) parallel to the \( y \)-axis as shown in Fig. 5.10. If we transform to a reference frame where the source charge is at rest we have that

\[
\begin{align*}
    \vec{F}'_{x'} &= \frac{k q_t q_s}{\gamma' y'^3} \vec{x}' \\
    \vec{F}'_{y'} &= \frac{k q_t q_s}{y'^2} \vec{y}' \\
    \vec{F}'_{z'} &= 0.
\end{align*}
\]  (5.23)

As in the previous case we have

\[
\begin{align*}
    u'_x &= -v \\
    u'_y &= \frac{u_y}{\gamma}
\end{align*}
\]
Using the force transformations we obtain

\begin{align*}
F_x &= \gamma \frac{kq_t q_s}{r'^3} \left( x + \frac{vu_t}{c^2} y \right) \\
F_y &= \gamma \frac{kq_t q_s}{r'^3} \\
F_z &= 0.
\end{align*}

We can immediately recognize the first summand of the \(x\)-component and the \(y\)-component of the force to be

\begin{equation}
\vec{F}_e = \frac{kq_t q_s}{((\gamma^2 x^2) + y^2)^\frac{3}{2}} \gamma \vec{r}
\end{equation}

\begin{equation}
\vec{F}_m = \gamma k q_s \frac{q_t}{r'^2} \left( \frac{v u_t}{c^2} \right).
\end{equation}

We can state the force due to a magnetic field in terms of what we learned in an introductory physics course as follows

\begin{equation}
\vec{F}_m = q_t \vec{u}_t \times \vec{B}.
\end{equation}

Since \(\vec{u}_t\) is parallel to the \(y\)-axis and \(\vec{F}_m\) is parallel to the \(x\)-axis, \(\vec{B}\) would have to be parallel to the \(z\)-axis. Also its magnitude would have to be given by

\begin{equation}
B = \frac{\gamma k q_s v_y}{r'^3}.
\end{equation}

We know \(v_y\) has the same magnitude as \(v \times r\) which also happens to be in the \(\hat{z}\) direction. Therefore, we can finally derive an elegant and, more importantly, correct expression for \(\vec{B}\)! We have that

\begin{equation}
\vec{B} = \frac{1}{c^2} \vec{v} \times \left( \frac{\gamma k q_s}{r'^3} \vec{r} \right)
\end{equation}

but the term in the parenthesis is equal to the electric field. Thus we obtain

\begin{equation}
\vec{B} = \frac{1}{c^2} \left( \vec{v} \times \vec{E} \right)
\end{equation}
where
\[ \vec{E} = \frac{\gamma k q_s}{r^3} \vec{r}. \]  \hspace{1cm} (5.32)

We started out discussing Maxwell’s Equations and then discussed equations for light waves and how they transform under different transformations. We have concluded this section by revisiting Coulomb’s Law and showing, finally, how one of the laws we originally thought fundamental actually changes relativistically. This is a very important and profound result of the concept of relativity.
Chapter 6

Black-Body Radiation - The Need for Quantization

This chapter will introduce the concept of quantization by analyzing spectra of energy emitted by a black-body as a function of temperature.

6.1 Energy radiated by a Black-body at a given temperature

One of the most revolutionary experiments in physics resulted in the discovery of black-body radiation. A black body is an opaque, non-reflective object with a uniform and constant temperature. The radiation that surrounds this object has a specific spectrum and intensity that changes with temperature. Black-bodies are by definition able to absorb energy from all frequencies and reflect none. Through the study of this process, the theory of quantization of energy was developed which is, perhaps, the most significant result of black-body radiation.

The most important result discovered about black bodies was presented in 1879 by the Austrian physicist Josef Stefan who found experimentally that the total energy radiated per unit volume at all frequencies by a black-body was not linearly proportional to the temperature as everybody thought at that time. Instead, he discovered the following relation

$$e_{\text{total}} = \int_0^\infty u(f, T) df = aT^4$$ \hspace{1cm} (6.1)

where $u(f, T)$ is the energy density at a given frequency and $a$ is the radiation constant. In order to understand the physics behind this expression, we need to understand the emission process of a black-body. In particular, we need to obtain a formula for $u(f, T)$ and determine whether or not it agrees with the Stefan-Boltzmann Law. We will first consider an approach based on Classical Statistical Physics, and observe what happens.

\footnote{The value of $a$ is $7.5657 \cdot 10^{-16} \text{Jm}^{-3} \text{K}^{-4}$. The radiation constant is related to the Stefan-Boltzmann constant by the relation $\sigma = \frac{a}{\pi}$. The factor of 4 is due to two things. If we think of a cylinder picture and we look from one side, the energy coming out of that side is only half of the total energy emitted because the other half goes the other way. The other factor of two comes from the fact that then not all of it is emitted straight but there is a direction to the emissions. We need the component that comes in our direction which involves a cosine. When we then talk about the intensity, we get a $\cos^2$ and integrating that over all directions gives the other factor of one half.}
6.2 Classical Approach

In order to find $u(f, T)$, we must first find expressions for the average energy per wave and the number of waves per unit volume with a given energy as a function of temperature. Therefore, we will need Boltzmann’s Law from statistical physics which states

$$P(E) \, dE \propto e^{\frac{-E}{k_B T}} \, dE$$  \hspace{1cm} (6.2)

where $P(E)$ is the probability of finding a wave or particle in a system on the the interval $[E, E+dE]$, and $k_B$ is the Boltzmann constant whose value is $1.3806488 \cdot 10^{-23} \text{m}^2 \cdot \text{kg} \cdot \text{s}^{-2} \cdot \text{K}^{-1}$. This formula involves an integral thus assuming that the energy spectra is continuous. In other words, it acts under the assumption that the energy for a wave is continuous. The previous equation can then be written as

$$P(E) \, dE = N \, e^{\frac{-E}{k_B T}} \, dE$$  \hspace{1cm} (6.3)

where $N$ is a normalization\footnote{Remember that the probability of finding a particle at all energies should be equal to 1.} constant so that

$$\int_0^\infty P(E) \, dE = 1.$$  \hspace{1cm} (6.4)

Therefore,

$$\int_0^\infty N \, e^{\frac{-E}{k_B T}} \, dE = 1$$

$$N \int_0^\infty e^{\frac{-E}{k_B T}} \, dE = 1$$

$$N \, k_B T = 1 \quad \longrightarrow \quad N = \frac{1}{k_B T}$$  \hspace{1cm} (6.5)

Recall from classical electrodynamics that the energy of a wave is proportional to the square of the magnitude of the electric field and the square of the magnitude of the magnetic field

$$e(x, t) = \frac{1}{2} \varepsilon_0 E^2(x, t) + \frac{1}{2\mu_0} B^2(x, t)$$  \hspace{1cm} (6.6)

where $e$ is the energy of the wave. Using (1.73) and (1.74), we obtain

$$e(x, t) = \frac{1}{2} \varepsilon_0 E_0^2 + \frac{1}{2\mu_0} \frac{E_0^2}{c^2} = \varepsilon_0 E_0^2.$$  \hspace{1cm} (6.7)

Interestingly enough, we find that the energy of the wave does not depend on $x$ or $t$ but that it is actually a constant equal to the square of the magnitude of the electric field! This comes from the fact that we chose to model the electric and magnetic field waves as being circularly polarized. The nice property of circular polarization is that whereas the direction of the wave changes with time, the magnitude of the wave stays a constant. We can see from this equation that any energy is allowed since $E_0$ can acquire any value. Thus, we can use the Stefan-Boltzmann Law to evaluate the energy per wave.
The average energy or expectation value for the energy is defined as

\[ \langle E \rangle = \int_0^\infty EP(E)dE \quad (6.8) \]

or

\[ \langle E \rangle = \frac{\int_0^\infty EP(E)dE}{\int_0^\infty P(E)dE} \quad (6.9) \]

where \( \langle E \rangle \) is the expectation value for the energy. Applying the Stefan-Boltzmann Law as in \( 6.3 \),

\[ \langle E \rangle = \frac{\int_0^\infty NEe^{\frac{-E}{k_BT}}dE}{\int_0^\infty Ne^{\frac{-E}{k_BT}}dE} \quad (6.10) \]

Since \( N \) has no dependence on \( E \) we can take it out of the integral on top and bottom and it cancels. Then making the substitution \( x = \frac{E}{k_BT} \), we obtain

\[ \langle E \rangle = k_BT\frac{\int_0^\infty xe^{-x}dx}{\int_0^\infty e^{-x}dx}. \quad (6.11) \]

Both integrals are gamma functions\(^3\) so

\[ \langle E \rangle = k_BT\frac{\Gamma(2)}{\Gamma(1)} \quad (6.12) \]

which reduces to

\[ \langle E \rangle = k_BT \quad \text{(per wave)}. \quad (6.13) \]

Thus regardless of the frequency of the wave, the average energy is linearly proportional to the temperature. Notice that this means the average energy is independent of the frequency of the wave. This fact, that \textbf{every wave regardless of frequency carries the same energy} is referred to as the \textbf{Equipartition Theorem}.\(^4\) Now we need to obtain an expression for the number of waves at a given frequency or at a given wavelength per unit volume. Let us consider the electric field component of a plane wave written in the following form,

\[ \vec{E}(\vec{r},t) = \cos \left( \frac{2\pi f}{c} (\hat{n} \cdot \vec{r} - ct) \right) \hat{k} \quad (6.14) \]

where \( f \) is the frequency, \( \hat{n} \) is a unit vector in the direction of the wave, and \( \vec{r} \) the position vector of a point on the wave. Consider the \( n \)-space of the wave that is shown in Fig. \[ 6.1 \]. It is easily observed that the number of electromagnetic waves between \( \vec{n} \) and \( \vec{n} + d\vec{n} \) is proportional to the volume of the spherical shell

\[ N(n)dn = 4\pi n^2dn \quad (6.15) \]

\(^3\)Recall that the gamma function for natural numbers has the property \( \Gamma(n) = (n - 1)! \).

\(^4\)What we have here is actually the relativistic version of the equipartition theorem. Normally we say that \( \langle E \rangle = \frac{1}{2}k_BT \) per particle per polarization. Non-relativistically we have only one polarization but as we have discussed, in the relativistic case we have two polarizations. As a result that \( \frac{1}{2} \) becomes a 1.
Figure 6.1: n-space of the electromagnetic wave

$dn$ represents the thickness of the spherical shell between $\vec{n}$ and $\vec{n} + \vec{dn}$

or in terms of frequency we have

$$N(f)df = \frac{4\pi f^2}{c^3} df.$$  \hspace{1cm} (6.16)

However, we have to account for the extra factor of two that comes from the fact that the electromagnetic wave can have left or right circular polarization. Therefore,

$$N(f)df = \frac{8\pi f^2}{c^3} df.$$ \hspace{1cm} (6.17)

Notice that $N(f)df$ has units of inverse volume. Then, the energy density per frequency is the average energy per wave times the number of waves with that frequency

$$u(f, T)df = \langle E \rangle N(f)df.$$ \hspace{1cm} (6.18)

Using (6.13) and (6.17), we obtain

$$u(f, T)df = k_B T \frac{8\pi f^2}{c^3} df.$$ \hspace{1cm} (6.19)

Integrating over all frequencies this becomes

$$T \int_0^\infty k_B \frac{8\pi f^2}{c^3} df = \infty.$$ \hspace{1cm} (6.20)

Notice that the coefficient, represented by the integral, diverges and the dependence on $T$ is not as predicted. Therefore, there must be something wrong with the assumption that any energy is allowed. This result is known as the **ultraviolet catastrophe** because the energy density increases without bound on the ultraviolet region of the electromagnetic spectrum. Feynman usually refers to this as the failure of classical physics, because the energy cannot be infinity.
6.3 Quantum Approach

In order to fix the problem, we have to restrict the energy values that an electromagnetic wave can possess. We will propose that the energy of a wave is proportional to its frequency in the following way

\[ E = n hf \] (6.21)

where \( n \) is a positive integer. This time, we cannot use an integral to find the average energy per wave since the energy values are no longer continuous. Instead, an infinite sum is required. This process can be thought of as dividing the possible energies into bins where each bin has length \( hf \)

\[ \langle E \rangle = \frac{\sum_{n=0}^{\infty} EP(E)\Delta E}{\sum_{n=0}^{\infty} P(E)\Delta E}. \] (6.22)

Using (6.3) and (6.21),

\[ \langle E \rangle = \frac{\sum_{n=0}^{\infty} Nnhf e^{-\frac{hf}{k_BT}} hf}{\sum_{n=0}^{\infty} Ne^{-\frac{hf}{k_BT}} hf}. \] (6.23)

As we did before, \( N \) should be such that

\[ \sum_{n=0}^{\infty} Ne^{-\frac{nhf}{k_BT}} hf = 1 \] (6.24)

to obey normalization conditions. Since \( N \) and \( hf \) are not involved in the sum, they can be separated

\[ Nh \sum_{n=0}^{\infty} e^{-\frac{nhf}{k_BT}} = 1. \] (6.25)

Now we let \( s = \frac{hf}{k_BT} \). Then, the infinite series becomes an infinite geometric series of the form

\[ \sum_{n=0}^{\infty} e^{-\frac{nhf}{k_BT}} = \sum_{n=0}^{\infty} e^{-sn} = \frac{1}{1 - e^{-s}} = \frac{1}{1 - e^{-\frac{hf}{k_BT}}}. \] (6.26)

Substituting this expression into (6.25) we have

\[ \frac{Nh f}{1 - e^{-\frac{hf}{k_BT}}} = 1 \quad \rightarrow \quad N = \frac{1 - e^{-\frac{hf}{k_BT}}}{hf}. \] (6.27)

We can now calculate the limit \( h \to 0 \) of this expression using L'Hôpital’s rule

\[ \lim_{h \to 0} \frac{1 - e^{-\frac{hf}{k_BT}}}{hf} = \lim_{h \to 0} \frac{f e^{-\frac{hf}{k_BT}}}{k_BT} = \frac{1}{k_BT}. \] (6.28)

So,

\[ \lim_{h \to 0} N = \frac{1}{k_BT}. \] (6.29)

\( ^{5} \Delta E = hf \) is the width of each bin.
This equation says that as the energy bin becomes smaller and smaller, we return to the continuous classical case. Now we return to solving (6.23). Canceling the normalization constants and some factors of $hf$

$$\langle E \rangle = hf \sum_{n=0}^{\infty} \frac{ne^{-\frac{n hf}{k_B T}}}{\sum_{n=0}^{\infty} e^{-\frac{n hf}{k_B T}}}.$$  

(6.30)

The infinite sum in the denominator is the infinite series we encountered in (6.26). To calculate the numerator we will use the following mathematical trick.

**A Math Trick:** Take the derivative with respect to $s$ of the third element in (6.26)

$$\frac{\partial}{\partial s} \sum_{n=0}^{\infty} e^{-sn} = \frac{\partial}{\partial s} \left( \frac{1}{1 - e^{-s}} \right).$$  

(6.31)

Applying term-wise differentiation on the left hand side, and evaluating the derivative on the right hand side we obtain

$$\sum_{n=0}^{\infty} -ne^{-sn} = \frac{-e^{-s}}{(1 - e^{-s})^2}. $$  

(6.32)

Therefore,

$$\sum_{n=0}^{\infty} -ne^{-\frac{n hf}{k_B T}} = \frac{-e^{-\frac{hf}{k_B T}}}{(1 - e^{-\frac{hf}{k_B T}})^2}. $$  

(6.33)

Combining (6.33) and (6.30)

$$\langle E \rangle = hf \left( 1 - e^{-\frac{hf}{k_B T}} \right) \frac{e^{\frac{hf}{k_B T}}}{\left( 1 - e^{\frac{hf}{k_B T}} \right)^2}. $$  

(6.34)

Simplifying,

$$\langle E \rangle = hf \frac{e^{\frac{hf}{k_B T}}}{1 - e^{\frac{hf}{k_B T}}}. $$  

(6.35)

Finally multiplying the numerator and denominator by $e^{\frac{hf}{k_B T}}$ we obtain

$$\langle E \rangle = \frac{hf}{e^{\frac{hf}{k_B T}} - 1}. $$  

(6.36)

If we take the limit as $h \to 0$, then we will return to (6.13).

$$\lim_{h \to 0} \frac{hf}{e^{\frac{hf}{k_B T}} - 1} = k_B T. $$  

(6.37)

Note that our calculations are consistent with what we did before. However, we will not take the limit since it has been pointed out that then the energy for all frequencies will diverge. Therefore, the new energy density function per frequency at a given temperature using (6.17) and (6.36) is

$$u(f,T) = \frac{hf}{e^{\frac{hf}{k_B T}} - 1}. $$  

(6.38)
which is known as Planck’s Law of radiation. We can perform the integration over all frequencies,

\[ e_{\text{total}} = \int_0^\infty \frac{hf}{e^{hf/k_BT} - 1} \frac{8\pi f^2 df}{c^3}. \] (6.39)

Separating out the constants and making the change of variables \( u = \frac{hf}{k_BT} \),

\[ e_{\text{total}} = \frac{8\pi h}{c^3} \int_0^\infty \frac{f^3 df}{e^{hf/k_BT} - 1} = \frac{8\pi h}{c^3} \left( \frac{k_BT}{h} \right)^4 \int_0^\infty \frac{u^3}{e^u - 1} du. \] (6.40)

It will be stated without proof\(^6\) that

\[ \int_0^\infty \frac{u^3}{e^u - 1} du = \frac{\pi^4}{15}. \] (6.41)

It follows that

\[ e_{\text{total}} = \frac{8\pi h}{c^3} \left( \frac{k_BT}{h} \right)^4 \frac{\pi^4}{15} \] (6.42)

which can be simplified to

\[ e_{\text{total}} = \frac{8\pi^5 k_B^4}{15h^3 c^3} T^4. \] (6.43)

This expression models the relation of energy density emitted as a function of temperature that was initially proposed by the Stefan-Boltzmann Law, where

\[ \frac{8\pi^5 k_B^4}{15h^3 c^3} = a. \] (6.44)

Solving for \( h \) we obtain

\[ h = \left( \frac{8\pi^5 k_B^4}{15hc^3} \right)^{1/3}. \] (6.45)

After substituting the corresponding numerical values we obtain \( h = 6.6261 \cdot 10^{-34} \text{J} \cdot \text{s} \). This is one of the most profound results in physics, and it comes from the observation of a black body. It implies that light is quantized; its energy can only take integer multiples of \( hf \). In the same manner, momentum is quantized because its energy can only take integer multiples of \( \frac{hf}{c} \). It is important to mention that the assumption that the energy of light is linearly proportional to its frequency is correct. This assumption is supported by the Stefan-Boltzmann Law and other experimental data such as Wien’s Displacement Law. For instance, energy density emitted as given by (6.38) can be written in terms of \( \lambda \) as

\[ u(\lambda, T) = \frac{8\pi h c}{\lambda^5} \frac{1}{e^{\frac{h c}{k_B T \lambda}} - 1}. \] (6.46)

The plot shown in Fig. 6.2 shows the wavelength dependence of energy density for a given temperature. Furthermore, the energy density has a maximum wavelength at fixed temperature when the wavelength derivative of (6.46) is zero, (as we know from elementary calculus)

\[ \frac{\partial u(\lambda, T)}{\partial \lambda} = 0. \] (6.47)

\(^6\)This integral uses Fourier Transforms and Fourier Series and thus will not be worked out in detail here. It is a rather tedious calculation.
Figure 6.2: Energy density emitted by a blackbody.

Notice how the peak shifts toward smaller values as the temperature increases! This is known as Wien’s Law.

Taking the derivative of the function we obtained,

$$8\pi hc \left( \frac{hc}{k_B T \lambda^4} \left( e^{\frac{hc}{k_B T \lambda}} - 1 \right)^2 - \frac{1}{\lambda^6} \frac{5}{e^{\frac{hc}{k_B T \lambda}} - 1} \right) = 0. \quad (6.48)$$

After, some simplifications,

$$\frac{hc}{\lambda k_B T} \frac{e^{\frac{hc}{k_B T \lambda}}}{e^{\frac{hc}{k_B T \lambda}} - 1} - 5 = 0. \quad (6.49)$$

Solving this numerically results in

$$\lambda_{\text{max}} = \frac{2.8978 \times 10^{-3} \text{m} \cdot \text{K}}{T} \quad (6.50)$$

which is exactly what Wien found experimentally.

Finally, if we think of the energy of the wave carried by \(n\) bundles of energy, then the energy per bundle (or quanta) is given by

$$E = hf \quad \text{(per quanta)}. \quad (6.51)$$

Similarly, the momentum per quanta is given by

$$p = \frac{h}{\lambda} \quad \text{(per quanta)}. \quad (6.52)$$
We can find the energy per quanta that carries radio waves, visible light and x-rays at some frequency. For radio waves \((f_r = 3.00 \cdot 10^9\text{Hz})\)

\[
E_r = 1.99 \cdot 10^{-24}\text{J}. \quad (6.53)
\]

For visible light \((f_v = 7.50 \cdot 10^{14}\text{Hz})\)

\[
E_v = 4.97 \cdot 10^{-19}\text{J}. \quad (6.54)
\]

For x-rays \((f_x = 7.50 \cdot 10^{19}\text{Hz})\)

\[
E_x = 1.99 \cdot 10^{-14}\text{J}. \quad (6.55)
\]

These are intuitive frequencies and we find that the bin-width, \(hf\), is extremely small. Whereas we found that \(h\) cannot be taken to 0, we know that for everyday frequency values \(hf\) is incredibly tiny. This conclusion that energy and momentum are quantized, of \(h\) not being able to be taken to 0 but only to very small values are both ideas that fed the idea of quantum mechanics. Quantum Mechanics was founded with the hope of being able to answer the questions that Classical Mechanics clearly didn’t have a plausible answer for.
Setting Up Wave Mechanics

There is enough experimental data to support the model that describes subatomic particles such as electrons behaving as waves and not as classical particles (rigid balls). One experiment done to this effect was the diffraction of electrons. This section deals with the mechanics of waves in order to understand the behavior at the atomic level.

7.1 Interference of Light Waves

We first introduce the interference of light waves coming from multiple correlated sources. This means that light is emitted from each source simultaneously. Consider $2N+1$ sources of light located in such a way that the distance of each source to the detector is equally spaced and increasing in steps of $\delta$ symmetrically as illustrated in Fig. 7.1. The waves emitted by the sources can be approximated as plane waves in the vicinity of the detector. Therefore the electric field of each wave has the form

$$\vec{E}(x,t) = \vec{E}_0 \cos(kx - \omega t). \tag{7.1}$$

In order to simplify the calculations, we will work with complex exponentials since they have the nice property

$$\cos(kx - \omega t) = \text{Re}(e^{i(kx-\omega t)}) \tag{7.2}$$

where $\text{Re}(z)$ denotes the real part of a complex number $z$. Then 7.1 can be written as

$$\vec{E}(x,t) = \vec{E}_0 \text{Re}\left(e^{i(kx-\omega t)}\right). \tag{7.3}$$

Since each source is at a different distance from the detector, the electric field for each source will be phase shifted. The phase shift is related to the distance by

$$\phi = \frac{2\pi \delta}{\lambda} \tag{7.4}$$

where $\lambda$ is the wavelength of light coming from the source. Then, the electric field that arrives at the detector is given by

$$E_n(x,t,\phi) = \text{Re}\left(e^{i(kx-\omega t + \phi_n)}\right) \tag{7.5}$$

which can be written as

$$E_n(x,t,\phi) = \text{Re}\left(e^{i(kx-\omega t)}\right) \text{Re}\left(e^{i\phi_n}\right). \tag{7.6}$$
Figure 7.1: Interference of light coming from $2N + 1$ correlated sources.

The amplitude of the electric field at the detector is given by the sum of amplitudes of the electric fields of the waves coming from each source

$$E_0(\phi) = \text{Re} \left( \sum_{n=-N}^{N} e^{i\phi n} \right).$$

(7.7)

Let us calculate the following geometric sum,

$$\sum_{n=1}^{N} e^{i\phi n} = e^{i\phi} \left( \frac{1 - e^{iN\phi}}{1 - e^{i\phi}} \right).$$

(7.8)

Thus, the closed form of (7.7) is

$$E_0(\phi) = 1 + e^{i\phi} \left( \frac{1 - e^{iN\phi}}{1 - e^{i\phi}} \right) + e^{-i\phi} \left( \frac{1 - e^{-iN\phi}}{1 - e^{-i\phi}} \right).$$

(7.9)

Expanding the above equation and making common denominators we obtain

$$E_0(\phi) = \frac{1 - e^{i(N+1)\phi}}{1 - e^{i\phi}} + \frac{e^{-i\phi} - e^{-i(N+1)\phi}}{1 - e^{-i\phi}}$$

$$= \frac{1 - e^{i\phi}}{1 - e^{i\phi}} \cdot \frac{e^{-i(N+\frac{1}{2})\phi} - e^{i(N+\frac{1}{2})\phi}}{e^{-i\phi} - e^{i\phi}}$$

$$= \frac{\sin((N + \frac{1}{2})\phi)}{\sin(\frac{\phi}{2})}.$$ 

(7.10)

Finally, the amplitude of the intensity of light detected is the square of the magnitude of the electric field,

$$I_0(\phi) = \left( \frac{\sin((N + \frac{1}{2})\phi)}{\sin(\frac{\phi}{2})} \right)^2.$$ 

(7.11)
If we graph this intensity function, as we have below, we see there are several places where the intensity (amplitude) goes to 0. These points are what we know from classical optics as **destructive interference**. Likewise, there are visible places of **constructive interference** as well. Each plot shows the intensity function for various values of $N$, i.e. the number of sources. Each graph has a peak at $\phi = 0$ where the intensity looks like $(2N + 1)^2$. (This can be figured out analytically through the use of derivatives and limits.) We can see that as the number of sources increases to be very large, the intensity function gets closer and closer to having one large peak at $\phi = 0$. Physically, this means that as we take the number of sources to be larger and larger, what is received by the detector begins to look more and more like a point source or a localized object. It has been observed that electrons behave the same way as light. That they have the same distribution for a multiple number of sources was proven in an experiment about electron diffraction from crystals.

### 7.2 Wave Mechanics - A Motivation

Now let us look at the non-relativistic version of Newton’s Second Law

$$ F = ma $$

or

$$ F = \frac{dp}{dt}. $$

This equation is the what we used in classical physics to describe the behavior of a particle. In order to solve this equation we need a formula for the force like Coulomb’s Law $\frac{kq_1q_2}{r^2}$ or the Universal Law of Gravitation $\frac{Gm_1m_2}{r^2}$. At this point we would be able to solve the problem but we can touch upon something a bit nicer. If we have a conservative force, which we usually do in nature, we know that the work done by this force is independent of the path taken. As a result we know that the force can be written as the negative gradient of some scalar function which we will call the potential that obeys $\vec{F} = -\nabla U(x)$. In this particular case it makes sense to define a quantity we can call as the total energy of a system

$$ E = K + U. $$

It then follows rather simply from this that total energy is conserved, $\frac{dE}{dt} = 0$.

---

**Relation between Energy and Force:** Total energy in one dimension can be written as

$$ E(x) = \frac{p^2}{2m} + U(x). $$

We can take the time derivative of this equation

$$ \frac{dE(x)}{dt} = \frac{2p}{2m} \frac{dp}{dt} + \frac{dU}{dx} \frac{dx}{dt} $$

which, non-relativistically, can be written as

$$ \frac{dE(x)}{dt} = vF + \frac{dU}{dx}v. $$

---

1We are assuming that $U$ only depends on $x$. 

84
Since the field is conservative, we know that \( \frac{dU}{dx} = -F \). Using this relation and (7.17) we conclude
\[
\frac{dE(x)}{dt} = 0 \tag{7.18}
\]
which implies that energy is conserved.

We can talk about this in connection with wave mechanics. We will first consider a specific case. As we found previously, the equation that describes light for a plane wave is
\[
\vec{E}(x, y, z, t) = \vec{E}_0 e^{i(k_x x + k_y y + k_z z - \omega t)} \tag{7.19}
\]
and from the results of black body radiation and \( E = pc \) from relativity, we know that the energy and momentum for a plane wave are given by
\[
E = \hbar \omega \tag{7.20}
\]
\[\vec{p} = \hbar \vec{k} \tag{7.21}\]
where \( k^2 = k_x^2 + k_y^2 + k_z^2 \), and \( \hbar = \frac{h}{2\pi} \). We need to find some way to get energy and momentum from (7.19). If we associate the following operators with energy and momentum
\[
\hat{E} = i\hbar \frac{\partial}{\partial t} \tag{7.22}
\]
\[
\hat{p} = -i\hbar \vec{\nabla} \tag{7.23}
\]
then it is possible to obtain expressions for energy and momentum from (7.19)
\[
\hat{E}\vec{E}(x, y, z, t) = \hbar \omega \vec{E}(x, y, z, t) \tag{7.24}
\]
\[
\hat{p}\vec{E}(x, y, z, t) = \hbar \vec{k}\vec{E}(x, y, z, t) \tag{7.25}
\]
It simply follows that
\[
\hat{p}^2 = -\hbar^2 \vec{\nabla} \cdot \vec{\nabla} \tag{7.26}
\]
We know that for conservative forces, energy is conserved when we are considering particles. In the case of wave mechanics, we consider the operators and we say that they must abide by the same relation. This is known as the Schrödinger Equation. We considered the simple case where we discussed a plane wave. We can postulate that this is the case for any wave. Then, the equation that describes the mechanics of waves can be obtained by combining (7.15), (7.22) and (7.26)
\[
-\frac{\hbar^2}{2m} \vec{\nabla} \cdot \vec{\nabla} \Psi(x, y, z, t) + U(x)\Psi(x, y, z, t) = i\hbar \frac{\partial \Psi(x, y, z, t)}{\partial t} \tag{7.27}
\]

\[\text{Note that the energy operator is Hermitian which means that all of its eigenvalues are real.}\]
\[\text{It might seem like there should be dot products in the relations with the momentum operator here. This is not the case because the momentum operator actually acts individually on each component of the electric field vector.}\]
\[\text{This equation is a PDE. In particular, it is a heat equation with complex diffusivity. It is known that the solution to this equation plus an initial condition (Cauchy Problem) is unique.}\]
\[\text{Even though it is not so clear to talk about the mass in terms of wavefunction, you might consider mass as a parameter.}\]
We can then discuss various properties of these fairly arbitrary wave functions we described. Recall from electrodynamics that the equation describing the electric field is related to the intensity of the wave as follows

\[ I(x, t) \propto \vec{E}(x, t) \cdot \vec{E}(x, t). \quad (7.28) \]

Analogous to that statement, the probability density of a particle being at a position at a specific time is

\[ P(x, t) = \overline{\Psi}(x, t) \Psi(x, t) \quad (7.29) \]

where \overline{\Psi}(x, t) = |\Psi(x, t)|^2. (7.30)

This guarantees that the probability density has a positive real output. We also know that since the intensity must be bounded, the operators must act only on bounded functions. In particular we will say that \Psi(x, y, z, t) vanishes to 0 as \(x, y, z, t\) go to \(\infty\). Before we continue it is important to note the square integrability of the wave function.

**Square Integrability:** Let us consider the Schrödinger Equation where the potential function\(^6\) we consider is real and independent of time. First, we multiply the Schrödinger Equation (7.27) by \(\Psi(x, y, z, t)\),

\[-\frac{\hbar^2}{2m} \Psi(x, y, z, t) \frac{\partial^2 \Psi(x, y, z, t)}{\partial x^2} + \Psi(x, y, z, t)V(x)\Psi(x, y, z, t) = i\hbar \frac{\partial \Psi(x, y, z, t)}{\partial t}. \]

(7.31)

Taking the complex conjugate of this equation and using the fact that the complex conjugate of a derivative is the derivative of the complex conjugate of the function\(^8\),

\[-\frac{\hbar^2}{2m} \Psi(x, y, z, t) \frac{\partial^2 \Psi(x, y, z, t)}{\partial x^2} + \Psi(x, y, z, t)V(x)\Psi(x, y, z, t) + i\hbar \Psi(x, y, z, t) \frac{\partial \Psi(x, y, z, t)}{\partial t}. \]

(7.32)

Subtracting (7.32) from (7.31), we get

\[-\frac{\hbar^2}{2m} \left( \frac{\partial^2 \Psi}{\partial x^2} - \frac{\partial^2 \overline{\Psi}}{\partial x^2} \right) = i\hbar \left( \frac{\partial \overline{\Psi}}{\partial t} - \frac{\partial \Psi}{\partial t} \right). \]

(7.33)

Integrating both sides from \(-\infty\) to \(\infty\),

\[-\frac{\hbar^2}{2m} \int \int \int \left( \frac{\partial^2 \Psi}{\partial x^2} - \frac{\partial^2 \overline{\Psi}}{\partial x^2} \right) dx dy dz = i\hbar \int \int \int \left( \frac{\partial \overline{\Psi}}{\partial t} - \frac{\partial \Psi}{\partial t} \right) dx dy dz. \]

(7.34)

Using integration by parts (for one of the derivatives because the other two variables behave the same way), the left hand side of (7.34) becomes

\[-\frac{\hbar^2}{2m} \left[ \frac{\partial \overline{\Psi}}{\partial x} \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{\partial \overline{\Psi}}{\partial x} \frac{\partial \Psi}{\partial x} dx - \frac{\hbar^2}{2m} \left[ \frac{\partial \Psi}{\partial x} \right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{\partial \Psi}{\partial x} \frac{\partial \overline{\Psi}}{\partial x} dx. \]

(7.35)

Using the fact that the wavefunctions vanish to 0 as \(x \to \infty\), the first and fourth terms are zero. Also, since integration is commutative, the second and third terms cancel each other out. Then, the right hand side of (7.34) is zero. Therefore,

\[i\hbar \int \int \int \left( \frac{\partial \overline{\Psi}}{\partial t} - \frac{\partial \Psi}{\partial t} \right) dx dy dz = 0. \]

(7.36)

\(^6\)Recall the notation that \(\overline{\Psi}(x, t)\) is the complex conjugate of \(\Psi(x, t)\).

\(^7\)Notice that we can then use \(\overline{V}(x) = V(x)\) since \(V(x)\) is real valued.

\(^8\)This can be proven using the definition of derivative.
Applying the product rule to the integrand,

\[ i\hbar \int \int \int \frac{\partial}{\partial t} [\bar{\Psi} \Psi] \, dxdydz = 0. \tag{7.37} \]

Interchanging the order of differentiation and integration according to the Leibniz Rule we have

\[ i\hbar \frac{d}{dt} \int \int \int \bar{\Psi} \Psi \, dxdydz = 0. \tag{7.38} \]

Thus we obtain

\[ \int \int \int \bar{\Psi} \Psi \, dxdydz = \text{constant.} \tag{7.39} \]

Since \( \Psi \) is square integrable, i.e. the integral of the square of \( \Psi \) is finite, the following normalization condition can be met

\[ \int_{-\infty}^{\infty} |\Psi(x,t)|^2 \, dx = 1 \tag{7.40} \]

since \( P(x,t) \) represents probability density. This means that the total probability of finding a particle somewhere (rather than at a specific position) is equal to 1. In general, the expectation value associated with an operator is given by

\[ \langle O(t) \rangle = \int \int \int \bar{\Psi}(x,y,z,t) \hat{O} \Psi(x,y,z,t) \, dxdydz \tag{7.41} \]

where \( \hat{O} \) is an operator acting on \( \Psi(x,t) \) and \( \langle O \rangle \) is the associated expectation value of \( \hat{O} \). For example, the expectation value of \( x \) is calculated as follows

\[ \langle x(t) \rangle = \int_{-\infty}^{\infty} \bar{\Psi}(x,t) x \Psi(x,t) \, dx \tag{7.42} \]

and the expectation value for momentum is

\[ \langle p(t) \rangle = \int_{-\infty}^{\infty} \bar{\Psi}(x,t) \left[ -i\hbar \frac{\partial}{\partial x} \Psi(x,t) \right] \, dx. \tag{7.43} \]

Moreover, since we are discussing probability distributions, the associated standard deviation is given by

\[ \Delta O(t) = \sqrt{\langle O^2(t) \rangle - \langle O(t) \rangle^2}. \tag{7.44} \]

All of these ideas will be useful in understanding the behavior of different systems of particles that are influenced by a potential. Proceed to use several concepts we have just discussed to understand the behavior of a free particle at rest.
Chapter 8

Quantum Mechanics of a Free Particle at Rest

The aim of this section is to understand the behavior of a free particle \((U(x) = 0)\) at rest \((p = 0)\). From the previous section, we know that the particle should obey this relation

\[
-\frac{\hbar^2}{2m} \frac{\partial^2 \Psi(x,t)}{\partial x^2} = i\hbar \frac{\partial \Psi(x,t)}{\partial t}.
\]

(8.1)

Associated with any physical system, there is usually an initial condition. From a physical standpoint the initial condition should have the functional form of a Gaussian. For instance if we can measure the position of the particle at \(t = 0\), we will obtain a histogram of the position of the particle. Statistically, for a large number of measurements \((\approx 10^6)\) this histogram will look like a Gaussian distribution. Thus, the initial condition for a free particle can be given by

\[
\Psi(x,0) = \left(\frac{1}{2\pi\sigma^2}\right)^{1/4} e^{-\frac{x^2}{4\sigma^2}}.
\]

(8.2)

If this initial condition represents a probability distribution then it should follow the normalization condition given by (7.40).

Does it obey the normalization criteria? We can check that as follows

\[
\int_{-\infty}^{\infty} |\Psi(x,0)|^2 \, dx = \left(\frac{1}{2\pi\sigma^2}\right)^{1/2} \int_{-\infty}^{\infty} e^{\frac{-t^2}{4\sigma^2}} \, dt.
\]

(8.3)

In order to solve the indefinite integral, we will need to define a function \(I(\alpha)\)

\[
I(\alpha) = \int_{-\infty}^{\infty} e^{-\alpha x^2} \, dx, \quad I(\alpha) > 0, \quad \alpha > 0.
\]

(8.4)

Squaring both sides we obtain

\[
I^2(\alpha) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\alpha(x^2+y^2)} \, dx \, dy.
\]

(8.5)
Then switching from Cartesian to Polar coordinates and separating integrals,
\[ I^2(\alpha) = \int_0^{2\pi} d\theta \int_0^\infty e^{-\alpha r^2} r dr. \] (8.6)
This is easier to solve and we can then obtain by a simple integration that
\[ I^2(\alpha) = \frac{\pi}{\alpha} \rightarrow I(\alpha) = \sqrt{\frac{\pi}{\alpha}} = \int_{-\infty}^{\infty} e^{-\alpha x^2} dx. \] (8.7)
Therefore, (8.3) becomes
\[ \int_{-\infty}^{\infty} |\Psi(x,0)|^2 dx = \left( \frac{1}{2\pi\sigma^2} \right)^{1/2} (2\pi\sigma^2)^{1/2} = 1 \] (8.8)
satisfying the normalization criteria.

Next, we can find the expectation value for the position at \( t = 0 \). Using (7.41), we obtain
\[ \langle x(0) \rangle = \int_{-\infty}^{\infty} x \left( \frac{1}{2\pi\sigma^2} \right)^{1/2} e^{-\frac{x^2}{2\sigma^2}} dx. \] (8.9)
Noticing that the integral is symmetric about the origin and that the integrand is an odd function we can easily conclude
\[ \langle x(0) \rangle = 0. \] (8.10)
This shows that our chosen initial condition represents a particle lying at the origin (\( x = 0 \)).

Now we can calculate the error in measurement of position (standard deviation) using (7.44),
\[ \Delta x(0) = \sqrt{\langle x^2(0) \rangle + \langle x(0) \rangle^2} = \sqrt{\langle x^2(0) \rangle}. \] (8.11)
We can evaluate the expectation value of \( x^2 \)
\[ \langle x^2(0) \rangle = \left( \frac{1}{2\pi\sigma^2} \right)^{1/2} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2\sigma^2}} dx. \] (8.12)
Since the integrand is even we can simplify the integral to
\[ \langle x^2(0) \rangle = 2 \left( \frac{1}{2\pi\sigma^2} \right)^{1/2} \int_{0}^{\infty} x^2 e^{-\frac{x^2}{2\sigma^2}} dx. \] (8.13)
Using a change of variables \( u = \frac{x^2}{2\sigma^2} \) we obtain
\[ \langle x^2(0) \rangle = \frac{2\sigma^2}{\sqrt{\pi}} \int_{0}^{\infty} \sqrt{u} e^{-u} du. \] (8.14)
Recognizing that the integral is a gamma function and that \( \Gamma \left( \frac{3}{2} \right) = \frac{\sqrt{\pi}}{2} \),
\[ \langle x^2(0) \rangle = \frac{2\sigma^2}{\sqrt{\pi}} \Gamma \left( \frac{3}{2} \right) = \sigma^2. \] (8.15)
Substituting this back into (8.11) we conclude that
\[ \Delta x(0) = \sigma. \] (8.16)
Therefore, the intrinsic uncertainty of the particle\(^1\) is \( \sigma \). Furthermore, if we take \( \sigma \) and divide by \( \sqrt{N} \) we obtain the error in the measurement of the average position. It is then evident that the more measurements we make, the more accurate the average becomes. In order to understand the evolution of the wavefunction describing the particle, we need to solve (8.1). It is easy to show that a solution for (8.1) is of the form
\[ \Psi(x, t) = e^{i(kx - \omega t)}. \] (8.17)

**Confirmation of Solution:** Taking the second derivative with respect to \( x \)
\[ \frac{\partial^2 \Psi(x, t)}{\partial x^2} = -k^2 \Psi(x, t). \] (8.18)
The derivative with respect to \( t \) is
\[ \frac{\partial \Psi(x, t)}{\partial t} = -i\omega \Psi(x, t). \] (8.19)
Then (8.1) implies that
\[ \frac{\hbar^2 k^2}{2m} = \hbar \omega \quad \rightarrow \quad \omega = \frac{\hbar k^2}{2m}. \] (8.20)
Substituting this relation back into (8.17),
\[ \Psi_k(x, t) = e^{i(kx - \frac{\hbar k^2 t}{2m})}. \] (8.21)
We use the subscript \( k \) since this solution is valid for any value of \( k \).

Before analyzing the evolution of the wave we consider the set of solutions at \( t = 0 \) given by
\[ \Psi_k(x, 0) = e^{ikx}. \] (8.22)
As we discerned earlier we have an initial condition to be satisfied given by (8.2). This initial condition is satisfied by a linear combination of \( k \)-dependent solutions. Since any real value of \( k \) is allowed it is not enough to write this linear combination as a discrete sum. Instead, the general solution\(^2\) to (8.1) is an integral\(^3\)
\[ \Psi(x, 0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{\Psi}(k)e^{ikx} \, dk \] (8.23)
where \( \tilde{\Psi}(k) \) can be interpreted as the coefficient for each value of \( k \) or as the wavenumber distribution of the waves. Furthermore, since \( p \propto k \), \( \tilde{\Psi}(k) \) can also be thought of as the momentum distribution. Now we need to Fig. out if there exists such \( \tilde{\Psi}(k) \) that satisfies these conditions.

---

1. This is a well know fact in statistics for equations of the form (8.3).
2. This is also known as the Inverse Fourier Transform of \( \Psi(x, 0) \).
3. Recall that for an \( n^{th} \) order linear ODE, there exist \( n \) independent solutions \( y(x)_n \), and the general solution is \( y_0 = \sum_{n=1}^{\infty} c_n y(x)_n \). In this case, the sum is replaced by an integral and the coefficient by a continuous function since any \( k \) is allowed, \( y_0 = \int_{-\infty}^{\infty} A(k)y(x, k) \, dk \). The factor of \( \frac{1}{\sqrt{2\pi}} \) could be absorbed into the coefficient but it is there to make future equations look nicer due to additions of other factors of \( 2\pi \).
Existence of \( \Psi(k) \): If we multiply both sides of (8.23) by \( e^{-ik'x} \) and integrate we obtain

\[
\int_{-\infty}^{\infty} \Psi(x,0)e^{-ik'x}dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{\Psi} e^{ikx} dk e^{-ik'x}dx. \tag{8.24}
\]

Assuming that the order of integration can be swapped and pulling out a factor of \( \sqrt{2\pi} \) from the second term,

\[
\int_{-\infty}^{\infty} \Psi(x,0)e^{-ik'x}dx = \sqrt{2\pi} \int_{-\infty}^{\infty} \tilde{\Psi}(k) \left[ \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ix(k-k')} dx \right] dk. \tag{8.25}
\]

Next, we recognize that,

\[
\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ix(k-k')} dx = \delta(k-k') \tag{8.26}
\]

where \( \delta(k-k') \) is the Dirac delta function. It follows that (8.25) becomes

\[
\int_{-\infty}^{\infty} \Psi(x,0)e^{-ik'x}dx = \sqrt{2\pi} \tilde{\Psi}(k) \delta(k-k')dk. \tag{8.27}
\]

From the definition of the Dirac delta function, the above equation can be re-written as

\[
\int_{-\infty}^{\infty} \Psi(x,0)e^{-ik'x}dx = \frac{1}{\sqrt{2\pi}} \tilde{\Psi}(k') \tag{8.28}
\]

Then we can solve this to obtain

\[
\tilde{\Psi}(k') = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Psi(x,0)e^{-ik'x}dx. \tag{8.29}
\]

Combining the initial condition from (8.2) with the previous equation we have

\[
\tilde{\Psi}(k') = \frac{1}{\sqrt{2\pi}} \left( \frac{1}{2\pi\sigma^2} \right)^{1/4} \int_{-\infty}^{\infty} e^{-\frac{x^2}{4\sigma^2}} e^{-ik'x} dx. \tag{8.30}
\]

Completing the square in the exponent and pulling out the parts that are not involved in the integration,

\[
\tilde{\Psi}(k') = \frac{1}{\sqrt{2\pi}} \left( \frac{1}{2\pi\sigma^2} \right)^{1/4} e^{-\sigma^2 k'^2/2} \int_{-\infty}^{\infty} e^{-\frac{1}{4\sigma^2} (x+2\sigma^2 ik')^2} dx. \tag{8.31}
\]

Notice that the exponent involves a complex number. In order to solve this problem, we need to perform a complex integration. In order to do this we make the change of variable \( z = x + 2\sigma^2 ik' \) where \( z \) is complex. Then,

\[
\tilde{\Psi}(k') = \frac{1}{\sqrt{2\pi}} \left( \frac{1}{2\pi\sigma^2} \right)^{1/4} e^{-\sigma^2 k'^2/2} \int_{-\infty}^{\infty} e^{-\frac{1}{4\sigma^2} z^2} e^{\frac{z^2}{4\sigma^2}} dz. \tag{8.32}
\]

Observe that this is a line integral over the complex plane. It is pictorially represented in Fig. 8.1. In order to solve this complex integration problem, consider a closed square loop whose top and bottom boundaries are line integrals as shown and whose side boundaries are line integrals located at infinity as shown in Fig. 8.2. Next, we will use the following result from complex analysis known as the **Cauchy Integral Theorem**,

\[
\oint_{\gamma} f(z)dz = 0 \tag{8.33}
\]

\(^4\)This is known as the Fourier Transform of \( \Psi(x,0) \).
Figure 8.1: Line Integral over a complex plane.

Figure 8.2: Closed Loop over a complex plane.
where \( f(z) \) is a complex analytic function on a closed path \( \gamma \).

Applying Cauchy’s Theorem to the closed loop shown in Fig. 8.2 broken into four line integrals we obtain (without writing the constant coefficients which we will add in later)
\[
\int_{L_1} e^{\frac{-z^2}{4\sigma^2}} dz + \int_{L_2} e^{\frac{-z^2}{4\sigma^2}} dz + \int_{L_3} e^{\frac{-z^2}{4\sigma^2}} dz + \int_{L_4} e^{\frac{-z^2}{4\sigma^2}} dz = 0. 
\] (8.34)

Since for such integrals the real part dominates and \( \lim_{z \to \infty} f(z) = \lim_{z \to -\infty} f(z) = 0 \), we find that
\[
\int_{L_2} e^{\frac{-z^2}{4\sigma^2}} dz + \int_{L_3} e^{\frac{-z^2}{4\sigma^2}} dz = 0. 
\] (8.35)

This result applied to (8.34) implies that
\[
\int_{L_1} e^{\frac{-z^2}{4\sigma^2}} dz = -\int_{L_4} e^{\frac{-z^2}{4\sigma^2}} dz 
\] (8.36)

which in this case can be re-written as
\[
\int_{-\infty+2\sigma^2ik'}^{\infty+2\sigma^2ik'} e^{\frac{-z^2}{4\sigma^2}} dz = \int_{-\infty}^{\infty} e^{\frac{-z^2}{4\sigma^2}} dz. 
\] (8.37)

The right hand side integral only involves real limits of integration so
\[
\int_{-\infty+2\sigma^2ik'}^{\infty+2\sigma^2ik'} e^{\frac{-z^2}{4\sigma^2}} dz = \int_{-\infty}^{\infty} e^{\frac{-z^2}{4\sigma^2}} dx. 
\] (8.38)

Substituting this result into (8.32),
\[
\tilde{\Psi}(k') = \frac{1}{\sqrt{2\pi}} \left( \frac{1}{2\pi\sigma^2} \right)^{1/4} e^{-\sigma^2 k'^2} \int_{-\infty}^{\infty} e^{\frac{-x^2}{4\sigma^2}} dx. 
\] (8.39)

This integral is a Gaussian that we previously encountered so
\[
\tilde{\Psi}(k') = \left( \frac{2\sigma^2}{\pi} \right)^{1/4} e^{-\sigma^2 k'^2}. 
\] (8.40)

This is the momentum distribution of the wavefunction. For consistency we will check if (8.40) satisfies (8.23).

**Consistency Check:** We have that
\[
\Psi(x, 0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left( \frac{2\sigma^2}{\pi} \right)^{1/4} e^{-\sigma^2 k^2} e^{ikx} dk. 
\] (8.41)

Pulling out the constants and completing the square in the exponents as we did before we obtain
\[
\Psi(x, 0) = \frac{1}{\sqrt{2\pi}} \left( \frac{2\sigma^2}{\pi} \right)^{1/4} e^{\frac{-x^2}{4\sigma^2}} \int_{-\infty}^{\infty} e^{-\sigma^2 (k - \frac{ikx}{2\sigma^2})^2} dk. 
\] (8.42)
Once again using the method of complex integration,

$$\Psi(x,0) = \frac{1}{\sqrt{2\pi}} \left(\frac{2\sigma^2}{\pi}\right)^{1/4} e^{-\frac{x^2}{4\sigma^2}} \int_{-\infty}^{\infty} e^{-\sigma^2 k^2} dk$$  \hspace{1cm} (8.43)

which then, using the answer for the Gaussian integral, results in \[8.2\].

This confirms that the result for the momentum distribution of the plane wave is given by \[8.40\].

Next, we can find the expectation value of the momentum at \(t = 0\)

$$\langle p(0) \rangle = \int_{-\infty}^{\infty} \Psi(x,0) \left[ -ih \frac{\partial}{\partial x} \Psi(x,0) \right] dx.$$  \hspace{1cm} (8.44)

Substituting in the initial condition given by \[8.2\],

$$\langle p(0) \rangle = \left(\frac{1}{2\pi\sigma^2}\right)^{1/2} \int_{-\infty}^{\infty} e^{-\frac{x^2}{4\sigma^2}} -ih \frac{\partial}{\partial x} e^{-\frac{x^2}{4\sigma^2}} dx.$$  \hspace{1cm} (8.45)

Evaluating the derivative and pulling out constant terms,

$$\langle p(0) \rangle = ih \left(\frac{1}{2\pi\sigma^2}\right)^{1/2} \left(\frac{1}{2\sigma^2}\right) \int_{-\infty}^{\infty} xe^{-\frac{x^2}{2\sigma^2}} dx.$$  \hspace{1cm} (8.46)

The integral is symmetric about the origin and the integrand is an odd function so the integral is equal to zero. Therefore, we conclude

$$\langle p(0) \rangle = 0$$  \hspace{1cm} (8.47)

which means that the expectation value for momentum is zero. The results obtained in \[8.10\] and \[8.47\] show that the wavefunction at \(t = 0\) discussed in this problem represents a particle lying at rest at the origin. From these results we can also find the intrinsic uncertainty of the momentum of the particle

$$\Delta p(0) = \sqrt{\langle p^2(0) \rangle}.$$  \hspace{1cm} (8.48)

The expectation value of the square of the momentum is given by

$$\langle p^2(0) \rangle = -\int_{-\infty}^{\infty} \Psi(x,0) \left[ h^2 \frac{\partial^2}{\partial x^2} \Psi(x,0) \right] dx.$$  \hspace{1cm} (8.49)

Taking the derivative, multiplying the functions and splitting the integral, we obtain

$$\langle p^2(0) \rangle = -h^2 \left(\frac{1}{2\pi\sigma^2}\right)^{1/2} \left( -\frac{1}{2\sigma^2} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}} dx + \frac{1}{4\sigma^2} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2\sigma^2}} dx \right).$$  \hspace{1cm} (8.50)

The first integral is a Gaussian integral while the second is similar to the one encountered in \[8.14\], so we conclude

$$\langle p^2(0) \rangle = -h^2 \left( -\frac{1}{2\sigma^2} + \frac{1}{4\sigma^2} \right) = \frac{h^2}{4\sigma^2}.$$  \hspace{1cm} (8.51)

\footnote{Do not forget that there is still an uncertainty associated with position and momentum.}
Combining this with (8.48)
\[ \Delta p(0) = \frac{\hbar}{2\sigma}. \]  
(8.52)
This is the **intrinsic uncertainty in momentum**. We will now compute the product of the uncertainties in position and momentum
\[ \Delta x(0)\Delta p(0) = \frac{\hbar}{2}. \]  
(8.53)
This equation says that the better we know the position of the particle, the worse we know its momentum. Additionally, this is true regardless of whether the situation is Quantum Mechanical or not. This is due to the fact that the intrinsic uncertainty in position is inversely proportional to the intrinsic uncertainty in the momentum. This is a consequence of Fourier Analysis and wave properties because the sharper a distribution is, the larger the number of waves we need to add up and hence, the broader the spectrum of momentum values. This concept is similar to that of wave interference we encountered in the previous section (refer to Fig. 7.1). Often this conclusion is referred to as **Heisenberg’s Uncertainty Principle**. Returning to the evolution of the wavefunction, we have that
\[ \Psi(x,t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{\Psi}(k)e^{ik(x - \frac{\hbar k t}{2m})} dk. \]  
(8.54)

**Obtaining a Solution:** Using the momentum distribution given in (8.40),
\[ \Psi(x,t) = \frac{1}{\sqrt{2\pi}} \left( \frac{2\sigma^2}{\pi} \right)^{1/4} \int_{-\infty}^{\infty} e^{ik(x - \frac{\hbar k t}{2m}) - \sigma^2k^2} dk. \]  
(8.55)
Rearranging terms in the exponent,
\[ \Psi(x,t) = \frac{1}{\sqrt{2\pi}} \left( \frac{2\sigma^2}{\pi} \right)^{1/4} \int_{-\infty}^{\infty} e^{ikx - \frac{\hbar k t}{2m} - \sigma^2k^2} dk \]  
(8.56)
which can be written as
\[ \Psi(x,t) = \frac{1}{\sqrt{2\pi}} \left( \frac{2\sigma^2}{\pi} \right)^{1/4} \int_{-\infty}^{\infty} e^{-\frac{\sigma^2 + \frac{\hbar t}{2m}}{2\frac{\sigma^2}{\pi}}} \left( k^2 - \frac{ikx}{\sigma^2 + \frac{\hbar t}{2m}} \right)^2 \]  
(8.57)
By completing the square in the exponent,
\[ \Psi(x,t) = \frac{1}{\sqrt{2\pi}} \left( \frac{2\sigma^2}{\pi} \right)^{1/4} \int_{-\infty}^{\infty} e^{-\frac{\sigma^2 + \frac{\hbar t}{2m}}{2\frac{\sigma^2}{\pi}}} \left( k - \frac{ix}{2(\sigma^2 + \frac{\hbar t}{2m})} \right)^2 \]  
(8.58)
Pulling out terms not involved in the integration,
\[ \Psi(x,t) = \frac{1}{\sqrt{2\pi}} \left( \frac{2\sigma^2}{\pi} \right)^{1/4} e^{-\frac{\sigma^2 + \frac{\hbar t}{2m}}{4(\sigma^2 + \frac{\hbar t}{2m})}} \int_{-\infty}^{\infty} e^{-\frac{\sigma^2 + \frac{\hbar t}{2m}}{2(\sigma^2 + \frac{\hbar t}{2m})}} \left( k - \frac{ix}{2(\sigma^2 + \frac{\hbar t}{2m})} \right)^2 \]  
(8.59)
Once again using complex integration in a similar fashion as before, we obtain
\[ \Psi(x,t) = \frac{1}{\sqrt{2\pi}} \left( \frac{2\sigma^2}{\pi} \right)^{1/4} e^{-\frac{\sigma^2 + \frac{\hbar t}{2m}}{4(\sigma^2 + \frac{\hbar t}{2m})}} \int_{-\infty}^{\infty} e^{-\frac{\sigma^2 + \frac{\hbar t}{2m}}{2m}} k^2 dk. \]  
(8.60)
Evaluating the integral and combining terms,

\[ \Psi(x, t) = \left( \frac{2\sigma^2 \pi^2}{4\pi^3 (\sigma^2 + \frac{\hbar t}{2m})^2} \right)^{1/4} e^{-\frac{x^2}{4\sigma^2(1 + \frac{\hbar t}{2m})}}. \quad (8.61) \]

Finally, simplifying terms,

\[ \Psi(x, t) = \left( \frac{1}{2\pi \sigma^2 \left( 1 + \frac{\hbar t}{2m\sigma^2} \right)^2} \right)^{1/4} e^{-\frac{x^2}{4\sigma^2}} \left( 1 + \frac{\hbar t}{2m\sigma^2} \right)^{1/2}. \quad (8.62) \]

This final equation shows the evolution of the wavefunction with respect to time and has information about the evolution of a particle. First, we need to verify that it is properly normalized for all time. Integrating the square of the norm of the wavefunction for all \( x \),

\[ \int_{-\infty}^{\infty} |\Psi(x, t)|^2 \, dx = \left( \frac{1}{2\pi \sigma^2 \left( 1 + \left( \frac{\hbar t}{2m\sigma^2} \right)^2 \right)^2} \right)^{1/2} \int_{-\infty}^{\infty} e^{-\frac{x^2}{4\sigma^2}} \left( 1 + \frac{\hbar t}{2m\sigma^2} \right)^{1/2} dx. \quad (8.63) \]

Evaluating this apparently Gaussian integral, we obtain

\[ \int_{-\infty}^{\infty} |\Psi(x, t)|^2 \, dx = \left( \frac{2\pi \sigma^2}{1 + \left( \frac{\hbar t}{2m\sigma^2} \right)^2} \right)^{1/2} \left( 2\pi \sigma^2 \left( 1 + \left( \frac{\hbar t}{2m\sigma^2} \right)^2 \right) \right)^{1/2} = 1. \quad (8.64) \]

Hence the wave function is normalized for any value of \( t \). We will later show that this is in fact always the case. As previously done we calculate the expectation value of position and the associated intrinsic uncertainty. The expectation value of position is

\[ \langle x(t) \rangle = \int_{-\infty}^{\infty} x |\Psi(x, t)|^2 \, dx = 0 \quad (8.65) \]

since the integral is symmetric about the origin and the integrand is odd. Therefore the wavefunction represents a particle lying at the origin at all time \( t \). We can calculate the uncertainty in the position to be

\[ \langle x^2(t) \rangle = \int_{-\infty}^{\infty} x^2 |\Psi(x, t)|^2 \, dx \\
= \left( \frac{1}{2\pi \sigma^2 \left( 1 + \left( \frac{\hbar t}{2m\sigma^2} \right)^2 \right)^2} \right)^{1/2} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{4\sigma^2}} \left( 1 + \frac{\hbar t}{2m\sigma^2} \right)^{1/2} dx. \quad (8.66) \]

Once again this integral is similar to the one we encountered in (8.14) so

\[ \langle x^2(t) \rangle = \left( \frac{1}{2\pi \sigma^2 \left( 1 + \left( \frac{\hbar t}{2m\sigma^2} \right)^2 \right)^2} \right)^{1/2} \left( 2\sigma^2 \left( 1 + \left( \frac{\hbar t}{2m\sigma^2} \right)^2 \right) \right)^{3/2} \frac{\sqrt{\pi}}{2}. \quad (8.67) \]

\[ \text{Notice that } |\Psi(x, t)|^2 \text{ is not simply } \Psi^2(x, t), \text{ but } \Psi(x, t)\Psi(x, t). \]
Simplifying terms, we finally obtain

\[ \langle x^2(t) \rangle = \sigma^2 \left( 1 + \left( \frac{\hbar t}{2m\sigma^2} \right)^2 \right). \] (8.68)

Therefore,

\[ \Delta x(t) = \sigma \left( 1 + \left( \frac{\hbar t}{2m\sigma^2} \right)^2 \right)^{1/2}. \] (8.69)

We can conclude that the intrinsic uncertainty increases with time. This means that as time progresses, we have to increase the number of measurements we make in order to obtain the same precision in our results. The expectation value of momentum and its associated uncertainty at any time are given by (8.52) and (8.47) since the distribution of momentum does not depend on time, but only on the initial conditions. So,

\[ \langle p(t) \rangle = 0, \quad \Delta p(t) = \frac{\hbar}{2\sigma}. \] (8.70)

The results obtained for the expectation values of position and momentum show that the wavefunction \( \Psi(x, t) \) given by (8.62) represents a particle lying at rest at the origin. The product of the uncertainty in position and uncertainty in momentum at any time is

\[ \Delta x(t) \Delta p(t) = \frac{\hbar}{2} \left( 1 + \left( \frac{\hbar t}{2m\sigma^2} \right)^2 \right)^{1/2} \geq \frac{\hbar}{2} \] (8.71)

with equality holding only for \( t = 0 \). As we mentioned before this is often referred to as Heisenberg’s Uncertainty Principle.

We will now try to understand the physics of the results we have obtained. We observed that the particle remained at rest at the origin but the error increased with time. We can consider the evolution of the intrinsic uncertainty of the position of a particle whose initial uncertainty is a micron or \( 10^{-6} \text{m} \) \( (\Delta x(0) = \sigma = 10^{-6}\text{m}) \). We can solve for the time it will take for this uncertainty to increase by a factor of ten. If the uncertainty increases by a factor of ten at time \( t_{10} \), then

\[ \Delta x(t_{10}) = 10 \cdot \Delta x(0). \] (8.72)

Using (8.16) and (8.69),

\[ \sigma \left( 1 + \left( \frac{\hbar t_{10}}{2m\sigma^2} \right)^2 \right)^{1/2} = 10\sigma \] (8.73)

and solving for \( t_{10} \),

\[ t_{10} = \frac{2m\sigma^2 \sqrt{99}}{\hbar}. \] (8.74)

Notice that \( t_{10} \) is proportional to the mass of the particle \( m \). For example, in the case of an electron, \( m = 9.109 \cdot 10^{-31} \text{ kg} \), which results in

\[ t_{10} = 1.73 \cdot 10^{-7} \text{s} = 173 \text{ ns}. \] (8.75)
It is evident that the localization of a free electron is lost in an extremely short amount of time which is modeled by Fig. 8.3 On the other hand, for a 70-kg physicist,

\[ t_{10} = 1.32 \cdot 10^{25} \text{s} \approx 4.19 \cdot 10^{17} \text{yrs}. \]  

This is an incredibly large amount of time as it is seven orders of magnitude greater than the age of the universe! Thus, the physicist remains localized as described by wave mechanics.

To briefly summarize this section, we have analyzed the behavior of a free particle at rest from the wave mechanical point of view. The results acquired in the last section show how important the effects of wave mechanics are for objects in a range of atomic sizes including electrons, protons, neutrons, etc. It is important to point out that what really matters here is the mass and not the size or density. On the other hand, negligible effects are observed for classical objects. Thus we conclude that classical mechanics holds good for most objects. Furthermore, the relative size \( \hbar \) plays an important role in the dynamical evolution of a system as observed in (8.74).

---

7However, this is not a formal proof of the Uncertainty Principle, it is only an example.
Chapter 9

Attractive Square Well Potential in One Dimension: Bound States

We will not engage in a detailed discussion of a demonstrative example. Consider the simple system consisting of an attractive center where a particle behaves in a distinctive way. When the particle is far away up to a certain distance from the center it doesn’t feel any attraction but once it comes to a certain distance it feels the attraction. For reference, this is exactly what happens with a gravitational potential. This example is useful when analyzing simplified physical models and approximations. Fig. 9.1 pictorially depicts such a potential. Before proceeding to use the wave mechanics ideas we have developed we will briefly discuss the classical approach.

9.1 Classical Approach

It is important to note that we have defined the total energy as $E = \frac{p^2}{2m} + V(x)$. There are three cases to consider in the classical approach depending on the total energy of the particle. We will call them the forbidden, the bounded and the unbounded cases.

- **Forbidden Case:** The case where a particle has energy $E < -V_0$ is forbidden since the total energy must be greater than the potential energy.

- **Bounded Case:** If $-V_0 < E < 0$, then the particle is bounded (it can not escape from the square well) since its total energy is not enough to overcome the potential barrier. In other words, it can never have a positive energy. Therefore, the particle will bounce back and forth as observed in Fig. 9.2.

- **Unbounded Case:** If $E > 0$, then the particle is unbounded and it can move freely outside the potential. However its kinetic energy and speed are reduced when located outside the well

\[
  v = \begin{cases} 
    \sqrt{\frac{2E}{m}}, & \text{if } -a < x < a. \\
    \sqrt{\frac{2(E-V_0)}{m}}, & \text{if } |x| > a. 
  \end{cases} 
\]  

(9.1)

This basic analysis shows that depending on the total energy of the particle, this will be bounded (confined) or unbounded.
Figure 9.1: Finite Square Well Potential

Figure 9.2: Classical behavior of Particle Inside the Well.
9.2 Wave Mechanical Approach

Now that we have considered the more familiar classical approach, we can analyze this problem using wave mechanics. We begin by writing the Schrödinger Equation associated with this potential

\[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Psi(x, t) + V(x) \Psi(x, t) = i\hbar \frac{\partial}{\partial t} \Psi(x, t)\]  \hfill (9.2)

where

\[V(x) = \begin{cases} -V_0, & \text{if } -a < x < a. \\ 0, & \text{if } |x| > a. \end{cases}\]  \hfill (9.3)

Before proceeding to find solutions, we should consider what the conditions for the wavefunction are. We know that the potential is continuous and we want to assume that the time derivative of \(\Psi\) is also continuous. In order to preserve the continuity of the Schrödinger Equation the second position derivative of \(\Psi\) must be continuous as well. Thus \(\Psi(x, t)\) must be continuous and smooth.\footnote{Recall that a function is called smooth if its derivative is continuous.}

The second condition is the \textbf{normalization criteria} from before which states that

\[\int_{-\infty}^{\infty} |\Psi(x, t)|^2 \, dx = 1.\]  \hfill (9.4)

Having seen some of the properties and conditions of the wavefunction, we are in place to solve the differential equation. The method we will use is one of the basic tools to solve partial differential equations: the method of separation of variables. First we assume that the wavefunction has separated to the form

\[\Psi(x, t) = \psi(x) \phi(t).\]  \hfill (9.5)

Combining (9.5) and (9.2)

\[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} (\psi(x)\phi(t)) + V(x)(\psi(x)\phi(t)) = i\hbar \frac{\partial}{\partial t} (\psi(x)\phi(t)).\]  \hfill (9.6)

Differentiating\footnote{The symbol ‘\(\prime\)’ denotes ordinary derivative.} accordingly and dividing the equation by \(\psi(x)\phi(t)\) we obtain

\[-\frac{\hbar^2}{2m} \frac{\psi''(x)}{\psi(x)} + V(x) = i\hbar \frac{\phi'(t)}{\phi(t)}.\]  \hfill (9.7)

These equations have separated variables because the left hand side only depends on \(x\) and the right hand side only depends on \(t\). The equality can only hold if both sides are equal to a constant that for convenience we denote as \(E\)

\[-\frac{\hbar^2}{2m} \frac{\psi''(x)}{\psi(x)} + V(x) = i\hbar \frac{\phi'(t)}{\phi(t)} = E.\]  \hfill (9.8)

This partial differential equation can be broken up into two ordinary differential equations,

\[i\hbar \frac{\phi'(t)}{\phi(t)} = E, \quad -\frac{\hbar^2}{2m} \frac{\psi''(x)}{\psi(x)} + V(x) = E.\]  \hfill (9.9)
The solution to the first is such that
\[ \phi(t) = e^{-iE \frac{t}{\hbar}} \rightarrow \Psi(x,t) = e^{iE \frac{t}{\hbar}} \psi(x). \] (9.10)

This is the solution with separated variables. In order to obtain \( \psi \), the second differential equation can be written as
\[ -\frac{\hbar^2}{2m} \frac{d^2 \psi(x)}{dx^2} + V(x) \psi(x) = E \psi(x). \] (9.11)

This equation is known as the time-independent Schrödinger equation. The solution to this equation depends on the form of the potential.

Regardless of the particular solution obtained for \( \psi(x) \) there is an interesting result for the probability density of the wavefunction,
\[ P(x,t) = \overline{\psi(x)} \overline{\phi(t)} \psi(x) \phi(t). \] (9.12)

Using the solution obtained for \( \phi \),
\[ P(x,t) = \overline{\psi(x)} e^{iE \frac{t}{\hbar}} \psi(x) e^{-iE \frac{t}{\hbar}} = \overline{\psi(x)} \psi(x). \] (9.13)

This equation explicitly reveals that the probability density is independent of time. Thus, the wavefunction represents a stationary state of a particle. This is the case for potentials that only depend on the spatial variables since the previous process of separation of variables works.\(^3\) Now we can completely solve for the wavefunction. The following equation shows the two differential equations that arise. For convenience, we set \( E = -\mathcal{E} \)
\[ -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x)}{\partial x^2} = \begin{cases} -\mathcal{E} \psi(x), & \text{if } |x| > a. \\ (V_0 - \mathcal{E}) \psi(x), & \text{if } |x| < a. \end{cases} \] (9.14)

We can now separately analyze the solutions in each region.

### 9.3 Outside the well: Regions I and III \((|x| > a)\)

If we let
\[ \alpha = \frac{\sqrt{2m\mathcal{E}}}{\hbar} \] (9.15)
then outside of the well the Schrödinger Equation can be expressed as
\[ \frac{d^2 \psi(x)}{dx^2} = \alpha^2 \psi(x). \] (9.16)

Notice that if \( \mathcal{E} < 0 \) then \( \alpha \) is a complex number and the corresponding solution will involve sinusoidal functions as such
\[ \psi(x) = C_1 \cos \left( \frac{\sqrt{2m\mathcal{E}}}{\hbar} x \right) + C_2 \sin \left( \frac{\sqrt{2m\mathcal{E}}}{\hbar} x \right). \] (9.17)

\(^3\)It is important to point out that the wavefunction forms a stationary state only if it is a basis solution of the Schrödinger Equation. If the wavefunction is a superposition of basis solutions, then the probability density will involve cross terms of different angular frequencies that will not cancel. Thus the wavefunction will not represent stationary states anymore.
However, we must meet the limit conditions for $\Psi$ and these sinusoidal functions do not (because the limit as $x$ goes to $\pm \infty$ is not 0). Therefore, we must have $E > 0$ and the general solution must be of the form

$$\psi(x) = C_1 e^{\alpha x} + C_2 e^{-\alpha x}. \quad (9.18)$$

Applying the conditions at infinity we obtain

$$\psi(x) = \begin{cases} 
A e^{-\alpha x}, & \text{if } x > a \\
B e^{\alpha x}, & \text{if } x < a
\end{cases} \quad (9.19)$$

where $A$ and $B$ are constant coefficients.

### 9.4 Inside the well: Region II ($|x| < a$)

If we consider the behavior inside the square well and define

$$\beta = \sqrt{\frac{2m(V_0 - E)}{\hbar}} \quad (9.20)$$

then the Schrödinger Equation inside the well is

$$\frac{\partial^2 \psi(x)}{\partial x^2} = -\beta^2 \psi(x). \quad (9.21)$$

Solving this differential equation we get that the wavefunction has the form

$$\psi(x) = C \cos(\beta x) + D \sin(\beta x). \quad (9.22)$$

For convenience we multiply the solution by $e^{-\alpha a}$

$$\psi(x) = C e^{-\alpha a} \cos(\beta x) + D e^{-\alpha a} \sin(\beta x). \quad (9.23)$$

Combining the results in (9.19) and (9.22), we have that

$$\psi(x) = \begin{cases} 
A e^{-\alpha x}, & \text{if } x > a \\
C e^{-\alpha a} \cos(\beta x) + D e^{-\alpha a} \sin(\beta x), & \text{if } -a < x < a \\
B e^{\alpha x}, & \text{if } x < -a.
\end{cases} \quad (9.24)$$

### 9.5 Matching Conditions and Boundary Conditions

In order to obtain the coefficients $A, B, C, D$ we must impose certain conditions. As we explained earlier, the wavefunction must be continuous and smooth. So far the functions that we have encountered are continuous and smooth in their respective regions. However, we must assure continuity and smoothness at the transitions. Mathematically, this means

$$\lim_{x \to -a^-} \psi(x) = \lim_{x \to a^+} \psi(x), \quad \lim_{x \to a^-} \psi(x) = \lim_{x \to a^+} \psi(x)$$

This is an important result because it implies that the particle can only be in bound states since $E < 0$.

We must have $V_0 > \mathcal{E}$ because otherwise $\beta$ will be complex and then the solutions will be exponentials that will not satisfy boundary conditions.

Because of linearity of the differential equation, we are allowed to multiply the solution by a constant.
\[
\lim_{x \to a^-} \psi'(x) = \lim_{x \to a^+} \psi'(x), \quad \lim_{x \to a^-} \psi'(x) = \lim_{x \to a^+} \psi'(x).
\]

**Continuity Relations**

At \(x = a\),

\[
Ae^{-\alpha a} = (C \cos(\beta a) + D \sin(\beta a))e^{-\alpha a} \tag{9.25}
\]

At \(x = -a\),

\[
Be^{-\alpha a} = (C \cos(\beta a) - D \sin(\beta a))e^{-\alpha a} \tag{9.26}
\]

**Smoothness Relations**

At \(x = a\)

\[-\alpha Ae^{-\alpha a} = (-\beta C \sin(\beta a) + \beta D \cos(\beta a))e^{-\alpha a} \tag{9.27}\]

At \(x = -a\)

\[\alpha Be^{-\alpha a} = (\beta C \sin(\beta a) + \beta D \cos(\beta a))e^{-\alpha a} \tag{9.28}\]

Adding (9.25) and (9.26) and cancelling the terms \(e^{-\alpha a}\) we obtain

\[A + B = 2C \cos(\beta a). \tag{9.29}\]

Similarly, subtracting (9.25) and (9.26),

\[A - B = 2D \sin(\beta a). \tag{9.30}\]

Repeating the same process for (9.27) and (9.28) we obtain

\[A + B = 2C \frac{\beta}{\alpha} \sin(\beta a), \quad A - B = -2D \frac{\beta}{\alpha} \cos(\beta a). \tag{9.31}\]

Combining the two sum equations

\[C \cos(\beta a) = C \frac{\beta}{\alpha} \sin(\beta a). \tag{9.32}\]

This condition is satisfied if either \(C = 0\) or \(\alpha = \beta \tan(\beta a)\). Combining the two difference equations

\[D \sin(\beta a) = -D \frac{\beta}{\alpha} \cos(\beta a). \tag{9.33}\]

It follows then that either \(D = 0\) or \(\alpha = -\beta \cot(\beta a)\). At this point, the conditions reduce to two meaningful cases.\(^7\)

---

\(^7\)The third case is that \(C = D = 0\). However, this is the trivial solution. Furthermore, notice that the conditions \(\alpha = \beta \tan(\beta a)\) and \(\alpha = -\beta \cot(\beta a)\) cannot be simultaneously satisfied.
• If $\alpha = \beta \tan(\beta a)$, then $D = 0$, and $A = B = C \cos(\beta a)$. Then (9.24) becomes

$$
\psi(x) = \begin{cases} 
C \cos(\beta a)e^{-\alpha x}, & \text{if } x > a. \\
C \cos(\beta x)e^{-\alpha a}, & \text{if } -a < x < a. \\
C \cos(\beta a)e^{\alpha x}, & \text{if } x < -a.
\end{cases}
$$

(9.34)

Notice that this case represents a symmetric wave function (even function).

• If $\alpha = -\beta \cot(\beta a)$, then $C = 0$, and $A = -B, A = D \sin(\beta a)$. Then (9.24) yields

$$
\psi(x) = \begin{cases} 
D \sin(\beta a)e^{-\alpha x}, & \text{if } x > a. \\
D \sin(\beta x)e^{-\alpha a}, & \text{if } -a < x < a. \\
-D \sin(\beta a)e^{\alpha x}, & \text{if } x < -a.
\end{cases}
$$

(9.35)

This case represents an antisymmetric wavefunction (odd function). The constant $C$ is determined by normalizing the wavefunction.

### 9.6 Finding $\alpha$ and $\beta$

It becomes useful to discuss the units of the products $\alpha a$ and $\beta a$. It can quickly be verified that the products are unitless. If we now multiply $\beta^2$ and $\alpha^2$ by $a^2$, we obtain the following relations (that should also be unitless)

$$
\alpha^2 a^2 = \frac{2m\mathcal{E}a^2}{h^2}, \quad \beta^2 a^2 = \frac{2mV_0a^2}{h^2} - \frac{2m\mathcal{E}a^2}{h^2}.
$$

(9.36)

Then putting these two equations together

$$
\beta^2 a^2 = \frac{2mV_0a^2}{h^2} - \alpha^2 a^2.
$$

(9.37)

We can make the following change of variables

$$
u = \alpha a, \quad v = \beta a, \quad r = a\sqrt{\frac{2mV_0}{h^2}}.
$$

(9.38)

It follows that (9.37) becomes

$$
v^2 + u^2 = r^2 \quad \forall u, v > 0
$$

(9.39)

which is the equation of a quarter of circle. Also, the conditions for $\alpha$ and $\beta$ for the symmetric and antisymmetric cases become

$$
u = v \tan(v) \quad \text{(symmetric function)}
$$

(9.40)

$$
u = -v \cot(v) \quad \text{(antisymmetric function)}
$$

(9.41)

Therefore, our matching conditions reduce to needing to satisfy (9.39) and either (9.40) (symmetric) or (9.41) (antisymmetric). Fig. 9.3 shows these conditions graphically where the intersections of the functions are where both conditions are satisfied.

There are two interesting pieces of information in Fig. 9.3
The intersections alternate between corresponding symmetric and antisymmetric wavefunctions.

The number of intersections depends on \( r \) (the radius of the circle) that we defined in (9.38). Thus, the larger the value of \( V_0 \) (the deeper the potential well) or larger the value of \( a \) (the wider the well), the greater the number of intersections. This means there are more possible \( v_n \)'s, where \( v_n \) is the \( n \)th intersection point. It can be observed from Fig. 9.3 that \( v_n \in \left( \frac{n}{2}, \frac{n+1}{2} \right) \) for \( n \in \mathbb{Z}^+ \).

Furthermore, using the relations from (9.20), (9.37) and (9.39)

\[
E_n = \frac{-v_n^2 \hbar^2}{2ma^2} + V_0. \quad (9.42)
\]

Recalling the substitution \( E = -\mathcal{E} \) we have

\[
E_n = \frac{v_n^2 \hbar^2}{2ma^2} - V_0. \quad (9.43)
\]

This equation shows that only certain values of energies are allowed. These \( E_n \)'s are called the energy eigenvalues for the particle. We can now find the expectation value and error in measurement of the energy of the particle

\[
\langle E_n \rangle = i\hbar \int_{-\infty}^{\infty} \overline{\Psi}(x,t) \frac{\partial \Psi(x,t)}{\partial t} \, dx. \quad (9.44)
\]

Substituting (9.10) into (9.44) and taking the derivative

\[
\langle E_n \rangle = E_n \int_{-\infty}^{\infty} \overline{\Psi}(x,t) \Psi(x,t) \, dx. \quad (9.45)
\]
Then due to normalization criteria we must have

$$\langle E_n(t) \rangle = E_n.$$ \hspace{1cm} (9.46)

For the error in measurement we have

$$\Delta E_n(t) = \sqrt{\langle E^2(t) \rangle - \langle E(t) \rangle^2}$$ \hspace{1cm} (9.47)

where

$$\langle E^2_n(t) \rangle = -\hbar^2 \int_{-\infty}^{\infty} \Psi(x,t) \frac{\partial^2 \Psi(x,t)}{\partial t^2} dx.$$ \hspace{1cm} (9.48)

Once again, combining (9.10) with (9.44) and taking the derivative

$$\langle E^2_n(t) \rangle = E_n^2.$$ \hspace{1cm} (9.49)

Therefore (9.47) reduces to

$$\Delta E_n(t) = 0.$$ \hspace{1cm} (9.50)

This means that for stationary states, the energy distribution is sharp (there is no error associated with it). In particular the lowest energy state is given by

$$E_0 = \frac{v_0^2 \hbar^2}{2ma^2} - V_0$$ \hspace{1cm} (9.51)

where the first term is given by the kinetic energy of the particle. This is also called the ground state.

So far, we have seen that in the classical approach to the square well potential any energy state is allowed for a particle. However, in the quantum mechanical model, only certain energy states (the eigenvalues) are permitted (all of them bounded), and the number of states depends on the depth and the width of the well, as well as on the mass of the particle. Finally, the lowest energy state is non-zero which means that the particle has some kinetic energy, and therefore it is not completely at rest.

---

8On the other hand, for wavefunctions that involve superposition of basis solutions the energy will not be sharp since there will be cross terms in the calculations.
Chapter 10

Harmonic Oscillator in One Dimension: Bound States

One of the most important and useful systems to analyze is that of the harmonic oscillator. This simple system consists of an attractive potential whose strength is proportional to the square of the separation distance from the center of the potential. In this section we will show that it is possible to find an exact solution to this problem and that once again quantization arises. We will start by revisiting the familiar classical approach to this problem.

10.1 The Classical Approach

The potential energy for the harmonic oscillator is given by $V(x) = \frac{1}{2} k x^2$. This corresponds to the force $\vec{F}(x) = -k x \hat{i}$ where $k$ is a positive constant of proportionality characterizing the strength of the force. (This constant $k$ is often called the “spring constant.”) The total energy of the particle is then given by

$$E(t) = \frac{1}{2} m \left( \frac{dx}{dt} \right)^2 + \frac{1}{2} k x^2. \quad (10.1)$$

Since we are considering only conservative forces (no dissipative forces such as air resistance) we know that total energy is conserved ($\frac{dE}{dt} = 0$.) Using our expression for total energy we obtain

$$m \frac{d x}{dt} \frac{d^2 x}{dt^2} + k x \frac{dx}{dt} = 0 \quad (10.2)$$

which simplifies to

$$\frac{d^2 x}{dt^2} = -\frac{k}{m} x. \quad (10.3)$$

We can define a new variable $\omega^2 = \frac{k}{m}$. Then the above differential equation has solutions of the form

$$x(t) = A \cos(\omega t + \phi) \quad (10.4)$$

where $A$ is the amplitude of the oscillations and $\phi$ is the phase angle. Fig. 10.2 depicts the time evolution of the position of the oscillator for some particular choice of parameters. It follows that

---

1 The energy approach has been preferred rather than the more common force-balancing approach in order to emphasize the relationship to the Schrödinger equation that will be used in the next subsection.
Figure 10.1: Representation of the energy of the harmonic oscillator as a function of position for $k = 1$ and $E = 8$

![Energy of the Classical Harmonic Oscillator as function of Position](image1)

Figure 10.2: Motion of the Classical Harmonic Oscillator as a function of time for $\omega = 1$, $\phi = 0$ and $A = 4$

![Position of the Oscillator as a function of Time](image2)
the velocity of the oscillator is given by

$$v(t) = \frac{dx}{dt} = -A\omega \sin(\omega t + \phi) \quad (10.5)$$

and the total energy of the system is given by

$$E = \frac{1}{2} k A^2. \quad (10.6)$$

This expression for the total energy should resemble that of the total energy of classical electromagnetic waves given by (6.7).

**Note:** Any energy is allowed for a particular frequency since $E$ is independent of $\omega$.

Another important characteristic of the classical harmonic oscillator is that the system is always bounded regardless of how high the energy is. This means that the oscillator is confined to a region $|x| \leq A$ which is usually referred to as the “classically allowed region”. The region $|x| > A$ is appropriately referred to as the “classically forbidden region”. Here $A$ is simply the amplitude of oscillations. We can solve for $A$ from (10.6) to obtain

$$A = \sqrt{\frac{2E}{k}} \quad (10.7)$$

which implies that $A$ only depends on the energy $E$ and the spring constant $k$ of the system. Since $k$ remains a constant and $E$ changes, the amplitude changes like $\sqrt{E}$. Fig. 10.1 shows the potential, kinetic and total energy of the harmonic oscillator as a function of position for specific values of $\omega, \phi$ and $A$. The points where the kinetic energy go to zero are positions of maximum amplitude.

### 10.2 The Quantum Mechanical Approach

Having refreshed our memory of the classical approach to the harmonic oscillator problem, it is interesting to look at this problem from a quantum mechanical perspective. As we did previously we will start with the now familiar Schrödinger equation. Similar to the one dimensional potential problem we considered earlier, we can separate variables and work with the time independent Schrödinger Equation given by

$$-\frac{\hbar^2}{2m} \frac{d^2 \psi(x)}{dx^2} + \frac{1}{2} m \omega^2 x^2 \psi(x) = E \psi(x) \quad (10.8)$$

where we have used the relation $k = m \omega^2$.

Rearranging the equation,

$$\frac{d^2 \psi}{dx^2} + \left( \frac{2mE}{\hbar^2} - \frac{m^2 \omega^2}{\hbar^2} x^2 \right) \psi = 0 \quad (10.9)$$

Before attempting to solve the differential equation at hand, it will be useful to introduce the concept of **natural units** to simplify the differential equation and understand the intrinsic scale of the system. Natural units will also prove to be useful in later problems we will consider.

---

Recall that the harmonic oscillator system is defined as one where the potential energy is a quadratic function of position.
Finding Natural Units: In order to find the natural units of the system we need to make a change of variable that will make the coefficients of the differential equation dimensionless. Let \( x = bz \), where \( b \) has units of length and \( z \) is a dimensionless quantity that tells us the number of units of the natural length \( b \). Thus, we can rewrite (10.9) as
\[
\frac{d^2 \psi}{dx^2} + \left( \frac{2mE}{\hbar^2} - \frac{m^2 \omega^2}{\hbar^2} x^2 \right) \psi = 1\frac{b^2}{b^2} \frac{d^2 \chi(z)}{dz^2} + \left( \frac{2mE}{\hbar^2} - \frac{m^2 \omega^2}{\hbar^2} (bz)^2 \right) \chi(z) = 0
\]
using the substitution \( \chi(z) = \psi(x) \). Then we can simplify that to make the leading coefficient equal to 1 and we obtain
\[
\frac{d^2 \chi(z)}{dz^2} + \left( \frac{2mEb^2}{\hbar^2} - \frac{m^2 \omega^2 b^4 z^2}{\hbar^2} \right) \chi(z) = 0.
\]
It seems reasonable that our first choice for natural units should be to pick \( b \) such that \( \frac{m^2 \omega^2 b^4}{\hbar^2} = 1 \) so we obtain
\[
b = \sqrt{\frac{\hbar}{m\omega}}.
\]
Notice that \( b \) is completely defined by the system since both \( m \) and \( \omega \) are intrinsic properties. This is what we will use as the natural length unit for the quantum harmonic oscillator and can be verified by dimensional analysis. Our previous differential equation then simplifies to
\[
\frac{d^2 \chi(z)}{dz^2} + \left( \frac{2E}{\hbar \omega} - z^2 \right) \chi(z) = 0.
\]
From this equation we can infer that as our second choice we should define the natural unit for energy as \( \epsilon \) where \( E = \epsilon \hbar \omega \).

Our final differential equation in natural units of length and energy is given by
\[
\frac{d^2 \chi(z)}{dz^2} + \left( 2\epsilon - z^2 \right) \chi(z) = 0.
\]
This process of choosing natural units has not only allowed us to simplify our equation but also revealed what the scale of the system is.

We will now shift our focus towards determining how to solve the differential equation.

Solving the Differential Equation: Unlike the square well potential, we do not have boundary conditions to satisfy. However, we need to fulfill the normalization criteria
\[
\int_{-\infty}^{\infty} |\chi(z)|^2 dz < \infty.
\]

3This concept of natural units will be used frequently.

4Notice that the normalization requirement is for \( \psi(x) \) but our change of variables implies that \( \chi(z) \) is square integrable if and only if \( \psi(x) \) is square integrable.
This requires that the function vanishes as \( z \) approaches infinity. We can try the following weighted gaussian function as a solution

\[
\chi(z) = u(z)e^{-\frac{z^2}{2}}. \tag{10.15}
\]

Evaluating the first and second derivatives using the product and chain rule we obtain

\[
\begin{align*}
\chi'(z) &= (u'(z) - zu(z))e^{-\frac{z^2}{2}} \\
\chi''(z) &= [u''(z) - 2zu'(z) - (1 - z^2)u(z)]e^{-\frac{z^2}{2}}. \tag{10.16}
\end{align*}
\]

Substituting \( \chi(z), \chi'(z) \) and \( \chi''(z) \) into (10.13), we obtain a new differential equation

\[
u''(z) - 2zu'(z) + (2\epsilon - 1)u(z) = 0. \tag{10.17}
\]

Now we can use the power series method to find a solution to this differential equation. We first assume that \( u(z) \) has a power series representation given by

\[
u(z) = \sum_{n=0}^{\infty} a_n z^n. \tag{10.18}
\]

Taking its first and second derivatives,

\[
\begin{align*}
u'(z) &= \sum_{n=1}^{\infty} a_n nz^{n-1} \\
u''(z) &= \sum_{n=2}^{\infty} a_n n(n-1)z^{n-2}. \tag{10.19}
\end{align*}
\]

Substituting this into (10.17) gives

\[
\sum_{n=2}^{\infty} a_n n(n-1)z^{n-2} - 2 \sum_{n=1}^{\infty} a_n nz^n + (2\epsilon - 1) \sum_{n=0}^{\infty} a_n z^n = 0. \tag{10.20}
\]

Shifting the index of the first sum we have

\[
\sum_{n=0}^{\infty} a_{n+2}(n+2)(n+1)z^n - 2 \sum_{n=1}^{\infty} a_n nz^n + (2\epsilon - 1) \sum_{n=0}^{\infty} a_n z^n = 0. \tag{10.21}
\]

Separating out the first terms of the first and second series and combining the remaining terms we obtain

\[
2a_2 - (2\epsilon - 1)a_0 + \sum_{n=1}^{\infty} [a_{n+2}(n+2)(n+1) - 2a_n n + (2\epsilon - 1)a_n]z^n = 0. \tag{10.22}
\]

The power series will be always equal to 0 if and only if each of its coefficients is identically 0. This leads to the following recursive relation for the coefficients,

\[
a_{n+2} = \frac{(2n + 1 - 2\epsilon)}{(n+2)(n+1)} a_n. \tag{10.23}
\]

Since we have a second order differential equation we have two constants \((a_0, a_1)\) that are given by initial conditions. Therefore, the power series representation will give an infinite polynomial. The solution to (10.13) is then given by

\[
\chi(z) = e^{-\frac{z^2}{2}} \sum_{n=0}^{\infty} a_n z^n. \tag{10.24}
\]

\(^5\)This equation is called a Hermite equation. The polynomial in front of the zeroth order term is called a Hermite polynomial.
We have obtained a solution but it is not yet clear why we need to talk about quantization. Recall that the normalization criteria (square integrability) requires the wavefunction to vanish at infinity. This only happens when the Hermite polynomial is finite because otherwise the integral cannot converge. Thus, we need the condition

\[ 2n + 1 - 2\epsilon = 0 \rightarrow \epsilon = n + \frac{1}{2} \]  

(10.25)

where \( n \) is some arbitrary non-negative integer. If this is the case, the polynomial will terminate at degree \( n \).

Let us examine this condition a bit further. Combining it with our expression for energy we obtain

\[ E = \left( n + \frac{1}{2} \right) \hbar \omega \]  

(10.26)

This result is strikingly similar to the quantization of light \( (6.51) \) since the difference between consecutive energy levels is precisely \( \hbar \omega \) (or \( hf \)) and is independent of the energy levels themselves. Recall our earlier observation that the classical harmonic oscillator resembled classical electromagnetic waves. Upon noticing this parallel it seems natural to treat the quantum harmonic oscillator as a model for the emission of light and of light itself (photons). \textbf{Light can be thought of as an oscillatory system!}

Tracing back to the implications of quantization, it can be seen that the polynomial will terminate at \( n \) where \( n + 2k = 0 \) and \( k \) is a positive integer. If \( n \) is even then all even terms larger than \( n \) will vanish. Thus we require \( a_1 = 0 \) so that no infinite odd terms appear. Likewise, if \( n \) is odd, all odd terms larger than \( n \) will be equal to zero and we will require \( a_0 = 0 \).

For instance, using the recursive relation given in \( (10.23) \) these are the first four wavefunctions for the harmonic oscillator

<table>
<thead>
<tr>
<th>( n ) value</th>
<th>Wavefunction</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>( \chi(z) = Ce^{-\frac{z^2}{2}} )</td>
</tr>
<tr>
<td>1</td>
<td>( \chi(z) = Cze^{-\frac{z^2}{2}} )</td>
</tr>
<tr>
<td>2</td>
<td>( \chi(z) = C \left( 1 - 2z^2 \right) e^{-\frac{z^2}{2}} )</td>
</tr>
<tr>
<td>3</td>
<td>( \chi(z) = C \left( z - \frac{2}{3} z^3 \right) e^{-\frac{z^2}{2}} )</td>
</tr>
</tbody>
</table>

where \( C \) is a constant determined by normalizing the wavefunctions. We can now compute the normalization for the wavefunctions.

---

\(^6\)It follows easily that every term in the series vanishes after the \( n \)th.
Computing the Normalization: Recall our normalization condition

$$\int_{-\infty}^{\infty} |\psi(x)|^2 dx = 1$$  \hspace{1cm} (10.27)

or in terms of \( \chi \) and \( z \),

$$b \int_{-\infty}^{\infty} |\chi(z)|^2 dz = 1.$$  \hspace{1cm} (10.28)

The first one is easy since we can recognize that it is a gaussian function. Then we know that the normalized wavefunction is

$$\chi_0(z) = \left( \frac{1}{\pi b^2} \right)^{1/4} e^{-\frac{z^2}{2b^2}}.$$  \hspace{1cm} (10.29)

Similarly,

$$\chi_1(z) = \left( \frac{1}{\pi b^2} \right)^{1/4} \sqrt{2} ze^{-\frac{z^2}{2b^2}},$$

$$\chi_2(z) = \left( \frac{1}{\pi b^2} \right)^{1/4} \frac{1}{\sqrt{2}} (1 - 2z^2) e^{-\frac{z^2}{2b^2}},$$

$$\chi_3(z) = \left( \frac{1}{\pi b^2} \right)^{1/4} \frac{1}{\sqrt{3}} \left( z - \frac{2}{3} z^3 \right) e^{-\frac{z^2}{2b^2}}.$$  \hspace{1cm} (10.30)

It is useful to analyze these results graphically. Fig. 10.3 shows the wavefunctions in their natural unit length scale. Notice that the only parameter in these functions is \( b \) which is given by the system. This supports the idea of \( b \) being an intrinsic scale for the harmonic oscillator.

The first and third plots correspond to \( n = 0 \) and \( n = 2 \) respectively. We can see that these wavefunctions are odd, whereas the second and fourth plots corresponding to \( n = 1 \) and \( n = 3 \) respectively, show that the wavefunctions are even. Furthermore, it can be observed that for larger \( n \) the probability distributions are more spread out. This can be partially understood using the fact that a larger \( n \) corresponds to larger energy and hence a larger degree of freedom. Another interesting result is that for \( n \neq 0 \) we observe nodes indicating that in the neighborhood around the nodes (points where the graph touches the x-axis) it is difficult to locate the particle. Finally for \( n = 0 \) we have a nice gaussian distribution similar to the result obtained for a free particle. However, an important difference is that there is no spreading of the wavefunction. This is due to the fact that the potential we have chosen localizes the particle.

In terms of \( \psi(x) \) the wavefunctions are

$$\psi_0(x) = \left( \frac{1}{\pi b^2} \right)^{1/4} e^{-\frac{x^2}{2b^2}},$$

$$\psi_1(x) = \left( \frac{1}{\pi b^2} \right)^{1/4} \frac{\sqrt{2}}{b} xe^{-\frac{x^2}{2b^2}},$$

$$\psi_2(x) = \left( \frac{1}{\pi b^2} \right)^{1/4} \frac{1}{\sqrt{2}} \left( 1 - \frac{2}{b^2} x^2 \right) e^{-\frac{x^2}{2b^2}}.$$  \hspace{1cm} (10.31)

\(^7\)However, unlike the classical case the particle can be found virtually anywhere.
The unit on the vertical axis representing the probability density is \((\frac{1}{\pi b^2})^{1/2}\).

\[
\psi_3(x) = \left(\frac{1}{\pi b^2}\right)^{1/4} \sqrt{3} \left(\frac{x}{b} - \frac{2}{3b^3} x^3\right) e^{-\frac{x^2}{2b^2}}.
\]  

(10.31)

We can now calculate some of the basic **observables** of the distributions such as average position, momentum, kinetic and potential energy.

---

**Observables:** The average position as given in (7.42) for the lowest energy wavefunction is given by

\[
\langle x_0 \rangle = \left(\frac{1}{\pi b^2}\right)^{1/2} \int_{-\infty}^{\infty} x e^{-\frac{x^2}{2b^2}} dx.
\]  

(10.32)

Note that this is an odd function so

\[
\langle x_0 \rangle = 0.
\]  

(10.33)

Similarly, the average momentum as given in (7.43) is given by

\[
\langle p_0 \rangle = -i\hbar \left(\frac{1}{\pi b^2}\right)^{1/2} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2b^2}} \frac{d}{dx} \left( e^{-\frac{x^2}{2b^2}} \right) dx.
\]  

(10.34)

Evaluating the derivative and simplifying we obtain

\[
\langle p_0 \rangle = i\hbar \left(\frac{1}{\pi b^2}\right)^{1/2} \int_{-\infty}^{\infty} \frac{x}{b^2} e^{-\frac{x^2}{2b^2}} dx
\]  

(10.35)
which by odd parity reduces to
\[ \langle p_0 \rangle = 0. \] (10.36)

The average kinetic energy is given by
\[ \langle K_0 \rangle = -\frac{\hbar^2}{2m} \left( \frac{1}{\pi\hbar^2} \right)^{1/2} \int_{-\infty}^{\infty} e^{\frac{-x^2}{2\hbar^2}} d^2x \left( e^{\frac{-x^2}{2\hbar^2}} \right) dx. \] (10.37)

After some algebraic manipulation and simplification we obtain
\[ \langle K_0 \rangle = \frac{1}{4} \hbar \omega. \] (10.38)

Likewise, we find the average potential energy to be
\[ \langle V_0 \rangle = \frac{1}{4} \hbar \omega. \] (10.39)

Now we can analyze our results. We notice that the average position and momentum of the oscillator are 0. This could have been directly seen from the plots and their symmetries. The oscillator spends half the time on either side of the potential. It is also important to notice that the sum of the expectation values of the energy is precisely \( \frac{1}{2} \hbar \omega \) which agrees with the result for the total energy given by quantization, guaranteeing conservation of energy. This energy conservation result is naturally consistent with the Schrödinger Equation. We can briefly illustrate this by multiplying the Schrödinger Equation in (7.27) by the complex conjugate of the wavefunction and integrating both sides over real space,
\[ -\int \int \int \frac{\hbar^2}{2m} \frac{\nabla^2 \Psi}{\Psi} dxdydz + \int \int \int \Psi V \Psi dxdydz = i \int \int \int \hbar \frac{\partial \Psi}{\partial t} dxdydz. \] (10.40)

This is equivalent to
\[ \langle K(t) \rangle + \langle V(t) \rangle = \langle E(t) \rangle. \] (10.41)

From symmetry arguments it is evident that the expectation value of position and momentum for the remaining energy levels is 0. It can also be seen that the average potential and kinetic energies for the 4th lowest wavefunction are equal to \( \frac{1}{4} \hbar \omega \).
Chapter 11

Single Delta Function Potential

We have worked with the quantum mechanics of one dimensional potentials once before when talking about the square well. Another important potential usually encountered in quantum mechanics is that of the Delta Function. We will denote this function by \( \delta(x) \) which you might have seen in a course on differential equations. This function is special in that it is zero everywhere except at a single point where it is infinite.

11.1 Dirac Delta Function

Recall the definition of the Dirac Delta Function which is not a true function but rather a generalized function or distribution defined by the following integral

\[
\int_{-\infty}^{\infty} \delta(x) f(x) dx = f(0). \tag{11.1}
\]

This can be identified as a mapping from the space of functions to the real numbers. It is easy to show by some simple substitution that the following properties are true

\[
\begin{align*}
\int_{-\infty}^{\infty} \delta(x-a) f(x) dx &= f(a) \\
\int_{-\infty}^{\infty} \delta(ax) f(x) dx &= \frac{f(0)}{|a|}. \tag{11.2}
\end{align*}
\]

Some limits of sequences of functions behave as the Dirac Delta Function such as

\[
\begin{align*}
\delta_{\alpha} &= \sqrt{\frac{\pi}{\alpha}} e^{-\alpha x^2} = \delta(x) \\
\delta_{\alpha} &= \begin{cases} 
0, & \text{if } |x| > \alpha \\
\frac{1}{2\alpha}, & \text{if } |x| \leq \alpha \end{cases} \tag{11.3}
\end{align*}
\]

when \( \alpha \) tends to infinity.

The second function is more common and is plotted in Fig. 11.1. The spike at \( x = 0 \) represents it being infinitely large at that point. However the integral over the entire real line is exactly 1.
Another important property of the Dirac Delta function is that if we let \( S(x) \) be a function with compact support, the function is zero everywhere except in an open interval around a given \( x \) value, around \( x = 0 \) such that \( S(0) = 1 \), then

\[
\int_{-\infty}^{\infty} \delta(x)S(x)f(x)\,dx = f(0).
\]  

This property allows us to reformulate the Dirac Delta Function in terms of the following limit of a definite integral

\[
\int_{-\infty}^{\infty} \delta(x)f(x)\,dx = \lim_{\epsilon \to 0} \int_{-\epsilon}^{\epsilon} \delta(x)f(x)\,dx.
\]  

11.2 Schrödinger Equation of the Attractive Delta Function Potential.

The **attractive Delta Function Potential** is defined by \( V(x) = -V_0 a \delta(x) \) where \( V_0 a \) is a positive constant that represents the strength of the potential. We will now analyze what happens in a quantum mechanical system containing this potential. The first step is to write the time independent Schrödinger Equation for the potential,

\[
-\frac{\hbar^2}{2m} \frac{d^2 \psi(x)}{dx^2} - V_0 a \delta(x) \psi(x) = E \psi(x).
\]  

Since \( \delta(x) \) by itself has no meaning whatsoever, we must integrate this equation around an open interval \( (-\epsilon, \epsilon) \) and take the limit as \( \epsilon \) tends to 0,

\[
\lim_{\epsilon \to 0} \left[ -\frac{\hbar^2}{2m} \int_{-\epsilon}^{\epsilon} \frac{d^2 \psi(x)}{dx^2} \,dx - V_0 a \int_{-\epsilon}^{\epsilon} \delta(x) \psi(x) \,dx \right] = E \lim_{\epsilon \to 0} \int_{-\epsilon}^{\epsilon} \psi(x) \,dx.
\]
Using the property of the Delta Function in (11.2) the previous equation reduces to

\[- \frac{\hbar^2}{2m} \lim_{\epsilon \to 0} \left( \frac{d\psi}{dx} \bigg|_{x=\epsilon} - \frac{d\psi}{dx} \bigg|_{x=-\epsilon} \right) - V_0 a \psi(0) = \lim_{\epsilon \to 0} 2 \epsilon E \psi(0). \tag{11.8} \]

The right hand side vanishes as we take the limit and

\[\lim_{\epsilon \to 0} \left( \frac{d\psi}{dx} \bigg|_{x=\epsilon} - \frac{d\psi}{dx} \bigg|_{x=-\epsilon} \right) = -2 \frac{m V_0 a}{\hbar^2} \psi(0). \tag{11.9}\]

If we let \( \beta^2 = -2 \frac{m V_0 a}{\hbar^2} \), then

\[\lim_{\epsilon \to 0} \left( \frac{d\psi}{dx} \bigg|_{x=\epsilon} - \frac{d\psi}{dx} \bigg|_{x=-\epsilon} \right) = -\beta^2 \psi(0). \tag{11.10}\]

Thus we see that there is a discontinuity of the first derivative with respect to \( x \) of the wavefunction. Recall that this quantity is associated with the momentum of the particle. The physical meaning of this discontinuity is that the particle receives an impulse when it reaches \( x = 0 \), the strength of which is precisely \( V_0 a \). For this reason this is sometimes called the **jump condition**.

Now that we know the physical result of this particular potential we can solve the differential equation.

---

**Solving the Differential Equation:** For \( x = 0 \), we have

\[- \frac{\hbar^2}{2m} \frac{d^2 \psi(x)}{dx^2} = E \psi(x). \tag{11.11}\]

Notice that the potential term vanishes at \( x \neq 0 \). If we conveniently change variables using \( E = -\mathcal{E} \), then we obtain

\[\frac{d^2 \psi(x)}{dx^2} = 2m \frac{\mathcal{E}}{\hbar^2} \psi(x). \tag{11.12}\]

We can define \( \alpha^2 = -2m \frac{\mathcal{E}}{\hbar^2} \) and the previous equation tranforms to

\[\frac{d^2 \psi(x)}{dx^2} = \alpha^2 \psi(x). \tag{11.13}\]

The solution for this equation that satisfies square integrability is given by

\[\psi(x) = \begin{cases} Ae^{-\alpha x}, & \text{if } x > 0 \\ Be^{\alpha x}, & \text{if } x < 0 \end{cases} \tag{11.14}\]

where \( \alpha > 0 \). Hence we must have \( \mathcal{E} > 0 \) or \( E < 0 \), which means that we have bound states. **How many are there?** To answer this question we solve the remaining conditions. In particular we need to satisfy the continuity of the wavefunction; and the derivative with respect to \( x \) must follow the jump condition. The continuity at \( x = 0 \) implies

\[\lim_{x \to 0^-} Ae^{-\alpha x} = \lim_{x \to 0^+} Be^{\alpha x} = \psi(0) \tag{11.15}\]

\(^1\)If the wavefunction is not continuous, momentum will be infinite at \( x = 0 \).
which means
\[ A = B = \psi(0). \] (11.16)

The jump condition gives
\[ \lim_{\epsilon \to 0} \left( -\alpha A e^{-\alpha \epsilon} - \alpha A e^{-\alpha \epsilon} \right) = -\beta^2 a A \] (11.17)
which, after taking the limits, reduces to
\[ 2\alpha = \beta^2 a. \] (11.18)

This implies that there is a single value of \( \alpha \) that satisfies all conditions and this is the only bound state for the particle under the Delta Function Potential. Solving for \( E \) we obtain
\[ E = -\mathcal{E} = -\frac{m V_0^2 a^2}{2\hbar^2}. \] (11.19)

The only remaining constant to be found is \( A \) which comes from the normalization criteria
\[ \int_{-\infty}^{\infty} \psi^2(x) dx = A^2 \int_{-\infty}^{0} e^{2\alpha x} dx + A^2 \int_{0}^{\infty} e^{-2\alpha x} dx = 1. \] (11.20)

Evaluating the integrals, we obtain
\[ \frac{A^2}{\alpha} = 1 \rightarrow A = \sqrt{\alpha}. \] (11.21)

Finally, we observe that the wavefunction has the form
\[ \psi(x) = \begin{cases} \sqrt{\alpha} e^{-\alpha x}, & \text{if } x > 0 \\ \sqrt{\alpha}, & \text{if } x = 0 \\ \sqrt{\alpha} e^{\alpha x}, & \text{if } x < 0 \end{cases} \] (11.22)

where \( \alpha = \sqrt{\frac{2m\mathcal{E}}{\hbar^2}} \), with energy eigenvalue \( E = -\frac{m V_0^2 a^2}{2\hbar^2} \). The probability distribution \( \psi^2(x) \) is plotted in Fig. 11.2.

Now we will solve for some basic observables and expectations values for the unique bound states of the wavefunction corresponding to the Delta Function Potential.

---

**Finding Observables:** The expectation values of position and momentum can be computed using formulas we previously obtained. It is easier to infer that they should be zero simply by looking at the symmetry of the probability distribution. We conclude that
\[ \langle x \rangle = 0, \quad \langle p \rangle = 0. \] (11.23)

The expectation value for the potential is easy to compute
\[ \langle V \rangle = -\int_{-\infty}^{\infty} V_0 \alpha \psi^2(x) \delta(x) dx. \] (11.24)
Using the property in (11.2) and the relations given in (11.16), (11.18) and (11.21)

\[ \langle V \rangle = -\frac{V_0 a^2 \beta^2}{2} \]  

(11.25)

or using \( \beta^2 = -\frac{2mV_0 a}{\hbar^2} \) we obtain

\[ \langle V \rangle = -\frac{mV_0^2 a^2}{\hbar^2}. \]  

(11.26)

Using conservation of energy, we find that

\[ \langle K \rangle = \frac{mV_0^2 a^2}{2\hbar^2}. \]  

(11.27)

It is important to emphasize that only one energy eigenvalue has been found and that it is a bound state. The Delta Function Potential is the simplest potential that gives rise to quantization, and it can model the behavior of a one dimensional proton-electron atom such as Hydrogen. In the following sections, we will use this potential to model more complex systems that could be interpreted as molecules. We will see that quantization arises for a system with several delta function potentials, and that not only one but rather several allowed states naturally appear. We will then briefly extend our model to understand the behavior and structure of crystalline solids.
In the previous section we investigated the quantum mechanical nature of a single delta function potential which could be thought of as a simple representation of an atomic potential. In this section, we will extend that model to one with many equally spaced delta function potentials. This might prove to be a good first model to use in understanding the behavior of composite systems such as molecules or solids.

12.1 Many Delta Function Potentials

We will start by setting up our problem. Consider a set of \( N \) equally separated attractive delta function potentials that are also equal in strength. Let the separation between them be \( d \) and the strength of each potential be \( V_0a \). Furthermore, for the sake of simplicity we choose aforementioned constants such that \( \frac{mV_0a}{\hbar^2} = 1 \).
We will set the first one to be at \( x = 0 \). Then it follows that the \( n^{th} \) potential will be at \( x = (n-1)d \). Naturally the problem of many potentials is harder than that of a single one, but we can use the result from the last section to significantly simplify the problem.

We will focus on the behavior of the wavefunction around the potential at \( x = nd \) which is located in the open interval \( ((n-1)d, (n+1)d) \). This is exactly the problem we solved in the last section where we found the solutions to be linear combinations of exponentials on the left and right sides of the potential. However, unlike the single delta function potential we can have both exponentially growing and decaying functions on either side since we are dealing with a finite region and we are not requiring square integrability over that finite region. Therefore, the solutions on either side will be of the form \( Ae^{-\alpha x} + Be^{\alpha x} \).

For computational convenience, we will center the left and right parts of the wavefunctions at \( x = nd - \frac{d}{2} \) and \( x = nd + \frac{d}{2} \) respectively. Thus, for the left side we have

\[
\psi_L(x) = A_L e^{-\alpha [x-(nd-d/2)]} + B_L e^{\alpha [x-(nd-d/2)]}.
\]

(12.1)

Similarly for the right side we have

\[
\psi_R(x) = A_R e^{-\alpha [x-(nd+d/2)]} + B_R e^{\alpha [x-(nd+d/2)]}
\]

(12.2)

where \( \alpha^2 = -\frac{2mE}{\hbar^2} \).

These solutions can be physically interpreted as incident, reflected and transmitted waves that were scattered off of the potential. On the left side the growing and decaying exponentials represent the incident and reflected wave respectively, whereas on the right side the decaying and growing exponentials represent the transmitted and the reflected wave respectively off of the next scattering potential. Therefore, the coefficients \( A_L, B_L, A_R \) and \( B_R \) quantify the amplitude of these processes.

---

**Obtaining the Transfer Matrix:** We are presented with the task of solving for the coefficients. Similar to the case we considered in the last section, we will use continuity of the wavefunction at \( x = nd \) and similarly we need to satisfy the jump condition at \( x = nd \). Mathematically this is given by

\[
\lim_{x \to nd} \psi_L(x) = \lim_{x \to nd} \psi_R(x) = \psi(nd)
\]

(12.3)

where \( \beta^2 = \frac{2mV_0 a}{\hbar^2} \). Recall that \( \frac{mV_0 a}{\hbar^2} = 1 \). Using this the second condition above becomes

\[
\lim_{\epsilon \to 0} \left( \frac{d\psi_R}{dx} \bigg|_{x=nd+\epsilon} - \frac{d\psi_L}{dx} \bigg|_{x=nd-\epsilon} \right) = -\beta^2 \psi(nd).
\]

(12.4)

We can always shift exponentials by \( x_0 \) to the right by adding a proportionality constant \( C = e^{-x_0} \) since \( e^{(x-x_0)} = e^x e^{-x_0} = Ce^x \).
Substituting (12.1) and (12.2) into (12.3), the continuity requirement becomes

\[ A_L e^{-\alpha d/2} + B_L e^{\alpha d/2} = A_Re^{\alpha d/2} + B_Re^{-\alpha d/2} = \psi(n d). \] (12.5)

Likewise, the jump condition becomes

\[ (-A_L \alpha e^{-\alpha d/2} + B_L \alpha e^{\alpha d/2}) - (-A_R \alpha e^{\alpha d/2} + B_R \alpha e^{-\alpha d/2}) = 2(A_L e^{-\alpha d/2} + B_L e^{\alpha d/2}) \] (12.6)

which after dividing by \( \alpha \) and rearranging we obtain

\[ -A_L e^{-\alpha d/2} \left( 1 + \frac{2}{\alpha} \right) + B_L e^{\alpha d/2} \left( 1 - \frac{2}{\alpha} \right) = -A_R e^{\alpha d/2} + B_R e^{-\alpha d/2}. \] (12.7)

After the substitution \( v = e^{\alpha d/2} \) (12.5) and (12.7) transform to

\[ -A_L \left( \frac{1}{v} + \frac{2}{\alpha v} \right) + B_L \left( v - \frac{2}{\alpha v} \right) = -A_R v + B_R \frac{v}{\alpha}. \] (12.8)

Adding and subtracting both equations, we obtain

\[ \frac{2B_R}{v} = -\frac{2A_L}{\alpha v} + B_L \left( 2v - \frac{2v}{\alpha} \right) \]
\[ 2A_R v = A_L \left( \frac{2v}{v} + \frac{2}{\alpha v} \right) + B_L \frac{2v}{\alpha}. \] (12.9)

Solving for \( A_R \) and \( B_R \) we obtain the following relations

\[ A_R = \frac{1}{v^2} \left( 1 + \frac{1}{\alpha} \right) A_L + \frac{1}{\alpha} B_L \]
\[ B_R = -\frac{1}{\alpha} A_L + v^2 \left( 1 - \frac{1}{\alpha} \right) B_L. \] (12.10)

In matrix notation this can be expressed as

\[
\begin{bmatrix}
A_R \\
B_R
\end{bmatrix} = \begin{bmatrix}
\frac{1}{v^2} \left( 1 + \frac{1}{\alpha} \right) & \frac{1}{\alpha} \\
-\frac{1}{\alpha} & v^2 \left( 1 - \frac{1}{\alpha} \right)
\end{bmatrix}
\begin{bmatrix}
A_L \\
B_L
\end{bmatrix}.
\] (12.11)

The antisymmetric matrix we will call \( T \) is

\[
T = \begin{bmatrix}
\frac{1}{v^2} \left( 1 + \frac{1}{\alpha} \right) & \frac{1}{\alpha} \\
-\frac{1}{\alpha} & v^2 \left( 1 - \frac{1}{\alpha} \right)
\end{bmatrix}.
\] (12.12)

This matrix is known as the **Transfer Matrix** since it recursively relates the coefficients of the wavefunction on either side of the attractive potential.

### 12.2 Transfer Matrix

Notice that the transfer matrix is independent of \( n \). This is a consequence of the translational invariance of the problem (all potentials are equally spaced by a distance \( d \)). This property turns
out to be extremely important in solving this problem since each pair of sides of the wavefunction
with respect to any potential will be related by the same transfer matrix \( T \). Therefore, if we know
the wavefunction on the left of the first potential we can recursively calculate how the wavefunction
will look all the way to the right by multiplying repeatedly by \( T \).

Now for the more complex problem of \( N \) attractive centers we need to know the coefficients for
the wavefunction on the left of the first potential and the right of the last potential (effectively our
initial conditions.) We will label these \( A_{L1} \), \( A_{R1} \), \( A_{LN} \) and \( B_{RN} \) respectively.

At this point we have to consider our square integrability conditions since there are no poten-
tials either at the left of the first one or at the right of the last one. Thus, we require \( A_{L1} = 0 \) and
\( B_{RN} = 0 \). Hence, the matrix condition is given by

\[
\begin{bmatrix}
A_{RN} \\
0
\end{bmatrix} = T^N \begin{bmatrix}
0 \\
B_{L1}
\end{bmatrix}
\]

(12.13)

If we let \( T^N_{n,m} \) be the entry corresponding to the \( n^{th} \) row and \( m^{th} \) column of \( T^N \) then we obtain the relation

\[
\begin{bmatrix}
A_{RN} \\
0
\end{bmatrix} = \begin{bmatrix}
T^N_{1,1} & T^N_{1,2} \\
T^N_{2,1} & T^N_{2,2}
\end{bmatrix} \begin{bmatrix}
0 \\
B_{L1}
\end{bmatrix}
\]

(12.14)

which can be written as

\[
A_{RN} = T^N_{1,2}B_{L1}, \quad 0 = T^N_{2,2}B_{L1}.
\]

(12.15)

We know that \( B_{L1} \neq 0 \) because if it were equal to zero, then the wavefunction would be zero
everywhere. Therefore, the previous equations can only be satisfied if

\[
T^N_{2,2} = 0.
\]

(12.16)

We will see that this consequence leads to quantization in this problem.

### 12.3 Analysis for various values of \( N \)

It is interesting to observe how the physics of this system changes as we vary \( N \). We will start
by considering the simple case where \( N = 1 \). The single delta function potential case that we
previously discussed.

---

\( N = 1 \): Using the transfer matrices we have the relation

\[
\begin{bmatrix}
A_{R1} \\
0
\end{bmatrix} = \begin{bmatrix}
\frac{1}{\pi^2}(1 + \frac{1}{\alpha}) \\
-\frac{1}{\pi^2}
\end{bmatrix} \begin{bmatrix}
1 \\
\frac{1}{\pi^2}(1 - \frac{1}{\alpha})
\end{bmatrix} \begin{bmatrix}
0 \\
B_{L1}
\end{bmatrix}.
\]

(12.17)

In this case \( T^1_{2,2} = v^2(1 - \frac{1}{\alpha}) \). It follows from (12.16) that the following must be true

\[
v^2 \left(1 - \frac{1}{\alpha}\right) = 0.
\]

(12.18)

---

\(^2\)Keep in mind that the wavefunction on the right of the potential at \( x = nd \) is the same as the wavefunction on
the left of potential at \( x = (n+1)d \).

\(^3\)Otherwise the wavefunction will blow up at infinity and negative infinity. This follows from (12.1) and (12.2).
Thus, \( \alpha = 1 \) which implies that the energy is given by

\[
E = -\frac{\hbar^2}{2m} = -\left(\frac{mV_0a}{\hbar^2}\right)^2 \frac{\hbar^2}{2m} = -\frac{mV_0^2a^2}{2\hbar^2}.
\]  
(12.19)

This agrees with the result obtained in the previous section. We see this by comparing it to the expression in (11.19). Since our method seems to be working, we can attempt to test other values of \( N \).

\( N = 2: \) Let \( N = 2 \), then our transfer matrices give us the relation

\[
\begin{bmatrix}
A_{R2} \\
0
\end{bmatrix} = \begin{bmatrix}
\frac{1}{v^2}(1 + \frac{1}{\alpha}) \\
-\frac{1}{\alpha}
\end{bmatrix} \begin{bmatrix}
\frac{1}{v^2} \\
\frac{1}{\alpha}
\end{bmatrix}^2 \begin{bmatrix}
0 \\
B_{L1}
\end{bmatrix}.
\]  
(12.20)

Then

\[
T_{2,2}^2 = -\left(\frac{1}{\alpha}\right)^2 + v^4 \left(1 - \frac{1}{\alpha}\right)^2 = 0.
\]  
(12.21)

So,

\[
\frac{1}{\alpha} = \pm v^2 \left(1 - \frac{1}{\alpha}\right),
\]

\[
\frac{1}{v^2} = \pm \alpha \left(1 - \frac{1}{\alpha}\right).
\]  
(12.22)

Using a substitution \( v = e^{\alpha d/2} \) and rearranging we obtain

\[
\alpha = 1 \pm e^{-\alpha d}.
\]  
(12.23)

Notice that this equation has \( d \) as a parameter. We can analyze this condition graphically for \( d = 1 \) and \( d = 2 \). This is shown in Fig. [12.2]

For \( d = 1 \) we have a single solution given by \( \alpha = 1.47 \), while for \( d = 2 \) we have two solutions given by \( \alpha_1 = 0.80 \) and \( \alpha_2 = 1.11 \). We can show analytically that for \( d > 1 \) we will always have two solutions and for \( d \leq 1 \) we can only obtain a single solution. Notice that the function \( f(\alpha) = \alpha \) always intersects the function \( g_+(\alpha) = 1 + e^{-\alpha d} \). It is enough to show that for \( d > 1 \), \( f(\alpha) \) intersects \( g_-(\alpha) = 1 - e^{-\alpha d} \) once and for \( d \leq 1 \) there are no intersections.

\[\text{Notice that this equation has } d \text{ as a parameter. We can analyze this condition graphically for } d = 1 \text{ and } d = 2. \text{ This is shown in Fig. [12.2]}
\]

For \( d = 1 \) we have a single solution given by \( \alpha = 1.47 \), while for \( d = 2 \) we have two solutions given by \( \alpha_1 = 0.80 \) and \( \alpha_2 = 1.11 \). We can show analytically that for \( d > 1 \) we will always have two solutions and for \( d \leq 1 \) we can only obtain a single solution. Notice that the function \( f(\alpha) = \alpha \) always intersects the function \( g_+(\alpha) = 1 + e^{-\alpha d} \). It is enough to show that for \( d > 1 \), \( f(\alpha) \) intersects \( g_-(\alpha) = 1 - e^{-\alpha d} \) once and for \( d \leq 1 \) there are no intersections.

\[\text{Notice that this equation has } d \text{ as a parameter. We can analyze this condition graphically for } d = 1 \text{ and } d = 2. \text{ This is shown in Fig. [12.2]}
\]

\[\text{For } d = 1 \text{ we have a single solution given by } \alpha = 1.47 \text{, while for } d = 2 \text{ we have two solutions given by } \alpha_1 = 0.80 \text{ and } \alpha_2 = 1.11. \text{ We can show analytically that for } d > 1 \text{ we will always have two solutions and for } d \leq 1 \text{ we can only obtain a single solution. Notice that the function } f(\alpha) = \alpha \text{ always intersects the function } g_+(\alpha) = 1 + e^{-\alpha d}. \text{ It is enough to show that for } d > 1, f(\alpha) \text{ intersects } g_-(\alpha) = 1 - e^{-\alpha d} \text{ once and for } d \leq 1 \text{ there are no intersections.}
\]

\[\text{Notice that this equation has } d \text{ as a parameter. We can analyze this condition graphically for } d = 1 \text{ and } d = 2. \text{ This is shown in Fig. [12.2]}
\]

\[\text{For } d = 1 \text{ we have a single solution given by } \alpha = 1.47 \text{, while for } d = 2 \text{ we have two solutions given by } \alpha_1 = 0.80 \text{ and } \alpha_2 = 1.11. \text{ We can show analytically that for } d > 1 \text{ we will always have two solutions and for } d \leq 1 \text{ we can only obtain a single solution. Notice that the function } f(\alpha) = \alpha \text{ always intersects the function } g_+(\alpha) = 1 + e^{-\alpha d}. \text{ It is enough to show that for } d > 1, f(\alpha) \text{ intersects } g_-(\alpha) = 1 - e^{-\alpha d} \text{ once and for } d \leq 1 \text{ there are no intersections.}
\]

\[\text{Notice that this equation has } d \text{ as a parameter. We can analyze this condition graphically for } d = 1 \text{ and } d = 2. \text{ This is shown in Fig. [12.2]}
\]

\[\text{For } d = 1 \text{ we have a single solution given by } \alpha = 1.47 \text{, while for } d = 2 \text{ we have two solutions given by } \alpha_1 = 0.80 \text{ and } \alpha_2 = 1.11. \text{ We can show analytically that for } d > 1 \text{ we will always have two solutions and for } d \leq 1 \text{ we can only obtain a single solution. Notice that the function } f(\alpha) = \alpha \text{ always intersects the function } g_+(\alpha) = 1 + e^{-\alpha d}. \text{ It is enough to show that for } d > 1, f(\alpha) \text{ intersects } g_-(\alpha) = 1 - e^{-\alpha d} \text{ once and for } d \leq 1 \text{ there are no intersections.}
\]
Figure 12.2: Condition for Alpha for \( N = 2 \)

The red line represents the equation \( f(\alpha) = \alpha \) while the blue curve represents \( g(\alpha) = 1 \pm e^{-\alpha d/2} \).

The top function is \( 1 + e^{-\alpha d/2} \), and the bottom one is \( 1 - e^{-\alpha d/2} \).

**Claim:** For \( d > 1 \), \( f(\alpha) \) intersects \( g_{-}(\alpha) = 1 - e^{-\alpha d} \) once and for \( d \leq 1 \) there are no intersections.

**Proof.** Observe that \( g_{-}(\alpha) \leq 1 \) and that \( g'_{-}(0) = d \). Thus for \( d > 1 \) we see that \( g_{-}(\alpha) \) is strictly larger than \( f(\alpha) \) for sufficiently small \( \alpha \). Therefore, there exists an intersection point for some \( \alpha < 1 \). On the other hand for \( d \leq 1 \) we see that \( g_{-}(\alpha) \) is less than \( f(\alpha) \) for all \( \alpha > 0 \) and there are no intersections.

It is also useful to look at the probability distributions for \( N = 2 \) and \( d = 2 \) as shown in Fig. 12.3. Notice that for smaller values of \( \alpha \) the particle seems to be more localized in comparison to larger values of \( \alpha \). This can be explained using the fact that smaller \( \alpha \) correspond to larger energies that arise from the particle being confined to a particular region in space.

The results for \( N = 2 \) could be said to model a diatomic molecule. In fact, we have shown that for small values of \( d \) the molecule loses a bound state which might be an indication of why certain molecules have specific bond lengths.

### 12.4 Numerical Computations of Eigenvalues and Probability Distributions for large \( N \): Band Structures

**Note:** All calculations and plots have been made using Matlab and we encourage the reader to reproduce them.
Recall an earlier statement that quantization comes from the square integrability condition,

\[
T_{2,2}^N = 0. \quad (12.24)
\]

Clearly \(T_{2,2}^N\) depends on \(N\) because that is the number of times we exponentiate the matrix. The matrix itself depends on the parameter \(d\), the distance between neighboring potentials, as well as the value of \(\alpha\) which determines the energy. Therefore,

\[
T_{2,2}^N = f(N, d, \alpha). \quad (12.25)
\]

Fixing \(N\) and \(d\), we can find \(\alpha\)’s that are solutions to (12.25). These are simply the eigenvalues of the matrix. Computing them can be done efficiently using computer iterative methods. We proceed as follows.

1. Create a vector of equally spaced possible values for \(\alpha\) and define a function \(f(N, d, \alpha)\) that takes the (second row, second column) entry of the matrix \(T^N\) given parameters \(N\) and \(d\) using \(\alpha\) as the variable. You will have values for this function for various \(\alpha\).

2. Next we construct a code that takes in values of \(\alpha\) where the function is close to zero. This can be done using a bisection method.

3. Finally, we obtain the eigenvalues.

4. Now we can find these eigenvalues as a function of \(d\) and \(N\) when one of them is a dynamic parameter and analyze the physics of the results we obtain.

The plots in Fig. 12.4 show all eigenvalues as a function of the number of potentials for a fixed distance between neighboring potentials \((d = 4)\).
There are two important features to mention about these plots.

1. Notice that the values for the eigenvalues are bounded above and below, this is hard to prove analytically but we can see it from the plots.

2. The number of eigenvalues is directly proportional to the number of potentials.

Combining these two features we observe that the density of eigenvalues increases as $N$ increases and a band of energies arises. You can infer from these results that for a solid having about $10^{23}$ atoms, the quantization of eigenvalues becomes a continuous band of allowed energies. This interesting result is the starting point for an area of physics called Solid State Physics.

In the previous section we observed that for $N = 2$ we either have one eigenvalue or two eigenvalues depending on the value for $d$. Now, we can numerically determine at which values of $d$ we lose a bound state for a given $N$. In order to do this, we have plotted the eigenvalues as a function of $d$ for $N = 3, 4, 5, 10$ in Fig. [12.5]

- For $N = 2$, we lose a bound state at $d = 1$.
- For $N = 3$, we lose a bound state at $d = 0.5$ and $d = 1.5$.
- For $N = 5$, we lose a bound state at $d = 0.4$, $d = 0.85$, $d = 1.6$ and $d = 2$.

It is also useful to point out that as $d$ increases all the eigenvalues get closer and closer to 1. This is an expected result since as we increase the distance between the potentials we decrease the chance for the wavefunction interactions to overlap, hence they act as independent single delta function potentials each with a corresponding eigenvalue of $\alpha = 1$. 

130
Figure 12.5: Eigenvalues of $\alpha$ as a function of $d$ for various fixed values of $N$. 
Finally we can plot all the probability distributions corresponding to the various values of $\alpha$ for $d = 2.5$ and $N = 5$. This is shown in Fig. [12.6]
Figure 12.6: Probability Distributions for various $\alpha$ for $d = 2.5$ and $N = 5$
Chapter 13

Solid State Physics Using the Delta Function Potential

In the previous section we discussed the behavior of a particle under the influence of many equally spaced delta function potentials and we observed that the number of eigenenergies grows linearly with the number of potentials. Furthermore, the allowed energies were confined to a finite region of space and hence we obtained a band of possible energies. In this section, we will do a similar problem that cleverly extends the number of potentials to infinity by considering equally spaced delta potentials in a periodic box. We will observe some interesting features that arise and are related to our observations from the previous section. Moreover, we will understand a powerful tool in Solid State Physics known as Bloch’s Theorem that can be used whenever we have periodicity conditions in a problem.

The easiest problem to handle is that of a periodic box with no potential i.e a free particle in a periodic box so we will begin with that.

13.1 Energy levels for a free particle in a one dimensional periodic box

Consider a free particle inside a one dimensional periodic box of length $L$. The Schrödinger Equation for the particle is the same as the one that we discussed previously

$$-rac{\hbar^2}{2m} \frac{d^2 \psi(x)}{dx^2} = E \psi(x).$$

(13.1)

We already found that the solutions to the differential equation are

$$\psi(x) = Ae^{i\alpha x} + Be^{-i\alpha x}$$

(13.2)

where $\alpha^2 = \frac{2mE}{\hbar^2}$ because exponentials do not satisfy periodicity conditions. Unlike the previous case of a free particle, here the periodicity of the box requires that the wavefunction obey the following conditions

$$\psi(x) = \psi(x + L)$$
\[
\left( \frac{d\psi}{dx} \right) \bigg|_{x=0} = \left( \frac{d\psi}{dx} \right) \bigg|_{x=L}.
\]

(13.3)

Therefore, we have that

\[
\begin{align*}
A + B &= Ae^{i\alpha L} + Be^{-i\alpha L} \\
A - B &= Ae^{i\alpha L} + Be^{-i\alpha L}
\end{align*}
\]

(13.4)

which in matrix notation can be written as

\[
\begin{bmatrix}
1 - e^{i\alpha L} & 1 - e^{-i\alpha L} \\
1 - e^{i\alpha L} & -1 + e^{-i\alpha L}
\end{bmatrix}
\begin{bmatrix}
A \\
B
\end{bmatrix}
= \begin{bmatrix}
0 \\
0
\end{bmatrix}.
\]

(13.5)

If we want to obtain non-trivial values\(^1\) for \(A\) and \(B\), we must require the determinant to be identically equal to zero. This is satisfied if

\[
\cos(\alpha L) = 0
\]

(13.6)

or equivalently if

\[
\alpha L = 2\pi n
\]

(13.7)

where \(n\) is an arbitrary integer. This implies that the energy of the particle is given by

\[
E = \frac{\pi^2 \hbar^2 n^2}{2mL^2}.
\]

(13.8)

Thus we conclude that the periodicity of the wavefunction gives rise to the quantization of energy.

13.2 Derivation of Bloch’s Theorem

We will take this opportunity to discuss the problem of a periodic potential in a periodic box. We are not going to bother discussing this process in three dimensions because the one dimensional version already contains all the interesting physical results. We will revisit the three dimensional case later.

First, we will set up the problem. Consider a set of \(N\) equally separated delta function potentials with equal strength in a periodic box of length \(L\). This can be represented as shown in Fig. 13.1. Due to the rotational symmetry of the problem we can arbitrarily choose \(x = 0\) to lie at the position of one of the potentials as shown in the Fig.. It follows that the other potentials are located at \(x = d\), \(x = 2d\) and so forth. The symmetry tells us that the potential must have the following property

\[
V(x) = V(x + d).
\]

(13.9)

The Schrödinger Equation at an arbitrary point \(x\) is given by

\[
-\frac{\hbar}{2m} \frac{d^2}{dx^2} \psi(x) + V(x) \psi(x) = E \psi(x).
\]

(13.10)

\(^1\)Remember from Linear Algebra that a homogeneous system of linear equations has non trivial solutions if the determinant is equal to zero.
Figure 13.1: Here, the periodic box is represented by a circular boundary of length $L$

Each arrow represents a delta potential function. The distance between each potential is $d = \frac{L}{N}$.

Similarly the Schrödinger Equation at a position $d$ units to the left of the origin, that is at a point $x + d$, is given by

$$-\hbar \frac{d^2}{2m} \psi(x + d) + V(x + d)\psi(x + d) = E\psi(x + d).$$  \hspace{1cm} (13.11)

Using the relation (13.9), the previous equation can be written as

$$-\hbar \frac{d^2}{2m} \psi(x + d) + V(x)\psi(x + d) = E\psi(x + d).$$  \hspace{1cm} (13.12)

Notice that that (13.10) and (13.12) are the same differential equation. Therefore, if the wave function has a solution at $x$ then it must also have a solution at $x + d$. This solution can at most differ by a multiplicative constant that we will call $C$ so we have

$$\psi(x) = C\psi(x + d).$$  \hspace{1cm} (13.13)

Notice that this relation is true for all values of $x$. Therefore, it follows that

$$\psi(x) = C\psi(x + d) = C^2\psi(x + 2d).$$  \hspace{1cm} (13.14)

If we repeat this process several times we find that in general

$$\psi(x) = C^k\psi(x + kd).$$  \hspace{1cm} (13.15)

Notice this implies that

$$\psi(x) = C^N\psi(x + L).$$  \hspace{1cm} (13.16)

However, if we go around the circle once (we move a distance $L$), we arrive at the same physical point (because of the periodicity of our problem.) Thus,

$$\psi(x) = \psi(x + L).$$  \hspace{1cm} (13.17)
This implies that

$$C^N = 1.$$  \hspace{1cm} (13.18)

We can ask what the probability distribution looks like for this problem. Once again, because of the symmetry of the problem we have that

$$P(x) = P(x + d)$$  \hspace{1cm} (13.19)

or by definition of the probability distribution,

$$\psi^*(x)\psi(x) = \psi^*(x + d)\psi(x + d).$$  \hspace{1cm} (13.20)

Substituting (13.13) into the previous equation, we obtain

$$C^*C = 1.$$  \hspace{1cm} (13.21)

So $C$ must be a constant that has the properties given by (13.18) and (13.21). Thus, $C$ must be an $N^{th}$ root of unity given by $C = e^{i2\pi k/N}$ where $k = 0, 1, 2, ..., N - 1$. We can check that $C^N = 1$ as follows

$$C^N = e^{i2\pi k} = \cos(2\pi k) + i\sin(2\pi k) = 1.$$  \hspace{1cm} (13.22)

Thus (13.13) leads to the following relation between neighboring wavefunctions

$$\psi(x) = e^{i2\pi k/N} \psi(x + d).$$  \hspace{1cm} (13.23)

This last equation is known as Bloch’s Theorem. This relation holds for any periodic potential where $d$ is the period and $N$ is the number of potentials. Notice that each $k$ will give rise to a different relation. Therefore, we actually have $N$ independent relations. This will turn out to be an important result when we solve for the energy eigenvalues.

### 13.3 Bound and unbound energy levels in the presence of a periodic delta function potential

Now that we have derived the useful relation that is Bloch’s Theorem, we can use it to conveniently solve our problem. It is enough to solve the problem in a neighborhood around a single potential because then we can use Bloch’s Theorem to recursively obtain the wavefunction everywhere. We are going to focus on the energy eigenvalues instead of on the wavefunction itself.

---

**Solving the Periodicity Problem:** Consider the wavefunction on the interval $-d/2 \leq x \leq d/2$. We found in the previous two sections that we have solutions corresponding to negative energies $E < 0$ (bound states) and to positive energies $E > 0$ (unbound states). The corresponding solutions are circular (i.e. sines and cosines), and hyperbolic (i.e. exponential) functions, respectively. First we consider the hyperbolic solutions

$$\psi_L(x) = A_L e^{\alpha x} + B_L e^{-\alpha x} \quad \text{for} \quad x < 0$$

\(^2\)Notice that the possible values of $k$ stop at $N - 1$ since $k = 0$ and $k = N$ give the same solution.

\(^3\)Recall that we do not have to worry about normalization since we are working in a finite interval. Normalization would be an overall scaling factor if anything.
\[ \psi_R(x) = A_R e^{\alpha x} + B_R e^{-\alpha x} \quad \text{for} \quad x > 0 \] (13.24)

where \( \alpha^2 = -\frac{2mE}{\hbar^2} \).

We now must satisfy certain conditions. First, (13.13) requires that

\[ \psi(-d/2) = e^{\frac{i\pi k}{N}} \psi(d/2). \] (13.25)

In addition to this condition we also need to satisfy periodicity of the first derivative by continuity and the jump condition. Mathematically we have

\[ \left. \frac{d\psi}{dx} \right|_{x=-d/2} = e^{\frac{i\pi k}{N}} \left. \frac{d\psi}{dx} \right|_{x=d/2}. \]

\[ \lim_{\epsilon \to 0} \left( \left. \frac{d\psi}{dx} \right|_{x=-\epsilon} - \left. \frac{d\psi}{dx} \right|_{x=\epsilon} \right) = 2\psi(0). \] (13.26)

Notice that we have five variables: \( A_L, B_L, A_R, B_R \) and \( \alpha \) and four independent conditions, so there will be a unique solution to this problem if an additional condition is satisfied. This extra condition comes from the need for the wavefunction to be normalized.

For computational convenience, we can define

\[ \upsilon = e^{\alpha d/2}. \] (13.27)

After substituting this variable into (13.24) and (13.24) the previous conditions become

\[ A_L \frac{1}{\upsilon} + B_L \upsilon = C \left( A_R \upsilon + B_R \frac{1}{\upsilon} \right) \]
\[ A_L \frac{1}{\upsilon} - B_L \upsilon = C \left( A_R \upsilon - B_R \frac{1}{\upsilon} \right) \]
\[ A_L + B_L = A_R + B_R \]
\[ \alpha(A_L - B_L - A_R + B_R) = 2(A_L + B_L). \] (13.28)

These equations can be written in matrix notation as

\[
\begin{bmatrix}
\frac{1}{\upsilon} & \upsilon & -C \upsilon & -\frac{C}{\upsilon} \\
\frac{1}{\upsilon} & -\upsilon & -C \upsilon & \frac{C}{\upsilon} \\
1 & 1 & 1 & -1 \\
\alpha - 2 & -(\alpha + 2) & -\alpha & \alpha
\end{bmatrix}
\begin{bmatrix}
A_L \\
B_L \\
A_R \\
B_R
\end{bmatrix}
= \begin{bmatrix}
0 \\
0 \\
0
\end{bmatrix}.
\] (13.29)

There exist non-trivial solutions to this system for certain values of \( \alpha \) such that the determinant of the matrix is identically equal to zero. The problem then reduces to the following

\[ C + \frac{1}{C} = \upsilon^2 + \frac{1}{\upsilon^2} - \frac{1}{\alpha} \left( \upsilon^2 - \frac{1}{\upsilon^2} \right). \] (13.30)

Replacing this with our original functions, we have the following relation

\[ e^{\frac{i\pi k}{N}} + e^{-\frac{i\pi k}{N}} = e^{\alpha d} + e^{-\alpha d} - \frac{1}{\alpha} \left( e^{\alpha d} - e^{-\alpha d} \right) \] (13.31)

which can be written as

\[ \cos \left( \frac{2\pi k}{N} \right) = \cosh (\alpha d) - \frac{1}{\alpha} \sinh (\alpha d). \] (13.32)
We can now consider the circular solutions corresponding to $E > 0$ given by

$$
\psi_L = A_L e^{i\alpha x} + B_L e^{-i\alpha x} \quad \text{for} \quad x < 0
$$

$$
\psi_R = A_R e^{i\alpha x} + B_R e^{-i\alpha x} \quad \text{for} \quad x > 0.
$$

(13.33)

If we perform the substitution $\gamma = i\alpha$, we obtain

$$
\psi_L = A_L e^{\gamma x} + B_L e^{-\gamma x}
$$

$$
\psi_R = A_R e^{\gamma x} + B_R e^{-\gamma x}.
$$

(13.34)

This solution needs to satisfy the same conditions as before so

$$
\cos\left(\frac{2\pi k}{N}\right) = \cosh(\gamma d) - \frac{1}{\alpha} \sinh(\gamma d)
$$

(13.35)

or in terms of $\alpha$ we obtain through the use of hyperbolic functions

$$
\cos\left(\frac{2\pi k}{N}\right) = \cos(\alpha d) - \frac{1}{\alpha} \sin(\alpha d).
$$

(13.36)

In summary the conditions the solutions must satisfy for both positive and negative energies are

$$
S_k = \begin{cases} 
\cos\left(\frac{2\pi k}{N}\right) = \cosh(\alpha d) - \frac{1}{\alpha} \sinh(\alpha d), & \text{if } E > 0. \\
\cos\left(\frac{2\pi k}{N}\right) = \cos(\alpha d) - \frac{1}{\alpha} \sin(\alpha d), & \text{if } E < 0.
\end{cases}
$$

(13.37)

We have intentionally included a subscript $k$ to emphasize the fact that each $k$ will give rise to a different set of eigenenergies.

### 13.4 Numerical Calculations of Energy Eigenvalues for Periodic Delta Potentials

The eigenenergies for the bounded case ($E < 0$) as a function of $N$ are shown in Fig. 13.2. Notice that Fig. 13.2 resembles the plot that we obtained in the previous section almost identically. In fact, we can now explain the mathematical reason behind the lower and upper bounds for the eigenenergies. In order to do so let us consider the following function

$$
f(\alpha) = \cosh(\alpha d) - \frac{1}{\alpha} \sinh(\alpha d).
$$

(13.38)

For large values of $N$ the possible values for $\cos\left(\frac{2\pi k}{N}\right)$ will lie between $-1$ and $1$. The allowed values for $\alpha$ will be those such that $f(\alpha)$ is between $-1$ and $1$. Fig. 13.3 shows $f(\alpha)$ for various values of $d$. In Fig. 13.3 we observe that for $d = 3$ and $d = 4$ the possible $\alpha$’s have an upper and a lower bound that are both different from 0. However, for $d = 1.5$ and $d = 2$ the lower bound is precisely equal to 0. In general, different values of $d$ will give rise to different sets of bounds.

Now we can plot the energy eigenvalues for the unbounded case ($E > 0$) as a function of $N$ as shown in Fig. 13.4. In Fig. 13.4 we observe bands in the eigenenergies that possess their own upper and lower bounds. The periodicity of the circular functions is the reason we observe this band structure.
Figure 13.2: Eigenenergies for bounded case as a function of $N$ for $d = 4$

![Figure 13.2: Eigenenergies for bounded case as a function of $N$ for $d = 4$](image)

Figure 13.3: Plot of $f(\alpha)$ for different values of $d$ showing bounds for $\alpha$

![Figure 13.3: Plot of $f(\alpha)$ for different values of $d$ showing bounds for $\alpha$](image)
Figure 13.4: Eigenenergies for unbounded case as a function of $N$ for $d = 4$

Figure 13.5: Eigenenergies for both bound and unbounded case as a function of $d$ for $N = 100$
Finally, we plot the eigenenergies as a function of $d$ for $N = 100$ in Fig. 13.5. Fig. 13.5 beautifully shows the band structure of a solid as a function of $d$. Notice that for $d < 2$ the bound and unbounded energies merge, as might be the case for a conductor. For $d > 2$, we observe an energy gap between the bands, as in the case of an insulator.

Even this simple model of a periodic solid helps us understand its band structure. This shows the power of Bloch’s Theorem in solving some analytically complex problems. In the next sections we will discuss problems involving different types of potentials that will allows as to “tune” the energy band gaps and create a more versatile model of conductors and eventually lead to the creation of a semiconductor.
Chapter 14

Semiconductors: A Problem in Tunable Band Gaps

In this section we will present a model for a one dimensional solid with a tunable parameter which, when tuned appropriately, makes the solid a semiconductor.

In the previous two sections we encountered the concepts of transfer matrix and Bloch’s theorem\(^1\). Both of these tools will be used to solve the semiconductor problem. Recall that in our previous models we could only vary the distance between the potentials. A more complicated model will be discussed in the following pages.

14.1 Model Definitions

Consider a solid made of periodic two delta function potentials. The potentials are set apart by a distance \(b\) and the period is \(d\). The Schrödinger Equation is

\[
-\frac{\hbar^2}{2m} \frac{d^2 \psi}{dx^2} + V(x)\psi(x) = E\psi(x).
\] (14.1)

The potential is given by

\[
V(x) = -aV_0\delta(x - \frac{b}{2}) - aV_0\delta(x + \frac{b}{2}), \quad -\frac{d}{2} \leq x \leq \frac{d}{2}
\] (14.2)

such that \(V(x + d) = V(x)\).

As in our previous problems we must impose the following conditions in the solution. Continuity and jump conditions

\[
\psi(x)\bigg|_{\pm\frac{b}{2}^-} = \psi(x)\bigg|_{\pm\frac{b}{2}^+},
\] (14.3)

\[
\left. \frac{d\psi}{dx} \right|_{\pm\frac{b}{2}^+} - \left. \frac{d\psi}{dx} \right|_{\pm\frac{b}{2}^-} = -\psi(\pm\frac{b}{2})
\] (14.4)

\(^1\)The reader is strongly encouraged to review these ideas before proceeding.
where we have set $\beta^2 = \frac{2amV_0}{R^2} = 1$.

According to Bloch’s theorem, which we discussed previously, the wave function obeys the following quasi-periodic conditions,

$$
\frac{d\psi}{dx} \bigg|_{\frac{d}{2}} = C \frac{d\psi}{dx} \bigg|_{-\frac{d}{2}}; \quad \psi\left(-\frac{d}{2}\right) = C\psi\left(\frac{d}{2}\right) \quad (14.5)
$$

where $C = e^{\frac{2\pi k}{N}}$.

### 14.2 Outline of method.

First, we study the wavefunction around two neighboring delta potentials and consider the following solutions for the bound states.

$$
\psi_L = A_L e^{-\alpha x} + B_L e^{\alpha x} \quad (14.6)
$$

$$
\psi_M = A_M e^{-\alpha x} + B_M e^{\alpha x} \quad (14.7)
$$

$$
\psi_R = A_R e^{-\alpha x} + B_R e^{\alpha x} \quad (14.8)
$$

We solve (14.3) through (14.5) using standard methods. The first two conditions can be recast using transfer matrices:

$$
\begin{bmatrix} A_M \\ B_M \end{bmatrix} = \begin{bmatrix} (1 - \frac{1}{v^2 \alpha}) & -\frac{1}{v^2 \alpha} \\ \frac{v^2}{\alpha} & (1 + \frac{1}{\alpha}) \end{bmatrix} \begin{bmatrix} A_L \\ B_L \end{bmatrix} \quad (14.9)
$$

where $v = e^{\alpha b/2}$.

Similarly,

$$
\begin{bmatrix} A_R \\ B_R \end{bmatrix} = \begin{bmatrix} (1 - \frac{1}{v^2 \alpha}) & -\frac{v^2}{\alpha} \\ \frac{v^2}{\alpha} & (1 + \frac{1}{\alpha}) \end{bmatrix} \begin{bmatrix} A_M \\ B_M \end{bmatrix} \quad (14.10)
$$

Bloch’s condition can also be written using a matrix

$$
\begin{bmatrix} A_R \\ B_R \end{bmatrix} = \begin{bmatrix} \frac{C}{v^2} & 0 \\ 0 & Cu^2 \end{bmatrix} \begin{bmatrix} A_L \\ B_L \end{bmatrix} \quad (14.11)
$$

where $u = e^{\alpha d/2}$.

From these matrix conditions and to obtain non-trivial solutions we need that

$$
\det(B - T_2 T_1) = 0 \quad (14.12)
$$

where

$$
B = \begin{bmatrix} \frac{C}{v^2} & 0 \\ 0 & Cu^2 \end{bmatrix}, \quad T_1 = \begin{bmatrix} (1 - \frac{1}{v^2 \alpha}) & -\frac{1}{v^2 \alpha} \\ \frac{v^2}{\alpha} & (1 + \frac{1}{\alpha}) \end{bmatrix} \quad \text{and} \quad T_2 = \begin{bmatrix} (1 - \frac{1}{\alpha}) & -\frac{v^2}{\alpha} \\ \frac{1}{v^2 \alpha} & (1 + \frac{1}{\alpha}) \end{bmatrix} \quad (14.13)
$$

The construction is entirely analogous to that performed on the previous section. However, in this case we do not center the wavefunctions; thus, obtaining two different transfer matrices.
After some algebra, the following condition is obtained,

$$\cos\left(\frac{2\pi k}{N}\right) = -\frac{\cosh(\alpha(d - 2b))}{\alpha^2} + \cosh(\alpha d) - \frac{2\sinh(\alpha d)}{\alpha} + \frac{\cosh(\alpha d)}{\alpha^2}$$  \hspace{1cm} (14.14)$$

Likewise for the unbound states we have \(^3\)

$$\cos\left(\frac{2\pi k}{N}\right) = \frac{\cos(\alpha(d - 2b))}{\alpha^2} + \cos(\alpha d) - \frac{2\sin(\alpha d)}{\alpha} - \frac{\cos(\alpha d)}{\alpha^2}$$  \hspace{1cm} (14.15)$$

It is useful to perform the changes of variables \(y = \alpha d\) and let \(r = \frac{b}{a}\). Then we have

$$\cos\left(\frac{2\pi k}{N}\right) = -\frac{d^2 \cosh(y(1 - 2r))}{y^2} + \cosh y - \frac{2d \sinh y}{y} + \frac{d^2 \cosh y}{y^2}$$  \hspace{1cm} (14.16)$$

$$\cos\left(\frac{2\pi k}{N}\right) = \frac{d^2 \cos(y(1 - 2r))}{y^2} + \cos y - \frac{2d \sin y}{y} - \frac{d^2 \cos y}{y^2}$$  \hspace{1cm} (14.17)$$

The following plots show the eigenenergies as a function of \(r\), the ratio period to distance between adjacent potentials. In each figure \(r\) ranges from \([0, \frac{1}{2}]\). The variable \(r = 0\) corresponds to the case where the distance between two potentials is 0. This leads to a periodic single delta function potential with double the strength and the same period. For \(r = \frac{1}{2}\) the parameters \(a\) and \(b\) are equal which corresponds to another case of the single delta function potential where the strength is the same but the period has now been halved. Recall that for a single delta function potential the gaps between bands could not be closed to make a semiconductor. By varying the parameter \(r\), thus going from one special case of a single delta function potential to another, we were able to close the bands and form conditions for a semiconductor.

\(^3\)This result can be easily obtained by performing the change of variables \(\gamma = i\alpha\)
Figure 14.2: Eigenenergies for both bound and unbounded case as a function of $r$ for $N = 100$ and $d = 2.9$. 

![Graph showing eigenenergies as a function of ratio period to distance between neighboring delta potentials for N=100.](image-url)
Chapter 15

Quantum Mechanics in Two Dimensions

So far most of our discussion of wave mechanics has been confined to one dimensional examples. In the next two sections, we will continue our discussion to include two and three dimensions. One of the important consequences of extending our theory to higher dimensions is that new quantities will come into play such as angular momentum. Also, the idea of symmetry will be discussed in more detail as well as its importance to various problems. To facilitate our discussion we will start with an introduction to Lagrangian Mechanics.

15.1 Lagrangian Mechanics

Lagrangian Mechanics is a reformulation of Newtonian Mechanics that is used to simplify problems involving constraints and symmetries. The Lagrangian is an explicit function of the coordinates of a particle and its first derivatives defined as

$$L = T - V = \frac{1}{2} m (\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2) - V(x_1, x_2, x_3). \quad (15.1)$$

The various partial derivatives of the Lagrangian are

$$\frac{\partial L}{\partial \dot{x}_i} = m \dot{x}_i = p_i,$$

$$\frac{\partial L}{\partial x_i} = -\frac{\partial V}{\partial x_i} = F_i. \quad (15.2)$$

Therefore, the equation of motion for the particle $F_i = \dot{p}_i$ can be expressed as

$$\frac{\partial L}{\partial x_i} = \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}_i} \right). \quad (15.3)$$

This last equation is known as the Euler-Lagrange Equation.

\footnote{It is important to notice the minus sign in this equation. This is not the total energy of the particle.}
15.2 The Schrödinger Equation in Two Dimensions

Before discussing quantum mechanics in two dimensions we will revisit the associated classical problem but now using Lagrangian Mechanics. The Lagrangian of a particle confined to a plane under the influence of a central potential is written in polar coordinates as

\[ L = \frac{1}{2} m \left( \dot{r}^2 + r^2 \dot{\phi}^2 \right) - V(r). \]  

(15.4)

The motion of the particle is found using the Euler-Lagrange equation given by

\[ \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}_i} \right) = \frac{\partial L}{\partial x_i}. \]  

(15.5)

It follows that the equation of motion for \( \phi \) is

\[ \frac{d}{dt} \left( m r^2 \dot{\phi} \right) = \frac{dL}{dt} = 0. \]  

(15.6)

Recall that the quantity \( L = m r^2 \dot{\phi} \) is the angular momentum. Hence, we observe that angular momentum is a constant,

\[ L = \text{constant}. \]  

(15.7)

This is a natural consequence of having a central potential and a direct consequence of the rotational symmetry of the problem. The equation of motion for \( r \) is

\[ m \ddot{r} = m \dot{r}^2 - \frac{dV}{dr} \]  

(15.8)

or in terms of angular momentum

\[ m \ddot{r} = \frac{L^2}{m r^3} - \frac{dV}{dr} = - \frac{d}{dt} \left( V + \frac{L^2}{2m r^2} \right) = - \frac{dU}{dt} \]  

(15.9)

where \( U = V + \frac{L^2}{2m r^2} \) and is labeled as our effective potential.\(^2\) Using Lagrangian Mechanics we have successfully reduced our two dimensional problem to a one dimensional one. This will always be the case when we discuss motion in a central potential.

Now we return to our original discussion of Quantum Mechanics. The time independent Schrödinger Equation in two dimensions is

\[ - \frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \psi(x, y) + V(x, y)\psi(x, y) = E\psi(x, y). \]  

(15.10)

If we consider the isotropic harmonic oscillator potential in two dimensions, the Schrödinger Equation is

\[ - \frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \psi(x, y) + \frac{1}{2} m \omega^2 \left( x^2 + y^2 \right) \psi(x, y) = E\psi(x, y). \]  

(15.11)

\(^2\)This extra term in the potential comes from the fact that we have motion in two dimensions along with the fact we have angular momentum conservation.
This equation is separable so we can consider a solution of the form $\psi(x,y) = X(x)Y(y)$. Substituting this solution into (15.11) and after separating variables we obtain

$$-\frac{\hbar^2}{2m} \frac{1}{X(x)} \frac{d^2 X(x)}{dx^2} + \frac{1}{2} m \omega^2 x^2 = E + \frac{\hbar^2}{2m} \frac{1}{Y(y)} \frac{d^2 Y(y)}{dy^2} - \frac{1}{2} m \omega^2 y^2. \quad (15.12)$$

The left hand side of the equation depends only on $x$, while the right hand side depends only on $y$. Since this equation is valid for all $x$ and $y$, each side must equal a constant that we will call $E_1$. Then we obtain two ordinary differential equations

$$-\frac{\hbar^2}{2m} \frac{1}{X(x)} \frac{d^2 X(x)}{dx^2} + \frac{1}{2} m \omega^2 x^2 = E_1$$
$$-\frac{\hbar^2}{2m} \frac{1}{Y(y)} \frac{d^2 Y(y)}{dy^2} + \frac{1}{2} m \omega^2 y^2 = E - E_1 = E_2 \quad (15.13)$$

where we let $E_2 = E - E_1$. The previous equations can be written in a more familiar form as

$$\frac{d^2 X(x)}{dx^2} + \left( \frac{2mE_1}{\hbar^2} - \frac{m \omega^2 x^2}{\hbar^2} \right) X(x) = 0$$
$$\frac{d^2 Y(y)}{dy^2} + \left( \frac{2mE_2}{\hbar^2} - \frac{m \omega^2 y^2}{\hbar^2} \right) Y(y) = 0. \quad (15.14)$$

In (10.26) we observed that the possible energy eigenvalues are

$$E_1 = \left( n_1 + \frac{1}{2} \right) \hbar \omega$$
$$E_2 = \left( n_2 + \frac{1}{2} \right) \hbar \omega. \quad (15.15)$$

Then, the total energy of the oscillator is given as

$$E = (1 + n_1 + n_2) \hbar \omega \quad (15.16)$$

where $n_1$ and $n_2$ are positive integers.

In Table 15.1 we show the energies for different combinations of values of $n_1$ and $n_2$. The degeneracy refers to the number of possible states with the same energy. This is a consequence of

<table>
<thead>
<tr>
<th>$E$</th>
<th>$n_1$</th>
<th>$n_2$</th>
<th>Energy Degeneracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hbar \omega$</td>
<td>0</td>
<td>0</td>
<td>1-fold</td>
</tr>
<tr>
<td>$2\hbar \omega$</td>
<td>1</td>
<td>0</td>
<td>2-fold</td>
</tr>
<tr>
<td>$2\hbar \omega$</td>
<td>0</td>
<td>1</td>
<td>2-fold</td>
</tr>
<tr>
<td>$3\hbar \omega$</td>
<td>1</td>
<td>1</td>
<td>3-fold</td>
</tr>
<tr>
<td>$3\hbar \omega$</td>
<td>2</td>
<td>0</td>
<td>3-fold</td>
</tr>
<tr>
<td>$3\hbar \omega$</td>
<td>0</td>
<td>2</td>
<td>3-fold</td>
</tr>
</tbody>
</table>

Table 15.1: Eigenenergies for the two dimensional isotropic harmonic oscillator
the symmetry of the problem and we will better understand this result when we transfer to polar coordinates. That said, we can write the two dimensional Schrödinger equation in polar coordinates for a general central potential. It is given by

\[ -\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \phi^2} \right) \psi(r, \phi) + V(r)\psi(r, \phi) = E\psi(r, \phi). \]  

(15.17)

If we consider a solution with separated variables \( \psi(r, \phi) = R(r)\Phi(\phi) \) then we can substitute it into (15.17) and separate variables to obtain

\[ r^2 \left[ \frac{2mE}{\hbar^2} + \frac{1}{R} \left( \frac{d^2R}{dr^2} + \frac{1}{r} \frac{dR}{dr} \right) - \frac{2mV(r)}{\hbar^2} \right] = \frac{1}{\Phi} \frac{d^2\Phi}{d\phi^2} = -m_z^2 \]  

(15.18)

where we have conveniently equated both sides to a constant \( m_z^2 \) since each side is completely independent of the other. Once again, the problem reduces to two ordinary differential equations.

- The angular part

\[ \frac{1}{\Phi} \frac{d^2\Phi}{d\phi^2} = -m_z^2 \]  

(15.19)

- and the radial part

\[ r^2 \left[ \frac{2mE}{\hbar^2} + \frac{1}{R} \left( \frac{d^2R}{dr^2} + \frac{1}{r} \frac{dR}{dr} \right) - \frac{2mV(r)}{\hbar^2} \right] = -m_z^2. \]  

(15.20)

Observe that regardless of the potential, the angular part always has the solution

\[ \Phi(\phi) = Ae^{im_z\phi} + Be^{-im_z\phi}. \]  

(15.21)

Due to periodicity conditions we require

\[ A + B = Ae^{i2\pi m_z} + Be^{-i2\pi m_z} \]

\[ A - B = Ae^{i2\pi m_z} + Be^{-i2\pi m_z}. \]  

(15.22)

We previously found that this set of equations has nontrivial solutions if

\[ 2\pi m_z = 2\pi n \implies m_z = n \]  

(15.23)

where \( n \) is an integer. Thus, we conclude

\[ \Phi(\phi) = Ce^{im_z\phi} \]  

(15.24)

where \( m_z = 0, \pm 1, \pm 2, \pm 3, ... \). This result is independent of the potential and it is a direct consequence of the symmetry of the problem as in the classical case. Therefore, we might infer that this quantity is related to the angular momentum of the particle and that it is quantized in a similar way to how energy is quantized. This result is due to the conservation of both quantities and the
rotational and time symmetry of the problem.

In order to solve the radial part of the equation we will revisit the isotropic harmonic oscillator potential: \( V(r) = \frac{1}{2} m \omega^2 r^2 \). Then our radial equation (15.20) after some algebraic manipulation becomes

\[
\frac{d^2 R}{dr^2} + \frac{1}{r} \frac{dR}{dr} - \left( \frac{m^2}{r^2} + \frac{2mE}{\hbar^2} - \frac{m^2 \omega^2}{\hbar^2} r^2 \right) R = 0. \tag{15.25}
\]

If we wish to simplify our problem, we should use natural units. We conveniently let \( r = b \rho \) then the natural choice is \( b = \sqrt{\frac{\hbar}{m \omega}} \). Furthermore, the natural change of variables for energy is \( E = \epsilon \hbar \omega \). Then our equation becomes

\[
\frac{d^2 P}{d\rho^2} + \frac{1}{\rho} \frac{dP}{d\rho} - \left( 2\epsilon - \rho^2 + \frac{m^2 z^2}{\rho^2} \right) P = 0 \tag{15.26}
\]

At this point in our discussion it shouldn’t be surprising that this equation is similar to the one obtained for the harmonic oscillator in one dimension. At the same time, the additional terms in the kinetic and potential energy are similar to those that appear in the classical case. Once again, we have successfully reduced our two dimensional problem into a one dimensional one since the motion in the angular direction is trivially fixed by the angular momentum.

It can be shown that the energy eigenvalues of this equation are given by the following relation

\[
\epsilon = 2k + |m_z| + 1 \tag{15.27}
\]

or in terms of energy

\[
E = (2k + |m_z| + 1) \hbar \omega \tag{15.28}
\]

where \( k \) is a positive integer. This is a very plausible answer because it is telling us that the energy depends on the magnitude of the angular momentum as in the classical case. However in this particular case \( k \) is a measurement of the radial kinetic energy so \( |m_z| \) a measurement of the angular kinetic energy.

As we did previously, we can make a table of the energy eigenvalues for different values of \( k \), \( n \) and \( m_z \). Moreover, we can define \( n = 2k + |m_z| \) and then it follows that \( E = (n + 1) \hbar \omega \). This agrees with the degeneracy obtained in Table 15.1. However, now we can see why we have degeneracy: it comes from the fact that energy is provided by both the radial and angular parts of the wavefunction. This is another reason to believe that \( m_z \) is related to the angular momentum of the particle.

---

3If you do not remember how to use natural units, go back to the section where we discussed the one dimensional harmonic oscillator.
Table 15.2: Eigenenergies for the two dimensional isotropic harmonic oscillator in Polar Coordinates.

<table>
<thead>
<tr>
<th>$E$</th>
<th>$n$</th>
<th>$k$</th>
<th>$m_z$</th>
<th>Energy Degeneracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h\omega$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1-fold</td>
</tr>
<tr>
<td>$2h\omega$</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2-fold</td>
</tr>
<tr>
<td>$2h\omega$</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>2-fold</td>
</tr>
<tr>
<td>$3h\omega$</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>3-fold</td>
</tr>
<tr>
<td>$3h\omega$</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>3-fold</td>
</tr>
<tr>
<td>$3h\omega$</td>
<td>2</td>
<td>0</td>
<td>-2</td>
<td>3-fold</td>
</tr>
</tbody>
</table>
Chapter 16
Quantum Mechanics in Three Dimensions

In this section we will discuss quantum mechanics in three dimensions again under central potentials. We mentioned that the most convenient way to describe quantum mechanics in three dimensions is using spherical coordinates. Thus, our first objective is to write the Schrödinger Equation in spherical coordinates.

16.1 Schrödinger Equation in Spherical Coordinates

It is important to start by specifying our convention for spherical coordinates. This is shown in Fig. [16.1].

Spherical Coordinates: The conversion between \((x, y, z)\) and \((r, \theta, \phi)\) is

\[
\begin{align*}
x &= r \sin \theta \cos \phi \\
y &= r \sin \theta \sin \phi \\
z &= r \cos \theta
\end{align*}
\]

(16.1)

and

\[
\begin{align*}
r^2 &= x^2 + y^2 + z^2 \\
\sin \theta &= \frac{\sqrt{x^2 + y^2}}{r} \\
\tan \phi &= \frac{y}{x}
\end{align*}
\]

(16.2)

Now we can calculate the various derivatives: \(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\) and \(\frac{\partial}{\partial z}\) in terms of \(\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}\) and \(\frac{\partial}{\partial \phi}\). By the chain rule we have that

\[
\begin{align*}
\frac{\partial}{\partial x} &= \left(\frac{\partial r}{\partial x}\right) \frac{\partial}{\partial r} + \left(\frac{\partial \theta}{\partial x}\right) \frac{\partial}{\partial \theta} + \left(\frac{\partial \phi}{\partial x}\right) \frac{\partial}{\partial \phi} \\
\frac{\partial}{\partial y} &= \left(\frac{\partial r}{\partial y}\right) \frac{\partial}{\partial r} + \left(\frac{\partial \theta}{\partial y}\right) \frac{\partial}{\partial \theta} + \left(\frac{\partial \phi}{\partial y}\right) \frac{\partial}{\partial \phi} \\
\frac{\partial}{\partial z} &= \left(\frac{\partial r}{\partial z}\right) \frac{\partial}{\partial r} + \left(\frac{\partial \theta}{\partial z}\right) \frac{\partial}{\partial \theta} + \left(\frac{\partial \phi}{\partial z}\right) \frac{\partial}{\partial \phi}.
\end{align*}
\]

(16.3)
In order to complete our description we need to find $\frac{\partial r}{\partial x}$, $\frac{\partial \theta}{\partial x}$, $\frac{\partial \phi}{\partial x}$, $\frac{\partial r}{\partial y}$, $\frac{\partial \theta}{\partial y}$, $\frac{\partial \phi}{\partial y}$, $\frac{\partial r}{\partial z}$, $\frac{\partial \theta}{\partial z}$ and $\frac{\partial \phi}{\partial z}$. We can obtain these by using implicit partial differentiation on the relations given in (16.2). The results are the following

$$
\begin{align*}
\frac{\partial r}{\partial x} &= \sin \theta \cos \phi \\
\frac{\partial \theta}{\partial x} &= \frac{\cos \theta \cos \phi}{r} \\
\frac{\partial \phi}{\partial x} &= -\frac{\sin \phi}{r} \\
\frac{\partial r}{\partial y} &= \sin \theta \sin \phi \\
\frac{\partial \theta}{\partial y} &= \frac{\cos \theta \sin \phi}{r} \\
\frac{\partial \phi}{\partial y} &= \frac{\cos \phi}{r} \\
\frac{\partial r}{\partial z} &= \cos \theta \\
\frac{\partial \theta}{\partial z} &= -\frac{\sin \theta}{r} \\
\frac{\partial \phi}{\partial z} &= 0.
\end{align*}
$$

(16.4)

It follows that

$$
\begin{align*}
\frac{\partial^2}{\partial x^2} &= \sin \theta \cos \phi \left( \frac{\partial^2}{\partial r^2} \right) + \frac{\cos \theta \cos \phi}{r} \frac{\partial}{\partial r} \frac{\partial}{\partial \theta} - \frac{\sin \phi}{r \sin \theta} \frac{\partial}{\partial \phi} \\
\frac{\partial^2}{\partial y^2} &= \sin \theta \sin \phi \left( \frac{\partial^2}{\partial r^2} \right) + \frac{\cos \theta \sin \phi}{r} \frac{\partial}{\partial r} \frac{\partial}{\partial \theta} + \frac{\cos \phi}{r \sin \theta} \frac{\partial}{\partial \phi} \\
\frac{\partial^2}{\partial z^2} &= \cos \theta \left( \frac{\partial^2}{\partial r^2} \right) - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta}.
\end{align*}
$$

(16.5)

Since $\frac{\partial^2}{\partial x^2} = \frac{\partial}{\partial x} \left( \frac{\partial}{\partial x} \right)$, $\frac{\partial^2}{\partial y^2} = \frac{\partial}{\partial y} \left( \frac{\partial}{\partial y} \right)$ and $\frac{\partial^2}{\partial z^2} = \frac{\partial}{\partial z} \left( \frac{\partial}{\partial z} \right)$, we have

$$
\begin{align*}
\frac{\partial^2}{\partial x^2} &= \left( \sin \theta \cos \phi \frac{\partial}{\partial r} + \frac{1}{r} \cos \theta \cos \phi \frac{\partial}{\partial \theta} - \frac{1}{r \sin \theta \sin \phi} \frac{\partial}{\partial \phi} \right)^2 \\
\frac{\partial^2}{\partial y^2} &= \left( \sin \theta \sin \phi \frac{\partial}{\partial r} + \frac{1}{r} \cos \theta \sin \phi \frac{\partial}{\partial \theta} + \frac{1}{r \sin \theta \sin \phi} \frac{\partial}{\partial \phi} \right)^2 \\
\frac{\partial^2}{\partial z^2} &= \left( \cos \theta \frac{\partial}{\partial r} - \frac{1}{r \sin \theta \sin \phi} \frac{\partial}{\partial \theta} \right)^2.
\end{align*}
$$

(16.6)
After some simplifications we obtain the following formula for the Laplacian in spherical coordinates

\[ \frac{\partial^2}{\partial r^2} = \left[ \sin^2 \theta \cos^2 \phi \frac{\partial^2}{\partial r^2} - \frac{1}{r^2} \cos \theta \sin \theta \cos^2 \phi \frac{\partial^2}{\partial \theta \partial r} + \frac{1}{r} \cos \theta \sin \theta \cos^2 \phi \frac{\partial^2}{\partial r^2} + \frac{1}{r^2} \sin \phi \cos \phi \frac{\partial}{\partial r} - \frac{1}{r} \sin \phi \cos \phi \frac{\partial^2}{\partial \phi \partial r} \right] \]

By convention the Laplacian is defined as

\[ \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \] (16.10)

After some simplifications we obtain the following formula for the Laplacian in spherical coordinates

\[ \nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \left( \frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \csc^2 \theta \frac{\partial^2}{\partial \phi^2} \right). \] (16.11)

In spherical coordinates we have that the **Hamiltonian Operator** (Total Energy Operator) is given by

\[ \hat{H} = -\frac{\hbar^2}{2m} \left[ \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \left( \frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \csc^2 \theta \frac{\partial^2}{\partial \phi^2} \right) \right] + V(r, \theta, \phi). \] (16.12)

Then the Shrödinger Equation in spherical coordinates for a spherically symmetric potential is

\[ -\frac{\hbar^2}{2m} \left[ \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \left( \frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \csc^2 \theta \frac{\partial^2}{\partial \phi^2} \right) \right] \psi(r, \theta, \phi) + V(r) \psi(r, \theta, \phi) = E \psi(r, \theta, \phi). \] (16.13)
We can attempt to find a solution for an arbitrary spherically symmetric potential. As we have done previously we can assume that the solution has the form \( \psi(r, \theta, \phi) = R(r)\Theta(\theta)\Phi(\phi) \). We then obtain

\[
-\frac{\hbar^2}{2m} \left[ \frac{1}{2mR} \left( \frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} \right) + \frac{1}{\Theta} \left( \frac{d^2 \Theta}{d\theta^2} + \cot \theta \frac{d\Theta}{d\theta} \right) + \frac{1}{\Phi} \frac{1}{r^2} \frac{d^2 \Phi}{d\phi^2} \right] + V(r) = E. \quad (16.14)
\]

It is important to remember that the method of separation of variables aims to separate our partial differential equation into several ordinary differential equations. We can then isolate the terms involving \( \phi \) as follows

\[
\frac{1}{\Phi} \frac{d^2 \Phi}{d\phi^2} = -\sin^2 \theta \left[ \frac{r^2}{R} \left( \frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} \right) + \frac{1}{\Theta} \left( \frac{d^2 \Theta}{d\theta^2} + \cot \theta \frac{d\Theta}{d\theta} \right) + \frac{2mr^2}{\hbar^2} (E - V(r)) \right]. \quad (16.15)
\]

Since the left hand side depends only on \( \phi \) and the right hand side depends only on \( r \) and \( \theta \), both sides must be equal to a constant that we can conveniently call \(-m_l^2\). Thus we obtain our first ordinary differential equation

\[
\frac{1}{\Phi} \frac{d^2 \Phi}{d\phi^2} = -m_l^2 \quad (16.16)
\]

which has the solution

\[
\Phi(\phi) = C e^{im_l \phi} \quad (16.17)
\]

for \( m_l = 0, \pm 1, \pm 2, \pm 3 \ldots \) where we have accounted for periodicity.

The second differential equation after some rearrangement is

\[
\frac{r^2}{R} \left( \frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} \right) + \frac{2mr^2}{\hbar^2} (E - V(r)) = -\frac{1}{\Theta} \left( \frac{d^2 \Theta}{d\theta^2} + \cot \theta \frac{d\Theta}{d\theta} \right) + \frac{m_l^2}{\sin^2 \theta}. \quad (16.18)
\]

Once again one side is dependent only on \( r \) and the other side depends only on \( \theta \). Thus both sides must be equal to a constant that we can set arbitrarily to \( l(l + 1) \). Thus, we obtain two differential equations

\[
\frac{d^2 \Theta}{d\theta^2} + \cot \theta \frac{d\Theta}{d\theta} - \frac{m_l^2}{\sin^2 \theta} \Theta = -l(l + 1)\Theta \quad (16.19)
\]

and

\[
-\frac{\hbar^2}{2m} \left[ \frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} \right] + \frac{l(l + 1)\hbar^2}{2mr^2} R(r) + V(r)R = ER(r). \quad (16.20)
\]

The second differential equation depends on the specific potential we are considering.\(^2\) For the moment we will shift our focus to the first differential equation.

---

\(^1\) A potential that only depends on \( r \) is also called central potential.

\(^2\) For instance we could discuss the isotropic oscillator in three dimensions \( V(r) = \frac{1}{2}kr^2 \) or an attractive inverse square potential such as that which is experienced by an electron under the influence of a proton in the hydrogen atom \( V(r) = -\frac{k}{r} \).
Solving the Differential Equation: In order to solve this seemingly nontrivial ordinary differential equation we will first make the change of variables $t = \cos \theta$ and $w(t) = \Theta(\theta)$. Using the chain rule and some trigonometry we have

\[
\begin{align*}
\frac{d\Theta}{d\theta} &= -\sin \theta \frac{dw}{dt} = -\sqrt{1-t^2} \frac{dw}{dt} \\
\frac{d^2\Theta}{d\theta^2} &= \sin^2 \theta \frac{d^2w}{dt^2} - \cos \theta \frac{dw}{dt} = (1-t^2) \frac{d^2w}{dt^2} - t \frac{dw}{dt}.
\end{align*}
\]  

(16.21)

We can then rewrite (16.19) as

\[
(1-t^2) \frac{d^2w}{dt^2} - 2t \frac{dw}{dt} + \left[l(l+1) - \frac{m_l^2}{1-t^2}\right] w = 0.
\]

(16.22)

The easiest way to solve this equation is to first find a solution for the case $m_l = 0$. We have that

\[
(1-t^2) \frac{d^2w}{dt^2} - 2t \frac{dw}{dt} + l(l+1)w = 0.
\]

(16.23)

A standard approach is to consider a power series solution of the form

\[
w(t) = \sum_{k=0}^{\infty} a_k t^k.
\]

(16.24)

Substituting this solution into (16.23) yields

\[
\sum_{k=0}^{\infty} \left[(k+2)(k+1)a_{k+2} - (k^2 - k - l(l+1))a_k\right] = 0.
\]

(16.25)

Thus we obtain the following recurrence relation for the coefficients

\[
a_{k+2} = \frac{k(k+1)-l(l+1)}{(k+2)(k+1)} a_k.
\]

(16.26)

We need to impose the condition\(^3\) $l = 0, 1, 2, 3...$ If we do not assume this condition then we will obtain an infinite series that diverges at $t = \pm 1$ or $\theta = 0, \pi$. Obviously, this does not correspond to any physical system.\(^4\) Thus for each $l$ we obtain a unique polynomial of degree $l$. These are called Legendre Polynomials. Some of the lower order polynomials are given as follows:\(^5\)

\[
\begin{align*}
P_0(t) &= 1 \\
P_1(t) &= t \\
P_2(t) &= (1-3t^2) \\
P_3(t) &= (t-5t^3) \\
P_4(t) &= (1-10t^2 + \frac{35}{3} t^4).
\end{align*}
\]

(16.27)

By convention, $P_l(t)$ is a solution for the corresponding $l$ where $m_l = 0.$

---

\(^3\)That is the reason we conveniently set our separation constant equal to $l(l+1)$, so we can obtain a simple relation.

\(^4\)This can be easily verified using the ratio test.

\(^5\)This might differ from other textbooks by a constant.
Obviously $m_l$ is not always equal to 0 so we proceed to find solutions for $m_l \neq 0$. It is enough to find a solution\textsuperscript{6} for $m_l > 0$. We first differentiate \textsuperscript{16.23} $m_l$ times using the Leibniz Rule and then substituting $P_l$ into the new equation we arrive at

\begin{equation}
(1 - t^2)P_l^{(m_l+2)} - 2(m_l + 1)tP_l^{(m_l+1)} + (l - m_l)(l + m_l + 1)P_l^{(m_l)} = 0.
\end{equation}

Thus we find that the function $P_l^{(m_l)}$ satisfies the equation

\begin{equation}
(1 - t^2)y'' - 2(m_l + 1)ty' + (l - m_l)(l + m_l + 1)y = 0.
\end{equation}

We can now try the function $y(t) = (1 - t^2)^{-m_l/2}w(t)$. We reduce \textsuperscript{16.29} to \textsuperscript{16.22}. Thus we have

\begin{equation}
w(t) = (1 - t^2)^{m_l/2} = (1 - t^2)^{m_l/2} \frac{d^{m_l}P_l}{dt^{m_l}}.
\end{equation}

We then obtain the following convention

\begin{equation}
P_l^{m_l}(t) = (1 - t^2)^{m_l/2} \frac{d^{m_l}P_l}{dt^{m_l}}.
\end{equation}

Notice that since $P_l(t)$ is a polynomial of degree $l$ that for $|m_l| > l$ we have $P_l^{m_l}(t) = 0$. Thus for nontrivial solutions we can impose that $|m_l| \leq l$. Some of these polynomials, called \textbf{Associated Legendre Polynomials}, are shown below\textsuperscript{7}

\begin{align*}
P_0^0(t) &= 1 \\
P_1^0(t) &= t \\
P_1^{\pm 1}(t) &= (1 - t^2)^{1/2} \\
P_2^0(t) &= (1 - 3t^2) \\
P_2^{\pm 1}(t) &= 3t(1 - t^2)^{1/2} \\
P_2^{\pm 2}(t) &= 3(1 - t^2).
\end{align*}

Thus we obtain

\begin{equation}
\Phi(\phi)\Theta(\theta) = Ce^{im_l\theta}P_l^{m_l}(\cos \phi)
\end{equation}

or equivalently

\begin{equation}
\Phi(\phi)\Theta(\theta) = CY_l^{m_l}(\theta, \phi)
\end{equation}

where $Y_l^{m_l}(\theta, \phi) = e^{im_l\theta}P_l^{m_l}(\cos \phi)$ is called a \textbf{Spherical Harmonic}.

Notice that the angular part of the wavefunction has been found and that it is completely independent of the type of central potential. In other words, this is an inherent property of wavefunctions that arises from the symmetry of the problem. The general form of the wavefunction is then given by

\begin{equation}
\psi(r, \theta, \phi) = R(r)Y_l^{m_l}(\theta, \phi).
\end{equation}

\textsuperscript{6}Notice that if we find a solution for a given $l$ and $m_l$ then this is also a solution to the corresponding $l$ and $-m_l$ since the differential equation involves $m_l^2$. This is where we get degeneracy from.

\textsuperscript{7}Note that $P_0^0(t) = P_1(t)$.
16.2 Angular Momentum Operator

We found, interestingly enough, that in two and three dimensions, angular momentum is conserved. It seems natural to discuss the role of a new associated operator: the angular momentum operator. First we will motivate its construction from its classical angular momentum counterpart. In classical mechanics the angular momentum is defined to be the vector obtained by the cross product of the position and momentum

\[ \mathbf{L} = \mathbf{r} \times \mathbf{p} = \dot{x}(yp_z - zp_y) + \dot{y}(zp_x - xp_z) + \dot{z}(xp_y - yp_x). \] (16.36)

This suggests that the angular momentum operator components in quantum mechanics are given by

\[ \hat{L}_x = y\hat{p}_z - z\hat{p}_y, \quad \hat{L}_y = z\hat{p}_x - x\hat{p}_z, \quad \hat{L}_z = x\hat{p}_y - y\hat{p}_x. \] (16.37)

Replacing the \( \hat{p}_i \) with the corresponding coordinate operators we obtain

\[
\begin{align*}
\hat{L}_x &= -i\hbar \left( y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right), \\
\hat{L}_y &= -i\hbar \left( z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right), \\
\hat{L}_z &= -i\hbar \left( x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right).
\end{align*}
\] (16.38)

After substitution and some algebraic manipulation these expressions can be converted to spherical coordinates,

\[
\begin{align*}
\hat{L}_x &= i\hbar \left( \sin \phi \frac{\partial}{\partial \theta} + \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right), \\
\hat{L}_y &= i\hbar \left( -\cos \phi \frac{\partial}{\partial \theta} + \cot \theta \sin \phi \frac{\partial}{\partial \phi} \right), \\
\hat{L}_z &= -i\hbar \frac{\partial}{\partial \phi}.
\end{align*}
\] (16.39)

Now we define the square of the total angular momentum operator to be

\[ \hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2. \] (16.40)

This formula can be converted to spherical coordinates by squaring the formulas given in (16.39). Simplifying we obtain

\[ \hat{L}^2 = -\hbar^2 \left[ \frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right]. \] (16.41)

The operators \( \hat{L}_z \) and \( \hat{L}^2 \) should resemble parts of the differential equations we previously solved. By acting on (16.34) by \( \hat{L}_z \) and \( \hat{L}^2 \) we obtain

\[
\begin{align*}
\hat{L}_z Y_l^m(\theta, \phi) &= \hbar m Y_l^m(\theta, \phi), \\
\hat{L}^2 Y_l^m(\theta, \phi) &= \hbar^2 l(l+1) Y_l^m(\theta, \phi).
\end{align*}
\] (16.42)

These results follow from the differential equations given in (16.16) and (16.20) that \( Y_l^m(\theta, \phi) \) obey. Thus the Spherical Harmonics form an eigenbasis that diagonalizes \( \hat{L}_z \) and \( \hat{L}^2 \) simultaneously. We will soon observe that this result is mathematically consistent due to the properties of the commutator.
16.3 Algebra of Angular Momentum Operator

This section largely deals with the concept of the commutator which measures the degree of commutation of two operators. We will now explore the properties of the commutator a bit further. Given two operators $\Omega$ and $\Lambda$ the commutator is defined as

$$[\Omega, \Lambda] = \Omega \Lambda - \Lambda \Omega.$$ \hspace{1cm} (16.43)

Two operators $\Omega$ and $\Lambda$ are said to commute if

$$[\Omega, \Lambda] = 0.$$ \hspace{1cm} (16.44)

There are several interesting properties of commuting operators.

---

**Theorem 1.** Assume $\Omega$ and $\Lambda$ are operators such that $[\Omega, \Lambda] = 0$. If $\psi$ is an eigenfunction of $\Omega$ with eigenvalue $\omega$ then $\Lambda \psi$ is an eigenfunction of $\Omega$ with the same eigenvalue $\omega$.

**Proof.**

$$\Omega \psi = \omega \psi$$

$$\Lambda (\Omega \psi) = \Lambda (\omega \psi)$$

$$\Omega (\Lambda \psi) = \omega (\Lambda \psi).$$ \hspace{1cm} (16.45)

The last step used the fact that $[\Omega, \Lambda] = 0$.

---

**Theorem 2.** Assume $\Omega$ and $\Lambda$ are two Hermitian operators such that $[\Omega, \Lambda] = 0$. Then there exists a common orthogonal eigenbasis that diagonalizes $\Omega$ and $\Lambda$ simultaneously.

**Proof.** We will prove this result for at least one non-degenerate operator, $\Omega$. Since $\Omega$ is Hermitian there is an orthogonal eigenbasis that diagonalizes $\Omega$. Let $\psi$ be an arbitrary element of this basis. From the previous theorem we have that if $\psi$ is an eigenfunction of $\Omega$ with eigenvalue $\omega$ then

$$\Omega (\Lambda \psi) = \omega (\Lambda \psi).$$ \hspace{1cm} (16.46)

However, since $\Omega$ is non-degenerate then $\Lambda \psi$ is a multiple of $\psi$ i.e.

$$\Lambda \psi = \lambda \psi.$$ \hspace{1cm} (16.47)

Thus $\psi$ is also an eigenvector of $\Lambda$. Thus we have found a common eigenbasis that diagonalizes $\Omega$ and $\Lambda$ simultaneously.

---

These results have important physical consequences. In particular, the first one tells us that if an operator commutes with the Hamiltonian then given an eigenfunction of the Hamiltonian, the
second operator gives us a new eigenfunction of the Hamiltonian. This turns out to be the origin of energy degeneracy!

The angular momentum operator has very interesting commutation properties. We will discuss some of them by finding the commutation relations for $\hat{L}_x$, $\hat{L}_y$ and $\hat{L}_z$. First, we have to compute $\hat{L}_x\hat{L}_y$ and $\hat{L}_y\hat{L}_x$ as follows

$$\hat{L}_x\hat{L}_y = -\hbar^2 \left( y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \left( z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right). \quad (16.48)$$

Proceeding with some simple algebraic manipulations we obtain

$$\hat{L}_x\hat{L}_y = -\hbar^2 \left( zx \frac{\partial^2}{\partial y \partial z} - z^2 \frac{\partial^2}{\partial y \partial z} - yx \frac{\partial^2}{\partial z^2} + y \frac{\partial}{\partial x} + yz \frac{\partial^2}{\partial z^2} \right). \quad (16.49)$$

Similarly,

$$\hat{L}_y\hat{L}_x = -\hbar^2 \left( x \frac{\partial}{\partial y} - xz \frac{\partial^2}{\partial z \partial y} + x^2 \frac{\partial^2}{\partial z^2} + y \frac{\partial}{\partial x} + xy \frac{\partial^2}{\partial z \partial x} \right). \quad (16.50)$$

Then, substituting these relations into the commutator of $\hat{L}_x$ and $\hat{L}_y$ we obtain

$$[\hat{L}_x, \hat{L}_y] = -\hbar^2 \left( y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} \right) = i\hbar\hat{L}_z. \quad (16.51)$$

Similar arguments can be derived for the other components and we obtain the following results

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z, \quad [\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x, \quad [\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y. \quad (16.52)$$

These relations are extremely important since they summarize the algebraic structure of the operators. We can now discuss the commutator $[\hat{L}_z, \hat{L}^2]$. Recall that the Spherical Harmonics form an eigenbasis of both operators. This might indicate that both operators commute. We proceed to prove this fact in a rather careful manner.

---

**An Important Commutator Relation:** Trivially we have

$$[\hat{L}_z, \hat{L}^2] = [\hat{L}_z, \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2] = [\hat{L}_z, \hat{L}_x^2] + [\hat{L}_z, \hat{L}_y^2] + [\hat{L}_z, \hat{L}_z^2]. \quad (16.53)$$

The last term vanishes. Using the properties of commutators we have

$$[\hat{L}_z, \hat{L}^2] = \hat{L}_x [\hat{L}_z, \hat{L}_x] + [\hat{L}_z, \hat{L}_x] \hat{L}_x + \hat{L}_y [\hat{L}_z, \hat{L}_y] + [\hat{L}_z, \hat{L}_y] \hat{L}_y. \quad (16.54)$$

Now using the commutation relations derived in (16.52) we can simplify this result to

$$[\hat{L}_z, \hat{L}^2] = i\hbar \hat{L}_x \hat{L}_y + i\hbar \hat{L}_y \hat{L}_x - i\hbar \hat{L}_y \hat{L}_x - i\hbar \hat{L}_x \hat{L}_y = 0. \quad (16.55)$$

Likewise, we obtain

$$[\hat{L}_x, \hat{L}^2] = [\hat{L}_y, \hat{L}^2] = [\hat{L}_z, \hat{L}^2] = 0. \quad (16.56)$$
As a consequence of Theorem 2, it is possible to find an eigenbasis that diagonalizes \( \hat{L}^2 \) and any other angular momentum operator component. However, it is not possible to find an eigenbasis that diagonalizes more than two of the angular momentum operator components.

Finally, we can obtain the commutation relations of the Hamiltonian for a spherically symmetric potential. Rewriting this combining (16.12) and (16.41) we obtain

\[
\hat{H} = -\frac{\hbar^2}{2m} \left[ \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} - \frac{\hat{L}^2}{\hbar^2 r^2} \right] + V(r). \tag{16.57}
\]

Notice that this equation depends only on \( r \) and \( \hat{L}^2 \). Since \( \hat{L}^2 \) does not depend on \( r \) it follows that

\[ [\hat{H}, \hat{L}^2] = 0. \tag{16.58} \]

The same applies to \( \hat{L}_z \) since it commutes with \( \hat{L}^2 \). Thus,

\[ [\hat{H}, \hat{L}_z] = 0. \tag{16.59} \]

The last two equations along with (16.55) show that there is an eigenbasis that diagonalizes \( \hat{H}, \hat{L}^2 \) and \( \hat{L}_z \) simultaneously. Notice that we have obtained this result from the algebra of the operators and we did not need to solve the differential equations! This demonstrates the usefulness and the importance of understanding the algebraic structure of operators.

### 16.4 Generalization of the Angular Momentum Operator

Previously we observed the importance of the algebra of operators. We will now investigate a particular type of abstract Hermitian operator that obeys certain commutation relations given in (16.52). We choose these relations because they seem to summarize all the features of the angular momentum operator components. Let \( \hat{J}_x, \hat{J}_y \) and \( \hat{J}_z \) be Hermitian operators that obey the following commutation relations,

\[
\begin{align*}
[\hat{J}_x, \hat{J}_y] &= i\hbar \hat{J}_z, \\
[\hat{J}_y, \hat{J}_z] &= i\hbar \hat{J}_x, \\
[\hat{J}_z, \hat{J}_x] &= i\hbar \hat{J}_y. 
\end{align*} \tag{16.60}
\]

We have the following definition

\[ \hat{j}^2 = \hat{j}_x^2 + \hat{j}_y^2 + \hat{j}_z^2 \tag{16.61} \]

and it follows that

\[ [\hat{j}_z, \hat{j}^2] = 0. \tag{16.62} \]

By Theorem 2 we can conclude that they have a common eigenbasis. At this point we choose to introduce a very useful notation accredited to Dirac. We will denote the eigenvector of \( \hat{j}^2 \) and \( \hat{J}_z \) with eigenvalues \( j \) and \( m_j \) respectively as \( |j, m_j\rangle \) i.e.

\[
\begin{align*}
\hat{j}^2 |j, m_j\rangle &= j |j, m_j\rangle \\
\hat{J}_z |j, m_j\rangle &= m_j |j, m_j\rangle. 
\end{align*} \tag{16.63}
\]
We now introduce two new operators called raising and lower operators

\[
\hat{J}_+ = \hat{J}_x + i\hat{J}_y, \\
\hat{J}_- = \hat{J}_x - i\hat{J}_y.
\] (16.64)

Notice that these are self-adjoint to each other. Let us examine some commutation relations of these new operators.

**Relevant Commutation Relations:** We start with the following

\[
\left[ \hat{J}_z, \hat{J}_\pm \right] = \hat{J}_z \hat{J}_\pm - \hat{J}_\pm \hat{J}_z = \hat{J}_z \hat{J}_x \pm i\hat{J}_z \hat{J}_y - \left( \hat{J}_x \hat{J}_z \pm i\hat{J}_y \hat{J}_z \right).
\] (16.65)

Rearranging the terms we obtain

\[
\left[ \hat{J}_z, \hat{J}_\pm \right] = \hat{J}_z \hat{J}_x - \hat{J}_x \hat{J}_z \pm i \left( \hat{J}_z \hat{J}_y - \hat{J}_y \hat{J}_z \right) = \left[ \hat{J}_z, \hat{J}_x \right] \pm i \left[ \hat{J}_z, \hat{J}_y \right].
\] (16.66)

Using the commutation relations we have

\[
\left[ \hat{J}_z, \hat{J}_\pm \right] = \pm \mathbb{h} \left( \hat{J}_z \hat{J}_y \pm \mathbb{h} \hat{J}_x \pm i \left( \hat{J}_z \hat{J}_y - \hat{J}_y \hat{J}_z \right) \right).
\] (16.67)

Finally we obtain

\[
\left[ \hat{J}_z, \hat{J}_\pm \right] = \pm \mathbb{h} \hat{J}_\pm.
\] (16.68)

On the other hand,

\[
\left[ \hat{J}^2, \hat{J}_\pm \right] = \left[ \hat{J}^2, \hat{J}_x \right] \pm i \left[ \hat{J}^2, \hat{J}_y \right].
\] (16.69)

Thus,

\[
\left[ \hat{J}^2, \hat{J}_\pm \right] = 0.
\] (16.70)

We can now prove the relations (16.67) and (16.70) to see how powerful they are.

\[
\hat{J}^2 \hat{J}_\pm |j, m_j\rangle = \hat{J}_\pm \hat{J}^2 |j, m_j\rangle = j \hat{J}_\pm |j, m_j\rangle
\] (16.71)

\[
\hat{J}_z \left( \hat{J}_\pm |j, m_j\rangle \right) = \left( \hat{J}_\pm \hat{J}_z \pm \mathbb{h} \hat{J}_\pm \right) |j, m_j\rangle = (m_j \pm \mathbb{h}) \hat{J}_\pm |j, m_j\rangle.
\] (16.72)

This shows that the effect of \( \hat{J}_\pm \) is to raise or lower the eigenvalues of the eigenvectors of \( \hat{J}_z \) by a factor of \( \mathbb{h} \) and to keep the eigenvalues of \( \hat{J}^2 \) invariant. You can think of the vector space as split into blocks of eigenspaces according to their eigenvalues corresponding to the operators \( \hat{J}^2 \) and \( \hat{J}_z \). Thus the effect of \( \hat{J}_\pm \) is to move up and down from different eigenspaces of \( \hat{J}_z \) while remaining in the same eigenspace of \( \hat{J}^2 \). This might hint towards the fact that the eigenspaces of \( \hat{J}_z \) are "smaller" and perhaps inside those of \( \hat{J}^2 \). Mathematically, we have the following result

\[
\hat{J}_\pm |j, m_j\rangle = C_{\pm(j,m_j)} |j, m_j \pm \mathbb{h}\rangle
\] (16.73)

where \( C_{\pm(j,m_j)} \) is a scalar.
This clearly indicates that the dimension of the eigenbasis of $\hat{J}_z$ is infinite for any value of $j$ since we can arbitrarily apply $\hat{J}_\pm$ to $|j, m_j\rangle$. However, since all these operators are Hermitian they all have real eigenvalues. Furthermore, the square of these operators have non-negative eigenvalues. In particular,
\[
\left(\hat{J}_x^2 + \hat{J}_y^2\right) |j, m_j\rangle = \left(\hat{J}_-^2 - \hat{J}_z^2\right) |j, m_j\rangle = (j - m_j^2) |j, m_j\rangle.
\] (16.74)

Thus
\[
j - m_j^2 \geq 0.
\] (16.75)

This means that the values of $m_j$ are bounded given $j$. The only way this can happen is that the eigenvalues are bounded above and below i.e. there exist $m_{\text{max}}$ and $m_{\text{min}}$ such that
\[
\hat{J}_+ |j, m_{\text{max}}\rangle = 0 \\
\hat{J}_- |j, m_{\text{min}}\rangle = 0.
\] (16.76)

Now we will consider a few relations of the raising and lowering operators
\[
\hat{J}_- \hat{J}_+ = \left(\hat{J}_x - i\hat{J}_y\right) \left(\hat{J}_x + i\hat{J}_y\right) = \hat{J}_x^2 + \hat{J}_y^2 + i \left[\hat{J}_x, \hat{J}_y\right] = \hat{J}_2 - \hat{J}_z^2 - \hbar \hat{J}_z
\] (16.77)

Applying $\hat{J}_-$ to the first relation in (16.76)
\[
\left(\hat{J}_2 - \hat{J}_z^2 - \hbar \hat{J}_z\right) |j, m_{\text{max}}\rangle = (j - m_{\text{max}}^2 - \hbar m_{\text{max}}) |j, m_{\text{max}}\rangle = 0.
\] (16.79)

Hence,
\[
j = m_{\text{max}}(m_{\text{max}} + \hbar).
\] (16.80)

Applying $\hat{J}_+$ to the second relation in (16.76) we obtain
\[
\left(\hat{J}_2 - \hat{J}_z^2 + \hbar \hat{J}_z\right) |j, m_{\text{min}}\rangle = (j - m_{\text{min}}^2 + \hbar m_{\text{min}}) |j, m_{\text{min}}\rangle = 0
\] (16.81)

\[
j = m_{\text{min}}(m_{\text{min}} - \hbar).
\] (16.82)

Combining (16.80) and (16.82) we obtain
\[
m_{\text{max}} = -m_{\text{min}}.
\] (16.83)

This symmetrical result is expected from the nature of the operators.

Furthermore, we can obtain $m_{\text{min}}$ from $m_{\text{max}}$ by a finite application of $\hat{J}_-$. Thus,
\[
m_{\text{max}} - m_{\text{min}} = \hbar s
\] (16.84)

where $s$ is a non-negative integer. Using (16.83) and (16.84) we get
\[
m_{\text{max}} - m_{\text{min}} = 2m_{\text{max}} = \hbar s
\] (16.85)

\footnote{This is the solution to a quadratic equation whose other solution is $m_{\text{max}} = m_{\text{min}} - \hbar$. However, this is not possible since $m_{\text{max}} \geq m_{\text{min}}$.}
then
\[ m_{\text{max}} = \frac{hs}{2} \quad s = 0, 1, 2, \ldots \]  
(16.86)

and
\[ j = h^2 \left( \frac{s}{2} \right) \left( \frac{s}{2} + 1 \right). \]  
(16.87)

We have successfully divided the vector space into blocks of eigenspaces of \( \hat{J}^2 \) and these blocks have been observed to split into cells of eigenspaces of \( \hat{J}_z \).

### 16.5 Representation of the Algebra for \( s = 1 \)

As an example we will look at the representation of the algebra for a specific case: \( s = 1 \). In this case we have that

\[ j = \frac{3}{4} h^2, \quad m_j = \frac{1}{2} h, -\frac{1}{2} h. \]  
(16.88)

The operators \( \hat{J}^2 \) and \( \hat{J}^z \) have a simple representation since they are diagonal, the entries are simply the eigenvalues

\[ \hat{J}^2 = \frac{3}{4} h^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \]  
(16.89)

\[ \hat{J}^z = \frac{1}{2} h \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \]  
(16.90)

We need to do some work to find the representation of \( \hat{J}_x \) and \( \hat{J}_y \). First let us find the representation of the raising and lower operators.

---

**Finding the Representation of the Raising and Lowering Operators:** We need to determine how \( \hat{J}_+ \) and \( \hat{J}_- \) act on \( | \frac{3}{4} h^2, \frac{1}{2} h \rangle \) and \( | \frac{3}{4} h^2, -\frac{1}{2} h \rangle \). We have already discussed that

\[ \hat{J}_+ \begin{bmatrix} \frac{3}{4} h^2, \frac{1}{2} h \end{bmatrix} = 0 \]  
(16.91)

and

\[ \hat{J}_- \begin{bmatrix} \frac{3}{4} h^2, \frac{1}{2} h \end{bmatrix} = 0 \]  
(16.92)

because they have the maximum and minimum eigenvalues respectively. What about \( \hat{J}_+ \begin{bmatrix} \frac{3}{4} h^2, -\frac{1}{2} h \end{bmatrix} \) and \( \hat{J}_- \begin{bmatrix} \frac{3}{4} h^2, \frac{1}{2} h \end{bmatrix} \)? We know from (16.73) that

\[ \hat{J}_+ \begin{bmatrix} \frac{3}{4} h^2, -\frac{1}{2} h \end{bmatrix} = C_+ \left( \frac{3}{4} h^2, -\frac{1}{2} h \right) \begin{bmatrix} \frac{3}{4} h^2, \frac{1}{2} h \end{bmatrix} \]  
(16.93)

and

\[ \hat{J}_- \begin{bmatrix} \frac{3}{4} h^2, \frac{1}{2} h \end{bmatrix} = C_- \left( \frac{3}{4} h^2, \frac{1}{2} h \right) \begin{bmatrix} \frac{3}{4} h^2, -\frac{1}{2} h \end{bmatrix}. \]  
(16.94)
Multiplying each equation by its adjoint we obtain
\[ C^2_{+\left(\frac{3}{4}\hbar, -\frac{1}{2}\hbar\right)} = \left\langle \frac{3}{4}\hbar, -\frac{1}{2}\hbar \right| \hat{J} - \hat{J} \left| \frac{3}{4}\hbar, -\frac{1}{2}\hbar \right\rangle \] (16.95)
\[ C^2_{-\left(\frac{3}{4}\hbar, \frac{1}{2}\hbar\right)} = \left\langle \frac{3}{4}\hbar, \frac{1}{2}\hbar \right| \hat{J} - \hat{J} \left| \frac{3}{4}\hbar, \frac{1}{2}\hbar \right\rangle. \] (16.96)

Using (16.77) and (16.78) we obtain
\[ C^2_{+\left(\frac{3}{4}\hbar, -\frac{1}{2}\hbar\right)} = \left\langle \frac{3}{4}\hbar, -\frac{1}{2}\hbar \right| \hat{J}^2 - \frac{3}{4}\hbar^2 - \frac{1}{2}\hbar \left| \hat{J}^2 - \frac{3}{4}\hbar^2, -\frac{1}{2}\hbar \right\rangle = \left( \frac{3}{4} - \frac{1}{4} + \frac{1}{2} \right) \hbar^2 = \hbar^2 \] (16.97)
and
\[ C^2_{-\left(\frac{3}{4}\hbar, \frac{1}{2}\hbar\right)} = \left\langle \frac{3}{4}\hbar, \frac{1}{2}\hbar \right| \hat{J}^2 - \frac{3}{4}\hbar^2 + \frac{1}{2}\hbar \left| \hat{J}^2 - \frac{3}{4}\hbar^2, \frac{1}{2}\hbar \right\rangle = \left( \frac{3}{4} - \frac{1}{4} + \frac{1}{2} \right) \hbar^2 = \hbar^2. \] (16.98)

Thus we have that
\[ C^2_{+\left(\frac{3}{4}\hbar, -\frac{1}{2}\hbar\right)} = C^2_{-\left(\frac{3}{4}\hbar, \frac{1}{2}\hbar\right)} = \hbar. \] (16.99)

Substituting this into (16.93) and (16.94) we get
\[ \hat{J}_+ \left| \frac{3}{4}\hbar^2, -\frac{1}{2}\hbar \right\rangle = \left| \frac{3}{4}\hbar^2, \frac{1}{2}\hbar \right\rangle \] (16.100)
and
\[ \hat{J}_- \left| \frac{3}{4}\hbar^2, \frac{1}{2}\hbar \right\rangle = \left| \frac{3}{4}\hbar^2, -\frac{1}{2}\hbar \right\rangle. \] (16.101)

These two equations along with (16.91) and (16.92) complete our understanding of how the eigenvectors transform under the raising and lower operators. Thus we have
\[ \hat{J}_+ = h \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \] (16.102)
\[ \hat{J}_- = h \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \] (16.103)

Finally we can compute \( \hat{J}_x \) and \( \hat{J}_y \) as follows
\[ \hat{J}_x = \frac{1}{2} \left( \hat{J}_+ + \hat{J}_- \right) = h \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \] (16.104)
\[ \hat{J}_y = \frac{1}{2i} \left( \hat{J}_+ - \hat{J}_- \right) = h \frac{1}{2} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}. \] (16.105)

The final result is
\[ \hat{J}_x = \frac{\hbar}{2} \sigma_x, \quad \hat{J}_y = \frac{\hbar}{2} \sigma_y, \quad \hat{J}_z = \frac{\hbar}{2} \sigma_z, \quad \hat{J}^2 = \frac{3}{4} \hbar^2 I_2 \] (16.106)
where
\[ \sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \] (16.107)

The matrices \( \sigma_x, \sigma_y \) and \( \sigma_z \) are known as Pauli Matrices. These are extremely useful in the theory of Spin and the Dirac equation which describes relativistic quantum mechanical phenomena.
Chapter 17

Statistical Mechanics

In this section we will briefly discuss the quantum mechanics of several particles. We will begin by discussing the Schrödinger equation for a system of particles. Evaluating the symmetry in such a system will help us reach the conclusion that there are two types of fundamental particles: fermions and bosons, that differ by the type of symmetry they obey. Then, we will be set up to discuss the Statistical Mechanics of these particles as well as some interesting consequences.

17.1 Particle Exchange

Recall that in classical mechanics if we have two identical interacting particles that do not experience an external potential then their energy is conserved and their motion can be described by their total energy

\[ E = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} + V(r_1 - r_2). \] (17.1)

If we perform some simple mathematical manipulations we obtain

\[ E = \frac{(p_1 + p_2)^2}{4m} + \frac{(p_1 - p_2)^2}{4m} + V(r_1 - r_2). \] (17.2)

Notice that the first term represents the center of mass energy of both particles while the other two terms represent the energy associated with the relative position of the particle. If we let

\[ R = \frac{1}{2}(r_1 + r_2), \quad r = r_1 - r_2, \quad M = 2m \text{ and } \mu = \frac{1}{2}m \]

then\(^1\) we have that the associated momenta are \( P = M\dot{R} \) and \( p = \mu\dot{r} \) and the previous equation can be rewritten as

\[ \frac{P^2}{2M} = E_{cm} \]

\[ \frac{p^2}{2\mu} + V(r) = E_{rel} \] (17.3)

where \( E = E_{cm} + E_{rel} \). This way we have changed the two particle problem into a single particle problem since the only interesting equation is the one involving the relative coordinates.

\(^1\)You should immediately realize that these are the center of mass coordinates and the relative coordinates.
We can take a similar approach when describing the quantum mechanics of the system and express the wavefunction of two identical interacting particles as a product of a wavefunction dependent on the center of mass coordinates and one dependent on the relative coordinates. This will be possible only if the potential depends on the relative positions of the particles, such as an electrostatic or elastic potential. Let us briefly discuss how this is done.

Two-Body problem to One-Body problem: We will need to write the Schrödinger Equation for a system containing two identical particles is given by

\[
-\frac{\hbar^2}{2m} \left[ \nabla^2_{r_1} + \nabla^2_{r_2} + V(r_1 - r_2) \right] \psi(r_1, r_2) = E\psi(r_1, r_2).
\] (17.4)

Notice that the first two terms represent the total kinetic energy. (The subscripts \( r_1 \) and \( r_2 \) indicate the coordinates used in each case.) Using the chain rule we have that

\[
\frac{\partial}{\partial x_1} = \frac{\partial}{\partial x} + \frac{1}{2} \frac{\partial}{\partial X}, \quad \frac{\partial}{\partial x_2} = \frac{\partial}{\partial x} - \frac{1}{2} \frac{\partial}{\partial X}.
\] (17.5)

Thus,

\[
\frac{\partial^2}{\partial x_1^2} = \frac{\partial^2}{\partial x^2} + \frac{1}{2} \frac{\partial^2}{\partial X^2} + \frac{1}{4} \frac{\partial^2}{\partial X^2}
\]
\[
\frac{\partial^2}{\partial x_2^2} = \frac{\partial^2}{\partial x^2} - \frac{1}{2} \frac{\partial^2}{\partial X^2} + \frac{1}{4} \frac{\partial^2}{\partial X^2}.
\] (17.6)

\[
\Psi(R, r) = \psi_R(R)\psi_r(r).
\] (17.7)

Similar results hold for \( y \) and \( z \). Then we obtain the following result

\[
\nabla^2_{r_1} + \nabla^2_{r_2} = 2\nabla^2_r + \frac{1}{2} \nabla^2_R.
\] (17.8)

Finally, substituting this relation into (17.4) and changing the variables of \( \psi \)

\[
\left[ -\frac{\hbar^2}{2\mu} \nabla^2_r - \frac{\hbar^2}{2M} \nabla^2_R + V(r) \right] \psi(R, r) = E\psi(R, r).
\] (17.9)

This is a separable equation in the variables \( R \) and \( r \). If we try the solution \( \psi(R, r) = \psi_R(R)\psi_r(r) \) we obtain

\[
E_{cm} = -\frac{\hbar^2}{2M} \nabla^2_R \psi_R(R)
\]
\[
E_{rel} = -\frac{\hbar^2}{2\mu} \nabla^2_r \psi_r(r) + V(r)\psi_r(r).
\] (17.10)
If we look at $E_{cm}$ and $E_{rel}$ we realize that $E_{cm}$ is only dependent on $R$ and that $E_{rel}$ is only dependent on $r$. Thus, we have verified that a two body problem in quantum mechanics can also be reduced to a one body problem by choosing the relative and center of mass coordinates. Luckily it turns out that usually we are interested in the relative coordinate problem.

If we happen to find a solution $\psi(r_1 - r_2)$, we might ask the question: **Is the wavefunction $\psi(r_2 - r_1)$ a physical solution?** At first glance it should be since both particles are indistinguishable hence exchanging them should result in the same wavefunction. However, what matters here is not the wavefunction but the probability distribution. Hence, we should have under particle exchange that the wavefunction acquires a constant phase, that is

$$\psi(r_2 - r_1) = e^{i\phi} \psi(r_1 - r_2) \quad (17.11)$$

Notice that $|\psi(r_1 - r_2)|^2 = |\psi(r_2 - r_1)|^2$ meaning their probability distributions wouldn’t be altered since the phase integral won’t make a difference.

We would now like to consider free particles. For free particles, since there is no potential, we know that $\psi(r_1, r_2) = \psi_i(r_1)\psi_j(r_2)$. In general we can have a wavefunction of the form

$$\alpha\psi_i(r_1)\psi_j(r_2) + \beta\psi_i(r_2)\psi_j(r_1) \quad (17.12)$$

where $\alpha, \beta$ are arbitrary constants. If we now switch $r_1$ and $r_2$ we obtain

$$\alpha\psi_i(r_2)\psi_j(r_1) + \beta\psi_i(r_1)\psi_j(r_2).$$

These two functions should only differ by a phase $\phi$ so we obtain the following relations

$$\beta = \alpha e^{i\phi}, \quad \alpha = \beta e^{i\phi}. \quad (17.13)$$

This means that

$$e^{2i\phi} = 1$$

so the only possibilities for a system of noninteracting particles are $\phi = 0, \pi$.

- If $\phi = 0$, we will label the particles **bosons** (they have now been experimentally observed in nature at the fundamental level.)
- If $\phi = \pi$, we will label the particles **fermions** (they as well are experimentally observed in nature at the fundamental level.)
- What if $\phi$ does not equal $0$ or $\pi$? We will call these particles **anyons**. Anyons cannot be fundamental particles because we see from the phase that they cannot exist without interactions.

We are capable of experimentally verifying the existence of bosons and fermions. This is what we call **statistics**. Why is it called statistics? This name depends on how you find out whether a
particle is one or the other.

Since anyons cannot exist as free particles, we will restrict our discussion to fermions and bosons.

We can consider our generalized wavefunction as in (17.12) for both fermions and bosons respectively

\[ \psi_i(r_1) \psi_j(r_2) \pm \psi_i(r_2) \psi_j(r_1). \]

The + sign corresponds to fermions and the − sign corresponds to bosons so we have the following if we switch \( r_1 \) and \( r_2 \)

\[ \psi(r_2, r_1) = \psi_i(r_2) \psi_j(r_1) + \psi_i(r_1) \psi_j(r_2) = \psi(r_1, r_2) \]

\[ \psi(r_2, r_1) = \psi_i(r_2) \psi_j(r_1) - \psi_i(r_1) \psi_j(r_2) = -\psi(r_1, r_2). \] (17.14)

The last equation has a deep consequence. Fermions are not allowed to be in the same state; otherwise, the wavefunction vanishes. This is more familiarly known as the Pauli Exclusion Principle.

17.2 Statistical Mechanics: Partition Function

We will not solely concentrate on understanding the development of statistical mechanics of a system. Consider a system \( T \) with a fixed energy \( U_T \). Having a fixed energy is characteristic of what we call an “isolated” system. Now suppose that we split up this system into two systems \( X \) and \( R \) where

\[ U_R >> U_X, \quad V_R >> V_X \quad \text{and} \quad N_R >> N_X. \]

We use the common conventions where \( U_k \) is the energy of system \( k \), \( V_k \) is the volume of system \( k \) and \( N_k \) is the number of particles in system \( k \).

We call \( R \) the reservoir for the smaller system \( X \). The term used to label the smaller system is the grand canonical system. We could consider the entire system \( T \) which we label as the microcanonical system. We could also tailor our systems such that \( X \) and \( R \) could only exchange energy and not particles. Then we would have the canonical system.

Returning to our discussion of the system in question, by construction we know

\[ U_R + E_X = U_T, \quad V_R + V_X = V_T, \quad N_R + N_X = N_T. \]

We can now introduce one of the elementary postulates of statistical mechanics.

The probability \( P \) of system \( X \) having energy \( U_X \), volume \( V_X \) and number of particles \( N_X \) is proportional to the multiplicities \( \Omega \) of the macrostate of the larger system having energy \( U_0 - U_X = U_R \), volume \( V_0 - V_X = V_R \) and number of particles \( N_0 - N_X = N_R \).

Therefore we have

\[ P(U_X, V_X, N_X) \propto \Omega(U_R, V_R, N_R). \]

We can now define the quantity \( S_R = k_B \ln(\Omega(U_R, V_R, N_R)) \) to be the entropy. Then we can restate the above postulate as

170
The system is more likely to be in the state with the most entropy.

Now assume that the system is in equilibrium so \( dS_T = dS_X + dS_R \). We define the temperature \( T \) as \( \frac{\partial S_R}{\partial U_R} = \frac{1}{T} \), the chemical potential \( \mu \) as \( \frac{\partial S_R}{\partial N_R} = \frac{\beta}{T} \), and the pressure \( P \) as \( \frac{\partial S_R}{\partial V_R} = \frac{P}{T} \). It can be verified that these definitions of temperature and pressure correspond to the operational definitions given by the ideal gas law.\(^2\) Therefore we have that

\[
dS_R = \frac{\partial S_R}{\partial U_R} dU_R + \frac{\partial S_R}{\partial V_R} dV_R + \frac{\partial S_R}{\partial N_R} dN_R = \frac{1}{T} (dU_R + PdV_R - \mu dN_R). \tag{17.15}
\]

It is important to introduce the concept of a **microstate**. A microstate is a given configuration of a system of particles. Any system can be in any number of microstates and the total number of microstates is what we referred to as the multiplicity. Now, let \( X_i \) be the microstate where small system \( X \) has energy \( U(X_i) \), volume \( V(X_i) \), and number of particles \( N(X_i) \). If we assume that the volume does not change, then we have that

\[
P(X_2) = \frac{\Omega_R(X_2)}{\Omega_R(X_1)} = \frac{e^{S_R(X_2)/k_B}}{e^{S_R(X_1)/k_B}} = e^{\frac{s_R(X_2) - s_R(X_1)}{k_B}} = e^{\frac{dS_R}{k_B}} = e^{\beta(dU_R - \mu dN_R)}
\]

where \( \beta = \frac{1}{k_BT} \). How do we know that \( S_R(X_2) - S_R(X_1) = dS_R \)? We are assuming that any change of states in the system \( X \) will lead to a very small change in the system \( R \) and therefore an even smaller change in \( S_R \) (because \( S_R \) is the log of \( \Omega \).) Therefore we can approximate this difference with a differential. We also know that since \( U_T, V_T \) and \( N_T \) are a constant that

\[
dU_R = -dU_X, \quad dV_R = -dV_X, \quad dN_R = -dN_X.
\]

Thus we have that

\[
P(X_2) = e^{-\beta(U_X - \mu N_X)} = e^{-\frac{\beta(U(X_2) - \mu N(X_2))}{k_B}}.
\]

So

\[
P(X_1) \propto e^{-\beta(U(X_1) - \mu N(X_1))}
\]

and finally \( P(X_1) = \frac{1}{Z} e^{-\beta(U(X_1) - \mu N(X_1))} \), where \( \frac{1}{Z} \) is the constant of proportionality which we shall denote \( Z \) as the **partition function**. We can find the value of \( Z \) by summing over all the possible \( P(X_i) \) and setting that equal to 1. We obtain that

\[
\Sigma_i P(X_i) = 1 \implies \frac{1}{Z} = \Sigma_i B(X_i) \implies Z = \Sigma_i P(X_i)
\]

(where \( B(X_i) \) is the Boltzmann factor \( e^{-\beta(U_i - \mu N_i)} \).) Therefore we have

\[
Z(\beta, \beta \mu) = \sum_{n_0} \sum_{n_1} ... \sum_{n_k} ... e^{-\beta \sum_{k} n_k (\epsilon_k - \mu)} \tag{17.16}
\]

\(^2\)The chemical potential is the energy needed to keep the entropy fixed while putting more particles in the system. It would seem that entropy increases when you put in particles (because more particles should mean more microstates) and the chemical potential should therefore always be negative, but as we shall see this does not always hold, can you see why? (Hint: The Fermi Exclusion Principle)
where \( n_i \) is equal to the number of particles with energy \( \epsilon_i \) and \( \Sigma_{n_i} (\square) \) is sum of \( \square \) over all the possible values of \( n_i \). Now we can simplify the expression for \( Z \). Using induction on \( e^{a+b} = e^a e^b \) we find that
\[
e^{-\beta \sum_k n_k (\epsilon_k - \mu)} = \prod_k e^{-\beta n_k (\epsilon_k - \mu)} \tag{17.17}
\]
and using induction once again on \( \sum_i \sum_j f(i) f(j) = \sum_i f(i) \sum_j f(j) \) we find that
\[
Z(\beta, \beta \mu) = \prod_k \sum_n e^{-\beta n_k (\epsilon_k - \mu)}. \tag{17.18}
\]

### 17.3 Expectation Values

The partition function is more than just a normalization factor. In fact, it contains a great deal of information that will help in our following analysis.

We will proceed to calculate some relevant expectation values.

**Expectation Values-the importance of the partition function:** Recall that to compute the expectation value of a discrete variable we use
\[
\langle A \rangle = \sum_i A(X_i) P(X_i). \tag{17.19}
\]
We wish to find the expectation values \( \langle N \rangle \) and \( \langle E \rangle \). The total number of particles and the total energy for a given microstate is
\[
N = \sum_0^\infty n_k \tag{17.20}
\]
\[
E = \sum_0^\infty n_k \epsilon_k. \tag{17.21}
\]
Thus,
\[
\langle N \rangle = \sum_i P(X_i) \left( \sum_0^\infty n_k \right) = \frac{1}{Z} \sum_0^\infty n_k \ldots \sum_0^\infty n_k \ldots \sum_0^\infty n_k \ldots \ldots \ldots \ldots e^{-\beta \sum_k n_k (\epsilon_k - \mu)} \tag{17.22}
\]
but we recognize that
\[
\sum_0^\infty n_k \ldots \sum_0^\infty n_k \ldots \sum_0^\infty n_k \ldots \ldots \ldots \ldots e^{-\beta \sum_k n_k (\epsilon_k - \mu)} = \frac{\partial}{\partial \beta \mu} \sum_0^\infty n_k \ldots \sum_0^\infty n_k \ldots \sum_0^\infty n_k \ldots \ldots \ldots \ldots e^{-\beta \sum_k n_k (\epsilon_k - \mu)} = \frac{\partial Z}{\partial \beta \mu}. \tag{17.23}
\]
Similarly for \( \langle E \rangle \) we have that
\[
\langle E \rangle = \sum_i P(X_i) \left( \sum_0^\infty n_k \epsilon_k \right) = \frac{1}{Z} \sum_0^\infty n_k \ldots \sum_0^\infty n_k \ldots \sum_0^\infty n_k \ldots \ldots \ldots \ldots e^{-\beta \sum_k n_k (\epsilon_k - \mu)} \tag{17.24}
\]
and that
\[
\sum_0^\infty n_k \ldots \sum_0^\infty n_k \ldots \sum_0^\infty n_k \ldots \ldots \ldots \ldots e^{-\beta \sum_k n_k (\epsilon_k - \mu)} = -\frac{\partial}{\partial \beta} \sum_0^\infty n_k \ldots \sum_0^\infty n_k \ldots \sum_0^\infty n_k \ldots \ldots \ldots \ldots e^{-\beta \sum_k n_k (\epsilon_k - \mu)} = \frac{\partial Z}{\partial \beta}. \tag{17.25}
\]
We therefore obtain
\[
\langle N \rangle = \frac{1}{Z} \frac{\partial Z}{\partial \beta \mu} = \frac{\partial \ln Z}{\partial \beta \mu}, \quad \langle E \rangle = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{\partial \ln Z}{\partial \beta}.
\] (17.26)

These two results emphasize the importance of the Partition Function in calculating fundamental quantities of a system. We will find this useful when we try to compute the thermodynamic quantities for a given system.

### 17.4 Free Fermions in a Periodic Box

In this section we will discuss the behavior of free fermions in a periodic box of length \(L\) in one, two and three dimensions and analyze our results. We know that no more than one fermion can occupy the same state. Hence, the partition function for fermions is given by
\[
Z(\beta, \beta \mu) = \prod_{k=0}^{\infty} \frac{1}{\sum_{n_k=0}^{\infty} e^{-\beta n_k (\epsilon_k - \mu)}} = \prod_{k=0}^{\infty} \frac{1}{1 + e^{-\beta (\epsilon_k - \mu)}}
\] (17.27)
which results in
\[
\langle N \rangle = \frac{\partial}{\partial \beta \mu} \ln Z = \frac{\partial}{\partial \beta \mu} \sum_{k=0}^{\infty} \ln (1 + e^{-\beta (\epsilon_k - \mu)}) = \sum_{k=0}^{\infty} \frac{e^{-\beta (\epsilon_k - \mu)}}{1 + e^{-\beta (\epsilon_k - \mu)}}
\] (17.28)
so
\[
\langle N \rangle = \sum_{k=0}^{\infty} \frac{1}{1 + e^{\beta (\epsilon_k - \mu)}},
\] (17.29)
Likewise through similar means we obtain that the average energy is given by
\[
\langle E \rangle = \sum_{k=0}^{\infty} \frac{\epsilon_k}{1 + e^{\beta (\epsilon_k - \mu)}},
\] (17.30)
The question is how we know that fermions exist in nature. In order to determine their existence, we need to know their energy states. In this section we consider the simplest case, that of free non-interacting fermions in a periodic box of length \(L\).

#### 17.4.1 Free Fermions in a Three Dimensional Periodic Box

In three dimensions each fermion must satisfy the Schrödinger Equation and boundary conditions respectively given by
\[
\frac{-\hbar^2}{2m} \left[ \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right] \psi(x, y, z) = E \psi(x, y, z)
\]
\[
\psi(x, y, z) = \psi(x + L, y, z) = \psi(x, y + L, z) = \psi(x, y, z + L).
\] (17.31)
It is easy to show that the eigenenergies are given by
\[
\epsilon_{k_1,k_2,k_3} = \frac{\hbar^2}{2m} \left( \frac{2\pi}{L} \right)^2 (k_1^2 + k_2^2 + k_3^2) \tag{17.32}
\]
where \(k_1, k_2\) and \(k_3\) are integers.\(^3\)

Once we know the different energy states we can find the expectation value of the number of particles and the expectation value of energy using (17.29) and (17.30).

**More Expectation Values:** We obtain
\[
\langle N \rangle = \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \sum_{k_3=-\infty}^{\infty} \frac{1}{1 + e^{\beta(\epsilon_{k_1,k_2,k_3} - \mu)}}. \tag{17.33}
\]

The trick to solve this problem is to turn these ordinary sums into Riemann sums. Let
\[
p_i = \frac{2\pi \hbar}{L} k_i \quad \text{for} \quad i = 1, 2, 3 \tag{17.34}
\]
and notice that
\[
\Delta p_i = p_{i+1} - p_i = \frac{2\pi \hbar}{L} \quad \text{for} \quad i = 1, 2, 3. \tag{17.35}
\]
Substituting these in we obtain
\[
\langle N \rangle = \frac{L^3}{(2\pi \hbar)^3} \sum_{p_1=-\infty}^{\infty} \sum_{p_2=-\infty}^{\infty} \sum_{p_3=-\infty}^{\infty} \frac{\Delta p_1 \Delta p_2 \Delta p_3}{1 + e^{\beta(p_2^2/m - \mu)}} \tag{17.36}
\]
where \(p^2 = p_1^2 + p_2^2 + p_3^2 = \left( \frac{2\pi \hbar}{L} \right)^2 (k_1^2 + k_2^2 + k_3^2)\).

To proceed we consider the particle density defined by \(\rho = \frac{\langle N \rangle}{L^3}\), then
\[
\rho(\beta) = \frac{1}{(2\pi \hbar)^3} \sum_{p_1=-\infty}^{\infty} \sum_{p_2=-\infty}^{\infty} \sum_{p_3=-\infty}^{\infty} \frac{\Delta p_1 \Delta p_2 \Delta p_3}{1 + e^{\beta(p_2^2/m - \mu)}}. \tag{17.37}
\]

Now, if we let \(L \to \infty\) we have that \(\Delta p_i \to 0\). Thus our Riemann Sum becomes a definite integral,
\[
\rho(\beta) = \frac{1}{(2\pi \hbar)^3} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{dp_1 dp_2 dp_3}{1 + e^{\beta(p_2^2/m - \mu)}}. \tag{17.38}
\]
This integral is easier to handle if we transform to spherical polar coordinates which results in
\[
\rho(\beta) = \frac{1}{(2\pi \hbar)^3} \int_{-1}^{1} \int_{0}^{2\pi} \int_{0}^{\infty} p^2 dp d\phi d\cos \theta \frac{1}{1 + e^{\beta(p_2^2/m - \mu)}}. \tag{17.39}
\]
Since our integral is independent of \(\phi\) and \(\cos \theta\), we obtain
\[
\rho(\beta) = \frac{4\pi}{(2\pi \hbar)^3} \int_{0}^{\infty} \frac{p^2 dp}{1 + e^{\beta(p_2^2/m - \mu)}}. \tag{17.40}
\]

\(^3\)Notice that each energy state depends on three variables corresponding to the three degrees of freedom expected in three dimensions.
We have obtained a formula for the particle density as a function of temperature. We would like to see what happens to the density as the temperature goes to zero and the chemical potential is kept fixed. This will be our next task.

If $T \to 0$, then $\beta \to \infty$ so we obtain

$$
\lim_{\beta \to \infty} \rho = \lim_{\beta \to \infty} \frac{4\pi}{(2\pi\hbar)^3} \int_0^\infty \frac{p^2 dp}{1 + e^{\beta \left( \frac{p^2}{2m} - \mu \right)}}.
$$

(17.41)

Now we need to evaluate this limit cleverly.

### Evaluating the Limit:

We can conveniently split the integral into two parts

$$
\lim_{\beta \to \infty} \rho = \lim_{\beta \to \infty} \frac{4\pi}{(2\pi\hbar)^3} \left[ \int_0^{\sqrt{2m\mu}} \frac{p^2 dp}{1 + e^{\beta \left( \frac{p^2}{2m} - \mu \right)}} + \int_{\sqrt{2m\mu}}^\infty \frac{p^2 dp}{1 + e^{\beta \left( \frac{p^2}{2m} - \mu \right)}} \right].
$$

(17.42)

Notice that the second integral is zero

$$
\lim_{\beta \to \infty} \int_{\sqrt{2m\mu}}^\infty \frac{p^2 dp}{1 + e^{\beta \left( \frac{p^2}{2m} - \mu \right)}} = 0
$$

(17.43)

because when $p > \sqrt{2m\mu}$ we have that $\frac{p^2}{2m} - \mu$ is positive and therefore that $\lim_{\beta \to \infty} e^{\beta \left( \frac{p^2}{2m} - \mu \right)} \to e^\infty$.

On the other hand the first integral is

$$
\lim_{\beta \to \infty} \int_0^{\sqrt{2m\mu}} \frac{p^2 dp}{1 + e^{\beta \left( \frac{p^2}{2m} - \mu \right)}} = \int_0^{\sqrt{2m\mu}} p^2 dp
$$

(17.44)

because when $p < \sqrt{2m\mu}$ we have that $\frac{p^2}{2m} - \mu$ is negative and therefore that $\lim_{\beta \to \infty} e^{\beta \left( \frac{p^2}{2m} - \mu \right)} \to e^{-\infty} = 0$.

Then,

$$
\lim_{\beta \to \infty} \rho = \lim_{\beta \to \infty} \frac{4\pi}{(2\pi\hbar)^3} \int_0^{\sqrt{2m\mu}} p^2 dp = \frac{4\pi}{3(2\pi\hbar)^3} (2m\mu)^{3/2}.
$$

(17.45)

At this point we introduce the term **Fermi Energy** which is denoted using $\mu_F$ and defined as the chemical potential at $T = 0$[^4].

Thus we have

$$
\rho_0 = \frac{4\pi}{3(2\pi\hbar)^3} (2m\mu_F)^{3/2}
$$

(17.46)

[^4]: Fermi Temperature is defined as $T_F = \frac{\mu_F}{k_B}$.
where \( \rho_0 = \lim_{\beta \to \infty} \rho \).

This is an important formula because it can be used to find the Fermi energy of a system if we know the particle density. Rearranging we have

\[
\mu_F = \frac{\hbar^2}{2m} \left( \frac{3\pi^2 \rho}{2} \right)^{2/3}.
\] (17.47)

Both the Fermi Energy and Fermi Temperature depend on the system we consider. For instance, we might calculate the Fermi Energy for electrons in a conductor or for neutrons in a neutron star. To put this into perspective we have a couple of examples.

- For electrons in copper the conduction electron density is about \( 8.5 \times 10^{28} \frac{1}{m^3} \) and the mass of an electron is 0.51 MeV, thus we obtain \( \mu_F = 4.5 \text{ eV} \) and \( T_F = 52000 \text{ K} \).
- For neutrons in a neutron star the neutron density is about \( 2.2 \times 10^{44} \frac{1}{m^3} \) and the mass of a neutron 940 MeV, hence we get \( \mu_F = 46 \text{ MeV} \) and \( T_F = 5.3 \times 10^{11} \text{ K} \).

So why define a Fermi Energy? The Fermi Energy is the energy required to add a particle to the system at zero temperature. These numerical examples reveal the enormous range of Fermi Energies that occur in nature. In this case the difference is 7 orders of magnitude!

We move on to discuss the relationship between the chemical potential and temperature. Suppose that we have some new temperature \( T \neq 0 \). We will consider the simplest case where we have a fixed particle density, \( \rho_0 \), for all \( T \), so \( \rho(\beta) = \rho_0 \). Then, using (17.40) and (17.46) we have

\[
\frac{4\pi}{3(2\pi\hbar)^3} (2m\mu_F)^{3/2} = \frac{4\pi}{(2\pi\hbar)^3} \int_0^\infty \frac{p^2 dp}{1 + e^{\beta \left( \frac{p^2}{2m} - \mu \right)}}
\] (17.48)

and simplifying

\[
\frac{(2m\mu_F)^{3/2}}{3} = \int_0^\infty \frac{p^2 dp}{1 + e^{\beta \left( \frac{p^2}{2m} - \mu \right)}}.
\] (17.49)

Using the fact that \( \epsilon = \frac{p^2}{2m} \), we have \( dp = \frac{\sqrt{2m}}{2} \frac{1}{\sqrt{\epsilon}} d\epsilon \) and we obtain

\[
\frac{(2\mu_F)^{3/2}}{3} = \int_0^\infty \frac{\sqrt{\epsilon} d\epsilon}{1 + e^{\beta (\epsilon - \mu)}}.
\] (17.50)

If we now consider the Fermi Energy \( \mu_F \) as our natural unit of energy and the Fermi Temperature \( T_F \) as our natural unit for temperature, we can perform the change of variables \( \epsilon = x \mu_F, \mu = s \mu_F \) and \( T = t T_F \). We then obtain

\[
\int_0^\infty \frac{\sqrt{x} dx}{1 + e^{\frac{\beta}{s} (x-s)}} = \frac{2}{3}.
\] (17.51)

\footnote{The following numbers differ from the experimentally acquired ones since we are not accounting for spin.}
The units used are natural units for both the chemical potential and temperature.

This is the implicit relation between chemical potential and temperature. We should expect from physical ground that chemical potential is a function of temperature. That means that given a certain temperature there is a unique chemical potential that satisfies the previous relation. \((17.51)\) is difficult to solve analytically for all \(t\) but we can consider the simple case where \(s = 0\), that is when the chemical potential is 0, and label the corresponding temperature \(t_c\), the critical temperature. After the simple substitution \(y = \frac{x}{t}\) we obtain

\[
\frac{2}{3} = t_c^3 \int_0^\infty \frac{\sqrt{y}dy}{1 + e^y}.
\]

(17.52)

This is a standard integral that can be found in any integral table which evaluates to

\[
\frac{2}{3} = t_c^3 \left(1 - \frac{1}{\sqrt{2}}\right) \Gamma \left(\frac{3}{2}\right) \zeta \left(\frac{3}{2}\right)
\]

(17.53)

The numerical value for \(t_c\) is 0.98873. This value is independent of the system considered as it is dependent on our choice of natural units. Note that the critical temperature is close to the Fermi Temperature.

In order to analyze the general behavior of the chemical potential as a function of temperature, we employed numerical methods using \((17.51)\) and including numerical integration and a nested analysis of variables. Fig. 17.1 shows the behavior of chemical potential as a function of temperature for fermions in a three dimensional box. The Fig. shows that the relation in \((17.51)\) is indeed a function where the independent variable is temperature. Furthermore, it show that chemical potential

\[\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}\] is the Riemann Zeta function.
potential decreases as temperature increases.

Up until now we have discussed in detail the particle density and we have derived various related results. We will now discuss the energy density defined as $\epsilon(\beta) = \langle E \rangle_{L^3}$.

**Discussing Energy Density:** Utilizing the same clever tricks we used for particle density we arrive at

$$
\epsilon(\beta) = \frac{4\pi}{(2\pi\hbar)^3} 2m \int_0^\infty \frac{p^4 dp}{1 + e^{\beta \left( \frac{p^2}{2m} - \mu(\beta) \right)}}.
$$

(17.54)

Simplifying that we obtain

$$
\epsilon(\beta) = \frac{4\pi}{(2\pi\hbar)^3} \sqrt{2} m^{3/2} \int_0^\infty \frac{e^{3/2} dx}{1 + e^{\beta (E - \mu(\beta))}}
$$

(17.55)

which is

$$
\epsilon(\beta) = \frac{4\pi}{(2\pi\hbar)^3} \sqrt{2} m^{3/2} \int_0^\infty \frac{x^{3/2} dx}{1 + e^{\beta (x - s(\beta))}}
$$

(17.56)

in natural units. We then define $u(t) = \frac{\epsilon(\beta)}{\rho_F}$ which represents the average energy in natural units. Substituting that variable into the above equation we get

$$
u(t) = \frac{4\pi}{(2\pi\hbar)^3} \sqrt{2} m^{3/2} \int_0^\infty \frac{x^{3/2} dx}{1 + e^{\beta (x - s(t))}}.
$$

(17.57)

An interesting quantity to consider is the average energy per particle as a function of temperature at fixed density which is given by

$$
\frac{u(t)}{\rho_0} = \frac{3}{2} \int_0^\infty \frac{x^{3/2} dx}{1 + e^{\frac{1}{1}(x - s(t))}}.
$$

(17.58)

This is a very complicated function because $s$ depends on $t$ rather nontrivially. However, we can loop at the case where $t = 0$. Then we have $s(0) = 1$ and

$$
\frac{u(0)}{\rho_0} = \frac{3}{2} \lim_{t \to 0} \int_0^\infty \frac{x^{3/2} dx}{1 + e^{\frac{1}{1}(x - 1)}}.
$$

(17.59)

We can take the limit as $t \to 0$ and we obtain

$$
\lim_{t \to 0} \int_0^\infty \frac{x^{3/2} dx}{1 + e^{\frac{1}{1}(x - 1)}} = \frac{2}{5}.
$$

(17.60)

Thus we have the relation

$$
\frac{u(0)}{\rho_0} = \frac{3}{5}.
$$

(17.61)

---

7 Notice that we are implicitly using the fact that we found the dependance of the chemical potential as function of temperature of $\beta$.

8 This is found by splitting the integral into two parts and noticing that one of the integrals vanishes leaving $\int_0^\infty x^{3/2} dx$. 

178
This value is a constant independent of the system! Hence, the average energy per fermion at zero temperature is $3\frac{3}{2}\mu_F$. Usually, we would expect that the average energy per particle would be zero at zero temperature, but this cannot be the case. The reason is that since fermions are not able to occupy the same energy state, it prevents the fermions to recede in the ground state.

With the aid of a computer we can plot the average energy per particle as a function of the temperature parameter $t$ by numerically analyzing the relation given by (17.58). The result is shown in Fig. 17.2. This is an important result because it is showing that for large temperatures, the average energy per fermion goes linearly proportional to the temperature with a proportionality constant of $\frac{3}{2}$. This is exactly what classical physics predicts with the non-relativistic equipartition theorem for a system with 3 degrees of freedom!

### 17.4.2 Free fermions in Two Dimensional Periodic Box

In a similar fashion now we will discuss fermions in a two dimensional periodic box. Most of the calculations are very similar to those performed in three dimensions with a couple of exceptions. We start with the claim that the energy eigenvalues are

$$
\epsilon_{k_1,k_2} = \frac{\hbar^2}{2m} \left(\frac{2\pi}{L}\right)^2 (k_1^2 + k_2^2)
$$

(17.62)

where $k_1$ and $k_2$ are integers and there are two for the sake of having two degrees of freedom.
The expectation value of the number of particles is given by
\[ \langle N \rangle = \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \frac{1}{1 + e^{\beta (\epsilon_{k_1,n_2} - \mu)}}. \] (17.63)

Using the same trick as before and defining our density as \( \rho = \frac{\langle N \rangle}{L^2} \) we have
\[ \rho(\beta) = \frac{1}{2\pi \hbar} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{dp_1 dp_2}{1 + e^{\beta \left( \frac{p_1^2}{2m} - \mu \right)}} \] (17.64)
where \( p^2 = p_1^2 + p_2^2 \). Now we make a change of variables to polar coordinates and obtain
\[ \rho(\beta) = \frac{1}{(2\pi)^2} \int_0^{2\pi} \frac{p d\theta dp}{1 + e^{\beta \left( \frac{p^2}{2m} - \mu \right)}}. \] (17.65)

The particle density at zero temperature is
\[ \rho_0 = \frac{2\pi}{(2\pi)^2} m \mu_F \] (17.66)
where \( \mu_F \) is the chemical potential at zero temperature.

Thus we obtain that the chemical potential as a function of particle density is given by
\[ \mu_F = \frac{\hbar^2}{2m} \frac{4\pi \rho}{2\pi \hbar}. \] (17.67)

The relationship between chemical potential and temperature is in this case given by
\[ \int_0^{\infty} \frac{dx}{1 + e^{\beta (x-s)}} = 1. \] (17.68)

The critical temperature is found by setting \( s = 0 \). After the substitution \( y = x/t_c \), we obtain
\[ t_c \int_0^{\infty} \frac{dy}{1 + e^{y}} = 1. \] (17.69)

The integral can be found in a table or can be computed using a trigonometric substitution. Its value is given by
\[ \int_0^{\infty} \frac{dy}{1 + e^{y}} = \ln 2. \] (17.70)

Therefore,
\[ t_c = \frac{1}{\ln(2)}. \] (17.71)

The numerical value for the critical temperature in this case is 1.4427. Fig. 17.3 shows the behavior of the chemical potential as a function of temperature for fermions in two dimensions. We can go further and define the average energy density as \( \epsilon(\beta) = \frac{\langle E \rangle}{L^2} \). Then,
\[ \epsilon(\beta) = \frac{2\pi}{(2\pi)^2} \frac{1}{2m} \int_0^{\infty} \frac{p^3 dp}{1 + e^{\beta \left( \frac{p^2}{2m} - \mu(\beta) \right)}} \] (17.72)
or in natural units its given by
\[
\epsilon(\beta) = \frac{2\pi}{(2\pi\hbar)^2 m\mu_F} \int_0^\infty \frac{xdx}{1 + e^{\beta(x-s(\beta))}}.
\] (17.73)

Defining \( u(t) = \frac{\epsilon(\beta)}{\mu_F} \), we have that the average energy per particle is
\[
\frac{u(t)}{\rho_0} = \int_0^\infty \frac{xdx}{1 + e^{\frac{1}{4}(x-s(t))}}.
\] (17.74)

The case \( t = 0 \) results in
\[
\frac{u(0)}{\rho_0} = \lim_{t \to 0} \int_0^\infty \frac{xdx}{1 + e^{\frac{1}{4}(x-1)}} = \frac{1}{2}.
\] (17.75)

Thus, the average energy per particle is \( \frac{1}{2} \mu_F \). This value is slightly lower than what we obtained in the three dimensional case. Fig. 17.4 shows the average energy per fermion in two dimensions as a function of the temperature. Observe that for large \( t \), the average energy per fermion is proportional to \( t \) with proportionality constant 1, corresponding to the two degrees of freedom of the fermions. This result also agrees with the equipartition theorem! Once again to put this into perspective we consider our electron and neutron examples.

- For conducting electrons is a copper sheet of one atom thickness the electron density is \( 2.3 \cdot 10^{19} \frac{1}{\text{m}^2} \) which results in \( \mu_F = 11 \text{ eV} \) and \( T_F = 130000 \text{ K} \).
- For neutrons in a two dimensional neutron star the particle density is \( 2.2 \cdot 10^{29} \frac{1}{\text{m}^2} \) which results in \( \mu_F = 34 \text{ MeV} \) and \( T_F = 3.9 \cdot 10^{11} \text{ K} \).
The blue curve is the function being analyzed while the red curve is $t$.

### 17.4.3 Free fermions in One Dimensional Periodic Box

Finally we will consider fermions confined to a one dimensional periodic box.

The energy eigenvalues are

$$
\epsilon_{k_1} = \frac{\hbar^2}{2m} \left(\frac{2\pi}{L}\right) k_1^2 \tag{17.76}
$$

where $k_1$ is an integer corresponding to one degree of freedom.

The average number of particles is

$$
\langle N \rangle = \sum_{k_1=-\infty}^{\infty} \frac{1}{1 + e^{\beta(\epsilon_{k_1} - \mu)}} \tag{17.77}
$$

Likewise, defining $\rho = \frac{\langle N \rangle}{L}$

$$
\rho(\beta) = \frac{1}{2\pi\hbar} \int_0^{\infty} \frac{dp}{1 + e^{\beta(\frac{p^2}{2m} - \mu)}}. \tag{17.78}
$$

The density at zero temperature is found to be

$$
\rho_0 = \frac{1}{2\pi\hbar} \sqrt{2m\mu_F} \tag{17.79}
$$

where $\mu_F$ is the Fermi Energy.
The units used are natural units for both the chemical potential and temperature.

The relation between chemical potential and temperature is now given by

\[ 2 = \int_{0}^{\infty} \left( \frac{1}{\sqrt{x}} \right) \frac{dx}{1 + e^{\frac{1}{T}(x-s)}}. \quad (17.80) \]

Setting \( s = 0 \) and performing the change of variables \( y = \frac{x}{t_c} \) we find that the critical temperature is given by

\[ 2 = \sqrt{t_c} \int_{0}^{\infty} \left( \frac{1}{\sqrt{y}} \right) \frac{dy}{1 + e^{y}}. \quad (17.81) \]

This complicated integral is solved using an integral table. Solving for \( t_c \) we obtain\(^9\)

\[ t_c = \left( \frac{2}{(1 - \sqrt{2}) \sqrt{\pi} \zeta \left( \frac{1}{2} \right)} \right)^2. \quad (17.82) \]

The numerical value for \( t_c \) is 3.4798. The chemical potential dependence on temperature is shown in Fig. 17.5. Moreover, the average energy density defined as \( \epsilon(\beta) = \frac{\langle E \rangle}{L} \) so

\[ \epsilon(\beta) = \frac{1}{2\pi\hbar} \frac{1}{2m} \int_{0}^{\infty} \frac{p^2 dp}{1 + e^{\beta\left(\frac{p^2}{2m} - \mu(\beta)\right)}}. \quad (17.83) \]

or in natural units we have

\[ \epsilon(\beta) = \frac{1}{2\pi\hbar} \frac{\sqrt{2m}}{2} \mu_F^{3/2} \int_{0}^{\infty} \frac{\sqrt{x} dx}{1 + e^{\beta(x-s(\beta))}}. \quad (17.84) \]

\(^9\)The Riemann Zeta function presented here is in its generalized complex valued form that preserves analyticity. Its value is -1.4604.
The blue curve is the function being analyzed while the red curve is $\frac{1}{2}t$.

It follows that

$$\frac{u(t)}{\rho_0} = \frac{1}{2} \int_0^\infty \frac{\sqrt{x}dx}{1 + e^{\frac{1}{2}(x-s(t))}}.$$  \hspace{1cm} (17.85)

For the case $t = 0$, we have that

$$\frac{u(0)}{\rho_0} = \lim_{t \to 0} \frac{1}{2} \int_0^\infty \frac{\sqrt{x}dx}{1 + e^{\frac{1}{2}(x-1)}} = \frac{1}{3}. \hspace{1cm} (17.86)$$

This means that the average energy per particle is $\frac{1}{3} \mu_F$. Fig. 17.6 shows the average energy per fermion in one dimension as a function of the temperature. Notice that for large $t$, we have that the average energy per fermion is proportional to $t$ with proportionality constant $\frac{1}{2}$, corresponding to one degree of freedom of the fermions. Once again, this result also agrees with the equipartition theorem! It might be interesting to compare this to the result obtained in (13.8), which described the energy for a free particle in a periodic box of length $L$.

### 17.5 Free Bosons in a Periodic Box

Recall our discussion about the existence of two particles: fermions and bosons. We have already discussed some behaviors of fermions so in this section we will discuss the behavior of bosons in one, two and three dimensions. We will see that different physical phenomena occur for bosons which we did not observe for fermions.

As we previously determined, one of the main features that distinguishes bosons from fermions is that there is no restriction on the number of bosons that can occupy a single energy state. The
partition function in this case is then given by\[10\]

\[
Z(\beta, \beta\mu) = \prod_{k=0}^{\infty} \sum_{n_k=0}^{\infty} e^{-\beta n_k (\epsilon_k - \mu)} = \prod_{k=0}^{\infty} \frac{1}{1 - e^{-\beta (\epsilon_k - \mu)}}
\]

which gives us

\[
\langle N \rangle = \frac{\partial}{\partial \beta \mu} \ln Z = \frac{\partial}{\partial \beta \mu} \sum_{k=0}^{\infty} \ln \left( \frac{1}{1 - e^{-\beta (\epsilon_k - \mu)}} \right) = \sum_{k=0}^{\infty} \frac{e^{-\beta (\epsilon_k - \mu)}}{1 + e^{-\beta (\epsilon_k - \mu)}}
\]

\[
\langle N \rangle = \sum_{k=0}^{\infty} \frac{1}{e^{\beta (\epsilon_k - \mu)} - 1}.
\]

Similarly,

\[
\langle E \rangle = \sum_{k=0}^{\infty} \frac{\epsilon_k}{e^{\beta (\epsilon_k - \mu)} - 1}.
\]

Now we would like to discuss free non interacting bosons in a periodic box of length \(L\). The possible energy eigenvalues are identical to those of fermions since both obey the Schrödinger Equation and the same boundary conditions.

### 17.5.1 Free Bosons in a Three Dimensional Periodic Box

It follows that the average number of particles is

\[
\langle N \rangle = \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} \sum_{k_3=-\infty}^{\infty} \frac{1}{e^{\beta (\epsilon_{k_1, k_2, k_3} - \mu)} - 1}.
\]

By the same arguments employed before and assuming that the density is constant, we obtain that

\[
\rho = \frac{4\pi}{(2\pi\hbar)^3} \int_0^{\infty} \frac{p^2 dp}{e^{\beta (p^2/2m) - \mu)} - 1}.
\]

If we attempt to find the chemical potential at zero temperature by taking the limit \(T \to 0\) we obtain

\[
\rho = -\frac{4\pi}{(2\pi\hbar)^3} \frac{(2m\mu)^{3/2}}{3}.
\]

This is a negative quantity which is impossible because average density cannot be physically negative! Therefore, our mathematical tricks must not be valid for bosons. Which step was the forbidden one? The answer is that the function inside the sum has a singularity at \(\epsilon_{k_1, k_2, k_3} - \mu = 0\).

If \(\mu < 0\), there is no problem because \(\epsilon_{k_1, k_2, k_3} \geq 0\).

---

\[10\] The second equality comes from the fact that the series inside the product is a geometric series so

\[
\sum_{n_k=0}^{\infty} e^{-\beta n_k (\epsilon_k - \mu)} = \sum_{n_k=0}^{\infty} (e^{-\beta (\epsilon_k - \mu)})^{n_k} = \frac{1}{1 - e^{-\beta (\epsilon_k - \mu)}}.
\]
The Mathematical Trick to Bosons: In order to solve this issue we will analyze the problem in a slightly different way. Let us assume that the sum to integral conversion is allowed. Then, multiplying both the numerator and denominator of the left hand side of (17.92) by \(e^{-\beta \frac{p^2}{2m} - \mu}\),

\[
\rho = \frac{4\pi}{(2\pi \hbar)^3} \int_0^\infty \frac{e^{-\beta \frac{p^2}{2m} \mu}}{1 - e^{-\beta \frac{p^2}{2m} - \mu}} dp.
\]  (17.94)

Making the substitution \(\alpha = e^{\beta \mu}\) and \(\frac{3p^2}{2m} = x^2\) we get

\[
\rho = \frac{1}{2\pi^2 \hbar^3} \left(\frac{2m}{\beta}\right)^{3/2} \int_0^\infty \frac{\alpha x^2 e^{-x^2}}{1 - \alpha e^{-x^2}} dx.
\]  (17.95)

The trick is to notice that this is just a geometric sum so

\[
\sum_{k=1}^\infty \left(\alpha e^{-x^2}\right)^k = \frac{\alpha e^{-x^2}}{1 - \alpha e^{-x^2}}
\]  (17.96)

and so (17.95) can be expressed as

\[
\rho = \frac{1}{2\pi^2 \hbar^3} \left(\frac{2m}{\beta}\right)^{3/2} \int_0^\infty x^2 \sum_{k=1}^\infty \left(\alpha e^{-x^2}\right)^k dx.
\]  (17.97)

Interchanging the sum and integral we have

\[
\rho = \frac{1}{2\pi^2 \hbar^3} \left(\frac{2m}{\beta}\right)^{3/2} \sum_{k=1}^\infty \alpha^k \int_0^\infty x^2 e^{-kx^2} dx.
\]  (17.98)

The integral is already familiar to us\(^{11}\) so it follows that

\[
\rho = \left(\frac{mk_B T}{2\pi \hbar^2}\right)^{3/2} \sum_{k=1}^\infty \alpha^k k^{3/2}.
\]  (17.99)

Often \(\sum_{k=1}^\infty \frac{\alpha^k}{k^{3/2}}\) is denoted as \(\zeta_{3/2}(\alpha)\) and called the Parametrized Riemann Zeta Function or the Polylogarithm Function. We finally obtain

\[
\rho = \left(\frac{mk_B T}{2\pi \hbar^2}\right)^{3/2} \zeta_{3/2}(1).
\]  (17.100)

Notice that \(\zeta_{3/2}(\alpha)\) only converges for \(\alpha \leq 1\) or equivalently when \(\beta \mu \leq 0\). Since \(\beta > 0\), then the chemical potential must be less than or equal to zero (\(\mu \leq 0\)). Physically, this means that a system of free non-interacting bosons is willing to accept more particles. We would like to investigate at what temperature the chemical potential is zero. This is the critical temperature and is denoted by \(T_c\). We have

\[
\rho = \left(\frac{mk_B T_c}{2\pi \hbar^2}\right)^{3/2} \zeta_{3/2}(1).
\]  (17.101)

\(^{11}\) \(\int_0^\infty x^2 e^{-x^2} dx = \frac{\sqrt{\pi}}{4x^{3/2}}\).

\(^{12}\) It is possible to have \(\beta < 0\) (negative temperature). However we will not consider that possibility.
Solving for \( T_c \) and using the fact that \( \zeta_{3/2}(1) = 2.6124 \), we obtain

\[
T_c = \frac{2\pi\hbar^2}{mk_B} (0.3828\rho)^{2/3}.
\]  (17.102)

For example, liquid Helium-4 has a particle density of \( 7.2 \cdot 10^{24} \ \text{m}^{-3} \), hence we obtain\(^{13}\) \( T_c = 2.9K \).

It is possible to numerically solve (17.100) for \( T < T_c \). First it will be useful to transform to natural units. In this case our natural unit for temperature will be \( T_c \) and our natural unit for energy will be \( T_c k_B \). If we now write (17.100) in terms of these units using the substitution \( T = tT_c \), \( \mu = sT_c k_B \) and equate the resulting equation with (17.101), we have the following relation

\[
1 = t^{3/2} \frac{\zeta_{3/2}(\alpha)}{2.6124} \]  (17.103)

where \( \alpha = e^{s/t} \). This equation is only valid for \( t > 1 \).

For the case \( 0 < t < 1 \), (17.103) cannot be solved. This can be reconciled if we refer back to what we said at the beginning of our discussion: there is a possible singularity at \( \epsilon = \mu \). In this case we hypothesize that the chemical potential is zero for \( t \), thus our singularity occurs at \( \epsilon = 0 \) (when particles are in the ground state.) We will state without proof that this singularity resembles that of a delta function, i.e that the contribution of \( \epsilon = 0 \) is finite. Therefore, our equation relating chemical potential and temperature is,

\[
\rho = \rho_0(t) + \left( \frac{mk_B T_c}{2\pi\hbar^2} \right)^{3/2} t^{3/2} \]  (17.104)

where \( \rho_0(t) \) is the finite contribution from the density of particles in the ground state.

Solving for \( \rho_0(t) \) using (17.104) we obtain the relation

\[
\rho_0(t) = \rho \left( 1 - t^{3/2} \right). \]  (17.105)

Fig. 17.7 shows the chemical potential as a function of temperature for bosons in three dimensions. Fig. 17.8 shows the fraction of bosons in the ground state as a function of temperature. This implies for temperature less than the critical temperature there are a finite non-zero number of bosons in the ground state. This is known as Bose-Einstein Condensation. Interestingly enough, at zero temperature all the bosons condense to the ground state. This phenomena had very interesting applications and consequences, one of them is superfluidity which we will not discuss in this text.

Similar to Fermions we can now discuss the average energy density \( \epsilon(\beta) = \frac{\langle E \rangle}{L^3} \). In this case we have

\[
\epsilon(\beta) = \frac{4\pi}{(2\pi\hbar)^3} \frac{1}{2m} \int_0^\infty e^{-\beta \left( \frac{p^2}{2m} - \mu \right)} p^3 dp \]  (17.106)

\(^{13}\) Again this might differ from the experimental values obtained since we are not taking spin into account.
Figure 17.7: Dimensionless chemical potential as a function of temperature parameter for free bosons in a three dimensional box.

The units used are natural units for both the chemical potential and temperature.

Figure 17.8: Fraction of bosons on the ground state as a function of temperature parameter for free bosons in a three dimensional box.

Notice that $\rho_0(0) = 1$ and $\rho_0(1) = 0$. 

\[ \rho = \frac{\rho_0(0)}{\rho} = 1 \]
Now we use the substitution $\alpha = e^{\beta \mu}$ and $\frac{\beta p^2}{2m} = x^2$. Then we have

$$\epsilon(\beta) = \frac{1}{2\pi^2 h^3} \frac{1}{2m} \left(\frac{2m}{\beta}\right)^{5/2} \sum_{k=1}^{\infty} \alpha^k \int_0^{\infty} x^4 e^{-kx^2} dx.$$

(17.107)

We already know how to evaluate this integral so we obtain

$$\int_0^{\infty} x^4 e^{-kx^2} dx = \frac{3\sqrt{\pi}}{8k^{5/2}}.$$

(17.108)

It follows that

$$\epsilon(\beta) = \frac{3}{2} \left(\frac{mk_B T}{2\pi h^2}\right)^{3/2} (k_B T) \zeta_{5/2}(\alpha).$$

(17.109)

If we define $u(t) = \frac{e(\beta)}{T_B(k_B T)}$ to represent the average energy in natural units, we have

$$u(t) = \frac{3}{2} \left(\frac{mk_B T_c}{2\pi h^2}\right)^{3/2} t^{5/2} \zeta_{5/2}(\alpha).$$

(17.110)

Notice that the last two equations are valid for any $t$. However in particular we have that for $t < 1$,

$$\epsilon(\beta) = \frac{3}{2} \left(\frac{mk_B T}{2\pi h^2}\right)^{3/2} (k_B T) \zeta_{5/2}(1)$$

$$u(t) = \frac{3}{2} \left(\frac{mk_B T_c}{2\pi h^2}\right)^{3/2} t^{5/2} \zeta_{5/2}(1).$$

(17.111)

We can consider the average energy per particle and we have

$$u(t) = \frac{3}{2 t^{5/2} \zeta_{5/2}(1)} \frac{\zeta_{5/2}(\alpha)}{\zeta_{5/2}(\alpha)}$$

for $t > 1$.

$$u(t) = \frac{3}{2} t^{5/2} \zeta_{5/2}(1) \zeta_{5/2}(\alpha)$$

for $t < 1$.

(17.112)

The following Fig. shows the average energy per particle as a function of temperature. Lo and behold, this agrees with the equipartition theorem and it also shows us that at zero temperature the energy is zero.

Another interesting quantity to calculate is the dimensionless heat capacity defined as $c = \frac{d}{dt} \left(\frac{u(t)}{\rho}\right)$.

Mathematical Identity 7: To help us with this task we have the relation

$$\frac{d\zeta_s(\alpha)}{d\alpha} = \frac{1}{\alpha} \zeta_{s-1}(\alpha).$$

(17.113)

Proof.

$$\frac{d\zeta_s(\alpha)}{d\alpha} = \frac{d}{dt} \sum_{k=1}^{\infty} \frac{\alpha^k}{k^s} = \sum_{k=1}^{\infty} \frac{d}{dt} \frac{\alpha^k}{k^s} = \sum_{k=1}^{\infty} \frac{\alpha^{k-1}}{k^{s-1}} = \frac{1}{\alpha} \sum_{k=1}^{\infty} \frac{\alpha^k}{k^{s-1}} = \frac{1}{\alpha} \zeta_{s-1}(\alpha).$$

(17.114)
To continue we will need to prove another useful result.

**Mathematical Identity 8:** Another useful result is that

\[
\frac{1}{\alpha} \frac{d\alpha}{dt} = \frac{3}{2} \frac{1}{\alpha} \zeta_{3/2}(\alpha) - \frac{3}{2} t \zeta_{1/2}(\alpha).
\]  

(17.115)

**Proof.** In order to obtain this relation we use the fact that \(\rho\) is constant so \(\frac{d\rho}{dt} = 0\). Then

\[
\frac{d\rho}{dt} = \frac{3}{2} \left( \frac{mk_B T_c}{2\pi \hbar^2} \right)^{3/2} \frac{d}{dt} \left( t^{3/2} \zeta_{3/2}(\alpha) \right) = 0.
\]  

(17.116)

Since the coefficients cannot be 0, we must have

\[
\frac{d}{dt} \left( t^{3/2} \zeta_{3/2}(\alpha) \right) = 0
\]  

(17.117)

which implies that

\[
\frac{3}{2} t^{1/2} \zeta_{3/2}(\alpha) + t^{3/2} \frac{d\zeta_{3/2}(\alpha)}{d\alpha} \frac{d\alpha}{dt} = 0.
\]  

(17.118)

Using the previous result we have

\[
\frac{3}{2} t^{1/2} \zeta_{3/2}(\alpha) + t^{3/2} \frac{1}{\alpha} \zeta_{1/2}(\alpha) \frac{d\alpha}{dt} = 0.
\]  

(17.119)

Solving for \(\frac{1}{\alpha} \frac{d\alpha}{dt}\) we obtain the desired result. \(\Box\)
We are interested in obtaining \( \frac{d\zeta_s(\alpha)}{dt} \). Using the chain rule and the two previously obtained results, we have

\[
\frac{d\zeta_s(\alpha)}{dt} = \frac{d\zeta_s(\alpha)}{d\alpha} \frac{d\alpha}{dt} = -\frac{3}{2} \frac{\zeta_{3/2}(\alpha)}{t \zeta_{1/2}(\alpha)} \zeta_{s-1}(\alpha). \tag{17.120}
\]

It follows that

\[
c(t) = \begin{cases} 
\frac{3}{2} \left[ \frac{\zeta_{3/2}(\alpha)}{\zeta_{1/2}(\alpha)} - \frac{3}{2} \left( \frac{\zeta_{3/2}(\alpha)}{\zeta_{1/2}(\alpha)} - \frac{\zeta_{5/2}(\alpha)}{\zeta_{3/2}(\alpha)} \right) \right] & \text{for } t > 1 \\
\frac{15}{4} \beta^{3/2} \frac{\zeta_{5/2}(1)}{\zeta_{3/2}(1)} & \text{for } t < 1
\end{cases} \tag{17.121}
\]

The following plot shows the dimensionless heat capacity as a function of temperature.

17.5.2 Free Bosons in a Two Dimensional Periodic Box

We will follow the same approach as we did for bosons in three dimensions. In this case it is easy to show that the particle density is given by

\[
\rho = \frac{2\pi}{(2\pi\hbar)^2} \int_0^{\infty} e^{-\beta(\frac{p^2}{2m} - \mu)} p dp. \tag{17.122}
\]

Making the substitution \( \alpha = e^{\beta \mu} \) and \( \frac{\beta \mu^2}{2m} = x^2 \) this simplifies to

\[
\rho = \frac{1}{2\pi \hbar^2 \left( \frac{2m}{\beta} \right)} \int_0^{\infty} \frac{\alpha x e^{-x^2}}{1 - \alpha e^{-x^2}} dx. \tag{17.123}
\]
Then,
\[ \rho = \frac{1}{2\pi h^2} \int_0^\infty \left( \frac{2m}{\beta} \right)^{1/2} \sum_{k=1}^\infty \alpha e^{-kx^2} \, dx = \frac{1}{2\pi h^2} \left( \frac{2m}{\beta} \right)^{1/2} \sum_{k=1}^\infty \alpha^k \int_0^\infty x e^{-kx^2} \, dx. \]  

The last integral is easy to solve\(^{14}\) so it follows that
\[ \rho = \left( \frac{mk_B T}{2\pi h^2} \right) \sum_{k=1}^\infty \frac{\alpha^k}{k} = \left( \frac{mk_B T}{2\pi h^2} \right) \zeta(\alpha). \]  

If we try to solve for the critical temperature we have to set \( \mu = 0 \) so \( \alpha = 1 \). However \( \zeta(1) \) diverges since it is a harmonic series. Thus there is no critical temperature for bosons in two dimensions. This is indeed the limiting case since the sum diverges. This means that it is not possible to get a Bose-Einstein condensate in a two dimensional space.

### 17.5.3 Free Bosons in a One Dimensional Box

Having observed that Bose-Einstein Condensation does not occur in two dimensions, we can now see if it occurs in one dimension. The particle density is given by
\[ \rho = \frac{1}{2\pi h} \int_0^\infty e^{-\beta \frac{x^2}{2m} - \mu} \frac{dp}{1 - e^{-\beta \frac{x^2}{2m} - \mu}}. \]  

Making the same substitution \( \alpha = e^{\beta \mu} \) and \( \frac{p^2}{2m} = x^2 \) we get that
\[ \rho = \frac{1}{2\pi h} \left( \frac{2m}{\beta} \right)^{1/2} \int_0^\infty \alpha e^{-x^2} \frac{dx}{1 - \alpha e^{-x^2}}. \]  

Then,
\[ \rho = \frac{1}{2\pi h} \left( \frac{2m}{\beta} \right)^{1/2} \int_0^\infty \sum_{k=1}^\infty \left( \alpha e^{-x^2} \right)^k \, dx = \frac{1}{2\pi h} \left( \frac{2m}{\beta} \right)^{1/2} \sum_{k=1}^\infty \alpha^k \int_0^\infty e^{-kx^2} \, dx. \]  

The last integral is a gaussian integral that is easy to carry out, and\(^{15}\) it follows that
\[ \rho = \left( \frac{mk_B T}{8\pi h^2} \right)^{1/2} \sum_{k=1}^\infty \frac{\alpha^k}{k^{1/2}} = \left( \frac{mk_B T}{8\pi h^2} \right)^{1/2} \zeta(1/2)(\alpha). \]  

Once again the process of finding a critical temperature fails since the \( \zeta(1/2)(1) \) diverges as can be easily verified using the integral test. Thus no condensation occurs in one dimension.

We arrive at the interesting conclusion that we require a three dimensional system in order to observe Bose-Einstein Condensation. Since that seems like the easiest and only way to distinguish bosons from fermions, it looks like in a two or one dimensional system we would not be able to tell whether a particle was a fermion or a boson.

\[^{14}\int_0^\infty xe^{-kx^2} \, dx = \frac{1}{2\pi k}.
\[^{15}\int_0^\infty e^{-kx^2} \, dx = \frac{\sqrt{\pi}}{2\sqrt{k}}.
\]