# Introduction to
# Relativistic Quantum Mechanics

James C. Rothwell

*Lycoming College*

# Introduction to
# Relativistic Quantum Mechanics

James C. Rothwell

*Lycoming College*

Presented to the faculty of Lycoming College in partial
fulfillment of the requirements for graduation
with Departmental Honors in Physics.

Approved by:

Dr. David Fisher

Richard R. Erickson

Dr. Richard Erickson

Dr. Richard Weida

Dr. Santu DeSilva

# Preface

This paper details much of the theory and some of the applications of Relativistic Quantum Mechanics as seen in an Honors Independent Study . The study was largely problem-based, some of which appear scattered throughout the text as examples of the type of problems discussed.

This text is divided into 4 parts. Part I, consisting of chapters 1 to 2, reviews mathematical and theoretical techniques with which the student should be familiar. Part II, consisting of chapters 3 to 6, is a review of theory that should have been covered previously. The topics are not dealt with in detail in either of these first two parts as they are meant purely as a refresher. If the student discovers that a topic does not come back to memory, he should go back to notes and books from previous sources. Part III, consisting of chapter 7 examines the Quantum Theory of Measurement, and deals with the mathematics involved with that theory. Part IV, consisting of chapters 8 to 10, deals with the extension of this previous material to include relativistic effects.

Part I reviews the mathematics necessary for an understanding of quantum theory. Chapter 1 deals with mathematics associated with observables and operators, including inner-products, norms, Dirac's Bra and Ket notation, expectation values, and Hermitean operators. Chapter 2 investigates eigenvalues and eigenfunctions.

Part II provides an overview of the development of quantum mechanical theory from classical Newtonian mechanics through the old semi-classical quantum mechanics to the non-classical quantum mechanics which is the main content of most undergraduate quantum mechanics courses. Chapter 3 gives a historical perspective for this part, and we then review some of the concepts of Classical Newtonian mechanics in chapter 4. Chapter 5 looks closer at the experiments and solutions which gave rise to the old quantum theory, while chapter 6 reviews non-classical quantum theory.

Part III is very brief, and looks at the mathematics and results of the quantum theory of measurement.

Part IV begins with chapter 8, which looks at the basics of matrices as they pertain to quantum mechanics. Chapter 9, dealing with orbital and spin angular momenta, utilizes this concept thoroughly, and pushes us into Dirac's formulation of quantum theory, which is covered in chapter 10.

The appendix summarizes two different approximation techniques that are useful for solving problems in quantum mechanics.

# Table of Contents

Preface i

Table of Contents ii

Part I Mathematical Techniques 1

Chapter 1 Observables and Operators 2

1.1 Observables, Variables, and Operators 2

1.2 Linear Operators 2

1.3 Inner Products and Norms 4

1.4 Dirac's Bra and Ket Notation 5

1.5 Expectation Values 6

1.6 Hermitean Operators 7

1.7 Anti-Hermitean Operators 8

1.8 Adjoint Operators 9

1.9 Commutators 10

Chapter 2 Eigenfunctions and Eigenvalues 12

2.1 The Eigenvalue Problem 12

2.2 Linear Dependence and Independence 13

2.3 The Expansion Problem 14

Part II Elementary Principles 16

Chapter 3 Historical Review 17

3.1 Overview of Dates 17

Chapter 4 Review of Concepts of Classical Mechanics 19

4.1 Generalized Coordinates 19

4.2 Hamiltonian Mechanics 20

Chapter 5 Semi-Classical Quantum Mechanics 24

5.1 Blackbody Radiation 24

5.2 The Photoelectric Effect 25

5.3 The Compton Effect 26

5.4 de Broglie Matter Waves and the Davisson-Germer Experiment 28

5.5 Bohr's Quantum Theory of Spectra 29

5.6 Wilson-Sommerfeld Quantum Conditions 31

5.7 Sommerfeld's Relativistic Treatment of the Hydrogen Atom 34

Chapter 6 Non-Classical Quantum Mechanics 36

6.1 The Heisenberg Uncertainty Principle 36

6.2 Operator Derivation of the Schrödinger Wave Equation 37

6.3 The Schrödinger and Klein-Gordon Equations 39

6.4 Stationary States 41

6.5 Physical Interpretation of the Wave Function 42

6.6 Statistical Interpretation of the Wavefunction, Born's Probability Waves 44

6.7 The Postulates of Quantum Mechanics 45

6.8 The Schrödinger Picture and The Heisenberg Picture 46

Part III Quantum Theory of Measurement 48

Chapter 7 Quantum Theory of Measurement 49

7.1 Simultaneous and Repeatable Measurements 49

7.2 Compatible Observables 50

7.3 Compatible Constants of the Motion 50

7.3 Complete Sets of Compatible Observables 51

Part IV Relativistic Quantum Mechanics 53

Chapter 8 Matrix Representations 54

8.1 Linear Algebraic Techniques 54

8.2 Matrix Representations of Operators 55

8.3 Eigenfunctions and Eigenvectors 58

8.4 The Hilbert Space H0 59

8.5 Matrix Representation of the Schrödinger Equation 59

8.6 Transformation from the Schrödinger to Heisenberg Picture 60

Chapter 9 Angular Momentum and Electron Spin 63

9.1 The Angular Momentum Operators 63

9.2 The Angular Momentum Shift Operators 64

9.3 Eigenvalues and Eigenfunctions of M2 and Mz 65

9.4 Angular Momentum and Magnetic Moment 69

9.5 The Stern-Gerlach Experiment 70

9.6 Stern-Gerlach Patterns and Electron Spin 71

9.7 Pauli's Theory of Electron Spin, Spin Wavefunctions and Operators 72

9.8 Inner Products and Expectation Values in 2-Component Formalism 77

9.9 Eigenfunctions of Hydrogen for the Stern-Gerlach Experiment 78

9.10 The Normal Zeeman Effect in the Pauli Theory 80

9.11 The Anomalous Zeeman Effect 82

Chapter 10 Relativistic Quantum Mechanics 84

10.1 The Klein-Gordon Equation 84

10.2 Dirac's Relativistic Wave Equation 85

10.3 The Dirac Matrices 86

10.4 The Dirac Equations 87

10.5 Pauli's Theory as a Nonrelativistic Approximation to Dirac's Theory 89

10.6 Plane Wave Solutions for a Free Dirac Particle 90

10.7 Dirac's Theory of the Positron 92

10.8 Inner Products and Expectation Values 93
10.9 Spin Angular Momentum and Magnetic Moment 93
10.10 Angular Momentum Eigenvalues and Eigenfunctions 95
10.11 Solution for the Hydrogen Atom 97
10.12 Physical and Statistical Interpretation of Dirac Wave Functions 99
Appendix I Approximation Techniques i
A.1 The WKB Approximation i
A.2 Time-Independent Perturbation Theory iv
A.3 The Stark Effect for the Plane Rotator vii
Bibliography ix

# Part I

# Mathematical Techniques

## Chapter 1

# Observables and Operators

*This chapter reviews the concepts of observables and their associated operators, as well as the properties of those operators.*

### 1.1 Observables, Variables, and Operators

An *observable* is any physical quantity that can be measured. In classical mechanics the principle observables are *position, linear momentum, angular momentum,* and *energy.* Quantum mechanics adds *parity* and *spin* to that list.

According to classical mechanics, observables can be simultaneously measured with any degree of precision by sufficient refinement of the measurement apparatus. According to quantum mechanics, by the Heisenberg Uncertainty Principle, this is not the case.

In classical mechanics, observables are *represented* by ordinary mathematical variables. In quantum mechanics, observables are also *associated* with an *operator.* One must now distinguish between the observable $A$, the variable $A$, and the operator $A$. Distinguishing symbols could be used, but as there is seldom confusion as the meanings are clear from the context. If there is cause for confusion in a context, the symbol $\hat{A}$ will be used for the operator, and a plain $A$ for either the observable or the variable.

### 1.2 Linear Operators

An operator is a law by which we associate with each member of a certain set, a member of the same set, or of a different set. For the moment, the set of objects considered will be functions.

Some common operators are the translation operator and the multiplicative operator. The translation operator is of the form $\hat{T}f(x) = f(x + c)$, and the multiplicative operator is of the form $\hat{M}f(x) = c \cdot f(x)$.

The wavefunctions on which the operators of quantum mechanics operate are called *linear manifolds*. A function is a linear manifold if, for any two functions in the set, say $f_1$ and $f_2$, then $f_3 = c_1 \cdot f_1 + c_2 \cdot f_2$ is also in the set.

An operator $A$ is said to be *linear* if for every $f$ and $g$ of a linear manifold $M$, and for all complex numbers $c$ and $d$,

(1.1a) $Af$ is defined, and

(1.1b) $$A(cf + dg) = c \cdot Af + d \cdot Ag.$$

This definition does not require that $Af$ belongs to $M$ for every $f$ of $M$. If however, $Af$ does belong to $M$, for every $f$ of $M$, then we say that $A$ is *strongly linear* over $M$.

When $A$ and $B$ are linear over a linear manifold $M$, linear operators $cA$ and $A \pm B$ are defined by

(1.2) $$(cA)f = c(Af), \text{ and}$$

(1.3) $$(A \pm B)f = Af \pm Bf.$$

If $A$ and $B$ are strongly linear over $M$, we can further say that

(1.4) $$(AB)f = A(Bf), \text{ and}$$

(1.5) $$A^n f = A\left(A^{n-1} f\right), n = 2,3,4,\dots.$$

We say that $A = 0$ over $M$ if and only if $Af = 0$ for all $f$ in $M$.

*Problem 1.1:* Show that the following equations define linear operators:

(a) $Af(x) = f(-x)$

(b) $Af(x) = f(x) + f(-x)$

(c) $Af(x) = f(x) - f(-x)$

(d) $Af(x) = f(x + c)$

(e) $Af(x) = p\dfrac{d^2 f(x)}{dx^2} + q\dfrac{df(x)}{dx} + rf(x)$

where $p$, $q$, and $r$ may be functions of $x$.[1]

---

[1] Ikenberry, p. 68, problem 6.1

*Problem 1.2:* Let $A = \frac{4d^2}{dx^2}$, $B = -\frac{d^4}{dx^4}$, and $C = \frac{d^2}{dx^2} + 4$. Show that $A = B$ and that $C = 0$ over the linear manifold of all linear combinations of $f_1 = \sin 2x$ and $f_2 = \cos 2x$.[2]

## 1.3 Inner Products and Norms

Many problems in quantum mechanics need the concept of an *inner product.* Generally, the inner product $(f,g)$ is defined by

$$(f,g) = \int_{-\infty}^{\infty} f^* g\,dx. \tag{1.6}$$

This is frequently modified in quantum mechanics, depending on the problem, most often by limiting the bounds of integration to the applicable area. Abstractly, the inner product $(f,g)$ is a complex number associated with a pair of functions $f$ and $g$ belonging to a linear manifold $M$ and having properties

$$(f,g) = (g,f)^*, \tag{1.7a}$$

$$(f,f) \geq 0 \text{ and } (f,f) = 0 \text{ iff } f = 0, \text{ and} \tag{1.7b}$$

$$(f, cg + dh) = c(f,g) + d(f,h). \tag{1.7c}$$

A linear manifold space for which an inner product is defined is called an *inner product space.*

The square-root of the inner product $(f,f)$ is called the *norm of $f$.* A function whose norm equals one is said to be *normalized.* When working with the Schrödinger Wave Equation, it is common to set the arbitrary constant equal to the value which normalizes the wavefunction.

---

[2] Ikenberry, p. 68, problem 6.2

Problem 1.3: Show that[3]

(a) $(cf,g) = c^*(f,g)$

(b) $(cf+dg,c'f'+d'g')$

$$= c^*c'(f,f') + c^*d'(f,g') + d^*c'(g,f') + d^*d'(g,g')$$

(c) $(cf+dg,cf+dg) = c^*c(f,f) + c^*d(f,g) + cd^*(g,f) + dd^*(g,g)$

(d) $(f,g)+(g,f) =$ a real number

(e) $i(f,g) - i(g,f) =$ a real number

(f) $(f,g)(g,f) = |(f,g)|^2 = |(g,f)|^2 =$ a real number

*Problem 1.4:* Find a constant $c$ so that $(cf,cf) = 1$, where $f \neq 0$.[4]

Answer: $c = e^{i\delta}/\sqrt{(f,f)}$

*Problem 1.5:* Show that if $(f,Ag) = (f,Bg)$ for all $f$ and $g$ in a linear manifold $M$, then $A = B$ over $M$.[5]

## 1.4 Dirac's Bra and Ket Notation

Dirac developed a very convenient and frequently used notation, in which the generic symbol $\psi$ for a wave function is replaced by a *ket* symbol $|\ \rangle$, or is sometimes written $|\psi\rangle$. The symbol $\psi_n$ for the $n$'th wavefunction of an ordered set of wavefunctions is replaced by the symbol $|n\rangle$, or $|\psi_n\rangle$. The complex conjugate $\psi^*$ is replaced by the *bra* symbol $\langle\ |$, or $\langle\psi|$. Similar to above, the symbol $\psi_n$ is replaced by $\langle n|$, or $\langle\psi_n|$. Also, the operated wavefunction $\hat{A}\psi_n$ becomes $\hat{A}|n\rangle$, or $\hat{A}|\psi_n\rangle$.

Obviously, the *bra-ket* notation for an inner product is

$$(\psi_m,\hat{A}\psi_n) = \langle m|\hat{A}|n\rangle = \langle\psi_m|\hat{A}|\psi_n\rangle. \tag{1.8}$$

---

[3] Ikenberry, p. 68, problem 6.3

[4] Ikenberry, p. 69, problem 6.4

[5] Ikenberry, p. 69, problem 6.6

This particular inner product is defined to be the matrix element $\hat{A}_{mn}$ of the operator $A$ between the states $m$ and $n$. The elements $A_{nn}$ are called the diagonal matrix elements of $A$. $A$ is said to be in diagonal form if $A_{mn} = 0$ for $m \neq n$. Matrix elements will play an important role in Heisenberg's formulation of quantum mechanics, and in many calculation techniques in Schrödinger's formulation.

If $a$ is any complex number, then the following rules hold:

(1.9a) $$\langle \psi | a\phi \rangle = a \langle \psi | \phi \rangle,$$

(1.9b) $$\langle a\psi | \phi \rangle = a^* \langle \psi | \phi \rangle,$$

(1.9c) $$\langle \psi | \phi \rangle^* = \langle \phi | \psi \rangle,$$

(1.9d) $$\langle \phi + \psi | = \langle \phi | + \langle \psi |, \text{ and}$$

(1.9e) $$| \phi + \psi \rangle = | \phi \rangle + | \psi \rangle.$$

## 1.5 Expectation Values

In classical mechanics, the state of a system of particles is described at any instant in time by giving the position and velocity of each particle in the system. Moreover, the classical observables can be measured to any desired degree of precision.

According to the basic postulates of quantum mechanics, all of the information that can be obtained about the quantum mechanical state of the system is contained in the specification of a wavefunction $\psi$. From this function can be calculated, for example, the expectation values of observables and the expected uncertainties in their measurement.

The *expectation value* $\langle A \rangle$ of an observable $A$, when the system is in the state specified by $\psi$, is:

(1.10) $$\langle A \rangle = \frac{\langle \psi | A | \psi \rangle}{\langle \psi | \psi \rangle}.$$

If $\psi$ is normalized, this simplifies to:

(1.11) $$\langle A \rangle = \langle \psi | A | \psi \rangle.$$

The *expectation value of the uncertainty* $\Delta A$, in a measurement of an observable $A$ is:

(1.12) $$\Delta A = \sqrt{\left\langle \left( A - \langle A \rangle \right)^2 \right\rangle},$$

or in another form:

$$(\Delta A)^2 = \left\langle A^2 \right\rangle - \langle A \rangle^2. \tag{1.13}$$

*Problem 1.6:* Obtain (1.13) from (1.12).[6]

## 1.6 Hermitean Operators

We know that the only possible results of measurements of real observables, such as position coordinates and components of momentum, must be real numbers. Hence, the expectation value of a real observable $A$ must be a real number. Since a number $z$ is real if and only if $z = z^*$, we readily see that $\langle A \rangle$ is real if and only if $\langle A\psi | \psi \rangle = \langle \psi | A\psi \rangle$. This equality imposes a stringent condition on the operator $A$, and leads to the *Hermitean Property* of linear operators. There are two definitions commonly used, and it is easy to show that they imply each other.

*First Definition of Hermiticity:* A linear operator $A$ is said to be Hermitean over an inner product space $M$ if and only if

$$\langle Af | f \rangle = \langle f | Af \rangle \tag{1.14}$$

for every $f$ in $M$.

*Second Definition of Hermiticity:* A linear operator $A$ is said to be Hermitean over an inner product space $M$ if and only if

$$\langle Af | g \rangle = \langle f | Ag \rangle \tag{1.15}$$

for every $f$ and $g$ in $M$.

The second definition implies the first by letting $g = f$, the first definition implies the second by replacing $f$ by $f + cg$. Since they are equivalent, both definitions are equally strong.

*Problem 1.7:* Show that the operator $A = c$ is Hermitean if and only if $c$ is real.[7]

*Problem 1.8:* Let $A$ be Hermitean. Show that $cA$ is Hermitean if and only if $c$ is real.[8]

---

[6] Ikenberry, p. 70, problem 6.10

[7] Ikenberry, p. 71, problem 6.11

[8] Ikenberry, p. 71, problem 6.12

*Problem 1.9:* Show that $A^n$ is Hermitean if $A$ is Hermitean.[9]

*Problem 1.10:* Let $A$ and $B$ be Hermitean. Show that $A^m B^n + B^n A^m$ is Hermitean. In particular, $AB + BA$ is Hermitean.[10]

*Problem 1.11:* Prove that if $Af = cf$ and $Ag = dg$ for certain $f$ and $g$ with finite norms, and certain constants $c$ and $d$, where $A$ is Hermitean and $c \neq d$, then $f$ and $g$ are orthogonal. Show also that $\langle f | Ag \rangle = 0$ and $\langle g | Af \rangle = 0$.[11]

*Problem 1.12:* Let a wave function $\psi$ with finite norm satisfy the equation $i\hbar \frac{\partial}{\partial t} \psi = H\psi$. Show that, if $H$ is Hermitean, then $\frac{\partial}{\partial t} \langle \psi | \psi \rangle = 0$.[12]

*Problem 1.13:* A projection operator $P$ which projects functions $f$ onto a given function $g$, which has unit norm, may be defined by $Pf = \langle g | f \rangle g$. Show that $P$ is Hermitean, and that $P^2 = P$.[13]

## 1.7 Anti-Hermitean Operators

There are certain operators in quantum mechanics for which the expectation value is always a purely imaginary number. A complex number $z$ is pure imaginary if and only if $z = -z^*$. From this we readily see that the expectation value of the operator $A$ is purely imaginary if and only if $\langle \psi | A\psi \rangle = -\langle A\psi | \psi \rangle$. This gives us two definitions as in the previous section.

*First Definition of Anti-Hermiticity:* A linear operator $A$ , $A \neq 0$, is said to be Anti-Hermitean over an inner product space $M$ if and only if, for every $f$ in $M$,

$$\langle f | Af \rangle = -\langle Af | f \rangle. \tag{1.16}$$

---

[9] Ikenberry, p. 71, problem 6.13

[10] Ikenberry, p. 71, problem 6.14

[11] Ikenberry, p. 71, problem 6.17

[12] Ikenberry, p. 71, problem 6.18

[13] Ikenberry, p. 71, problem 6.20

*Second Definition of Anti-Hermiticity:* A linear operator $A$ , $A \neq 0$, is said to be Anti-Hermitean over an inner product space $M$ if and only if

$$\langle f | Ag \rangle = -\langle Af | g \rangle \tag{1.17}$$

for every $f$ and $g$ in $M$.

*Problem 1.14:* Prove the equivalence of the two definitions of anti-Hermiticity.[14]

*Problem 1.15:* Show that the operator $A = c$ is anti-Hermitean if and only if $c$ is pure imaginary.[15]

*Problem 1.16:* Let $A$ be Hermitean, $A \neq 0$, and $B$ be anti-Hermitean. Show that $iA$ is anti-Hermitean, and that $iB$ is Hermitean.[16]

*Problem 1.17:* Let $A = \frac{d}{dx}$ and $B = i\frac{d}{dx}$ over the linear manifold of square-integrable wave functions. Show that $A$ is anti-Hermitean and $B$ is Hermitean.[17]

## 1.8 Adjoint Operators

A general theorem from the theory of linear operators states that if $A$ is any linear operator over a linear manifold $M$, there exists a linear operator $B$ over $M$ such that $\langle f | Ag \rangle = \langle Bf | g \rangle$ for all $f$ and $g$ in $M$. We call $B$ the *adjoint* of $A$, and denote it by the symbol $A^{\dagger}$. Notice that if $A$ is Hermitean, $A^{\dagger} = A$; while if $A$ is anti-Hermitean, $A^{\dagger} = -A$. This gives us useful working definitions of Hermiticity and Anti-Hermiticity. As with Hermiticity and Anti-Hermiticity, we have two equivalent definitions.

*First Definition of $A^{\dagger}$:* $A^{\dagger}$ is called the adjoint of $A$ over an inner product space $M$ if and only if

$$\langle A^{\dagger} f | f \rangle = \langle f | Af \rangle \tag{1.18}$$

for every $f$ in $M$.

---

[14] Ikenberry, p. 72, problem 6.21

[15] Ikenberry, p. 72, problem 6.22

[16] Ikenberry, p. 72, problem 6.23 and 6.24

[17] Ikenberry, p. 72, problem 6.25

*Second Definition of* $A^\dagger$*:* $A^\dagger$ is called the adjoint of $A$ over an inner product space $M$ if and only if

(1.19) $$\left\langle A^\dagger f \middle| g \right\rangle = \langle f | Ag \rangle$$

for every $f$ and $g$ in $M$.

Some useful properties of the adjoint follow. It is left to the reader to prove any that are not immediately obvious.

(1.19) $$A = c \text{ if and only if } A^\dagger = c^*,$$

(1.20) $$\left(A^\dagger\right)^\dagger = A,$$

(1.21) $$(AB)^\dagger = B^\dagger A^\dagger, \text{ and}$$

(1.22) $$L_+ \equiv A + iB \text{ and } L_- \equiv A - iB \text{ are adjoint operators.}$$

## 1.9 Commutators

Two linear operators $A$ and $B$ over a linear manifold $M$ are said to *commute* if

(1.23) $$A(Bf) = B(Af)$$

for every $f$ in $M$. This is generally written $AB = BA$ for brevity. Two linear operators $A$ and $B$ over a linear manifold $M$ are said to *anti-commute* if

(1.24) $$A(Bf) = -B(Af)$$

for every $f$ in $M$. This is generally written $AB = -BA$ for brevity.

Due to the importance of the commutation relation, we associate an operator with the concept. This operator is called the *commutator of A and B*, and is defined by

(1.25) $$[A,B] \equiv AB - BA$$

over $M$. $A$ and $B$ commute if $[A,B] = 0$, and they anti-commute if $[A,B] = 2AB$.

The basic properties of the commutator are as follows. It is left up to the reader to prove any that are not immediately obvious.

(1.26) $$[A,B] = -[B,A],$$

(1.27) $$[A,B+C] = [A,B] + [A,C],$$

(1.28) $$[A,BC] = B[A,C] + [A,B]C, \text{ and}$$

(1.29) $$[AB,C] = A[B,C] + [A,C]B.$$

*Problem 1.18:* Show that if A and B are noncommuting Hermitean operators, then $\left[A,B\right]$ is anti-Hermitean and $i\left[A,B\right]$ is Hermitean.[18]

*Problem 1.19:* Prove that the product of two Hermitean operators is Hermitean if and only if they commute.[19]

---

[18] Ikenberry, p. 74, problem 6.37

[19] Ikenberry, p. 74, problem 6.38

# Chapter 2

# Eigenfunctions and Eigenvalues

*This chapter looks at the usefulness of Eigenvalue problem-solving techniques, and their applications to operator-observable pairs in quantum mechanics. A central problem in quantum mechanics is the finding of eigenvalues and eigenfunctions of linear operators, usually associated with real observables. Of particular importance are the energy and angular momentum eigenvalues and eigenfunctions.*

## 2.1 The Eigenvalue Problem

If $A$ is an operator defined over a linear manifold $M$, and $f$ is any particular function in $M$, such that there exists a number $a$ such that

$$\hat{A}f = a \cdot f, \tag{2.1}$$

then $f$ is called the *eigenfunction of* $A$ and $a$ is called the *eigenvalue of* $A$. If the linear manifold M is also an inner product space, then

$$(f, Af) = a(f, f), \text{ or} \tag{2.2}$$

$$\langle f|A|f\rangle = a\langle f|f\rangle. \tag{2.3}$$

If $A$ is Hermitean, then it has only real eigenvalues, and the eigenfunctions belonging to different eigenvalues are orthogonal (see problem 1.12); further $\Delta A = 0$. It is also true that if $A$ is a Hermitean operator, and $\Delta A = 0$ for some $f$, then $f$ is an eigenfunction of $A$.

*Problem 2.1:* Let M be the linear manifold of functions $f$ which are bounded and differentiable for all $x$. Find the eigenfunctions of $p = -i\hbar\frac{d}{dx}$.[20]

Answer: Set of all $f = c\,e^{iax/\hbar}$, $a$ real.

---

[20] Ikenberry, p. 112, problem 10.1

*Problem 2.2:* Let $f$ be an eigenfunction of a linear operator $A$ for $A = a$. Show that for any $c$, $f' = cf$ is also an eigenfunction of $A = a$. Note that for any inner product space, $c$ may be chosen so that $\langle f' | f' \rangle = 1$.[21]

*Problem 2.3:* Let $f$ be an eigenfunction of $A$ for $A = a$, in an inner product space. Show that, for any polynomial $f$, $\langle f(A) \rangle = f(a)$. In particular, $\langle A^n \rangle = a^n$.[22]

## 2.2 Linear Dependence and Independence

Recall that a set of $k$ functions $f_i$ is said to be *linearly dependent* if there exist constants $c_i$, not all zero, such that

$$\sum_{i=1}^{k} c_i f_i = 0. \tag{2.4}$$

A set of $k$ functions $f_i$ is said to be *linearly independent* if there exist no constants $c_i$, except $c_i = 0$, such that (2.4) holds. Stated differently, $k$ functions $f_i$ are linearly independent if (2.4) implies that all $c_i = 0$; they are linearly dependent if (2.4) does not necessarily imply that all $c_i = 0$.

Frequently there exist more than one linear independent eigenfunction for a single eigenvalue. That is, there exist at least two functions $f$ and $g$ such that, for some $a$, $Af = af$ and $Ag = ag$, but $f \neq cg$. This is called *degeneracy*. We say that an eigenvalue with $k > 1$ linearly independent eigenfunctions is *k-fold degenerate*. A non-degenerate eigenvalue is said to be a *simple eigenvalue*.

*Problem 2.4:* Taking $M$ and $p$ as in problem 2.1, show that the non-zero eigenvalues of $p^2$ are doubly degenerate and that $c = 0$ is a simple eigenvalue.[23]

---

[21] Ikenberry, p. 113, problem 10.2

[22] Ikenberry, p. 113, problem 10.3

[23] Ikenberry, p. 114, problem 10.8

*Problem 2.5:* Prove that any linear combination of eigenfunctions of an operator $A$ for $A = a$ is also an eigenvalue of $A$ for $A = a$.[24]

*Problem 2.6:* Show that any $k$ mutually orthogonal functions are linearly independent.[25]

*Problem 2.7:* Show that any $k$ linearly dependent functions in an inner product space are not mutually orthogonal.[26]

## 2.3 The Expansion Problem

We can now determine a method for determining the wavefunction $\psi(\mathbf{r},t)$ at any time $t$, given the wavefunction $\psi(\mathbf{r})$ at time $t = 0$. The orthogonality of the Hamiltonian operator for a system, together with a property known as completeness, and Gram-Schmidt Orthogonalization (which provides a method to find $k$ mutually orthogonal eigenfunctions belonging to a $k$-fold degenerate eigenvalue) give us all of the necessary mathematical techniques to rigorously prove the results stated in this section. The proof is tedious and has been omitted.

We find that, given $\psi(\mathbf{r})$ at time $t = 0$,

$$\psi(\mathbf{r},t) = \sum_{j=0}^{\infty} c_j \psi_j(\mathbf{r}) e^{-iE_j t/\hbar}, \text{ where} \tag{2.5}$$

$$c_j = \langle \psi_j | \psi \rangle. \tag{2.6}$$

As an example, recall the orthonormal energy eigenfunctions for particle in a one-dimensional box $0 \le x \le a$ are found to be, for $n \in \{1,2,3,\ldots\}$,

$$\psi_n = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi x}{a}\right). \tag{2.7}$$

Given $\psi(x) = 1$ in the box, at $t = 0$, we find

$$c_n = \langle \psi_n | \psi \rangle = \langle \psi_n | 1 \rangle, \tag{2.8}$$

---

[24] Ikenberry, p. 114, problem 10.9

[25] Ikenberry, p. 114, problem 10.10

[26] Ikenberry, p. 114, problem 10.11

which, written out, is

$$c_n = \int_0^a \psi_n{}^*(1)\,dx. \tag{2.9}$$

This is solved, and we get

$$c_n = \frac{-1}{n\pi}\sqrt{2a}\cos\left(\frac{n\pi x}{a}\right)\Bigg|_0^a, \tag{2.10}$$

or, simplified,

$$c_n = \begin{cases} 0 & \text{if } n \text{ even} \\ \dfrac{2\sqrt{2a}}{n\pi} & \text{if } n \text{ odd} \end{cases}. \tag{2.11}$$

This we plug into (2.5) and solve, to get

$$\psi_n(x,t) = \sum_{n=0}^{\infty} \frac{2\sqrt{2a}}{(2n+1)\pi}\sqrt{\frac{2}{a}}\sin\left(\frac{(2n+1)\pi x}{a}\right)e^{-iE_{(2n+1)}t/\hbar}, \tag{2.12}$$

or, simplified,

$$\psi_n(x,t) = \sum_{n=0}^{\infty} \frac{4}{(2n+1)\pi}\sin\left(\frac{(2n+1)\pi x}{a}\right)e^{-iE_{(2n+1)}t/\hbar}. \tag{2.13}$$

# Part II

# Elementary Principles

# Chapter 3

# Historical Review

*This chapter provides an outline of the problems in classical Newtonian mechanics that prompted the evolution of quantum mechanics during the first three decades of this century.*

## 3.1 Overview of Dates

Physics at the turn of the century was in a state of turmoil. People were discovering that there were experimental observations which, on the basis of firmly established classical Newtonian mechanics and Maxwellian Electrodynamics, were totally inexplicable. Over the first three decades of the century these questions were answered, and the shocking solutions led to a total revision of the philosophy of science. A whole new way of thinking was needed, for at the core of natural law lay subjective probability, not objective determinism as according to old classical mechanics.

The observations which prompted the development of quantum mechanics were threefold. The first dilemma to be solved deals with blackbody radiation. This is the radiation emitted by a cavity whose walls are maintained at a constant temperature. Theory based on the wave nature of light was unable to account for the observed frequency distribution of the radiant energy.

Second, light exhibits interference, and so may be safely assumed to be a wave phenomenon. However, in examining the photoelectric effect (light hitting metal surfaces ejects electrons), it was determined that the energy of the ejected electrons is dependent only on the frequency of the incident radiation, not the intensity as should be expected by classical light theory.

The third problem involved an experiment by Rutherford in 1911. He established that an atom has a positive central core, and satellite electrons. However, an orbiting electron is accelerating, and hence must radiate energy. Due to this loss of energy an electron should spiral into the nucleus, giving off a burst of ultraviolet light as the electron and nucleus annihilate. Instead of the continuous spectrum we should see from this, we see a discrete line spectrum. Also, there is still matter, even though classical mechanics predicts it should have all annihilated itself in $10^{-8}$ seconds after it was created.

Following is a list of the developments during the beginning of the century which removed the enigmas posed by these problems:

| | | |
|---|---|---|
| 1901 | Planck | Blackbody Radiation |
| 1905 | Einstein | Photoelectric Effect |
| 1913 | Bohr | Quantum Theory of Spectra |
| 1922 | Compton | Scattering photons off electrons |
| 1924 | Pauli | Exclusion Principle |
| 1925 | de Broglie | Matter Waves |
| 1926 | Schrödinger | Wave Equation |
| 1927 | Heisenberg | Uncertainty Principle |
| 1927 | Davisson and Germer | Experiment on Wave Properties of Electrons |
| 1927 | Born | Interpretation of the Wavefunction |

In the following chapters we will look at the evolution of quantum mechanics in more detail.

# Chapter 4

# Review of Concepts of Classical Mechanics

*This chapter is intended to review some of the fundamental concepts of classical mechanics that are important to the development of quantum mechanics. Coordinate systems, cyclic coordinates, Hamiltonian mechanics, and constants of the motion are discussed.*

## 4.1 Generalized Coordinates

A point particle constrained to move in one dimension is said to have *one degree of freedom*. This means that one variable only is needed to uniquely specify the location of the particle in space. This variable is normally the displacement of the particle from an arbitrarily specified origin within the dimension.

A point particle constrained to move in two dimensions (a flat plane) has *two degrees of freedom*. In this case, two variables are needed to uniquely specify the location of the particle. Generally one of two systems are used for this: Cartesian coordinates $(x,y)$, or Polar coordinates $(r,\theta)$.

Another system with two degrees of freedom is the system of two point particles in one dimension. This is generally represented by the coordinates $(x_1,x_2)$.

A point particle moving in free space (three-dimensions) has *three degrees of freedom*. Appropriate coordinates systems for the representation of this system are: Cartesian coordinates $(x,y,z)$, Cylindrical coordinates $(r,\theta,z)$, and Spherical coordinates $(r,\theta,\phi)$.

There are definitely other systems with three degrees of freedom. For instance, a rigid rod in two dimensions, which normally uses two dimensional Cartesian coordinates along with a coordinate $\theta$ which describes the angular displacement of the rod from the horizontal, or a system of three particles in one dimension which uses a similar coordinate system to that of two particles in one dimension, with a third coordinate added.

Independent coordinates that uniquely determine the orientation and position of a system in space are called *generalized* (or *canonical*) coordinates. It should also be evident from the previous examples that a system of $n$ generalized coordinates has $n$

degrees of freedom, and visa versa. The one further stipulation on the coordinates is that they are all independent variables.

## 4.2 Hamiltonian Mechanics

Energy, the Hamiltonian, and Angular Momentum are all very important elements of classical mechanics, and all have direct counterparts in quantum mechanics that are as important.

We begin our discussion of the first two of these three vital elements with a review of *constants of the motion*. Recall that a constant of the motion is a dynamical function that is constant throughout time. For instance, in an isolated system (one that does not interact with other objects in the universe outside the system), the energy function, linear momentum, and angular momentum are constants of the motion; as is described by the *law of conservation of energy*, the *law of conservation of momentum*, and the *law of conservation of angular momentum.*

As an example, consider the hydrogen atom in free space. If we consider it as two point particles orbiting each other in the absence of a magnetic field, we discover the following: it has six degrees of freedom. If $(x_1, y_1, z_1)$ are the coordinates of the proton and $(x_2, y_2, z_2)$ are the coordinates of the electron, then the energy of the hydrogen atom appear as

$$E = \tfrac{1}{2}M(\dot{x}_1^2 + \dot{y}_1^2 + \dot{z}_1^2) + \tfrac{1}{2}m(\dot{x}_2^2 + \dot{y}_2^2 + \dot{z}_2^2) - \frac{q^2}{d_{12}} \tag{4.1}$$

where $M$ is the mass of the proton, $m$ is the mass of the electron, $q$ is the electric charge, and $d_{12}$ is the distance between the proton and electron (classically). Regardless of the time at which the energy is measured, it will be a constant.

The constants of the motion for complicated systems are not easily found. However, *Hamiltonian formalism* treats this problem directly and simply. Consider, for example, the energy expression for an electron between two capacitor plates:

$$E = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + \frac{q\Phi_0}{d}z \tag{4.2}$$

where we have an electron of mass $m$ and charge $q$, between two capacitor plates maintained at the potential difference $\Phi_0$ and separated by the distance $d$. The electron is a distance $z$ above the bottom plate. Rewriting this expression in terms of the linear momentum instead of velocity by means of

$$E(x, y, z, \dot{x}, \dot{y}, \dot{z}) \leftrightarrow H(x, y, z, p_x, p_y, p_z) \tag{4.3}$$

gives us

(4.4) $$H = \frac{1}{2m}(p_x{}^2 + p_y{}^2 + p_z{}^2) + \frac{q\Phi_0}{d} z$$

The energy, written as a function of the coordinates and momenta is called the Hamiltonian. One says $p_x$ is the momentum conjugate to $x$.

The equations of motion that replace Newton's second law in Hamiltonian theory are called *Hamilton's equations* and are (for a point particle moving in three-dimensional space):

(4.5a,b) $$\frac{\partial H}{\partial x} = -\dot{p}_x \qquad \frac{\partial H}{\partial p_x} = \dot{x}$$

(4.5c,d) $$\frac{\partial H}{\partial y} = -\dot{p}_y \qquad \frac{\partial H}{\partial p_y} = \dot{y}$$

(4.5e,f) $$\frac{\partial H}{\partial z} = -\dot{p}_z \qquad \frac{\partial H}{\partial p_z} = \dot{z}$$

Now, obviously, for the Hamiltonian corresponding to an electron between capacitor plates, one obtains:

(4.6) $$\frac{\partial H}{\partial x} = \frac{\partial H}{\partial y} = 0$$

This is because the expression for the Hamiltonian contains neither $x$ nor $y$. When coordinates are missing from the Hamiltonian like this, they are called *cyclic* or *ignorable*. The momentum conjugate of a cyclic coordinate is a constant of the motion. This property follows directly from Hamilton's Equations. In our example, we clearly see that $\partial H/\partial x = 0$ implies that $\dot{p}_x = 0$, so $p_x$ is constant. This also follows for $p_y$. The other four equations give

(4.7) $$\dot{p}_z = -\frac{q\Phi_0}{d} \qquad p_x = m\dot{x} \qquad p_y = m\dot{y} \qquad p_z = m\dot{z}.$$

The last three equations give us no new information, as they are just the definitions of the momenta in terms of velocity, but the first equation is the $z$ component of Newton's Second Law.

As a second example, consider a particle of mass $m$ in a constant gravitational field. We know that the Hamiltonian in this case is $H = \frac{1}{2m}(p_x{}^2 + p_y{}^2 + p_z{}^2) + mgz$.

From this we readily obtain Hamilton's equations:

$$\dot{p}_x = -\frac{\partial H}{\partial x} = 0 \qquad \dot{p}_y = -\frac{\partial H}{\partial y} = 0 \qquad \dot{p}_y = -\frac{\partial H}{\partial y} = -mg, \tag{4.8}$$

and

$$\dot{x} = \frac{1}{m} p_x \qquad \dot{y} = \frac{1}{m} p_y \qquad \dot{z} = \frac{1}{m} p_z\,. \tag{4.9}$$

The first set of equations yields the following:

$$p_x = p_x(0) \qquad p_y = p_y(0) \qquad F = mg, \tag{4.10}$$

while the second set gives:

$$mv_x = p_x \qquad mv_y = p_y \qquad mv_z = p_z. \tag{4.11}$$

This could also be written in cylindrical or spherical coordinates; we will now do the later. The Hamiltonian becomes:

$$H = \frac{1}{2m} p_r^2 + \frac{1}{2m}\frac{1}{r^2} p_\theta^2 + \frac{1}{2m}\frac{1}{r^2 \sin^2\theta} p_\theta^2 + mgr\cos\theta. \tag{4.12}$$

Exploring Hamilton's equations in spherical coordinates yields:

$$\dot{p}_r = \underbrace{\frac{1}{m}\frac{p_\theta^2}{r^3} + \frac{1}{m}\frac{p_\phi^2}{r^3 \sin^2\theta}}_{\text{Centripetal Force}} - \underbrace{mg\cos\theta}_{\substack{\text{Component} \\ \text{of gravity in} \\ \text{radial direction}}} \tag{4.13a}$$

$$\dot{p}_\theta = \underbrace{\frac{1}{m}\frac{p_\theta^2 \cos\theta}{r^2 \sin^3\theta}}_{\substack{\text{Torque from} \\ \text{Centripetal Force}}} - \underbrace{mgr\sin\theta}_{\substack{\text{Torque from} \\ \text{Gravity}}} \tag{4.13b}$$

$$\dot{p}_\phi = 0 \text{ so } p_\phi = \text{constant}\,. \tag{4.13c}$$

*Problem 4.1:* A particle of mass m is attracted to the origin by the force $\vec{F} = -K\vec{r}$. Write the Hamiltonian for this system in spherical and Cartesian coordinates. What are the cyclic coordinates in each of these frames?[28]

---

[28] Liboff, p. 18, problem 1.11

*Problem 4.2:* A particle of mass m is in the environment of a force field with components

$$\vec{F}_z = -K\vec{z} \qquad F_x = 0 \qquad F_y = 0$$

with K constant. (a) Write down the Hamiltonian of the particle in Cartesian coordinates. What are the constants of the motion? (b) Use the fact that the Hamiltonian itself is also constant to obtain the orbit. (c) What is the Hamiltonian in cylindrical coordinates? What are the constants of the motion?[29]

---

[29] Liboff, p. 18, problem 1.9

# Chapter 5

# Semi-Classical Quantum Mechanics

*This chapter looks at the closely at the problems that are unsolvable by classical Newtonian mechanics, and the inspired solutions to those problems. This is the beginning of Quantum Theory, from which it has evolved into a complex field of its own.*

## 5.1 Blackbody Radiation

The first of the three problems with classical mechanics to be solved was the problem with blackbody radiation. The problem was observed when a closed, evacuated container (with a small window in one side) was placed in an oven and held at a uniform temperature. At sufficiently high temperatures, visible light emerges from the window of the container. Other electromagnetic radiation is also being emitted, and blackbody radiation theory must account for how the radiation is distributed throughout the various frequencies. When the container is held at a fixed temperature, the inner walls emit and absorb photons at the same rate. Hence, the cavity contains radiant energy, which is in thermal equilibrium with the cavity walls. Classical theory correctly predicted the radiation for low frequencies, but at higher frequencies the experimental values rapidly diverged from the theoretical values. This is called the *UV Catastrophe*. By analyzing the radiation escaping through the window, experimenters found that:

a) the energy depends only on the frequency and temperature of the enclosure, and is not affected by the shape of the enclosure or the material its made from or contains,
b) for low frequencies, the energy is proportional to the frequency squared,
c) for high frequencies, the energy decreases exponentially as the frequency increases,
d) the total amount of radiation energy per unit volume is proportional to the temperature to the fourth power (the Stephen-Boltzmann Law),
e) The frequency at which the energy is maximal is proportional to the absolute temperature (Wein's Displacement Law),

f) the energy density function may be written in the form of Wein's Relation:

$$E_\nu = \frac{\nu^3}{c^2} f\left(\frac{\nu}{T}\right). \tag{5.1}$$

Many people tried to discover a classical solution that fit this data, but none could be found. Maxwell Planck took a wild shot, and assumed that the energy was *quantized* (has only discrete values). Remarkably, this assumption—that energy of radiation with frequency $\nu$ exists only in multiples of Planck's constant $h$—modified existing theory enough to make experimental results match prediction. ***Planck's constant*** is $h = 6.626x10^{-27}$ erg-s, and a quantum of radiation of energy $h\nu$ is called a ***photon***. (Recall that $\hbar = h/2\pi$.) Now, if we suppose that the total radiation energy per unit volume in the cavity is $U$, and $u(\nu)\,d\nu$ is the energy with frequency between $\nu$ and $\nu + d\nu$, then clearly $U = \int_0^\infty u(\nu)\,d\nu$. The correct formula for $u(\nu)$ which results from the assumption that $E = h\nu$ is:

$$u(\nu) = \frac{8\pi h\nu^3}{c^3}\left(e^{h\nu/K_bT} - 1\right)^{-1}. \tag{5.2}$$

This value precisely matches experimental evidence. Quantization of energy marked the end of classical Newtonian mechanics, and ushered in what we now call the old semi-classical quantum mechanics. Newtonian mechanics was still valid in its limited domain: large objects moving at reasonable speeds, but when systems got small, this semi-classical quantum mechanics needed to be invoked.

## 5.2 The Photoelectric Effect

The photoelectric effect is a phenomena in which radiative energy is transferred to bound electrons, liberating them from atoms in a metallic surface. According to classical theory

a) the velocity of the electrons liberated should increase as the intensity of the light increased,

b) a measurable time should elapse after the metal plate is illuminated before a current is detected, and

c) the effect should be observed with light of any frequency.

For several years after the discovery of the photoelectric effect it was assumed that these predictions were valid. In 1902 Lenard conclusively demonstrated that

a) the velocity of the electrons does not depend on the intensity,

b) the time elapsed before the effect is observed is immeasurable, and

c) below a certain frequency (called the threshold frequency) no electrons are emitted.

He further demonstrated that the velocity of the electrons increases as the frequency of the light is increased, and that the current increases with increased intensity. The wave theory of light provides us with no explanation of these observed characteristics.

With the introduction of Planck's quantized energy though, and Albert Einstein's realization that radiation has a dual nature everywhere—part wavelike and part corpuscular (Planck believed that radiation is continuous except in the immediate vicinity of matter)—Einstein could easily explain the photoelectric effect. It is for this work that Einstein received the Nobel Prize in 1921, not his theories of relativity as many people believe.

Einstein's arguments went as follows: Let $W$ be the energy required to free an electron from a metallic surface (the work function). Then the energy $h\nu - W$ is available, when a photon "hits" the electron, to give it kinetic energy $\frac{1}{2}mv^2$, where $m$ is the mass of the electron, and $v$ is its speed. We then have

$$\frac{1}{2}mv^2 = h\nu - W. \tag{5.3}$$

Electrons for which $W$ is smallest are liberated with the greatest speed:

$$E_{\max} = \frac{1}{2}m v_{\max}{}^2 = h\nu - W_{\min} = h(\nu - \nu_0) \tag{5.4}$$

where $\nu_0 = W_{\min}/h$ is the threshold frequency.

## 5.3 The Compton Effect

In the photoelectric effect, a photon gives all of its energy to a bound electron. It is also possible that only a part of the energy of the photon is transferred to the electron, resulting in a scattering of the photon. The quantum theory of this scattering of electromagnetic waves is known as the *Compton Effect.*

According to classical theory, when a monochromatic electromagnetic wave hits a charged particle whose radius is much smaller than the wavelength of the radiation, the charged particle will respond to this changing electric force by oscillating in simple harmonic motion at the frequency of the radiation (a driven oscillator problem). Since the charge is accelerated continuously, it will produce electromagnetic radiation of the same frequency in all directions. This is compatible with experimental evidence for wavelengths of visible light and all other longer wavelengths of radiation. However, the experimental evidence for x-rays is not compatible for this theory; it breaks down as the size of the particle gets closer to the wavelength.

In 1922 Arthur Compton used the particle-like, quantum nature of electromagnetic radiation to explain the scattering of x-rays. Because a photon may be regarded as a zero-rest-mass particle with speed $c$, the magnitude of the corresponding momentum is given by

(5.5) $$p = \frac{E}{c} = \frac{h\nu}{c} = \frac{h}{\lambda}.$$

The direction of $p$ is along the direction of propagation of the incident wave.

The quantization of energy thus leads to the quantization of momentum in integral multiples of the momentum of a single photon. From the above equation we see that a photon's momentum increases with frequency, just as its energy increases with frequency. Thus the momentum of a high-frequency, high-energy photon (such as a $\gamma$-ray) will far exceed the momentum of a low-frequency, low-energy photon (such as a radio photon, or even a visible light photon).

By all of this we can consider the scattering problem as a collision between a particle-like photon and a charged particle. Then the problem is solved merely by applying the laws of energy and momentum conservation. This needs relativistic tools because the resulting momentum of the charged particle may be extremely large, so we will go through the derivation here, as a review of relativistic principles.

We take the particle, of rest mass $m_0$ and rest energy $E_0 = m_0 c^2$, to be free and initially at rest. Then applying energy conservation to the collision gives

(5.6) $$h\nu + E_0 = h\nu' + E$$

where $E$ is the energy of the recoiling particle after the collision, $h\nu$ is the energy of the incident photons, and $h\nu'$ is the energy of the scattered photons. Since

(5.7) $$E = \gamma E_0 > E_0$$

we see immediately that $h\nu' < h\nu$. Consequently the scattered photon has less energy, a lower frequency, and a longer wavelength then the incident photon. Remember to avoid thinking that the scattered photon is merely the incident photon moving in a different direction. Instead the incident photon is annihilated, and the scattered photon is created.

Turning now to the law of conservation of momentum, we see that

(5.8) $$p = \gamma m_0 v$$

is the relativistic momentum of the recoiling particle,

(5.9) $$p_\lambda = \frac{h\nu}{c} = \frac{h}{\lambda}$$

is the momentum of the incident photon, and

(5.10) $$p_{\lambda'} = \frac{h\nu'}{c} = \frac{h}{\lambda'}$$

is the momentum of the scattered photon. The scattering angle $\theta$ is the angle between the directions of $p_\lambda$ and $p_{\lambda'}$.

Using the law of cosines leads to the basic equation of the Compton effect:

(5.11) $$\Delta\lambda = \lambda' - \lambda = \frac{h}{m_0 c}(1 - \cos\theta).$$

This gives the increase in the wavelength of the scattered photon. Although this gives the the wavelength increase unambiguously, we cannot predict in advance the angle any one photon will emerge.

For visible and other long-wavelength electromagnetic radiation, $\Delta\lambda$ is so small as to be masked by the thermal motion of the electrons. However, shorter wavelength radiation easily demonstrates this effect. Clear in any data though, is the existence of both types of interactions: the classical and the Compton scattering, implying that photons have both wave-like and corpuscular natures simultaneously.

## 5.4 de Broglie Matter Waves and the Davisson-Germer Experiment

Having just seen that light has a dualistic nature, we will explore the dualistic nature of matter, and then look at the hydrogen atom.

In 1924, in his doctoral thesis, Louis de Broglie observed that on the left-hand sides of the equations

(5.12) $$E = h\nu, \qquad p = \frac{h}{\lambda}$$

appear properties associated with particles, while on the right-hand side of the equations appear properties associated with waves. He reasoned that if these relations are valid, as the explanation of the photoelectric and Compton effects indicate, then it is impossible to obtain a purely wave or purely corpuscular theory of light. This is due to the fact that if any such theory existed, there could be no relationship equating wave and particle properties. Further, if the particle and wave aspects of light are inseparable, it might be that the particle and wave concepts of matter are inseparable also. That is, in order to explain some aspects of matter, it may be necessary to consider matter as a wave.

When discussing his work, de Broglie made the following remarks: "When I began to consider these difficulties I was chiefly struck by two facts. On the one hand the Quantum Theory of Light cannot be considered satisfactory, since it defines the energy of a light corpuscle by the equation $E = h\nu$ containing the frequency $\nu$. Now a purely corpuscular theory contains nothing that enables us to define a frequency; for this reason alone, therefore, we are compelled in the case of Light, to introduce the idea of a corpuscle and that of periodicity simultaneously.

"On the other hand, determination of the stable motion of electrons in the atom introduces integers; and up to this point the only phenomena involving integers in Physics were those of interference and of normal modes of vibration. This fact suggested to me the idea that electrons too could not be regarded simply as corpuscles, but that periodicity must be assigned to them.

"In this way, then, I obtained the following idea, in accordance with which I pursued my investigations: that it is necessary in the case of Matter, as well as of radiation generally and of Light in particular, to introduce the idea of the corpuscle and of the wave simultaneously; or in other words, in the one case as in the other, we must assume the existence of corpuscles accompanied by waves. But corpuscles and waves cannot be independent of each other; in Bohr's terms, they are two complementary aspects of reality; and it must consequently be possible to establish a certain parallelism between the motion of a corpuscle and the propagation of its associated wave. The first object at which to aim, therefore, was to establish the existence of this parallelism."[30]

Soon after de Broglie's dissertation appeared, Elsasser pointed out that, if de Broglie's ideas were correct, particles such as electrons should exhibit diffraction effects. Later that year Davisson and Germer succeeded in observing a diffraction pattern in low energy electrons reflected from a nickel crystal. This was the proof needed that matter, as well as light, has a dualistic nature.

## 5.5 Bohr's Quantum Theory of Spectra

We now return to 1913, when Niels Bohr began investigating another of the phenomena that classical mechanics could not explain. Most likely you have seen this phenomena yourself in the lab. When a closed tube is filled with gas, and a voltage is applied through the tube, the gas glows. If the gas is then examined with a spectroscope or optical grating, it is seen that only a discrete set of frequencies are emitted. Hence we have a line spectrum, as opposed to the continuous spectrum that classical mechanics predicts. Bohr was able to account for the discrete spectra in Hydrogen making two assumptions:

---

[30] Ikenberry, p. 27-28.

(1) Hydrogen exists in discrete values of the angular momentum as given by the relation

$$\oint p_\theta \, d\theta = nh \tag{5.13}$$

with $n$ an integer greater than zero. In these states the atom does not radiate. The line integral is performed on one complete orbit of the electron about the nucleus.

(2) When an atom undergoes a change in energy levels, from $E_n$ to $E_m$, electromagnetic radiation (a photon) is emitted at a frequency $\nu$ given by

$$h\nu = |E_n - E_m|. \tag{5.14}$$

Recall that condition (1) leads to a discrete set of energies of:

$$E_n = -\frac{\mathfrak{R}}{n^2} \tag{5.15}$$

where $\mathfrak{R}$ is the Rydberg constant:

$$\mathfrak{R} = \frac{me^4}{2\hbar^2} = 2.18x10^{-11} \text{ erg} = 13.6 \text{ eV} \tag{5.16}$$

Recall also that the calculated radius of the orbiting electron according to Bohr's theory is

$$\mathrm{r_n} = \frac{n^2\hbar^2}{me^2} \tag{5.17}$$

The negative quality of the energy reflects the fact that we are dealing with *bound states*. When $n = 1$, the atom is in the *ground state* and has energy $-\mathfrak{R}$. To ionize the atom when it is in this state takes $+\mathfrak{R}$ ergs of energy. The value of $r$ when the atom is in the ground state is

$$r_1 \equiv \frac{\hbar^2}{me^2} = 0.529\mathring{A}. \tag{5.18}$$

This is the fundamental length in physics and is called the *Bohr radius.*

The emission spectra of hydrogen is generated by the values for $E_n$ in the second assumption. The frequencies generated (with some minor refinements, eg: accounting for the motion of the proton) agree to a high degree of accuracy with the data. Characteristically, the spectrum divides into various series of lines: Lyman, Balmer, Paschen, etc. These are generated respectively by transitions to the ground state, to the second excited state, to the third excited state, etc.

### 5.6 Wilson-Sommerfeld Quantum Conditions

Wilson and Sommerfeld independently generalized the first assumption of Bohr's quantization rule (5.13). Their generalization is now called the Wilson-Sommerfeld quantization rule. Any periodic motion is quantized in such a way that the action integral

$$\oint p_i \, dq_i \tag{5.19}$$

is equal to an integral multiple of $h$. That is,

$$\oint p_i \, dq_i = nh \tag{5.20}$$

where $n$ is an integer.

As an application of this, consider a particle bouncing in a gravitational field of strength $g$, off a level and perfectly elastic floor. Given that:

$$E = \frac{p^2}{2m} + mgz, \tag{5.21}$$

we know that this can be rewritten in the form:

$$E = \frac{p_z{}^2}{2m} + mgz, \tag{5.22}$$

because the only motion is in the $\hat{z}$ direction. Now, solving this for $p_z$, we get:

$$p_z = \sqrt{2mE - 2m^2 gz}\,. \tag{5.23}$$

Using the Wilson-Sommerfeld rule we know that

$$nh = \oint p_z \, dz, \tag{5.24}$$

or, substituting in for $p_z$,

$$nh = \int_0^{h_0} \sqrt{2mE - 2m^2 gz} \, dz. \tag{5.25}$$

Making the substitution $u = 2mE - 2m^2 gz$ lets us rewrite this as

$$nh = \frac{-1}{2m^2 g} \int_{z=0}^{z=h_0} u^{1/2} \, du, \text{ or} \tag{5.26}$$

$$nh = \frac{-1}{3m^2 g} \left(2mE - 2m^2 gz\right)^{3/2} \Big]_0^{h_0 = E/mg} . \tag{5.27}$$

Hence,

$$\left(3m^2gnh\right)^{2/3} = 2mE - 2m^2gz\Big]_0^{E/mg}, \tag{5.28}$$

or, solving this for $E$, we get

$$E_n = \frac{\left(3m^2gnh\right)^{2/3}}{2m}. \tag{5.29}$$

This example serves to demonstrate how the Wilson-Sommerfeld rule is used. Looking now at the hydrogen atom, which is slightly more tricky, we see that, although it is more involved that the simpler Bohr model, it does give us some benefits.

Picture the electron as moving in an elliptical orbit around the nucleus. The Wilson-Sommerfeld rule gives

$$J_\theta = \oint p_\theta \, d\theta = \int_0^{2\pi} P \, d\theta = 2\pi P = kh \tag{5.30}$$

where $k$ is an integer. The choice $k = 0$ may be ignored, since $k = 0$ implies $P = 0$, or straight line motion into the nucleus. Further, negative values of $k$ need not be considered, since the motion corresponding to a negative value differs from that corresponding to a positive value only by the reversal of the direction of travel around the nucleus. The phase integral, in terms of points designated in Fig 5.1, is

$$J_r = \oint p_r \, dr = \int_{ABC} p_r \, dr + \int_{CDA} p_r \, dr. \tag{5.31}$$

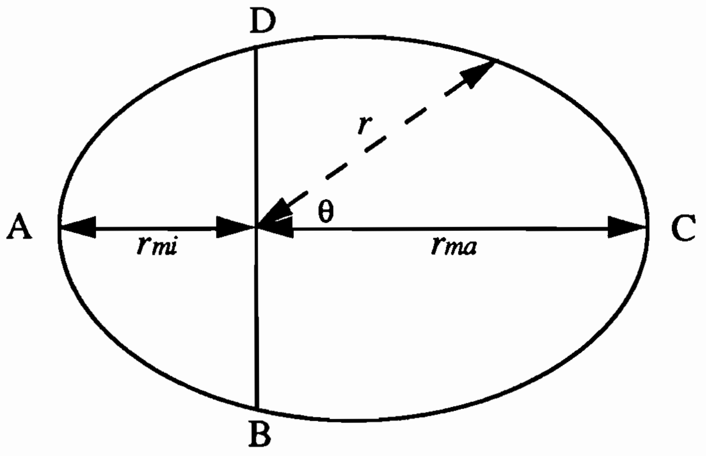


Fig 5.1: Geometry for Sommerfeld's treatment of the hydrogen atom

From the energy equation

$$H = E = \frac{1}{2m}\left(p_r{}^2 + \frac{P^2}{r^2}\right) - \frac{Ze^2}{r} = \text{ a constant}, \tag{5.32}$$

where an electron is attracted to the origin with force $\frac{Ze^2}{r^2}$, we obtain

$$p_r = \pm\sqrt{2m\left(E + \frac{Ze^2}{r} - \frac{P^2}{2mr^2}\right)} = \pm sqrt. \tag{5.33}$$

Now $r$ increases along ABC and decreases along CDA. Hence, by the Hamiltonian equations of motion we get

(5.34) $p_r = m\dot{r}$, $p_r = +sqrt$ along ABC, $p_r = -sqrt$ along CDA.

Therefore

$$J_r = \int_{r\,\min}^{r\,\max}(+sqrt)\,dr + \int_{r\,\max}^{r\,\min}(-sqrt)\,dr \tag{5.35}$$

$$= 2\int_{r\,\min}^{r\,\max}(+sqrt)\,dr \tag{5.36}$$

Sommerfeld evaluated this integral, and, with use of the Wilson-Sommerfeld rule obtained:

$$J_r = -2\pi P + 2\pi Ze^2\sqrt{\frac{m}{-2E}} = sh \tag{5.37}$$

where $s$ is the "radial quantum number" and is an element of the the non-negative integers. Eliminating P by means of equation 5.30 we find that

$$2\pi Ze^2\sqrt{\frac{m}{-2E}} = (s+k)h \tag{5.38}$$

now, setting $n = s + k$ = "the principle quantum number", and solving for $E$, we obtain the energy levels

$$E_n = -\frac{2\pi^2 mZ^2e^4}{h^2n^2} = -\frac{\Re Z^2}{n^2} \tag{5.39}$$

where $\Re$ is the Rydberg constant (see equation 5.16).

Comparing this with the energy levels gotten from Bohr's model, we see that the Wilson-Sommerfeld conditions give the same energy levels for the hydrogen atom as the simpler Bohr theory gave. The gain is that we are no longer restricted to circular orbits. However, we cannot yet consider non-periodic motions, nor can we predict transition probabilities, that is spectrum line intensities.

*Problem 5.1:* Determine the Wilson-Sommerfeld energy levels of a plane, rigid rotator of moment of inertia $I$ about a fixed axis. The classical energy is $E = \frac{1}{2}I\dot{\theta}^2 = \frac{1}{2I}p_\theta{}^2$.[31]

Answer: $E_n = n^2\hbar^2/2I$

*Problem 5.2:* Determine the Wilson-Sommerfeld energy levels of a one-dimensional harmonic oscillator. The classical energy is $E = \frac{p^2}{2m} + \frac{1}{2}kx^2$.[32]

Answer: $E_n = nh\nu$ where $\nu = \frac{1}{2\pi}\sqrt{\frac{k}{m}}$.

*Problem 5.3:* Determine the Wilson-Sommerfeld energy levels of a particle constrained to move along the $x$ axis in the region $0 \le x \le L$.[33]

Answer: $E_n = \frac{h^2n^2}{2mL^2}$.

## 5.7 Sommerfeld's Relativistic Treatment of the Hydrogen Atom

At the same time that Sommerfeld was working on the above theory, he realized that, according to the Bohr formula for the velocity of an electron in a circular orbit $v_n = e^2/n\hbar$, the speed of the electron is so great that the relativistic variation of the mass of the electron with its speed should not be disregarded. For example, taking $n = 1$, we find

$$\alpha \equiv \frac{v_1}{c} = \frac{e^2}{\hbar c} \approx \frac{1}{137}. \tag{5.40}$$

A Bohr electron in the ground state has a velocity almost equal to 0.01 $c$. With the Wilson-Sommerfeld elliptical orbits around the nucleus, the velocity of the electron varies, and the corresponding relativistic variation of the mass of the electron may be expected to affect the energy levels, resulting perhaps in detectable shifts in spectral lines.

---

[31] Ikenberry, p. 51, problem 4.8

[32] Ikenberry, p. 52, problem 4.9

[33] Ikenberry, p. 52, problem 4.10

Sommerfeld therefore calculated the phase integrals $J_\theta$ and $J_r$ taking into account the relativistic variation of the mass with speed as given by the formula

$$m = m_0\left(1 - \frac{v^2}{c^2}\right)^{-1/2}. \tag{5.41}$$

Setting $J_\theta = kh$, $J_r = sh$, as we have seen the Wilson-Sommerfeld quantum condition implies, and solving the relations obtained for $E$, he derived the formula

$$E = m_0c^2\left\{-1 + \left(1 + \frac{\alpha^2}{\left(s + \sqrt{k^2 - \alpha^2}\right)^2}\right)^{-1/2}\right\} \tag{5.42}$$

A sufficiently accurate approximate to this is given by the formula

$$E_{n,k} = -\frac{\Re}{n^2} - \frac{\Re\alpha^2}{n^4}\left(\frac{n}{k} - \frac{3}{4}\right) \tag{5.43}$$

where $n = s + k$.

Notice that the Sommerfeld energy levels depend on both $n$ and $k$, and not only on $n$ as with the Bohr energy levels. This result provides an example of a general rule: *The introduction of a perturbation tends to decrease the degeneracy of the energy levels.* In this example, the variation of mass with velocity may be considered as a perturbation. The degeneracy is not completely removed, since the energy levels do not depend on the magnetic quantum number $m$. When a magnetic field is present, it is found that the levels do depend on $m$ as well as on $n$ and $k$. The degeneracy is then completely removed, unless we consider electron spin.

*Problem 5.4:* Verify the correctness of equation (5.43) by obtaining the first two terms in the series expansion of the right-hand side of the energy equation in powers of $\alpha^2$. {Hint: use the approximation

$$(1 + x)^{-1/2} \approx 1 - \tfrac{1}{2}x + \tfrac{3}{8}x^2$$

Also use the definition of $\alpha$ and $\Re$ to introduce the Rydberg constant.}[34]

---

[34] Ikenberry, p. 53, problem 4.12

# Chapter 6

# Non-Classical Quantum Mechanics

*In this chapter we consider further developments of quantum mechanics, beginning with the Heisenberg Uncertainty Principle, then spending most of the chapter dealing with aspects of the Schrödinger Wave Equation. The work of Schrödinger begins the phase of quantum theory known as non-classical quantum mechanics, as opposed to the previous semi-classical quantum mechanics. It is this aspect of quantum mechanics to which the majority of an undergraduate course in quantum mechanics is directed, so what we are doing is reviewing some of the major topics, not covering them in detail.*

## 6.1 The Heisenberg Uncertainty Principle

Classical Newtonian mechanics is *deterministic*, in that if we are given all of the information about a particle, and the forces acting on it, it is possible to calculate precisely any information pertaining to that particle at any future time. Take for example Newton's Second Law. It states, that given the initial coordinates and the velocity of a particle, as well as any forces which affect the particle, the orbit is uniquely determined. This also holds for a system of particles, and hence, under Newtonian mechanics, the universe is deterministic.

Quantum mechanics dispels this "nice" deterministic philosophy. With Werner Heisenberg's paper in 1927, the world had to adopt a new, ambiguous philosophy, for at the heart of natural law is subjective probability. The implications of Heisenberg's paper were that if the momentum of a particle is known precisely, then it follows that the position of that particle is completely unknown.

Quantitatively, if an experiment is set up where the momentum of a particle is measured, then the measurement of the momentum disturbs the system in such a way that a subsequent measurement of position is random to a certain degree. Similarly, if the position is measured first, momentum will be random to an extent. As a derivation of the Heisenberg Uncertainty Principle, assume that we are preforming a measurement on the position of particles with a constant momentum. Let the average of these measurements be $\langle x \rangle$. Then we can form the mean-square deviation

$$(\Delta x)^2 \equiv \left\langle \left(x - \langle x \rangle\right)^2 \right\rangle. \tag{6.1}$$

The standard deviation is labeled $\Delta x$. This is also called the *uncertainty in* $x$, because, if $\Delta x$ is small compared to a typical length in the experiment, one is more certain to find the value $x = \langle x \rangle$, while, if $\Delta x$ is large, it is not certain what the measurement of $x$ will yield.

Similarly, one can find the uncertainty in any physically observable quantity. Heisenberg's Uncertainty relation for momentum and position appears as

$$\Delta x \, \Delta p_x \gtrsim \hbar, \tag{6.2}$$

which is read as "the uncertainty of $x$ multiplied by the uncertainty of $p_x$ is approximately greater than or equal to $\hbar$". (Remember that $\hbar \equiv h/2\pi$.) Observable quantities that follow this relation are called complementary variables. Examples of complementary variables include:

a) position and momenta $(x, p_x)$,
b) energy and time $(E, t)$, and
c) any two Cartesian components of angular momentum $(M_x, M_z)$.

You may recall that in order to determine whether any two operators are complementary variables, you check to see if they are commutable. If they are commutable, than they are compatible, and the Heisenberg Uncertainty relation does not apply. If, however, they are not commutable, then they are complementary, and the Heisenberg Uncertainty relation does apply. For a review of the commutator relation, see Section 1.9.

## 6.2 Operator Derivation of the Schrödinger Wave Equation

Recall that the Hamiltonian function for a particle in a conservative field is equal to the total energy:

$$\hat{H}(\mathbf{r}, \mathbf{p}) = \frac{p^2}{2m} + V(\mathbf{r}) = E. \tag{6.3}$$

This equality between observables implies not that $\hat{H} = E$ but that

$$\hat{H}\psi = E \cdot \psi. \tag{6.4}$$

The use of any convenient representation is implied by (6.4), the symbol $\psi$ not necessarily implying the position representation. For instance, one could be using the position representation, the momentum representation, or the $k$ representation.

As a slight divergence, recall that the correspondences between the previous three representations are:

(6.5) $$\psi_i \leftrightarrow \phi_i \leftrightarrow \varphi_i$$

(6.6) $$A_x \leftrightarrow A_p \leftrightarrow A_k$$

(6.7) $$x \leftrightarrow i\hbar\frac{\partial}{\partial p} \leftrightarrow i\frac{\partial}{\partial x}$$

(6.8) $$-i\hbar\frac{\partial}{\partial x} \leftrightarrow p \leftrightarrow \hbar k \cdot$$

Returning now to our derivation of the Schrödinger Wave Equation, recall that for a specified energy, (6.4) becomes the *Schrödinger Time Independent Wave Equation* (STIWE):

(6.9) $$H\psi = E\psi.$$

For an unspecified energy, (6.4) becomes the *Schrödinger Time Dependent Wave Equation* (STDWE):

(6.10) $$H\psi = i\hbar\frac{\partial\psi}{\partial t}.$$

Frequently (6.10) is written in the form of (6.9) with the understanding from the context that $E = i\hbar\frac{\partial}{\partial t}$.

Equations (6.9) and (6.10) become the Schrödinger wave equations for a system of $n$ particles when $H$ is the total Hamiltonian for the entire system, and the wavefunction $\psi$ (in the position representation) is a function of the time and of the $3n$ position coordinates of the $n$ particles.

*Problem 6.1:* Give the Schrödinger wave equation for a system of $n$ free particles, each of mass $m$.[35]

Answer: $$-\frac{\hbar^2}{2m}\sum_{j=1}^{n}\left(\frac{\partial^2\psi}{\partial x_j{}^2}+\frac{\partial^2\psi}{\partial y_j{}^2}+\frac{\partial^2\psi}{\partial z_j{}^2}\right) = i\hbar\frac{\partial\psi}{\partial t}$$

---

[35] Ikenberry, p. 84, problem 8.1

*Problem 6.2:* Give the wave equations in the position and in the k representations, for an isotropic linear oscillator in three-space dimensions.[36]

$$\text{Answers: } -\frac{\hbar^2}{2m}\nabla^2\psi + \frac{1}{2}ar^2\psi = E\psi, \quad \frac{\hbar^2 k^2}{2m}\varphi - \frac{1}{2}a\nabla_k^2\varphi = E\varphi.$$

*Problem 6.3:* Give the Schrödinger wave equations for a hydrogen atom, taking into account the motion of the nucleus.[37]

$$\text{Answer: } -\frac{\hbar^2}{2m}\nabla_1^2\psi - \frac{\hbar^2}{2M}\nabla_2^2\psi - \frac{e^2}{r_{12}}\psi = E\psi$$

where $m$ = mass of electron, $M$ = mass of proton, $r_{12}$ is the distance between the proton and electron, and $\psi = \psi(\mathbf{r}_1, \mathbf{r}_2, t)$.

*Problem 6.4:* Give the Schrödinger wave equation for the two electrons in a helium atom, neglecting the motion of the nucleus.[38]

$$\text{Answer: } -\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2)\psi + \left(\frac{e^2}{r_{12}} - \frac{2e^2}{r_1} - \frac{2e^2}{r_2}\right)\psi = E\psi.$$

## 6.3 The Schrödinger and Klein-Gordon Equations

Theorists often obtain differential equations which are satisfied by functions representing wave fields by applying relevant basic principles of physics and suitable simplifying assumptions. For instance, the equation which governs the motion of waves along a string is derived by applying Newton's Second Law and assuming transverse motion in a plane, small displacements, absence of friction, etc.

By using this method, theoreticians have derived the Schrödinger wave equations. When the correspondences (6.7) and (6.8) are applied to the nonrelativistic dispersion relation for de Broglie waves associated with free particles in one-space dimension

$$\omega = \frac{\hbar k^2}{2m}, \tag{6.11}$$

---

[36] Ikenberry, p. 85, problem 8.2

[37] Ikenberry, p. 85, problem 8.3

[38] Ikenberry, p. 85, problem 8.4

we find that

$$i\frac{\partial\psi}{\partial t} = -\frac{\hbar}{2m}\frac{\partial^2\psi}{\partial x^2} \tag{6.12}$$

is the STDWE for a free particle in one-space dimension. Similarly, when the correspondences (6.7) and (6.8) are applied to the relativistic dispersion relation for de Broglie waves associated with free particles in one-space dimension

$$\omega^2 = k^2c^2 + \frac{m_0^2c^4}{\hbar^2}, \tag{6.13}$$

leads to

$$\frac{\partial^2\psi}{\partial t^2} = c^2\frac{\partial^2\psi}{\partial x^2} - \frac{m_0^2c^4}{\hbar^2}\psi. \tag{6.14}$$

This is the *Klein-Gordon equation* for a free particle in one-space dimension. The Klein-Gordon equation is frequently called the *Schrödinger relativistic wave equation.*

*Problem 6.5:* From the nonrelativistic energy-momentum relation

$$E = \frac{p^2}{2m} + V(\mathbf{r}) \tag{6.15}$$

for a particle in a potential field V(r), obtain the Schrödinger time dependent wave equation:

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi = i\hbar\frac{\partial\psi}{\partial t}. \tag{6.16}$$[39]

*Problem 6.6:* From the relativistic energy-momentum relation

$$(E - e\Phi)^2 = (cp - eA)^2 + m_0^2c^4 \tag{6.17}$$

for a particle of charge e in an electromagnetic field, obtain the Klein-Gordon equation

$$\left(-\hbar^2\frac{\partial^2}{\partial t^2} - 2ie\hbar\Phi\frac{\partial}{\partial t} - ie\hbar\frac{\partial\Phi}{\partial t} + e^2\Phi^2\right)\psi = \left(-\hbar^2c^2\nabla^2 + 2i\hbar ceA\cdot\nabla + i\hbar ce(\nabla\cdot A) + e^2A^2 + m_0^2c^4\right)\psi. \tag{6.18}$$[40]

---

[39] Ikenberry, p. 59, problem 5.7

[40] Ikenberry, p. 59, problem 5.8

## 6.4 Stationary States

Consider a conservative system in which the potential function is a function of the coordinates only, and not the time. Then the Hamiltonian does not depend explicitly on the time either. We may thus expect to find solutions of the form

$$\psi(\mathbf{r},t) = \psi(\mathbf{r}) \cdot T(t). \tag{6.19}$$

Since $H$ commutes with $T(t)$ and $i\hbar\frac{\partial}{\partial t}$ commutes with $\psi(\mathbf{r})$, we may write

$$T(t) H \psi(\mathbf{r}) = i\hbar\, \psi(\mathbf{r}) \frac{d}{dt} T(t) \tag{6.20}$$

when we substitute (6.19) into (6.10). In (6.19) the variables $\mathbf{r}$ and $t$ are separable. Standard arguments lead to the separated equations

$$H\psi(\mathbf{r}) = \lambda\psi(\mathbf{r}), \text{ and} \tag{6.21a}$$

$$i\hbar \frac{d}{dt} T(t) = \lambda T(t), \tag{6.21b}$$

where $\lambda$ is some constant. This second equation is readily integrated, and omitting the constant of integration (since it can easily be absorbed by $\psi(\mathbf{r})$), we obtain

$$T(t) = e^{-i\lambda t/\hbar}. \tag{6.22}$$

From this and any solution $\psi(\mathbf{r})$ of the *eigenvalue equation* (6.21a), we obtain a solution

$$\psi(\mathbf{r},t) = \psi(\mathbf{r}) \cdot e^{-i\lambda t/\hbar} \tag{6.23}$$

of the Schrödinger wave equation.

Wave functions of the form (6.23) are said to represent *stationary states*. Such states play a central role in quantum mechanics.

*Problem 6.7:* Show that, for a wavefunction (6.23) with finite norm, $\lambda = \langle H \rangle$.[41]

---

[41] Ikenberry, p. 89, problem 8.12

*Problem 6.8:* In the Schrödinger wave equation (6.10) let

$$\psi = Ce^{iW/\hbar}$$

where $W = W(\mathbf{r},t)$. Show that, in the limit $\hbar \rightarrow 0$, W satisfies the Hamilton-Jacobi equation of classical mechanics:

$$\frac{1}{2m}(\nabla W)^2 + V + \frac{\partial W}{\partial t} = 0.$$[42]

### 6.5 Physical Interpretation of the Wave Function

In classical continuum mechanics the equation of conservation is

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0 \tag{6.24}$$

where $\rho$ is the density and $\mathbf{j}$ is the current. When integrated over a fixed region $R$ of space, by aid of Gauss's theorem for transforming the volume integral of a divergence into a surface integral, this gives

$$\frac{\partial}{\partial t}\iiint_R \rho \, dV + \iint_\sigma \mathbf{j} \cdot d\sigma = 0 \tag{6.25}$$

where $\sigma$ is the surface bounding the region $R$. There is no net flux through $\sigma$ and the total amount of matter (or charge) in $R$ remains constant when the surface integral vanishes.

In interpreting an equation of the form (6.24), we are inclined to call the vector whose divergence appears in the second term a current, and the quantity whose time-derivative appears in the first term a density.

Looking at the STDWE in three-space dimensions,

$$i\hbar\frac{\partial \psi}{\partial t} = \frac{\hbar^2}{2m}\nabla^2\psi + V\psi, \tag{6.26}$$

and its complex conjugate,

$$-i\hbar\frac{\partial \psi^*}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi^* + V\psi^*, \tag{6.27}$$

---

[42] Ikenberry, p. 89, problem 8.13

we multiply the former by $\psi^*$, the later by $\psi$, and subtract the resulting equations, to get

$$\frac{\partial}{\partial t}\left(\psi^*\psi\right)+\nabla\cdot\left(\frac{i\hbar}{2m}\left(\psi\nabla\psi^* - \psi^*\nabla\psi\right)\right)=0, \tag{6.28}$$

an equation of the form of (6.24): $\frac{\partial}{\partial t}J + \nabla\cdot\mathbf{S} = 0$. Hence $J = \psi^*\psi$ is a density, and

$\mathbf{S} = \frac{i\hbar}{2m}\left(\psi\nabla\psi^* - \psi^*\nabla\psi\right)$ is a current.

From this we can picture an electron as a *cloud of charge* with density $\rho$ varying from point to point, albeit highly concentrated in a small region. The current $\mathbf{j} = \rho\mathbf{u}$ and the density vanish at a great distance from the center of the cloud. We can represent the cloud density and current in terms of a wavefunction as follows.

Let $\psi_1$ satisfy the STDWE and have finite norm. Taking $\psi = c\psi_1$, where $c$ is chosen to normalize $\psi$, we obtain a wavefunction which satisfies the STDWE and has unit norm. Then, as stated above, we may take

$$\rho = eJ = e\psi^*\psi, \text{ and } \mathbf{j} = e\mathbf{S} = \frac{i\hbar e}{2m}\left(\psi\nabla\psi^* - \psi^*\nabla\psi\right) \tag{6.29}$$

as the charge density and the current density in our cloud picture of the electron.

*Problem 6.9:* Show that

$$\mathbf{S} = \frac{1}{m}\operatorname{Re}\left[\psi^*\mathbf{p}\psi\right] \tag{6.30}$$

where the symbol Re denotes real part.[43]

*Problem 6.10:* Writing $\psi$ as $\psi = |\psi|e^{i\theta}$, show that

$$J = |\psi|^2, \text{ and } \mathbf{S} = \frac{\hbar}{m}|\psi|^2\,\nabla\theta. \tag{6.31}$$

Hence the magnitude of $\psi$ determines the charge density, this and the gradient of the phase of $\psi$ determine the charge current.[44]

---

[43] Ikenberry, p. 87, problem 8.6

[44] Ikenberry, p. 87, problem 8.7

*Problem 6.11:* For a given $\psi$ with unit norm define a second wave function $\psi'$ with unit norm by

(6.32) $$\psi = \psi' e^{i\lambda t},$$

where $\lambda$ is a real constant. Show that:

(6.33) $$J = \psi'^{*}\psi', \text{ and } \mathbf{S} = \frac{i\hbar}{2m}\left(\psi'\nabla\psi'^{*} - \psi'^{*}\nabla\psi'\right).$$[45]

## 6.6 Statistical Interpretation of the Wavefunction, Born's Probability Wave

When dealing with waves, if one calls the amplitude function $\psi$, the intensity function will be $|\psi|^2$. Born suggested in 1927 that, when referring to the propagation of particles (since we now know that particles can have wave representations), $|\psi|^2$ is more appropriately considered a *probability density*. The function $\psi$ is called the wavefunction (or state function or state vector) of the particle. Quantitatively, the Born Postulate states (in Cartesian space): the wavefunction for a particle $\psi(x,y,z,t)$ is such that

(6.34) $$|\psi|^2\, dx\,dy\,dz = P\, dx\,dy\,dz$$

where $P\, dx\,dy\,dz$ is the probability that measurement of the particle's position at the time finds it in the volume element $dx\,dy\,dz$ about the point $(x,y,z)$.

This statement is consistent with what we know about the interference of photo or electrons. In all cases an interference pattern develops when we examine a large numb of particles. The wavefunction $\psi$ generates this interference pattern. Where $|\psi|^2$ is lar the probability that a particle is found there is large, and so the intensity is greater at t spot.

The rules of quantum mechanics (to be explored in the next section) giv technique for calculating the wavefunction $\psi$ to within an arbitrary multiplicative const As we have seen, the equation that one solves to find $\psi$ is one of the Schrödi equations, and is a homogeneous linear equation. Suppose we solve it and obta function $\psi$. Then $A\psi$ is also a solution, where $A$ is a constant. For problems whe

---

[45] Ikenberry, p. 87, problem 8.8

can be said with certainty that the particle is in a given region of space, we know that

$$\int_V |\psi|^2 \, dx\, dy\, dz = 1 \tag{6.35}$$

This is the standard property that probability densities must satisfy under the above conditions.

The wave mechanical expectation values of functions $f(x)$ and $g(p)$, defined by

$$\langle f(x) \rangle = \frac{\langle \psi | f | \psi \rangle}{\langle \psi | \psi \rangle} \tag{6.36a}$$

$$\langle g(p) \rangle = \frac{\langle \Phi | g | \Phi \rangle}{\langle \Phi | \Phi \rangle} \tag{6.36b}$$

for wave functions for which the integrals exist, are readily correlated with classical expectation values by the identifications

$$P(x) = \frac{\psi^* \psi}{\langle \psi | \psi \rangle} \tag{6.37a}$$

$$P(p) = \frac{\Phi^* \Phi}{\langle \Phi | \Phi \rangle}. \tag{6.37b}$$

When the wavefunction has been normalized, we have the simpler relations

$$P(x) = \psi^* \psi \tag{6.38a}$$

$$P(p) = \Phi^* \Phi. \tag{6.38b}$$

A normed wave function is sometimes called a *probability amplitude*, and the square of its magnitude a *probability density*.

## 6.7 The Postulates of Quantum Mechanics

In this section, an attempt is made to justify two postulates which provide the foundation for Born's statistical interpretation of the wave function. These two postulates, Postulates IV and V, are preceded by the statement of three postulates which have already been used, without statement.

*Postulate I:* A state of a mechanical system is completely specified by a wave function $\psi$. All possible information about the system can be derived from its wavefunction.

*Postulate II:* To every observable there corresponds a Hermitian operator with a complete set of eigenvalues.

*Postulate III:* For every system there exists a Hermitian operator $H$, the Hamiltonian operator, which determines the time variation of the wavefunction during any time interval in which the system is not disturbed, through the Schrödinger Time Dependent Wave Equation.

These three postulates need not be discussed further at this point. However, it should be noted that *superposition of states*, required by experimental evidence, implies that quantum mechanical operators are linear. The study of a linear operator results in its *spectral resolution*, or the determination of its eigenvalues and eigenfunctions.

*Postulate IV:* The only possible values which a precise measurement of an observable can yield are the eigenvalues of the operator associated with the observable.

*Postulate V:* The wavefunction of a compound system of noninteracting systems is the sum of the separated wavefunctions.

Postulate V tells us that, when any measuring process is applied to a wavefunction, and the process involves the separation of an assemblage of non-interacting systems, the separation process does not alter the wavefunction of any part of the assemblage. Further, any measurement of an observable in part of the separated system only gives the value of the separated part.

To show that the Heisenberg Uncertainty Principle does not violate these postualates, we assume that the wave function $\psi$ is known immediately before a measurement of an observable $A$. In the measurement of $A$, there is an interaction between the systems in the assemblage and the measurement apparatus. This interaction is not included in the Hamiltonian for the systems of the assemblage, hence the STDWE does not account for the change in the wavefunction produced by the interaction.

## 6.8 The Schrödinger Picture and The Heisenberg Picture

In the *Schrödinger Picture*, wavefunctions are time dependent while operators are considered time independent. In the *Heisenberg Picture*, though, wavefunctions are time independent and operators are time dependent. The basic elements of Heisenberg formalism may be developed by a transformation from the Schrödinger picture. In this development, wavefunctions and operators in the two pictures, Heisenberg and Schrödinger, will be distinguished by superscripts $H$ and $S$ respectively.

First we need a formal solution of the STDWE (6.10) in the Schrödinger picture. Assuming that the Hamiltonian does not depend explicitly on time, the STDWE may be solved formally by treating the Hamiltonian as a constant, resulting in:

(6.39) $\psi(\mathbf{r},t) = T_{-}\psi(\mathbf{r},0)$, where

(6.40a) $$T_{-} = e^{-iHt/\hbar},$$

(6.40b) $$T_{+} = e^{+iHt/\hbar} = (T_{-})^{\dagger},$$

(6.40c) $$T_{-}T_{+} = T_{+}T_{-} = 1.$$

Since wave functions are time independent in the Heisenberg picture, but time dependent in the Schrödinger picture, we obtain our starting point from (6.39):

(6.41a) $$\psi^{H} = \psi(\mathbf{r},0) = T_{+}\psi^{S}$$

(6.41b) $$\psi^{S} = \psi(\mathbf{r},t) = T_{-}\psi^{H}.$$

The relation between the two pictures $A^{H}$ and $A^{S}$ of an operator $A$ is now completely and uniquely determined by the requirement that

(6.42) $$\left\langle \psi^{H} \middle| A^{H} \middle| \psi^{H} \right\rangle = \left\langle \psi^{S} \middle| A^{S} \middle| \psi^{S} \right\rangle$$

for all wavefunctions. In fact, eliminating $\psi^{S}$ from (6.42) by means of (6.41b), we find that

(6.43) $$\left\langle \psi^{H} \middle| A^{H} \middle| \psi^{H} \right\rangle = \left\langle T_{-}\psi^{H} \middle| A^{S} \middle| T_{-}\psi^{H} \right\rangle = \left\langle \psi^{H} \middle| T_{+}A^{S}T_{-} \middle| \psi^{H} \right\rangle.$$

Hence, we find that

(6.44a) $$A^{H} = T_{+}A^{S}T_{-}$$

(6.44b) $$A^{S} = T_{-}A^{H}T_{+}.$$

*Problem 6.12:* Show that $A^{H} = A^{S}$ if A and H commute.[46]

*Problem 6.13:* Show that $\left[A^{H}, B^{H}\right] = T_{+}\left[A^{S}, B^{S}\right]T_{-}$.[47]

---

[46] Ikenberry, p. 93, problem 8.19

[47] Ikenberry, p. 93, problem 8.20

*Problem 6.14:* Let $\psi(x,0) = e^{ax}$. Calculate $\psi(x,t)$ from (6.39), using $H = \frac{p^2}{2m} = \frac{-\hbar^2}{2m}\frac{d^2}{dx2}$.[48]

Answer: $\psi(x,t) = e^{ax - i\hbar\alpha^2 t/2m}$ where $\alpha = \frac{d}{dx}$.

---

[48] Ikenberry, p. 93, problem 8.23

# Part III

# Quantum Theory of Measurement

# Chapter 7

# Quantum Theory of Measurement

*This chapter looks at one of the results of quantum theory – the Heisenberg Uncertainty Principle – in general form, and its consequences for measurement. We will find a way to quickly determine whether two observables may be determined precisely simultaneously on the same system.*

## 7.1 Simultaneous and Repeatable Measurements

In analyzing the results of repeated measurements, one must consider the effect each measurement has on the state of the system, and any changes in the state between measurements. Any measurement on a system is gotten by an interaction between the system and the measurement apparatus. Extreme cases occur when the alteration is negligible, or when the measurement radically alters the system. An example of the first is the measurement of the energy of a charged particle by measuring its radius of curvature in a magnetic field, and example of the second is the measurement of a particles energy by measuring the length of its track in a Wilson cloud chamber.

A measurement of an observable $A$ is repeatable only if the measurement does not affect the value of $A$, and the expectation value of $A$ does not change during the time between measurements. If $A$ is not a constant of the motion, in any interval when the system is not disturbed, the expectation value of $A$ will change at a rate of

$$\frac{d}{dt}\langle A\rangle = \frac{1}{i\hbar}\langle [A,H]\rangle + \left\langle \frac{dA}{dt}\right\rangle . \tag{7.1}$$

A disturbance is any interaction of the system with its environment that is not accounted for in the Hamiltonian of the system. Even in the absence of disturbances though, the rate of change will be zero only when $A$ is a constant of the motion. There may be times, even if $A$ is not a constant of the motion, when two measurements will be taken in such a short time interval that the system has not changed appreciably. We call such measurements *simultaneous measurements*.

## 7.2 Compatible Observables

Suppose a measurement is made of observable $A$, and is followed immediately by a measurement of observable $B$, and is followed by a measurement of observable $A$ again. The measurement of $B$ will have disturbed the system, and hence the second measurement of $A$ may be radically different than the first measurement. This suggests the following definition: the measurement of $B$ is *compatible* with the measurement of $A$ if the immediate remeasurement of $A$ necessarily gives a value of $a$ in the same range as the first measurement of $A$. If two observables are compatible, successive simultaneous measurements give repetitive results.

A practical test for compatibility is the commutivity (or lack thereof) of the observable's operators. Assuming that $A$ and $B$ possess complete sets of eigenfunctions, then the following are equivalent:

(7.2a) $A$ and $B$ are compatible,

(7.2b) $A$ and $B$ possess a complete set of simultaneous eigenfunctions,

(7,2c) $A$ and $B$ commute.

*Problem 7.1:* Show that if two noncommuting operators $A$ and $B$ possess a simultaneous eigenfunction, then the operator $C = i[A,B]$ possesses zero as an eigenvalue. Hence, show in particular that $x_i$ and $p_i$ do not possess a simultaneous eigenfunction.[49]

## 7.3 Compatible Constants of the Motion

If the expectation value of an observable $A$ changes with time, repetitions of measurements of $A$ would not be expected to give repetitive results. In the definition of compatible observables we avoid this complication by requiring measurements be simultaneous. Obviously, this requirement is not necessary for constants of the motion, whose expectation values do not change as long as the system remains undisturbed.

Classically, an observable $A$ is a constant of the motion if $\frac{dA}{dt} = 0$ for all possible motions. Quantum mechanically, an observable $A$ is a constant of the motion if $\frac{d\langle A\rangle}{dt} = 0$ for all states. Although these two concepts are not identical, we can apply our knowledge of classical constants of the motion to our search for quantum mechanical constants of the

---

[49] Ikenberry, p. 151, problem 13.1

motion. In fact, it can be proven that any classical constant of the motion is also a quantum mechanical constant of the motion. The converse is also true, if the quantum mechanical observables have classical counterparts.

When we consider two or more quantum mechanical constants of the motion we encounter a complication not encountered in classical mechanics. If we measure a constant of the motion $A$ twice, we expect the same answer both times. However, if we measure another observable $B$ between the two measurement of $A$, we have disturbed the system in such a way the the second measurement of $A$ much satisfy the Heisenberg Uncertainty Relation (Section 6.1) with $B$. However, if $A$ and $B$ are compatible constants of the motion, the Heisenberg Uncertainty Relation does not apply, and repetitive measurements will produce repetitive results.

We say that an eigenvalue of a constant of the motion is a "good quantum number" because it will produce the same results, regardless of the time development of an undisturbed system.

*Problem 7.2:* Show that $H$, $p_x$, $p_y$, and $p_z$ are good quantum numbers for a free particle and that these may be united into a compatible set of constants of the motion.[50]

## 7.3 Complete Sets of Compatible Observables

Classically, the initial state of a mechanical system is determined by precise measurements of the position coordinates and conjugate momenta. This will not be sufficient for quantum mechanical systems, since position and momentum are not compatible. Instead, we must try to find other measurements which could determine the state of the system (eg: determine the wavefunction $\psi$ uniquely).

The only possible results of a precise measurements of an observable $A$ are the eigenfunctions of the Hermitean operator associated with $A$. Suppose such a measurement results in a certain eigenvalue of $A$, say $a_j$. Then the wavefunction describing the state of the system is an eigenfunction of $A$ for $A = a_j$. If $a_j$ is a non-degenerate eigenvalue then the wavefunction is uniquely determined, except for arbitrary phase and normalization factors. If however, $a_j$ is a degenerate eigenvalue of $A$ (say $k$-fold) then the wavefunction may be a linear combination of $k$ linearly independent eigenfunctions. In this case, we

---

[50] Ikenberry, p. 157, problem 13.6

need to measure a second observable $B$ compatible with $A$. Again, this will result in some eigenvalue of $B$, say $b_j$. This may now or may not uniquely determine $\psi$. If it does, $A$ and $B$ form a *complete set of compatible observables*. If it does not the process can be continued until a complete set is formed. Of special importance are *complete sets of constants of the motion*, whose expectation values do not change with time.

# Part IV

# Relativistic Quantum Mechanics

# Chapter 8

# Matrix Representations

*This chapter introduces a technique that is essential to the development of the mathematics of relativistic quantum mechanics–the ability to represent operators and functions as matrices and vectors, respectively. It looks at matrix representations of eigenfunctions and eigenvalues, and the Hilbert Space $H_0$. It then uses these to look at the Schrödinger Equation, and the transformation to the Heisenberg Picture.*

## 8.1 Linear Algebraic Techniques

If $f_1, f_2, f_3, \cdots, f_n$ are a set of $n$ linearly independent functions in a linear manifold $M$ having the property that, given any $f$ in $M$, there are $n$ constants $c_1, c_2, c_3, \cdots, c_n$, real or complex such that

$$f = \sum_{i=1}^{n} c_i f_i \,, \tag{8.1}$$

then the $c_i$, which are determined by $f$ and the particular set of functions $f_1, f_2, f_3, \cdots, f_n$ are called *components of $f$ relative to the basis $f_1, f_2, f_3, \cdots, f_n$*. Since the components of $f$ relative to a given basis fully specify $f$, we may consider the $c_i$ as providing a *matrix representation of $f$* and write $f$ as the column vector

$$f \leftrightarrow c = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}. \tag{8.2}$$

If there exists only a finite number of functions $f_i$ in the linear manifold $M$ such that the representation (8.1) exists for all $f$ in $M$, then $M$ is *finite dimensional*. The dimensionality is the minimum number of basis functions required to do this. When $M$ is a finite dimensional inner-product space, such as will be important to quantum mechanics, an orthonormal basis is frequently chosen. Then we have

$$\langle f_i | f_j \rangle = \delta_{ij}, \tag{8.3a}$$

(8.3b) $$c_i = \langle f_i | f \rangle, \text{ and}$$

(8.3c) $$c_i^* = \langle f | f_i \rangle.$$

Further, when $f$ has the representation (8.1), its complex conjugate $f^*$ has the representation

(8.4) $$f^* = \sum_{i=1}^{n} c_i^* f_i^* .$$

This corresponds to the row vector

(8.5) $$f^* \leftrightarrow c' = \left( c_1^*, c_2^*, \cdots, c_n^* \right).$$

In general $c' \neq c^*$, unless the basis functions are real valued.

*Problem 8.1:* Obtain a basis for the linear manifold of all functions $f(x)$ which satisfy

$$\frac{d^2 f}{dx^2} + 4f = 0.$$[51]

Answer: $\{\cos 2x, \sin 2x\}$

*Problem 8.2:* Obtain an orthonormal basis for the linear manifold of all eigenfunctions of the squared angular momentum operator $M^2$ for $M^2 = l(l+1)\hbar^2$, $l$ fixed.[52]

Answer: $\left\{ Y_l^m \middle| -l \leq m \leq l \text{ and } Y_l^m \text{ is the simultaneous eigenfunction of } M^2 \text{ and } M_z \right\}$

## 8.2 Matrix Representations of Operators

If $A$ is a linear operator on a finite dimensional linear manifold $M$, such that $Af$ is an element of $M$ whenever $f$ is, and there is a basis set $f_1, f_2, f_3, \cdots, f_n$, then we can

\

---

[51] Ikenberry, p. 206, problem 17.3

[52] Ikenberry, p. 206, problem 17.4

find a representation of the form (8.1) for any $f$ in $M$. Moreover, since $Af_j$ belongs to $M$, we can write it as

$$(8.6) \qquad Af_j = \sum_{i=1}^{n} a_{ij} f_i \ , \ j = 1,\ 2,\ \ldots,\ n.$$

The sequence of $a_{ij}$ provides a representation of $Af_j$ as a column vector,

$$(8.7) \qquad Af_j \leftrightarrow a_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}.$$

Further, the sequence of column vectors $\mathcal{A} = (a_1, a_2, \cdots, a_n)$, which is the $n$x$n$ matrix

$$(8.8) \qquad \mathcal{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix},$$

provides a matrix representation of the operator $A$ relative to the basis $f_1, f_2, f_3, \cdots, f_n$. For brevity it is common to write $A \leftrightarrow \mathcal{A}$ and mention the basis set only when necessary.

When $M$ is an inner product space, and the basis functions are orthonormal, we have

$$(8.9) \qquad a_{ij} = \langle f_i | A | f_j \rangle.$$

However, (8.6) is often more useful in determining the matrix elements.

Returning to some Matrix Algebra, recall that a matrix whose diagonal elements are the only non-zero elements is called a diagonal matrix. It may be observed from (8.6) that the matrix (8.8) is diagonal if and only if the basis functions are eigenfunctions of $A$.

As an example of the technique, suppose we want to obtain the matrix representations $\mathcal{A}$ and $\mathcal{B}$ of $A = \frac{d}{dx}$ and $B = \frac{d^2}{dx^2}$ over the linear manifold $M$ determined by functions $f(x)$ which satisfy $\frac{d^2 f}{dx^2} + m^2 f = 0$, relative to $f_1 = \frac{1}{\sqrt{\pi}} \cos(mx)$ and

$f_2 = \frac{1}{\sqrt{\pi}}\sin(mx)$ as a basis. To do this, we look at $Af \leftrightarrow \mathcal{A}f$. Notice that

(8.10) $$Af = D\begin{pmatrix} \frac{1}{\sqrt{\pi}}\cos(mx) \\ \frac{1}{\sqrt{\pi}}\sin(mx) \end{pmatrix} = \begin{pmatrix} \frac{-m}{\sqrt{\pi}}\sin(mx) \\ \frac{m}{\sqrt{\pi}}\cos(mx) \end{pmatrix} \leftrightarrow \mathcal{A}f.$$

Consequently,

(8.11) $$\mathcal{A}f = \begin{bmatrix} 0 & -m \\ m & 0 \end{bmatrix}.$$

Similarly, we look at $Bf \leftrightarrow \mathcal{B}f$. Notice that

(8.12) $$Bf = D^2\begin{pmatrix} \frac{1}{\sqrt{\pi}}\cos(mx) \\ \frac{1}{\sqrt{\pi}}\sin(mx) \end{pmatrix} = \begin{pmatrix} \frac{-m^2}{\sqrt{\pi}}\cos(mx) \\ \frac{-m^2}{\sqrt{\pi}}\sin(mx) \end{pmatrix} \leftrightarrow \mathcal{B}f.$$

Consequently,

(8.13) $$\mathcal{B}f = \begin{bmatrix} -m^2 & 0 \\ 0 & -m^2 \end{bmatrix}.$$

*Problem 8.3:* Obtain the matrix representations $\mathcal{A}$ and $\mathcal{B}$ of $A = \frac{d}{dx}$ and $B = \frac{d^2}{dx^2}$ over the linear manifold $M$ determined by functions $f(x)$ which satisfy $\frac{d^2f}{dx^2} + m^2 f = 0$, relative to $f_1 = \frac{1}{\sqrt{2\pi}}e^{imx}$ and $f_2 = \frac{1}{\sqrt{2\pi}}e^{-imx}$ as a basis.[53]

Answer: $\mathcal{A} = \begin{bmatrix} im & 0 \\ 0 & -im \end{bmatrix}$ and $\mathcal{B} = \begin{bmatrix} -m^2 & 0 \\ 0 & -m^2 \end{bmatrix}$

---

[53] Ikenberry, p. 208, problem 17.13

*Problem 8.4:* Using the simultaneous eigenfunctions $Y_l^m(\theta,\phi)$ of $M^2$ and $M_z$, obtain matrix representations of $M_+, M_-, M_x, M_y, M_z, M^2$ when restriced to the linear manifold of functions for which $M_z = 2\hbar^2 \; (l = 1)$.[54]

Answer: $M^2 = 2\hbar^2\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$, $M_z = \hbar\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & -1\end{bmatrix}$, $M_+ = \sqrt{2}\hbar\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{bmatrix}$,

$M_- = \sqrt{2}\hbar\begin{bmatrix}0 & 0 & 0\\ 1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}$, $M_x = \frac{\hbar}{\sqrt{2}}\begin{bmatrix}0 & 1 & 0\\ 1 & 0 & 1\\ 0 & 1 & 0\end{bmatrix}$, $M_y = \frac{i\hbar}{\sqrt{2}}\begin{bmatrix}0 & -1 & 0\\ -1 & 0 & 1\\ 0 & 1 & 0\end{bmatrix}$

## 8.3 Eigenfunctions and Eigenvectors

The vector $u$ such that $\mathcal{A}u = au$ for some scalar $a$ is called an eigenvector of $\mathcal{A}$ for $\mathcal{A} = a$. When $\mathcal{A}$ is a square matrix of order $n$, the eigenvalue equation $\mathcal{A}u = \lambda u$ is equivalent to a system of $n$ linear homogeneous equations for the $n$ components of $u$. This system is consistent for only certain values of $\lambda$, those for which the $\det(\mathcal{A} - \lambda I) = 0$. The values of $\lambda$ for which this system is consistent are the eigenvalues of $\mathcal{A}$. Corresponding to any eigenvalue, call it $a$, is at least one eigenvector $u$ such that $\mathcal{A}u = au$. It is almost trivial that any matrix $\mathcal{A}$ of order $n$ having $n$ distinct eigenvalues has $n$ linearly independent eigenvectors. It is also easy to interpret previously deduced properties of linear operators, in particular of Hermitean operators, in terms of matrices.

As an example, suppose we want to find the eigenvalues and eigenvectors of the matrix $\mathcal{A}$ in the example in the previous section.

$$\mathcal{A}f = \begin{bmatrix}0 & -m\\ m & 0\end{bmatrix}, \tag{8.14}$$

so

$$\mathcal{A} - \lambda I = \begin{bmatrix}-\lambda & -m\\ m & -\lambda\end{bmatrix}. \tag{8.15}$$

---

[54] Ikenberry, p.209, problem 17.15

Now we set the determinant of this to zero to find the eigenvalues, which yields

$$\lambda^2 + m^2 = 0, \text{ or } \lambda = \pm im. \tag{8.16}$$

These are the eigenvalues. The eigenvectors are then

$$f = af_1 + bf_2 \text{ where } f_1 = \frac{1}{\sqrt{\pi}}\cos(mx) \text{ and } f_2 = \frac{1}{\sqrt{\pi}}\sin(mx), \tag{8.17}$$

the linear combinations of the basis vectors. In the case $\lambda = im$, $f = -if_1 + if_2$; while if $\lambda = -im$, $f = if_1 - if_2$.

## 8.4 The Hilbert Space $H_0$

The set of all column vectors c such that

$$c^2 = \lim_{n\to\infty} \sum_{i=1}^{n} c_i^{\,*} c_i\,, \tag{8.18}$$

is called the Hilbert Space $H_0$. There is a one-to-one correspondence between vectors in $H_0$ and functions in the linear manifold M of all square integrable functions, relative to an orthonormal basis in M.

It is easy to show that if $f \leftrightarrow c$ and $g \leftrightarrow d$, where $f$ and $g$ are two square integrable functions and $c$ and $d$ are their representations in $H_0$ then

$$\alpha f + \beta g \leftrightarrow \alpha c + \beta d \tag{8.19}$$

where the ith components are equal [eg: $(\alpha f + \beta g)_i = \alpha c_i + \beta d_i$].

All of the results and techniques of sections 8.1, 8.2, and 8.3 apply to $H_0$, with the realization that we are now talking about an infinite dimensional vector space, rather than a finite dimensional linear manifold or inner-product space.

## 8.5 Matrix Representation of the Schrödinger Equation

We have already seen that any square integrable function may be represented as an infinite series, or, more particularly, any such wave function may be written in the form

$$\psi(\mathbf{r},t) = \lim_{n\to\infty} \sum_{j=1}^{n} c_j(t) f_j(\mathbf{r}), \tag{8.20}$$

where $f_1, f_2, f_3, \ldots$ is an arbitrarily selected complete orthonormal set in the linear manifold of all square integrable functions (section 8.1 extended by 8.4), or in the Hilbert Space $H_0$. Substituting this into the STDWE we readily obtain

$$\lim_{n\to\infty}\sum_{j=1}^{n} c_j H f_j = i\hbar \lim_{n\to\infty}\sum_{j=1}^{n} f_j \frac{\partial c_j}{\partial t}. \tag{8.21}$$

But, because the $f_j$ are mutually orthogonal, this implies that

$$\sum_{j=1}^{\infty} H_{ij} c_j = i\hbar \frac{\partial c}{\partial t} \tag{8.22}$$
,

where $H_{ij}$ is the element in the $i$th row and $j$th column of the matrix representation $\mathcal{H}$ of the Hamiltonian operator $H$, relative to the basis. Writing this in matrix form yields

$$\mathcal{H} c = i\hbar \frac{\partial c}{\partial t} \tag{8.23}$$
.

This is the matrix representation of the STDWE relative to $f_1, f_2, f_3, \ldots$ as a basis.

The STIWE is written as

$$\mathcal{H} c = E c \tag{8.24}$$

in matrix representation relative to $f_1, f_2, f_3, \ldots$ as a basis, and is frequently the starting point for obtaining approximations to the energy eigenvalues of a system.

## 8.6 Transformation from the Schrödinger to Heisenberg Picture

Although Heisenberg matrix mechanics are less useful in solving complex quantum mechanical systems than Schrödinger wave mechanics, it is still an important concept that should be understood. By emphasizing the importance of observables in physical theory Heisenberg was led to the formulation of matrix mechanics in 1925, before Schrödinger's formulation of wave mechanics in 1926. Hence, in the matrix formulation of quantum theory, matrices were dealt with independent of any relation to a wave equation or wave function. The introduction of a vector representation for the wave function is necessary in order to equate the two pictures. Schrödinger did this in 1926.

In the Schrödinger picture, the time development of a mechanical system is imbedded in the wavefunction. The vector representation, relative to any complete set of time independent, orthonormal functions, has time dependent components. If for example we take the time dependent energy eigenfunctions $\psi_j(\mathrm{r}, t)$ as a basis, we write

$$\psi(\mathrm{r},t) \leftrightarrow c^{S} = \begin{pmatrix} c_0 e^{-iE_0 t/\hbar} \\ c_1 e^{-iE_1 t/\hbar} \\ c_2 e^{-iE_2 t/\hbar} \\ \vdots \end{pmatrix}. \tag{8.25}$$

Schrödinger operators, such as the Hamiltonian and the operators for position coordinates and the components of linear and angular momentum are in general time independent.

In the Heisenberg picture, the time development of a mechanical system is imbedded in matrices whose elements correspond to physical observables. The vectors on which the matrices operate have time independent components which are of secondary importance, so long as we stay in the Heisenberg Picture. If we switch to the Schrödinger picture, they become more important. This is easily seen by using the time independent energy eigenfunctions $\psi_j(\mathrm{r},t)$ as an orthonormal basis for a vector representation of the wavefunction. We readily see that

$$\psi(\mathrm{r},t) \leftrightarrow c^{H} = \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \end{pmatrix}, \tag{8.26}$$

where the c's are the same as in the Schrödinger picture. As in Section 6.8, the $S$ and $H$ denote the Schrödinger and Heisenberg pictures, respectively.

Comparing (8.25) and (8.25) it is easy to see that

$$\left(c^{H}\right)_j = e^{iE_j t/\hbar}\left(c^{S}\right)_j, \text{ or} \tag{8.27}$$

$$c^{H} = e^{i\mathcal{H}t/\hbar} c^{S}, \tag{8.28}$$

where $\mathcal{H}$ is the matrix representation of the Hamiltonian relative to the set of its orthonormal eigenfunctions as a basis. In this representation, both $\mathcal{H}$ and $e^{i\mathcal{H}t/\hbar}$ are diagonal matrices, the diagonal elements being $E_j$ and $e^{iE_j t/\hbar}$, respectively. This result appears to be specific to the choice of the energy eigenfunctions as a basis. It is not though, as can be easily demonstrated by letting $T$ be the matrix for transforming from the energy eigenfunctions to any other complete orthonormal basis. Then

$$c^{S} = T^{\dagger} d^{S}, \tag{8.29a}$$

$$c^{H} = T^{\dagger} d^{H}, \tag{8.29b}$$

where $d^S$ and $d^H$ are vector representations of the wavefunction in the appropriate picture relative to the new basis. Substituting these into (8.28) we readily obtain

$$d^H = e^{i\mathcal{H}'t/\hbar} d^S , \tag{8.30}$$

where

$$\mathcal{H}' = T\mathcal{H}T^{\dagger}, \tag{8.31a}$$

$$e^{i\mathcal{H}'t/\hbar} = Te^{i\mathcal{H}t/\hbar}T^{\dagger}. \tag{8.31b}$$

Since (8.28) and (8.30) are of the same form, the choice of basis is unimportant.

To relate matrix representations of operators in the two pictures, write for any operator $A$ and vector $c$

$$(Ac)^H = A^H c^H , \tag{8.32a}$$

$$(Ac)^S = A^S c^S . \tag{8.32b}$$

However, according to (8.28),

$$(Ac)^H = e^{i\mathcal{H}t/\hbar} (Ac)^S , \tag{8.33}$$

hence

$$A^H c^H = e^{i\mathcal{H}t/\hbar} A^S c^S . \tag{8.34}$$

Eliminating $c^H$ by using (8.28) we have

$$A^H e^{i\mathcal{H}t/\hbar} c^S = e^{i\mathcal{H}t/\hbar} A^S c^S . \tag{8.35}$$

But, since $c^S$ is arbitrary, this implies

$$A^H = e^{i\mathcal{H}t/\hbar} A^S e^{-i\mathcal{H}t/\hbar} . \tag{8.36}$$

These results are easily seen to be compatible with the results from Section 6.8.

*Problem 8.5:* Let $A^S$ be a time independent Schrödinger operator. Show that any eigenvalue of $A^S$ is also an eigenvalue of $A^H$.[55]

---

[55] Ikenberry, p. 218, problem 17.33

# Chapter 9

# Angular Momentum and Electron Spin

*This chapter looks briefly at the Angular Momentum Operators, and more closely at Pauli's Theory of Electron Spin.*

## 9.1 The Angular Momentum Operators

Classically, the angular momentum $M$ of a particle, with respect to the origin in a chosen rectangular coordinate system, is the vector

(9.1) $$M = r \times p = M_x \hat{i} + M_y \hat{j} + M_z \hat{k},$$

where

(9.2a) $$M_x = yp_z - zp_y,$$

(9.2b) $$M_y = zp_x - xp_z,$$

(9.2c) $$M_z = xp_y - yp_x.$$

The square of the magnitude of $M$ is

(9.3) $$M^2 = M_x{}^2 + M_y{}^2 + M_z{}^2.$$

In working with relations involving components of $r, p$, and $M$, it is frequently useful to obtain additional relations affecting cyclic permutations of $x$, $y$, and $z$.

Quantum mechanically, with each component of $M$ and with $M^2$ there are associated operators which are obtained by replacing in (9.1) and (9.2) the operators associated with the components of $r$ and $p$. For example, in the position representation (as opposed to the momentum representation) the operators for the components of $M$ are

(9.4a) $$M_x = -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right),$$

(9.4b) $$M_y = -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right),$$

(9.4c) $$M_z = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right).$$

The commutators of any two of the components can be easily found, and with a little algebra these relations are discovered

(9.5a) $$\left[M_x, M_y\right] = i\hbar M_z,$$

(9.5b) $$\left[M_y, M_z\right] = i\hbar M_x,$$

(9.5c) $$\left[M_z, M_x\right] = i\hbar M_y,$$

or, in symbolic form

(9.5d) $$M \times M = i\hbar M.$$

Similarly, we obtain

(9.6a) $$\left[M^2, M_x\right] = 0,$$

(9.6b) $$\left[M^2, M_y\right] = 0,$$

(9.6c) $$\left[M^2, M_z\right] = 0,$$

or, in combined form

(9.6d) $$\left[M^2, M\right] = 0.$$

Thus $M^2$ commutes with each component of $M$, but the components of $M$ do not commute among themselves.

*Problem 9.1:* Using the Hermiticity of the components of $r$ and $p$, and the Heisenberg commutation relations (the commutation of position components and momentum components), show that the operators for the components of $M$ are Hermitean. Then show $M^2$ is Hermitean.[56]

## 9.2 The Angular Momentum Shift Operators

Two more angular momentum operators that prove very useful in quantum mechanical theory are the shift operators $M_+$ and $M_-$, frequently called the raising and lowering operator, respectively. These two adjoint operators are defined by

(9.7a) $$M_+ = M_x + iM_y,$$

(9.7b) $$M_- = M_x - iM_y.$$

---

[56] Ikenberry, p. 163, problem 14.1

With some little bit of algebra it can be shown that

(9.8a) $$M_+M_- = M^2 - M_z{}^2 + \hbar M_z, \text{ and}$$

(9.8b) $$M_-M_+ = M^2 - M_z{}^2 - \hbar M_z.$$

With the help of the commutation relations, it can be further shown that

(9.9a) $$[M_+, M_-] = 2\hbar M_z,$$

(9.9b) $$[M_+, M_x] = \hbar M_z,$$

(9.9c) $$\left[M_+, M_y\right] = i\hbar M_z,$$

(9.9d) $$[M_+, M_z] = -\hbar M_+,$$

(9.9e) $$[M_-, M_x] = -\hbar M_z,$$

(9.9f) $$\left[M_-, M_y\right] = i\hbar M_z,$$

(9.9g) $$[M_-, M_z] = \hbar M_-,$$

(9.9h) $$\left[M^2, M_+\right] = 0, \text{ and}$$

(9.9i) $$\left[M^2, M_-\right] = 0.$$

## 9.3 Eigenvalues and Eigenfunctions of $M^2$ and $M_z$

Since $M^2$ and $M_z$ commute, we look for simultaneous eigenfunctions of both of them. Suppose $Y$ is one such simultaneous eigenfunction for $M^2 = \lambda$ and $M_z = \mu$. Then the following two equations hold

(9.10a) $$M^2 Y = \lambda Y, \text{ and}$$

(9.10b) $$M_z Y = \mu Y.$$

This implies that

(9.11a) $$M_+M^2 Y = \lambda M_+ Y, \text{ and}$$

(9.11b) $$M_+M_z Y = \mu M_+ Y.$$

From (9.9d) and (9.9h) we have

(9.12a) $$M_+M^2 Y = M^2 M_+ Y, \text{ and}$$

(9.12b) $$M_+M_z Y = M_z M_+ Y - \hbar M_z Y.$$

Substituting (9.12) into (9.11), we find that (9.10) implies

(9.13a) $$M^2 M_+ Y = \lambda M_+ Y, \text{ and}$$

(9.13b) $$M_z M_+ Y = (\mu + \hbar) M_+ Y.$$

From this we conclude that if $Y$ is a simultaneous eigenfunction of $M^2$ and $M_z$ for $M^2 = \lambda$ and $M_z = \mu$, then either $M_+ Y = 0$ or $M_+ Y$ is a simultaneous eigenfunction of $M^2$ and $M_z$ for $M^2 = \lambda$ and $M_z = \mu + \hbar$. From this we find by induction that, our supposition implies $(M_+)^k Y = 0$ or $(M_+)^k Y$ is a simultaneous eigenfunction of $M^2$ and $M_z$ for $M^2 = \lambda$ and $M_z = \mu + k\hbar$, where $k$ is any non-negative integer. It may be similarly shown that our supposition implies $(M_-)^k Y = 0$ or $(M_-)^k Y$ is a simultaneous eigenfunction of $M^2$ and $M_z$ for $M^2 = \lambda$ and $M_z = \mu - k\hbar$, where $k$ is any non-negative integer.

The next step in finding the eigenvalues and simultaneous eigenfunctions of $M^2$ and $M_z$ is to show that, corresponding to any particular simultaneous eigenfunction $Y$, there exist non-negative integers such that

(9.14a) $$(M_+)^{k_1+1} Y = 0, \text{ and}$$

(9.14b) $$(M_-)^{k_2+1} Y = 0.$$

To show this, we start with (9.3) which implies that

(9.15) $$\langle M^2 \rangle = \langle M_x{}^2 \rangle + \langle M_y{}^2 \rangle + \langle M_z{}^2 \rangle.$$

From this we readily see, since each of the terms on the right is non-negative, that

(9.16) $$\langle M^2 \rangle \geq \langle M_z{}^2 \rangle \geq 0.$$

For the simultaneous eigenfunctions $(M_+)^k Y$ and $(M_-)^k Y$ for which $M^2 = \lambda$ and $M_z = \mu \pm k\hbar$, this gives respectively

(9.17a) $$\lambda \geq (\mu + k\hbar)^2 \geq 0, \text{ and}$$

(9.17b) $$\lambda \geq (\mu - k\hbar)^2 \geq 0.$$

We get the two equations looked for in (9.14) by letting $k_1$ and $k_2$ be the greatest integers for which (9.17a) and (9.17b), respectively, are satisfied.

The next step is to obtain several relations between $\lambda$, $\mu$, $k_1$ and $k_2$ which enable us to determine the eigenvalues $\lambda$ and $\mu$. From (9.8) we see that

(9.18a) $M_-M_+\left(M_+^{k_1}Y\right) = M^2\left(M_+^{k_1}Y\right) - M_z^{\,2}\left(M_+^{k_1}Y\right) - \hbar M_z\left(M_+^{k_1}Y\right)$,

(9.18b) $M_+M_-\left(M_-^{k_2}Y\right) = M^2\left(M_-^{k_2}Y\right) - M_z^{\,2}\left(M_-^{k_2}Y\right) + \hbar M_z\left(M_-^{k_2}Y\right)$.

Using (9.14) and the eigenvalue relations for $M^2$ and $M_z$ we find that

(9.19a) $\lambda = (\mu + k_1\hbar)(\mu + k_1\hbar + \hbar)$, and

(9.19b) $\lambda = (\mu - k_2\hbar)(\mu - k_2\hbar - \hbar)$.

At this point, skipping some algebra, we eventually end up with

(9.20a) $\mu = (k_1 - k_2)\hbar/2 = k_0\hbar$, and

(9.20b) $\lambda = (\mu + k_1\hbar)(\mu + k_1\hbar + \hbar) = l(l+1)\hbar^2$.

Of the four integers $k_0$, $k_1$, $k_2$, and $l$, we may select any two (consistent with the requirements that $k_1 + k_2$ be an even integer and $k_1$ and $k_2$ are non-negative), and the other two can be determined from the equations $k_1 = l - k_0$ and $k_2 = l + k_0$. Two convenient choices are: (1) pick any $l \geq 0$ and let $k_2 = 0$, then $k_0 = -l$, $k_1 = 2l$, and $\mu = -l\hbar$; and (2) pick any $l \geq 0$ and let $k_1 = 0$, then $k_0 = l$, $k_2 = 2l$, and $\mu = l\hbar$. For the same $l$, the two choices lead to the same sequence of normalized eigenfunctions and eigenvalues of $M^2$ and $M_z$. The first starts with $M_z = -l\hbar$ and by $2l$ applications of $M_+$ leads to $M_z = l\hbar$. The second starts with $M_z = l\hbar$ and by $2l$ applications of $M_-$ leads to $M_z = -l\hbar$. The variable $l$ has special significance, and is called the *orbital quantum number.*

Let us look in more depth at the first choice, ascending the $M_z$ scale. We choose any $l \geq 0$ and let $k_2 = 0$. The system of equations

(9.21a) $M_z Y_l^{-l} = -l\hbar Y_l^{-l}$, and

(9.21b) $M_- Y_l^{-l} = 0$,

determines a simultaneous eigenfunction of $M^2$ and $M_z$ for $M^2 = l(l+1)\hbar^2$ and $M_z = -l\hbar$. Let $Y_l^{-l}$ be normalized, inner products and norms being defined by integration over the surface of the unit sphere. Define $Y_l^{-l+k}$ by

$$d_l^k Y_l^{-l+k} = M_+^k Y_l^{-l}, \tag{9.22}$$

where $d_l^k$ is a constant to retain normalization. $Y_l^{-l+k}$, where $k$ goes from 0 to $2l$, is a simultaneous eigenfunction of $M^2$ and $M_z$, $M^2 = l(l+1)\hbar^2$ and $M_z = (-l+k)\hbar$. Letting $m = -l + k$, for $k$ between 0 and $2l$, we may write the above equation as

$$d_l^{l+m} Y_l^m = M_+^{l+m} Y_l^{-l}, \tag{9.23}$$

for $-l \le m \le l$. Then, making the next step,

$$d_l^{l+m+1} Y_l^{m+1} = M_+^{l+m+1} Y_l^{-l} = M_+\left(M_+^{l+m} Y_l^{-l}\right) = M_+\left(d_l^{l+m} Y_l^{m+1}\right), \tag{9.24}$$

or

$$c_l^m Y_l^{m+1} = M_+ Y_l^m, \text{ for } -l \le m \le l. \tag{9.25}$$

The $Y_l^m$ are thus determined in sequence, which is why $M_+$ is called the raising operator, beginning with $Y_l^{-l}$ and ending with $Y_l^l$. These eigenfunctions are called surface spherical harmonics and are given explicitly in terms of associated Legendre polynomials. The derivation of such is not relevant to our discussion, and can be found in most undergraduate quantum mechanics texts.

*Problem 9.2:* Show that $Y_l^m$ is a simultaneous eigenfunction of $M_-M_+$ and $M_+M_-$ for $M_-M_+ = (l-m)(l+m+1)\hbar^2$ and $M_+M_- = (l+m)(l-m+1)\hbar^2$.[57]

---

[57] Ikenberry, p. 170, problem 14.19

## 9.4 Angular Momentum and Magnetic Moment

What follows for the next few sections is a non-classical development of magnetic quantum number and spin quantum numbers. We will later see how these two appear as a consequence of relativistic quantum mechanics.

Picture an electron as a particle of mass $m_0$, charge $e$, bound to a fixed nucleus by a central force arising from a Coulomb potential $V(r)$. The orbit of the electron is circular if there is equilibrium between the centripetal force $m_0 r\omega^2$ and the Coulombic attraction $Ze^2/r^2$. From the equality between these two forces we derive an expression for the classical frequency of orbital motion

$$\nu = \frac{\omega}{2\pi} = \frac{1}{2\pi}\sqrt{\frac{Ze^2}{m_0 r^3}}. \tag{9.26}$$

The mean electric current due to the motion of the electron in its orbit is $J = e/T$ where $T = 1/\nu$ is the period of the motion. According to the laws of electrodynamics, a closed current $J$ is equivalent in external effects to a magnetic moment

$$\mu_{\mathrm{M}} = \frac{|J|A}{c} = \frac{A|e|}{cT}, \tag{9.27}$$

where $A$ is the area enclose by the current, and $c$ is the velocity of light. We wish to find the relation between the magnentic moment and the angular momentum. According to classical mechanics $M$, which is perpendicular to the plane of the motion, is a constant of the motion, in particular, $M = m_0 r^2\dot{\phi}$ = a constant. Hence

$$A = \frac{1}{2}\int_0^{2\pi} r^2 d\phi = \frac{MT}{2m_0}. \tag{9.28}$$

Elimination of $A$ from (9.27) and (9.28) gives the proportionality between $\mu_{\mathrm{M}}$ and $M$

$$\mu_{\mathrm{M}} = \frac{eM}{2m_0 c}. \tag{9.29}$$

A magnetic field $H$ (or, more properly, magnetic induction $B$), interacting with a magnetic dipole of moment $\mu_{\mathrm{M}}$ produces a torque, and hence the causes the angular momentum $M$ to change at a rate

$$\frac{d}{dt}M = \mu_{\mathrm{M}} \times H = \frac{e}{2m_0 c} M \times H. \tag{9.30}$$

From this expression, we see that the time derivative of the angular momentum is perpendicular to both the angular momentum and the magnetic field. Hence, in the presence of a magnetic field, the angular momentum vector precesses around the direction of $H$ with a frequency that can be found to be

$$\omega_L = \frac{e}{2m_0c}H. \tag{9.31}$$

Classically, $\theta$, the angle between the directions of $M$ and $H$, may have any value between 0 and $\pi$, inclusive. Quantum mechanically, the only observable values of $M^2$ and $M_z$, are the eigenvalues $M^2 = l(l+1)\hbar^2$ and $M_z = m\hbar$, found in the last section. From these a discrete set of $2l+1$ values of $\theta$ is calculable from the formula

$$\cos\theta = \frac{M_z}{M} = \frac{m}{\sqrt{l(l+1)}}, -l \leq m \leq l. \tag{9.32}$$

The quantity $\frac{|e|\hbar}{2m_0c}$, which is the natural unit of magnetic moment, is called the *Bohr magneton.* The quantum number $m$, which is directly observable is called the *magnetic quantum number.*

*Problem 9.3:* Show that $(\Delta M_x)(\Delta M_y) \geq \frac{1}{2}m\hbar^2$ when $M_z = m\hbar$. [Hint: $M_x$ and $M_y$ do not commute, so you may use the Heisenberg Uncertainty Relation.][58]

## 9.5 The Stern-Gerlach Experiment

In the Stern-Gerlach experiment, a spatially inhomogeneous magnetic field serves to separate atoms for which values of $M_z$ for the valence electrons differ. This experiment would serve to confirm or deny the spatial quantization predicted by (9.32). When the experiment was performed in 1922, it did indeed confirm the spatial quantization, but there were some anomalous observations which we will look at in the next section.

The potential energy of a magnetic dipole of moment $\mu$ in a magnetic field $H$ is

$$U = -\mu \cdot H. \tag{9.33}$$

In an inhomogeneous field the dipole experiences a force

$$F = -\nabla U = \nabla(\mu \cdot H). \tag{9.34}$$

---

[58] Ikenberry, p. 173, problem 14. 25

To avoid considering the rapid variation of $F$ as $\mu$ precesses around $H$, we take the mean value of $F$. Since the mean value of $\mu$ is antiparallel to $H$, we obtain

(9.35) $$F = \nabla(\mu_z H) = \mu_z \nabla H,$$

when we make this replacement. Here $\mu_z$ is the component of $\mu$ in the direction of $H$. Finally, using (9.29) we find that the mean force on an atom due to the interaction between $H$ and the magnetic dipole produced by an orbital electron is

(9.36) $$F = \frac{e}{2m_0c} M_z \nabla H.$$

In the experiment, magnetic pole pieces are arranged to give a magnetic field $H$ which varies rapidly in the direction of $H$. In a typical experiment, $\partial H/\partial z = 250{,}000$ Gauss per centimeter. Particles injected at 90° to $H$ are deflected by a force whose component parallel to $H$ is

(9.37) $$F_z = KM_z,$$

where $K = \frac{e}{2m_0c}\frac{\partial H}{\partial z}$. From this we find, by integrating twice, that

(9.38) $$z = \frac{KM_z t^2}{2m_0},$$

where $z$ is the deflection in the direction of $H$ and $t$ is the time required for an atom to traverse the field.

Classically, $M_z$ and thus $z$ may have any value in a continuum. Hence a spreading out of the beam toward each side, with a gradual decrease in intensity from the maximum at the center, would be expected. This broadening into a single wide beam is not observed, rather there is observed a splitting into a definite number of distinct beams, the number depending upon the atomic species and its state of excitation. Quantum mechanically, $M_z$ has only discrete values, $M_z = m\hbar$ where $m$ is an integer. The number of beams into which the original beam is split gives the number of possible values of $m$. Further, the magnitude of the magnetic moment of orbital electrons can be calculated from the magnitude of the splitting and the constants of the apparatus.

## 9.6 Stern-Gerlach Patterns and Electron Spin

According to the theory developed to this point, the number of components to be expected in a Stern-Gerlach experiment is equal to the number of possible values of the magnetic quantum number $m$. For orbital angular momentum $M$ for which

$M^2 = l(l+1)\hbar^2$, the number of possible values of $m$ is $2l+1$, an odd integer. Thus, the pattern should always have an odd number of components and should always contain an undeflected component corresponding to $m = 0$. However, for many elements, notably H, Ag, and the alkalis, there are an even number of components, all of them deflected.

To account for this, and for the observed splitting of excited energy levels by a magnetic field (called the anomalous Zeeman effect), Uhlenbeck and Goudsmit were led in 1925 to propose that the "point electron" model should be replaced with a "spinning electron" model. Similar to the orbital angular momentum $M$ for which $M^2 = l(l+1)\hbar^2$, they proposed that an electron also has a spin angular momentum $|s|$ for which $|s|^2 = s(s+1)\hbar^2$. Note that the spin angular momentum is denoted by $|s|$ while the spin quantum number is denoted by $s$.

To discover the values of the spin quantum numbers, we must consider the elements for which the Stern-Gerlach experimental values did not agree with the predicted $2l+1$ values of $M_z$. For H, Ag, or an alkali in the ground state, for which $n = 1$, $l = 0$, and $m = 0$, we would expect one undisplaced component. However, two displaced components are observed. These two components must be explained entirely in terms of electron spin. Corresponding to the $2l+1$ values of $M_z$, let us suppose that there are $2s+1$ values of $s_z$, each of which accounts for one of the components in the experiment for H, Ag, or an alkali, in the ground state. Since there are two observed components, we write $2s+1 = 2$ and find that $s = \frac{1}{2}$. Then

(9.39a) $$|s|^2 = s(s+1)\hbar^2 = \tfrac{3}{4}\hbar^2, \text{ and}$$

(9.39b) $$s_z = \pm s\hbar = \pm\tfrac{1}{2}\hbar.$$

Measurement of the deflection of the components in a Stern-Gerlach pattern leads to the value of the magnetic moment being associated with the spin angular momentum. It is found to be roughly equal to the magnetic moment associated with the orbital angular momentum (a slight difference was observed in 1947 in very precise measurements by Lamb, and are now called the *Lamb shift*).

## 9.7 Pauli's Theory of Electron Spin, Spin Wavefunctions and Operators

We have seen how the results of the Stern-Gerlach experiment require the electron spin hypothesis. Other experimental evidence for its existence are the anomalous Zeeman effect, the existence of doublets in the spectra of alkalis, and the gyromagnetic effect.

The mathematical formalism for the treatment of the vector representing electron spin has been developed by modification of the mathematical relations for orbital angular momentum. The Pauli theory of electron spin is of importance because of its historical place (between Schrödinger nonrelativistic one-component wave theory and Dirac's relativistic four-component wave theory) and the introduction it provides to the study of Dirac's theory.

Wave functions describing electron spin may be introduced in various ways. Assuming that our knowledge of the state of an electron includes information about the position of the electron and about $s_z$, the component of the spin vector in some specified direction, we write

$$\psi = \psi(x,y,z,\sigma_z), \tag{9.40}$$

where $\sigma_z = \frac{2}{\hbar} s_z$. We have seen that $s_z$ is allowed to be $\pm\frac{\hbar}{2}$, so $\sigma_z$ is $\pm 1$. Since $\sigma_z$ is restricted to two values, we find it convenient to introduce the two-component wavefunction

$$\psi = \begin{pmatrix} \psi_+(x,y,z) \\ \psi_-(x,y,z) \end{pmatrix}, \text{ where} \tag{9.41}$$

$$\psi_\pm(x,y,z) = \psi(x,y,z,\pm 1). \tag{9.42}$$

The wave functions $\begin{pmatrix} \psi_+ \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ \psi_- \end{pmatrix}$ are to be interpreted that $\sigma_z$ is definitely +1 or -1 respectively. More generally, when there is a probability distribution for $\sigma_z$, we may normalize that so

$$\langle\psi|\psi\rangle = \langle\psi_+|\psi_+\rangle + \langle\psi_-|\psi_-\rangle = \int_\infty \left(\psi_+^*\psi_+ + \psi_-^*\psi_-\right) dr = 1. \tag{9.43}$$

Then $\langle\psi_+|\psi_+\rangle$ gives the probability that $\sigma_z = +1$ and $\langle\psi_-|\psi_-\rangle$ gives the probability that $\sigma_z = -1$.

Now, Hermitean operators with complete sets of orthonormal eigenfunctions are to be assigned to spin observables. We will take these operators to be $s^2$ and some component of $s$, say $s_z$. The eigenvalues of $s_z$ are the observed values $\pm\hbar/2$. Hence the eigenvalues of $\sigma_z$ are $+1$ and $-1$. Taking the states $\begin{pmatrix}1\\0\end{pmatrix}$ and $\begin{pmatrix}0\\1\end{pmatrix}$ to be those in which $\sigma_z$ are certainly $+1$ and $-1$, respectively, then we have

(9.44a) $$\sigma_z\begin{pmatrix}1\\0\end{pmatrix}=1\cdot\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}1\\0\end{pmatrix}, \text{ and}$$

(9.44b) $$\sigma_z\begin{pmatrix}0\\1\end{pmatrix}=-1\cdot\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}0\\-1\end{pmatrix}.$$

From this we can see that, in the two-component wavefunction formalism, $\sigma_z$ is to be represented by a $2\times2$ matrix

(9.45) $$\sigma_z=\begin{bmatrix}p & r\\ q & s\end{bmatrix}.$$

Substituting this into (9.44) we find

(9.46a) $$\sigma_z\begin{pmatrix}1\\0\end{pmatrix}=\begin{bmatrix}p & r\\ q & s\end{bmatrix}\cdot\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}p\\q\end{pmatrix}=\begin{pmatrix}1\\0\end{pmatrix}, \text{ and}$$

(9.46b) $$\sigma_z\begin{pmatrix}0\\1\end{pmatrix}=\begin{bmatrix}p & r\\ q & s\end{bmatrix}\cdot\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}r\\s\end{pmatrix}=\begin{pmatrix}0\\-1\end{pmatrix}.$$

Thus,

(9.47a) $$\sigma_z=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}, \text{ and}$$

(9.47b) $$\sigma^2=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\cdot\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=1.$$

To find $\sigma_x$ and $\sigma_y$, we need to consider $\alpha,\beta,\gamma$ as the direction cosines of a $z'$ axis in relation to an $xyz$ coordinate system. Thus we have

(9.48) $$\sigma_{z'}=\alpha\sigma_x+\beta\sigma_y+\gamma\sigma_z, \text{ and, hence}$$

(9.49) $$\sigma_{z'}{}^2=\alpha^2\sigma_x{}^2+\beta^2\sigma_y{}^2+\gamma^2\sigma_z{}^2+\alpha\beta(\sigma_x\sigma_y+\sigma_y\sigma_x)$$
$$+\beta\gamma(\sigma_y\sigma_z+\sigma_z\sigma_y)+\gamma\alpha(\sigma_z\sigma_x+\sigma_x\sigma_z).$$

From this and (9.47b), we readily see that

(9.50a) $$\sigma_x{}^2 = 1,$$

(9.50b) $$\sigma_y{}^2 = 1,$$

(9.50c) $$\sigma_z{}^2 = 1,$$

(9.50d) $$\sigma_x\sigma_y + \sigma_y\sigma_x = 0,$$

(9.50e) $$\sigma_y\sigma_z + \sigma_z\sigma_y = 0, \text{ and}$$

(9.50f) $$\sigma_z\sigma_x + \sigma_x\sigma_z = 0.$$

Thus the components of $\sigma$, and therefore of $s$, anticommute. We also find that

(9.51a) $$\sigma^2 = \sigma_x{}^2 + \sigma_y{}^2 + \sigma_z{}^2 = 3, \text{ so}$$

(9.51b) $$s^2 = \tfrac{3}{4}\hbar^2.$$

In analogy with the relation $M \times M = i\hbar M$ satisfied by the orbital angular momentum operators, assume that the spin angular momentum operator $s$ satisfies

(9.52a) $$s \times s = i\hbar s,$$

and consequently

(9.52b) $$\sigma \times \sigma = 2i\sigma.$$

Using this in a method very similar to that used to find $\sigma_z$, we find that the *Pauli spin matrices* are

(9.53a) $$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

(9.53b) $$\sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \text{ and}$$

(9.53c) $$\sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Since the Pauli spin matrices are Hermitean, their eigenvalues are real.

The total angular momentum $J$ is defined in terms of the orbital angular momentum $M$ and the spin angular momentum $s$, and is given in terms of vector addition

(9.54) $$J = M + s = M + \frac{\hbar}{s}\sigma,$$

where, in the two-component Pauli formalism, the components of M are represented by the $2 \times 2$ matrices

(9.55a) $$M_x = \begin{bmatrix} M_x & 0 \\ 0 & M_x \end{bmatrix},$$

(9.55b) $$M_y = \begin{bmatrix} M_y & 0 \\ 0 & M_y \end{bmatrix}, \text{ and}$$

(9.55c) $$M_z = \begin{bmatrix} M_z & 0 \\ 0 & M_z \end{bmatrix}.$$

The addition in (9.54) then becomes simple matrix addition. It is also important to note that spin-independent operators, such as those for the orbital angular momentum, commute with the spin operators.

*Problem 9.4:* Using the Pauli spin matrices, obtain a matrix representation of $\sigma_{z'}$.[59]

Answer: $$\sigma_{z'} = \begin{bmatrix} \gamma & \alpha - \beta i \\ \alpha + \beta i & -\gamma \end{bmatrix}.$$

*Problem 9.5:* Obtain the eigenvalues and normalized eigenfunctions of $\sigma_x$ in the Pauli representation.[60]

Answer: $$c = \pm 1, p = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, q = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

*Problem 9.6:* Using the Pauli matrices, obtain a matrix representation of

$$M \cdot \sigma = M_x\sigma_x + M_y\sigma_y + M_z\sigma_z.$$[61]

Answer: $$M \cdot \sigma = \begin{bmatrix} M_z & M_- \\ M_+ & -M_z \end{bmatrix}.$$

---

[59] Ikenberry, p. 225, problem 18.5

[60] Ikenberry, p. 226, problem 18.7

[61] Ikenberry, p. 226, problem 18.10

## 9.8 Inner Products and Expectation Values in 2-Component Formalism

If we consider two two-component functions $f = \begin{pmatrix} f_+ \\ f_- \end{pmatrix}$ and $g = \begin{pmatrix} g_+ \\ g_- \end{pmatrix}$, it is easily seen that the inner product, defined by

(9.56) $$\langle f \,|\, g \rangle = \langle f_+ \,|\, g_+ \rangle + \langle f_- \,|\, g_- \rangle = \int_\infty \left( f_+{}^* g_+ + f_-{}^* g_- \right) dr,$$

satisfies the requirements in (1.7). In the case that the two components of $f$ depend on $xyz$ in the same way, and the two components of $g$ depend on $xyz$ in the same way, that is,

(9.57) $$f = f_1(x,y,z) \begin{pmatrix} c_+ \\ c_- \end{pmatrix}, \text{ and } g = g_1(x,y,z) \begin{pmatrix} d_+ \\ d_- \end{pmatrix},$$

then the inner-product becomes

(9.58) $$\langle f \,|\, g \rangle = \left( c_+{}^* d_+ + c_-{}^* d_- \right) \int_\infty f_1{}^* g_1 dr.$$

Then, for spin vectors $u = \begin{pmatrix} c_+ \\ c_- \end{pmatrix}$ and $v = \begin{pmatrix} d_+ \\ d_- \end{pmatrix}$ in spin space, we may define the inner product $\langle u | v \rangle$ by

(9.59) $$\left( c_+{}^* d_+ + c_-{}^* d_- \right).$$

Two spin vectors $u$ and $v$ are orthogonal if $\langle u | v \rangle = 0$, and the norm squared of a function in spin space is $\langle u | u \rangle$.

In a state represented by a spin vector $u$, the expectation value of a spin observable $A$ is

(9.60) $$\langle A \rangle = \frac{\langle u | A | u \rangle}{\langle u | u \rangle},$$

where, on the right side, $A$ is the matrix representation of the spin operator $A$.

If $u_1$ and $u_2$ are orthonormal eigenfunctions of a Hermitean spin operator $A$, and $a_1$ and $a_2$ are their respective eigenvalues, then any two-component spin function $u$ may be represented as

(9.61) $$u = c_1 u_1 + c_2 u_2,$$

where

(9.61a) $$c_1 = \langle u_1 | u \rangle,$$

(9.61b) $$c_2 = \langle u_2 | u \rangle, \text{ and}$$

(9.61c) $$\langle u | u \rangle = |c_1|^2 + |c_2|^2 = c^2.$$

The probability of being in either eigenstate is

(9.62a) $$P(a_1) = |c_1|^2 / c^2, \text{ and}$$

(9.62b) $$P(a_2) = |c_2|^2 / c^2.$$

*Problem 9.7:* Calculate the expectation value and the probability distribution for $\sigma_x$ for a state in which $\sigma_z = -1$.[62]

Answers: $\langle \sigma_x \rangle = 0, P(-1) = \frac{1}{2}, P(1) = \frac{1}{2}.$

## 9.9 Eigenfunctions of Hydrogen for the Stern-Gerlach Experiment

In a Stern-Gerlach experiment with atomic hydrogen, atoms in the first or second excited states are separated according to the values of a component of $J = M + s$, call it $J_z$. Rather than simultaneous eigenfunctions of $H = \frac{p^2}{2m} + \frac{e^2}{r}$, $M^2$, $M_z$, and $s_z$ we are interested in simultaneous eigenfunctions of $H$, $J^2$, $M^2$, and $J_z$. We can easily see that the two-component wavefunctions are not eigenfunctions of $J^2$, since

(9.63) $$J^2 = \left(M + \tfrac{\hbar}{2}\sigma\right)^2 = M^2 + \hbar M \cdot \sigma + \tfrac{3\hbar^2}{4},$$

is a mixed operator.

We instead take our complete set of commuting operators to be $H, J, M^2$, and $J_z$. Calculations could be shortened by also including $M \cdot \sigma$ when we look for simultaneous eigenfunctions. Starting with $M^2$ and $J_z$ we must have for some $\lambda$ and $\mu$,

(9.64a) $$M^2 \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} = \lambda \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix}, \text{ and}$$

(9.64b) $$J_z \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} = \begin{pmatrix} M_z + \frac{\hbar}{2} & 0 \\ 0 & M_z - \frac{\hbar}{2} \end{pmatrix} \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} = \mu \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix}.$$

This gives the system of four equations

(9.65a) $$M^2 \psi_+ = \lambda \psi_+,$$

---

[62] Ikenberry, p. 227, problem 18.11

(9.65b)
$$M^2\psi_- = \lambda\psi_-,$$

(9.65c)
$$M_z\psi_+ = \left(\mu - \tfrac{\hbar}{2}\right)\psi_+, \text{ and}$$

(9.65d)
$$M_z\psi_- = \left(\mu + \tfrac{\hbar}{2}\right)\psi_-.$$

Hence $\psi_+$ and $\psi_-$ are eigenfunctions of $M^2$ for $M^2 = \lambda$ and of $M_z$ for $M_z = \left(\mu - \tfrac{\hbar}{2}\right)$ and $M_z = \left(\mu + \tfrac{\hbar}{2}\right)$, respectively. Since the eigenvalues of $M^2$ and of $M_z$ are $l(l+1)\hbar^2$ and $m\hbar$, we must take $\lambda = l(l+1)\hbar^2$ and $\mu - \tfrac{\hbar}{2} = m\hbar$, $\mu + \tfrac{\hbar}{2} = (m+1)\hbar$, $\mu = \left(m + \tfrac{1}{2}\right)\hbar$, for some $m$, $-l \le m \le l$. Then

(9.66)
$$\psi = \begin{pmatrix} f_+(r)Y_l^m(\theta,\phi) \\ f_-(r)Y_l^{m+1}(\theta,\phi) \end{pmatrix},$$

where $f_+(r)$ and $f_-(r)$ remain to be determined, is a simultaneous eigenfunction of $M^2$ and $J_z$ for $M^2 = l(l+1)\hbar^2$ and $J_z = \left(m + \tfrac{1}{2}\right)\hbar$.

A relation between $f_+(r)$ and $f_-(r)$ may be obtained by requiring that (9.66) be an eigenfunction of $M \cdot \sigma$. Writing $M \cdot \sigma\psi = u\hbar\psi$ and using the solution to problem 9.6, we have

(9.67)
$$\begin{bmatrix} M_z & M_- \\ M_+ & -M_z \end{bmatrix}\begin{pmatrix} f_+Y_l^m \\ f_-Y_l^m \end{pmatrix} = u\hbar\begin{pmatrix} f_+Y_l^m \\ f_-Y_l^m \end{pmatrix}.$$

Writing this out as two equations, taking into account the fact that the components of the matrix all commute with $f_+(r)$ and $f_-(r)$, and using the standard eigenvalues of the momentum operators, we obtain

(9.68a)
$$m\hbar f_+ + c_l^m f_- = u\hbar f_+, \text{ and}$$

(9.68b)
$$c_l^m f_+ + (m+1)\hbar f_- = u\hbar f_-.$$

These are only consistent if $u = l$ or $u = -l - 1$.

For the case $u = l$, we find that the two-component wavefunction

(9.69)
$$\psi_1 = g_1(r)\begin{pmatrix} \sqrt{l+m+1}\, Y_l^m(\theta,\phi) \\ \sqrt{l-m}\, Y_l^{m+1}(\theta,\phi) \end{pmatrix}$$

is a simultaneous eigenfunction of $J^2 = j(j+1)\hbar^2$, $M^2 = l(l+1)\hbar^2$, and $J_z = \left(m + \frac{1}{2}\right)\hbar$, where $j = l + \frac{1}{2}$ and $-l \leq m \leq l$.

For the case $u = -l - 1$, we find that the two-component wavefunction

(9.70)
$$\psi_2 = g_2(r)\begin{pmatrix} \sqrt{l-m}\,Y_l^m(\theta,\phi) \\ -\sqrt{l+m+1}\,Y_l^{m+1}(\theta,\phi) \end{pmatrix}$$

is a simultaneous eigenfunction of $J^2 = j(j+1)\hbar^2$, $M^2 = l(l+1)\hbar^2$, and $J_z = \left(m + \frac{1}{2}\right)\hbar$, where $j = l - \frac{1}{2}$ and $-l \leq m \leq l$.

For the Hamiltonian $H$, we take the form

(9.71)
$$H = \frac{M^2}{2m_0 r^2} + \frac{1}{2m_0 r^2}\left(r\, p_r^{\,2}\, r\right) + V(r),$$

which the radial form for the Hamiltonian of the hydrogen atom. Note that $H$ is an even operator and, since the two two-component wave functions are eigenfunctions of $M^2 = l(l+1)\hbar^2$, we see that they are also eigenfunctions of $H$.

## 9.10 The Normal Zeeman Effect in the Pauli Theory

On the basis of classical theory, we expect that the frequencies of spectral lines correspond to frequencies of rotation of electrons in their orbits. These frequencies though, are affected by magnetic fields. The frequency shifts of spectral lines by a magnetic field is what is called the *Zeeman effect*. In the *normal* Zeeman effect, the observed shifts are in full agreement with classical theory and with the Schrödinger wave theory. However, in a far greater number of cases, an *anomalous* effect is observed, which can only be accurately predicted by inclusion of the spin angular momentum or with Dirac's relativistic wave equation.

An explanation of the anomalous Zeeman effect requires that interaction between the spin and orbital magnetic momenta be taken into account. In this section we ignore this interaction, but include the interaction between the spin magnetic moment and an external magnetic field $B$. This may be done by adding to the Hamiltonian for a hydrogen atom in a magnetic field a term giving the effect of the spin on the energy. We write the energies of interaction between the magnetic field $B$ and the orbital and spin magnetic moments $\mu_M$ and $\mu_S$ as

(9.72a)
$$U_M = -\mu_M \cdot B = GM \cdot B, \text{ and}$$

(9.72b) $$U_s = -\mu_s \cdot B = G_s\, s \cdot B,$$

where $G = -e/2m_0c$ and $G_s = -e/m_0c$. We then have

(9.73) $$H = H_0 + GM \cdot B + G_s\, s \cdot B = H^{(1)} + G_s\, s \cdot B,$$

where

(9.74) $$H^{(1)} = H_0 + GM \cdot B = \frac{p^2}{2m_0} + eV + GM \cdot B.$$

Now, in the Pauli two-component wave theory, any spin independent operator, such as $H^{(1)}$, is represented by an even $2 \times 2$ matrix

(9.75) $$H^{(1)} = \begin{bmatrix} H^{(1)} & 0 \\ 0 & H^{(1)} \end{bmatrix}.$$

Replacing $M$ by $B$ in problem 9.6, we also have

(9.76) $$s \cdot B = \frac{\hbar}{2}\sigma \cdot B = \frac{\hbar}{2}\begin{bmatrix} B_z & B_- \\ B_+ & -B_z \end{bmatrix},$$

where $B_\pm = B_x \pm iB_y$. Combining these two by (9.64) gives us

(9.77) $$H = \begin{bmatrix} H^{(1)} - \frac{e\hbar}{2m_0c}B_z & B_- \\ B_+ & H^{(1)} + \frac{e\hbar}{2m_0c}B_z \end{bmatrix}.$$

To simplify further calculations, we will consider the $x$ axis as parallel to $B$. Then $B_x = 0$, $B_y = 0$, $B_+ = 0$, $B_- = 0$, and $B_z = B$. In this coordinate system the Hamiltonian is an even operator since (9.68) becomes

(9.78) $$H = \begin{bmatrix} H^{(1)} - \frac{e\hbar}{2m_0c}B & 0 \\ 0 & H^{(1)} + \frac{e\hbar}{2m_0c}B \end{bmatrix}.$$

For this Hamiltonian the energy eigenvalue equation $H\psi = E\psi$ separates into two independent equations, since $\psi$ is an two-component wavefunction. These two are

(9.79a) $$\left(H^{(1)} - \frac{e\hbar}{2m_0c}B\right)\psi_+ = E\psi_+, \text{ and}$$

(9.79b) $$\left(H^{(1)} + \frac{e\hbar}{2m_0c}B\right)\psi_- = E\psi_-.$$

Rewriting these as

(9.80a) $$H^{(1)}\psi_+ = \left(E + \frac{e\hbar}{2m_0c}B\right)\psi_+ \text{ and}$$

(9.80b) $$H^{(1)}\psi_- = \left(E - \frac{e\hbar}{2m_0c}B\right)\psi_-,$$

we see that both $\psi_+$ and $\psi_-$ are eigenfunctions of $H^{(1)}$, but for different eigenvalues.

For the hydrogen atom, the simultaneous eigenfunction of $H^{(1)}$, $M^2$, and $M_z$, for $H^{(1)} = E_{nm}$, $M^2 = l(l+1)\hbar^2$, and $M_z = m\hbar$ is denoted $\psi_{nl}^m$. But since we are working in Pauli space, this will be a two-component wavefunction, equal to

(9.81a) $$\psi_{+nl}^{\ m} = \begin{pmatrix} \psi_{nl}^m \\ 0 \end{pmatrix}, \text{ and}$$

(9.81b) $$\psi_{-nl}^{\ m} = \begin{pmatrix} 0 \\ \psi_{nl}^m \end{pmatrix}.$$

The first is the eigenfunction of $H = E_{nm} - e\hbar B/2m_0c$ and $s_z = \hbar/2$, and the second is the eigenfunction of $H = E_{nm} + e\hbar B/2m_0c$ and $s_z = -\hbar/2$, and they are both eigenfunctions for $M^2 = l(l+1)\hbar^2$ and $M_z = m\hbar$.

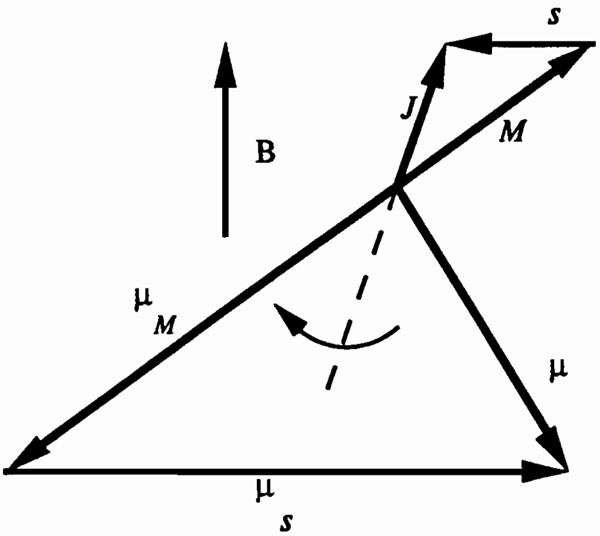
## 9.11 The Anomalous Zeeman Effect

In the anomalous Zeeman effect, a magnetic field splits spectral lines into more than the three components observed in the normal Zeeman effect. The electron spin hypothesis that we have been dealing with was introduced just to explain this effect.

In addition to the orbital magnetic moment associated with the orbital angular momentum $M$, an electron is attributed an intrinsic magnetic moment associated with its spin angular momentum $s$. In the last section, we looked for a solution while ignoring the interaction between the magnetic moments $\mu_M$ and $\mu_s$. However, except in the presence of a very strong magnetic field, this coupling cannot be ignored. It is found that this coupling can be properly accounted for by a very simple vector model relating the orbital and spin angular moments and magnetic momenta. The model (shown below) is a rigid vector structure in which $M$ and $s$ precess around $J = M + s$ so rapidly that the only effective part of $\mu$ is its component along -$J$. Then we can write

(9.82) $$U = -\mu \cdot B = GM \cdot B + G_s s \cdot B = gGJ \cdot B,$$

where $gGJ = \mu\cos(\mu, J)$ is the component of $\mu$ along -$J$. The little $g$ is called the *Landé g factor*, and was discovered before the introduction of electron spin.



To evaluate $g$ with $G_s = 2G$, we write

$$gGJ = -\mu \cdot J = \left(GM + G_s s\right) \cdot J = G(M + 2s) \cdot (M + s) \tag{9.83}$$
$$= G\left(M^2 + 3M \cdot s + 2|s|^2\right).$$

Eliminating $M \cdot s$ by using $J^2 = M^2 + 2M \cdot s + 2|s|^2$, we get

$$g = \left(3J^2 - M^2 + |s|^2\right)/2J^2. \tag{9.84}$$

Upon introducing the eigenvalues for $J^2$, $M^2$, and $|s|^2$, we get the standard form for $g$:

$$g = 1 + \frac{j(j+1) + s(s+1) - l(l+1)}{2j(j+1)}. \tag{9.85}$$

For the hydrogen atom, $j = l \pm \frac{1}{2}$ ($j = \frac{1}{2}$ when $l = 0$) and $s = \frac{1}{2}$.

# Chapter 10

# Relativistic Quantum Mechanics

*This chapter looks at relativistic corrections to Schrödinger's wave equation, and the short-comings of this approach. It then deals extensively with Dirac's relativistic wave equations, which involve matrices of order 4. It also looks at some of the consequences of Dirac's formulation of quantum theory.*

## 10.1 The Klein-Gordon Equation

Schrödinger's discovery in 1925 of a non-relativistic wave equation led immediately to the suggestion that a relativistic wave equation could be obtained by making the substitutions

(10.1a) $p \leftrightarrow i\hbar\nabla$, (which implies

(10.1b) $p^2 \leftrightarrow -\hbar^2\nabla^2$), and

(10.1c) $E \leftrightarrow i\hbar\frac{\partial}{\partial t}$

into the relativistic energy-momentum relation

(10.2) $$(E - e\Phi)^2 = (cp - eA)^2 + m_0{}^2c^4.$$

This substitution gives the Klein-Gordon Equation

(10.3) $$\left(-\hbar^2\frac{\partial^2}{\partial t^2} - 2ie\hbar\Phi\frac{\partial}{\partial t} - ie\hbar\frac{\partial\Phi}{\partial t} + e^2\Phi^2\right)\psi$$
$$= \left(-\hbar^2c^2\nabla^2 + 2i\hbar ceA\cdot\nabla + i\hbar ce(\nabla\cdot A) + e^2A^2 + m_0{}^2c^4\right)\psi.$$

This was developed in more detail in chapter 6, section 3.

The Klein-Gordon equation fails to account for electron spin and the formalism of nonrelativistic quantum mechanics can not be applied since the Hamiltonian operator no longer plays a central role. For a time the equation was ignored because of the difficulty of interpreting solutions, but since Dirac theory has been formulated, it has been noted that it is accurate and reasonable for particles with no spin.

## 10.2 Dirac's Relativistic Wave Equation

In order to obtain a relativistic wave equation which avoids the difficulties encountered with the Klein-Gordon equation, we might solve the relativistic energy-momentum relation for E:

$$E = c\sqrt{\left(p - \tfrac{e}{c}A\right)^2 + m_0c^2} + V, \tag{10.4}$$

where $V = e\phi$, and then introduce the operator representations (10.1) for $p$ and $E$. The resulting equation

$$c\sqrt{\left(i\hbar\nabla - \tfrac{e}{c}A\right)^2 + m_0c^2}\,\psi + V\psi = E\psi, \tag{10.5}$$

is of the first order in differentiation with respect to time. Thus, one of the earlier objections to the Klein-Gordon equation has been eliminated. The square-root operator still causes some difficulty, but progress has been made since this may be written in the form of Schrödinger's equation by letting the Hamiltonian be

$$H = c\sqrt{\pi^2 + m_0c^2} + V, \text{ where} \tag{10.6}$$

$$\pi = p - \tfrac{e}{c}A. \tag{10.6a}$$

This suggests that the formalism of nonrelativistic quantum mechanics might still be used, provided some means were discovered to interpret the square-root operator.

Dirac accomplished this in 1928 by writing

$$\sqrt{\pi^2 + m_0c^2} = \alpha\cdot\pi + m_0c\beta, \tag{10.7}$$

where $\alpha$ and $\beta$ are not ordinary numbers. We will investigate their nature next.

Squaring both sides of (10.7), and assuming that $\beta$ and the components of $\alpha$ commute with the components of $\pi$, but not with each other, we obtain (upon setting $A = 0, \pi = p$ for brevity of writing only):

$$\begin{aligned}(10.8)\; p_x^2 + p_y^2 + p_z^2 + m_0c^2 &= \alpha_1^2p_x^2 + \alpha_2^2p_y^2 + \alpha_3^2p_z^2 + m_0^2c^2\beta^2\\ &+(\alpha_1\alpha_2 + \alpha_2\alpha_1)p_xp_y + (\alpha_2\alpha_3 + \alpha_3\alpha_2)p_yp_z + (\alpha_3\alpha_1 + \alpha_1\alpha_3)p_zp_x\\ &+m_0c\left\{(\alpha_1\beta + \beta\alpha_1)p_x + (\alpha_2\beta + \beta\alpha_2)p_y + (\alpha_3\beta + \beta\alpha_3)p_z\right\}\end{aligned}$$

This is an identity if

$$\alpha_i^2 = 1, \tag{10.9a}$$

$$\alpha_i\alpha_j + \alpha_j\alpha_i = 0, \tag{10.9b}$$

$$\alpha_i\beta + \beta\alpha_i = 0, \text{ and} \tag{10.9c}$$

(10.9d) $$\beta^2 = 1.$$

The next step is to develop Dirac's matrix representations for $\alpha$ and $\beta$. In this representation the Pauli spin matrices play a central role.

### 10.3 The Dirac Matrices

Dirac found that the relations in (10.9) can be satisfied by matrices of order 4, but not by matrices of lower order. We start with the Pauli spin matrices and the identity matrix $I_2$:

(10.10a) $$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

(10.10b) $$\sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix},$$

(10.10c) $$\sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \text{ and}$$

(10.10d) $$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

It can be found that $\beta$ and the components of $\alpha$ depend on these, and are

(10.11a) $$\alpha_1 = \begin{bmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

(10.11b) $$\alpha_2 = \begin{bmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{bmatrix},$$

(10.11c) $$\alpha_3 = \begin{bmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \text{ and}$$

(10.11d)
$$\beta=\begin{bmatrix} I_2 & 0 \\ 0 & -I_2 \end{bmatrix}=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

We may readily verify that the conditions of (10.9) are satisfied by these matrices. Notice also that each of the matrices are Hermitean.

## 10.4 The Dirac Equations

From (10.6) and (10.7) we see that the Hamiltonian $H$ in Dirac's relativistic wave equation $H\psi = E\psi$ is

(10.12)
$$H = c\alpha \cdot \pi + m_0 c^2 \beta + V.$$

By means of (10.11) we find that

(10.13)
$$\alpha \cdot \pi = \begin{bmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{bmatrix}\pi_x + \begin{bmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{bmatrix}\pi_y + \begin{bmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{bmatrix}\pi_z = \begin{bmatrix} 0 & \sigma\cdot\pi \\ \sigma\cdot\pi & 0 \end{bmatrix}.$$

Next, using the Pauli spin matrices and $I_2$ we find

(10.14)
$$\sigma \cdot \pi = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\pi_x + \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\pi_y + \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\pi_z = \begin{bmatrix} \pi_z & \pi_- \\ \pi_+ & -\pi_z \end{bmatrix},$$

where $\pi_\pm = \pi_x \pm i\pi_y$. Substituting this into (10.13) we obtain

(10.15)
$$\alpha \cdot \pi = \begin{bmatrix} 0 & 0 & \pi_z & \pi_- \\ 0 & 0 & \pi_+ & -\pi_z \\ \pi_z & \pi_- & 0 & 0 \\ \pi_+ & -\pi_z & 0 & 0 \end{bmatrix}.$$

In (10.12) $V$ is to be represented by a $4 \times 4$ scalar matrix in which the only non-zero elements are $V$'s along the main diagonal. Then, substituting into (10.15) and recalling the form of $\beta$, we get

(10.16)
$$\frac{H}{c} = \begin{bmatrix} \frac{V}{c} + m_0 c & 0 & \pi_z & \pi_- \\ 0 & \frac{V}{c} + m_0 c & \pi_+ & -\pi_z \\ \pi_z & \pi_- & \frac{V}{c} - m_0 c & 0 \\ \pi_+ & -\pi_z & 0 & \frac{V}{c} - m_0 c \end{bmatrix}.$$

Sometimes it is more convenient to write this as a $2 \times 2$ matrix whose elements are themselves $2 \times 2$ matrices. In this representation we would have

$$\frac{H}{c} = \begin{bmatrix} \left(\frac{V}{c} + m_0 c\right) I_2 & \sigma \cdot \pi \\ \sigma \cdot \pi & \left(\frac{V}{c} - m_0 c\right) I_2 \end{bmatrix}. \tag{10.17}$$

Now, because Dirac operators are $4 \times 4$ matrices, Dirac wavefunctions must be 4-vectors. Thus, the wave equation becomes

$$\begin{bmatrix} \frac{V}{c} + m_0 c & 0 & \pi_z & \pi_- \\ 0 & \frac{V}{c} + m_0 c & \pi_+ & -\pi_z \\ \pi_z & \pi_- & \frac{V}{c} - m_0 c & 0 \\ \pi_+ & -\pi_z & 0 & \frac{V}{c} - m_0 c \end{bmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = \frac{E}{c} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}. \tag{10.18}$$

This is equivalent to a system of four equations, whose simultaneous solution will provide the four components of the wavefunction $\psi$.

This could also be written using (10.17), in which case $\psi$ would be a 2-vector whose components are 2-vectors:

$$\psi = \begin{pmatrix} \phi \\ \varphi \end{pmatrix}, \phi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \varphi = \begin{pmatrix} \psi_3 \\ \psi_4 \end{pmatrix}. \tag{10.19}$$

The wave equation is then written as

$$\begin{bmatrix} \left(\frac{V}{c} + m_0 c\right) I_2 & \sigma \cdot \pi \\ \sigma \cdot \pi & \left(\frac{V}{c} - m_0 c\right) I_2 \end{bmatrix} \begin{pmatrix} \phi \\ \varphi \end{pmatrix} = \frac{E}{c} \begin{pmatrix} \phi \\ \varphi \end{pmatrix}, \tag{10.20}$$

which would be equivalent to a system of two equations which could be solved simultaneously for the two components of $\psi$, $\phi$ and $\varphi$.

In concluding this section, we will show that if a four-component wavefunction $\psi$ satisfies Dirac's wave equation for a free particle, then each component of $\psi$ satisfies the Klein-Gordon equation for a free particle (10.3) with $A = 0, \phi = 0$. The elimination necessary to prove this is most easily done by operating on both sides of Dirac's equation in the form $c(\alpha \cdot p + m_0 c \beta)\psi = E\psi$ with the operator $c(\alpha \cdot p + m_0 c \beta)$. Recalling that $c(\alpha \cdot p + m_0 c \beta)$ commutes with $E$, we obtain

$$c^2 (\alpha \cdot p + m_0 c \beta)^2 \psi = cE(\alpha \cdot p + m_0 c \beta)\psi. \tag{10.21}$$

Substituting from (10.7) where $\pi = p$ into the the left hand side of (10.21) and from our formulation of the equation into the right give us

(10.22) $$c^2\left(p^2+m_0{}^2c^2\right)\psi=E^2\psi.$$

Since these are scalar matrices, this is equivalent to the four equations

(10.23) $$c^2\left(p^2+m_0{}^2c^2\right)\psi_i=E^2\psi_i, i=1,2,3,4,$$

each of which is Klein-Gordon's equation.

## 10.5 Pauli's Theory as a Nonrelativistic Approximation to Dirac's Theory

Upon substituting

(10.24a) $$\psi=e^{-im_0c^2t/\hbar}\psi' \text{ and}$$

(10.24b) $$E\psi=i\hbar\frac{\partial\psi}{\partial t}=e^{-im_0c^2t/\hbar}\left(E+m_0c^2\right)\psi',$$

where $\psi$ and $\psi'$ are four-component wave functions, into the Dirac wave equation we obtain

(10.25) $$H\psi'=\left(E+m_0c^2\right)\psi'.$$

Now, using (10.17) as the form of the Dirac Hamiltonian, we write this as

(10.26) $$\begin{bmatrix}\left(\frac{V}{c}+m_0c\right)I_2 & \sigma\cdot\pi\\ \sigma\cdot\pi & \left(\frac{V}{c}-m_0c\right)I_2\end{bmatrix}\begin{pmatrix}\phi'\\ \varphi'\end{pmatrix}=\left(\frac{E}{c}+m_0c\right)\begin{pmatrix}\phi'\\ \varphi'\end{pmatrix}.$$

This is equivalent to

(10.27a) $$\left(\frac{V-E}{c}\right)\phi'+\sigma\cdot\pi\varphi'=0, \text{ and}$$

(10.27b) $$\sigma\cdot\pi\phi'+\left(\frac{V-E}{c}-2m_0c\right)\varphi'=0.$$

Assuming now that $2m_0c^2 >> E-V$, we get

(10.28a) $$\sigma\cdot\pi\varphi'=\left(\frac{E-V}{c}\right)\phi' \text{ and}$$

(10.28b) $$\sigma\cdot\pi\phi'=2m_0c\varphi'$$

as the nonrelativistic approximation to (10.27). Eliminating $\varphi'$ from this, we find that $\phi'$ must satisfy

(10.29) $$\frac{1}{2m_0}(\sigma\cdot\pi)^2\phi'+V\phi'=E\phi',$$

in the nonrelativistic limit. Substituting in an alternate form for $(\sigma\cdot\pi)^2$ we get

(10.30) $$\left(\frac{1}{2m_0}\pi^2+V+G_s s\cdot B\right)\phi'=E\phi'.$$

The Hamiltonian now contains a spin dependent part, in addition to the spin independent part, in full agreement with the Pauli theory. Hence, in the nonrelativistic limit, the Dirac four-component wave theory reduces to the Pauli two-component wave theory. Unfortunately the spin-orbit interaction is lost in the process of taking the nonrelativistic limit.

## 10.6 Plane Wave Solutions for a Free Dirac Particle

For the Dirac equation for a free particle,

(10.31) $$H\psi = i\hbar\frac{\partial\psi}{\partial t}, \text{ where}$$

(10.31b) $$H = c\alpha\cdot p + m_0c^2\beta,$$

we try to get solutions of the form

(10.32) $$\psi(r,t) = f(t)\psi(r),$$

where $\psi$ is a four-component time independent wave function. Separation of the time variable leads to

(10.33a) $$H\psi(r) = E\psi(r) \text{ and}$$

(10.33b) $$i\hbar\frac{\partial f(t)}{\partial t} = Ef(t),$$

where $E$ is a separation constant. Solving this, we find that from any four-component energy eigenfunction $\psi$ and energy eigenvalue $E$, we can form a stationary state wave function

(10.34) $$\psi(r,t) = e^{-iEt/\hbar}\,\psi(r)\cdot$$

We proceed to determine the simultaneous eigenfunctions of $p$ and $H$, which are readily verified to be commuting operators. Starting with $p$, we find that the eigenvalue equation

(10.35) $$p\psi = \hbar k\psi,$$

where $p = -i\hbar\nabla$, has solutions

(10.36) $$\psi = Be^{ik\cdot r},$$

Where $B$ is a four-component function with constant components.

Substituting this into (10.18), in which we let $V = 0$ and $\pi = -i\hbar\nabla$, we find that the energy eigenvalue equation becomes

$$\begin{bmatrix} m_0c & 0 & p_z & p_- \\ 0 & m_0c & p_+ & -p_z \\ p_z & p_- & -m_0c & 0 \\ p_+ & -p_z & 0 & -m_0c \end{bmatrix} \begin{pmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \end{pmatrix} = \frac{E}{c} \begin{pmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \end{pmatrix}, \tag{10.37}$$

where now $p = \hbar k$ represents the observable $p$, rather than the differential operator $p = -i\hbar\nabla$.

The system of four linear, homogeneous equations has a solution other than $B = 0$ if and only if $p$ and $E$ are related so that the determinant of $E$ is $0$. To avoid the labor of expanding the determinant, recall that any solution of the Dirac equation for a free particle satisfies the Klein-Gordon equation for a free particle. Substituting any component

$$\psi_j(r,t) = e^{i(p \cdot r - Et)/\hbar} B_j, j = 1,2,3,4, \tag{10.38}$$

into the Klein-Gordon equation (10.3) where $A = 0$ and $\phi = 0$, we find that

$$E^2 = c^2p^2 + m_0{}^2c^4. \tag{10.39}$$

Thus $E$ and $p$ must satisfy the relativistic energy-momentum relation, which is to be expected.

In observing that, in order to obtain a complete set of eigenfunctions, the negative as well as the positive roots of the right-hand side of (10.39) must be included as energy eigenvalues, Dirac was led to predict the existence of particles called *positrons*. This theory will be discussed in the next section, but we will continue now to complete the calculations for both the positive and the negative energy solutions.

From (10.39) we must have $E = \pm\varepsilon$, where $\varepsilon = \sqrt{c^2p^2 + m_0{}^2c^4}$. For the positive energy solutions, we write $E = +\varepsilon$ in (10.37) and obtain the equations

$$\left(m_0c - \tfrac{\varepsilon}{c}\right)B_1 + p_zB_3 + p_-B_4 = 0, \tag{10.40a}$$

$$\left(m_0c - \tfrac{\varepsilon}{c}\right)B_2 + p_+B_3 - p_zB_4 = 0, \tag{10.40b}$$

$$p_zB_1 + p_-B_2 - \left(m_0c + \tfrac{\varepsilon}{c}\right)B_3 = 0, \tag{10.40c}$$

$$p_+B_1 - p_zB_2 - \left(m_0c + \tfrac{\varepsilon}{c}\right)B_4 = 0. \tag{10.40d}$$

For e as given, it may be verified that every third-order determinant formed from the coefficients in (10.40) vanishes. That is, the matrix of the coefficients is of rank two, and hence any two of the $B$'s may be expressed in terms of the other two. For example, we could choose to solve $a,b$ for $B_1$ and $B_2$. Or, we could solve $c,d$ for $B_3$ and $B_4$.

The first choice involves a division by a smaller number than the second choice, so $B_1$ and $B_2$ are called the large components, while $B_3$ and $B_4$ are called the small components. It is customary to solve for the small components $B_3$ and $B_4$ in terms of the large components. Doing this for the positive energy solutions we get

$$B_3 = \frac{p_z B_1 + p_- B_2}{m_0 c + \frac{\varepsilon}{c}}, \text{ and} \tag{10.41a}$$

$$B_4 = \frac{p_+ B_1 - p_z B_2}{m_0 c + \frac{\varepsilon}{c}}. \tag{10.41b}$$

## 10.7 Dirac's Theory of the Positron

In classical relativistic mechanics, negative energy values $E = -\varepsilon$ are not considered, since transitions between positive and negative energy values may not occur. However, in quantum theory such discontinuous transitions may take place, and further, negative energy eigenfunctions are needed to obtain complete sets of energy eigenfunctions. Hence, it is not possible to ignore negative energy states.

By a calculation known as *charge conjugation*, Dirac showed that the negative energy solutions refer to the motion of a particle having the mass of an electron and the opposite charge. Such particles have been observed experimentally and are called *positrons*. However, as Dirac points out, the negative energy solutions do not represent positrons, since this interpretation would lead to ludicrous conclusions, such as negative kinetic energy and a relativistic energy $E$ decreasing as $p^2$ increases. It is equally incorrect to interpret the negative energy solutions as representing particles of negative mass.

In order to obtain an interpretation of the negative energy solutions in agreement with experimental evidence, Dirac introduced two basic assumptions regarding these solutions: (1) Nearly all negative energy states are occupied by electrons, (2) A electron in a negative energy state is not observable, but unoccupied negative energy states are observed as positrons.

According to Dirac's first assumption, negative energy solutions represent electrons in unobservable states. Consider an electron in a negative energy state $E = -\varepsilon \le -m_0 c^2$. Raising it to a positive energy state, $E' = +\varepsilon \ge +m_0 c^2$, requires energy $E' - E \ge 2 m_0 c^2$. The previously unobservable electron becomes observable as a positive energy electron, and the unoccupied *hole* which it vacates is observable as a positron. This is a brief synopsis of *Dirac's Hole Theory*, which accounts for observed *pair production.*

## 10.8 Inner Products and Expectation Values

In order to give full physical content to Dirac's relativistic wave equation, we need to define inner products and expectation values in terms of four-component wave functions and Dirac operators ($4 \times 4$ matrices). Then the formalism of the Schrödinger nonrelativistic wave mechanics may be carried over completely. In particular, we can say that observables associated with time independent operators which commute with the Dirac Hamiltonian will be quantum mechanical constants of the motion.

If we let $f$ and $g$ be two four-component wave functions such that each of the components of $f$ and each of the components of $g$ are square-integrable, then the inner product is defined by

$$\langle f | g \rangle = \int_{\infty} \left( f_1^* g_1 + f_2^* g_2 + f_3^* g_3 + f_4^* g_4 \right) dr, \tag{10.42}$$

and the norm squared of $f$ is given by $\langle f | f \rangle$. Further, a constant $c$ may be found as before such that the norm of $f' = cf$ is one, in which case $f'$ is ***normalized.*** If $\langle f | g \rangle = 0$, $f$ and $g$ are *orthogonal.*

The expectation value $\langle A \rangle$ of an observable $A$ is defined for a represented by a four-component wave function $\psi$ as

$$\langle A \rangle = \frac{\langle \psi | A | \psi \rangle}{\langle \psi | \psi \rangle}, \tag{10.43}$$

where on the right-hand side, $A$ is the $4 \times 4$ matrix representation of the operator $A$ associated with observable $A$.

## 10.9 Spin Angular Momentum and Magnetic Moment

One of the most remarkable successes of Dirac's theory is the natural way in which spin angular momentum and spin magnetic moment are correctly predicted by the Dirac equation.

For a particle in a central field $V(r)$, we have seen that the Dirac Hamiltonian is

$$H = c\alpha \cdot p + m_0 c^2 \beta + V(r). \tag{10.44}$$

It might be expected, as in the Schrödinger theory, that $H$ and $M$ commute. In determining if this is the case, it is only necessary to work with one component of $M$, say $M_z$. Since $M_z$ commutes with the second two terms of (10.44), we need only examine

$$[M_z, H] = c\alpha_1[M_z, p_x] + c\alpha_2\left[M_z, p_y\right] + c\alpha_3[M_z, p_z]. \tag{10.45}$$

Using the commutation relations between $M_z$ and the components of $p$, we may write

$$[M_z, H] = i\hbar c\left(\alpha_1 p_y - \alpha_2 p_x\right). \tag{10.46}$$

From this we see that

$$[M, H] = i\hbar c\alpha \times p. \tag{10.47}$$

Hence $M$ does not commute with $H$, and is not a constant of the motion, as it was in the Schrödinger theory.

We still expect there exists some vector, similar to $M$, which commutes with $H$. Working with angular momentum still, we try $\sigma$. Again, for convenience sake, we work with one component, say $\sigma_3$.Working with this we will find that

$$[\sigma, H] = -2ic\alpha \times p. \tag{10.48}$$

So $\sigma$ also does not commute with $H$.

Not getting frustrated though, we notice that the two commutators above have a similar form, and their addition will give us an operator that will commute with the Hamiltonian, namely,

$$J = M + \tfrac{\hbar}{2}\sigma. \tag{10.49}$$

Since $M$ represents orbital angular momentum, it is quite natural to interpret $s = \frac{\hbar}{2}\sigma$ as the spin angular momentum and $J = M + s$ as the total angular momentum.

To show that

$$\mu = \frac{e}{m_0 c}s = \frac{\hbar e}{2m_0 c}\sigma \tag{10.50}$$

represents the intrinsic magnetic moment of an electron, Dirac squared the Hamiltonian for a particle in a magnetic field $B = \nabla \times A$. From (10.6) and (10.7) we find that

$$\left(\tfrac{H}{c}\right)^2 = (\alpha \cdot \pi + m_0 c\beta)^2 = (\alpha \cdot \pi)^2 + m_0 c\{\beta(\alpha \cdot \pi) + (\alpha \cdot \pi)\beta\} + m_0^2 c^2. \tag{10.51}$$

By using (10.9), this can be reduced to

$$\left(\tfrac{H}{c}\right)^2 = (\alpha \cdot \pi)^2 + m_0^2 c^2. \tag{10.52}$$

However, we also know that

$$(\alpha \cdot \pi)^2 = \pi^2 - \tfrac{\hbar e}{c}\sigma \cdot B, \text{ so} \tag{10.53}$$

$$\left(\frac{H_1}{c} + m_0 c\right)^2 = \pi^2 - \tfrac{\hbar e}{c}\sigma \cdot B + m_0^2 c^2, \tag{10.54}$$

where $H = H_1 + m_0c^2$. Assuming that $H_1 << m_0c^2$, we obtain

$$H_1 = \frac{\pi^2}{2m_0} - \frac{\hbar e}{2m_0c}\sigma \cdot B. \tag{10.55}$$

Since the energy of a dipole of magnetic moment m in a magnetic field $B$ is -$\mu B$, it is natural to attribute an electron an intrinsic magnetic moment (10.50). This agrees with experimental evidence, and with the Uhlenbeck-Goudsmit hypothesis. In the Dirac theory, no extraneous hypothesis is needed for the introduction of electron spin.

## 10.10 Angular Momentum Eigenvalues and Eigenfunctions

In preparation for solving the problem of an electron in a radially symmetric electrostatic field, one of the most important early results of the Dirac theory, we want to obtain simultaneous eigenfunctions of$^2$ and $J_z$. These two operators commute, and it can also be shown that any eigenfunction of $K = \beta\sigma \cdot M + \hbar\beta$ is also an eigenfunction of $J^2$, since $J^2 = K^2 - \frac{\hbar^2}{4}$.

From $J_z = M_z + \frac{\hbar}{2}\sigma_z$ we find that a four-component wave function $\psi$ is an eigenfunction of $J_z$ for $J_z = \mu\hbar$ only if

$$M_z\psi_{1,3} = \left(\mu - \tfrac{1}{2}\right)\hbar\psi_{1,3}, \text{ and} \tag{10.56a}$$

$$M_z\psi_{2,4} = \left(\mu + \tfrac{1}{2}\right)\hbar\psi_{2,4}. \tag{10.56b}$$

Now, if we also let $\psi$ be an eigenfunction of $K$ for $K = k\hbar$, then we can show that

$$M^2\psi = \left(k^2 - k\beta\right)\hbar^2\psi, \tag{10.57}$$

since $M^2 = K^2 - \hbar\beta K$. Using the $4 \times 4$ matrix representation of $\beta$, in (10.11), we see that

$$\begin{bmatrix} M^2 & 0 & 0 & 0 \\ 0 & M^2 & 0 & 0 \\ 0 & 0 & M^2 & 0 \\ 0 & 0 & 0 & M^2 \end{bmatrix}\begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = \hbar^2\begin{bmatrix} k^2-k & 0 & 0 & 0 \\ 0 & k^2-k & 0 & 0 \\ 0 & 0 & k^2+k & 0 \\ 0 & 0 & 0 & k^2+k \end{bmatrix}\begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}. \tag{10.58}$$

From this we find that

$$M^2\psi_{1,2} = k(k-1)\hbar^2\psi_{1,2}, \text{ and} \tag{10.59a}$$

$$M^2\psi_{3,4} = k(k+1)\hbar^2\psi_{3,4}, \tag{10.59b}$$

if $K = k\hbar$. Hence, $\psi_{1,2}$ must be an eigenfunction of $M^2 = k(k-1)\hbar^2$ and $\psi_{3,4}$ an eigenfunction of $M^2 = k(k+1)\hbar^2$.

The simultaneous eigenfunctions of $M^2$ and $M_z$ are the surface spherical harmonics $Y_l^m(\theta,\phi)$, multiplied by arbitrary functions of the radial coordinate $r$. For $Y_l^m$ we have $M^2 = l(l+1)\hbar^2$ and $M_z = m\hbar$. Writing $k(k-1) = l(l+1)$ we obviously have either $k = -l$ or $k = l+1$. Writing $k(k+1) = l(l+1)$ we obviously have either $k = l$ or $k = -l-1$. Thus we obtain all possible solutions of (10.56) and (10.59) by taking $\mu = m + \frac{1}{2}$ and $k = -l$ or $k = l$.

Taking $\mu = m + \frac{1}{2}$ and $k = -l$ we find, by using our knowledge of surface spherical harmonics, that (10.56) and (10.59) are satisfied by

$$\psi_{-l} = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = \begin{pmatrix} f_1(r)Y_l^m(\theta,\phi) \\ f_2(r)Y_l^{m+1}(\theta,\phi) \\ f_3(r)Y_{l-1}^m(\theta,\phi) \\ f_4(r)Y_{l-1}^{m+1}(\theta,\phi) \end{pmatrix}. \tag{10.60}$$

Next, from substituting into a system of equations derived from the Dirac representation of $K$, that we may set the functions $f$ to

$$f_1(r) = \sqrt{l-m}\,f(r), \tag{10.61a}$$

$$f_2(r) = -\sqrt{l+m+1}\,f(r), \tag{10.61b}$$

$$f_3(r) = \sqrt{l+m}\,F(r), \text{ and} \tag{10.61c}$$

$$f_4(r) = \sqrt{l-m-1}\,F(r), \tag{10.61d}$$

where $f$ and $F$ are arbitrary still. Now, putting it all together, we get

$$\psi_{-l} = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = \begin{pmatrix} \sqrt{l-m}\,f(r)Y_l^m(\theta,\phi) \\ -\sqrt{l+m+1}\,f(r)Y_l^{m+1}(\theta,\phi) \\ \sqrt{l+m}\,F(r)Y_{l-1}^m(\theta,\phi) \\ \sqrt{l-m-1}\,F(r)Y_{l-1}^{m+1}(\theta,\phi) \end{pmatrix}, \tag{10.62}$$

is a simultaneous eigenfunction of $J_z = \left(m + \frac{1}{2}\right)\hbar$ and $K = -l\hbar$.

In a similar manner, taking $\mu = m + \frac{1}{2}$ and $k = l$, it can be found that

(10.63)
$$\psi_l = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = \begin{pmatrix} \sqrt{l+m}\, g(r) Y_l^m(\theta,\phi) \\ \sqrt{l-m-1}\, g(r) Y_l^{m+1}(\theta,\phi) \\ \sqrt{l-m}\, G(r) Y_{l-1}^m(\theta,\phi) \\ \sqrt{l+m+1}\, G(r) Y_{l-1}^{m+1}(\theta,\phi) \end{pmatrix},$$

where $g$ and $G$ are arbitrary functions of $r$, is a simultaneous eigenfunction of $K = l\hbar$ and $J_z = \left(m + \frac{1}{2}\right)\hbar$.

From these we can see that the four-component wave functions $\psi_l$ and $\psi_{-l}$ are simultaneous eigenfunctions of $J_z = \left(m + \frac{1}{2}\right)\hbar$ and $J^2 = j(j+1)\hbar^2$ where $j = l - \frac{1}{2}$.

The functions $f$, $F$, $g$, and $G$ remain to be determined by the requirement that the $\psi$'s are eigenfunctions of the Dirac Hamiltonian. The solution for $V(r) = \frac{-e^2}{r}$, giving the energy eigenvalues and four-component wave functions for the hydrogen atom, is done in the next section.

## 10.11 Solution for the Hydrogen Atom

Because there is a spherical symmetry to the hydrogen atom, the eigenfunctions and eigenvalues of the Dirac Hamiltonian are most readily found by using a spherical coordinate system. The first thing to do is express the terms in our equations in spherical coordinates. Doing so we find that

(10.64)
$$\alpha \cdot p = \frac{1}{r^2}\left\{(\alpha \cdot r)(r \cdot p) + i(\alpha \cdot r)(\sigma \cdot M)\right\}.$$

By substituting this into (10.44) and using the fact that $\sigma \cdot M = \beta K - \hbar$, we obtain a representation of the Dirac Hamiltonian in spherical coordinates

(10.65)
$$H = \frac{c}{r^2}(\alpha \cdot r)\left\{(r \cdot p) + i\beta K - i\hbar\right\} + m_0 c^2 \beta + V(r).$$

The next major step is to obtain a $2 \times 2$ nested matrix representation of $H$. By replacing $\pi$ in (10.15) by $r$ we find first that

(10.66)
$$\alpha \cdot r = \begin{pmatrix} 0 & r\tau \\ r\tau & 0 \end{pmatrix},$$

where

(10.67) $$\tau = \begin{pmatrix} \frac{z}{r} & \frac{x-iy}{r} \\ \frac{x+iy}{r} & \frac{-z}{r} \end{pmatrix} = \begin{pmatrix} \cos\theta & e^{-i\phi}\sin\theta \\ e^{i\phi}\sin\theta & -\cos\theta \end{pmatrix}.$$

Through substitution of the other identities that we know, we find that for an eigenfunction of $K = k\hbar$ that

(10.68) $$\frac{H}{c} = \begin{pmatrix} \left(m_0c + \frac{V}{c}\right)I_2 & \frac{\tau}{r}\left(r\cdot p - i(k+1)\hbar\right) \\ \frac{\tau}{r}\left(r\cdot p + i(k-1)\hbar\right) & \left(-m_0c + \frac{V}{c}\right)I_2 \end{pmatrix}.$$

By writing the energy eigenvalue equation as a system of two equations depending on the two-component wavefunctions $\phi$ and $\varphi$, and performing some elementary matrix algebra, along with the substitutions $\phi = \frac{u}{r}$ and $\varphi = \frac{i\tau v}{r}$, we obtain

(10.69a) $$\left(a_1 + \frac{V}{\hbar c}\right)u + \left(v' + \frac{kv}{r}\right) = 0, \text{ and}$$

(10.69b) $$\left(u' - \frac{ku}{r}\right) + \left(a_2 - \frac{V}{\hbar c}\right)v = 0, \text{ where}$$

(10.69c) $$a_1 = \frac{m_0c^2 - E}{\hbar c} \text{ and } a_2 = \frac{m_0c^2 + E}{\hbar c}.$$

The solution to this set of equations is more easily found by using its approximate solution for large values of $r$. If $V(r) \rightarrow 0$ as $r \rightarrow \infty$, as is the case of the Coulomb potential $V(r) = \frac{-e^2}{r}$, the system of equations can be approximated for large $r$ by

(10.70) $$a_1u + v' = 0 \text{ and } u' + a_2v = 0.$$

Now, the bounded solutions to this system is proportional to $e^{ar}$, where $a$ is such that $a^2 = a_1a_2 = m_0c^2 - \frac{E^2}{c^2}$. Thus we will make the substitutions $u = e^{ar}P(r)$ and $v = e^{ar}Q(r)$, and also make a change of variable to $s = ar$. Then (10.69) becomes

(10.71a) $$\left(\frac{a_1}{a} - \frac{\alpha}{s}\right)P + \left(\frac{\partial}{\partial s} - 1 + \frac{k}{s}\right)Q = 0, \text{ and}$$

(10.71b) $$\left(\frac{\partial}{\partial s} - 1 - \frac{k}{s}\right)P + \left(\frac{a_2}{a} + \frac{\alpha}{s}\right)Q = 0,$$

where $\alpha = \frac{e^2}{\hbar c}$ is the *fine structure constant.*

Considering P and Q to be able to be written as an infinite series, we let

(10.72) $$P = \sum_{j=0}^{\infty} p_j s^{j+\lambda} \text{ and } Q = \sum_{j=0}^{\infty} q_j s^{j+\lambda}.$$

Substituting these into our system of equations, collecting like powers of $s$, and dividing by $s^{\lambda}$ we end up with a system of equations which can be satisfied only if the coefficients of $s$ vanish separately. Equating the coefficients of $s^{-1}$ to zero, we find that we must have

(10.73a) $$\alpha p_0 - (\lambda + k) q_0 = 0, \text{ and}$$

(10.73b) $$(\lambda - k) p_0 + \alpha q_0 = 0.$$

These two equations are consistent only if $\lambda^2 = k^2 - \alpha^2$. To obtain a solution that is finite at $r = 0$, we take $\lambda = +\sqrt{k^2 - \alpha^2}$. Building a recurrence relation by a similar method, eg: equating the coefficients of $s^{j-1}$ to zero, we get

(10.74) $$\left((j + \lambda - k)a - \alpha a_2\right) p_j + \left((j + \lambda + k) a_2 + \alpha a\right) q_j = 0.$$

From this recurrence relation we can show that the solutions being obtained are square integrable only if the series (10.72) terminate, which they will do with the term $s^{N+\lambda}$.

then we let $j=N$, and we use $a^2 = a_1 a_2 = m_0 c^2 - \frac{E^2}{c^2}$ again, to get

(10.75) $$2(N + \lambda - k)a = \alpha(a_2 - a_1).$$

This will determine the energy eigenvalues. Using $\lambda = +\sqrt{k^2 - \alpha^2}$ we obtain

(10.76) $$E = m_0 c^2 \sqrt{1 + \left(\frac{\alpha}{N + \sqrt{k^2 - \alpha^2}}\right)^2}.$$

When we subtract the rest energy $m_0 c^2$ and replace $N$ by a symbol $s$, we observe that the resulting expression is identical with (5.42) which was derived using the Wilson-Sommerfeld quantization. Hence, Sommerfeld's formula is a consequence of Dirac's relativistic wave equation.

## 10.12 Physical and Statistical Interpretation of Dirac Wave Functions

A natural extension of Schrödinger's physical interpretation of the nonrelativistic wave function, section 6.5, to the Dirac four-component wave function is to assume that

(10.77) $$\rho = e\psi^*\psi = e\left(\psi_1^*\psi_1 + \psi_2^*\psi_2 + \psi_3^*\psi_3 + \psi_4^*\psi_4\right)$$

represents the density of the charge in a cloud picture of the electron. In a statistical interpretation we assume that $J = \psi^*\psi$ represents the probability density. These assumptions are valid only if a conservation equation is satisfied.

If $\psi$ satisfies the Dirac equation

$$i\hbar\frac{\partial\psi}{\partial t} = -i\hbar\nabla(c\alpha\psi) - \frac{e}{c}\alpha\cdot A\psi + m_0c^2\beta\psi, \tag{10.78}$$

then the conjugate wavefunction satisfies the conjugate equation

$$-i\hbar\frac{\partial\psi^*}{\partial t} = i\hbar\nabla\left(\psi^*c\alpha\right) - \frac{e}{c}\psi^*\alpha\cdot A + m_0c^2\psi^*\beta. \tag{10.79}$$

In interpreting these, we remind ourselves that each term in the former is a column vector, while each term in the later is a row vector, and the operators are Hermitean matrices. Multiply (10.78) from the left with $\psi^*$, and (10.79) from the right with $\psi$, subtract the two, and divide by $i\hbar$. The resulting equation is

$$\frac{\partial}{\partial t}\left(\psi^*\psi\right) = -\nabla\left(\psi^*c\alpha\psi\right), \tag{10.80}$$

and it is seen that $J = \psi^*\psi$ may be interpreted as a probability density, provided

$$S = \psi^*c\alpha\psi \tag{10.81}$$

is interpreted as the probability current. In the cloud picture of the electron, $\rho = eJ$ and $j = eS$, as in section 6.5.

By considering the motion of a free electron in the Heisenberg picture, Dirac was led to a very surprising and interesting result. First, he observed that the momentum $p$ commutes with the Hamiltonian, and is thus a constant of the motion. Next he found that the operator for the velocity is

$$\dot{r} = c\alpha. \tag{10.82}$$

Dirac concluded that, since $\pm 1$ are the only eigenvalues of the components of $\alpha$, "*a measure of a component of the velocity of a free electron is certain to lead to the result* $\pm c$." Dirac explained this surprising result in terms of the uncertainty principle and a "trembling motion" in which the average velocity, which is observed, is less than $c$, but the instantaneous velocity is always $\pm c$.

Although it may be paradoxical on first thought, on reflection it is natural that Dirac's theory, which contributes so much to our concept of the electron, leaves us wondering just what an electron really is. Advanced quantum electrodynamics is beset with difficulties much more serious that interpreting this trembling motion. Dirac suggests that "the difficulties, being of a profound character, can be removed only by some drastic changes in the foundations of the theory, probably a change as drastic as the passage from Bohr's orbit theory to the present quantum mechanics." This change may be contingent on discoveries in mathematics, as drastic as the step made by Newton and Leibnitz in discovering the calculus.

# Appendix I

# Approximation Techniques

*This chapter looks at some of the approximation techniques more frequently used in quantum mechanics. The first to be looked at is the Wentzel-Kramers-Brillouin method, or WKB Method for short. This is a scheme for determining approximate energy eigenvalues. Then we look at perturbation theory, a general method for establishing approximate wavefunctions and energies for potentials that can be expressed as slight departures from potentials for which the Schrödinger equations can be solved exactly.*

## A.1 The WKB Approximation

The Wentzel-Kramers-Brillouin method, or WKB Method, arises from *barrier penetration* problems. These are the non-zero probability of finding a particle outside its classical turning points due to the exponential-decay characterizing solutions of the Schrödinger equation when $E < V(x)$. If the penetration of the wave into the potential barrier is slight, one can say that $\psi(x)$ approximately goes to zero where $E = V(x)$.

This situation is similar to that of a vibrating string clamped at both ends. The fact that the ends of the string cannot oscillate restricts the string to oscillating in modes where integral numbers of half-wavelengths fill the intervening space between the ends. The essence of the WKB approximation is that if $\psi(x)$ is characterized by a wavelength $\lambda$, then the particle energy $E$ must be such that approximately an integral number of half-wavelengths fit between the turning points of the motion. This approximation arises from ignoring the presumably negligent barrier penetration.

To derive the WKB approximation, recall that Schrödinger's wave equation is

(A.1a) $$H\psi(x) = E\psi(x),$$

or, in the momentum representation,

(A.1b) $$\left[\frac{p^2}{2m} + V(x)\right]\psi = E\psi.$$

This gives us the momentum representation of the energy,

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

Rearranging this to solve for p, we get

(A.3) $$p=\sqrt{2m(E-V(x))}.$$

From de Broglie's work (see section 2.4), we know that

(A.3) $$p=\frac{h}{\lambda},$$

and, given that a standing wave with an integral number of half-wavelengths will approximate our solution and give us an energy that is constant in time, we get

(A.4) $$\frac{h}{\lambda}\approx\sqrt{2m(E-V(x))}\ .$$

Integrating this with respect to $x$ gives us

(A.5) $$n\hbar\approx 2\int\sqrt{2m(E-V(x))}\,dx=\sqrt{8m}\int\sqrt{E-V(x)}\,dx,$$

which is the WKB approximation. Remember when using this that the turning points occur where the kinetic energy is zero, hence the momentum is zero, and $E=V(x=\text{turning points})$.

As an example of how to use the WKB method, consider a free particle in an infinite well of width $L$, where the potential function is given by

(A.6) $$V(x)=\begin{cases}\infty & x<0\\ 0 & 0\le x\le L\\ \infty & L<x\end{cases}.$$

Using the WKB approximation, we know that

(A.7) $$n\hbar\approx\sqrt{8m}\int_{-\infty}^{\infty}\sqrt{E-V(x)}\,dx,$$

which for our problem simplifies to

(A.8) $$n\hbar\approx\sqrt{8m}\int_{0}^{L}\sqrt{E}\,dx.$$

Integration of this gives us

(A.9) $$n\hbar\approx\sqrt{8mE}\,L,$$

or, rewriting

$$E_n \approx \frac{n^2h^2}{8mL^2}. \tag{A.10}$$

As you may recall, this is the exact solution, not approximate. This means that our assumption, that there is no penetration of the barrier, is true in this case. Keep in mind that this is not necessarily true, and in most cases it won't be. It does however validate the method.

*Problem A.1:* Consider a potential function given by

$$V(x) = \begin{cases} \infty & x \le 0 \\ ax & 0 < x < \infty \end{cases}.$$

Find $E$ using the WKB approximation, if $V(x) < E$ for $0 < x < a$. In other words, $x = 0$ and $x = a$ are the turning points of the motion.[1]

Answer: $E_n \approx \left[\frac{3nha}{4\sqrt{2m}}\right]^{2/3}$

*Problem A.2:* Consider a potential function given by

$$V(x) = \begin{cases} \infty & x \le 0 \\ 0 & 0 < x \le a \\ V_0 & a < x \le b \\ \frac{1}{2}V_0 & b < x \le c \\ \infty & c < x \end{cases}.$$

Find $E$ using the WKB approximation, if $V_0 < E < \infty$ for $0 < x < c$.[2]

Answer: $E_n \approx \frac{n^2h^2}{8mc^2} - \frac{V_0}{c^2}\left(a^2 - \frac{b^2}{2} - \frac{c^2}{2}\right)$

---

[1] Fisher, Notes, WKB Homework, problem 1

[2] Fisher, Notes, WKB Homework, problem 2

*Problem A.3:* Find the energy of a hydrogen atom using spherical coordinates. Hint: $V(x) = \frac{-1}{4\pi\varepsilon_0}\frac{e}{r}$ and $E = \frac{-1}{4\pi\varepsilon_0}\frac{e^2}{a}$.[3]

Answer: $E_n \approx \dfrac{me^4}{8n^2h^2\varepsilon_o{}^2}$

## A.2 Time-Independent Perturbation Theory

The WKB approximation only works with one dimensional systems. To solve problems in more than one dimension, *perturbation theory* is needed. Perturbation theory is applicable when the system being looked at is similar to a system with a closed form solution, and the the closed form solution is of the form desired for the system in question. Some of the important applications of perturbation theory include the solutions to the photoelectric effect, the Compton effect, pair production and annihilation, the hydrogen atom in an electric field (the Stark effect), the helium atom, the hydrogen molecule, many particle systems, problems in scattering and diffraction, and many more.

To derive the equations of perturbation theory, consider a system in which the energy $E$ is constant in time. We then make the assumption that the Schrödinger time independent wave equation (6.9) is exactly solvable for a Hamiltonian $H^0$, where $H^0$ does not differ greatly from the Hamiltonian $H$ for a similar problem. Thus we can say that

(A.11) $$H^0\psi_n{}^0 = E_n{}^0\psi_n{}^0,$$

and

(A.12) $$H = H^0 + \lambda H^{(1)},$$

where $\lambda$ is the *perturbation parameter*, and is introduced to aid in distinguishing orders of magnitude. Making this substitution in the STIWE gives us

(A.13) $$H\psi = \left(H^0 + \lambda H^{(1)}\right)\psi = E\psi.$$

Note that when $\lambda = 0$, (A.3) is satisfied by any eigenfunction $\psi_n{}^0$ of $H^0$ for $E = E_n{}^0$. Thus, it is reasonable to expect that when $\lambda \neq 0$, (A.3) is satisfied by a

---

[3] Fisher, Notes, WKB Homework, problem 3

solution $\psi_n$ belonging to an eigenvalue $E = E_n$ which does not differ greatly from $E_n{}^0$. We may therefore write

(A.14a) $$\psi_n = \psi_n{}^0 + \lambda\psi_n{}^{(1)} + \lambda^2\psi_n{}^{(2)} + \ldots$$

(A.14b) $$E_n = E_n{}^0 + \lambda E_n{}^{(1)} + \lambda^2 E_n{}^{(2)} + \ldots.$$

Substituting (A.14) into (A.13), and equating coefficients of corresponding powers of $\lambda$, we obtain, up to the coefficients of $\lambda^2$:

(A.15a) $$H^0\psi_n{}^0 = E_n{}^0\psi_n{}^0$$

(A.15b) $$H^0\psi_n{}^{(1)} + H^{(1)}\psi_n{}^0 = E_n{}^0\psi_n{}^{(1)} + E_n{}^{(1)}\psi_n{}^0$$

(A.15c) $$H^0\psi_n{}^{(2)} + H^{(1)}\psi_n{}^{(1)} = E_n{}^0\psi_n{}^{(2)} + E_n{}^{(1)}\psi_n{}^{(1)} + E_n{}^{(2)}\psi_n{}^0.$$

Note that the zero order approximation (A.15a) is the equation for the unperturbed system. The first order approximation (A.15b) can be found by first rearranging so that like terms of $\psi$ are together:

(A.16) $$\left(H^0 - E_n{}^0\right)\psi_n{}^{(1)} = -\left(H^{(1)} - E_n{}^{(1)}\right)\psi_n{}^0,$$

and then forming the inner product of each term with $\psi_n{}^0$:

(A.17) $$\left\langle \psi_n{}^0 \middle| H^0 - E_n{}^0 \middle| \psi_n{}^{(1)} \right\rangle = -\left\langle \psi_n{}^0 \middle| H^{(1)} - E_n{}^{(1)} \middle| \psi_n{}^0 \right\rangle.$$

When this is solved, the left hand side of the equation goes to zero because of the Hermiticity of $H^0$ and the fact that $E_n{}^0$ is a real number. Likewise, $E_n{}^{(1)}$ is also a real number, and hence (A.17) becomes:

(A.18) $$0 = -\left\langle \psi_n{}^0 \middle| H^{(1)} \middle| \psi_n{}^0 \right\rangle + E_n{}^{(1)},$$

or, with some rearranging,

(A.19) $$E_n{}^{(1)} = H_{nn}{}^{(1)} = \left\langle \psi_n{}^0 \middle| H^{(1)} \middle| \psi_n{}^0 \right\rangle.$$

This is the first order approximation of $\psi_n$ belonging to an eigenvalue $E = E_n$.

From (A.15c), and following the above technique, it is easy to show that

$$E_n^{(2)} = \left\langle \psi_n^0 \middle| H^{(1)} \middle| \psi_n^{(1)} \right\rangle - E_n^{(1)} \left\langle \psi_n^0 \middle| \psi_n^{(1)} \right\rangle. \tag{A.20}$$

This form lacks some general utility though, because $\psi_n^{(1)}$ is not always known. Another form of the second order approximation may be derived by first making the assumption (valid, but not proven here) that

$$\psi_k^{(1)} = \sum_j a_j \psi_j^0, \tag{A.21a}$$

where

$$a_j = \frac{H_{jk}^{(1)}}{E_k^0 - E_j^0}. \tag{A.21b}$$

Now, we know from (A.15c), and the Hermiticity of $H^0$ that

$$H^{(1)} \psi_n^{(1)} = E_n^{(1)} \psi_n^{(1)} + E_n^{(2)} \psi_n^0. \tag{A.22}$$

With appropriate substitutions from (A.19) and (A.21), we get

$$\left[ H^{(1)} - \left\langle \psi_n^0 \middle| H^{(1)} \middle| \psi_n^{(1)} \right\rangle \right] \sum_j a_j \psi_j^0 = E_n^{(2)} \psi_n^0. \tag{A.23}$$

Forming the inner product with $\psi_n^0$ gives:

$$\left\langle \psi_n^0 \middle| H^{(1)} \sum_j a_j \middle| \psi_j^{(1)} \right\rangle = E_n^{(2)}, \tag{A.24}$$

or

$$E_n^{(2)} = \sum_{j \neq n} \frac{H_{nj}^{(1)} \cdot Hjn^{(1)}}{E_k^0 - E_j^0}. \tag{A.25}$$

This is sometimes an easier way to obtain the second order perturbation because it does not need a wavefunction in order to be evaluated.

### A.3 The Stark Effect for the Plane Rotator

As an example of the use of perturbation theory, consider a rigid rotator of moment of inertia $I$ and electric moment $\mu$. If this object is constrained to rotate about an axis perpendicular to a uniform electric field $F$, its potential energy is $U = -\mu F \cos\theta$ where $\theta$ is the angle between the polar axis axis and the direction of $F$. We know that the Hamiltonian for a rigid rotator with moment of inertia $I$ is

$$H^0 = \frac{L^2}{2I}, \tag{A.26}$$

and the energy is

$$E_n{}^0 = \frac{n^2\hbar^2}{2I}. \tag{A.27}$$

When the STIWE is solved with these values, we get

$$\psi_{\pm n}{}^0 = \frac{1}{\sqrt{2\pi}}\, e^{\pm in\theta}. \tag{A.28}$$

Note that there is a double degeneracy in the absence of the electric field (except in the ground state); this corresponds to the two possible directions of rotation.

Now, applying perturbation theory to take into consideration the electric field, set

$$H^{(1)} = -\mu F \cos\theta. \tag{A.29}$$

Using this and the first order approximation (A.19), we readily get

$$E_{kj}{}^{(1)} = \langle \psi_k \,|-\mu F \cos\theta|\, \psi_j \rangle. \tag{A.30}$$

Solving the inner product, we obtain

$$E_{kj}{}^{(1)} = H_{kj}{}^{(1)} = \frac{-\mu F}{2\pi}\int_0^{2\pi} e^{-ik\theta}\cos\theta\, e^{ij\theta}\, d\theta. \tag{A.31a}$$

Note that

$$H_{kj}{}^{(1)} = \begin{cases} \frac{-\mu F}{2} & j = k \pm 1 \\ 0 & j \neq k \pm 1 \end{cases}. \tag{A.31b}$$

We are interested in $E_n{}^{(1)} = E_{nn}{}^{(1)}$, though, which clearly is 0. Hence,

$$E_n = E_n{}^0 + E_n{}^{(1)} = E_n{}^0. \tag{A.32}$$

Thus the first order approximation makes no changes to the energy, and we must look at the second order approximation. Plugging (A.31b) into (A.25) gives us

$$E_n^{(2)} = \frac{\left(\frac{-1}{2}\mu F\right)^2}{E_k^0 - E_{k+1}^0} + \frac{\left(\frac{-1}{2}\mu F\right)^2}{E_k^0 - E_{k-1}^0}, \tag{A.33}$$

which, with (A.27) becomes:

$$E_n^{(2)} = \frac{\frac{1}{4}\mu^2 F^2}{\hbar^2/2I}\left[\frac{1}{n^2-n^2-2n-1} + \frac{1}{n^2-n^2+2n-1}\right], \tag{A.34}$$

or, simplifying,

$$E_n^{(2)} = \frac{\mu^2 F^2 I}{\hbar^2\left(4n^2-1\right)}. \tag{A.35}$$

Hence the second order approximation is required to account for the observed displacement of the energy levels.

# Bibliography

Arfleen, George. Mathematical Methods for Physicists. New York: Academic Press, Inc, 1985.

Butkov, Eugene. Mathematical Physics. Reading, MA: Addison-Wesley Publishing Company, 1968.

Dirac, P.A.M. The Principles of Quantum Mechanics. Oxford: Clarendon Press, 1958.

Fisher, David. Independent Study Research Notes. Unpublished, 1976.

Ikenberry, Ernest. Quantum Mechanics for Mathematicians and Physicists. New York: Oxford University Press, 1962.

Liboff, Richard. Introductory Quantum Mechanics. Reading, MA: Addison-Wesley Publishing Company, 1980.

Rice, Bernard and Strange, Jerry. Ordinary Differential Equations with Applications. California: Brooks/Cole Publishing Company, 1989.

Schneider, Dennis. Linear Algebra: A Concrete Introduction. New York: Macmillan Publishing Company, 1987.