

Quantum Mechanics – Concepts and Applications

Tarun Biswas

October 3, 2003

Copyright

Copyright ©1990, 1994, 1995, 1998, 1999 by Tarun Biswas, Physics Department, State University of New York at New Paltz, New Paltz, New York 12561.

Copyright Agreement

This online version of the book may be reproduced freely in small numbers (10 or less per individual) as long as it is done in its entirety including the title, the author's name and this copyright page. For larger numbers of copies one must obtain the author's written consent. Copies of this book may *not* be sold for profit.

Contents

1	Mathematical Preliminaries	1
1.1	The state vectors	1
1.2	The inner product	2
1.3	Linear operators	5
1.4	Eigenstates and eigenvalues	6
1.5	The Dirac delta function	13
2	The Laws (Postulates) of Quantum Mechanics	16
2.1	A lesson from classical mechanics	16
2.2	The postulates of quantum mechanics	17
2.3	Some history of the postulates	19
3	Popular Representations	20
3.1	The position representation	20
3.2	The momentum representation	23
4	Some Simple Examples	25
4.1	The Hamiltonian, conserved quantities and expectation value	25
4.2	Free particle in one dimension	30
4.2.1	Momentum	31

4.2.2	Energy	31
4.2.3	Position	32
4.3	The harmonic oscillator	34
4.3.1	Solution in position representation	36
4.3.2	A representation free solution	39
4.4	Landau levels	42
5	More One Dimensional Examples	45
5.1	General characteristics of solutions	45
5.1.1	$E < V(x)$ for all x	46
5.1.2	Bound states	47
5.1.3	Scattering states	49
5.2	Some oversimplified examples	53
5.2.1	Rectangular potential well (bound states)	55
5.2.2	Rectangular potential barrier (scattering states)	58
6	Numerical Techniques in One Space Dimension	64
6.1	Finite differences	65
6.2	One dimensional scattering	66
6.3	One dimensional bound state problems	72
6.4	Other techniques	74
6.5	Accuracy	75
6.6	Speed	75
7	Symmetries and Conserved Quantities	78
7.1	Symmetry groups and their representation	78
7.2	Space translation symmetry	82

7.3	Time translation symmetry	83
7.4	Rotation symmetry	84
7.4.1	Eigenvalues of angular momentum	85
7.4.2	Addition of angular momenta	89
7.5	Discrete symmetries	92
7.5.1	Space inversion	92
7.5.2	Time reversal	93
8	Three Dimensional Systems	96
8.1	General characteristics of bound states	96
8.2	Spherically symmetric potentials	97
8.3	Angular momentum	99
8.4	The two body problem	100
8.5	The hydrogen atom (bound states)	102
8.6	Scattering in three dimensions	104
8.6.1	Center of mass frame vs. laboratory frame	106
8.6.2	Relation between asymptotic wavefunction and cross section	107
8.7	Scattering due to a spherically symmetric potential	108
9	Numerical Techniques in Three Space Dimensions	112
9.1	Bound states (spherically symmetric potentials)	112
9.2	Bound states (general potential)	114
9.3	Scattering states (spherically symmetric potentials)	116
9.4	Scattering states (general potential)	118
10	Approximation Methods (Bound States)	121
10.1	Perturbation method (nondegenerate states)	122

10.2	Degenerate state perturbation analysis	126
10.3	Time dependent perturbation analysis	128
10.4	The variational method	132
11	Approximation Methods (Scattering States)	136
11.1	The Green's function method	137
11.2	The scattering matrix	142
11.3	The stationary case	146
11.4	The Born approximation	147
12	Spin and Atomic Spectra	149
12.1	Degenerate position eigenstates	150
12.2	Spin-half particles	153
12.3	Spin magnetic moment (Stern-Gerlach experiment)	155
12.4	Spin-orbit coupling	158
12.5	Zeeman effect revisited	160
13	Relativistic Quantum Mechanics	162
13.1	The Klein-Gordon equation	163
13.2	The Dirac equation	166
13.3	Spin and the Dirac particle	169
13.4	Spin-orbit coupling in the Dirac hamiltonian	170
13.5	The Dirac hydrogen atom	172
13.6	The Dirac particle in a magnetic field	177
A	'C' Programs for Assorted Problems	180
A.1	Program for the solution of energy eigenvalues for the rectangular potential well	180

A.2	General Program for one dimensional scattering off arbitrary barrier	181
A.3	Function for rectangular barrier potential	182
A.4	General energy eigenvalue search program	183
A.5	Function for the harmonic oscillator potential	185
A.6	Function for the hydrogen atom potential	186
B	Uncertainties and wavepackets	189

Preface

The fundamental idea behind any physical theory is to develop predictive power with a minimal set of experimentally tested postulates. However, historical development of a theory is not always that systematic. Different theorists and experimentalists approach the subject differently and achieve successes in different directions which gives the subject a rather “patchy” appearance. This has been particularly true for quantum mechanics. However, now that the dust has settled and physicists know quantum mechanics reasonably well, it is necessary to consolidate concepts and put together that minimal set of postulates.

The minimal set of postulates in classical mechanics is already very well known and hence it is a much easier subject to present to a student. In quantum mechanics such a set is usually not identified in text books which, I believe, is the major cause of fear of the subject among students. Very often, text books enumerate the postulates but continue to add further assumptions while solving individual problems. This is particularly disconcerting in quantum mechanics where, physical intuition being nonexistent, assumptions are difficult to justify. It is also necessary to separate the postulates from the sophisticated mathematical techniques needed to solve problems. In doing this one may draw analogies from classical mechanics where the physical postulate is Newton’s second law and everything else is creative mathematics for the purpose of using this law in different circumstances. In quantum mechanics the equivalent of Newton’s second law is, of course, the Schrödinger equation. However, before using the Schrödinger equation it is necessary to understand the mathematical meanings of its components e.g. the wavefunction or the state vector. This, of course, is also true for Newton’s law. There one needs to understand the relatively simple concept of particle trajectories.

Some previous texts have successfully separated the mathematics from the physical principles. However, as a consequence, they have introduced so much mathematics that the physical content of the theory is lost. Such books are better used as references rather than textbooks. The present text will attempt a compromise. It will maintain the separation of the minimal set of postulates from the mathematical techniques. At the same time close contact with experiment will be maintained to avoid alienating the physics student. Mathematical rigor will also be maintained barring some exceptions where it would take the reader too far afield into mathematics.

A significantly different feature of this book is the highlighting of numerical methods. An unavoidable consequence of doing practical physics is that most realistic problems do not have analytical solutions. The traditional approach to such problems has been a process of approximation of the complex system to a simple one and then adding appropriate numbers of correction terms. This has given rise to several methods of finding correction terms and some of them will be discussed in this text. However, these techniques were originally meant for hand computation. With the advent of present day computers more direct approaches to solving complex problems are available. Hence, besides learning to solve standard analytically solvable problems, the student needs to learn general numerical techniques that would allow one to solve any problem that has a solution. This would serve two purposes. First, it makes the student confident that every well defined problem is solvable and the world does not have to be made up of close approximations of the harmonic oscillator and the hydrogen atom. Second, one very often comes up with a problem that is so far from analytically solvable problems that standard approximation methods would not be reliable. This has been my motivation in including two chapters on numerical techniques and encouraging the student to use such techniques at every opportunity. The goal of these chapters is not to provide the most accurate algorithms or to give a complete discussion of all numerical techniques known (the list would be too long even if I were to know them all). Instead, I discuss the intuitively obvious techniques and encourage students to develop their own tailor-made recipes for specific problems.

This book has been designed for a first course (two semesters) in quantum mechanics at the graduate level. The student is expected to be familiar with the physical principles behind basic ideas like the Planck hypothesis and the de Broglie hypothesis. He (or she) would also need the background of a graduate level course in classical mechanics and some working knowledge of linear algebra and differential equations.

Chapter 1

Mathematical Preliminaries

1.1 The state vectors

In the next chapter we shall consider the complete descriptor of a system to be its *state vector*. Here I shall define the state vector through its properties. Some properties and definitions that are too obvious will be omitted. I shall use a slightly modified version of the convenient notation given by Dirac [1]. A state vector might also be called a *state* or a *vector* for short. In the following, the reader is encouraged to see analogies from complex matrix algebra.

A state vector for some state s can be represented by the so called *ket* vector $|s\rangle$. The label s can be chosen conveniently for specific problems. $|s\rangle$ will in general depend on all degrees of freedom of the system as well as time. The space of all possible kets for a system will be called the linear vector space \mathcal{V} . In the following, the term linear will be dropped as all vector spaces considered here will be linear. The fundamental property (or rule) of \mathcal{V} is

Rule 1 If $|s\rangle, |r\rangle \in \mathcal{V}$ then

$$a|s\rangle + b|r\rangle \in \mathcal{V},$$

where $a, b \in \mathcal{C}$ (set of complex numbers)

The meaning of addition of kets and multiplication by complex numbers will become obvious in the sense of components of the vector once components are defined. The physical content of the state vector is purely in its “direction”, that is

Rule 2 The physical contents of $|s\rangle$ and $a|s\rangle$ are the same if $a \in \mathcal{C}$ and $a \neq 0$.

At this stage the following commonly used terms can be defined.

Definition 1 A *LINEAR COMBINATION* of state vectors is a sum of several vectors weighted by complex numbers e.g.

$$a|p\rangle + b|q\rangle + c|r\rangle + d|s\rangle + \dots$$

where $a, b, c, d \in \mathcal{C}$.

Definition 2 A set of state vectors is called *LINEARLY INDEPENDENT* if no one member of the set can be written as a linear combination of the others.

Definition 3 A subset U of linearly independent state vectors is called *COMPLETE* if any $|s\rangle \in \mathcal{V}$ can be written as a linear combination of members of U .

1.2 The inner product

The *inner product* is defined as a mapping of an ordered pair of vectors onto \mathcal{C} , that is, the inner product is a complex number associated to an ordered pair of state vectors. It can be denoted as $(|r\rangle, |s\rangle)$ for the two states $|r\rangle$ and $|s\rangle$. The following property of the inner product is sometimes called sesquilinearity.

Rule 3

$$\begin{aligned} (a|r\rangle + b|u\rangle, c|s\rangle + d|v\rangle) = \\ a^*c(|r\rangle, |s\rangle) + b^*c(|u\rangle, |s\rangle) + a^*d(|r\rangle, |v\rangle) + b^*d(|u\rangle, |v\rangle). \end{aligned}$$

This indicates that the inner product is linear in the right argument in the usual sense but *antilinear* in the left argument. The meaning of antilinearity is obvious from rule 3. For compactness of notation one defines the following.

Definition 4 \mathcal{V}^\dagger , is called the *adjoint* of \mathcal{V} . For every member $|s\rangle \in \mathcal{V}$ there is a corresponding member $|s\rangle^\dagger \in \mathcal{V}^\dagger$ and vice versa. The $\langle s|$ (bra of s) notation is chosen such that

$$|s\rangle^\dagger \equiv \langle s|, \quad \langle s|^\dagger \equiv |s\rangle$$

The one-to-one correspondence of \mathcal{V} and \mathcal{V}^\dagger is specified as follows through the corresponding members $|r\rangle$ and $\langle r|$.

$$|r\rangle^\dagger |s\rangle \equiv \langle r||s\rangle \equiv \langle r|s\rangle \equiv (|r\rangle, |s\rangle) \tag{1.1}$$

where $|s\rangle$ is an arbitrary ket.

The names “bra” and “ket” are chosen because together they form the “bracket” of the inner product. From rule 3 and definition 4 it can be seen that

$$(a|r\rangle + b|u\rangle)^\dagger = a^*\langle r| + b^*\langle u|, \quad (1.2)$$

$$(a\langle r| + b\langle u|)^\dagger = a^*|r\rangle + b^*|u\rangle. \quad (1.3)$$

Using this new notation, rule 3 can now be written as

Rule 3

$$\begin{aligned} & (a|r\rangle + b|u\rangle)^\dagger (c|s\rangle + d|v\rangle) \\ &= (a^*\langle r| + b^*\langle u|)(c|s\rangle + d|v\rangle) \\ &= a^*c\langle r|s\rangle + b^*c\langle u|s\rangle + a^*d\langle r|v\rangle + b^*d\langle u|v\rangle. \end{aligned}$$

Another property of the inner product that is necessary for our applications is

Rule 4

$$\langle r|s\rangle^* \equiv \langle r|s\rangle^\dagger = |s\rangle^\dagger \langle r|^\dagger = \langle s|r\rangle.$$

At this stage it might have occurred to the student that state vectors are a generalization of vectors in arbitrary dimensions. In fact they will be seen to be of infinite dimensionality in most cases. The kets are like column vectors and the bras like row vectors of complex matrix algebra. The inner product is the equivalent of the scalar or dot product.

Extending the analogy one can define orthogonality, and norm.

Definition 5 *Two nonzero vectors represented by the kets $|r\rangle$ and $|s\rangle$ are defined to be ORTHOGONAL if $\langle r|s\rangle = 0$.*

Definition 6 *The NORM of a vector $|s\rangle$ is defined as its inner product with itself viz. $\langle s|s\rangle$. Note that, for convenience, this is chosen to be the square of the usual definition of the norm.*

From rule 4 it is obvious that the norm of any vector must be real. Another rule that one needs can now be introduced.

Rule 5 *The norm of every vector in \mathcal{V} is positive definite except for the zero vector (the additive identity) which has a norm of zero.*

Now one can prove two useful theorems relating orthogonality and linear independence of a set of vectors.

Theorem 1.1 *A set of mutually orthogonal nonzero vectors is linearly independent.*

Proof: Let the set of mutually orthogonal vectors be $\{|f_i\rangle\}$ where the label i distinguishes different members of the set. Here I shall choose i to be a positive integer. But the proof presented here can be readily generalized for i belonging to any set of integers or even a continuous set of real numbers.

We shall prove the theorem by contradiction. Hence, let us assume that the set is not linearly independent i.e. some member $|f_k\rangle$ of the set can be written as a linear combination of the others. Then

$$|f_k\rangle = \sum_{i \neq k} a_i |f_i\rangle. \quad (1.4)$$

Multiplying (i.e. taking an inner product) from the left by $\langle f_j|$ ($j \neq k$), one obtains

$$\langle f_j|f_k\rangle = \sum_{i \neq k} a_i \langle f_j|f_i\rangle. \quad (1.5)$$

From the mutual orthogonality condition the left side vanishes and the right side has only one term remaining i.e.

$$0 = a_j \langle f_j|f_j\rangle. \quad (1.6)$$

From rule 5 we conclude that $\langle f_j|f_j\rangle$ cannot be zero and hence

$$a_j = 0 \quad \forall j. \quad (1.7)$$

This leads to the right side of equation 1.4 being zero. But the vector $|f_k\rangle$ is not zero. This contradiction completes the proof.

Theorem 1.2 *Members of a set of n linearly independent nonzero vectors can be written as a linear combination of a (nonunique) set of n mutually orthogonal nonzero vectors.*

Proof: Let the given set of linearly independent vectors be $\{|g_i\rangle\}$. For convenience the label i can be considered to be a positive integer ($i = 1, 2, \dots, n$). However, a generalization for i belonging to any set of integers or even a continuous set of real numbers is possible.

We shall prove this theorem by construction. Let us define a set of vectors $\{|f_i\rangle\}$ ($i = 1, 2, \dots, n$) by

$$|f_k\rangle = |g_k\rangle - \sum_{i=1}^{k-1} \frac{\langle f_i|g_k\rangle}{\langle f_i|f_i\rangle} |f_i\rangle. \quad (1.8)$$

This set can be seen to be a mutually orthogonal set (by induction). If the $|g_k\rangle$'s are linearly independent then all the $|f_k\rangle$'s can be shown to be nonzero. Also it is evident from equation 1.8 that the $|g_k\rangle$'s can be written as a linear combination of the $|f_k\rangle$'s. This completes the proof.

Definition 7 *A linear transformation from a linearly independent nonzero set $\{|g_i\rangle\}$ to a mutually orthogonal nonzero set $\{|f_i\rangle\}$ is called ORTHOGONALIZATION. This is not a unique transformation and the one shown in equation 1.8 is just an example.*

1.3 Linear operators

An operator defined on the space \mathcal{V} is an object that maps the space \mathcal{V} onto itself. If Q is an operator then its operation on a ket $|s\rangle$ is written as $Q|s\rangle$ and $Q|s\rangle \in \mathcal{V}$. An operator Q is a linear operator if

Rule 6

$$Q(a|r\rangle + b|s\rangle) = aQ|r\rangle + bQ|s\rangle,$$

where $a, b \in \mathcal{C}$ and $|r\rangle, |s\rangle \in \mathcal{V}$.

The addition of two operators and multiplication by a complex number is defined by the following.

Definition 8

$$(aP + bQ)|s\rangle \equiv a(P|s\rangle) + b(Q|s\rangle), \quad (1.9)$$

where $a, b \in \mathcal{C}$, $|s\rangle \in \mathcal{V}$ and P and Q are linear operators (to be called just operators from here on as nonlinear operators will never be used).

Product of two operators P and Q is defined to be PQ in an obvious way.

Definition 9

$$(PQ)|s\rangle \equiv P(Q|s\rangle), \quad (1.10)$$

where $|s\rangle \in \mathcal{V}$.

In general $PQ \neq QP$. Hence, we define:

Definition 10 *The COMMUTATOR BRACKET (or just COMMUTATOR) of two operators P and Q is defined as*

$$[P, Q] = PQ - QP \quad (1.11)$$

The following identities involving commutators can be readily proved from the above definition.

$$[P, Q] = -[Q, P], \quad (1.12)$$

$$[P, Q + R] = [P, Q] + [P, R], \quad (1.13)$$

$$[P, QR] = [P, Q]R + Q[P, R], \quad (1.14)$$

$$[P, [Q, R]] + [R, [P, Q]] + [Q, [R, P]] = 0. \quad (1.15)$$

These are the same as the properties of the Poisson bracket in classical mechanics. Postulate 2 in the next chapter uses this fact.

Operation of an operator Q on a bra $\langle s|$ is written as $\langle s|Q$ and is defined as follows.

Definition 11

$$(\langle s|Q)|r\rangle \equiv \langle s|Q|r\rangle \equiv \langle s|(Q|r)\rangle \quad (1.16)$$

where $|r\rangle \in \mathcal{V}$.

Another useful definition is:

Definition 12 *The adjoint of an operator Q is called Q^\dagger and defined as*

$$Q^\dagger|s\rangle \equiv (\langle s|Q)^\dagger \quad (1.17)$$

where $|s\rangle \in \mathcal{V}$.

For the description of observables the following kind of operators will be needed.

Definition 13 *An operator H is said to be HERMITIAN (or SELF ADJOINT) if*

$$H^\dagger = H \quad (1.18)$$

1.4 Eigenstates and eigenvalues

Definition 14 *If for some operator Q , there exists a state $|q\rangle$ and a complex number q such that*

$$Q|q\rangle = q|q\rangle, \quad (1.19)$$

then q is called an EIGENVALUE of Q and $|q\rangle$ the corresponding EIGENSTATE.

It is in general possible for more than one eigenstate to have the same eigenvalue.

Definition 15 When $n(> 1)$ linearly independent eigenstates have the same eigenvalue, they are said to be (n -FOLD) DEGENERATE.

For our purposes the eigenvalues and eigenstates of hermitian operators are of particular interest. If H is a hermitian operator, some useful theorems can be proved for its eigenstates and corresponding eigenvalues.

Theorem 1.3 All eigenvalues of a hermitian operator H are real.

Proof: If $|h\rangle$ is the eigenstate corresponding to the eigenvalue h then

$$H|h\rangle = h|h\rangle \quad (1.20)$$

The adjoint of this relation is (see problem 4)

$$\langle h|H^\dagger = h^*\langle h|.$$

As H is hermitian this is the same as

$$\langle h|H = h^*\langle h|. \quad (1.21)$$

Multiplying (that is taking the inner product) equation 1.20 from the left by $\langle h|$ one gets

$$\langle h|H|h\rangle = h\langle h|h\rangle. \quad (1.22)$$

Multiplying equation 1.21 from the right by $|h\rangle$ one gets

$$\langle h|H|h\rangle = h^*\langle h|h\rangle. \quad (1.23)$$

Hence, barring the trivial case of $|h\rangle$ being the zero vector, equations 1.22 and 1.23 lead to

$$h = h^*. \quad (1.24)$$

This completes the proof.

Theorem 1.4 Eigenstates $|h_1\rangle$ and $|h_2\rangle$ of a hermitian operator H are orthogonal (i.e. $\langle h_1|h_2\rangle = 0$) if the corresponding eigenvalues h_1 and h_2 are not equal.

Proof: By definition

$$H|h_1\rangle = h_1|h_1\rangle, \quad (1.25)$$

$$H|h_2\rangle = h_2|h_2\rangle. \quad (1.26)$$

As H is hermitian, using theorem 1.3, the adjoint of equation 1.25 is seen to be

$$\langle h_1|H = h_1\langle h_1|. \quad (1.27)$$

Multiplying equation 1.26 from the left by $\langle h_1|$ one gets

$$\langle h_1|H|h_2\rangle = h_2\langle h_1|h_2\rangle. \quad (1.28)$$

Multiplying equation 1.27 from the right by $|h_2\rangle$ one gets

$$\langle h_1|H|h_2\rangle = h_1\langle h_1|h_2\rangle. \quad (1.29)$$

Subtracting equation 1.28 from equation 1.29 gives

$$(h_1 - h_2)\langle h_1|h_2\rangle = 0. \quad (1.30)$$

As $h_1 \neq h_2$ this means

$$\langle h_1|h_2\rangle = 0. \quad (1.31)$$

This completes the proof.

Corollary 1.1 *From theorem 1.2 it can be shown that the orthogonalization of a set of n -fold degenerate eigenstates produces a set of mutually orthogonal n -fold degenerate eigenstates with the same common eigenvalue.*

Corollary 1.2 *From theorem 1.4 and corollary 1.1, one can readily see that any set of linearly independent eigenstates of a hermitian operator can be linearly transformed (only the degenerate eigenstates need be transformed) to a set of mutually orthogonal eigenstates with the same eigenvalues.*

Definition 16 *A set of eigenvalues is called DISCRETE if it has a one to one correspondence with some subset of the set of integers and any real number between two successive members of the set is not an eigenvalue.*

Definition 17 *A set of eigenvalues is called CONTINUOUS if it has a one to one correspondence with the set of points on a segment of the real line.*

Hence, for a discrete set of eigenvalues (of a hermitian operator) the eigenstates can be labelled by integers and chosen such that

$$\langle h_i|h_j\rangle = n_i\delta_{ij}$$

where i and j are integers, $|h_i\rangle$ and $|h_j\rangle$ are eigenstates and δ_{ij} is the Kronecker delta (equation 1.61 gives a definition). Rule 2 can be used to choose n_i , the norm of the i -th eigenstate, to be unity. With this choice we obtain

$$\langle h_i|h_j\rangle = \delta_{ij}, \quad (1.32)$$

where i and j are integers. For continuous eigenvalues one cannot use equation 1.32 as the eigenstates cannot be labelled by integers. They will have real numbers as labels. It is very often convenient to use the eigenvalue itself as a label (unless there is a degeneracy). Hence, for continuous eigenvalues one writes the equivalent of equation 1.32 as its limiting case of successive eigenvalues getting indefinitely close. In such a limit the Kronecker delta becomes the Dirac delta function (see equation 1.63 for a definition). So, once again, using rule 2 suitably one gets

$$\langle h|h'\rangle = \delta(h - h') \quad (1.33)$$

where $|h\rangle$ and $|h'\rangle$ are the eigenstates with real number labels h and h' and $\delta(h - h')$ is the Dirac delta function. Note that in this case the norm of an eigenstate is infinite.

Definition 18 *The choice of suitable multipliers for the eigenstates (using rule 2) such that the right sides of equations 1.32 and 1.33 have only delta functions, is called NORMALIZATION and the corresponding mutually orthogonal eigenstates are called NORMALIZED EIGENSTATES or ORTHONORMAL EIGENSTATES. From here on, all eigenstates of hermitian operators will be assumed to be normalized according to equation 1.32 or equation 1.33. However, very often for brevity equation 1.32 might be used symbolically to represent both cases. As these are mutually exclusive cases there would be no confusion.*

The completeness definition of section 1.1 can now be written in terms of discrete and continuous labels.

Definition 19 *A set of states $\{|h_i\rangle\}$ with label i is said to be COMPLETE if any $|s\rangle \in \mathcal{V}$ can be written as a linear combination of the $|h_i\rangle$ i.e.*

$$|s\rangle = \sum_i a_i |h_i\rangle \quad (1.34)$$

where a_i are complex coefficients. For continuous eigenvalues the above summation is to be understood to be its obvious limit of an integral over the continuous label (or labels).

$$|s\rangle = \int a(h) |h\rangle dh \quad (1.35)$$

where $a(h)$ is a complex function of the label h .

Now one can state and prove the completeness theorem for the eigenstates of a hermitian operator. The proof presented here is not for the most general case. However, it illustrates a method that can be generalized. In a first reading this proof may be omitted.

Theorem 1.5 *An orthonormal (not necessary but convenient) set of all linearly independent eigenstates of a hermitian operator is complete.*

Proof: Let the hermitian operator be H and let the orthonormal set of all linearly independent eigenstates of H be $\{|h_i\rangle\}$ with i as the label. For convenience, the label will be chosen to be discrete ($i = 1, 2, \dots$). However, the proof can be readily extended for other discrete sets of labels as well as continuous labels.

The theorem will be proved by contradiction. Hence, it is assumed that the set $\{|h_i\rangle\}$ be not complete. From theorem 1.2 it then follows that there exists a complementary set $\{|g_i\rangle\}$ of orthonormal states that together with $\{|h_i\rangle\}$ will form a complete orthonormal set. This would mean that all $|g_i\rangle$'s are orthogonal to all $|h_i\rangle$'s. The operation of H on $|g_i\rangle$ can then be written as a linear combination of the complete set:

$$H|g_i\rangle = \sum_j a_{ij}|g_j\rangle + \sum_j b_{ij}|h_j\rangle. \quad (1.36)$$

Multiplying from the left by $\langle h_k|$ one gets

$$\langle h_k|H|g_i\rangle = b_{ik}, \quad (1.37)$$

where use is made of the orthonormality of the $|g_i\rangle$'s and $|h_i\rangle$'s. As $\langle h_k|$ is the bra adjoint to the eigenket $|h_k\rangle$ with eigenvalue h_k and H is hermitian,

$$\langle h_k|H = h_k\langle h_k|. \quad (1.38)$$

Using this in equation 1.37 one gets (using orthonormality)

$$b_{ik} = h_k\langle h_k|g_i\rangle = 0. \quad (1.39)$$

Hence, equation 1.36 becomes

$$H|g_i\rangle = \sum_j a_{ij}|g_j\rangle. \quad (1.40)$$

Now let us consider the set \mathcal{V}_c of all states that are linear combinations of the $|g_i\rangle$'s i.e.

$$|k\rangle \in \mathcal{V}_c \iff |k\rangle = \sum_i c_i|g_i\rangle, \quad (1.41)$$

for some set of complex numbers c_i . It can be readily shown (problem 4) that $\langle k|H|k\rangle/\langle k|k\rangle$ is a real number and hence would have some minimum value for all $|k\rangle \in \mathcal{V}_c$. If e is this minimum value¹ then for any $|k\rangle \in \mathcal{V}_c$

$$\langle k|H|k\rangle/\langle k|k\rangle \geq e. \quad (1.42)$$

¹If $e = -\infty$ one needs to be more careful, but the proof of the theorem still holds in an appropriate limiting sense. To be rigorous, one also needs to consider the possibility that the range of $\langle k|H|k\rangle/\langle k|k\rangle$ for all $|k\rangle$ is an open set. Then equation 1.42 does not have the possibility of equality. Here again a limiting choice is to be made for $|g_1\rangle$ such that $(a_{11} - e) \rightarrow 0$.

Without loss of generality the first of the set $\{|g_i\rangle\}$, viz. $|g_1\rangle$, could be chosen to be the one for which equation 1.42 becomes an equality (theorem 1.2) i.e.

$$\langle g_1|H|g_1\rangle = e, \quad (1.43)$$

where it is noted that $\langle g_1|g_1\rangle = 1$ from orthonormalization. Also from equations 1.40 and 1.43 one sees that

$$a_{11} = e. \quad (1.44)$$

If $|k\rangle \in \mathcal{V}_c$ then from equations 1.40, 1.41 and 1.42 one obtains

$$\sum_{ij} c_i c_j^* a_{ij} \geq e \sum_i |c_i|^2. \quad (1.45)$$

As the c_i 's are arbitrary, one may choose them all to be zero except

$$c_1 = 1, c_m = \epsilon + i\delta, \quad (1.46)$$

where ϵ and δ are real and $m \neq 1$. Then from equations 1.45 and 1.44 it follows that

$$\epsilon(a_{m1} + a_{1m}) + i\delta(a_{m1} - a_{1m}) + (\epsilon^2 + \delta^2)(a_{mm} - e) \geq 0. \quad (1.47)$$

For small enough ϵ and δ , it can be seen that the last term on the left hand side will contribute negligibly and hence, the inequality can be violated with suitable choices for the signs of ϵ and δ , unless

$$a_{m1} + a_{1m} = 0, \quad a_{m1} - a_{1m} = 0. \quad (1.48)$$

This gives

$$a_{1m} = a_{m1} = 0. \quad (1.49)$$

This being true for any $m \neq 1$, one concludes from equation 1.40 that

$$H|g_1\rangle = a_{11}|g_1\rangle. \quad (1.50)$$

This, of course, means that $|g_1\rangle$ is an eigenstate of H thus contradicting the original statement that the $|g_i\rangle$'s are not eigenstates of H . Hence, the set $\{|g_i\rangle\}$ must be empty and the set $\{|h_i\rangle\}$ must be complete. This completes the proof.

From the completeness theorem 1.5, we see that if $\{|h_i\rangle\}$ is a set of all orthonormal eigenstates of H then any state $|s\rangle$ can be written as

$$|s\rangle = \sum_i c_{si} |h_i\rangle. \quad (1.51)$$

Definition 20 The coefficient c_{si} in equation 1.51 is called the *COMPONENT* of $|s\rangle$ along $|h_i\rangle$.

Multiplying equation 1.51 from the left by $\langle h_j |$ and using orthonormality one obtains

$$c_{sj} = \langle h_j | s \rangle. \quad (1.52)$$

Replacing this in equation 1.51 we get

$$|s\rangle = \sum_i |h_i\rangle \langle h_i | s \rangle. \quad (1.53)$$

Symbolically this can be written as

$$|s\rangle = \left(\sum_i |h_i\rangle \langle h_i| \right) |s\rangle, \quad (1.54)$$

giving the object in parenthesis the meaning of an operator in an obvious sense. But this operator operated on any state produces the same state. Hence, it is the *identity operator*

$$I = \sum_i |h_i\rangle \langle h_i|. \quad (1.55)$$

Equation 1.55 can be seen to be a compact mathematical statement of the completeness of the eigenstates $\{|h_i\rangle\}$.

Very often it is useful to define the projection operators corresponding to each $|h_i\rangle$.

Definition 21 *The projection operator for $|h_i\rangle$ is defined to be*

$$P_i = |h_i\rangle \langle h_i| \quad (1.56)$$

which selects out the part of the vector $|s\rangle$ in the “direction” $|h_i\rangle$.

$$P_i |s\rangle = c_{si} |h_i\rangle. \quad (1.57)$$

Also from equations 1.55 and 1.56

$$I = \sum_i P_i. \quad (1.58)$$

We shall sometimes use equations 1.55 and 1.58 symbolically in the same form for continuous eigenvalues as well. However, it should be understood to mean

$$I = \int |h\rangle \langle h| dh \quad (1.59)$$

for the real valued label h . In the same spirit equation 1.53 will also be used for continuous eigenvalues and would be interpreted as

$$|s\rangle = \int |h\rangle \langle h | s \rangle dh. \quad (1.60)$$

In fact in future chapters, as a general rule, a summation over indices of a complete set of eigenvalues will be understood to be an integral over eigenvalues for continuous eigenvalues.

1.5 The Dirac delta function

The Kronecker delta is usually defined as

Definition 22

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \quad (1.61)$$

where i and j are integers.

However, the following equivalent definition is found to be useful for the consideration of a continuous index analog of the Kronecker delta.

$$\sum_j \delta_{ij} f_j = f_i, \quad (1.62)$$

where i and j are integers and f_i represents the i -th member of an arbitrary sequence of finite numbers.

The Dirac delta is an analog of the Kronecker delta with continuous indices. For continuous indices the i and j can be replaced by real numbers x and y and the Dirac delta is written as $\delta(x - y)$. The difference $(x - y)$ is used as the argument because the function can be seen to depend only on it. Likewise f_i is replaced by a function $f(x)$ of one real variable. $f(x)$ must be finite for all x . Hence, the continuous label analog of equation 1.62 produces the following definition of the Dirac delta function.

Definition 23

$$\int \delta(x - y) f(y) dy = f(x), \quad (1.63)$$

where $f(x)$ is finite everywhere. An integral with no limits shown explicitly is understood to have the limits $-\infty$ to $+\infty$.

From this definition it is seen that, $f(x)$ being an arbitrary function, the only way equation 1.63 is possible is if $\delta(x - y)$ is zero everywhere except at $x = y$. At $x = y$, $\delta(x - y)$ would have to be infinite as dy is infinitesimal. Hence, the following are true for the Dirac delta function.

$$\delta(0) = \infty, \quad (1.64)$$

$$\delta(x) = 0 \quad \text{if } x \neq 0. \quad (1.65)$$

Because of the infinity in equation 1.64, the Dirac delta has meaning only when multiplied by a finite function and integrated. Some identities involving the Dirac delta (in the same

integrated sense) that can be deduced from the defining equation 1.63 are

$$\delta(x) = \delta(-x), \quad (1.66)$$

$$\int \delta(x) dx = 1, \quad (1.67)$$

$$x\delta(x) = 0, \quad (1.68)$$

$$\delta(ax) = \frac{1}{|a|}\delta(x) \quad \text{for } a \neq 0, \quad (1.69)$$

$$\delta(x^2 - a^2) = \frac{1}{2|a|}[\delta(x - a) + \delta(x + a)] \quad \text{for } a > 0, \quad (1.70)$$

$$\int \delta(a - x)\delta(x - b) dx = \delta(a - b), \quad (1.71)$$

$$f(x)\delta(x - a) = f(a)\delta(x - a). \quad (1.72)$$

The derivatives of a Dirac delta can be defined once again in the sense of an integral. I shall consider only the first derivative $\delta'(x)$.

$$\int \delta'(x - y)f(y) dy = -f(y)\delta(x - y)|_{-\infty}^{+\infty} + \int \delta(x - y)f'(y) dy, \quad (1.73)$$

where a prime denotes a derivative with respect to the whole argument of the function. Thus

$$\int \delta'(x - y)f(y) dy = f'(x). \quad (1.74)$$

Some identities involving the $\delta'(x)$ can be derived in the same fashion.

$$\delta'(x) = -\delta'(-x), \quad (1.75)$$

$$x\delta'(x) = -\delta(x). \quad (1.76)$$

To understand the Dirac delta better it is very often written as the limit of some better known function. For example,

$$\delta(x) = \lim_{g \rightarrow \infty} \frac{\sin(gx)}{\pi x}, \quad (1.77)$$

$$\delta(x) = \lim_{a \rightarrow 0} \frac{1}{a\sqrt{\pi}} \exp\left(-\frac{x^2}{a^2}\right), \quad (1.78)$$

$$\delta(x) = \frac{1}{2\pi} \int \exp(ikx) dk. \quad (1.79)$$

Problems

1. The norm $\langle s|s \rangle$ of a vector $|s\rangle$ is sometimes written as $||s||^2$. In this chapter the norm has been defined from the inner product. However, it is possible to first define the

norm and then the inner product as its consequence. Such an approach needs fewer rules but is more unwieldy. The inner product is then defined as:

$$\begin{aligned} \langle r|s \rangle = & \frac{1}{2}[||r\rangle + |s\rangle|^2 - i||r\rangle + i|s\rangle|^2 \\ & + (i-1)(||r\rangle|^2 + ||s\rangle|^2)]. \end{aligned}$$

Prove this result using the definition of inner product and norm as given in this chapter.

2. In equation 1.8, show that a linearly *dependent* set $\{|g_i\rangle\}$ would give some of the $|f_i\rangle$'s to be the zero vector.
3. Using the defining equation 1.11 of the commutators prove the identities in equations 1.12 through 1.15.
4. Prove the following operator relations (for all operators P and Q , $|s\rangle \in \mathcal{V}$, and $a, b \in \mathcal{C}$)
 - (a) $(Q|s\rangle)^\dagger = \langle s|Q^\dagger$
 - (b) $Q^{\dagger\dagger} = Q$
 - (c) $(aP + bQ)^\dagger = a^*P^\dagger + b^*Q^\dagger$
 - (d) $(PQ)^\dagger = Q^\dagger P^\dagger$
 - (e) PQ is hermitian if P and Q are hermitian and $[P, Q] = 0$.
 - (f) For a hermitian operator H and $|s\rangle \in \mathcal{V}$, $\langle s|H|s\rangle$ is real.
5. Prove the corollary 1.1.
6. Using the defining equation 1.63 of the Dirac delta, prove the identities in equations 1.66 through 1.72. For the derivative of the Dirac delta prove the identities in equations 1.75 and 1.76. [Hint: Remember that these identities have meaning only when multiplied by a finite function and integrated.]

Chapter 2

The Laws (Postulates) of Quantum Mechanics

In the following, the term postulate will have its mathematical meaning i.e. an assumption used to build a theory. A law is a postulate that has been experimentally tested. All postulates introduced here have the status of laws.

2.1 A lesson from classical mechanics

There is a fundamental difference in the theoretical structures of classical and quantum mechanics. To understand this difference, one first needs to consider the structure of classical mechanics independent of the actual theory given by Newton. It is as follows.

1. The fundamental measured quantity (or the descriptor) of a system is its trajectory in *configuration space* (the space of all independent position coordinates describing the system). The configuration space has dimensionality equal to the number of degrees of freedom (say n) of the system. So the trajectory is a curve in n dimensional space parametrized by time. If x_i is the i -th coordinate, then the trajectory is completely specified by the n functions of time $x_i(t)$. These functions are all observable.
2. A predictive theory of classical mechanics consists of equations that describe some initial value problem. These equations enable us to determine the complete trajectory $x_i(t)$ from data at some initial time. The Newtonian theory requires the x_i and their time derivatives as initial data.
3. The $x_i(t)$ can then be used to determine other observables (sometimes conserved quantities) like energy, angular momentum etc.. Sometimes the equations of motion

can be used directly to find such quantities of interest.

The above structure is based on the nature of classical measurements. However, at small enough scales, such classical measurements (like the trajectory) are found to be experimentally meaningless. Thus, a different theoretical structure becomes necessary. This structure is that of quantum mechanics. The structure of quantum mechanics, along with the associated postulates, will be stated in the following section. It is itemized to bring out the parallels with classical mechanics.

The reader must be warned that without prior experience in quantum physics the postulates presented here might seem rather ad hoc and “unphysical”. But one must be reminded that in a first course in classical physics, Newton’s laws of motion might seem just as ad hoc. Later, a short historical background will be given to partially correct this situation. About the “unphysical” nature of these postulates, very little can be done. Phenomena like the falling of objects due to gravity are considered “physical” due to our long term exposure to their repeated occurrence around us. In contrast, most of the direct evidence of quantum physics is found at a scale much smaller than everyday human experience. This makes quantum phenomena inherently “unphysical”. Hence, all one can expect of the following postulates is their self consistency and their ability to explain all observed phenomena within their range of applicability. To make quantum phenomena appear as “physical” as classical phenomena, one needs to repeatedly experience quantum aspects of nature. Hence, this text (like most others) tries to provide as many examples as possible.

At first sight, the reader might also find the postulates to be too abstract and computationally intractable. The next two chapters should go a long way in correcting this problem.

2.2 The postulates of quantum mechanics

In the following, the postulates of quantum mechanics are presented within a theoretical structure that has a flavor similar to classical mechanics. The reader is encouraged to observe similarities and differences of the two theories.

1. The descriptor is given by the zeroth postulate. Its relation to measurements is somewhat indirect (see postulates 2 through 5).

Postulate 0 *The complete descriptor (but not a measured quantity) of a system is its state vector $|s\rangle$ and the complete descriptor of an observable q is a hermitian operator Q defined to operate on any $|s\rangle \in \mathcal{V}$. $|s\rangle$, in general, depends on as many variables as there are degrees of freedom and time.*

2. Predictive power comes from the initial value problem described by the following somewhat generalized Schrödinger equation.

Postulate 1

$$i\hbar \frac{\partial}{\partial t} |s\rangle = H|s\rangle, \quad (2.1)$$

where H is the hamiltonian operator obtained by replacing the classical position and momentum in the classical expression of the hamiltonian by their quantum operator analogs. \hbar ($\approx 1.0545 \times 10^{-34}$ Joule.sec.) is a universal constant. $2\pi\hbar (= h)$ is usually called Planck's constant.

3. Quantum measurements of observables are conceptually distinctly different from classical measurements. Classical measurements can be theoretically predicted with indefinite accuracy (but the theory fails completely at smaller scales). Quantum mechanics can predict only the probabilities of measurements (at all scales but would very often be impractical at larger scales). Every observable has an associated operator that operates on the space of state vectors \mathcal{V} .

Postulate 2 All measurable aspects of observables are completely determined by the mutual commutators of their respective operators. These commutators are determined by the following transition from classical to quantum.

$$\{q, p\} \longrightarrow \frac{[Q, P]}{i\hbar}, \quad (2.2)$$

where q and p are classical observables with quantum operator analogs given by Q and P respectively and $\{, \}$ is the Poisson bracket.

Postulate 3 The possible results of measurement of an observable, represented by the operator Q , are its eigenvalues q_i only.

Postulate 4 If a system is in a state $|s\rangle$ and a measurement of an observable represented by the operator Q is made on it, the probability that the result will be the eigenvalue q_i is proportional to

$$\sum_{deg} \langle q_i | s \rangle \langle s | q_i \rangle = \sum_{deg} |\langle q_i | s \rangle|^2 \quad (2.3)$$

where $|q_i\rangle$ is an eigenstate corresponding to q_i and the summation is over all degenerate states with the same eigenvalue q_i . The proportionality constant can be chosen arbitrarily for computational convenience. It is very often chosen to keep the total probability of all outcomes as unity.

Postulate 5 If the result of the measurement is indeed q_i , then after the measurement the system will collapse into a corresponding eigenstate $|q_i\rangle$.

This completes the set of postulates necessary in a theory of quantum mechanics. To understand the theory we need to use these postulates in physical examples. The rest of the book will be seen to be creative applications of mathematics to do just this.

2.3 Some history of the postulates

At the end of the nineteenth century one of the major experimental results that baffled classical physicists was the blackbody radiation spectrum. Classical physics had been highly successful in explaining and predicting a large variety of phenomena but for the radiation spectrum of a blackbody it gave the absurd result of infinite emission at infinite frequency (sometimes called the “ultraviolet catastrophe”). Planck was able to suitably “explain” the experimentally observed spectrum by a hitherto arbitrary assumption that the energies of oscillators producing radiation of frequency ν can have energies only in integer multiples of $h\nu$. This was the origin of the constant h . At the same time discreteness was also noticed in the frequency spectra of atoms. Such observations led to the hypothesis (by de Broglie) that just as Planck had noticed that electromagnetic waves have a discrete (particle) nature, particles (like electrons) have a wave nature. The wavelength λ of a particle of momentum p is given by h/p . This led Schrödinger to the equation in postulate 1 in a particular representation that will be discussed as the position representation in the next chapter. The “wavefunction” of Schrödinger’s equation is the equivalent of the state vector (postulate 0). The generalized properties of the state vector (linearity etc.) have their origin in the wavefunction. The state vector was later chosen as the descriptor to allow greater generality, mathematical convenience, and economy in concepts.

The postulates 2 through 5 were discovered in the process of consolidating experimental observations with a theory of wavefunctions (or state vectors).

After this rather short and oversimplified narration of the history of quantum physics, in general, the historical approach will be avoided in this text. This is not to downplay the role of history, but to avoid many of the confusions that arose in the historical development of the subject. As we now have the advantage of twenty-twenty hindsight, we shall use it.

Chapter 3

Popular Representations

To get numerical results from the laws discussed in the previous chapter it is very often convenient to use some less abstract “representations” of state vectors and operators.

We first note (from postulate 0) that operators corresponding to observables are hermitian and therefore have real eigenvalues (theorem 1.3). Hence, from postulate 3, we see that measured quantities will be real as expected. It was shown in theorem 1.5 that the eigenstates of a hermitian operator form a complete orthonormal set. This makes it natural to expand any state as a linear combination of the eigenstates of an operator corresponding to an observable. For example, if Q is such an operator with $|q_i\rangle$ as its eigenstates, then an arbitrary state $|s\rangle$ can be written as

$$|s\rangle = \sum_i c_{si} |q_i\rangle, \quad (3.1)$$

where the components c_{si} are given by (definition 20 on page 11)

$$c_{si} = \langle q_i | s \rangle. \quad (3.2)$$

Definition 24 *The set of components, c_{si} , of $|s\rangle$ along the eigenstates of the operator Q completely describes the state $|s\rangle$ and hence, will be called the Q REPRESENTATION of $|s\rangle$.*

Two popular representations are discussed below.

3.1 The position representation

The most popular choice of representation is that of the position vector operator \mathbf{R} for a single particle system. The eigenvalues of \mathbf{R} are known to be continuous as every value of

position is measurable. The corresponding eigenstates are assumed to be nondegenerate for now¹. Hence, they can be uniquely labeled by the eigenvalues \mathbf{r} i.e. $|\mathbf{r}\rangle$. Then, from equation 3.2, the position representation of the state $|s\rangle$ could be written as the components

$$\langle \mathbf{r}|s\rangle \equiv \Psi_s(\mathbf{r}). \quad (3.3)$$

So this set of components can now be seen as values of a function at different positions \mathbf{r} . This function, $\Psi_s(\mathbf{r})$, is conventionally known as the wavefunction because of its wave nature in some typical situations. Historically, this representation of the state has been the most popular for two reasons. First, the wave nature of this representation had early experimental consequences. Second, it will be seen to reduce most problems to differential equations. The mathematics of differential equations, including methods of approximation, is very well known and makes problem solving easier.

The wavefunction $\Psi_s(\mathbf{r})$ can be seen from another point of view. Applying postulate 4 for the position operator, one sees that the probability of finding the particle at the position \mathbf{r} is proportional to

$$\langle \mathbf{r}|s\rangle\langle s|\mathbf{r}\rangle = \Psi_s(\mathbf{r})\Psi_s^*(\mathbf{r}). \quad (3.4)$$

This is usually known as the *probability density* which when integrated over a finite volume gives the probability of finding the particle in that volume.

We shall now derive the forms of the position and momentum operators and their eigenstates in the position representation just defined. We shall do this in one space dimension. Extension to three dimensions is straightforward (problem 1).

The eigenstates of X , the position operator in one dimension, are $|x\rangle$ with corresponding eigenvalues x i.e.

$$X|x\rangle = x|x\rangle. \quad (3.5)$$

For an arbitrary state $|s\rangle$ the result of operation by X is $X|s\rangle$. Its position representation is $\langle x|X|s\rangle$. Using equations 3.3 and 3.5 and the hermiticity of X one gets

$$\langle x|X|s\rangle = x\langle x|s\rangle = x\Psi_s(x). \quad (3.6)$$

This shows that in the position representation the effect of operating by X is just multiplication of the wavefunction by the corresponding eigenvalue x .

To find the position representation of the momentum operator we note that the representations of the operators X (position) and P (momentum) must satisfy postulate 2. Hence, as their classical Poisson bracket is 1, we may write

$$[X, P] \equiv XP - PX = i\hbar, \quad (3.7)$$

¹A degeneracy would mean that there are degrees of freedom other than just position. Such internal degrees of freedom have no classical analog and can be ignored for now. However, quantum theory allows such degrees of freedom and experiments have verified their existence. Hence, they will be discussed separately in the later chapter on particle spin.

or

$$[XP - PX]|s\rangle = i\hbar|s\rangle, \quad (3.8)$$

for any $|s\rangle$. If the eigenkets of X are $|x\rangle$, then the position representation of equation 3.8 is obtained by multiplying on the left by the corresponding eigenbra $\langle x|$.

$$\langle x|[XP - PX]|s\rangle = i\hbar\langle x|s\rangle. \quad (3.9)$$

Inserting the identity operator of equation 1.59 in two places in the equation we get

$$\int \int [\langle x|X|x'\rangle\langle x'|P|x''\rangle\langle x''|s\rangle - \langle x|P|x'\rangle\langle x'|X|x''\rangle\langle x''|s\rangle] dx' dx'' = i\hbar\langle x|s\rangle. \quad (3.10)$$

Using the fact that $|x\rangle$ are eigenstates of X and the orthonormality of continuous eigenstates one obtains

$$\int \int [x\delta(x-x')\langle x'|P|x''\rangle - \langle x|P|x'\rangle x''\delta(x'-x'')]\langle x''|s\rangle dx' dx'' = i\hbar\langle x|s\rangle. \quad (3.11)$$

Integrating over x' gives

$$\int \langle x|P|x''\rangle(x-x'')\langle x''|s\rangle dx'' = i\hbar\langle x|s\rangle. \quad (3.12)$$

Using the defining equation 1.63 of the Dirac delta² and equation 1.76, one notices that the above equation is satisfied by

$$\langle x|P|x''\rangle = i\hbar \frac{\delta(x-x'')}{x-x''} = -i\hbar\delta'(x-x''). \quad (3.13)$$

Now in the position representation the operation of P on any state $|s\rangle$ would be (inserting an identity operator)

$$\langle x|P|s\rangle = \int \langle x|P|x'\rangle\langle x'|s\rangle dx'. \quad (3.14)$$

Then using equations 3.13 and 1.74, one obtains

$$\langle x|P|s\rangle = -i\hbar \frac{\partial}{\partial x} \langle x|s\rangle = -i\hbar \frac{\partial}{\partial x} \Psi_s(x). \quad (3.15)$$

Here a partial derivative is used as the function, in general, depends on time as well. Hence, in the position representation the effect of operating by P on a state is equivalent to operating the corresponding wavefunction by the differential operator $-i\hbar\partial/\partial x$.

²Mathematical note: To use equation 1.63 one must have $\langle x|s\rangle$ to be strictly a finite function of x . We do not know this for a fact even though from the probability interpretation it might seem reasonable. However, it is possible to extend equation 1.63 to also apply for a certain class of $f(x)$ that are infinite at some points. For example, from equation 1.71 one sees that the $f(x)$ of equation 1.63 itself could be a Dirac delta. For fear of getting too deep into mathematics, I shall restrict the discussion to only those wavefunctions (finite or infinite) that satisfy equation 1.63 when substituted for $f(x)$. Problem 5 demonstrates how some types of infinite wavefunctions cannot be allowed. The mathematically oriented reader might try to solve for $\langle x|P|x''\rangle$ from equation 3.12 for a more general class of wavefunctions.

It is now possible to find the position representations of the eigenstates of position and momentum. The position eigenstates are $|x\rangle$. Their position representation at the position x' is, by definition, $\langle x'|x\rangle$. As the position eigenstates must be orthonormal

$$\langle x'|x\rangle = \delta(x - x'). \quad (3.16)$$

The eigenstates of momentum are $|p\rangle$ (with eigenvalue p) and their position representation is, by definition, $\langle x|p\rangle$. From the definition of eigenstates and equation 3.15

$$-i\hbar \frac{\partial}{\partial x} \langle x|p\rangle = p \langle x|p\rangle. \quad (3.17)$$

The solution is

$$\langle x|p\rangle = A \exp(ixp/\hbar). \quad (3.18)$$

Normalization gives the value of A . That is

$$\begin{aligned} \delta(p - p') = \langle p|p'\rangle &= \int \langle p|x\rangle \langle x|p'\rangle dx \\ &= \int \langle x|p\rangle^* \langle x|p'\rangle dx \\ &= A^* A \int \exp[ix(p' - p)/\hbar] dx \\ &= A^* A 2\pi\hbar \delta(p - p'). \end{aligned}$$

This gives

$$A = (2\pi\hbar)^{-1/2} \exp(i\alpha), \quad (3.19)$$

where α is an arbitrary real number that could depend on time. But it has no physical significance due to rule 2. Hence, choosing $\alpha = 0$, equation 3.18 gives

$$\langle x|p\rangle = (2\pi\hbar)^{-1/2} \exp(ixp/\hbar). \quad (3.20)$$

3.2 The momentum representation

Another popular representation is the momentum representation. It is analogous to the position representation. The momentum representation of a state $|s\rangle$ would be a function of momentum eigenvalues given by the components

$$\Phi_s(p) = \langle p|s\rangle. \quad (3.21)$$

The effect of operating $|s\rangle$ by momentum P in the momentum representation would be like multiplying by p .

$$\langle p|P|s\rangle = p \langle p|s\rangle. \quad (3.22)$$

The effect of operating $|s\rangle$ by X would be

$$\langle p|X|s\rangle = i\hbar \frac{\partial}{\partial p} \langle p|s\rangle = i\hbar \frac{\partial}{\partial p} \Phi_s(p). \quad (3.23)$$

The momentum representation of the eigenstates of momentum are

$$\langle p'|p\rangle = \delta(p' - p). \quad (3.24)$$

The momentum representation of the eigenstates of position are

$$\langle p|x\rangle = (2\pi\hbar)^{-1/2} \exp(-ipx/\hbar). \quad (3.25)$$

Problems

1. Generalize equations 3.6, 3.15, 3.16 and 3.20 for three dimensions.
2. Derive equations 3.22, 3.23, 3.24 and 3.25.
3. Generalize the results of problem 2 for three dimensions.
4. For any state $|s\rangle$, show that its momentum representation is a Fourier transform of its position representation. [Hint: Use equation 1.59]
5. If the position representation (wavefunction) of a state $|s\rangle$ goes to infinity (monotonically) at infinity, show that its momentum representation is infinite for all p .
6. Consider two arbitrary state vectors $|r\rangle$ and $|s\rangle$. Let their respective position representations be $\Psi_r(x)$ and $\Psi_s(x)$ and their respective momentum representations be $\Phi_r(p)$ and $\Phi_s(p)$. Show that the inner product $\langle r|s\rangle$ is given by

$$\langle r|s\rangle = \int \Psi_r^* \Psi_s dx = \int \Phi_r^* \Phi_s dp.$$

Chapter 4

Some Simple Examples

4.1 The Hamiltonian, conserved quantities and expectation value

The observational philosophy of quantum mechanics is so different from that of classical mechanics that we need to discuss it in more concrete terms before considering examples. From the laws of quantum mechanics (postulate 4) we have learnt that predictions are only probabilistic. Hence, given a system in a state $|s\rangle$, the result of a measurement on it will in general be different at different times. Furthermore, as a result of the first measurement the state of the system might change violently as it has to transform into an eigenstate of the operator just measured (postulate 5). What, then, would be the use of such a measurement? It seems that a measurement made at any time will say very little about later measurements and without such predictive power a theory has little use.

However, the situation is not that hopeless. Certain measurements can still be predicted rather well by the quantum theory. For example consider a conservative system¹ with a hamiltonian (same as energy for our purposes) operator H . The following theorem shows that energy measurements in such a system *are* predictable.

Theorem 4.1 *For a conservative system an energy eigenstate changes with time only by a multiplicative factor and hence, stays in the same physical state.*

Proof: Let the eigenstates, $|E\rangle$, of the hamiltonian, H , be labeled by E , the eigenvalues.

¹Note: This is usually not a restrictive assumption in quantum mechanics as most quantum systems of interest are microscopic in nature where all forms of energy loss can be accounted for and included in the system to make it conservative. Hence, most of the text will deal with conservative systems and when a nonconservative system is to be studied, special care will be taken.

As the system is conservative, H has no explicit time dependence and hence, E will be time independent. Let the eigenstate $|E\rangle$ change to some state $|E\rangle_t$ in time t . $|E\rangle_t$ is not necessarily an eigenstate of H . From the Schrödinger equation (postulate 1)

$$i\hbar \frac{\partial}{\partial t} |E\rangle_t = H |E\rangle_t. \quad (4.1)$$

The known initial condition at time $t = 0$ is

$$|E\rangle_0 = |E\rangle \quad (4.2)$$

From equation 4.1 we see that in an infinitesimal time dt after $t = 0$ the state $|E\rangle$ changes to $|E\rangle_{dt}$ given by

$$|E\rangle_{dt} = (1 - iHdt/\hbar)|E\rangle = (1 - iEdt/\hbar)|E\rangle$$

as E is the energy eigenvalue of $|E\rangle$. If n such increments in time are made successively such that $n \rightarrow \infty$, $dt \rightarrow 0$ and $ndt = t$ (finite t), then one obtains

$$|E\rangle_t = \lim_{n \rightarrow \infty} (1 - iEdt/\hbar)^n |E\rangle, \quad (4.3)$$

which can be seen to give (see problem 1)

$$|E\rangle_t = \exp(-iEt/\hbar)|E\rangle. \quad (4.4)$$

So an eigenstate of energy changes in time only by a multiplicative factor which means that it remains the same eigenstate and from rule 2, it is evident that there is no physical change. This completes the proof.

Now if a measurement of energy yields the value E , we know from postulate 5, that the system collapses into the eigenstate $|E\rangle$. Theorem 4.1 states that once this happens there is no more temporal change in the state of the system (unless otherwise disturbed). If another measurement of energy is made on the system after some time (with no other disturbance) the probability of obtaining a value E' is given by postulate 4 to be related to $|\langle E'|E\rangle|^2$. From the orthogonality of eigenstates of H , this is seen to give zero probability of E' being anything other than E . This is perfect predictability and is restricted by experimental errors alone (like in classical mechanics).

Such predictability of repeated energy measurements makes the hamiltonian a very special operator in quantum mechanics. However, for specific problems, one may find other observables which have the same predictability in repeated measurements. Such observables are called conserved quantities and are defined as follows.

Definition 25 *An observable is a CONSERVED QUANTITY if repeated measurements of it at different times result in the same value as long as the system is not disturbed in any way between measurements.*

To identify such observables we shall use the following theorem.

Theorem 4.2 *For a conservative system with hamiltonian H , an observable Q (with no explicit time dependence) is a conserved quantity if and only if $[Q, H] = 0$ (i.e. Q and H commute).*

Proof: We shall first prove that if Q is a conserved quantity $[Q, H] = 0$. Suppose a measurement of Q results in the eigenvalue q . Hence, the resulting eigenstate of Q is one of a set of some n_q -fold degenerate eigenstates with eigenvalue q . This state will be labelled as $|qi\rangle$, q giving the eigenvalue and $i (= 1, 2, \dots, n_q)$ labelling the different degenerate states. At a time t after the measurement, the same state will change to $|qi\rangle_t$. $|qi\rangle$ can be expanded in eigenstates of H as

$$|qi\rangle = \sum_E a_{qiE} |E\rangle. \quad (4.5)$$

The sum over all energies E is really a shorthand for sum or integral over all energy eigenstates which means that degenerate eigenstates are individually included. From the result of problem 2 one can also see that

$$|qi\rangle_t = \sum_E a_{qiE} |E\rangle_t. \quad (4.6)$$

Using equation 4.4 this gives

$$|qi\rangle_t = \sum_E a_{qiE} \exp(-iEt/\hbar) |E\rangle. \quad (4.7)$$

As Q does not depend on time explicitly, its set of eigenvalues are unchanged with time. So it is meaningful to say that if Q is a conserved quantity then q should be its only possible measured value in the time developed state $|qi\rangle_t$ as well. In other words, $|qi\rangle_t$ must be a linear combination of the n_q degenerate eigenstates of Q with eigenvalue q .

$$|qi\rangle_t = \sum_{j=1}^{n_q} u_{ij}(t) |qj\rangle. \quad (4.8)$$

Hence, $|qi\rangle_t$ is an eigenstate of Q at all times and to satisfy equation 4.7 at all times the energy eigenstates must also be eigenstates of Q (for degenerate energy eigenstates some suitable linear combination may have to be chosen). Thus we conclude that for Q to be a conserved quantity, there must exist a complete set of simultaneous eigenstates of Q and H . We shall label these eigenstates by the corresponding eigenvalues of both operators i.e. $|qE\rangle$ (the labels for distinguishing degenerate states will be suppressed for convenience).

Now let us expand an arbitrary state $|s\rangle$ as a linear combination of this complete set of eigenstates.

$$|s\rangle = \sum_{qE} c_{qE} |qE\rangle. \quad (4.9)$$

Then it follows that

$$\begin{aligned} QH|s\rangle &= \sum_{qE} c_{qE} qE |qE\rangle \\ &= \sum_{qE} c_{qE} HQ |qE\rangle \\ &= HQ|s\rangle. \end{aligned}$$

Hence, a necessary condition for Q to be a conserved quantity is

$$[Q, H]|s\rangle = 0.$$

As $|s\rangle$ is an arbitrary state, one may write the condition as

$$[Q, H] = 0. \quad (4.10)$$

Now it will be shown that equation 4.10 is also a sufficient condition. Once again, let $|qi\rangle$, ($i = 1, 2, \dots, n_q$) denote a set of n_q -fold degenerate eigenstates of Q with eigenvalue q . As Q is not explicitly dependent on time, q will be independent of time. Then, for all i

$$Q|qi\rangle = q|qi\rangle. \quad (4.11)$$

Operating this with H gives

$$HQ|qi\rangle = qH|qi\rangle. \quad (4.12)$$

Given that $[Q, H] = 0$, this leads to

$$QH|qi\rangle = qH|qi\rangle. \quad (4.13)$$

This means that the state $H|qi\rangle$ is also an eigenstate of Q with eigenvalue q . Hence, it must be some linear combination of the n_q degenerate states i.e.

$$H|qi\rangle = \sum_{j=1}^{n_q} c_{ij} |qj\rangle. \quad (4.14)$$

The above equation can be used to show that repeated operations (meaning higher powers) of H on $|qi\rangle$ will still produce some linear combination of the n_q degenerate states. This leads to the conclusion that the time development operator of problem 2, operated on $|qi\rangle$ will also produce a linear combination of the same n_q degenerate states i.e.

$$|qi\rangle_t = U_t(t)|qi\rangle = \sum_{j=1}^{n_q} u_{ij}(t)|qj\rangle. \quad (4.15)$$

So the time developed state $|qi\rangle_t$ is an eigenstate of Q with eigenvalue q at all times. Hence, the measured value remains q at all times thus showing Q to be a conserved quantity. This completes the proof.

Conserved quantities are known to be of importance in classical mechanics as they are often used to label specific trajectories. Correspondingly, in quantum mechanics state vectors are labeled by eigenvalues of conserved quantities (e.g. energy, angular momentum etc.). Further, there is a classical result that has the same physical significance as theorem 4.2:

$$\frac{dq}{dt} = \{q, H\} + \frac{\partial q}{\partial t}, \quad (4.16)$$

where q represents the classical observable corresponding to the quantum operator Q . The last term, in the above equation, is nonzero only when q depends explicitly on time. In quantum mechanics explicit time dependence of observables is uncommon². Hence, the vanishing of the commutator brackets in quantum mechanics would classically mean the vanishing of the Poisson brackets (postulate 2) giving q to be a classically conserved quantity. One can also prove the following quantum result that looks more like the classical relation of equation 4.16.

Theorem 4.3 *If Q is an observable and $|r\rangle$ and $|s\rangle$ are two arbitrary states then*

$$\frac{d}{dt}\langle r|Q|s\rangle = \frac{\langle r|[Q, H]|s\rangle}{i\hbar} + \langle r|\frac{\partial Q}{\partial t}|s\rangle$$

Proof: Using the product rule for derivatives and postulate 1,

$$\begin{aligned} \frac{d}{dt}\langle r|Q|s\rangle &= \left(\frac{\partial}{\partial t}\langle r|\right)Q|s\rangle + \langle r|Q\left(\frac{\partial}{\partial t}|s\rangle\right) + \langle r|\frac{\partial Q}{\partial t}|s\rangle \\ &= \left(-\frac{\langle r|H}{i\hbar}\right)Q|s\rangle + \langle r|Q\left(\frac{H|s\rangle}{i\hbar}\right) + \langle r|\frac{\partial Q}{\partial t}|s\rangle \\ &= \frac{\langle r|[Q, H]|s\rangle}{i\hbar} + \langle r|\frac{\partial Q}{\partial t}|s\rangle. \end{aligned}$$

This completes the proof.

In quantum mechanics theorem 4.3 is not as useful as theorem 4.2 because it does not give the actual measured values explicitly. However, theorem 4.3 can be used to find the time

²For explicitly time dependent observables, $\partial q/\partial t \neq 0$. In general, $\{q, H\}$ depends on the properties of the specific system through H , but $\partial q/\partial t$ does not. This means $\{q, H\}$ cannot exactly cancel $\partial q/\partial t$ on the right side of equation 4.16. So, explicitly time dependent observables cannot be classically conserved quantities. In quantum mechanics such nonconserved quantities have limited predictability and thus are of lesser importance

dependence of the average measured value of any operator. For this we need to define the average measured value in quantum mechanics which is the so called expectation value. In giving meaning to an average value in quantum mechanics, one has to be careful. Making a measurement on a state can change it so radically that making repeated measurements over time and then averaging (time average) has no physically useful meaning. Hence, the meaning of an average must be that of an ensemble average as stated below.

Definition 26 *The EXPECTATION VALUE $\langle Q \rangle_s$ of an observable Q in a state $|s\rangle$ is defined as the average of Q measurements made on a large number (tending to infinity) of identical systems all in the same state $|s\rangle$ with no two measurements made on the same system.*

Consider a state $|s\rangle$. For an observable Q , the probability of measuring its eigenvalue q , in this state, is known from postulate 4. Hence, using this postulate and the definition of an average, the expectation value measured for a large number (tending to infinity) of systems all in state $|s\rangle$ would be

$$\langle Q \rangle_s = \frac{\sum_q q \langle s|q\rangle \langle q|s\rangle}{\sum_q \langle s|q\rangle \langle q|s\rangle}, \quad (4.17)$$

where $|q\rangle$ denotes an eigenstate with eigenvalue q and \sum_q a sum over all eigenstates (degenerate states considered individually). Then using equation 1.55 twice

$$\langle Q \rangle_s = \frac{\sum_q \langle s|Q|q\rangle \langle q|s\rangle}{\langle s|s\rangle} = \frac{\langle s|Q|s\rangle}{\langle s|s\rangle}. \quad (4.18)$$

Then theorem 4.3 will give

$$\frac{d}{dt} \langle Q \rangle_s = \frac{\langle [Q, H] \rangle_s}{i\hbar} + \left\langle \frac{\partial Q}{\partial t} \right\rangle_s \quad (4.19)$$

This is a generalization of what is known as Ehrenfest's theorem. It provides a means of comparison of classical and quantum measurements. It is seen that averages of measurements (expectation values) in quantum mechanics obey the classical equations of motion given by equation 4.16. This is in accordance with the idea that classical measurements are on larger scale objects and hence so inaccurate that only averages of quantum measurements adequately agree with them.

4.2 Free particle in one dimension

To understand the principles discussed in chapter 2 and to use some of the mathematical results obtained in chapter 3 and this chapter, we will study the simplest possible system viz. the one dimensional free particle. The classical case of this problem is quite trivial as

it would give the solution to be a constant velocity trajectory. In quantum the problem is not as trivial and does merit discussion. It is to be noted that for a particle to show quantum behavior it must be small enough e.g. an electron.

The form of the Schrödinger equation tells us that the system is described completely by the hamiltonian H . From classical physics the form of the free particle hamiltonian is known to be

$$H = \frac{P^2}{2m}, \quad (4.20)$$

where P is the momentum and m the mass. In quantum, P is known to be an operator. We shall now try to predict the measurement of three common observables viz. momentum, energy and position.

4.2.1 Momentum

We already know P has continuous eigenvalues that can take values from minus to plus infinity. So if we start with some state $|s\rangle$ the result of a P measurement will be p with probability $|\langle p|s\rangle|^2$ (postulate 4) if $|p\rangle$ is the eigenstate corresponding to the eigenvalue p . As a result of the measurement the system will collapse into the state $|p\rangle$. As an operator commutes with itself i.e. $[P,P] = 0$, it is easy to see that (equation 1.14)

$$[P, H] = 0. \quad (4.21)$$

Hence, from theorem 4.2, P is a conserved quantity and subsequent measurement of momentum on this system will continue to give the same value p as long as the system is not disturbed in any other way. The state of the system stays $|p\rangle$.

4.2.2 Energy

If $|E\rangle$ is an energy eigestate with eigenvalue E then the probability of measuring E in a state $|s\rangle$ would be $|\langle E|s\rangle|^2$. As we are considering only conservative systems, energy is of course conserved and hence every subsequent measurement of energy will produce the same value E as long as the system is not otherwise disturbed. Now it can be seen that $|p\rangle$ is also an eigenstate of H (see problem 3).

$$H|p\rangle = \frac{P^2}{2m}|p\rangle = \frac{p^2}{2m}|p\rangle = E|p\rangle. \quad (4.22)$$

Hence, the set of states $|p\rangle$ are the same as the set of states $|E\rangle$. However, we choose to label these simultaneous eigenstates with p and not E . This is because, in E , they are degenerate. Two states with opposite momenta have the same value for $E(= p^2/2m)$.

In chapter 3 we saw that the position representation of $|p\rangle$ i.e. its wavefunction (for fixed time) is

$$\Psi_p(x) = \langle x|p\rangle = A \exp(ixp/\hbar). \quad (4.23)$$

As this is also an eigenstate of energy, from equation 4.4 we see that the time dependence of this wavefunction is given by

$$\Psi_p(x, t) = A \exp[i(xp - Et)/\hbar]. \quad (4.24)$$

This function is seen to be a wave with wavelength $2\pi\hbar/p$ and angular frequency E/\hbar . Historically, it was this wave form of the position representation of the energy eigenstates of a free particle that inspired the name wavefunction. In early interference type experiments this relationship between wave properties (wavelength and frequency) and particle properties (momentum and energy) was discovered.

Now one can see why the position representation has been historically preferred. Experiments like electron diffraction basically make position measurements on some given state. By making such measurements on several electrons (each a different system) in the same state the probability distribution of position measurements is obtained. And this probability distribution is directly related to the position representation of a state as given by equation 3.4.

4.2.3 Position

The position operator X is not a conserved quantity as it is seen not to commute with the hamiltonian. Using equation 3.7 and the properties of commutator brackets, one obtains

$$[X, H] = [X, \frac{P^2}{2m}] = i\hbar \frac{P}{m}. \quad (4.25)$$

Hence, position measurements are meaningful only in certain types of experiments. A position measurement on a system can predict very little about subsequent position measurements on the same system even if it is not disturbed in any other way. However, in particle scattering type experiments, a position measurement is made only once on each particle. In such experiments position is measured for different particles each in the same state to obtain information on the probability distribution of position measurements³. This is the kind of situation we will be interested in.

If a particle is in a state $|s\rangle$, its position probability distribution is $|\langle x|s\rangle|^2$. In particular if $|s\rangle$ is a momentum (or energy) eigenstate, this distribution is $|\langle x|p\rangle|^2$ which, from equation 3.20, is seen to be independent of x . Hence, a particle with its momentum known

³For this, each particle needs to behave like a separate isolated system which means the density of particles must be low enough to ignore interactions amongst them.

exactly, is equally likely to be anywhere in space! This is a special case of the celebrated Heisenberg uncertainty principle (see appendix B).

To understand the unpredictable nature of nonconserved quantities, it is instructive to further analyze this specific example of the position operator for a free particle. So we shall see what happens if repeated position measurements are made on the same free particle. Whatever the initial state, the first measurement results in a value, say x_1 . This collapses the state to $|x_1\rangle$. Due to nonconservation of position, this state starts changing with time right after the measurement. If the first measurement is made at time $t = 0$, at later times t the state will be $|x_1\rangle_t$ which can be found from the Schrödinger equation:

$$i\hbar \frac{\partial}{\partial t} |x_1\rangle_t = H|x_1\rangle_t. \quad (4.26)$$

To observe the time development of $|x_1\rangle$, as given by the above equation, it is convenient to expand it in energy eigenstates⁴ which in this case are the momentum eigenstates $|p\rangle$. So, for $t = 0$ we write

$$|x_1\rangle = \int \langle p|x_1\rangle |p\rangle dp. \quad (4.27)$$

The time development of this state is given by (problem 2)

$$|x_1\rangle_t = \int \langle p|x_1\rangle |p\rangle_t dp, \quad (4.28)$$

where $|p\rangle_t$ is the time development of the energy eigenstate as given by equation 4.4. Then using equation 3.25 we get

$$|x_1\rangle_t = (2\pi\hbar)^{-1/2} \int \exp(-ipx_1/\hbar) \exp(-iEt/\hbar) |p\rangle dp, \quad (4.29)$$

where $E = p^2/2m$. Hence, at time t the probability of measuring a value x for position is given by $|\langle x|x_1\rangle_t|^2$ where

$$\langle x|x_1\rangle_t = (2\pi\hbar)^{-1/2} \int \exp(-ipx_1/\hbar) \exp(-iEt/\hbar) \langle x|p\rangle dp. \quad (4.30)$$

Using equation 3.20 and $E = p^2/2m$ this gives

$$\langle x|x_1\rangle_t = (2\pi\hbar)^{-1} \int \exp \left[\frac{-i}{\hbar} \left(p(x_1 - x) + \frac{p^2 t}{2m} \right) \right] dp. \quad (4.31)$$

This integral has meaning only in a limiting sense. If

$$u(x_1 - x, t) = \int \exp \left[-ap^2 - \frac{i}{\hbar} \left(p(x_1 - x) + \frac{p^2 t}{2m} \right) \right] dp, \quad (4.32)$$

⁴Expanding in known eigenstates of a conserved quantity is convenient because its time dependence is simple. In particular, the time dependence of the energy eigenstates is already known

with $a > 0$, then

$$\langle x|x_1 \rangle_t = (2\pi\hbar)^{-1} \lim_{a \rightarrow 0} u(x_1 - x, t). \quad (4.33)$$

Computing the integral in equation 4.32 gives

$$u(x_1 - x, t) = \left[-\frac{2\pi im\hbar(t + ib)}{t^2 + b^2} \right]^{1/2} \exp \left[\frac{im(x_1 - x)^2 t}{2\hbar(t^2 + b^2)} \right] \exp \left[-\frac{mb(x_1 - x)^2}{2\hbar(t^2 + b^2)} \right], \quad (4.34)$$

where $b = 2m\hbar a$. Hence, the probability of finding the particle at x after time t is

$$|\langle x|x_1 \rangle_t|^2 = \lim_{b \rightarrow 0} (2\pi\hbar)^{-1} m(t^2 + b^2)^{-1/2} \exp \left[-\frac{mb(x_1 - x)^2}{\hbar(t^2 + b^2)} \right]. \quad (4.35)$$

To understand this physically we first consider a nonzero value for b . In that case we notice that the probability decreases with time if

$$|x_1 - x| < \left[\frac{\hbar(t^2 + b^2)}{2mb} \right]^{1/2}, \quad (4.36)$$

and increases at points farther out from x_1 . We shall call the right hand side in the above inequality the inversion point. With time, the inversion point moves outwards from x_1 . This is sometimes interpreted as probability “flowing” outwards from the initial point x_1 somewhat like in diffusion.

If $b = 0$, equation 4.35 gives unusual results. At $t = 0$ the probability is still zero everywhere other than $x = x_1$. But even an infinitesimal time later, the probability becomes a nonzero value constant over all space and decreases with time as $1/t!$. This happens because with $b = 0$ the initial delta function wavefunction has infinite momentum (and hence, infinite velocity) components in finite amounts. Therefore, parts of the probability can go to infinity instantaneously and then get lost giving a decreasing overall probability. In the light of special relativity infinite velocity is not possible. This issue can be resolved only by introducing a relativistic quantum mechanics as will be done later.

4.3 The harmonic oscillator

In both classical and quantum mechanics a common practical problem is that of the behavior of a system around an equilibrium point. A classical example of this is a bridge. It is in an equilibrium state but for the small oscillations caused by traffic, strong winds or sometimes even an earthquake.

Classically the standard method for analyzing this is to expand the potential energy in a Taylor series about the position of equilibrium in terms of all degrees of freedom [2]. This is not as complex as it may sound because such an expansion need not have a zeroth

order (i.e. a constant) term as the potential is known only upto an arbitrary constant. The first order term also vanishes as it has the derivative of the potential (i.e. the force) at the equilibrium point as a coefficient. This means that the lowest nonzero term for the potential is the second order term which is quadratic in displacement coordinates. Considering only small oscillations, one can now ignore all higher order terms. This approximation simplifies the problem and still gives useful solutions in many situations.

A similar problem at the atomic level is that of atoms in a molecule that have equilibrium distances between each other. These atoms also can oscillate about their equilibrium position. Once again, a small oscillations approximation leads to a quadratic form for the potential. However, in the atomic case classical mechanics will be inadequate and quantum analysis is required. The results obtained can be verified experimentally.

As in the classical case, a suitable choice of coordinates can separate the problem into several one dimensional problems each with a potential energy given by [2]

$$V = \frac{1}{2}kX^2, \quad (4.37)$$

where k is a constant and X the linear displacement from equilibrium. We shall now proceed to solve this one dimensional problem. This system is usually called the harmonic oscillator. Using equation 4.37, we write the hamiltonian to be

$$H = \frac{P^2}{2m} + V = \frac{P^2}{2m} + \frac{1}{2}kX^2, \quad (4.38)$$

where m is a parameter that is determined by the fact that P is the momentum conjugate to X . The determination of m is no different from the classical problem as the hamiltonian, in both classical and quantum, is of the same form. For example, for a diatomic molecule $m = m_1m_2/(m_1 + m_2)$ with m_1 and m_2 being the masses of the two atoms. This m is called the reduced mass. As P is the momentum conjugate to X , postulate 2 gives

$$[X, P] = i\hbar. \quad (4.39)$$

Now it is easy to see from theorem 4.2 that neither P nor X is conserved. The only conserved quantity is H . Direct position measurements, like in scattering experiments, are not possible as that would mean directly measuring interatomic distances in molecules. This makes the measurement of P or X experimentally uninteresting. Hence, we shall discuss the measurement of H alone. These measurements are actually made indirectly in molecular spectra. Unlike the free particle, this system has a discrete set of energy eigenvalues. These eigenvalues, being the only possible results of energy measurements, need to be found.

So the problem at hand is to find all possible values of E such that

$$H|E\rangle = E|E\rangle, \quad (4.40)$$

where H is given by equation 4.38 and $|E\rangle$ is the corresponding eigenstate. There happens to be no simple recipe for solving such a problem directly. So we shall draw from vector algebra where, very often, problem solving is more straightforward in a suitably chosen coordinate system. The analog of a coordinate system, in this case, is a representation of state vectors. The two representations discussed in chapter 3 are both suitable for our purposes. We shall choose the position representation in the following only because of historical popularity (see problem 4).

4.3.1 Solution in position representation

The position representation reduces equation 4.40 to a differential equation that can be solved by standard methods. Using the results of chapter 3, equations 4.38 and 4.40 would give

$$\left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} kx^2\right) u_E(x) = E u_E(x), \quad (4.41)$$

where $u_E(x) = \langle x|E\rangle$ will be called the energy eigenfunction. As $u_E(x)$ is related to the probability of finding the particle at position x , it is reasonable to believe that it cannot be infinity at infinite x that is:

$$\lim_{|x| \rightarrow \infty} u_E(x) < \infty. \quad (4.42)$$

This can also be seen from the result of problem 5 in chapter 3 which says that if the above condition is not satisfied, the momentum representation is meaningless and hence, $|E\rangle$ would not be a state belonging to the set \mathcal{V} . Besides, the position representation of the momentum operator, used in equation 4.41, was derived in chapter 3 under the same condition. So, this boundary condition given by equation 4.42 will have to be imposed on the solutions of equation 4.41.

It is convenient to write equation 4.41 in terms of a dimensionless independent variable

$$y = \alpha x, \quad (4.43)$$

where a choice of

$$\alpha^4 = mk/\hbar^2, \quad (4.44)$$

is seen to simplify equation 4.41 to the following.

$$\frac{d^2 u_E}{dy^2} + (e - y^2) u_E = 0, \quad (4.45)$$

where

$$e = \frac{2\alpha^2 E}{k} = \frac{2E}{\hbar} \left(\frac{m}{k}\right)^{1/2} = \frac{2E}{\hbar\omega}, \quad (4.46)$$

and ω is the classical angular frequency of natural oscillation.

To solve equation 4.45 we first obtain a solution at large $|y|$ to make it easier to impose the boundary condition of equation 4.42. For large $|y|$, $e \ll y^2$ and hence, equation 4.45 would give

$$\frac{d^2v}{dy^2} \approx y^2v, \quad (4.47)$$

where v is the large $|y|$ limit of u_E . For large $|y|$, an approximate solution of v is seen to be

$$v \approx \exp(\pm y^2/2). \quad (4.48)$$

The positive exponent is not possible due to the boundary condition in equation 4.42. Hence, using the negative exponent solution in equation 4.48, we could write a solution of u_E to be

$$u_E = H_E(y) \exp(-y^2/2). \quad (4.49)$$

Replacing this in equation 4.45 would give the equation for $H_E(y)$ to be

$$\frac{d^2H_E}{dy^2} - 2y\frac{dH_E}{dy} + (e-1)H_E = 0. \quad (4.50)$$

The solution of equation 4.50 can be obtained by the standard series method i.e. H_E is written as a power series in y .

$$H_E(y) = y^s \sum_{i=0}^{\infty} a_i y^i, \quad \text{with } a_0 \neq 0. \quad (4.51)$$

Replacing this in equation 4.50 would give the following conditions on the coefficients a_i .

$$s(s-1)a_0 = 0, \quad (4.52)$$

$$s(s+1)a_1 = 0, \quad (4.53)$$

$$(i+s+2)(i+s+1)a_{i+2} - (2i+2s+1-e)a_i = 0, \quad i = 0, 1, 2, \dots \quad (4.54)$$

As $a_0 \neq 0$, from the first equation above, we see that

$$s = 0 \text{ or } 1. \quad (4.55)$$

We must now analyze the behavior of H_E for large $|y|$ to make sure the boundary condition of equation 4.42 is satisfied. The large $|y|$ behavior, obviously, is governed by the large powers of y in the series. So, from the above we notice that for large values of i ,

$$a_{i+2} \approx 2a_i/i. \quad (4.56)$$

A Taylor series expansion of the function $\exp(y^2)$ shows that its coefficients also satisfy the above relation for large powers. Hence, for large $|y|$, H_E behaves like $\exp(y^2)$. But, from equation 4.49, we see this to violate the boundary condition of equation 4.42. However, if the series were to terminate beyond a certain power of y , equation 4.56 would be irrelevant

and the boundary condition would be satisfied due to the exponential in equation 4.49. From equation 4.54 we notice that if

$$e = 2j + 2s + 1, \quad (4.57)$$

for some integer j , then a_{j+2} vanishes and all subsequent coefficients related to it (like a_{j+4} , a_{j+6} etc.) also vanish. This implies that if j is odd, all a 's with odd indices higher than j vanish and if j is even, all a 's with even indices higher than j vanish. Hence, if j is odd the series will still have large powers from the even indexed terms that will not vanish as they are all related to a_0 (from equation 4.54) and a_0 is chosen to be nonzero in equation 4.51. So, for the complete termination of the series, j must be even to ensure the termination of the even indexed terms and a_1 must be chosen to be zero to ensure the termination of the odd indexed terms. This shows that H_E must be an even polynomial if $s = 0$ and an odd polynomial if $s = 1$. Now, from equations 4.46 and 4.57, we find that H_E is a suitable solution only if

$$E = \hbar\omega(j + s + 1/2). \quad (4.58)$$

If $n = j + s$, then n is odd if $s = 1$ and it is even if $s = 0$. This is because j is always even. Hence, H_E is an odd or even polynomial depending on whether n is odd or even. The corresponding energy eigenvalues are

$$E = (n + 1/2)\hbar\omega, \quad n = 0, 1, 2, \dots \quad (4.59)$$

This is seen to be a discrete set. It is this kind of discontinuity of possible observed values that attracted attention in early investigations and inspired the name “quantum”. It is to be noted that zero energy is not possible. The lowest, or the so called ground state, energy is $\hbar\omega/2$. However, this energy is in no way measurable. In actual spectroscopic measurements only the differences in energy of pairs of states are measured when the system jumps from a higher to a lower energy state releasing a photon (particle or quantum of electromagnetic radiation) carrying the energy difference.

The solutions for H_E are very often labelled by the integer n of equation 4.59 rather than by E i.e. $H_E \equiv H_n$. The H_n are called the Hermite polynomials in mathematics. Properties of the H_n are to be found in standard texts [3]. The property which is most useful comes from the orthonormality of eigenstates i.e. if we label the states $|E\rangle$ also by the integer n and call them $|n\rangle$ and correspondingly, u_E is called u_n then

$$\delta_{nl} = \langle n|l\rangle = \int \langle n|x\rangle \langle x|l\rangle dx = \int u_n^* u_l dx$$

or

$$\delta_{nl} = \alpha^{-1} \int H_n^* H_l \exp(-y^2) dy. \quad (4.60)$$

Here we have used equations 4.43 and 4.49. The orthonormality condition is used to find the a_0 coefficient for each H_n .

4.3.2 A representation free solution

The previous section illustrated a “brute force” method of solving a problem. Such a method is very useful when one has no guidelines for approaching a problem. However, sometimes (and only sometimes) long periods of reflection can reveal more elegant solutions. This happens to be true for the harmonic oscillator case. In the following, I shall present this elegant, representation free, solution.

For convenience, let us first define the following dimensionless quantities proportional to position and momentum respectively.

$$Q = (m\omega/\hbar)^{1/2}X \quad (4.61)$$

$$K = (m\hbar\omega)^{-1/2}P \quad (4.62)$$

From equation 4.39 one sees that

$$[Q, K] = i, \quad (4.63)$$

and from equation 4.38

$$H = \hbar\omega(K^2 + Q^2)/2 = \hbar\omega G, \quad (4.64)$$

where

$$G = (K^2 + Q^2)/2, \quad (4.65)$$

is a dimensionless operator proportional to H . As G is the sum of two squared hermitian operators, one might expect that all its eigenvalues g are positive. This of course needs to be proved. The following theorem will lead to the result.

Theorem 4.4 *For any arbitrary state $|s\rangle$, the following is true.*

$$\langle s|G|s\rangle > 0$$

Proof: Let

$$a = 2^{-1/2}(K - iQ). \quad (4.66)$$

Its hermitian adjoint is

$$a^\dagger = 2^{-1/2}(K + iQ). \quad (4.67)$$

Then, using equations 4.63 and 4.65

$$a^\dagger a = G + i[Q, K]/2 = G - 1/2. \quad (4.68)$$

Now, if $|r\rangle = a|s\rangle$, then from rule 5 of state vectors

$$0 \leq \langle r|r\rangle = \langle s|a^\dagger a|s\rangle = \langle s|G|s\rangle - \langle s|s\rangle/2$$

Hence,

$$\langle s|G|s\rangle > 0$$

This completes the proof.

In the above theorem, if $|s\rangle$ were replaced by one of the eigenstates of G , it would follow that the corresponding eigenvalue

$$g > 0.$$

The eigenstates $|n\rangle$ of the operator

$$N = a^\dagger a, \quad (4.69)$$

are the same as those of G , as G and N differ by a constant number. They are labelled by the eigenvalues n of N . Also, theorem 4.4 can be seen to show that

$$n \geq 0. \quad (4.70)$$

From equation 4.63, the operators a and a^\dagger can be seen to have the following commutator

$$[a, a^\dagger] = 1. \quad (4.71)$$

Now consider the state $a|n\rangle$ where $|n\rangle$ is the energy eigenstate corresponding to the eigenvalue n of N . Using equation 4.71 we get

$$\begin{aligned} Na|n\rangle &= a^\dagger a a|n\rangle = (a a^\dagger - 1)a|n\rangle \\ &= (aN - a)|n\rangle = (an - a)|n\rangle = (n - 1)a|n\rangle. \end{aligned}$$

This shows that $a|n\rangle$ is also an eigenstate of N with eigenvalue $n - 1$. Consequently, a is called the lowering operator as it lowers an eigenstate to another with eigenvalue less by one than that is

$$a|n\rangle = c_n|n - 1\rangle, \quad (4.72)$$

where c_n is a complex number. Similarly, a^\dagger can be seen to be the raising operator.

$$\begin{aligned} Na^\dagger|n\rangle &= a^\dagger a a^\dagger|n\rangle = a^\dagger(a^\dagger a + 1)|n\rangle \\ &= a^\dagger(n + 1)|n\rangle = (n + 1)a^\dagger|n\rangle. \end{aligned}$$

Hence,

$$a^\dagger|n\rangle = d_n|n + 1\rangle, \quad (4.73)$$

where d_n is a complex number. Now, due to equations 4.70 and 4.72, if any eigenstate is repeatedly operated on by the operator a , then after a finite number of steps one obtains the state with the lowest n (say n_0). n_0 must be less than 1 as otherwise it would be possible to lower it further. Hence, from equation 4.70

$$0 \leq n_0 < 1. \quad (4.74)$$

As $|n_0\rangle$ is the lowest eigenstate, further lowering by a should lead to zero.

$$a|n_0\rangle = 0. \quad (4.75)$$

Hence, using the definition of eigenstates,

$$n_0|n_0\rangle = N|n_0\rangle = a^\dagger a|n_0\rangle = 0. \quad (4.76)$$

Consequently, $n_0 = 0$ and as all higher values of n differ by positive integer values, they must all be positive integers. This gives the complete set of eigenstates to be $|n\rangle$ ($n = 0, 1, 2, \dots$) (see problem 6). Using equations 4.68 and 4.69, one can find the eigenvalues of G .

$$G|n\rangle = (N + 1/2)|n\rangle = (n + 1/2)|n\rangle, \quad (4.77)$$

and hence, from equation 4.64, the energy eigenvalues are found to be

$$E = (n + 1/2)\hbar\omega, \quad (4.78)$$

which is the same result as found by the earlier method.

It is to be noticed that here the eigenstates are not given in any functional form as they have no direct observational consequence. However, for future use the constants c_n and d_n need to be found. Finding the norm of both sides in equation 4.72 we get

$$|c_n|^2 \langle n-1|n-1\rangle = \langle n|a^\dagger a|n\rangle = \langle n|N|n\rangle = n \langle n|n\rangle. \quad (4.79)$$

As all eigenstates are normalized, this would mean $c_n = \sqrt{n}$ and

$$a|n\rangle = \sqrt{n}|n-1\rangle. \quad (4.80)$$

Similarly, from equations 4.73 and 4.71

$$\begin{aligned} |d_n|^2 \langle n+1|n+1\rangle &= \langle n|aa^\dagger|n\rangle = \langle n|(a^\dagger a + 1)|n\rangle \\ &= \langle n|(N+1)|n\rangle = (n+1) \langle n|n\rangle. \end{aligned} \quad (4.81)$$

This gives $d_n = \sqrt{n+1}$ and

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \quad (4.82)$$

It is also possible to relate the states $|n\rangle$ to their position representations by using equation 4.80. As $u_n = \langle x|n\rangle$,

$$\langle x|a|n\rangle = \sqrt{n} \langle x|n-1\rangle = \sqrt{n} u_{n-1}. \quad (4.83)$$

Then, using equations 4.61, 4.62 and 4.66

$$\langle x|(2m\omega\hbar)^{-1/2}[P - im\omega X]|n\rangle = \sqrt{n} u_{n-1}. \quad (4.84)$$

Using the position representations of X and P , this gives

$$\left(\frac{\hbar}{2m\omega}\right)^{1/2} \frac{du_n}{dx} + \left(\frac{m\omega}{2\hbar}\right)^{1/2} x u_n = i\sqrt{n} u_{n-1}. \quad (4.85)$$

This equation can be used to find all the u_n as we know that $u_{-1} = 0$. Here the boundary condition of equation 4.42 is seen to be automatically satisfied.

4.4 Landau levels

An electron in a uniform magnetic field is another simple system which is of considerable practical interest. In particular, in some solid state materials the electron is restricted to a surface and the magnetic field is applied perpendicular to it. In such a system the quantum nature is reflected in Hall effect measurements at low temperatures and high magnetic fields. The electron in this two dimensional system has discrete energy eigenvalues just like those of the harmonic oscillator. These energies are known as the Landau levels.

To find the Landau levels we first write down the hamiltonian of the system in SI units.

$$H = \frac{1}{2m}(\mathbf{P} + e\mathbf{A})^2, \quad (4.86)$$

where \mathbf{P} is the two dimensional momentum, m the electron mass, e the magnitude of the electron charge and \mathbf{A} the vector potential due to the uniform magnetic field. We shall choose the x - y plane to be the surface to which the electron is restricted. So the uniform magnetic field will be in the z direction. In a simple gauge choice the vector potential for this field would be given by

$$A_x = 0, \quad A_y = BX, \quad A_z = 0. \quad (4.87)$$

where B is the magnitude of the magnetic induction and X is the x component of the position operator. Hence, equation 4.86 gives

$$H = \frac{1}{2m}[P_x^2 + (P_y + eBX)^2]. \quad (4.88)$$

At this stage we take a second look at the representation free solution of the harmonic oscillator problem. It can be seen that the problem is completely specified by the hamiltonian in equation 4.64 and the commutator in equation 4.63. If the operators K and Q were replaced by any two operators with the same commutator, the solutions would remain the same. From inspection it can be seen that the present problem can be transformed to look exactly like equations 4.63 and 4.64 and hence, the solutions would be the same. The necessary transformation to dimensionless variables is

$$Q_x = (eB\hbar)^{-1/2}P_y + (eB/\hbar)^{1/2}X, \quad (4.89)$$

$$K_x = (eB\hbar)^{-1/2}P_x, \quad (4.90)$$

$$Q_y = (eB\hbar)^{-1/2}P_x + (eB/\hbar)^{1/2}Y, \quad (4.91)$$

$$K_y = (eB\hbar)^{-1/2}P_y, \quad (4.92)$$

where Y is the y component of the position operator. So

$$[Q_x, K_x] = [Q_y, K_y] = i, \quad (4.93)$$

and all other possible commutators vanish. The hamiltonian would then be

$$H = \frac{eB\hbar}{2m} [K_x^2 + Q_x^2]. \quad (4.94)$$

By putting

$$K = K_x, \quad Q = Q_x, \quad \omega = eB/m \quad (4.95)$$

in equations 4.63 and 4.64 we see them to be the same as given in equations 4.93 and 4.94. Hence, the corresponding energy eigenvalues (i.e. the Landau levels) must be

$$E = (n + 1/2)\hbar eB/m, \quad n = 0, 1, 2, \dots \quad (4.96)$$

Here eB/m is seen to be the classical cyclotron frequency. These energy levels have been indirectly observed in quantum Hall effect measurements [4].

Problems

1. Derive equation 4.4 from equation 4.3.
2. Show that the time development operator is

$$U_t(t) \equiv \exp(-iHt/\hbar)$$

i.e. a state $|s\rangle$ at zero time develops to $|s\rangle_t$ at time t where

$$|s\rangle_t = U_t(t)|s\rangle.$$

Hint: An analytical function of an operator is defined by its Taylor series i.e.

$$f(H) = \sum_{n=0}^{\infty} \frac{H^n}{n!} \left. \frac{d^n f(x)}{dx^n} \right|_{x=0}$$

3. Prove that two observables Q and P can have simultaneous eigenstates if and only if $[Q, P] = 0$. [Hint: Part of the proof can be found in the proof of theorem 4.2.]
4. Find the energy eigenstates and eigenvalues for the harmonic oscillator in the momentum representation.
5. Find the Hermite polynomials, H_n , for $n = 0, 1, 2$, and 3. For the same values of n , show by direct integration that the u_n are mutually orthogonal. Find a_0 in each of the four cases by normalizing according to equation 4.60.
6. Show that noninteger eigenvalues are not possible for the number operator N . [Hint: Assume such an eigenstate to exist and see what happens on repeatedly lowering it by using a .]

7. For $n = 0, 1, 2$, and 3 , find u_n using equation 4.85.
8. For the Landau level problem find the raising and lowering operators in terms of momentum and position operators. Then find the equivalent of equation 4.85 in this case and solve for wavefunctions of the four lowest levels.

Chapter 5

More One Dimensional Examples

The examples of quantum systems presented in chapter 4 gave some hint as to what to expect in quantum mechanics. The next step will be to understand general properties of solutions of some physically interesting classes of problems. Then we shall go through some oversimplified examples to illustrate these general properties. In this chapter we shall work with only a single particle in one space dimension. Extension to three dimensions will be later seen to be quite straightforward (chapter 8). However, extension to multiparticle systems, in general, is more tricky and will not be discussed in this text.

5.1 General characteristics of solutions

In chapter 4 we saw that energy as an observable has some special importance. In particular, being conserved in closed systems, it can be used to describe the states of such systems. Hence, we would want to identify the energy eigenstates and the corresponding eigenvalues of any conservative system. The relevant equation would be

$$H|E\rangle = E|E\rangle, \quad (5.1)$$

where H is the hamiltonian, $|E\rangle$ an eigenstate of H and E the corresponding eigenvalue. Equation 5.1 is often referred to as the time independent Schrödinger equation due to its similarity to equation 2.1 (the time dependent Schrödinger equation). The total energy operator i.e. the hamiltonian of a single particle is derived from its classical form to be

$$H = \frac{P^2}{2m} + V(X), \quad (5.2)$$

where P and X are the momentum and position operators respectively and m is the mass of the particle. V , the potential energy, is assumed to be a function of position alone. However, a more general form of V is not difficult to handle.

It is to be noted that the kinetic energy term ($P^2/2m$) in H is the same for all single particle one dimensional problems and hence, it is V , the potential energy, that characterizes a specific problem. In most practical problems, V is found to be a function of position. This makes the position representation a natural choice¹. Hence, we write equation 5.1 in the position representation as follows.

$$Hu_E(x) = Eu_E(x), \quad (5.3)$$

where $u_E(x) = \langle x|E\rangle$ is the position representation of $|E\rangle$ and the position representation of H can be seen to be

$$H = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \quad (5.4)$$

As in equations 5.3 and 5.4, in the future we shall use the same symbol for both an operator and its representation. This does not cause any confusion as long as the representations of the state vectors are suitably labelled (as in equation 5.3).

Equation 5.3 is seen to be a differential equation and should not, in general, be difficult to solve if $V(x)$ is given to be some physically plausible potential energy function. But before solving specific problems we shall study the general characteristics of solutions for certain classes of $V(x)$.

5.1.1 $E < V(x)$ for all x

In classical physics we know that $E < V$ has no meaning as it would give a negative kinetic energy and hence, an imaginary velocity. However, in quantum mechanics we have already seen nonzero values of the wavefunction $u_E(x)$ in regions of space where $E < V$ (the harmonic oscillator as discussed in chapter 4). This would result in nonzero probabilities of finding the particle in such a region. Such an unexpected result prompts us to be more cautious in dealing with quantum mechanics. Consequently, we shall first consider the extreme case of $E < V$ for all x . Fig. 5.1a shows such a potential energy and fig. 5.1b shows the possible wavefunctions. Consider an arbitrary point x_0 . As the overall sign of $u_E(x)$ can be chosen arbitrarily, we can choose $u_E(x_0)$ to be positive. From equations 5.3 and 5.4 we get

$$\frac{d^2}{dx^2}u_E(x) = \frac{2m}{\hbar^2}[V(x) - E]u_E(x). \quad (5.5)$$

As the second derivative is related to the curvature, we notice from equation 5.5 that at x_0 the curvature is positive. Consequently, it could not decrease in both directions. In a direction that it increases it will stay positive and hence, from equation 5.5, continue to have a positive curvature. This would make it increase to infinity. But such a solution is not allowed (problem 5 of chapter 3). So the conclusion is that $E < V$ for all x is not possible for any system.

¹Later we shall see that the description of internal degrees of freedom of a particle (e.g. spin) needs a different representation.

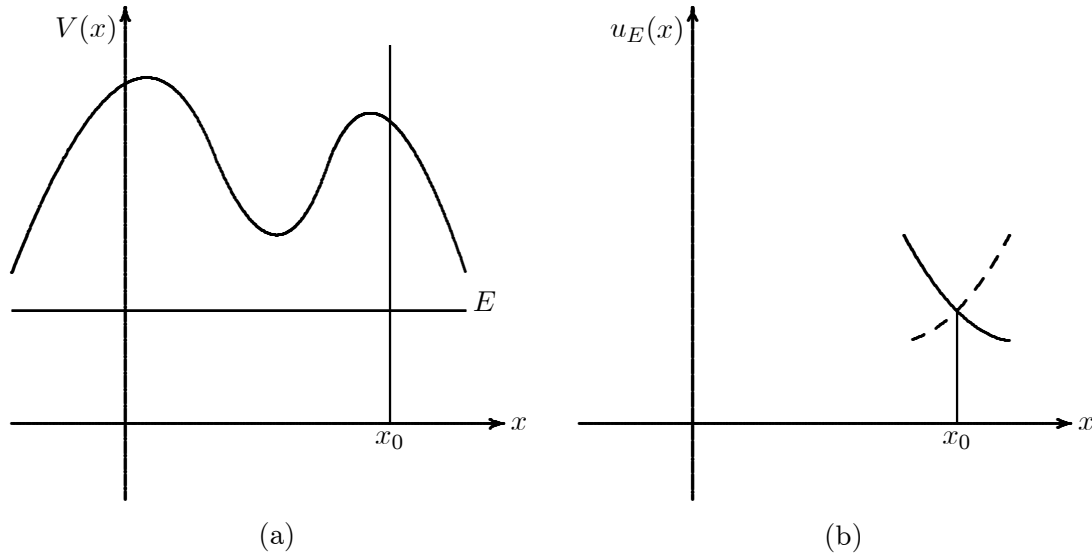


Figure 5.1: Potential energy and wavefunction for $E < V(x)$ case.

5.1.2 Bound states

This is the case where $E > V(x)$ only in some regions of finite x and $E < V(x)$ for $|x| \rightarrow \infty$. States satisfying these conditions are called bound states because the corresponding classical case is that of a particle restricted to a certain region of space. The harmonic oscillator is an example where these conditions are satisfied for all finite values of E . Fig. 5.2a shows a somewhat arbitrary example of such a potential. Given the nature of the potential, one can locate two points, x_1 and x_2 , such that $E < V(x)$ for $x < x_1$ and $x > x_2$.

We shall now consider the behavior of $u_E(x)$ for this potential (fig. 5.2b). As in the last section one can choose the function to be positive at x_2 . Then from equation 5.5 one concludes that for all $x > x_2$ the curvature is positive if the function remains positive. This allows three possibilities for the value of the function at infinity. First, it could curve upwards and go to infinity. Second, it could continue to curve upwards but have a negative slope all the way upto infinity without changing sign. This would require the function to go to zero at infinity. Third, the function could change sign even while it curves upwards. But once it changes sign equation 5.5 would require it to have a negative curvature. This would force it to go to negative infinity. Only the second possibility is allowed due to reasons given earlier. Equation 5.5, being a second order differential equation, allows two initial conditions. One can choose these conditions to be the values of the function and its derivative at x_2 ($u_E(x_2)$ and $u'_E(x_2)$). The choice can be made to ensure that the second of the above three possibilities is true.

Next we consider the behavior of the function to the left of x_2 . In the regions where

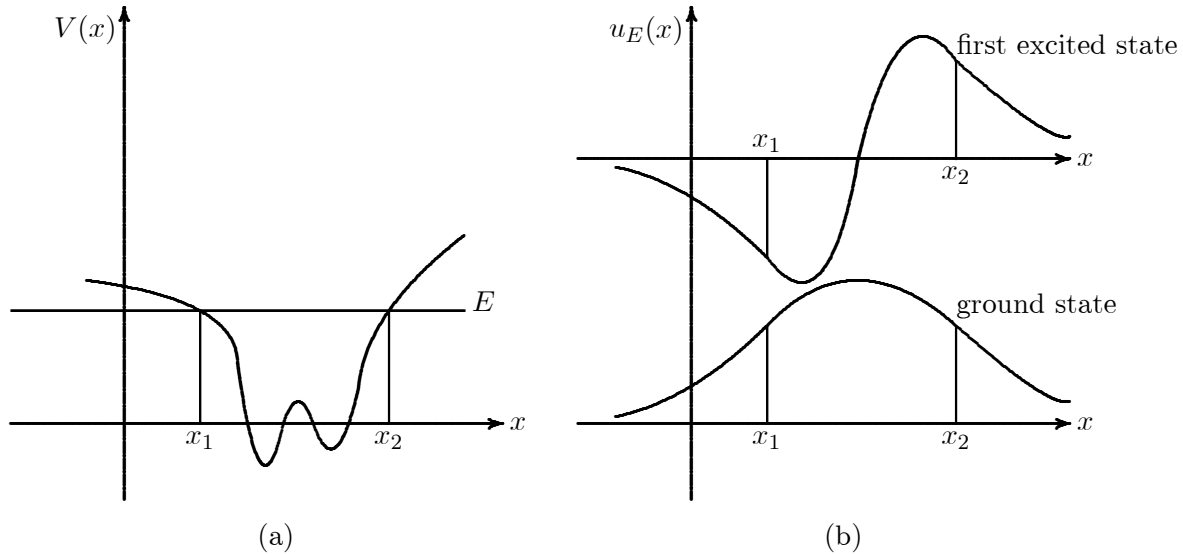


Figure 5.2: Potential energy and wavefunction for bound state case.

$E > V$, using equation 5.5 once again, one finds the curvature to be negative (as long as the function is positive). However, if it is not sufficiently negative, it might still go to infinity at negative infinity as the curvature would be positive again for $x < x_1$. The values of E that will yield such a result are not allowed. If E is increased to some critical value (say E_0) the curvature of the function would be sufficiently negative between x_1 and x_2 to make it go to zero at negative infinity. However, if E is increased beyond E_0 the curvature is so highly negative that it makes the function drop below zero. Once, the function is less than zero equation 5.5 would give it a negative curvature for $x < x_1$ which would make it go to negative infinity at negative infinity. This is not allowed. This shows that E_0 is an allowed energy but energies in its immediate neighborhood are not. E_0 is seen to be the lowest allowed energy and hence, it is called the ground state energy and the corresponding state is called the ground state. If one continues to increase E beyond E_0 , the function will drop below zero between x_1 and x_2 and hence produce a positive curvature between these points. If the curvature is sufficiently positive it can pull up the function from negative infinity to zero for $x \rightarrow -\infty$. The critical energy E_1 at which this happens is the next allowed energy (the first excited state). Energies immediately beyond E_1 will once again be disallowed as they would send the function to positive infinity at $x = -\infty$. A repetition of this argument shows that the allowed energies are a discrete sequence of energies (E_0, E_1, E_2, \dots) . The subscript of the allowed energy can be seen to correspond to the number of times the corresponding wavefunction (eigenfunction) changes sign. This discrete nature of the set of eigenvalues (which are the observed values) is unique to the quantum mechanics of bound states and has no parallel in classical mechanics.

Experimental observation of these eigenvalues is a little indirect. A general approach

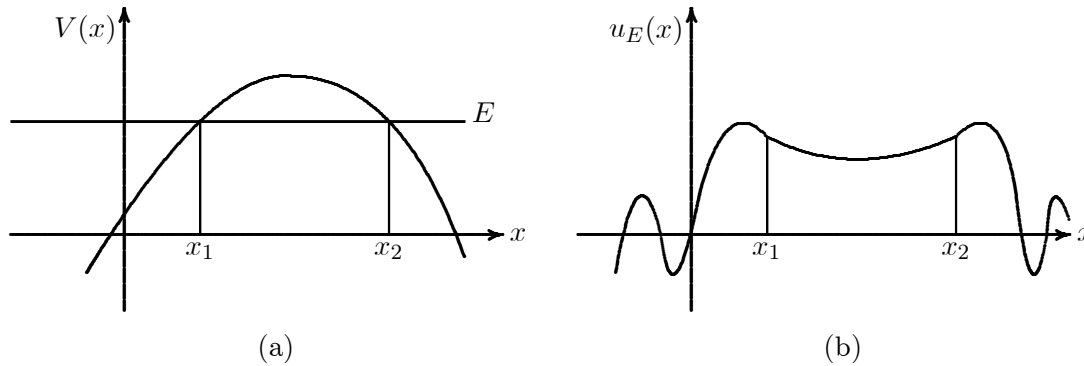


Figure 5.3: Potential energy and wavefunction for scattering state case.

is to give an ensemble of the system under study (e.g. a large collection of identical atoms), a random amount of energy (maybe heat). This would make different members of the ensemble (atoms) rise to different energy levels (eigenvalues). Then, by mechanisms that will be discussed later, each member “jumps” to lower energy levels by emitting the balance of energy in some form (usually electromagnetic particles called photons). Hence, the emitted energy gives the difference between initial and final energies. Different members of the ensemble will have different initial and final energies and would emit different amounts of energy. An analysis of these emitted energies (photons) can then verify the eigenvalues of the system. Specific examples will be discussed later.

5.1.3 Scattering states

This is the case where $E > V(x)$ at $x = -\infty$, or $x = +\infty$ or both. In addition, one may use the fact that for most physical cases E and V are known to be finite at infinity. The corresponding states are called scattering states as classical scattering problems have $E > V$ at large distances from the scatterer. For bound states one noticed that the function had to go to zero at large distances in both directions to prevent it from going to infinity. In the present case we shall see that at least in the direction that has $E > V$ at infinity, no such condition is necessary.

Consider, for example, that $E > V$ for all $x > x_2$ (fig. 5.3a). Then, from equation 5.5, one concludes that for positive values of the function the curvature is negative and vice versa. This forces the function to curve down when it is positive and curve up when it is negative (fig. 5.3b). As a result, the function becomes oscillatory and does not go to infinity for any value of E as long as $E > V$ at infinity. Hence, there are no disallowed states as long as $E > V$ at infinity. If $E > V$ for all $x < x_1$, the solution becomes oscillatory in that region too. It can be seen that there are no disallowed energies even if the solution is oscillatory only at one of the infinities. This is because the function can be held at finite

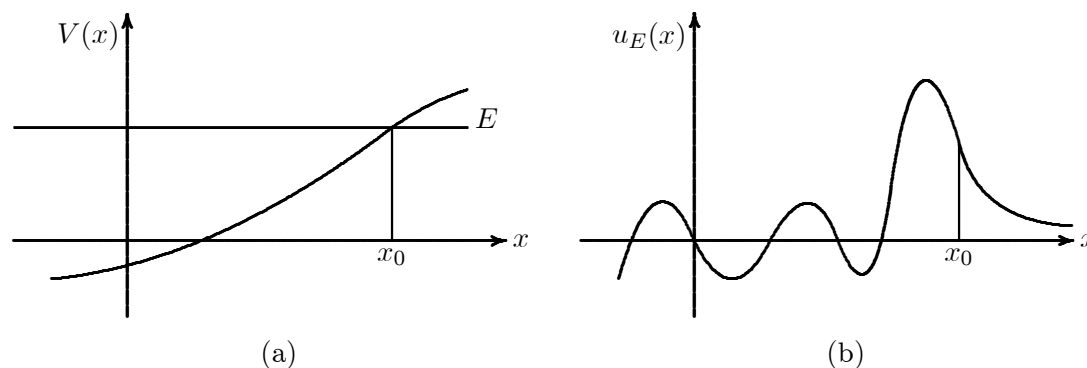


Figure 5.4: Potential energy and wavefunction for $E < V(x)$ at $x = +\infty$ and $E > V(x)$ at $x = -\infty$.

values in one direction simply by an appropriate choice of initial conditions (fig. 5.4).

The scattering states are often said to have a “continuous spectrum” (see definition 17 on page 8) of energies and the bound states are said to have a “discontinuous spectrum” or a “discrete spectrum” (see definition 16 on page 8) of energies. These terms are due to their relation to the electromagnetic emission or absorption spectra of materials (chapter 8). A system can often have both bound and scattering states in different ranges of E (e.g. the hydrogen atom).

For scattering states theoretical prediction of possible energies is trivialized by the continuous nature of the spectrum. However, computations of probabilities of scattering of a particle in different directions is meaningful and nontrivial. For bound states theoretical prediction of possible energies is nontrivial due to the discrete nature of the spectrum but scattering probabilities have no meaning as no particle can escape to infinity (i.e. scatter).

Consequently, for scattering states we need to study the meaning of scattering experiments as pertaining to quantum mechanics. A standard experimental situation is that of a beam of particles impinging on a target and then scattering in all directions (fig. 5.5). Although at present we are discussing only one dimensional problems, for future use, the following general analysis of such scattering processes will be in three dimensions.

It is intuitive to conclude that the information on scattering probabilities in different directions must be contained in the wavefunction as it gives the probability of finding the particle at some given position. However, it must be noted that the wavefunction, as discussed till now, describes a single particle and not a whole beam of particles. Hence, it can provide information about the scattering of the beam only if all particles of the beam are in the same state and the beam is not dense enough to require the consideration of interparticle forces. All forces (or potential energies) are due to the target. Under these conditions we need to find a relation between the single particle wavefunction and

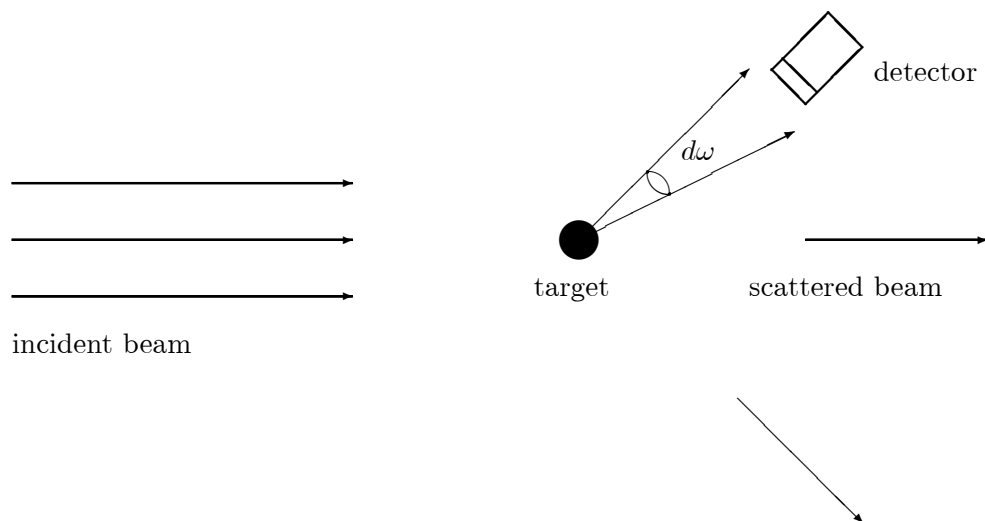


Figure 5.5: A particle scattering experiment.

the measured scattering probabilities for a beam of particles. The measure of scattering probabilities is called the scattering cross section and it is defined as follows.

Definition 27 *The (DIFFERENTIAL) SCATTERING CROSS SECTION is defined as*

$$\sigma = \frac{1}{N_p} \frac{dn_p}{d\omega}, \quad (5.6)$$

where dn_p is the number of particles scattered per unit time into the infinitesimal solid angle $d\omega$ (measured with the target as center) and N_p is the number of incident particles per unit time per unit cross sectional area of the incident beam.

σ would of course be a function of the direction (often conveniently given by the 3-dimensional polar angles (θ, ϕ)) in which the measurement is made. dn_p would have to be measured with some particle detector placed in that direction for a known N_p . The target is usually small enough compared to its distance from the detector to be considered a single point. Let \mathbf{S}_p denote the particle current density at any point in space. Then for the incident beam

$$|\mathbf{S}_p| = N_p, \quad (5.7)$$

and for scattered particles (in 3-dimensions) near the detector

$$dn_p = |\mathbf{S}_p| r^2 d\omega, \quad (5.8)$$

where r is the distance from the target to the detector and hence $r^2 d\omega$ must be the surface area of the particle detector. As particle number is conserved, the particle current density

must obey the following conservation equation (see problem 1).

$$\nabla \cdot \mathbf{S}_p + \frac{\partial \rho_p}{\partial t} = 0, \quad (5.9)$$

where ρ_p is the particle density. We can also show that a similar relation holds for the single particle probability density $\rho = \Psi^* \Psi$ where Ψ is the single particle wavefunction (i.e. the position representation of the state vector).

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial t}(\Psi^* \Psi) = \Psi^* \frac{\partial \Psi}{\partial t} + \Psi \frac{\partial \Psi^*}{\partial t}. \quad (5.10)$$

The position representation of equation 2.1 is

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi + V \Psi, \quad (5.11)$$

where the three dimensional hamiltonian H has been replaced by its position representation

$$H = \frac{P^2}{2m} + V = -\frac{\hbar^2}{2m} \nabla^2 + V. \quad (5.12)$$

The complex conjugate of equation 5.11 is

$$-i\hbar \frac{\partial \Psi^*}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \Psi^* + V \Psi^*. \quad (5.13)$$

Now multiplying equation 5.11 by Ψ^* and equation 5.13 by Ψ and then subtracting the two, one obtains from equation 5.10

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= -\frac{i\hbar}{2m} (\Psi \nabla^2 \Psi^* - \Psi^* \nabla^2 \Psi) \\ &= -\frac{i\hbar}{2m} \nabla \cdot (\Psi \nabla \Psi^* - \Psi^* \nabla \Psi). \end{aligned} \quad (5.14)$$

Hence,

$$\nabla \cdot \mathbf{S} + \frac{\partial \rho}{\partial t} = 0, \quad (5.15)$$

where

$$\mathbf{S} = \frac{i\hbar}{2m} (\Psi \nabla \Psi^* - \Psi^* \nabla \Psi). \quad (5.16)$$

Equation 5.15 can be seen to look exactly like equation 5.9. Moreover, under the present assumption of all particles of the beam being in the same state and there being no interaction among them, it can be seen that the particle density of the beam must be proportional to the single particle probability density i.e.

$$\rho_p = \alpha \rho, \quad (5.17)$$

for some constant α . Hence, from equations 5.9 and 5.15 it can be concluded that

$$\mathbf{S}_p = \alpha \mathbf{S}, \quad (5.18)$$

and from equations 5.7 and 5.8 this would give

$$N_p = \alpha N, \quad dn_p = \alpha dn, \quad (5.19)$$

where

$$N = |\mathbf{S}|, \quad dn = |\mathbf{S}|r^2 d\omega, \quad (5.20)$$

near the source and the detector respectively. Now from equations 5.6 and 5.19, we can rewrite σ in terms of quantities depending only on the single particle wavefunction.

$$\sigma = \frac{1}{N} \frac{dn}{d\omega}. \quad (5.21)$$

For the one dimensional case, there are only two directions of scattering (forward and backward). $dn/d\omega$ in the forward direction would simply be the forward scattered (or transmitted) current n_t and in the backward direction it would be the reflected current n_r . N , the incident current density, will be replaced by the incident current, because in one dimension the cross sectional area of the beam has no meaning. Hence, σ would have the meaning of a transmission and a reflection coefficient (T and R) given by

$$T = n_t/N, \quad R = n_r/N. \quad (5.22)$$

5.2 Some oversimplified examples

Before working out some standard examples, I shall state and prove two useful theorems. Theorem 5.1 will not be directly useful in working out problems, but it will explain why most energy eigenfunctions are seen to be real.

Theorem 5.1 *Any eigenfunction of energy (i.e. a solution $u_E(x)$ of equation 5.3) can be written as a linear combination of real eigenfunctions of energy.*

Proof: Consider an eigenfunction i.e. a solution $f(x)$ of equation 5.3 that is

$$Hf(x) = Ef(x). \quad (5.23)$$

As the differential operator H and the eigenvalue E are both real, the complex conjugate of equation 5.23 would be

$$Hf^*(x) = Ef^*(x). \quad (5.24)$$

Hence, $f^*(x)$ is also an eigenfunction. Consequently, the two real functions

$$u_1(x) = f(x) + f^*(x), \quad u_2(x) = i[f(x) - f^*(x)], \quad (5.25)$$

are also eigenfunctions. Now $f(x)$ can be written as a linear combination of these real eigenfunctions:

$$f(x) = [u_1(x) - iu_2(x)]/2. \quad (5.26)$$

This proves the theorem.

Theorem 5.2 *An energy eigenfunction and its first derivative are continuous if the potential energy V is finite.*

Proof: Let us integrate equation 5.5 as follows (for positive ϵ).

$$\int_{x-\epsilon}^{x+\epsilon} \frac{d^2u(x)}{dx^2} dx = -\frac{2m}{\hbar^2} \int_{x-\epsilon}^{x+\epsilon} [E - V(x)]u(x) dx, \quad (5.27)$$

where the subscript E of the eigenfunction is omitted. E and V are given to be finite and u can be shown to be finite (see problem 2). Hence, the right hand side of equation 5.27 goes to zero in the limit $\epsilon \rightarrow 0$. This gives

$$\lim_{\epsilon \rightarrow 0} \frac{du}{dx} \Big|_{x+\epsilon} - \frac{du}{dx} \Big|_{x-\epsilon} = 0. \quad (5.28)$$

This proves the continuity of the derivative of u . Next consider the double integral of equation 5.5 with the following limits.

$$\int_{x''-\epsilon}^{x''+\epsilon} \left[\int_0^{x'} \frac{d^2u(x)}{dx^2} dx \right] dx' = -\frac{2m}{\hbar^2} \int_{x''-\epsilon}^{x''+\epsilon} \left[\int_0^{x'} [E - V(x)]u(x) dx \right] dx'. \quad (5.29)$$

Performing the x integral on the left side gives

$$\int_{x''-\epsilon}^{x''+\epsilon} \left[\frac{du(x')}{dx'} - K \right] dx' = -\frac{2m}{\hbar^2} \int_{x''-\epsilon}^{x''+\epsilon} \left[\int_0^{x'} [E - V(x)]u(x) dx \right] dx', \quad (5.30)$$

where K is the constant value of the derivative of u at $x = 0$. Once again, as E , V and u are finite, the right hand side of equation 5.30 has a limit of zero when $\epsilon \rightarrow 0$. Similarly, K (on the left hand side) can be seen to be finite and hence its integral goes to zero in the limit of $\epsilon \rightarrow 0$. This leaves us with the following.

$$\lim_{\epsilon \rightarrow 0} [u(x'' + \epsilon) - u(x'' - \epsilon)] = 0. \quad (5.31)$$

Hence, u is found to be continuous. This completes the proof.

Now we are ready to solve some illustrative examples. The potentials used in these examples will be oversimplified to bring out the qualitative aspects with very little mathematical manipulations.

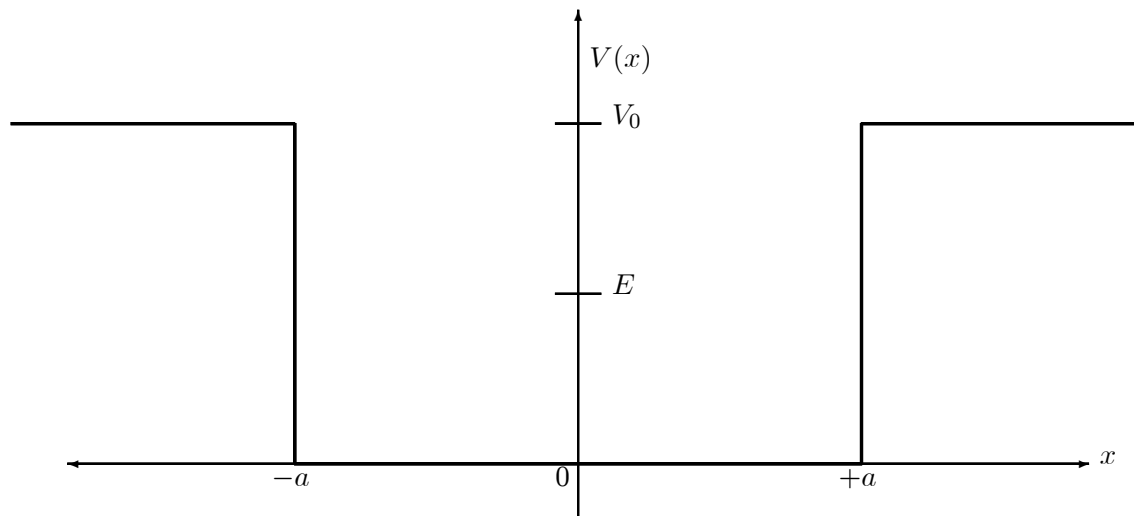


Figure 5.6: Rectangular potential well.

5.2.1 Rectangular potential well (bound states)

The one dimensional rectangular potential well is given as follows (fig. 5.6).

$$V(x) = \begin{cases} 0 & \text{for } |x| < a \\ V_0 & \text{for } |x| > a \end{cases} \quad (5.32)$$

where V_0 is a constant. Here we shall study only the bound states of such a potential (see problem 5). Hence, we need $0 < E < V_0$. In the region $|x| < a$, the time independent Schroedinger equation (equation 5.5) gives

$$\frac{d^2u}{dx^2} = -\frac{2mE}{\hbar^2}u, \quad (5.33)$$

where the subscript E of u is suppressed. A general solution of equation 5.33 is

$$u = A \sin(kx) + B \cos(kx), \quad (5.34)$$

where A and B are arbitrary constants and

$$k = +(2mE/\hbar^2)^{1/2}. \quad (5.35)$$

In the regions where $|x| > a$, equations 5.5 and 5.32 give

$$\frac{d^2u}{dx^2} = \frac{2m}{\hbar^2}(V_0 - E)u. \quad (5.36)$$

A general solution of this equation is

$$u = C \exp(-Kx) + D \exp(Kx), \quad (5.37)$$

where C and D are arbitrary constants and

$$K = +[2m(V_0 - E)/\hbar^2]^{1/2}. \quad (5.38)$$

As u cannot be allowed to go to infinity for $|x| \rightarrow \infty$,

$$u = C \exp(-Kx), \quad \text{for } x > a, \quad (5.39)$$

and

$$u = D \exp(Kx), \quad \text{for } x < -a. \quad (5.40)$$

To determine the four unknown constants A , B , C , and D of equations 5.34, 5.39, and 5.40 we use the continuity conditions of u and du/dx (theorem 5.2) at the two boundaries $x = a$ and $x = -a$. This gives

$$A \sin(ka) + B \cos(ka) = C \exp(-Ka), \quad (5.41)$$

$$Ak \cos(ka) - Bk \sin(ka) = -CK \exp(-Ka), \quad (5.42)$$

$$-A \sin(ka) + B \cos(ka) = D \exp(-Ka), \quad (5.43)$$

$$Ak \cos(ka) + Bk \sin(ka) = DK \exp(-Ka), \quad (5.44)$$

These are four homogeneous linear algebraic equations for the four constants A , B , C , and D . Hence, a nonzero solution will exist only if the following determinant vanishes.

$$\begin{vmatrix} \sin(ka) & \cos(ka) & -\exp(-Ka) & 0 \\ k \cos(ka) & -k \sin(ka) & K \exp(-Ka) & 0 \\ -\sin(ka) & \cos(ka) & 0 & -\exp(-Ka) \\ k \cos(ka) & k \sin(ka) & 0 & -K \exp(-Ka) \end{vmatrix} = 0. \quad (5.45)$$

A direct solution of equation 5.45 is a little tedious. However, a slight manipulation of equations 5.41 through 5.44 can produce the same results more readily i.e.

$$2A \sin(ka) = (C - D) \exp(-Ka), \quad (5.46)$$

$$2Ak \cos(ka) = -(C - D)K \exp(-Ka), \quad (5.47)$$

$$2B \cos(ka) = (C + D) \exp(-Ka), \quad (5.48)$$

$$2Bk \sin(ka) = (C + D)K \exp(-Ka). \quad (5.49)$$

One kind of solution of these four equations is given by

$$A = 0, \quad C = D, \quad B/C = \exp(-Ka)/\cos(ka), \quad (5.50)$$

and

$$k \tan(ka) = K. \quad (5.51)$$

The other kind of solution is given by

$$B = 0, \quad C = -D, \quad A/C = \exp(-Ka)/\sin(ka), \quad (5.52)$$

and

$$k \cot(ka) = -K. \quad (5.53)$$

It is easily seen that the first kind of solution is symmetric in x and the second antisymmetric. The conditions in equations 5.51 and 5.53 can be seen to be solutions of equation 5.45. These conditions are responsible for the discreteness of the set of eigenvalues as discussed in subsection 5.1.2. Solving these equations to find the allowed values of E is not possible in a closed form. However, numerical solutions can be readily obtained. To this end it is convenient to write the equations in terms of the dimensionless parameter $\xi = ka$. Then equations 5.51 and 5.53 would become (using equations 5.35 and 5.38)

$$\xi \tan \xi = (\gamma^2 - \xi^2)^{1/2}, \quad (5.54)$$

and

$$\xi \cot \xi = -(\gamma^2 - \xi^2)^{1/2}, \quad (5.55)$$

where

$$\gamma = (2mV_0a^2/\hbar^2)^{1/2}. \quad (5.56)$$

Once a solution for ξ is found, equation 5.35 will give

$$E = \frac{\hbar^2 \xi^2}{2ma^2}. \quad (5.57)$$

Consider equation 5.54 first. The $\tan(\xi)$ function being periodic there is a possibility of multiple solutions. By definition ξ is positive. In each interval of ξ given by $(n+1/2)\pi < \xi < (n+1)\pi$ (n a non-negative integer) there are no solutions as $\tan(\xi)$ is negative and the right hand side of equation 5.54 is positive. In each interval of ξ given by $n\pi < \xi < (n+1/2)\pi$ (n a non-negative integer) there can be at most one solution as in these intervals the left hand side increases and the right hand side decreases (fig. 5.7). This fact can be used to numerically approximate the solution by the two point bisection method.

The bisection method involves choosing an interval in which one and only one solution exists. Let $[\xi_b, \xi_t]$ be such an interval and let ξ , the solution, belong to this interval. In the present case such an interval would be $[n\pi, (n+1/2)\pi]$ where $\gamma > n\pi$ and n is a non-negative integer. The midpoint $\xi_1 = (\xi_b + \xi_t)/2$ is chosen as the first trial solution. Using equation 5.54, if one finds that $\xi_1 < \xi$ then ξ_1 is taken to be the new lower bound ξ_b and if $\xi_1 > \xi$ then it is taken to be the new upper bound ξ_t . This process shrinks the size of the interval while ensuring that the solution is still within it. Repeating the process can reduce the size of the interval to that of tolerable error and then ξ_b (or ξ_t) could be accepted as the solution ξ . A listing of a computer program implementing this process is given in appendix A.

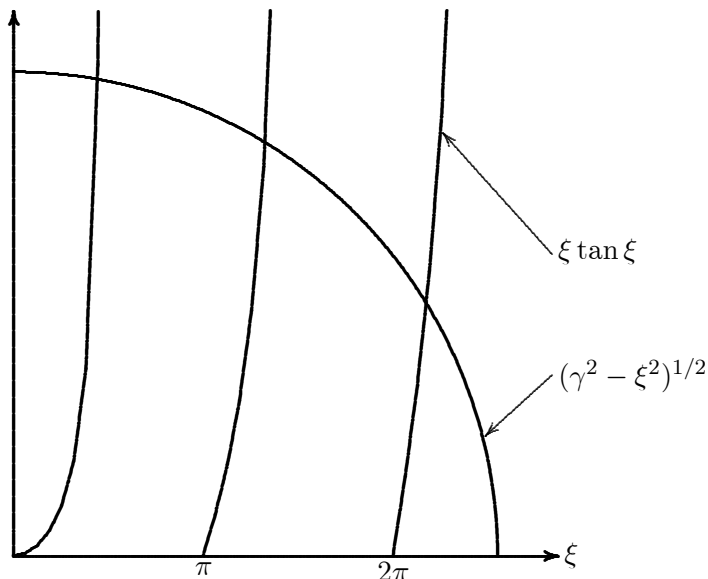


Figure 5.7: Rectangular potential well – graphical solution.

Solutions of equation 5.55 can be obtained in a similar fashion. Once the energy eigenvalues are known, equations 5.50 and 5.52 would give the corresponding eigenfunctions to be in agreement with the qualitative discussions of subsection 5.1.2 (see problem 3).

The solutions of equations 5.54 and 5.55 become much simpler in the limit of $V_0 \rightarrow 0$. The corresponding eigenfunctions are also simpler (see problem 4).

5.2.2 Rectangular potential barrier (scattering states)

The one dimensional rectangular potential barrier is given by the following potential (fig. 5.8).

$$V(x) = \begin{cases} V_0 & \text{for } 0 < x < a \\ 0 & \text{otherwise} \end{cases} \quad (5.58)$$

From the discussion in subsection 5.1.2 we conclude that this potential does not allow any bound states. Hence, all possible states are scattering states. This requires that the form of the incident beam (i.e. its wavefunction) be known from the experimental setup. In most scattering experiments the incident beam has a well-defined momentum i.e. each particle in the beam is in a momentum eigenstate. If the beam is assumed to be incident from the

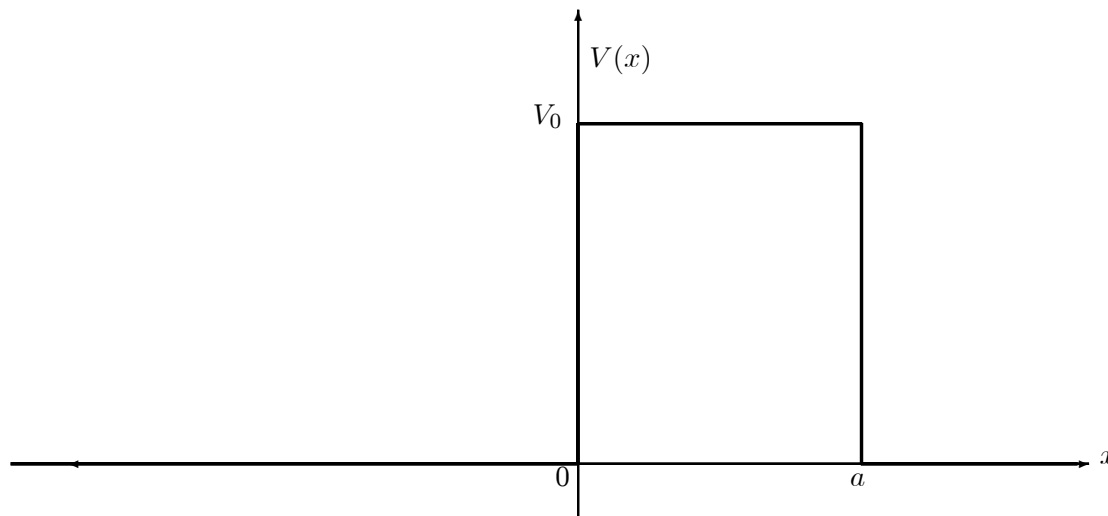


Figure 5.8: Rectangular potential barrier.

left, the corresponding eigenvalue, p , is positive and the energy is

$$E = \frac{p^2}{2m}. \quad (5.59)$$

The position representation of the momentum eigenstate (equation 3.18) would then be

$$u_i = A \exp(ipx/\hbar). \quad (5.60)$$

However, the hamiltonian of the system can be seen not to commute with the momentum operator and hence this initial momentum eigenstate will change. The state that it will change into must still be an energy eigenstate of eigenvalue E (equation 5.59) as energy is conserved. Hence, the position representation, u , of this solution must still satisfy the time independent Schrödinger equation

$$Hu = Eu$$

or

$$-\frac{\hbar^2}{2m} \frac{d^2u}{dx^2} + Vu = Eu. \quad (5.61)$$

In the region where $x < 0$, this would become

$$\frac{d^2u}{dx^2} = -\frac{2mE}{\hbar^2}u. \quad (5.62)$$

A general solution of equation 5.62 is

$$u = A \exp(ikx) + B \exp(-ikx), \quad (5.63)$$

where A and B are arbitrary constants yet to be determined and

$$k = (2mE/\hbar^2)^{1/2} = p/\hbar. \quad (5.64)$$

Hence, we see that the first term in the solution is exactly the incident beam given by equation 5.60. The additional term is by itself also a momentum eigenfunction, but with a momentum equal and opposite to that of the incident beam. Hence, it must be interpreted as the beam reflected by the potential barrier.

$$u_r = B \exp(-ikx) = B \exp(-ipx/\hbar). \quad (5.65)$$

In the region where $x > a$, the potential is again zero and hence the Schrödinger equation still has the form given by equation 5.62 and consequently its solution is

$$u = C \exp(ikx) + D \exp(-ikx). \quad (5.66)$$

The first term of this solution is a momentum eigenstate with positive momentum and hence must be the transmitted beam. The second term is a momentum eigenstate with negative momentum. But that would mean a beam coming in from the far right. As there is no such beam, we must set $D=0$. This leaves us with only the transmitted beam on the right side of the barrier i.e.

$$u = u_t = C \exp(ikx). \quad (5.67)$$

The one dimensional probability current would be the x component of \mathbf{S} as given in equation 5.16.

$$S = \frac{i\hbar}{2m} \left(\Psi \frac{d\Psi^*}{dx} - \Psi^* \frac{d\Psi}{dx} \right). \quad (5.68)$$

Hence, from equation 5.60, the magnitude of the incident current is found to be

$$N = \frac{i\hbar}{2m} \left(u_i \frac{du_i^*}{dx} - u_i^* \frac{du_i}{dx} \right) = \frac{\hbar k}{m} |A|^2, \quad (5.69)$$

and similarly from u_r of equation 5.65 and u_t of equation 5.67 we find

$$n_r = \frac{\hbar k}{m} |B|^2, \quad n_t = \frac{\hbar k}{m} |C|^2. \quad (5.70)$$

Now, from the definitions of the reflection and the transmission coefficients given in equation 5.22, we get

$$R = \frac{|B|^2}{|A|^2}, \quad T = \frac{|C|^2}{|A|^2}. \quad (5.71)$$

To find these coefficients we need to find the solution to equation 5.61 in the region where $0 < x < a$. The general solution in this region can be seen to be

$$u = F \exp(-Kx) + G \exp(Kx), \quad (5.72)$$

where F and G are arbitrary constants and

$$K = [2m(V_0 - E)/\hbar^2]^{1/2}. \quad (5.73)$$

K is real if $E < V_0$ and imaginary if $E > V_0$. The boundary conditions of theorem 5.2 applied to the solutions in equations 5.63, 5.67 and 5.72 at the two boundaries $x = 0$ and $x = a$ will give

$$A + B = F + G, \quad (5.74)$$

$$ikA - ikB = -KF + KG, \quad (5.75)$$

$$F \exp(-Ka) + G \exp(Ka) = C \exp(ika), \quad (5.76)$$

$$-KF \exp(-Ka) + KG \exp(Ka) = ikC \exp(ika). \quad (5.77)$$

These are four equations for the five unknown constants. Hence, all five constants cannot be determined. However, as seen from equation 5.71, for the observable quantities all we need are some ratios of the constants. So we divide the equations 5.74 through 5.77 by A . This gives us four equations for the four unknowns B/A , C/A , F/A and G/A . After some straightforward but tedious algebraic manipulations we find the two relevant solutions to be

$$\frac{B}{A} = -\frac{(K^2 + k^2)[1 - \exp(2Ka)]}{(K + ik)^2 - (K - ik)^2 \exp(2Ka)}, \quad (5.78)$$

$$\frac{C}{A} = \frac{4iKk \exp[(K - ik)a]}{(K + ik)^2 - (K - ik)^2 \exp(2Ka)}. \quad (5.79)$$

If $V_0 > E$ then K is real. Then from equation 5.71 and some more tedious algebraic manipulations we get

$$R = \left[1 + \frac{4E(V_0 - E)}{V_0^2 \sinh^2(Ka)} \right]^{-1}, \quad (5.80)$$

$$T = \left[1 + \frac{V_0^2 \sinh^2(Ka)}{4E(V_0 - E)} \right]^{-1}. \quad (5.81)$$

For $V_0 < E$, K is imaginary such that $j = iK$ is real. Then R and T would have the forms

$$R = \left[1 + \frac{4E(E - V_0)}{V_0^2 \sin^2(ja)} \right]^{-1}, \quad (5.82)$$

$$T = \left[1 + \frac{V_0^2 \sin^2(ja)}{4E(E - V_0)} \right]^{-1}. \quad (5.83)$$

For all energies it can be verified that

$$R + T = 1. \quad (5.84)$$

This is due to the conservation of probability.

The above results are in direct contradiction with classical results. Classically, for $E < V_0$, there can be no transmission at all. But here we see that T is nonzero. This effect has been called “quantum mechanical tunnelling” for want of a better name². It has been observed experimentally in a variety of systems and has also been used in the design of electronic devices. Also, classical physics would not allow any reflection if $E > V_0$. Once again the above quantum results show otherwise.

Problems

1. Show that equation 5.9 holds for any system with a fixed number of particles if ρ_p is the particle density and S_p is the particle current density.
2. Show that for a solution, u , of equation 5.5 to be allowed, it must be finite at all points. [Hint: Use results of section 5.1 showing that a solution can go to infinity only in a region where $E < V$ and in that case it is disallowed.]
3. Sketch the energy eigenfunctions for the four lowest eigenvalues of the rectangular potential well problem. Assume V_0 to be large enough to allow four bound states.
4. Find the energy eigenvalues and eigenfunctions for an infinitely deep rectangular potential well i.e. $V_0 \rightarrow \infty$.
5. Find the reflection and transmission coefficients for the scattering states ($E > V_0$) of the rectangular potential well.
6. Find the reflection and transmission coefficients for the step potential given by

$$V(x) = \begin{cases} 0 & \text{for } x < 0 \\ V_0 & \text{for } x > 0 \end{cases}$$

where V_0 is a positive constant.

7. Consider the following one dimensional potential

$$V(x) = \begin{cases} V_1 & \text{for } x < 0 \\ 0 & \text{for } 0 < x < a \\ V_2 & \text{for } x > a \end{cases}$$

where V_1 and V_2 are positive and constant and $V_2 > V_1$. For this potential find

²Here our classical mindset is seen to hinder the understanding of quantum phenomena. Classically, particles must have continuous trajectories. Hence, we understand that if a particle travels from one side of a barrier to the other, it must have “tunnelled” through. In other words, for some period of time, it must have moved through the barrier. But this is energetically impossible! On the other hand, in a quantum theory the concept of a trajectory is nonexistent. So the particle does not actually have to go through the barrier to be on the other side. In fact, the probability current inside the barrier can be seen to be zero.

- (a) the conditions for energy eigenstates to be bound or scattering,
- (b) the bound state energy eigenvalues,
- (c) and the reflection and transmission coefficients for the scattering states.

Chapter 6

Numerical Techniques in One Space Dimension

For our purposes we shall assume that for any well-defined problem in quantum mechanics it is possible to obtain a solution of acceptable accuracy using some computing machine. Of course, for some problems the required computing machine might not be available presently. Hence, the greatest challenge for the numerical analyst is to find methods that can solve practical problems in reasonable amounts of time using present day computers. This has resulted in some rather involved numerical techniques. A complete discussion of these techniques would divert our attention away from quantum mechanics. Hence, we shall not attempt such a task here, hoping to delegate it to trained numerical analysts. However, even for the physicist that is going to delegate the task, it is important to be familiar with the principles. It helps in two ways. First, it allows physicists to solve simple numerical problems on their own. Second, it lets them communicate better with numerical analysts while solving physics problems.

In the following, some simple and intuitive numerical methods will be discussed to build the groundwork for future development. Some sample programs written in the C language are provided to illustrate the methods (see appendix A). The language C is chosen rather than the more popular FORTRAN to provide greater flexibility. Conversions to other languages should be straightforward once the material in this chapter is understood. Although FORTRAN or BASIC could do the job, structured languages like PASCAL or C should be preferred. Some later versions of FORTRAN allow structured programming. But they still require programming discipline to avoid unstructured programs. If BASIC is used, it should be noted that some of the programs may take several hours on a microcomputer.

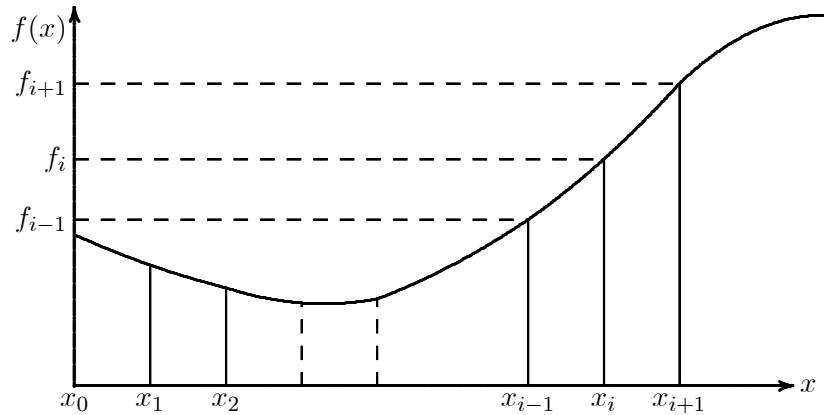


Figure 6.1: Discretization of a function.

6.1 Finite differences

In solving differential equations numerically, it must be noted that the basic concept of limits as defined in calculus, needs to be approximated. Infinitesimal quantities need to be replaced by small but finite quantities. This would, of course, introduce errors. However, such errors can be made indefinitely small by running the computer for a long enough time. A little reflection on the nature of the so called analytical solutions shows that even they cannot provide exact numerical results for real life applications. For example, the value of a trigonometric function of an arbitrary angle can be evaluated only upto a certain accuracy in a finite amount of time!

Approximate forms of derivatives that are needed for numerical computations are called FINITE DIFFERENCES. The resulting equivalents of differential equations are called DIFFERENCE EQUATIONS. To determine the form of finite differences, we shall consider the function $f(x)$ which depends only on x . The values of $f(x)$ are expected to be known or computed only at a discrete set of values of x . For simplicity, this set of values of x will be assumed to be equally spaced and members of the set will be denoted by x_i ($i = 0, 1, 2, \dots$) (see fig. 6.1). The corresponding values of $f(x)$ will be denoted by f_i ($i = 0, 1, 2, \dots$). The interval between two consecutive values of x_i will be given by

$$x_i - x_{i-1} = w. \quad (6.1)$$

The value of w , in principle, can be reduced indefinitely to obtain more accurate results. However, as w is reduced, computation time increases. Also, a reduction in w must be matched with a suitable increase in the number of significant figures to make sure that roundoff errors do not spoil the accuracy.

It can be intuitively seen that the value of the first derivative (the slope) of $f(x)$ at a

point x_i can be approximated by the expression

$$\Delta f_i = \frac{f_{i+1} - f_{i-1}}{x_{i+1} - x_{i-1}} = \frac{f_{i+1} - f_{i-1}}{2w}. \quad (6.2)$$

We shall call this the first difference. To confirm this intuitive result we shall expand f_{i+1} and f_{i-1} each in a Taylor series as follows.

$$f_{i+1} = f_i + wf'_i + \frac{w^2}{2}f''_i + \frac{w^3}{6}f'''_i + \dots, \quad (6.3)$$

$$f_{i-1} = f_i - wf'_i + \frac{w^2}{2}f''_i - \frac{w^3}{6}f'''_i + \dots, \quad (6.4)$$

where f'_i , f''_i and f'''_i denote the first three derivatives of $f(x)$ at x_i . Subtracting f_{i-1} from f_{i+1} and then dividing by $2w$ gives

$$f'_i = \frac{f_{i+1} - f_{i-1}}{2w} + O(w^2) = \Delta f_i + O(w^2) \quad (6.5)$$

where $O(w^2)$ has terms of order two or higher in w . For small enough w , $O(w^2)$ can be ignored.

To obtain an intuitive expression for the second difference (an approximation of the second derivative), we notice that the first difference at the point halfway between x_i and x_{i-1} is

$$f_{i-1/2} = (f_i - f_{i-1})/w, \quad (6.6)$$

and at the point halfway between x_{i+1} and x_i it is

$$f_{i+1/2} = (f_{i+1} - f_i)/w. \quad (6.7)$$

The second difference at x_i should then be

$$\Delta^2 f_i = (f_{i+1/2} - f_{i-1/2})/w = (f_{i+1} - 2f_i + f_{i-1})/w^2. \quad (6.8)$$

To verify that equation 6.8 is an approximate form of the second derivative of $f(x)$, we add the two expressions in equations 6.3 and 6.4. Then solving for f''_i would give

$$f''_i = (f_{i+1} - 2f_i + f_{i-1})/w^2 + O(w^2) = \Delta^2 f_i + O(w^2). \quad (6.9)$$

In the following, only second order differential equations will be discussed. Hence, equations 6.2 and 6.8 will be the only finite differences we will need.

6.2 One dimensional scattering

In chapter 5 we discussed one dimensional scattering from simple barriers (rectangular). Analytical solutions were possible in these cases. In a general case the form of the wavefunction is still the same very far to the left and right of the barrier as all physical forces

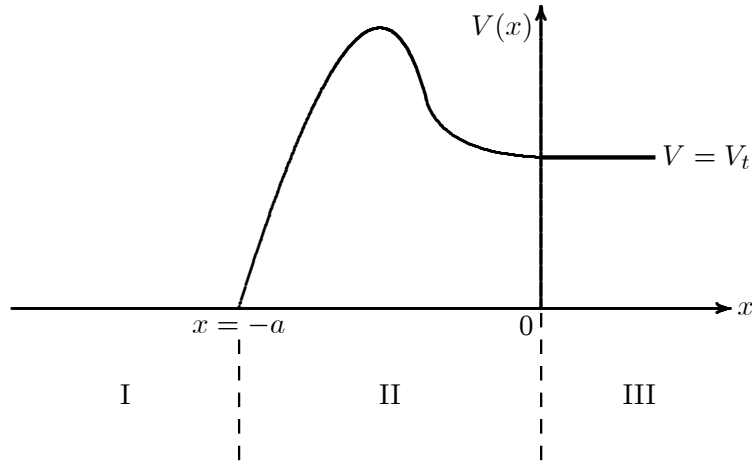


Figure 6.2: General one dimensional scattering potential.

are expected to die out at large distances. Also, in physical problems measurements of scattered and reflected waves are made at large distances where the wave is virtually free of forces. Hence, in the following, one dimensional space will be broken into three regions as follows (fig. 6.2).

- I. Region of incident (and reflected) beam ($V = 0$).
- II. Region of scattering ($V \neq 0$).
- III. Region of transmitted beam ($V = V_t$).

Region I is force free and hence, has a constant potential. With a suitable choice of reference, this potential can be taken to be zero. Region III is also force free. However, as the choice of reference has already been made in keeping the potential in region I as zero, the potential in region III cannot, in general, be zero. But it will still be a constant V_t . The region of nonzero force (region II) has been chosen to be between $x = -a$ and $x = 0$. This choice of coordinate origin is made for future convenience. For numerical solutions of problems the position representation is usually preferred as it leads to differential equations and standard numerical techniques for differential equations are well known. Hence, we shall use the same method as given in chapter 5 with appropriate changes made for the more general potential. The equation to be solved in the three regions is still the time independent Schrödinger equation as given in equation 5.61.

$$-\frac{\hbar^2}{2m} \frac{d^2 u}{dx^2} + Vu = Eu. \quad (6.10)$$

From equation 5.63, we already know the solution in region I to be

$$u = A \exp(ikx) + B \exp(-ikx), \quad \text{for } x < -a. \quad (6.11)$$

In region III $V (= V_t)$ is a constant and hence equation 6.10 can be written as

$$\frac{d^2u}{dx^2} = -k_t^2u, \quad (6.12)$$

where k_t is a constant given by

$$k_t = [2m(E - V_t)/\hbar^2]^{1/2}. \quad (6.13)$$

Hence, the solution of equation 6.12 would be

$$u = C \exp(ik_t x) + D \exp(-ik_t x), \quad \text{for } x > 0. \quad (6.14)$$

If $E < V_t$, then k_t is imaginary. This would make the second term in equation 6.14 go to infinity at $x = +\infty$. As this is not allowed, D must be zero. Such a solution can be seen to lead to a zero transmission probability irrespective of the potential in region II (see problem 6 of chapter 5). Hence, for a numerical solution we need to consider only the case $E > V_t$. So k_t must be real and can be chosen to be positive. In chapter 5 it was shown that the second term in equation 6.14 would be due to a particle moving backwards in region III. Such a particle is not possible as there is no source or reflecting force in region III. Hence, $D = 0$ and

$$u = C \exp(ik_t x), \quad \text{for } x > 0. \quad (6.15)$$

The solution in region II is to be found numerically for arbitrary potentials. Numerical solutions cannot be written in a general form as an arbitrary linear combination of two linearly independent solutions. The freedom of the two arbitrary constants needs to be removed through boundary or initial conditions before solving the equation. These conditions must come from the known solutions in regions I and III and the continuity conditions at the boundaries of the three regions (theorem 5.2). For the computation of the only measurable quantities, (the reflection and transmission coefficients) we have seen that only the ratios of the constants A , B and C are needed. Hence, without loss of generality, one can set any one of these constants to be equal to 1. Depending on which of the constants is chosen to be unity, we need different numerical methods. At present we are going to discuss the method in which $C = 1$. This will be later seen to be the simplest.

As a computing machine, in general, cannot handle complex numbers, one must write equation 6.10 as two separate equations for the real and imaginary parts. Separating u in its real and imaginary parts one writes

$$u = g + ih, \quad (6.16)$$

where g and h are real. From equation 6.10 one now obtains the two real equations for g and h .

$$\frac{d^2g}{dx^2} + \frac{2m(E - V)g}{\hbar^2} = 0, \quad (6.17)$$

$$\frac{d^2h}{dx^2} + \frac{2m(E - V)h}{\hbar^2} = 0. \quad (6.18)$$

For a numerical solution of these equations it is convenient to choose a dimensionless variable for the independent parameter. Let y be such a variable such that

$$x = cy. \quad (6.19)$$

A choice of c that will simplify equations 6.17 and 6.18 would be

$$c = \hbar/(2mE)^{1/2} = 1/k. \quad (6.20)$$

Then equations 6.17 and 6.18 become

$$\frac{d^2g}{dy^2} + (1 - V/E)g = 0, \quad (6.21)$$

$$\frac{d^2h}{dy^2} + (1 - V/E)h = 0. \quad (6.22)$$

To obtain numerical solutions of the above equations, we use the finite difference approximation for the second derivative as given in equation 6.8 with the x_i replaced by the y_i so that $w = y_i - y_{i-1}$. It leads to the following difference equations.

$$g_{i+1} - 2g_i + g_{i-1} + w^2(1 - V_i/E)g_i = 0, \quad (6.23)$$

$$h_{i+1} - 2h_i + h_{i-1} + w^2(1 - V_i/E)h_i = 0, \quad (6.24)$$

where V_i is the value of the potential V at y_i . These equations are recursion relations for g and h . If g_i , h_i , g_{i-1} and h_{i-1} are given, these equations can be solved to find g_{i+1} and h_{i+1} that is

$$g_{i+1} = [w^2(V_i/E - 1) + 2]g_i - g_{i-1}, \quad (6.25)$$

$$h_{i+1} = [w^2(V_i/E - 1) + 2]h_i - h_{i-1}. \quad (6.26)$$

Hence, if g_0 , h_0 , g_1 and h_1 are known, g and h can be found at all points. These initial values can be found from initial conditions of the differential equations. g_0 and h_0 are the initial values of the functions. To determine g_1 and h_1 we notice that approximate forms for the derivatives at the initial point are (see problem 1)

$$(\Delta g)_0 = (g_1 - g_0)/w, \quad (6.27)$$

$$(\Delta h)_0 = (h_1 - h_0)/w. \quad (6.28)$$

Then

$$g_1 = w(\Delta g)_0 + g_0, \quad (6.29)$$

$$h_1 = w(\Delta h)_0 + h_0. \quad (6.30)$$

If the initial values of the functions and their derivatives are known, equations 6.29 and 6.30 would give the approximations for g_1 and h_1 . Hence, in this formulation of the problem,

initial conditions are easier to handle than boundary conditions. A set of boundary conditions would give values of the functions at both ends of region II, but not their derivatives. So values for g_1 and h_1 cannot be found directly.

Initial conditions in region II can be determined from the results of the analytical solutions in either region I or region III if we use the conditions of theorem 5.2. However, the solution in region I has two unknowns (A and B). One of them can be set to unity, but the other still remains unknown. The solution in region III, on the other hand, has only one unknown (viz. C). This can be set equal to 1 as mentioned earlier.

$$C = 1. \quad (6.31)$$

Now equation 6.15 can be written as

$$u = \exp(ik_t x), \quad \text{for } x > 0. \quad (6.32)$$

Then the values of u and its derivative at $x = 0$ (or $y = 0$) would be

$$u_0 = 1, \quad (6.33)$$

$$\left. \frac{du}{dy} \right|_{y=0} = \frac{1}{k} \left. \frac{du}{dx} \right|_{x=0} = \frac{ik_t}{k} = i(1 - V_t/E)^{1/2}. \quad (6.34)$$

Using equation 6.16 this gives

$$g_0 = 1, \quad h_0 = 0, \quad (6.35)$$

$$\left. \frac{dg}{dy} \right|_{y=0} = 0, \quad \left. \frac{dh}{dy} \right|_{y=0} = (1 - V_t/E)^{1/2}. \quad (6.36)$$

As the region of computation (region II) is to the left of this initial point ($y = 0$), we number the indices of y , g and h in increasing order from right to left such that

$$y_0 = 0, \quad y_i = y_{i-1} - w. \quad (6.37)$$

This makes $(\Delta g)_0$ and $(\Delta h)_0$, the approximate forms of the derivatives (equations 6.27 and 6.28), change in sign. Hence, replacing approximate derivatives by the negatives of actual derivatives (equation 6.36), equations 6.29 and 6.30 will give

$$g_1 = 1, \quad h_1 = -w(1 - V_t/E)^{1/2}. \quad (6.38)$$

Now, equations 6.35 and 6.38 will provide all the initial values needed to solve the recursion relations in equations 6.25 and 6.26. If, in region II, there are n intervals of width w each, then the last computed value of g_i and h_i will be g_n and h_n respectively. This will complete the numerical solution of the differential equation.

To determine the reflection and transmission coefficients we need to compute the constants A and B . Once the numerical solution for u is known in region II, A and B

can be found by matching the values of u and its derivative (with respect to y) u_y at the boundary of region I and region II. In doing this we use equations 6.11, 6.19 and 6.20 to get

$$u = A \exp(iy) + B \exp(-iy) \quad (6.39)$$

$$u_y = iA \exp(iy) - iB \exp(-iy). \quad (6.40)$$

This leads to

$$|A|^2 = \frac{|u_y + iu|^2}{4} \quad (6.41)$$

$$|B|^2 = \frac{|u_y - iu|^2}{4}. \quad (6.42)$$

Then using equation 6.16 we get

$$|A|^2 = \frac{(g_y - h)^2 + (h_y + g)^2}{4} \quad (6.43)$$

$$|B|^2 = \frac{(g_y + h)^2 + (h_y - g)^2}{4}, \quad (6.44)$$

where g_y and h_y are the derivatives of g and h with respect to y . Due to theorem 5.2, the values of g , g_y , h and h_y must be the same at $x = -a$ ($y = -ka$) using either the solution in region I or the numerical solution in region II. Hence, the approximate values from region II can be introduced in equations 6.43 and 6.44 to obtain

$$|A|^2 \simeq \frac{(g_{yn} - h_n)^2 + (h_{yn} + g_n)^2}{4} \quad (6.45)$$

$$|B|^2 \simeq \frac{(g_{yn} + h_n)^2 + (h_{yn} - g_n)^2}{4}, \quad (6.46)$$

where g_n and h_n are the last computed values from the recursion relations of equations 6.25 and 6.26 i.e. they are the values of g and h at $x = -a$ ($y = -ka$) which also means $nw = ka$. g_{yn} and h_{yn} are the approximations of the derivatives of g and h at $y = -ka$.

$$g_{yn} = (g_{n-1} - g_n)/w, \quad h_{yn} = (h_{n-1} - h_n)/w. \quad (6.47)$$

Now, the transmission coefficient T and the reflection coefficient R can be found from the definitions in equation 5.22. The currents N and n_r are the same as in equations 5.69 and 5.70. However, as k_t is no longer the same as k , n_t is given by

$$n_t = \frac{\hbar k_t}{m} |C|^2. \quad (6.48)$$

Hence, we obtain

$$T = \frac{k_t |C|^2}{k |A|^2} = \frac{4(1 - V_t/E)^{1/2}}{(g_{yn} - h_n)^2 + (h_{yn} + g_n)^2} \quad (6.49)$$

$$R = \frac{|B|^2}{|A|^2} = \frac{(g_{yn} + h_n)^2 + (h_{yn} - g_n)^2}{(g_{yn} - h_n)^2 + (h_{yn} + g_n)^2}. \quad (6.50)$$

Thus an algorithm for the computation of transmission and reflection coefficients will contain the following parts.

- Determine desired values of E , V_t and w . For better results, the function should not change much in any interval w . Hence, the wavelength of the wavefunction must be large compared to w i.e. $w \ll 1/k$.
- Set the initial values g_0 , h_0 , g_1 and h_1 using equations 6.35 and 6.38.
- Use the recursion relations in equations 6.25 and 6.26 in a loop to compute all values of g_i and h_i upto g_n and h_n . In each step of the loop, y is to be decreased by w and the value of the potential is to be computed at that value of y . The loop is to be terminated when $y \leq -ka$.
- Reflection and transmission coefficients are to be computed using equations 6.47, 6.49 and 6.50.

6.3 One dimensional bound state problems

In bound state problems the energy eigenvalues and eigenfunctions need to be found. A rather simple method is possible for potentials that have a reflection symmetry. As will be seen in chapter 7, wavefunctions for such potentials are either symmetric or antisymmetric. Hence, solving the equation for positive x is sufficient. The antisymmetric wavefunctions vanish at the origin and the symmetric wavefunctions have zero slope at the origin. This initial condition, along with an arbitrary choice of normalization constant, is sufficient for the numerical solution of the equation. As the energy eigenvalue is not known a priori, one needs to solve the equation for a series of energy values and observe the behavior of the function at some large distance (viz. the “tail”) for each case. If the magnitude of the tail increases rapidly with distance, one concludes that the energy is not an energy eigenvalue. In the neighborhood of an energy eigenvalue a slow change in the chosen energy will show a rapid change in the tail. In fact the tail can be seen to change sign precisely at energies that are eigenvalues. This change in sign is so rapid that the solution for the wavefunction is extremely unstable and inaccurate at large distances. However, the rapid change in sign for small changes in energy can be used to locate the energy eigenvalue very precisely. To illustrate this method, we shall use it for the harmonic oscillator problem that has already been solved analytically. Hence, it will be a good test for the method.

As in the scattering case, the numerical solution of bound state problems is also facilitated by choosing an appropriate dimensionless independent variable. For the harmonic oscillator problem, this has already been done in chapter 4. The choice of the independent variable y , is given in equations 4.43 and 4.44. The resulting equation to be solved is

equation 4.45 which we write here without the subscript E of the function u .

$$\frac{d^2u}{dy^2} + (e - y^2)u = 0. \quad (6.51)$$

The eigenvalue e will now be computed numerically and then the energy eigenvalue E will be found using the relation in equation 4.46. Due to theorem 5.1 we know that only real eigenfunctions need be found and hence, we choose u to be real. By using the finite difference form of equation 6.8, we can write the difference equation corresponding to equation 6.51 to be

$$u_{i+1} = [(y_i^2 - e)w^2 + 2]u_i - u_{i-1}. \quad (6.52)$$

For even solutions, one knows that u is nonzero at the origin. This initial value can be arbitrarily chosen to be 1, as the function need be known only upto an undetermined multiplicative constant (see rule 2). For the initial condition on the derivative, one can write an equation similar to equation 6.29:

$$u_1 = w(\Delta u)_0 + u_0. \quad (6.53)$$

For even solutions, u must have a derivative of zero at the origin. Hence, the approximation of the derivative in equation 6.53 can be set to zero. Thus we get the following initial conditions.

$$u_1 = u_0 = 1 \quad (\text{even}). \quad (6.54)$$

For odd solutions the value at the origin must be zero. But the value of the slope at the origin can be chosen arbitrarily due to rule 2. We shall choose the approximate value of the initial slope to be 1. Hence, the initial conditions are

$$u_0 = 0, \quad u_1 = w \quad (\text{odd}). \quad (6.55)$$

Using the conditions of equation 6.54 or 6.55, one can solve the recursion relation of equation 6.52 using $e = 0$. The computation should be stopped at some value $i = n$ such that u_n is very large (say 100). The sign of u_n should be noted. Next, the process should be repeated several times for values of e incremented by a small amount (say e_d) each time. If the sign of u_n changes between two consecutive value of e (say e_0 and $e_0 + e_d$), then, for some energy between these two values, the wavefunction must go to zero at large distances – This energy would be the lowest eigenvalue. If e_d is chosen to be smaller than tolerable error in eigenvalue computation, then e_0 itself can be accepted as an approximation of the lowest eigenvalue. If the above process is continued for higher energies, higher eigenvalues can be obtained.

Very often the e_d that is needed for the desired accuracy is so small that computation time becomes unacceptably large. Hence, the search for eigenvalues needs to be done in a more sophisticated manner. A binary search would be faster. However, to conduct such a search, one would need to know intervals in e that have one and only one eigenvalue each. Thus a practical approach would be to conduct a rough linear search as before with an e_d

just small enough such that the interval $[e, e + e_d]$ can have at most one eigenvalue. Once the interval in which an eigenvalue is located is found, a binary search within the interval can be conducted. This search technique is similar to the bisection method outlined in chapter 5 for the solution of equations 5.54 and 5.55 (see problem 2).

As one knows that only the eigenvalue is the observed quantity, the above numerical technique is usually quite sufficient. However, very often, to estimate the changes in an eigenvalue due to small changes in the potential energy function, one needs the wavefunction. The above method gives the correct wavefunction except for the unstable tail. The tail region of the solution is easily identified by a catastrophic increase (or decrease) in the computed function. Setting this part of the solution to be identically zero gives a reasonably good approximation for the wavefunction. However, for better results one may use a matrix method (see problem 3) to compute the function in the tail region. To do this one may assign the boundary condition of the function at a large distance to be zero. The other condition is that of continuity with the solution already obtained before the tail region (see problem 5). The matrix method could, of course, have been used for the complete solution rather than just the tail. But that would require a larger computer memory.

If the potential function $V(x)$ is not symmetric, finding bound state energy eigenvalues can become more tricky. In such a situation the matrix method might be easier to use. The value $u(0)$ ($= u_0$) of the function at the origin can be chosen to be 1 at the origin. Then for some energy E , a matrix solution can be found from $x = -\infty$ to $x = 0$ using the boundary conditions $u(-\infty) = 0$ and $u(0) = 1$. Here infinity is understood to be a large enough value for computer usage. Next, for the same energy E , another matrix solution can be found from $x = 0$ to $x = +\infty$. The derivatives of the two solutions must match at $x = 0$. This matching condition can be used to search for the eigenvalues of energy. If by accident the origin is chosen at a point where the wavefunction vanishes, a shift in the coordinate system would be necessary.

6.4 Other techniques

The numerical methods discussed in this chapter are some of the more intuitive and theoretically straightforward ones that are known. They are to be considered as only a beginning. In many situations the physicist needs to develop more specialized and sometimes more involved methods to suit the needs. Numerical tricks, that improve accuracy and speed in special cases, are continually being developed by numerical analysts. Hence, the reader is also encouraged to develop his (or her) own methods whenever the need arises.

6.5 Accuracy

In any numerical technique it is crucial to know the degree of accuracy. Without an error estimate the numerical results are quite worthless. One of the simplest and practical methods of error estimation involves computing the change in the final results due to a change in the interval w used in the solution of the differential equation. The number of significant figures of the solution that do not change due to a decrease in w , give the accuracy of the solution.

The approximation methods used in this chapter are based on expanding a function in a Taylor series and selecting a suitable number of terms from it. The error introduced by ignoring higher order terms in the Taylor series is called a truncation error. Truncation errors can usually be reduced by choosing a small enough interval w . However, very small intervals can increase what are called roundoff errors. The source of roundoff errors is the following.

A machine computation is usually done with a certain fixed number of significant figures. When the interval w is chosen to be small, it very often requires the computation of the small difference between two large numbers (e.g. a first difference). The number of significant figures of the difference can be seen to be much smaller than that of the original numbers even though the computer will arbitrarily fill in the missing significant figures. For example, the two numbers, 4.02395 and 4.02392, have six significant figures each. However, their difference, 3×10^{-5} , has only one significant figure. The computer will fill in significant figures and might consider this difference to be 3.00000×10^{-5} . The error introduced by arbitrarily assigning these extra significant figures is called a roundoff error. Roundoff errors can be reduced by choosing a larger number of significant figures for all computations and thus paying the price through longer computation times. Most computer languages provide at least two different choices of significant figures (viz. single or double precision). Higher precision computation can be obtained by some custom programming for the underlying arithmetic operations.

Truncation errors can also be reduced by using higher order algorithms that use a larger number of terms of the Taylor series expansion. Such methods are beyond the scope of this book.

6.6 Speed

Computation speeds can of course be improved by better computer hardware. Such improvements are limited only by technology and the financial status of the physicist. On the other hand improving speed through efficient software is an art that is often learnt from experience. In any effort in machine computation one needs to strike the right balance in

accuracy, computation speed, and time (or money) spent on software. While discussing the accuracy of computations we have already shown that, in general, an increase in accuracy is accompanied by a decrease in speed.

Problems

1. Show that equations 6.27 and 6.28 give approximate forms of the derivatives of g and h at the origin.
2. Based on the bisection method outlined in chapter 5 for the solution of equation 5.54, describe a method for the binary search of an eigenvalue (for a bound state) when an interval containing one and only one eigenvalue has already been located.
3. Show that recursion relations like those of equations 6.25, 6.26 and 6.52 can be written as $n - 1$ linear algebraic equations for $n + 1$ unknowns. This set of equations can be solved using numerical matrix methods if two more independent equations are included. Show the following:
 - (a) Initial conditions can provide these two equations.
 - (b) Boundary conditions on both ends can also provide these two equations.
4. Show that the one dimensional scattering problem can be solved by the matrix method of problem 3, if in equation 6.11 A is chosen to be 1 while B and C (of equation 6.15) are computed from the solution. For such a choice, find the four extra equations (two for the g recursion relation and two for the h recursion relation) needed to solve the matrix equations.
5. Find a numerical method based on the matrix method of problem 3 for the computation of the tail part of a bound state eigenfunction. Use the boundary condition that the function goes to zero at large distances.
6. Develop a computer algorithm to compute the reflection and transmission coefficients for the junction potential of a semiconductor p-n junction. Such a potential is given by the following.

$$\begin{aligned}
 V &= 0 & x < -d_p \\
 &= K_p(x + d_p)^2 & -d_p < x < 0 \\
 &= V_t - K_n(x - d_n)^2 & 0 < x < d_n \\
 &= V_t & x > d_n.
 \end{aligned}$$

where d_p , d_n , K_p and K_n are given constants characterizing the potential and $V_t = K_p d_p^2 + K_n d_n^2$.

7. Develop a computer algorithm to compute the eigenvalues for a quartic potential:

$$V = Kx^4.$$

where K is a given constant.

Chapter 7

Symmetries and Conserved Quantities

7.1 Symmetry groups and their representation

A general transformation of a system can be visualized as a coordinate transformation in some arbitrary coordinate system. A symmetry transformation is a transformation that keeps the physical characteristics of the system unchanged (for example, a rotation of a spherical object). In classical mechanics a symmetry transformation is defined as follows.

Definition 28 *A CLASSICAL SYMMETRY TRANSFORMATION of a system keeps the form of the hamiltonian unchanged.*

A quantum symmetry transformation can be defined to be the same. However, one knows that the quantum hamiltonian (H), like any operator, is specified completely by its operations on all possible states. As the set of eigenstates of the hamiltonian form a complete set, it would then suffice to specify the operation of H on all its eigenstates. This would of course amount to specifying all eigenvalues of H . Hence, for convenience, the following alternative definition will be used for quantum mechanics.

Definition 29 *A QUANTUM SYMMETRY TRANSFORMATION keeps the set of all eigenvalues and eigenstates of the hamiltonian unchanged. (Note: degenerate states with the same eigenvalue can exchange places in such a transformation.)*

In both classical and quantum mechanics, it can be seen that symmetry transformations become important due to their relation to conserved quantities. However, in quantum

mechanics the importance of symmetries is further enhanced by the fact that observation of conserved quantities can be exactly predictable in spite of the probabilistic nature of quantum predictions (see chapter 4). Hence, in this chapter, we shall study the nature of symmetry transformations in quantum mechanics.

Let us consider an arbitrary transformation of an arbitrary state $|s\rangle$ to be given by the operator U such that the transformation gives

$$|s\rangle \rightarrow U|s\rangle. \quad (7.1)$$

If U were to produce a symmetry transformation, the following theorem can be proved.

Theorem 7.1 *If the operator U produces a symmetry transformation on all ket vectors, then it must commute with the hamiltonian.*

Proof: By definition of a symmetry transformation, the operator U could transform an energy eigenstate either to itself or another eigenstate degenerate to it. Hence, if $|E_i\rangle$ is an eigenstate of H with eigenvalue E_i then

$$\begin{aligned} HU|E_i\rangle &= H|E'_i\rangle = E_i|E'_i\rangle \\ &= E_iU|E_i\rangle = UE_i|E_i\rangle = UH|E_i\rangle \end{aligned} \quad (7.2)$$

where $|E_i\rangle$ and $|E'_i\rangle$ are either degenerate or the same. This gives the result

$$[H, U]|E_i\rangle = 0. \quad (7.3)$$

The above equation is true for all energy eigenstates $|E_i\rangle$. From the completeness theorem one knows that any arbitrary state $|s\rangle$ can be written as a linear combination of the eigenstates $|E_i\rangle$. Hence, it follows that

$$[H, U]|s\rangle = 0. \quad (7.4)$$

As $|s\rangle$ is an arbitrary ket vector, one concludes

$$[H, U] = 0. \quad (7.5)$$

This proves the theorem.

The following definitions are going to be useful in the future.

Definition 30 *An operator U is called UNITARY if*

$$U^\dagger U = I$$

where I is the identity operator defined in chapter 1.

Definition 31 An operator Q is called *ANTILINEAR* if

$$Q(a|r\rangle + b|s\rangle) = a^*Q|r\rangle + b^*Q|s\rangle$$

where $a, b \in \mathcal{C}$.

Definition 32 An operator U is called *ANTIUNITARY* if it is antilinear and

$$\langle r|U^\dagger U|s\rangle = \langle r|s\rangle^* = \langle s|r\rangle$$

Now the following theorem can also be proved for symmetry transformations.

Theorem 7.2 If a linear operator U produces a symmetry transformation then it is unitary.

Proof: Let the set of eigenstates of the hamiltonian H be $\{|E_i\rangle\}$ and let the operation of U on this set be given by

$$U|E_i\rangle = |E'_i\rangle. \quad (7.6)$$

By the definition of a symmetry transformation the set $\{|E'_i\rangle\}$ is also that of eigenstates of the hamiltonian. They can also be chosen to be normalized. Consider two arbitrary states $|r\rangle$ and $|s\rangle$. They can be written as linear combinations of the energy eigenstates $|E_i\rangle$.

$$|r\rangle = \sum_i a_i |E_i\rangle \quad (7.7)$$

$$|s\rangle = \sum_i b_i |E_i\rangle. \quad (7.8)$$

Then, as U is linear, equation 7.6 gives

$$U|r\rangle = \sum_i a_i U|E_i\rangle = \sum_i a_i |E'_i\rangle \quad (7.9)$$

$$U|s\rangle = \sum_j b_j U|E_j\rangle = \sum_j b_j |E'_j\rangle. \quad (7.10)$$

Hence,

$$\langle r|U^\dagger U|s\rangle = \sum_{ij} a_i^* b_j \langle E'_i|E'_j\rangle = \sum_i a_i^* b_i = \langle r|s\rangle. \quad (7.11)$$

Here the orthonormality of energy eigenstates is used. As equation 7.11 is true for any two states $|r\rangle$ and $|s\rangle$, it follows that $U^\dagger U$ is the identity operator I .

$$U^\dagger U = I. \quad (7.12)$$

This completes the proof.

A more general form of theorem 7.2 can also be proved. However, here we shall state it without proof as follows.

Theorem 7.3 *An operator U that produces a symmetry transformation must be either unitary or antiunitary.*

We shall later see that the time reversal symmetry operator is an example of an antiunitary operator.

The set of all symmetry transformations of a system is called its symmetry group as it satisfies the mathematical properties of a special type of set called a group. A group is defined as follows:

Definition 33 *If G is a GROUP and $A, B, C \in G$, then the following are true:*

1. *The product $AB \in G$ if the product is defined as two successive operations.*
2. *The identity transformation $I \in G$ where I is defined by $AI = IA = A$.*
3. *Every $A (\in G)$ has a unique inverse $A^{-1} (\in G)$ such that $AA^{-1} = A^{-1}A = I$.*
4. *The product is associative i.e. $A(BC) = (AB)C$.*

Definition 34 *A CONTINUOUS GROUP (e.g. rotation group) is one whose elements have a one-to-one correspondence with the values of a set of continuous variables called the GROUP PARAMETERS (e.g. angle of rotation).*

Definition 35 *If the algebra of a group can be realized by a set of operators, then this set is called its REPRESENTATION.*

Definition 36 *The operators on the quantum states of a system, that generate transformations of these states according to the group elements of the symmetry of the system, must form a representation of that group. This will be called the QUANTUM STATES REPRESENTATION of the group.*

If θ is a group parameter, a group element (in its quantum state representation) infinitesimally different from the identity can be written as

$$U(d\theta) = I + Qd\theta, \quad (7.13)$$

where Q is an operator. From theorem 7.3 we know that $U(d\theta)$ could be either unitary or antiunitary. If it were antiunitary, it would have to be so even in the limit of $d\theta$ going to zero.

But from equation 7.13, we see that for $d\theta = 0$, $U(d\theta) = I$ which is not antiunitary. Hence, we conclude that $U(d\theta)$ must be unitary and in general the quantum states representation of any continuous group must be unitary as a continuous variation of the group parameters cannot bring about the discontinuous change from unitary to antiunitary. Unitarity of $U(d\theta)$ gives

$$I = U^\dagger(d\theta)U(d\theta) = (I + Q^\dagger d\theta)(I + Qd\theta) = I + (Q^\dagger + Q)d\theta, \quad (7.14)$$

where second order terms in the infinitesimal $d\theta$ are dropped. Hence,

$$Q^\dagger = -Q. \quad (7.15)$$

If $Q = iJ$ then equation 7.15 would give J to be hermitian, and

$$U(d\theta) = I + iJd\theta. \quad (7.16)$$

To obtain a finite group operation one may operate on a state n times with $U(d\theta)$ and then let $n \rightarrow \infty$, such that $nd\theta = \theta$ is finite. This gives

$$U(\theta) = \lim_{n \rightarrow \infty} (I + iJ\theta/n)^n = \exp(iJ\theta), \quad (7.17)$$

where J , the hermitian operator, is called a GENERATOR of the group. Very often J can be found to be an observable.

Corollary 7.1 *From theorem 7.1 and theorem 4.2, it follows that a generator of a symmetry group must be a conserved quantity (see problem 1).*

In chapter 4 we had shown that conserved quantities are particularly important for quantum mechanics because they are the only quantities whose measurements can be predicted precisely in the absence of experimental error. Corollary 7.1 gives a way of identifying these conserved quantities through the symmetries of the system. Hence, it becomes important to study the symmetries of a system. In the following sections we shall investigate some common symmetries and the corresponding conserved quantities.

7.2 Space translation symmetry

A system that appears the same from different points in space is said to have space translation symmetry or just TRANSLATION SYMMETRY. Let $U_s(\mathbf{q})$ denote the quantum states representation of the translation group for a translation in space by the displacement vector \mathbf{q} . The operation of this symmetry operator on an arbitrary state $|s\rangle$ can be best understood in its position representation $\psi_s(\mathbf{r})$. First, considering a one dimensional system we give an infinitesimal translation of dq to the one dimensional wavefunction $\psi_s(x) = \langle x|s\rangle$.

The result of the translation should make the new function at x be equal in value to the old function at $x - dq$ i.e.

$$\begin{aligned}\langle x|U_s(dq)|s\rangle &= \langle x - dq|s\rangle = \psi_s(x - dq) \\ &= \psi_s(x) - dq \frac{\partial \psi_s(x)}{\partial x} = \left(1 - dq \frac{\partial}{\partial x}\right) \psi_s(x).\end{aligned}\quad (7.18)$$

Comparing with equation 7.16 we see the generator for translation in one dimension to be proportional to the momentum operator P ($= -i\hbar\partial/\partial x$ in the position representation) i.e.

$$\langle x|U_s(dq)|s\rangle = (1 - iPdq/\hbar)\psi_s(x). \quad (7.19)$$

For a finite translation of q we use equation 7.17 to obtain

$$\langle x|U_s(q)|s\rangle = \exp(-iPq/\hbar)\psi_s(x). \quad (7.20)$$

Hence, in general, the translation operation in one dimension is represented by

$$U_s(q) = \exp(-iPq/\hbar). \quad (7.21)$$

This result can be generalized for the three dimensional case to give

$$U_s(\mathbf{q}) = \exp(-i\mathbf{P} \cdot \mathbf{q}/\hbar), \quad (7.22)$$

where \mathbf{P} is the momentum vector operator. Now, from corollary 7.1 we can at once conclude that in a translationally symmetric system the momentum is conserved. The free particle is an example of such a system.

7.3 Time translation symmetry

The Schrödinger equation is postulated to generate time translation. Hence, the quantum states representation of the time translation operator must be directly derivable from the Schrödinger equation. This has been done in problem 2 of chapter 4. Thus we find the time translation operator to be

$$U_t(t) = \exp(-iHt/\hbar). \quad (7.23)$$

This can be seen to be a symmetry operation for all conservative systems as in chapter 4 we showed that, for such systems, energy eigenstates do not change with time.

7.4 Rotation symmetry

Under a rotation the transformation of the rectangular coordinates $(x_1, x_2, x_3) = (x, y, z)$ is given by a matrix with elements $a_{ij}(i, j = 1, 2, 3)$ such that the transformed coordinates $(x'_1, x'_2, x'_3) = (x', y', z')$ are given by [2]

$$x'_i = \sum_j a_{ij} x_j, \quad i = 1, 2, 3. \quad (7.24)$$

The scalar product of two vectors is unchanged by this transformation. This leads to the condition

$$\sum_j a_{ij} a_{jk}^T = \sum_j a_{ij} a_{kj} = \delta_{ik}. \quad (7.25)$$

where δ_{ik} is the Kronecker delta. The set of all 3×3 matrices that satisfy equation 7.25 form the rotation group. This group is also known as the orthogonal group in 3 dimensions or $O(3)$. From equation 7.25 it follows that the matrices a_{ij} have the determinant $+1$ or -1 . The subset of these matrices that have the determinant $+1$ is also a group and it is called the $SO(3)$ group. $SO(3)$ contains all elements of $O(3)$ that can be continuously transformed to the identity i.e. it does not include coordinate inversions. At present we are going to discuss only $SO(3)$ as it has no discrete transformation elements and hence, can be represented in the form of equation 7.17.

A rotation about the z direction by an angle θ can be seen to be given by the following element of the rotation group.

$$a_z(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (7.26)$$

Hence, an infinitesimal rotation about the z direction by an angle $d\theta$ is given by

$$a_z(d\theta) = \begin{pmatrix} 1 & -d\theta & 0 \\ d\theta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (7.27)$$

Now we can derive the quantum states representation U_R of $SO(3)$. Under an infinitesimal rotation of $d\theta$, about the z axis, the position representation $\psi_s(\mathbf{r})$ of the state $|s\rangle$ becomes

$$\langle \mathbf{r} | U_R(d\theta) | s \rangle = \psi_s(\mathbf{r}'), \quad (7.28)$$

where $\mathbf{r}' = a_z^{-1}(d\theta)\mathbf{r}$. Hence,

$$\psi_s(\mathbf{r}') = \psi_s(x + yd\theta, y - xd\theta, z). \quad (7.29)$$

A Taylor series expansion up to first order terms in $d\theta$ would give

$$\begin{aligned}\psi_s(\mathbf{r}') &= \psi_s(\mathbf{r}) + yd\theta \frac{\partial \psi_s(\mathbf{r})}{\partial x} - xd\theta \frac{\partial \psi_s(\mathbf{r})}{\partial y} \\ &= \left[1 + \left(y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} \right) d\theta \right] \psi_s(\mathbf{r}).\end{aligned}\quad (7.30)$$

In the position representation the quantity in parenthesis can be seen to be

$$\begin{aligned}\left(y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} \right) &= i(yP_x - xP_y)/\hbar \\ &= -iL_z/\hbar,\end{aligned}\quad (7.31)$$

where P_x and P_y are the respective x and y components of momentum and L_z is the z component of angular momentum (in its operator form in the position representation). Hence, we conclude that the operator form of $U_R(d\theta)$ in any representation is

$$U_R(d\theta) = I - iL_z d\theta/\hbar. \quad (7.32)$$

The corresponding finite rotation by an angle θ about the z axis is

$$U_R(\theta) = \exp(-iL_z\theta/\hbar). \quad (7.33)$$

This can be generalized, as follows, for any rotation θ about some direction given by the unit vector $\hat{\mathbf{n}}$.

$$U_R(\theta) = \exp(-i\mathbf{L} \cdot \hat{\mathbf{n}}\theta/\hbar). \quad (7.34)$$

Thus we see that angular momentum is the generator of rotation i.e. $SO(3)$. Hence, it must be conserved in a system that is rotationally symmetric.

7.4.1 Eigenvalues of angular momentum

As L_z is conserved in a spherically symmetric system it will have simultaneous eigenstates with the hamiltonian (see problem 3 in chapter 4). Similarly, L_x , or L_y could also individually share eigenstates with the hamiltonian. However, different components of \mathbf{L} cannot share eigenstates as they do not commute. For example, if the position and momentum operators are $\mathbf{R} = (X, Y, Z)$ and $\mathbf{P} = (P_x, P_y, P_z)$ then

$$\begin{aligned}[L_x, L_y] &= [YP_z - ZP_y, ZP_x - XP_z] \\ &= Y[P_z, Z]P_x + X[Z, P_z]P_y \\ &= i\hbar(XP_y - YP_x) \\ &= i\hbar L_z.\end{aligned}\quad (7.35)$$

Similarly,

$$[L_y, L_z] = i\hbar L_x, \quad (7.36)$$

$$[L_z, L_x] = i\hbar L_y. \quad (7.37)$$

The magnitude squared of the total angular momentum, $L^2 (= L_x^2 + L_y^2 + L_z^2)$ commutes with each component. Hence, we can find simultaneous eigenstates for any one of the following three sets of operators: $\{H, L^2, L_x\}$, $\{H, L^2, L_y\}$, $\{H, L^2, L_z\}$. In the following, without loss of generality, we shall choose the simultaneous eigenstates of $\{H, L^2, L_z\}$. These eigenstates, $|c, d\rangle$, are labelled by c , the eigenvalue of L^2 , and d , the eigenvalue of L_z . The energy eigenvalue is suppressed as it will have a fixed value for the following discussion. Hence,

$$L^2|c, d\rangle = c|c, d\rangle, \quad L_z|c, d\rangle = d|c, d\rangle \quad (7.38)$$

We can now find the possible values of the eigenvalues c and d . To this end we first define the operators

$$L_+ = L_x + iL_y, \quad (7.39)$$

$$L_- = L_x - iL_y. \quad (7.40)$$

The commutators of L_+ and L_- can be obtained from equations 7.35, 7.36 and 7.37 to be

$$[L_+, L_-] = 2\hbar L_z, \quad (7.41)$$

$$[L_+, L_z] = -\hbar L_+, \quad (7.42)$$

$$[L_-, L_z] = \hbar L_-. \quad (7.43)$$

Then,

$$\begin{aligned} L_z L_+ |c, d\rangle &= L_+ L_z |c, d\rangle + \hbar L_+ |c, d\rangle \\ &= L_+ d |c, d\rangle + \hbar L_+ |c, d\rangle \\ &= (d + \hbar) L_+ |c, d\rangle. \end{aligned} \quad (7.44)$$

As L^2 commutes with all angular momentum components we can also see that

$$L^2 L_+ |c, d\rangle = c L_+ |c, d\rangle. \quad (7.45)$$

Hence, it is seen that the operator L_+ “raises” the eigenstate $|c, d\rangle$ to another eigenstate with the L_z eigenvalue greater by \hbar while the L^2 eigenvalue remains the same. Thus one can write

$$L_+ |c, d\rangle = N_d |c, d + \hbar\rangle, \quad (7.46)$$

where N_d is chosen to maintain normalization of the eigenstates. Similarly, it can be shown that

$$L_- |c, d\rangle = M_d |c, d - \hbar\rangle. \quad (7.47)$$

For a given eigenvalue of L^2 , the eigenvalue of L_z is expected to have both an upper limit and a lower limit. This can be seen from the fact that

$$L_-L_+|c, d\rangle = (L_x^2 + L_y^2 + i[L_x, L_y])|c, d\rangle. \quad (7.48)$$

From equations 7.35, 7.46 and 7.47 this gives

$$\begin{aligned} M_{d+\hbar}N_d|c, d\rangle &= (L_x^2 + L_y^2 - \hbar L_z)|c, d\rangle \\ &= (L^2 - L_z^2 - \hbar L_z)|c, d\rangle \\ &= (c - d^2 - \hbar d)|c, d\rangle. \end{aligned} \quad (7.49)$$

We also know that $\langle c, d|L_-$ is the adjoint of $L_+|c, d\rangle$. Hence,

$$|N_d|^2 = \langle c, d|L_-L_+|c, d\rangle = M_{d+\hbar}N_d, \quad (7.50)$$

and thus

$$M_{d+\hbar} = N_d^*. \quad (7.51)$$

Then from equation 7.49 we see that

$$c - d^2 - \hbar d = |N_d|^2 > 0, \quad (7.52)$$

or

$$d^2 + \hbar d < c. \quad (7.53)$$

This shows that for a given c , d has a maximum and a minimum value. Let the maximum value for d be $l\hbar$ where l is dimensionless. Then

$$L_+|c, l\hbar\rangle = 0, \quad (7.54)$$

and from equation 7.49 we get

$$0 = L_-L_+|c, l\hbar\rangle = (c - l^2\hbar^2 - l\hbar^2)|c, l\hbar\rangle, \quad (7.55)$$

or

$$c = l(l+1)\hbar^2. \quad (7.56)$$

As the eigenvalue of L_z has a minimum value for a given c , it is evident that a finite number of operations on $|c, l\hbar\rangle$ by L_- should bring it down to the eigenstate of that minimum eigenvalue. If this finite number is n (a non-negative integer) then

$$L_-|c, l\hbar - n\hbar\rangle = 0. \quad (7.57)$$

Hence,

$$\begin{aligned} 0 = L_+L_-|c, (l-n)\hbar\rangle &= (L_x^2 + L_y^2 - i[L_x, L_y])|c, (l-n)\hbar\rangle \\ &= (L^2 - L_z^2 + \hbar L_z)|c, (l-n)\hbar\rangle \\ &= [c - (l-n)^2\hbar^2 + (l-n)\hbar^2]|c, (l-n)\hbar\rangle. \end{aligned} \quad (7.58)$$

It then follows that

$$c - (l - n)(l - n - 1)\hbar^2 = 0. \quad (7.59)$$

Using equations 7.56 and 7.59 we obtain

$$l(l + 1) - (l - n)(l - n - 1) = 0, \quad (7.60)$$

or (as n is non-negative)

$$l = n/2. \quad (7.61)$$

Hence, l can take either integer or half integer values and the eigenvalue of L^2 corresponding to it is $l(l + 1)\hbar^2$. If the eigenvalue of L_z is written as $m\hbar$, m can take a maximum value of l . All other value of m differ from l by some integer. The minimum value of m is $l - n$ which, from equation 7.61, can be seen to be $-l$. Now we can label the angular momentum eigenstates by the numbers l and m (instead of c and d) i.e $|l, m\rangle$ such that

$$L^2|l, m\rangle = l(l + 1)\hbar^2|l, m\rangle, \quad (7.62)$$

and

$$L_z|l, m\rangle = m\hbar|l, m\rangle, \quad (7.63)$$

where l is either an integer or a half integer and m takes the values $l, l - 1, l - 2, \dots, -l$. Hence, the total number of m values for a given l is $2l + 1$.

Thus we have shown that angular momentum eigenvalues are discrete. However, these results are obtained only from the commutators of the generators and further restrictions may apply on the allowed eigenvalues when finite rotations are considered. From equation 7.33 a rotation of 2π on a state $|l, m\rangle$ would give

$$\exp(-2\pi i L_z / \hbar)|l, m\rangle = \exp(-2\pi i m)|l, m\rangle. \quad (7.64)$$

If l (and consequently m) is a half integer this does not produce the expected identity transformation. Hence, half integer l is not allowed for the kind of state vectors we have discussed so far. It will later be seen that state vectors that include spin will allow half integer l values.

Two relations that will be seen to be useful later are as follows.

$$L_+|l, m\rangle = [l(l + 1) - m(m + 1)]^{1/2}\hbar|l, m + 1\rangle \quad (7.65)$$

$$L_-|l, m\rangle = [l(l + 1) - m(m - 1)]^{1/2}\hbar|l, m - 1\rangle \quad (7.66)$$

These can be obtained from equations 7.49, 7.50, 7.56 and 7.63 (see problem 3).

7.4.2 Addition of angular momenta

If a system has two different components (e.g. a two particle system), each of which has a measurable angular momentum, the relation between individual component angular momenta and total angular momentum is nontrivial in quantum mechanics. Hence, we shall discuss it here. Let the individual component angular momenta be \mathbf{L}_1 and \mathbf{L}_2 and the total angular momentum be \mathbf{L} such that

$$\mathbf{L} = \mathbf{L}_1 + \mathbf{L}_2. \quad (7.67)$$

This means

$$L_+ = L_{1+} + L_{2+}, \quad (7.68)$$

$$L_- = L_{1-} + L_{2-}, \quad (7.69)$$

$$L_z = L_{1z} + L_{2z}, \quad (7.70)$$

$$L^2 = L_1^2 + L_2^2 + 2\mathbf{L}_1 \cdot \mathbf{L}_2, \quad (7.71)$$

where the subscripts $+$, $-$, and z mean the same for the vectors \mathbf{L}_1 and \mathbf{L}_2 as they do for \mathbf{L} . If the eigenstates of L_1^2 and L_{1z} are labelled as $|l_1, m_1\rangle$ and that of L_2^2 and L_{2z} as $|l_2, m_2\rangle$, then the combined system can be represented by the direct product $|l_1, m_1\rangle \otimes |l_2, m_2\rangle$ of these states. A more compact notation for these direct products would be

$$|l_1, m_1\rangle \otimes |l_2, m_2\rangle \equiv |l_1, l_2, m_1, m_2\rangle \equiv |m_1, m_2\rangle. \quad (7.72)$$

In the last form l_1 and l_2 are suppressed. This is convenient when the values of l_1 and l_2 are fixed for some computation. It is possible to choose l_1 , m_1 , l_2 , and m_2 as labels as the corresponding operators L_1^2 , L_{1z} , L_2^2 , and L_{2z} commute with each other and hence have simultaneous eigenstates. Another set of commuting operators is $\{L_1^2, L_2^2, L^2, L_z\}$. Hence, a set of simultaneous eigenstates for these operators can be found. These eigenstates would be labelled as $|l_1, l_2, l, m\rangle$ such that

$$L_1^2 |l_1 l_2 l m\rangle = l_1(l_1 + 1)\hbar^2 |l_1 l_2 l m\rangle, \quad (7.73)$$

$$L_2^2 |l_1 l_2 l m\rangle = l_2(l_2 + 1)\hbar^2 |l_1 l_2 l m\rangle, \quad (7.74)$$

$$L^2 |l_1 l_2 l m\rangle = l(l + 1)\hbar^2 |l_1 l_2 l m\rangle, \quad (7.75)$$

$$L_z |l_1 l_2 l m\rangle = m\hbar |l_1 l_2 l m\rangle, \quad (7.76)$$

where the commas within the ket are omitted. We shall call the $|l_1 l_2 m_1 m_2\rangle$ states of equation 7.72, the individual angular momenta states and the $|l_1 l_2 l m\rangle$ states the total angular momentum states. As these two different sets of eigenstates describe the same system, they must be related as linear combinations of each other. The coefficients for such linear combinations are called the CLEBSCH-GORDAN COEFFICIENTS. If the common labels (l_1 and l_2) of the two sets are suppressed, then one may write

$$|lm\rangle = \sum_{m_1 m_2} |m_1 m_2\rangle \langle m_1 m_2 | lm\rangle \quad (7.77)$$

as $\sum_{m_1 m_2} |m_1 m_2\rangle \langle m_1 m_2|$ would be the identity in the subspace where l_1 and l_2 are fixed. From equation 7.77 we see that the coefficients $\langle m_1 m_2 | l m \rangle$ are the Clebsch-Gordan coefficients.

We shall discuss some useful general results before computing these coefficients. Operating equation 7.77 by $L_z (= L_{1z} + L_{2z})$ and then multiplying from the left by $\langle m_1 m_2 |$ one obtains

$$m \langle m_1 m_2 | l m \rangle = (m_1 + m_2) \langle m_1 m_2 | l m \rangle. \quad (7.78)$$

Hence, the coefficient $\langle m_1 m_2 | l m \rangle$ can be nonzero only if

$$m = m_1 + m_2. \quad (7.79)$$

So for fixed values of l_1 and l_2 , the largest value of m can be $l_1 + l_2$. Also, as l is the largest value of m , the largest value of l is $l_1 + l_2$. Hence, for $l = m = l_1 + l_2$, equation 7.77 would become

$$|l_1 + l_2, l_1 + l_2\rangle = |l_1 l_2\rangle \langle l_1 l_2 | l_1 + l_2, l_1 + l_2\rangle. \quad (7.80)$$

As all eigenstates are normalized, one could choose $\langle l_1 l_2 | l_1 + l_2, l_1 + l_2\rangle = 1$. This would give

$$|l_1 + l_2, l_1 + l_2\rangle = |l_1 l_2\rangle. \quad (7.81)$$

In the last two equations, one notices a possibility of confusion in notation of the individual angular momenta states and the total angular momentum states. To avoid such confusion, we shall always have the individual angular momenta kets on the right side of an equation unless they are part of an inner product. Similarly, the total angular momentum kets will always be placed on the left side of an equation. From equation 7.71 it can be seen that l can take different values for fixed values of l_1 and l_2 . However, as l is the maximum value of m (which is $m_1 + m_2$), it can be less than $l_1 + l_2$ only by a positive integer. The minimum value of l is $|l_1 - l_2|$. This can be seen from the fact that the number of eigenstates for the individual angular momenta must be the same as that for the total angular momentum (for fixed l_1 and l_2) as both sets of eigenstates form a complete set spanning the same space (see problem 5). We can now illustrate a general method of computing Clebsch-Gordan coefficients through an example. If $l_1 = 1$, and $l_2 = 1/2$, then from equation 7.81 we get

$$|3/2, 3/2\rangle = |1, 1/2\rangle. \quad (7.82)$$

Multiplying this by $L_- (= L_{1-} + L_{2-})$ gives

$$L_- |3/2, 3/2\rangle = L_{1-} |1, 1/2\rangle + L_{2-} |1, 1/2\rangle. \quad (7.83)$$

Using equation 7.66, for both individual angular momenta and total angular momentum, one obtains

$$\sqrt{3} |3/2, 1/2\rangle = \sqrt{2} |0, 1/2\rangle + |1, -1/2\rangle, \quad (7.84)$$

or

$$|3/2, 1/2\rangle = \sqrt{2/3} |0, 1/2\rangle + \sqrt{1/3} |1, -1/2\rangle. \quad (7.85)$$

This gives the following Clebsch-Gordan coefficients.

$$\langle 0, 1/2 | 3/2, 1/2 \rangle = \sqrt{2/3}, \quad (7.86)$$

$$\langle 1, -1/2 | 3/2, 1/2 \rangle = \sqrt{1/3}. \quad (7.87)$$

Operating on equation 7.85 by L_- once again would give

$$\begin{aligned} 2|3/2, -1/2\rangle &= \sqrt{2/3}\sqrt{2}|-1, 1/2\rangle + \sqrt{2/3}|0, -1/2\rangle \\ &+ \sqrt{1/3}\sqrt{2}|0, -1/2\rangle + 0, \end{aligned} \quad (7.88)$$

or

$$|3/2, -1/2\rangle = \sqrt{1/3}|-1, 1/2\rangle + \sqrt{2/3}|0, -1/2\rangle. \quad (7.89)$$

Operating again by L_- gives

$$|3/2, -3/2\rangle = |-1, -1/2\rangle. \quad (7.90)$$

The next possible value for l is $(3/2 - 1) = 1/2$. In this case the state of highest possible m is $|1/2, 1/2\rangle$. Due to equation 7.79, this state could be written as the linear combination

$$|1/2, 1/2\rangle = a|0, 1/2\rangle + b|1, -1/2\rangle. \quad (7.91)$$

This state must be orthonormal to all other total angular momentum states and in particular to $|3/2, 1/2\rangle$ as given by equation 7.85. Hence, we obtain

$$|1/2, 1/2\rangle = \sqrt{1/3}|0, 1/2\rangle - \sqrt{2/3}|1, -1/2\rangle. \quad (7.92)$$

Operating this by L_- gives

$$|1/2, -1/2\rangle = \sqrt{2/3}|-1, 1/2\rangle - \sqrt{1/3}|0, -1/2\rangle. \quad (7.93)$$

This completes the computation. The Clebsch-Gordan coefficients for $l_1 = 1$ and $l_2 = 1/2$ as obtained in the equations 7.82, 7.85, 7.89, 7.90, 7.92 and 7.93, can now be summarized in the following chart.

		l	$3/2$	$3/2$	$1/2$	$3/2$	$1/2$	$3/2$
		m	$3/2$	$1/2$	$1/2$	$-1/2$	$-1/2$	$-3/2$
m_1	m_2							
1	1/2	1						
1	-1/2		$\sqrt{1/3}$	$-\sqrt{2/3}$				
0	1/2		$\sqrt{2/3}$	$\sqrt{1/3}$				
0	-1/2				$\sqrt{2/3}$	$-\sqrt{1/3}$		
-1	1/2				$\sqrt{1/3}$	$\sqrt{2/3}$		
-1	-1/2							1

(7.94)

7.5 Discrete symmetries

A discrete subgroup of a continuous symmetry group can always be defined by choosing the group parameter at periodic intervals (see problem 7). However, here we are going to discuss some discrete symmetries that are not subgroups of continuous groups. Such symmetries are not associated to any conserved quantities, as they have no generators. The following two discrete symmetries are of general importance in physics.

7.5.1 Space inversion

The space inversion operator, I_s , has the following operation on the position vector \mathbf{r} .

$$I_s \mathbf{r} = -\mathbf{r}. \quad (7.95)$$

The quantum states representation of I_s will be called U_I and for an arbitrary state $|s\rangle$

$$\langle \mathbf{r} | U_I | s \rangle = j \langle -\mathbf{r} | s \rangle = j \psi_s(-\mathbf{r}). \quad (7.96)$$

The extra factor j is needed due to the discreteness of the symmetry. In continuous symmetry operations the value of j is unity as in the limit of all symmetry parameters going to zero the wavefunction must stay unchanged. For a discrete symmetry such a limit cannot be defined. For the spinless particles that we have discussed till now, two space inversions should produce the original wavefunction i.e.

$$U_I^2 |s\rangle = |s\rangle, \quad \text{for any } |s\rangle. \quad (7.97)$$

Hence, from equation 7.96 we get

$$j^2 = 1, \quad j = \pm 1. \quad (7.98)$$

The value of j is called the INTRINSIC PARITY of the system.

Theorem 7.4 *The energy eigenstates of an inversion symmetric system can be chosen such that they change at most by a sign under the inversion operation.*

Proof: If $|E\rangle$ is an eigenstate of energy with eigenvalue E , then from the definition of a quantum symmetry $U_I |E\rangle$ is also an eigenstate with the same eigenvalue. Hence, the following are also eigenstates of energy with the same eigenvalue.

$$|E_1\rangle = |E\rangle + U_I |E\rangle, \quad |E_2\rangle = |E\rangle - U_I |E\rangle. \quad (7.99)$$

From equations 7.97 and 7.99 it can be seen that

$$U_I |E_1\rangle = +|E_1\rangle, \quad U_I |E_2\rangle = -|E_2\rangle. \quad (7.100)$$

This proves the theorem.

Definition 37 *The energy eigenstates of the type $|E_1\rangle$ in equation 7.100 are called SYMMETRIC and they are also said to have POSITIVE (TOTAL) PARITY. The eigenstates of the type $|E_2\rangle$ in equation 7.100 are called ANTISYMMETRIC and they are also said to have NEGATIVE (TOTAL) PARITY.*

It should be noted that the intrinsic parity is included in the total parity. However, the intrinsic parity of particles cannot be absolutely determined. The intrinsic parities of some particles have to be assumed and then those of others can be determined if the system is inversion symmetric (i.e. total parity is conserved).

From equation 7.97 we see that $U_I^\dagger = U_I$ and hence, in the position representation (using equations 7.96 and 7.98)

$$U_I^\dagger \mathbf{r} U_I \psi_s(\mathbf{r}) = U_I^\dagger \mathbf{r} j \psi_s(-\mathbf{r}) = -\mathbf{r} \psi_s(\mathbf{r}). \quad (7.101)$$

As this is true for any state $\psi_s(\mathbf{r})$, the following must be true for the position operator \mathbf{R} .

$$U_I^\dagger \mathbf{R} U_I = -\mathbf{R}. \quad (7.102)$$

Similarly, for the momentum operator \mathbf{P}

$$U_I^\dagger \mathbf{P} U_I = -\mathbf{P}. \quad (7.103)$$

Hence, for the angular momentum operator, $\mathbf{L} = \mathbf{R} \times \mathbf{P}$, one obtains

$$U_I^\dagger \mathbf{L} U_I = \mathbf{L}. \quad (7.104)$$

7.5.2 Time reversal

The time reversal operator is expected to be different in nature from all other symmetry operators discussed upto now. This is due to the fact that the Schrödinger equation is first order in time and hence, a time reversal would change the sign of only the time derivative term. To be precise it can be seen that the time reversal operator, T , is antiunitary. We have seen this to be possible from theorem 7.3. However, we have also seen that continuous group transformations cannot be antiunitary. So, due to its discrete nature, it is possible for T to be antiunitary. To demonstrate the antiunitary nature of T , let us consider the energy eigenstate $|E\rangle$ at time $t = 0$ that has the eigenvalue E . The result of a time translation of t followed by a time reversal must be the same as that of a time reversal followed by a time translation of $-t$. Hence, from equation 7.23

$$\begin{aligned} T U_t(t) |E\rangle &= U_t(-t) T |E\rangle, \\ T \exp(-iHt/\hbar) |E\rangle &= \exp(iHt/\hbar) T |E\rangle, \\ T \exp(-iEt/\hbar) |E\rangle &= \exp(iEt/\hbar) T |E\rangle, \\ T \exp(-iEt/\hbar) |E\rangle &= [\exp(-iEt/\hbar)]^* T |E\rangle. \end{aligned} \quad (7.105)$$

As the above equation is true for any t and E , an arbitrary state $|s\rangle$ that can be written as the linear combination

$$|s\rangle = \sum_E a_E |E\rangle, \quad (7.106)$$

would be time reversed as

$$T|s\rangle = \sum_E a_E^* T|E\rangle. \quad (7.107)$$

Also the arbitrary state

$$|r\rangle = \sum_E b_E |E\rangle, \quad (7.108)$$

is time reversed as

$$T|r\rangle = \sum_E b_E^* T|E\rangle \quad (7.109)$$

Hence,

$$\langle r|T^\dagger T|s\rangle = \sum_{E'E} b_{E'} a_E^* \langle E'|T^\dagger T|E\rangle. \quad (7.110)$$

Due to time reversal symmetry the set of states $\{T|E\rangle\}$ for all E would be the same as the set of states $\{|E\rangle\}$ for all E . Hence,

$$\langle E'|T^\dagger T|E\rangle = \delta_{E'E} \quad (7.111)$$

and then from equation 7.110 it follows that

$$\langle r|T^\dagger T|s\rangle = \sum_E b_E a_E^* = \langle s|r\rangle. \quad (7.112)$$

This demonstrates, from definition, that T is an antiunitary operator.

Under a time reversal, one expects the position operator to stay unchanged and the momentum operator to change in sign. Thus

$$T\mathbf{R} = \mathbf{R}T \quad T\mathbf{P} = -\mathbf{P}T. \quad (7.113)$$

Hence, the angular momentum $\mathbf{L} = \mathbf{R} \times \mathbf{P}$ has the property

$$T\mathbf{L} = -\mathbf{L}T. \quad (7.114)$$

Problems

1. Prove corollary 7.1.
2. Show that the momentum eigenstates of a particle stay physically unchanged by a space translation.

3. Derive the equations 7.65 and 7.66.
4. For a system of two angular momenta with given magnitudes of individual angular momenta (l_1 and l_2 fixed) show that the number of angular momentum eigenstates is $(2l_1 + 1)(2l_2 + 1)$.
5. Show the minimum value of l for the total angular momentum states is $|l_1 - l_2|$. [Hint: For fixed values of l_1 and l_2 , the number of eigenstates of both the individual angular momenta and total angular momentum are the same.]
6. Find the Clebsch-Gordan coefficients for $l_1 = 1$ and $l_2 = 1$.
7. A periodic potential $V(\mathbf{r})$ has a three dimensional periodicity given by the vector $\mathbf{a} = (n_1 a_1, n_2 a_2, n_3 a_3)$ where a_1 , a_2 and a_3 are fixed lengths and n_1 , n_2 and n_3 can take any integer values such that

$$V(\mathbf{r} + \mathbf{a}) = V(\mathbf{r}).$$

- (a) Show that the discrete translation symmetry operators

$$U_d(\mathbf{a}) = \exp(-i\mathbf{P} \cdot \mathbf{a}/\hbar)$$

commute with the hamiltonian.

- (b) Show that the set $\{U_d(\mathbf{a})\}$ for all possible integers (n_1, n_2, n_3) in \mathbf{a} form a group.
 (c) The Bloch states are defined by their position representation

$$u_B = u(\mathbf{r}) \exp(i\mathbf{p} \cdot \mathbf{r}/\hbar),$$

where $u(\mathbf{r} + \mathbf{a}) = u(\mathbf{r})$ and \mathbf{p} gives three labels for such states. Show that these states are physically unchanged by $U_d(\mathbf{a})$.

Chapter 8

Three Dimensional Systems

The generalization of problem solving methods to three dimensions is conceptually simple. However, the mathematical details can be quite nontrivial. In this chapter we shall discuss some analytical methods in three dimensions. Quite obviously, such methods can have only limited applicability. However, the analytical solution of the hydrogen atom problem provides a better understanding of more complex atoms and molecules. Numerical methods in three dimensions can either be based on the analytical hydrogen atom solution or be independent of it, according to the nature of the system.

8.1 General characteristics of bound states

The characteristics of bound states in one dimension can be generalized to three dimensions. If $E < V$ at large distances in all directions, then the energy eigenvalues must be discrete. However, in three dimensions, there must be two other quantities, besides energy, that must also have discrete values. This is because in each dimension the condition of finiteness of the wavefunction will lead to some parameter being allowed only discrete values (using similar arguments as for energy in the one dimensional case discussed in chapter 5). For spherically symmetric potentials, it will be seen that the two extra discrete parameters are the eigenvalues of one angular momentum component (say L_z) and the magnitude of the total angular momentum (L^2). A study of numerical methods (chapter 9) will clarify the nature of these discrete parameters for more general cases.

8.2 Spherically symmetric potentials

A spherically symmetric potential is defined to be invariant under any rotation about one fixed center of rotation. The center of rotation defines the origin of a convenient coordinate system. Spherical polar coordinates (r, θ, ϕ) can be defined using this origin such that the corresponding rectangular coordinates are given by

$$\begin{aligned}x &= r \sin \theta \cos \phi \\y &= r \sin \theta \sin \phi \\z &= r \cos \theta.\end{aligned}\tag{8.1}$$

In three dimensions the time independent Schrödinger equation has the following form.

$$-\frac{\hbar^2}{2m}\nabla^2 u + Vu = Eu.\tag{8.2}$$

For a spherically symmetric potential, V is a function of r alone. Hence, in spherical polar coordinates equation 8.2 would have the form:

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]u + V(r)u = Eu.\tag{8.3}$$

As V depends only on r , the following separation of variables for the function u is useful.

$$u(r, \theta, \phi) = R(r)Y(\theta, \phi).\tag{8.4}$$

Inserting equation 8.4 in equation 8.3 and dividing by u gives

$$\begin{aligned}\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \frac{2mr^2}{\hbar^2}[E - V(r)] &= \\-\frac{1}{Y}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2}\right].\end{aligned}\tag{8.5}$$

As the left side of equation 8.5 depends only on r and the right side only on θ and ϕ , it is evident that both sides of the equation must be equal to a constant (say K). Then we have the following two equations.

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \frac{2m}{\hbar^2}[E - V(r)]R - \frac{K}{r^2}R = 0,\tag{8.6}$$

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2} + KY = 0.\tag{8.7}$$

Equation 8.7 is independent of the potential. Hence, we shall solve it first. The variables of $Y(\theta, \phi)$ can be further separated as follows.

$$Y(\theta, \phi) = \Theta(\theta)\Phi(\phi).\tag{8.8}$$

This would lead to the following two separated equations.

$$\frac{d^2\Phi}{d\phi^2} + m^2\Phi = 0, \quad (8.9)$$

$$\frac{1}{\sin\theta} \frac{d}{d\theta} \left(\sin\theta \frac{d\Theta}{d\theta} \right) + \left(K - \frac{m^2}{\sin^2\theta} \right) \Theta = 0. \quad (8.10)$$

The separation constant m^2 is like the K of the previous separation of the variable r . This m is not to be confused with the mass of the particle. The same symbol is used to maintain standard notation and the context of usage is seen to remove ambiguity. Equation 8.9 can be readily solved to give

$$\begin{aligned} \Phi &= A \exp(im\phi) + B \exp(-im\phi) \quad \text{for } m \neq 0, \\ \Phi &= A + B\phi \quad \text{for } m = 0. \end{aligned} \quad (8.11)$$

To maintain the continuity of the function, it is necessary to require that the value of the function be the same at $\phi = 0$ and $\phi = 2\pi$. This gives the only possible solutions to be

$$\Phi = \exp(im\phi), \quad (8.12)$$

where m is an integer. A constant coefficient is omitted here as it can be included in an overall normalization constant for Y .

The solution of equation 8.10 is more involved. The following change of variables makes it easier to handle.

$$w = \cos\theta. \quad (8.13)$$

The resulting form of equation 8.10 is

$$\frac{d}{dw} \left[(1-w^2) \frac{dP}{dw} \right] + \left(K - \frac{m^2}{1-w^2} \right) P = 0, \quad (8.14)$$

where $P(w) = \Theta(\theta)$. As θ belongs to the interval $[0, \pi]$, w must belong to the interval $[-1, +1]$. A standard series solution of equation 8.14 shows that $P(w)$ is finite in this interval only if

$$K = l(l+1), \quad (8.15)$$

where l is a non-negative integer. As the wavefunction is necessarily finite, equation 8.15 is a required condition for allowed solutions. The allowed solutions of equation 8.14 for $m = 0$ are the well known Legendre polynomials $P_l(w)$, where l , the order of the polynomial, is given by equation 8.15. For non-zero m , the allowed solutions of equation 8.14 are the associated Legendre functions P_l^m that are defined as follows.

$$P_l^m = (1-w^2)^{|m|/2} \frac{d^{|m|}}{dw^{|m|}} P_l(w). \quad (8.16)$$

The P_l^m can be seen to be non-zero only if

$$|m| < l. \quad (8.17)$$

Now, the possible solutions of equation 8.7 can be written as

$$Y_{lm}(\theta, \phi) = N_{lm} P_l^m(\cos \theta) \exp(im\phi), \quad (8.18)$$

where N_{lm} are the normalization constants. If the N_{lm} are chosen to be

$$N_{lm} = \sqrt{\frac{(2l+1)(l-|m|)!}{4\pi(l+|m|)!}}, \quad (8.19)$$

then the Y_{lm} are seen to be mutually orthonormal in the following sense.

$$\int_0^{2\pi} \int_0^\pi Y_{lm} Y_{l'm'} \sin \theta d\theta d\phi = \delta_{ll'} \delta_{mm'}. \quad (8.20)$$

The functions Y_{lm} are called the spherical harmonics. Some of the lower order spherical harmonics are as follows.

$$\begin{aligned} Y_{0,0} &= 1/\sqrt{4\pi} \\ Y_{1,0} &= \sqrt{\frac{3}{4\pi}} \cos \theta, \\ Y_{1,\pm 1} &= \sqrt{\frac{3}{8\pi}} \sin \theta \exp(\pm i\phi), \\ Y_{2,0} &= \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1), \\ Y_{2,\pm 1} &= \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta \exp(\pm i\phi), \\ Y_{2,\pm 2} &= \sqrt{\frac{15}{32\pi}} \sin^2 \theta \exp(\pm 2i\phi). \end{aligned} \quad (8.21)$$

8.3 Angular momentum

In chapter 7, we found that a spherically symmetric system must have the three components of angular momentum as conserved quantities. Also, the operators L_z and L^2 commute and hence have simultaneous eigenstates. In the position representation, the Y_{lm} can be seen to be these eigenstates. By definition,

$$\mathbf{L} = \mathbf{R} \times \mathbf{P}, \quad (8.22)$$

and hence, in the position representation

$$\begin{aligned} L_x &= -i\hbar \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right), \\ L_y &= -i\hbar \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right), \\ L_z &= -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right). \end{aligned} \quad (8.23)$$

A transformation to spherical polar coordinates gives

$$\begin{aligned} L_x &= i\hbar \left(\sin \phi \frac{\partial}{\partial \theta} + \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right), \\ L_y &= i\hbar \left(-\cos \phi \frac{\partial}{\partial \theta} + \cot \theta \sin \phi \frac{\partial}{\partial \phi} \right), \\ L_z &= -i\hbar \frac{\partial}{\partial \phi}. \end{aligned} \quad (8.24)$$

This also leads to the following expression for L^2 .

$$\begin{aligned} L^2 &= L_x^2 + L_y^2 + L_z^2 \\ &= -\hbar^2 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right]. \end{aligned} \quad (8.25)$$

From equations 8.7 and 8.15, it can now be seen that

$$L^2 Y_{lm} = l(l+1)\hbar^2 Y_{lm}, \quad (8.26)$$

and from equation 8.18 one obtains

$$L_z Y_{lm} = m\hbar Y_{lm}. \quad (8.27)$$

Hence, the eigenvalues of L_z and L^2 are in accordance with the general results of chapter 7. Half-integer values of l are not allowed here. Later, it will be seen that half-integer values of l are possible only when the position eigenstates are degenerate (see chapter 12).

8.4 The two body problem

In realistic situations a spherically symmetric potential does not appear as a background potential for a single particle system. What is commonly encountered is a two (point) particle system with the force acting along the line joining the two particles. Such a system can be shown to be mathematically equivalent to a combination of two independent systems – a free particle system and a spherically symmetric one particle system.

The hamiltonian for such a two particle system would be

$$H_t = \frac{P_1^2}{2m_1} + \frac{P_2^2}{2m_2} + V(r), \quad (8.28)$$

where \mathbf{P}_1 and \mathbf{P}_2 are the momenta of the two particles (with magnitudes P_1 and P_2), \mathbf{r}_1 and \mathbf{r}_2 are their positions and r is the magnitude of

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2. \quad (8.29)$$

In the position representation, equation 8.28 gives

$$H_t = -\frac{\hbar^2}{2m_1}\nabla_1^2 - \frac{\hbar^2}{2m_2}\nabla_2^2 + V(r), \quad (8.30)$$

where ∇_1^2 and ∇_2^2 have the meaning of the Laplacians for the position vectors \mathbf{r}_1 and \mathbf{r}_2 . The center of mass coordinates are defined by

$$\mathbf{r}_c = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2} \quad (8.31)$$

It can then be shown that the hamiltonian in terms of the center of mass coordinates \mathbf{r}_c and the relative coordinates \mathbf{r} can be written as

$$H_t = -\frac{\hbar^2}{2M}\nabla_c^2 - \frac{\hbar^2}{2m}\nabla^2 + V(r), \quad (8.32)$$

where ∇_c^2 is the Laplacian in \mathbf{r}_c and ∇^2 is the Laplacian in \mathbf{r} . Also,

$$M = m_1 + m_2, \quad m = \frac{m_1m_2}{m_1 + m_2}. \quad (8.33)$$

Now H_t can be written as

$$H_t = H_c + H, \quad (8.34)$$

where

$$H_c = -\frac{\hbar^2}{2M}\nabla_c^2, \quad H = -\frac{\hbar^2}{2m}\nabla^2 + V(r). \quad (8.35)$$

As H_c depends only on \mathbf{r}_c and H depends only on \mathbf{r} , one can find their eigenvalues E_c and E , independently. Then the eigenvalues E_t of H_t would be

$$E_t = E_c + E. \quad (8.36)$$

H_c is the three dimensional free particle hamiltonian. Hence, its eigenvalues are continuous (from a generalization of the one dimensional case). The eigenvalues E of H can be found by solving the differential eigenvalue problem

$$Hu = Eu. \quad (8.37)$$

This can be seen to be equivalent to a one particle spherically symmetric problem in the relative coordinate \mathbf{r} if the equivalent mass is taken to be m (the so-called reduced mass).

If H has a discrete spectrum, it can be “washed out” by the continuous spectrum of H_c if the energies E_c are large. This is observed in the case of the hydrogen atom. The energies E_c are due to the random motion of the centers of mass of the atoms. Hence, at high temperatures E_c is higher. So, to observe the discrete spectrum (of E) of the bound states formed by the electron and the proton of the atom, one needs to make observations at sufficiently low temperatures. Higher the temperature, broader will be the spectral lines observed.

8.5 The hydrogen atom (bound states)

The hydrogen atom is a two particle system (one electron and one proton) as discussed above. Its reduced mass can be seen to be almost equal to the electron mass as the proton mass is much larger. The electrostatic potential energy is

$$V(r) = -\frac{k_e e^2}{r}, \quad (8.38)$$

where $k_e = 1/(4\pi\epsilon_0)$, ϵ_0 is the permittivity of free space and e is the proton charge. Using this potential one can find the bound state energies E from equation 8.37. The form of the potential shows that bound states must have $E < 0$. With this condition one can solve equation 8.37 using the general method of section 8.2. The angular part of the eigenfunctions is already known. The radial part in the present case would be (from equations 8.6, 8.15, 8.33 and 8.38)

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) + \frac{2mk_e e^2}{\hbar^2 r} R + \frac{2mE}{\hbar^2} R - \frac{l(l+1)}{r^2} R = 0. \quad (8.39)$$

This equation can be solved in a manner very similar to the one dimensional harmonic oscillator problem. To simplify the form of the equation, the dimensionless parameter s is defined as

$$s = \alpha r, \quad \alpha^2 = -8mE/\hbar^2 \quad (8.40)$$

With the definition

$$\beta = \frac{2mk_e e^2}{\alpha \hbar^2} = \frac{k_e e^2}{\hbar} \sqrt{\frac{-m}{2E}} \quad (8.41)$$

equation 8.39 becomes

$$\frac{1}{s^2} \frac{d}{ds} \left(s^2 \frac{dR}{ds} \right) + \left[\frac{\beta}{s} - \frac{1}{4} - \frac{l(l+1)}{s^2} \right] R = 0. \quad (8.42)$$

The large s behavior of R can be seen to be of the form $\exp(-s/2)$. Hence, we choose

$$R(s) = F(s) \exp(-s/2). \quad (8.43)$$

From equations 8.42 and 8.43 one obtains the differential equation for F to be

$$\frac{d^2F}{ds^2} + \left(\frac{2}{s} - 1\right) \frac{dF}{ds} + \left[\frac{\beta - 1}{s} - \frac{l(l+1)}{s^2}\right] F = 0. \quad (8.44)$$

The following series form for F is chosen.

$$F = s^p \sum_{i=0}^{\infty} a_i s^i, \quad a_0 \neq 0. \quad (8.45)$$

On substituting this in equation 8.44 and equating the total coefficient of the lowest power of s to zero one obtains $p = l$ or $p = -(l+1)$. As $p = -(l+1)$ would make the eigenfunction at the origin infinite, we must choose $p = l$. In a manner similar to the harmonic oscillator problem, the substitution of the series solution in equation 8.44 results in a recursion relation for the a_i 's.

$$a_{i+1} = \frac{(i + l + 1 - \beta)a_i}{(i + 1)(i + 2l + 2)}. \quad (8.46)$$

Once again it can be shown that unless the series terminates, the function F rises fast enough at infinity to make R go to infinity at infinity. Hence, for an eigenstate the series must terminate. This will happen if for some non-negative integer n'

$$\beta = n' + l + 1 \equiv n. \quad (8.47)$$

Hence, from equation 8.41 the corresponding condition on the energy is seen to be (a subscript n is added to E to identify the energy level)

$$E_n = -\frac{mk_e^2 e^4}{2\hbar^2 n^2}, \quad n = 1, 2, 3, \dots \quad (8.48)$$

and $n > l$.

The differences of these energy eigenvalues are observed as the energies of photons emitted by excited atoms. The resulting discrete line spectrum produced by hydrogen is of great historical importance. The part of the spectrum caused by transitions to the $n = 2$ level, from higher levels, falls in the visible region and its distinct pattern was noticed early on (the Balmer series). Its dramatic explanation was a triumph of quantum mechanics. The extension of these quantum mechanical computations to multi-electron atoms turns out to be too complex for an exact treatment. However, phenomenologically motivated approximations are very useful in the study of atomic spectra[6]. Such spectra have been used for a long time to identify elements in a mixture – in particular, when the mixture is somewhat inaccessible for chemical analysis (as in stars!). The study of finer structure in atomic spectra has led to the understanding of other quantum phenomena like electron spin (which produces the so-called “fine structure”), proton spin (which produces the “hyperfine structure”), and also quantum field interactions (which produce the “Lamb shift”).

The functions, R , depend on both n and l and hence they are labelled as R_{nl} . They can be found from the recursion relation of equation 8.46 and substitution in equations 8.43 and 8.45. Some of the R_{nl} are as follows.

$$\begin{aligned} R_{10}(r) &= 2a^{-3/2} \exp(-r/a), \\ R_{20}(r) &= (2a)^{-3/2} (2 - r/a) \exp(-r/(2a)), \\ R_{21}(r) &= 3^{-1/2} (2a)^{-3/2} (r/a) \exp(-r/(2a)). \end{aligned} \quad (8.49)$$

where $a = \hbar^2 / (mk_e e^2)$.

One notices that the energy eigenvalues given by equation 8.48 depend only on the quantum number n and not on the other two quantum numbers m and l . The eigenfunctions u_{nlm} ($= R_{nl} Y_{lm}$), on the other hand, are different for different values of n , l , and m . This means that there must be some degeneracy in eigenstates. For a given l the number of m values is $(2l + 1)$. Also, for a given n , l takes values from 0 up to $n - 1$. This gives the total number of states for a given n to be

$$d = \sum_{l=0}^{n-1} (2l + 1) = n^2. \quad (8.50)$$

Hence, the energy eigenvalue E_n (sometimes called the n -th energy level) is n^2 -fold degenerate. The inclusion of spin (chapter 12) makes this degeneracy $2n^2$ -fold.

8.6 Scattering in three dimensions

In chapter 5, we defined scattering states as the ones for which $E > V$ at infinity in some direction. The same definition can be used in three dimensional problems as well. However, for the sake of simplicity, in this chapter, we shall consider only the scattering states that have $E > V$ in *all* directions at infinity. In chapter 5 it was also shown that, for scattering states, the experimentally significant quantity is the scattering cross section given by equation 5.6 or equation 5.21.

As discussed in section 8.4, realistic potentials are usually not static background potentials. For two-particle systems, the potential is due to the interaction of one particle with another. Such systems can be reduced (in the fashion of section 8.4) to two independent systems – one a free particle and the other a particle in some effective background potential. The free particle part of the energy is seen to be zero in the center of mass frame. Hence, in a scattering problem it is easier to compute the cross section in the center of mass (CM) frame. This makes it necessary to determine a conversion factor between the CM frame cross section and the laboratory frame cross section.

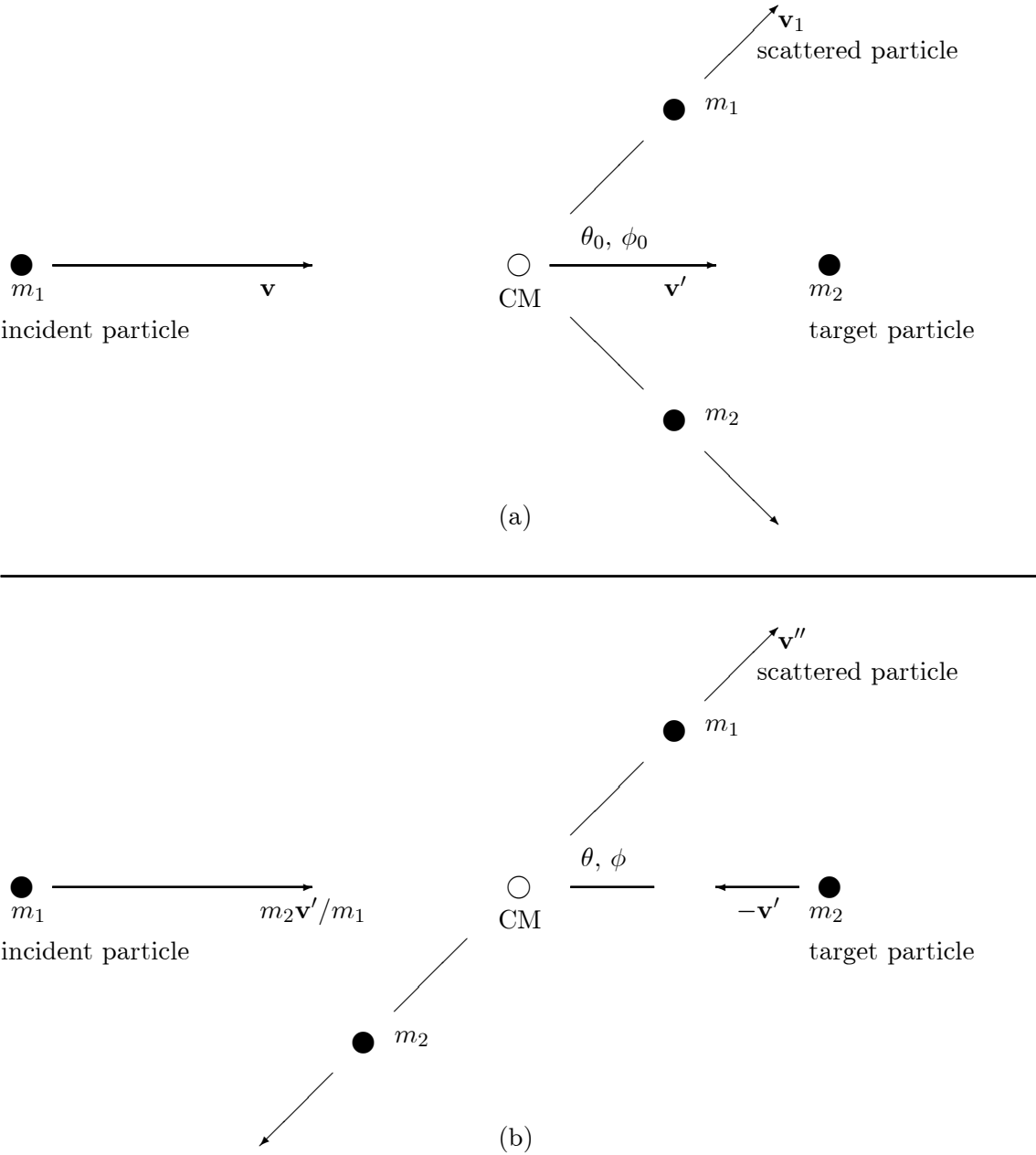


Figure 8.1: Two particle scattering in (a) laboratory frame (b) CM frame.

8.6.1 Center of mass frame vs. laboratory frame

Fig. 8.1a shows a standard laboratory frame setup for a scattering experiment. The target particle of mass m_2 is initially at rest. The incident particle of mass m_1 is moving along the positive z direction (horizontal) at a velocity \mathbf{v} (magnitude v) initially. After the collision, the incident particle moves in a direction given by the spherical polar angles (θ_0, ϕ_0) . Its velocity is \mathbf{v}_1 (magnitude v_1). The center of mass moves at the velocity \mathbf{v}' (magnitude v'). It can be shown that

$$\mathbf{v}' = \frac{m_1 \mathbf{v}}{m_1 + m_2}. \quad (8.51)$$

Fig. 8.1b shows the same experiment in the CM frame. In this frame, before the collision, the target particle moves at a velocity of $-\mathbf{v}'$ and the incident particle at $m_2 \mathbf{v}'/m_1$. After the collision, the incident particle moves at the velocity \mathbf{v}'' (magnitude v'') which has a direction given by the spherical polar angles (θ, ϕ) . This is the scattering direction in the CM frame. For an elastic collision, in the CM frame, it can be shown that the magnitude of the velocity of each particle must be the same before and after collision. Hence,

$$v'' = \frac{m_2 v'}{m_1} = \frac{m_2 v}{m_1 + m_2}. \quad (8.52)$$

From the definition of the CM frame we note that

$$\mathbf{v}_1 = \mathbf{v}' + \mathbf{v}''. \quad (8.53)$$

Writing equation 8.53 in component form gives us the relation between the laboratory frame angles (θ_0, ϕ_0) and the CM frame angles (θ, ϕ) .

$$\begin{aligned} v_1 \cos \theta_0 &= v' + v'' \cos \theta, \\ v_1 \sin \theta_0 &= v'' \sin \theta, \\ \phi_0 &= \phi. \end{aligned} \quad (8.54)$$

Hence,

$$\tan \theta_0 = \frac{\sin \theta}{\gamma + \cos \theta}, \quad (8.55)$$

where

$$\gamma = v'/v''. \quad (8.56)$$

For elastic collisions it can be seen that

$$\gamma = m_1/m_2. \quad (8.57)$$

If the collision is inelastic, it is not necessary that the number of particles be the same before and after collision. Hence, the above analysis, would not work in general. At this point we shall not digress into the analysis of general inelastic collisions. However, the case of the inelastic collision with two product particles of masses m_3 and m_4 ($m_1 + m_2 = m_3 + m_4$),

is not too different from the elastic case. In such a situation one needs to use the following expression for γ in equation 8.55.

$$\gamma = \sqrt{\frac{m_1 m_3 E}{m_2 m_4 (E + Q)}}, \quad (8.58)$$

where Q is the amount of internal energy converted into additional kinetic energy for the product particles. Q is negative for endothermic reactions.

It is clear that the number of particles scattered in the same element of solid angle should not appear different in different frames of reference. Hence, we conclude

$$\sigma_0(\theta_0, \phi_0) \sin \theta_0 d\theta_0 d\phi_0 = \sigma(\theta, \phi) \sin \theta d\theta d\phi. \quad (8.59)$$

Now using equation 8.55 and the last of the set of equations 8.54, it can be shown that

$$\sigma_0(\theta_0, \phi_0) = \frac{(1 + \gamma^2 + 2\gamma \cos \theta)^{3/2}}{|1 + \gamma \cos \theta|} \sigma(\theta, \phi). \quad (8.60)$$

8.6.2 Relation between asymptotic wavefunction and cross section

Now that we know the relation between the laboratory frame and the CM frame measurements of cross section, all computations can be done in the CM frame. Equation 8.37 would be the relevant equation i.e.

$$-\frac{\hbar^2}{2m} \nabla^2 u + V u = E u, \quad (8.61)$$

where $m = m_1 m_2 / (m_1 + m_2)$ and E (the energy in the CM frame) is related to E_0 (the energy in the laboratory frame) as follows.

$$E = \frac{m_2 E_0}{m_1 + m_2}. \quad (8.62)$$

In most scattering experiments, the region of space where a measuring instrument can be placed is at large values of r where V is effectively zero. Hence, one is interested in the solution of equation 8.61 in such ‘‘asymptotic’’ regions where r is very large. If the incident beam is a plane wave of fixed momentum moving in the positive z direction, then the expected form of the asymptotic solution is

$$u(r, \theta, \phi) = A[\exp(ikz) + r^{-1} f(\theta, \phi) \exp(ikr)], \quad (8.63)$$

where

$$z = r \cos \theta, \quad k = p/\hbar = \sqrt{2mE}/\hbar. \quad (8.64)$$

The first term in equation 8.63 is the incident beam which is a momentum eigenstate ($\exp(i\mathbf{p} \cdot \mathbf{r}/\hbar)$) with \mathbf{p} , the momentum, directed along the z direction. The second term is

the scattered wave in the lowest order of r^{-1} . Higher order terms of r^{-1} do not contribute, as the particle current due to them tends to zero at large distances within a fixed solid angle. Only inverse powers of r are expected in the scattered wave as it must die out at large distances.

Normally, the incident beam is collimated to have a width large enough to maintain the plane wave nature near the target particle, but not large enough to produce any significant readings on the detectors that are placed at an angle away from the direct beam. Hence, the detectors observe only the second term in equation 8.63 and the observed $dn/d\omega$ of equation 5.21 is seen to be (using equations 5.16 and 5.20)

$$\frac{dn}{d\omega} = \frac{\hbar k}{m} |A|^2 |f(\theta, \phi)|^2. \quad (8.65)$$

As the incident beam is only the first term in equation 8.63, the N of equation 5.21 is found to be (using equations 5.16 and 5.20)

$$N = \hbar k |A|^2 / m. \quad (8.66)$$

Hence, from equation 5.21, the scattering cross section evaluates to

$$\sigma = |f(\theta, \phi)|^2. \quad (8.67)$$

Thus, $f(\theta, \phi)$ is the quantity to be computed.

8.7 Scattering due to a spherically symmetric potential

For a spherically symmetric potential the solution for $u(r, \theta, \phi)$ is expected to have a cylindrical symmetry about the z axis and hence it must be independent of ϕ . Thus, from the analysis of section 8.2, one may write

$$u(r, \theta) = \sum_{l=0}^{\infty} (2l+1) i^l R_l(r) P_l(\cos \theta), \quad (8.68)$$

where the constant coefficient of each term is chosen for later convenience and $R_l(r)$ is a general solution of equation 8.6 which can also be written as

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dR_l}{dr} \right) + \left[k^2 - \frac{2mV(r)}{\hbar^2} - \frac{l(l+1)}{r^2} \right] R_l = 0. \quad (8.69)$$

where $k^2 = 2mE/\hbar^2$. Due to the scattering nature of the wavefunction, one knows that E and hence k^2 are positive and also V tends to zero at large r . One can usually consider V to be effectively zero if r is greater than some constant a . In this interaction free region

the solution for R_l can be found to be a linear combination of the spherical Bessel and the spherical Neumann functions:

$$R_l(r) = A_l[\cos \delta_l j_l(kr) - \sin \delta_l n_l(kr)]. \quad (8.70)$$

The spherical Bessel functions j_l can be written in terms of the Bessel functions J_l as follows.

$$j_l(kr) = (2kr/\pi)^{-1/2} J_{l+1/2}(kr). \quad (8.71)$$

The spherical Neumann functions n_l can be similarly written as

$$n_l(kr) = (-1)^{l+1} (2kr/\pi)^{-1/2} J_{-l-1/2}(kr). \quad (8.72)$$

The asymptotic forms of j_l and n_l are given by [7]

$$j_l(kr) \rightarrow (kr)^{-1} \cos[kr - (l+1)\pi/2], \quad (8.73)$$

$$n_l(kr) \rightarrow (kr)^{-1} \sin[kr - (l+1)\pi/2]. \quad (8.74)$$

Thus the asymptotic form of equation 8.70 is

$$R_l(r) \rightarrow (kr)^{-1} A_l \sin(kr - l\pi/2 + \delta_l). \quad (8.75)$$

To compare the actual solution to the form chosen in equation 8.63, one needs the following expansion of the incident wave in terms of the spherical Bessel functions.

$$\begin{aligned} \exp(ikz) &= \exp(ikr \cos \theta) \\ &= \sum_{l=0}^{\infty} (2l+1) i^l j_l(kr) P_l(\cos \theta). \end{aligned} \quad (8.76)$$

As the asymptotic form of equation 8.68 must be the same as equation 8.63, one can write (using equations 8.73, 8.75 and 8.76)

$$\begin{aligned} &\sum_{l=0}^{\infty} (2l+1) i^l (kr)^{-1} \sin(kr - l\pi/2) P_l(\cos \theta) + r^{-1} f(\theta) \exp(ikr) \\ &= \sum_{l=0}^{\infty} (2l+1) i^l A_l (kr)^{-1} \sin(kr - l\pi/2 + \delta_l) P_l(\cos \theta). \end{aligned} \quad (8.77)$$

Comparing the coefficients of $\exp(ikr)$ and $\exp(-ikr)$ one then obtains

$$\begin{aligned} &2ikf(\theta) + \sum_{l=0}^{\infty} (2l+1) i^l \exp(-il\pi/2) P_l(\cos \theta) \\ &= \sum_{l=0}^{\infty} (2l+1) i^l A_l \exp(i\delta_l - il\pi/2) P_l(\cos \theta), \end{aligned} \quad (8.78)$$

$$\begin{aligned} &\sum_{l=0}^{\infty} (2l+1) i^l \exp(il\pi/2) P_l(\cos \theta) \\ &= \sum_{l=0}^{\infty} (2l+1) i^l A_l \exp(-i\delta_l + il\pi/2) P_l(\cos \theta). \end{aligned} \quad (8.79)$$

As equation 8.79 is true for all θ , it can be true only if

$$A_l = \exp(i\delta_l). \quad (8.80)$$

Substituting equation 8.80 in equation 8.78 gives

$$f(\theta) = (2ik)^{-1} \sum_{l=0}^{\infty} (2l+1) [\exp(2i\delta_l) - 1] P_l(\cos \theta). \quad (8.81)$$

Hence, the cross section is

$$\begin{aligned} \sigma(\theta) &= |f(\theta)|^2 \\ &= k^{-2} \left| \sum_{l=0}^{\infty} (2l+1) \exp(i\delta_l) \sin \delta_l P_l(\cos \theta) \right|^2. \end{aligned} \quad (8.82)$$

The total cross section, which is defined as

$$\sigma_t = \int \sigma(\theta, \phi) d\omega = \int_0^{2\pi} \int_0^\pi \sigma(\theta, \phi) \sin \theta d\theta d\phi, \quad (8.83)$$

would then be

$$\sigma_t = 2\pi \int_0^\pi \sigma(\theta) \sin \theta d\theta = 4\pi k^{-2} \sum_{l=0}^{\infty} (2l+1) \sin^2 \delta_l. \quad (8.84)$$

The computation of the cross section in equations 8.82 and 8.84 would require the knowledge of the phase shift angles δ_l . These can be found from the continuity of $R_l(r)$ and its derivative at $r = a$. We have already defined a as the distance beyond which the potential $V(r)$ is effectively zero. If the solution for $r < a$ is found by some analytical or numerical method and then the corresponding value of $(1/R_l)(dR_l/dr)$ at $r = a$ is determined to be γ_l , then equation 8.70 and the continuity condition would give

$$\frac{k[j'_l(ka) \cos \delta_l - n'_l(ka) \sin \delta_l]}{j_l(ka) \cos \delta_l - n_l(ka) \sin \delta_l} = \gamma_l, \quad (8.85)$$

where j'_l and n'_l denote the derivatives (with respect to their arguments) of the spherical Bessel and Neumann functions respectively. Solving for δ_l gives

$$\tan \delta_l = \frac{kj'_l(ka) - \gamma_l j_l(ka)}{kn'_l(ka) - \gamma_l n_l(ka)}. \quad (8.86)$$

In summary, it should be noted that the computation of the scattering cross section by this method can become rather tedious if the sum of the series in equation 8.82 converges slowly. This and the fact that the solution of R_l for $r < a$ cannot always be found analytically, tells us that such a method is usually suitable for a computing machine. For specific potentials, it is sometimes possible to use other methods to directly obtain an analytical solution for the cross section. An example is the scattering due to a repulsive $1/r$ type of potential [8].

Problems

1. Derive equation 8.32 from equation 8.30.
2. From equation 8.42, show that, for large s , R has the form $\exp(-s/2)$.
3. Show that if the series expansion for F (equation 8.45) does not terminate, the solution is unacceptable.
4. Obtain the functions in equation 8.49 from the recursion relation of equation 8.46 and the condition of equation 8.47.
5. Derive equation 8.58.
6. Find the phase shifts δ_0 and δ_1 for scattering from a potential given as follows.

$$\begin{aligned}V &= V_0 \quad \text{for } r < a \\V &= 0 \quad \text{for } r > a\end{aligned}$$

where V_0 is a constant. Also determine the conditions under which these phase shifts become infinite. Explain such conditions physically.

$$\begin{aligned}j_0(s) &= s^{-1} \sin s, \\n_0(s) &= -s^{-1} \cos s, \\j_1(s) &= s^{-2} \sin s - s^{-1} \cos s, \\n_1(s) &= -s^{-2} \cos s - s^{-1} \sin s.\end{aligned}$$

Chapter 9

Numerical Techniques in Three Space Dimensions

For the numerical solution of three dimensional problems, one can very often draw from the methods developed for one dimensional problems. However, the increase in the number of space dimensions can just as often require qualitatively new numerical techniques. For example, the three dimensional scattering problem is complicated by angular momentum, which is not defined in one space dimension. In the following, for each of the two cases of bound states and scattering states, we shall first discuss the simpler spherically symmetric problem and then the general problem.

9.1 Bound states (spherically symmetric potentials)

For spherically symmetric bound state problems the angular part of the solution is already known analytically (chapter 8). The radial part of the wavefunction depends on the functional form of $V(r)$, the potential. For arbitrary $V(r)$, a numerical solution of equation 8.6 can be obtained in a manner similar to one dimensional problems. The method will be illustrated here using the hydrogen atom potential, so that results can be compared to analytical results.

For the hydrogen atom, equation 8.6 takes the form of equation 8.42 with the appropriate transformation to the dimensionless independent variable s . Equation 8.42 can be rewritten as

$$\frac{d^2 R}{ds^2} + \frac{2}{s} \frac{dR}{ds} + \left[\frac{\beta}{s} - \frac{1}{4} - \frac{l(l+1)}{s^2} \right] R = 0. \quad (9.1)$$

The initial conditions for this equation will be given at $s = 0$, and the solution need be found only for $s > 0$. Hence, the method of solution will be similar to the one dimensional

symmetric potential case. The potential still has reflection symmetries along each axis. But, as s is proportional to the magnitude of the displacement from the origin, it does not change sign in a reflection. Hence, the solution will not be exclusively odd or even in s . Nonetheless, one of the two independent solutions can be discarded as follows.

It is possible to choose the form of R to be

$$R(s) = s^p G(s), \quad G(0) \neq 0, \quad (9.2)$$

where $G(0)$ is finite. If equation 9.2 is substituted in equation 9.1, one obtains

$$G'' + \frac{2(p+1)}{s} G' + \left[\frac{\beta}{s} - \frac{1}{4} + \frac{p(p+1) - l(l+1)}{s^2} \right] G = 0, \quad (9.3)$$

where $G' = dG/ds$ and $G'' = d^2G/ds^2$. For $s \rightarrow 0$, the above equation can be true only if the coefficient of each negative power of s tends to zero. For the s^{-2} coefficient, this gives

$$[p(p+1) - l(l+1)]G(0) = 0. \quad (9.4)$$

As $G(0) \neq 0$, this gives

$$p = l \quad \text{or} \quad -(l+1). \quad (9.5)$$

Similarly, for the s^{-1} coefficient in equation 9.3 one obtains

$$2(p+1)G'(0) + \beta G(0) = 0. \quad (9.6)$$

The second choice in equation 9.5 would make R go to infinity at the origin. This is not allowed and hence

$$p = l. \quad (9.7)$$

Thus, one of the two independent solutions of R is discarded. Using equation 9.7 in equation 9.6 one obtains

$$G'(0) = -\frac{\beta G(0)}{2(l+1)}. \quad (9.8)$$

Using equation 9.7 in equation 9.3 gives

$$G'' + \frac{2(l+1)}{s} G' + \left[\frac{\beta}{s} - \frac{1}{4} \right] G = 0. \quad (9.9)$$

Using equations 6.2 and 6.8, the finite difference form of equation 9.9 can be seen to be

$$\frac{G_{i-1} - 2G_i + G_{i+1}}{w^2} + \frac{2(l+1)}{s_i} \left[\frac{G_{i+1} - G_{i-1}}{2w} \right] + (\beta/s_i - 1/4)G_i = 0, \quad (9.10)$$

where G_i and s_i are the values of G and s at the i -th point and w is the width of the interval. This can be rewritten as the recursion relation

$$G_{i+1} = \frac{[2 + (1/4 - \beta/s_i)w^2]G_i + [(l+1)w/s_i - 1]G_{i-1}}{1 + (l+1)w/s_i}. \quad (9.11)$$

As the normalization constant can be chosen arbitrarily without changing the eigenstates and the eigenvalues, one can choose

$$G_0 = G(0) = 1, \quad (9.12)$$

for convenience. A finite difference form of equation 9.8 would be

$$\frac{G_1 - G_0}{w} = -\frac{\beta G_0}{2(l+1)}. \quad (9.13)$$

Using equation 9.12, this gives

$$G_1 = 1 - \frac{\beta w}{2(l+1)}. \quad (9.14)$$

With equations 9.12 and 9.14 as initial conditions, one can solve the recursion relation of equation 9.11 for all G_i . In a fashion similar to the one outlined in chapter 6, one can solve for G for a series of values of E , the energy (or equivalently β), to locate the energy eigenvalues. The change in sign of the tail of the wavefunction identifies the location of the eigenvalues.

9.2 Bound states (general potential)

Physically interesting potentials can, in general, be seen to have some symmetry properties. This is due to two reasons. First, a system with symmetry attracts more attention. Second, systems that are built by physicist and engineers are built with symmetry in mind so that theoretical computations are easier. An example is the system of electrons in a uniform magnetic field (see chapter 4). Hence, in solving an arbitrary problem, the first step is to identify as many symmetries as possible. Next, a coordinate system that matches the symmetry needs to be chosen (like the spherical polar coordinates were chosen for the spherically symmetric potential). This allows the separation of the time independent Schrödinger equation to a certain degree. The resulting simplified equations can then be solved either analytically (if possible) or numerically. It is not practical to present numerical methods for all possible symmetry situations. Hence, in this section a method will be outlined for the problem with no identifiable symmetry. Although such a complete lack of symmetry is highly unlikely, this method can give useful hints for the construction of algorithms for systems with some arbitrary symmetry (see problem 5).

For the general numerical algorithm, we need to write the complete Schrödinger equation in its difference equation form. Using equation 6.8 in each rectangular coordinate direction, the Laplacian of u will have the following difference form.

$$\begin{aligned} \nabla_d^2 u(i, j, k) = w^{-2} [& u(i+1, j, k) + u(i-1, j, k) + u(i, j+1, k) \\ & + u(i, j-1, k) + u(i, j, k+1) + u(i, j, k-1) - 6u(i, j, k)], \end{aligned} \quad (9.15)$$

where the discrete indices i , j , and k for the three directions are written in parenthesis to avoid long subindices. These indices are assumed to take all integer values with an arbitrary origin defined by all zero indices. This will require some manipulations if the algorithm is implemented in the C language as C does not allow negative array indices. The difference form of the time independent Schrödinger equation would then be

$$\nabla_d^2 u(i, j, k) + \frac{2m}{\hbar^2} [E - V(i, j, k)] u(i, j, k) = 0. \quad (9.16)$$

Due to the three different indices, this will not reduce to any solvable recursion relation. However, it can be visualized as a set of linear algebraic equations for the unknown $u(i, j, k)$'s. As an illustration of the method of solution [5], consider a cubic region of space in which the solution is to be found. Let i , j , and k each run from 0 to a maximum of i_m , j_m and k_m respectively in this region. To find bound states, the boundaries of the region must be chosen at a large enough distance where u goes to zero. Hence, in the present case $u(i, j, k)$ must be zero if at least one of the three indices is 0 or the corresponding maximum value $-i_m$, j_m or k_m . The values of $u(i, j, k)$ at the interior (non-boundary) points are the unknowns that need to be determined. They are $(i_m - 1)(j_m - 1)(k_m - 1)$ in number. Equation 9.16 provides the same number of equations if written for all interior points. Thus, a solvable system is obtained. In order to write these equations in a matrix form, the interior points must be suitably numbered such that the $u(i, j, k)$'s appear as a one dimensional array. One possibility is illustrated by the following definition of the array U given for $i_m = j_m = k_m = 3$.

$$\begin{aligned} U_1 = u(1, 1, 1), \quad U_2 = u(1, 1, 2), \quad U_3 = u(1, 2, 1), \quad U_4 = u(1, 2, 2), \\ U_5 = u(2, 1, 1), \quad U_6 = u(2, 1, 2), \quad U_7 = u(2, 2, 1), \quad U_8 = u(2, 2, 2). \end{aligned} \quad (9.17)$$

Of course, practical choices for i_m , j_m and k_m must be much larger to achieve reasonable accuracy.

The boundary values being zero, the linear equations for U turn out to be homogeneous. Hence, for a non-zero solution, the related matrix must have a zero determinant. As this matrix contains the energy eigenvalue E , the zero determinant condition will provide the possible energy eigenvalues and the related eigenfunctions. This is a matrix eigenvalue problem for which standard numerical methods are available [5]. In general, for a large enough number of interior points, this will require significant amounts of computer time and memory. However, some efficiency is achieved by making use of the sparse nature of the matrix. It is seen to be tridiagonal with fringes. Besides, one is usually interested in the lower eigenvalues and some time can be saved by focussing only on them. It is to be noted that this method can provide only the lowest $(i_m - 1)(j_m - 1)(k_m - 1)$ eigenvalues and higher the eigenvalue the greater is its inaccuracy. This is due to the rapid oscillations of the higher energy eigenfunctions which require smaller values of w to maintain accuracy.

Another somewhat related method [9] is in some ways more intuitive. Instead of considering only the homogeneous boundary at infinity, one may include an arbitrary point

(normalization point) in the interior as part of the boundary. The value of the function at this point can be chosen to be an arbitrary non-zero constant using the freedom of the normalization constant. Now equation 9.16 can be written for all interior points other than the normalization point. This would result in an *inhomogeneous* matrix equation which can be solved for a given E . The search for the eigenvalues E requires solving the equation for a series of finely spaced values of E and then comparing them, in each case, to the value E_n of E evaluated from equation 9.16 written for the normalization point. Every E that is the same as the corresponding E_n , is an eigenvalue.

It is to be noted that the methods discussed here provide only one quantum number – E . For the spherically symmetric case, n was related to E . However, there were two other quantum numbers (l and m) that were a direct result of the spherical symmetry. The general methods discussed here cannot identify these quantum numbers because, in general, there are no physically meaningful quantities that relate to these numbers. For comparison to symmetric cases, one may consider the following quantities, evaluated at some arbitrarily chosen position, to be the quantum numbers.

$$k_x^2 = \frac{1}{u} \frac{\partial^2 u}{\partial x^2}, \quad k_y^2 = \frac{1}{u} \frac{\partial^2 u}{\partial y^2}, \quad k_z^2 = \frac{1}{u} \frac{\partial^2 u}{\partial z^2}, \quad (9.18)$$

where x , y and z are the usual rectangular coordinates. For a free particle system these quantum numbers are conserved quantities. The Schrödinger equation relates these quantum numbers to the energy eigenvalue E . Hence, one still has a total of three quantum numbers for a three dimensional problem.

9.3 Scattering states (spherically symmetric potentials)

It is to be noted that the computation method for scattering states described in section 8.7 is more suitable for numerical computation rather than analytical computation. The infinite sum in equations 8.82 and 8.84 will, in general, require several terms in the series to be computed to get significant accuracy. Also the computation of the γ_l of equation 8.86 would require the solution of a differential equation that is seldom expected to have analytical solutions for arbitrary potentials. Hence, here we shall discuss the numerical aspects of the same method.

As an example, we shall consider the scattering of an electron due to a neutral atom of atomic number Z . The scattering potential of such a system can very often be approximated by

$$V(r) = -\frac{Ze^2}{4\pi\epsilon_0} \frac{\exp(-r/B)}{r}, \quad (9.19)$$

where e is the magnitude of the electron charge, ϵ_0 the permittivity of free space, and B a measure of the range of the potential which is the “radius” of the atom. The first step

would be to solve equation 8.69 using this potential. The differential equation to be solved would be

$$\frac{d^2 R_l}{ds^2} + \frac{2}{s} \frac{dR_l}{ds} + \left[1 + A \frac{\exp(-s/b)}{s} - \frac{l(l+1)}{s^2} \right] R_l = 0, \quad (9.20)$$

where

$$s = kr, \quad k^2 = \frac{2mE}{\hbar^2}, \quad A = \frac{2mZe^2}{4\pi\epsilon_0\hbar^2k}, \quad b = kB. \quad (9.21)$$

As in section 9.1, we can write

$$R_l(s) = s^p G(s), \quad G(0) \neq 0. \quad (9.22)$$

The equation for G would then be

$$G'' + \frac{2(p+1)}{s} G' + \left[1 + A \frac{\exp(-s/b)}{s} + \frac{p(p+1) - l(l+1)}{s^2} \right] G = 0. \quad (9.23)$$

Using arguments similar to section 9.1, one finds

$$p = l, \quad (9.24)$$

$$G'(0) = -\frac{AG(0)}{2(l+1)}. \quad (9.25)$$

Inserting equation 9.24 in equation 9.23 gives

$$G'' + \frac{2(l+1)}{s} G' + \left[1 + A \frac{\exp(-s/b)}{s} \right] G = 0. \quad (9.26)$$

From equations 6.2 and 6.8 one obtains the finite difference form of equation 9.26 to be

$$\frac{G_{i-1} - 2G_i + G_{i+1}}{w^2} + \frac{2(l+1)}{s_i} \left[\frac{G_{i+1} - G_{i-1}}{2w} \right] + \left[1 + A \frac{\exp(-s_i/b)}{s_i} \right] G_i = 0. \quad (9.27)$$

where G_i and s_i are the values of G and s at the i -th point and w is the width of the interval. This leads to the recursion relation

$$G_{i+1} = \frac{[2 - (1 + A \exp(-s_i/b)/s_i)w^2]G_i + [(l+1)w/s_i - 1]G_{i-1}}{1 + (l+1)w/s_i}. \quad (9.28)$$

Using the freedom of choice of the normalization constant, one chooses

$$G_0 = G(0) = 1. \quad (9.29)$$

The finite difference form of equation 9.25 is

$$\frac{G_1 - G_0}{w} = -\frac{AG_0}{2(l+1)}. \quad (9.30)$$

From equations 9.29 and 9.30 one finds

$$G_1 = 1 - \frac{wA}{2(l+1)}. \quad (9.31)$$

The initial conditions in equations 9.29 and 9.31 allow us to compute G_i at all points using equation 9.28. One needs to know this solution up to a point ($r = a$ or $s = ka$) beyond which the potential is effectively zero. The choice of a depends on the desired accuracy. If this boundary point is reached at the n -th point in the numerical computation then $s_n = ka$. Now, the γ_l of equation 8.86 is defined as

$$\gamma_l = \frac{1}{R_l} \frac{dR_l}{dr} \Big|_{r=a} = \frac{k}{R_l} \frac{dR_l}{ds} \Big|_{s=ka} = \frac{l}{a} + \frac{kG'}{G} \Big|_{s=ka}. \quad (9.32)$$

Hence, the difference form of γ_l would be

$$\gamma_l = \frac{l}{a} + \frac{k}{w} \left(1 - \frac{G_{n-1}}{G_n} \right). \quad (9.33)$$

To find the phase shifts from equation 8.86 one needs equation 9.33, the spherical Bessel and Neumann functions and their derivatives. Standard series solutions for the Bessel and Neumann functions can be used for their numerical computation. Once the δ_l are known, equations 8.82 and 8.84 can be used to find the differential and the total cross sections.

9.4 Scattering states (general potential)

For arbitrary potentials, scattering cross sections are much more difficult to compute even numerically. Some approximate methods (see chapter 11) can handle such general potentials as long as they are small compared to the energy of the incident particle. The reach of such methods, which were originally designed for analytical computations, can be significantly increased by the use of a computing machine.

On the other hand, a non-spherically symmetric scattering potential is very unlikely in realistic situations. A scattering experiment with such a potential would require both the target and scattered particles to approach each other at a fixed orientation for every collision. Such an experimental setup is difficult to come by. So, even for non-spherically symmetric potentials, the actual scattering data will show a spherically symmetric character as the random orientations of the particles in each collision will create the appearance of an average potential over all angles which is spherically symmetric. In some realistic experiments one may produce spin aligned target and scattered particles. In such a situation the spherical symmetry is truly lost. A proper treatment of these systems requires quantum field theory which is beyond the scope of this book.

So, even though it is possible to come up with a numerical algorithm for scattering computations for non-spherically symmetric potentials, we shall not do it here.

Problems

1. A spherically symmetric linear potential is given by

$$V(r) = Ar$$

where A is a positive real number constant. For this potential do the following.

- (a) Find the recursion relation for the numerical solution of the Schrödinger equation and also the necessary initial conditions.
 - (b) Develop a computer algorithm to determine the eigenvalues and eigenfunctions of such a system.
2. A spherically symmetric quadratic potential (3-dimensional isotropic harmonic oscillator) is given by

$$V(r) = Ar^2$$

where A is a positive real number constant. For this potential, repeat the steps as in problem 1 and then compare the numerical results with analytical results (an extension of the one dimensional harmonic oscillator problem to 3 dimensions).

3. Phenomenological models of mesons often have the potential between the quark and the antiquark components to be a combination of a linear potential (as in problem 1) and a Coulomb type of potential as follows.

$$V(r) = Ar - B/r$$

where $A > 0$ and $B > 0$. For this potential, repeat the steps as in problem 1.

4. Some other phenomenological models of mesons replace the linear part of the potential in problem 3 by a quadratic potential (as in problem 2). Hence, the potential is given by

$$V(r) = Ar^2 - B/r$$

For this potential, repeat the steps as in problem 1.

5. A superlattice is fabricated by depositing several thin layers of two semiconductor materials alternated in a periodic fashion. If the layers are thin and uniform enough, quantum effects allow such structures to have electronic and optical properties that are not found in the constituent materials. An electron trapped by a positive ion impurity in a superlattice experiences the following potential.

$$V(r) = -\frac{e^2}{4\pi\epsilon r} + W(z)$$

where e is the electron charge, ϵ the permittivity of the material and $r = \sqrt{x^2 + y^2 + z^2}$. x , y and z are the usual rectangular coordinates with the z -axis oriented perpendicular

to the layers of the superlattice. The function $W(z)$ can sometimes be approximated as

$$W(z) = \begin{cases} W_0 & \text{for } |z| > a \\ 0 & \text{for } |z| < a \end{cases}$$

where W_0 and a are positive constants. For this cylindrically symmetric potential, develop a computer algorithm to compute the eigenvalues and eigenfunctions of energy [9].

6. Develop a computer algorithm for the computation of the scattering cross section for the following potential.

$$V(r) = A \exp(-ar^2)$$

where A and a are positive real number constants.

7. For the scattering of neutrons off an atomic nucleus the effective potential can be approximated by the Woods-Saxon potential which is given as follows.

$$V(r) = \frac{-V_0}{1 + \exp[(r - R)/a]}$$

where V_0 , R and a are positive real number constants. Develop a computer algorithm to compute the scattering cross section for this potential.

Chapter 10

Approximation Methods (Bound States)

Given a specific problem in quantum mechanics, one needs to find the quickest method of solving it. An analytical solution is usually the most desirable. However, as we have seen in the previous chapters, such solutions are not always possible. In such a situation, numerical methods can often be used successfully. But we have seen in chapter 9 that numerical methods can sometimes be very time consuming. Hence, one needs to look for alternative methods that would be quicker. Sometimes approximation methods are very handy. Some readers might find this last statement somewhat perplexing as numerical methods are usually considered to be approximation methods. But here we shall use the terms “numerical” and “approximate” with different meanings. A numerical method can produce results of indefinitely high accuracy provided enough computer time is spent. An approximate method has its accuracy (and sometimes even its validity) limited by the characteristics of the specific problem being solved. For example, some series expansions converge only under certain conditions.

Here, we shall discuss two of the most popular approximation methods – the perturbation method and the variational method ¹.

¹The so-called WKB method [8] will not be discussed even though, historically, it has been very popular. This is because, at present times, any problem that can be solved by the WKB method, can be solved with far greater accuracy and speed by numerical methods. Besides, the original mathematical justification for the WKB method was rather weak and later mathematical treatments that better justify it are too lengthy.

10.1 Perturbation method (nondegenerate states)

Let H_t , the hamiltonian of a system, be written as

$$H_t = H + H', \quad (10.1)$$

such that the eigenstates and eigenvalues of H are already known and written as $|n\rangle$ and E_n respectively. So,

$$H|n\rangle = E_n|n\rangle, \quad (10.2)$$

where n is the label that identifies a specific eigenstate. The corresponding eigenstates and eigenvalues of H_t will be called $|tn\rangle$ and E_{tn} respectively such that

$$H_t|tn\rangle = E_{tn}|tn\rangle. \quad (10.3)$$

The additional part, H' , is defined to be small if the differences $(E_{tn} - E_n)$ are suitably small compared to E_n . The present method is valid only if H' is small and hence it will be assumed to be so in the following. The correction terms necessary to obtain the solution of equation 10.3 from that of equation 10.2 can be written as a series of increasing order of smallness. To keep track of the order, one uses a parameter λ that is later set to be equal to 1. In terms of λ one can write

$$H_t = H + \lambda H', \quad (10.4)$$

$$|tn\rangle = \sum_{s=0}^{\infty} \lambda^s |ns\rangle, \quad (10.5)$$

$$E_{tn} = \sum_{s=0}^{\infty} \lambda^s E_{ns}, \quad (10.6)$$

where s is the order of the correction terms $|ns\rangle$ and E_{ns} for $|tn\rangle$ and E_{tn} respectively. Substituting this in equation 10.3 and equating terms of the same order on each side, one obtains

$$(H - E_{n0})|n0\rangle = 0, \quad (10.7)$$

$$(H - E_{n0})|n1\rangle = (E_{n1} - H')|n0\rangle, \quad (10.8)$$

$$(H - E_{n0})|n2\rangle = (E_{n1} - H')|n1\rangle + E_{n2}|n0\rangle, \quad (10.9)$$

$$(H - E_{n0})|n3\rangle = (E_{n1} - H')|n2\rangle + E_{n2}|n1\rangle + E_{n3}|n0\rangle, \quad (10.10)$$

and so on. The first of the above set of equations shows that $|n0\rangle$ is an eigenstate of H . As in the present case we are discussing only nondegenerate states, one can unambiguously identify

$$|n0\rangle = |n\rangle, \quad E_{n0} = E_n. \quad (10.11)$$

Using equations 10.2 and 10.11, one can see that multiplying each of the equations 10.7 through 10.10 from the left by $\langle n|$ makes their left hand sides vanish. This leaves one with the following equations.

$$E_{n1} = \langle n|H'|n\rangle, \quad (10.12)$$

$$E_{n2} = \langle n|H'|n1\rangle - E_{n1}\langle n|n1\rangle, \quad (10.13)$$

$$E_{n3} = \langle n|H'|n2\rangle - E_{n1}\langle n|n2\rangle - E_{n2}\langle n|n1\rangle, \quad (10.14)$$

and so on. Here it is assumed that the eigenstates $|n\rangle$ are normalized. As the eigenstates of H form a complete set, the correction terms $|ns\rangle$ in each order s can be written as a linear combination of the $|n\rangle$'s as follows.

$$|ns\rangle = \sum_i a_{nis}|i\rangle, \quad \text{for } s = 1, 2, 3, \dots \quad (10.15)$$

Substituting these into the equations 10.7 through 10.10 and using equations 10.12 through 10.14, it is possible to find all the coefficients a_{nis} except the a_{nns} . This is because all terms containing a_{nns} vanish identically. This leads to the conclusion that the a_{nns} can be chosen arbitrarily. The simplest choice, of course, would be zero. This choice can be written in two equivalent forms:

$$a_{nns} = 0, \quad \langle n|ns\rangle = 0. \quad (10.16)$$

Thus the equations 10.12 through 10.14 would simplify to

$$E_{ns} = \langle n|H'|n, s-1\rangle, \quad \text{for } s = 1, 2, 3, \dots \quad (10.17)$$

Hence, the first order correction to the energy eigenvalues would be

$$E_{n1} = \langle n|H'|n0\rangle = \langle n|H'|n\rangle. \quad (10.18)$$

The first order correction to energy eigenstates is given in the form of equation 10.15 to be

$$|n1\rangle = \sum_i a_{ni1}|i\rangle. \quad (10.19)$$

Substituting this in equation 10.8 and using equations 10.2 and 10.11, gives

$$\sum_i (E_i - E_n) a_{ni1}|i\rangle = (E_{n1} - H')|n\rangle. \quad (10.20)$$

Multiplying this from the left by $\langle j|$ and using the orthonormality of the $|i\rangle$ states, one obtains:

$$a_{nj1} = \frac{\langle j|H'|n\rangle}{E_n - E_j}, \quad \text{for } j \neq n. \quad (10.21)$$

It should be noted that the value for a_{nn1} cannot be determined from equation 10.20. However, it has already been chosen to be zero in equation 10.16. This completes the computation of first order correction terms.

The second order correction to energy eigenvalues is seen from equation 10.17 to be

$$E_{n2} = \langle n|H'|n1\rangle. \quad (10.22)$$

Using equations 10.19 and 10.21 one then obtains:

$$E_{n2} = \sum_{i \neq n} \frac{\langle n|H'|i\rangle \langle i|H'|n\rangle}{E_n - E_i}. \quad (10.23)$$

As H' is hermitian this gives

$$E_{n2} = \sum_{i \neq n} \frac{|\langle n|H'|i\rangle|^2}{E_n - E_i}. \quad (10.24)$$

From equation 10.15, the second order correction to the eigenstate is seen to be

$$|n2\rangle = \sum_i a_{ni2}|i\rangle. \quad (10.25)$$

Substituting this and equation 10.19 in equation 10.9 and then using equations 10.2 and 10.11, one obtains

$$\sum_i (E_i - E_n) a_{ni2}|i\rangle = (E_{n1} - H') \sum_i a_{ni1}|i\rangle + E_{n2}|n\rangle. \quad (10.26)$$

Multiplying this from the left by $\langle j|$ ($j \neq n$) and using orthonormality gives

$$a_{nj2}(E_j - E_n) = a_{nj1}E_{n1} - \sum_i a_{ni1} \langle j|H'|i\rangle. \quad (10.27)$$

Using equation 10.21 this gives

$$a_{nj2} = \sum_{i \neq n} \frac{\langle j|H'|i\rangle \langle i|H'|n\rangle}{(E_n - E_j)(E_n - E_i)} - \frac{\langle j|H'|n\rangle \langle n|H'|n\rangle}{(E_n - E_j)^2}. \quad (10.28)$$

Once again a_{nn2} cannot be found from equation 10.27. It was assumed to be zero in equation 10.16. This completes the computation of second order correction terms. Higher order corrections can be computed in a similar fashion. One could summarize the results upto second order as follows:

$$E_{tn} = E_n + \langle n|H'|n\rangle + \sum_{i \neq n} \frac{|\langle n|H'|i\rangle|^2}{E_n - E_i} + \dots \quad (10.29)$$

$$\begin{aligned} |tn\rangle = & |n\rangle + \sum_{i \neq n} \left[\frac{\langle i|H'|n\rangle}{E_n - E_i} \left(1 - \frac{\langle n|H'|n\rangle}{E_n - E_i} \right) \right. \\ & \left. + \sum_{j \neq n} \frac{\langle i|H'|j\rangle \langle j|H'|n\rangle}{(E_n - E_i)(E_n - E_j)} \right] |i\rangle + \dots \end{aligned} \quad (10.30)$$

It is to be noted that for higher order terms the computations get significantly more involved and hence, writing a computer algorithm for the general s -th order correction term might be useful. But we shall not do it here.

As an application of the above formalism one may compute the first order correction to the solution of the harmonic oscillator problem (see chapter 4) when a quartic perturbation is added. Then

$$H_t = \frac{P^2}{2m} + \frac{kX^2}{2} + \frac{KX^4}{6}, \quad (10.31)$$

where K must be small enough for the perturbation analysis to work. As the solution to the harmonic oscillator problem is already known, we identify

$$H = \frac{P^2}{2m} + \frac{kX^2}{2}, \quad H' = \frac{KX^4}{6}. \quad (10.32)$$

In terms of the raising and lowering operators one may write

$$H' = A(a - a^\dagger)^4, \quad A = \frac{K\hbar^2}{24m^2\omega^2}. \quad (10.33)$$

The expanded form of H' would then be

$$H' = A(a^4 - a^\dagger a^3 - aa^\dagger a^2 - a^2 a^\dagger a - a^3 a^\dagger + a^{\dagger 2} a^2 + a^\dagger aa^\dagger a + a^\dagger a^2 a^\dagger + aa^\dagger a^2 + aa^\dagger aa^\dagger + a^2 a^{\dagger 2} - a^{\dagger 3} a - a^{\dagger 2} aa^\dagger - a^\dagger aa^{\dagger 2} - aa^{\dagger 3} + a^{\dagger 4}). \quad (10.34)$$

Now from equations 4.80, 4.82, 10.18 and 10.34 one computes the first order correction to the ground state to be

$$E_{01} = 3A. \quad (10.35)$$

For the perturbation analysis to be valid one must have $3A \ll \hbar\omega/2$. Then from equation 10.21 one finds

$$a_{021} = \frac{3\sqrt{2}A}{\hbar\omega}, \quad a_{041} = -\frac{\sqrt{3}A}{\sqrt{2}\hbar\omega}, \quad (10.36)$$

and all other a_{0j1} vanish. Hence, from equation 10.19 we get the first order correction to the ground eigenstate to be

$$|01\rangle = \frac{A}{\sqrt{2}\hbar\omega}(6|2\rangle - \sqrt{3}|4\rangle). \quad (10.37)$$

These computations could also have been done in the position representation. However, it would require the computation of several integrals.

It should be noted that, however small K might be, for some higher excited states the correction terms will become too large for a perturbative computation to be valid. This is because a quartic potential rises faster than a quadratic potential as the position coordinate increases.

10.2 Degenerate state perturbation analysis

If the eigenstates of H , the unperturbed hamiltonian, have some degeneracy then the identification of $|n0\rangle$ in equation 10.11 need not be unique. The general method of analysis in such a situation can be understood by considering the particular case of a doubly degenerate energy level, say E_n , that corresponds to the two states $|n\rangle$ and $|m\rangle$. Hence, any linear combination of $|n\rangle$ and $|m\rangle$ will also be an eigenstate corresponding to E_n . As a result of the perturbation, these states may not stay degenerate. The two resulting nondegenerate eigenstates of H_t must tend to eigenstates of H corresponding to E_n as λ tends to zero. However, in general, these limiting eigenstates of H are linear combinations of $|n\rangle$ and $|m\rangle$ and are to be identified as $|n0\rangle$. Hence, one writes

$$|n0\rangle = c_n|n\rangle + c_m|m\rangle, \quad (10.38)$$

where the set of constants $\{c_n, c_m\}$ will be different for the limits of the two different eigenstates of H_t . c_n and c_m are yet to be determined. The choice of equation 10.16 must now be extended to give

$$\langle n|ns\rangle = 0, \quad \langle m|ns\rangle = 0, \quad \text{for } s = 1, 2, 3, \dots \quad (10.39)$$

Equation 10.8 can now be multiplied from the left by each of $\langle m|$ and $\langle n|$ to get two equations. These equations are (using equations 10.2, 10.11 and 10.38)

$$(\langle m|H'|m\rangle - E_{n1})c_m + \langle m|H'|n\rangle c_n = 0, \quad (10.40)$$

$$\langle n|H'|m\rangle c_m + (\langle n|H'|n\rangle - E_{n1})c_n = 0. \quad (10.41)$$

This is a set of linear homogeneous equations for c_m and c_n . A nonzero solution of this set is possible only if the determinant of the corresponding matrix vanishes. This gives

$$(\langle m|H'|m\rangle - E_{n1})(\langle n|H'|n\rangle - E_{n1}) - \langle m|H'|n\rangle\langle n|H'|m\rangle = 0. \quad (10.42)$$

The solution of this gives the two different correction terms for the two corresponding states of H_t :

$$E_{n1} = \frac{1}{2}(\langle m|H'|m\rangle + \langle n|H'|n\rangle) \pm \left[\frac{1}{4}(\langle m|H'|m\rangle - \langle n|H'|n\rangle)^2 + |\langle m|H'|n\rangle|^2 \right]^{1/2}. \quad (10.43)$$

The two correction terms are seen to be identical if and only if

$$\langle m|H'|m\rangle = \langle n|H'|n\rangle, \quad \text{and} \quad \langle m|H'|n\rangle = 0. \quad (10.44)$$

If the two correction terms are different, the degeneracy is said to have been “lifted”. The values for c_m and c_n can now be found for each value of E_{n1} by using equation 10.40 (or

equation 10.41) and a suitable normalization. This would give the two possible choices of $|n0\rangle$. The correction $|n1\rangle$ to each of these can be found by the same method as for the nondegenerate case using the two conditions in equation 10.39. Once the degeneracy is lifted, higher order corrections can be found as for the nondegenerate case.

If the degeneracy is not lifted in first order computations, a similar method can be employed in second order computations.

As an example of the method developed in this section, we shall study the effect of a relatively weak uniform magnetic field on the bound states of a spherically symmetric system (in particular the hydrogen atom). This is known as the Zeeman effect. The vector potential for a uniform magnetic field \mathbf{B} can be written as

$$\mathbf{A} = \frac{1}{2}(\mathbf{B} \times \mathbf{R}), \quad (10.45)$$

where \mathbf{R} is the position vector operator. The hamiltonian, in the presence of such a vector potential in addition to the scalar potential ϕ , would become [2]

$$H_t = \frac{(\mathbf{P} + e\mathbf{A})^2}{2m} - e\phi, \quad (10.46)$$

where m is the reduced mass, \mathbf{P} the canonical momentum and e the magnitude of the electron charge. As \mathbf{A} is the perturbing term, one can separate H_t in the form of equation 10.1 with

$$H = \frac{P^2}{2m} - e\phi, \quad (10.47)$$

$$H' = \frac{e}{2m}(\mathbf{P} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{P} + eA^2). \quad (10.48)$$

For the present computations, the A^2 term can be ignored as the magnitude of \mathbf{B} is assumed to be relatively small. Using equation 10.45 and the commutators of the components of \mathbf{P} and \mathbf{R} , one can show that $\mathbf{P} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{P}$. Then we have

$$H' = \frac{e}{m}\mathbf{A} \cdot \mathbf{P} = \frac{e}{2m}\mathbf{B} \times \mathbf{R} \cdot \mathbf{P} = \frac{e}{2m}\mathbf{B} \cdot \mathbf{R} \times \mathbf{P} = \frac{e}{2m}\mathbf{B} \cdot \mathbf{L}, \quad (10.49)$$

where \mathbf{L} is the angular momentum. Now, let us compute the first order corrections to the first excited states ($n = 2$) of the hydrogen atom. The 4 degenerate states can be written in the form $|l, m\rangle$ for the various possible angular momentum quantum numbers:

$$|0, 0\rangle, |1, -1\rangle, |1, 0\rangle, |1, 1\rangle.$$

Without loss of generality, if the magnetic field is considered to be in the x -direction, then

$$H' = \frac{eBL_x}{2m}. \quad (10.50)$$

Using equations 7.39 and 7.40, this would give

$$H' = \frac{eB}{4m}(L_+ + L_-). \quad (10.51)$$

Now, using equations 7.65, 7.66 and 10.51, the nonzero matrix elements of H' for the 4 degenerate states with $n = 2$ are found to be

$$\langle 1, -1|H'|1, 0\rangle = \langle 1, 0|H'|1, 1\rangle = \frac{\hbar eB}{2\sqrt{2}m}. \quad (10.52)$$

The method described earlier for a 2-fold degenerate level can be easily generalized for the 4-fold degenerate level in this example. Then, using equation 10.52, the 4 first order correction terms can be found to be

$$0, 0, +\frac{\hbar eB}{2m}, -\frac{\hbar eB}{2m}.$$

Hence, we notice that two of the original degenerate states are still degenerate. If the magnetic field were chosen to be in the z-direction, the computations would have been somewhat simpler. The computation of the corrections to the eigenstates will be left as an exercise (problem 2).

10.3 Time dependent perturbation analysis

All hamiltonians considered in this text have, until now, been time independent. Hence, energy has been conserved and energy eigenvalues have been meaningful measured quantities. However, the only direct way of measuring the energy of something like a hydrogen atom would be to measure its mass and use the relativistic mass energy equivalence. This being a particularly inaccurate method of measurement, one needs to find indirect methods of measuring the energy. The most common method would be to allow the system (for example, the hydrogen atom) to transfer from one energy level to another and release the difference in energy in the form of another particle (for example, a photon). If the released particle is a photon, its frequency ν can be measured spectroscopically and then the energy computed to be $2\pi\hbar\nu$. This is a result of the original Planck hypothesis. It will also become evident from our later discussion of the relativistic Schrödinger equation for the photon. Thus, by this method one can find the differences between energy levels. The energy levels can very often be reconstructed from these differences.

However, to use the above method, it would become necessary to transfer the system from one energy level to another. For a conservative system this is, by definition, not possible. Hence, a time dependent (thus nonconservative) perturbation, $H'(t)$, must be added to the system to achieve the transfer between levels. This $H'(t)$ must be small in order not to disturb the energy levels of the original system too much. If $H'(t)$ were too

large, energy would not be conserved even approximately and hence, energy measurements would be of no use. Very often $H'(t)$ is introduced through the sinusoidal variations of electromagnetic radiation. Hence, we shall choose

$$H'(t) = H_0 \sin(\omega t), \quad (10.53)$$

where H_0 is independent of time. The total hamiltonian would be

$$H_t = H + H'(t), \quad (10.54)$$

and the development of an arbitrary state $|s\rangle$ would be given by the Schrödinger equation to be

$$i\hbar \frac{\partial |s\rangle}{\partial t} = H_t |s\rangle. \quad (10.55)$$

As the eigenstates $|i\rangle$, of the original hamiltonian H , form a complete set, it must be possible to expand $|s\rangle$ as follows.

$$|s\rangle = \sum_i a_i(t) \exp(-iE_i t/\hbar) |i\rangle, \quad (10.56)$$

where E_i is the eigenvalue of H corresponding to $|i\rangle$. The exponential time dependence is written separately from the $a_i(t)$ as it is expected that for small $H'(t)$, $a_i(t)$ will vary slowly with time. This is because in the absence of $H'(t)$, $a_i(t)$ would be constant. Now, substituting equation 10.56 in equation 10.55 and then multiplying from the left by $\langle j|$ would give

$$i\hbar \exp(-iE_j t/\hbar) \frac{da_j}{dt} = \sum_i a_i \exp(-iE_i t/\hbar) \langle j|H'|i\rangle. \quad (10.57)$$

If we choose

$$\omega_{ji} = (E_j - E_i)/\hbar, \quad (10.58)$$

then equation 10.57 could be written as

$$\frac{da_j}{dt} = (i\hbar)^{-1} \sum_i a_i \exp(i\omega_{ji} t) \langle j|H'|i\rangle. \quad (10.59)$$

Once again, we can expand the solution in a perturbation series using the parameter λ to identify the order of smallness of a term. Hence, we write

$$a_i = \sum_{s=0}^{\infty} \lambda^s a_{is}. \quad (10.60)$$

As before, the H' in equation 10.59 must also be multiplied by a λ to keep track of orders. Then, substituting equation 10.60 in equation 10.59 and equating terms of the same order would give

$$\frac{da_{j0}}{dt} = 0, \quad (10.61)$$

$$\frac{da_{j(s+1)}}{dt} = (i\hbar)^{-1} \sum_i a_{is} \exp(i\omega_{ji} t) \langle j|H'|i\rangle. \quad (10.62)$$

Equation 10.61 reasserts the fact that a_j is independent of time in the absence of the perturbation H' . In a typical physical situation, the system is initially ($t = 0$) in an eigenstate of H . At $t = 0$ the perturbation H' is turned on and at a later time T it is turned off. If that initial eigenstate is $|n\rangle$, then for $t < 0$, $a_n = 1$ and all other a_j 's are zero. As the a_{j0} do not change with time, this means

$$a_{n0} = 1, \quad (10.63)$$

$$a_{j0} = 0 \quad \text{for } j \neq n. \quad (10.64)$$

Hence, from equation 10.62 the first order corrections for $t > T$ are

$$a_{j1} = (i\hbar)^{-1} \int_0^T \langle j|H'(t)|n\rangle \exp(i\omega_{jn}t) dt. \quad (10.65)$$

Then from equation 10.53 we obtain

$$a_{j1} = \frac{\langle j|H_0|n\rangle}{2i\hbar} \left[\frac{\exp[i(\omega_{jn} - \omega)T] - 1}{\omega_{jn} - \omega} - \frac{\exp[i(\omega_{jn} + \omega)T] - 1}{\omega_{jn} + \omega} \right]. \quad (10.66)$$

Hence, a_{j1} can be seen to have maximum values at $\omega_{jn} = \pm\omega$ that is

$$E_j = E_n - \hbar\omega, \quad \text{or} \quad E_j = E_n + \hbar\omega. \quad (10.67)$$

The first case is interpreted as a high probability for the system to emit the Planck energy $\hbar\omega$ and transfer to a suitable lower energy state. The second case shows a similar high probability for the system to absorb the energy $\hbar\omega$ and transfer to a suitable higher energy state. Of course, in each case, the corresponding lower or higher energy must be an eigenvalue of H . The first case is often known as stimulated emission where external electromagnetic radiation of frequency ω stimulates the system to radiate energy in the form of photons of the same frequency. Emission can also take place in the absence of external radiation (spontaneous emission). However, such a phenomenon can be explained only by a quantum field theory which is beyond the scope of this book. The second case is that of absorption of energy from the external radiation. Both the emission and absorption of specific frequencies of electromagnetic radiation can be observed spectroscopically. Thus it is often possible to reconstruct the energy eigenvalues of H from spectroscopic data.

It is to be noted that the peaks of a_{j1} , as a function of ω , become narrower and higher if T becomes larger. This is understood by considering the perturbation H' to be a sinusoidal wave within a rectangular envelope ranging from $t = 0$ to $t = T$. The Fourier transform of such a time function can be seen to have a larger range of frequencies for smaller T . This means that, for smaller T , the perturbing external radiation provides photons of a larger range of frequencies and hence, allows transitions of the system to a larger range of energy states. This prompts one to write what is known as the time-energy uncertainty relation:

$$\Delta E \Delta t \gtrsim \hbar, \quad (10.68)$$

where ΔE is a measure of the range of energies around E_j that it is possible to transfer to and Δt is a measure of the duration of time for which the perturbation is on. The symbol ' \gtrsim ' means 'approximately greater than or equal to'. In the present case $\Delta t = T$. This uncertainty relation can be written mathematically more precisely by defining ΔE and Δt more precisely. But we shall not do it here.

Another item to be noted is that a_{j1} also depends on the time independent factor $\langle j|H_0|n\rangle$. An extreme example would be that of $\langle j|H_0|n\rangle = 0$. In such a situation the probability of transition from $|n\rangle$ to $|j\rangle$ will be zero even if $\omega_{jn} = \pm\omega$. Such a condition gives the so called *selection rules*.

In a process like the ionization of an atom, the final state $|j\rangle$ is not a bound state and hence E_j is not a discrete energy level. There is in fact a continuum of energy eigenvalues E_j . This makes it more practical to compute a transition probability (or ionization probability) per unit time for the whole range of energies E_j . Of course, in such a situation $E_j > E_n$ and hence, the part of a_{j1} in equation 10.66 that describes the emission process is small enough to be ignored. This gives the transition probability to be

$$|a_{j1}|^2 = \frac{|\langle j|H_0|n\rangle|^2 \sin^2[(\omega_{jn} - \omega)T/2]}{\hbar^2(\omega_{jn} - \omega)^2}. \quad (10.69)$$

The total transition probability to all states with energies around $E_j = E_n + \hbar\omega$ (the peak of this probability function) can be obtained by integrating $|a_{j1}|^2$ over such states:

$$P = \int |a_{j1}|^2 \rho(j) dE_j, \quad (10.70)$$

where $\rho(j)$ is defined as the density of states at the energy eigenvalue E_j . This total transition probability is expected to increase with the duration T of the perturbation. Hence, one defines a convenient measurable quantity called the transition probability per unit time:

$$w = P/T = T^{-1} \int |a_{j1}|^2 \rho(j) dE_j. \quad (10.71)$$

For a large enough T , $|a_{j1}|^2$ would peak so sharply that $\rho(j)$ and $\langle j|H_0|n\rangle$ would be effectively independent of E_j in the range where $|a_{j1}|^2$ is not negligible. Hence, in the integral of equation 10.71, one can pull the term $\rho(j)|\langle j|H_0|n\rangle|^2$ outside the integral. This gives

$$w = \frac{\pi}{2\hbar} \rho(j) |\langle j|H_0|n\rangle|^2. \quad (10.72)$$

The above integral was done by substituting $x = (\omega_{jn} - \omega)T/2$ and using the following definite integral.

$$\int_{-\infty}^{\infty} x^{-2} \sin^2 x dx = \pi. \quad (10.73)$$

The infinite limits are justified by the fact that the integrand is sharply peaked for large T . Equation 10.72 is sometimes called Fermi's golden rule.

10.4 The variational method

A weakness of the perturbation method is that the total hamiltonian H_t must be a close approximation of some H for which the eigenvalues and eigenstates are already known. Hence, we shall now discuss a method (called the variational method) that can be used to find eigenvalues and eigenstates for any bound state problem. However, this method also has a shortcoming. It requires the knowledge of the eigenstate as a function of some undetermined parameters prior to finding the solution. Guessing such a functional form is usually possible if one has some physical understanding of the problem. It will also be seen that the variational method is most appropriate for the ground state. The computation of higher excited states becomes progressively more involved.

To understand the variational method we shall prove the following theorem.

Theorem 10.1 *For an arbitrary nonzero state $|s\rangle$, the expectation value of H , the hamiltonian, satisfies the following inequality.*

$$\langle H \rangle_s \equiv \frac{\langle s|H|s\rangle}{\langle s|s\rangle} \geq E_0$$

where E_0 is the lowest energy eigenvalue (ground state energy). On physical grounds, it is assumed that E_0 exists.

Proof: Let the set of eigenstates of H be $\{|i\rangle\}$ ($i = 0, 1, 2, \dots$) with the corresponding set of eigenvalues $\{E_i\}$. As $\{|i\rangle\}$ must be a complete set, it is possible to expand $|s\rangle$ as the following series.

$$|s\rangle = \sum_i a_{si}|i\rangle. \quad (10.74)$$

Hence, using the orthonormality and eigenstate property of $\{|i\rangle\}$

$$\frac{\langle s|H|s\rangle}{\langle s|s\rangle} = \frac{\sum_i |a_{si}|^2 E_i}{\sum_i |a_{si}|^2} = \frac{\sum_i |a_{si}|^2 (E_i - E_0)}{\sum_i |a_{si}|^2} + E_0 \geq E_0, \quad (10.75)$$

as E_0 is the lowest energy and hence $(E_i - E_0) \geq 0$ for all i . This completes the proof.

The following corollary of the above theorem can also be proved with little difficulty (see problem 5).

Corollary 10.1 *If the state $|s\rangle$ is orthogonal to the lowest n eigenstates of H , that is*

$$a_{si} = \langle i|s\rangle = 0 \quad \text{for } i = 0, 1, \dots, n,$$

then

$$\frac{\langle s|H|s\rangle}{\langle s|s\rangle} \geq E_{n+1}.$$

The variational method involves the choice of a trial ground state $|\alpha, 0\rangle$ that is a function of some arbitrary parameters α_i ($i = 1, 2, \dots$). The expectation value of H for this state is denoted by $\langle H \rangle_{\alpha 0}$. A minimum of $\langle H \rangle_{\alpha 0}$ with respect to all the parameters α_i can be found. From theorem 10.1, we know that this minimum must be greater than or equal to E_0 , and if the functional form of the trial ground state is chosen appropriately, it could be a good approximation for E_0 . The α_i determined in the process of minimization, would also give the corresponding approximation for the ground state $|0\rangle$. To find the first excited state, corollary 10.1 would be used. A trial first excited state is chosen to be orthogonal to the already determined approximate ground state. This trial state can be used to find the approximations of E_1 and $|1\rangle$ by repeating the procedure used for the ground state. In principle, this method can be used for the computation of any number of higher excited state. However, in practice, it becomes progressively more difficult to choose trial states for higher energies that must be orthogonal to *all* lower eigenstates.

To illustrate the method outlined above, we shall find the ground state $|0\rangle$ and the corresponding energy E_0 for a particle of mass m placed in the following spherically symmetric potential.

$$V(r) = -k \frac{\exp(-r/a)}{r}, \quad (10.76)$$

where k and a are positive constants. For this potential, it can be shown that the position representation u_0 , of the ground state $|0\rangle$, behaves as $\exp(-\alpha r)$ at large r . Hence, for the ground state we shall choose this functional form to be the trial eigenfunction.

$$u_0 = \exp(-\alpha r), \quad (10.77)$$

where α will be the only variational parameter used here for the minimizing of the expectation value of H . The expectation value of H for this wavefunction is:

$$\langle H \rangle_0 = \frac{\int u_0^* H u_0 dv}{\int u_0^* u_0 dv}, \quad (10.78)$$

where dv is the volume element and H is in its position representation. If H and dv are written in spherical polar coordinates, the above integrals are quite straightforward to compute. The result is:

$$\langle H \rangle_0 = \frac{\hbar^2 \alpha^2}{2m} - \frac{4k\alpha^3 a^2}{(2\alpha a + 1)^2}. \quad (10.79)$$

For the value of α that will minimize $\langle H \rangle_0$, one must have

$$\frac{d\langle H \rangle_0}{d\alpha} = 0, \quad (10.80)$$

and

$$\frac{d^2\langle H \rangle_0}{d\alpha^2} > 0. \quad (10.81)$$

From equations 10.79 and 10.80, one finds that the value of α that will minimize $\langle H \rangle_0$ must be a solution of the following cubic equation.

$$16(\alpha a)^2 - 12\alpha a(2\alpha a + 1) + b(2\alpha a + 1)^3 = 0, \quad (10.82)$$

where

$$b = \frac{\hbar^2}{mka}. \quad (10.83)$$

Hence, finding α involves the solution of equation 10.82. This can be achieved numerically by standard methods. For the specific case of $b = 20/27$, the solution is possible without resorting to numerical techniques. It is seen to be given by

$$\alpha a = 1, \quad 1/10, \quad -5/4. \quad (10.84)$$

The negative root is not possible as both α and a are positive. Of the other two roots, the first one is seen to give a lower value of $\langle H \rangle_0$. It can also be seen that for $\alpha a = 1$, the inequality 10.81 is satisfied. Hence, we conclude that the minimum in $\langle H \rangle_0$ is obtained by $\alpha a = 1$. This gives the estimate for the ground state energy to be

$$E_0 = -\frac{2k}{27a}, \quad (10.85)$$

and the corresponding normalized eigenfunction is

$$u_0 = \frac{1}{\sqrt{\pi a^3}} \exp(-r/a). \quad (10.86)$$

One may test this method by taking the limit of $a \rightarrow \infty$. In this limit the potential becomes the same as that for the hydrogen atom. The corresponding $\langle H \rangle_0$ is seen to be

$$\langle H \rangle_0 = \frac{\hbar^2 \alpha^2}{2m} - k\alpha. \quad (10.87)$$

Minimizing this gives the value of α to be mk/\hbar^2 . Hence, the estimated value for E_0 is

$$E_0 = -\frac{mk^2}{2\hbar^2}. \quad (10.88)$$

This and the corresponding eigenfunction u_0 can be seen to be the same as the exact results of chapter 8.

Problems

1. If the one dimensional harmonic oscillator is perturbed by a potential of the form

$$H' = A \exp(-aX^2),$$

where A and a are positive constants, find the first order correction to the ground state energy. (Hint: the position representation is convenient for this problem.)

2. Find the first order corrections to the energy eigenstates for the Zeeman effect case discussed in the text.
3. If the hydrogen atom hamiltonian is perturbed by the following:

$$H' = AL^2,$$

where A is a constant and L^2 is the magnitude squared of the angular momentum, find the first order perturbation corrections to the energies of the degenerate states corresponding to $n = 3$.

4. For the time dependent perturbation given in equation 10.53, let

$$H_0 = i\mathbf{k} \cdot \mathbf{R}$$

where \mathbf{k} is the wave vector of the perturbing electromagnetic wave and \mathbf{R} the position vector. If the unperturbed system given by H is the hydrogen atom, find the selection rules for the differences in the angular momentum quantum numbers (l, m) between the initial and final states of a transition.

5. Prove corollary 10.1.
6. For the perturbed harmonic oscillator of problem 1, find the ground state energy by the variational method. Compare the results to that of the perturbation computation. (Hint: One may consider the following as a trial wavefunction with α as the variational parameter: $u_0 = \exp(-\alpha x^2)$.)

Chapter 11

Approximation Methods (Scattering States)

In general, the quantum mechanical scattering problem is difficult to present in a well defined mathematical form. This, obviously, makes it more difficult to solve. The cause of this difficulty is in the inherent nature of the time independent Schrödinger equation. It is an equation that needs boundary conditions. These boundaries are usually at infinity. For a scattering problem one does not know these boundary conditions *a priori*. In fact, the solution of the problem amounts to *finding* the behavior of the wavefunction at infinity! Hence, in the special cases discussed in chapters 5 and 8, we have had to guess the form of the solution at infinity with some undetermined parameters. The solution of the problem had then amounted to the determination of these parameters. However, in a general situation, it is not always possible to guess the form of the solution at infinity. Hence, one sometimes needs to approach the problem differently.

In the following we shall discuss such a different approach. The problem will be considered to be a limiting form of an initial value problem. The potential will be assumed to be “turned on” at some initial time $-T/2$ ($T \rightarrow \infty$) and “turned off” at the later time $T/2$. This does not change the problem physically, but allows one to assume the initial state to be a free particle state with a fixed momentum (i.e. a momentum eigenstate) corresponding to the incident beam. The final state is also a free particle state, but it is expected to be a linear combination of different momentum eigenstates. The probability of the final state being in a momentum state in a certain direction will give the scattering cross section in that direction. Hence, we need a formalism that will allow us to find the state of the system at the time $T/2$, if it is known at the time $-T/2$. This is a standard initial value problem and can be solved by using the following Green’s function method.

11.1 The Green's function method

Let the Schrödinger equation be written in the following form.

$$i\hbar \frac{\partial}{\partial t} |s, t\rangle = [H_0 + V]|s, t\rangle, \quad (11.1)$$

where $|s, t\rangle$ is an arbitrary state for a system given by the hamiltonian

$$H = H_0 + V, \quad (11.2)$$

and

$$H_0 = \frac{P^2}{2m} \quad (11.3)$$

is the kinetic energy part of the hamiltonian. The state $|s, t\rangle$ is labelled by s and, for convenience in the present discussion, its dependence on time t is explicitly specified in the notation. We shall now state without proof that the Schrödinger equation specifies a well defined initial value problem i.e. if $|s, t\rangle$ is given at some time t , then the Schrödinger equation will give a unique $|s, t'\rangle$ at another time t' . Due to the linearity of the equation, it is expected that $|s, t\rangle$ and $|s, t'\rangle$ are linearly related and hence the following convenient form of the relation is chosen.

$$|s, t'\rangle = iG(t', t)|s, t\rangle, \quad (11.4)$$

where $i = \sqrt{-1}$ and $G(t', t)$ is an operator on the linear vector space \mathcal{V} of quantum states as defined in chapter 1. The position representation of $G(t', t)$, given by the following, is usually called the *Green's function*.

$$\langle \mathbf{r}' | G(t', t) | \mathbf{r} \rangle \equiv G(\mathbf{r}', t'; \mathbf{r}, t), \quad (11.5)$$

where $|\mathbf{r}\rangle$ and $|\mathbf{r}'\rangle$ are eigenkets of the position operator. To avoid superfluous nomenclature, we shall call $G(t', t)$ also the Green's function. From problem 2 of chapter 4, we notice that $G(t', t)$ is related to the time translation operator:

$$G(t', t) = -i \exp[-iH(t' - t)/\hbar]. \quad (11.6)$$

For our present application, it is convenient to separate $G(t', t)$ into two parts that propagate a state either forward or backward in time. Hence, we define the *retarded Green's function* or the *propagator* as

$$G^+(t', t) = \begin{cases} G(t', t) & \text{for } t' > t \\ 0 & \text{for } t' < t \end{cases} . \quad (11.7)$$

This gives

$$\theta(t' - t)|s, t'\rangle = iG^+(t', t)|s, t\rangle, \quad (11.8)$$

where $\theta(t' - t)$ is the step function:

$$\theta(t' - t) = \begin{cases} 1 & \text{for } t' > t \\ 0 & \text{for } t' < t \end{cases} . \quad (11.9)$$

Similarly, one defines the *advanced Green's function* as

$$G^-(t', t) = \begin{cases} -G(t', t) & \text{for } t' < t \\ 0 & \text{for } t' > t \end{cases}. \quad (11.10)$$

This gives

$$\theta(t - t')|s, t'\rangle = -iG^-(t', t)|s, t\rangle. \quad (11.11)$$

It can be shown that (see problem 3)

$$\frac{d}{dt'}\theta(t' - t) = \delta(t' - t). \quad (11.12)$$

Using equation 11.12, one can differentiate equation 11.8 to give

$$\delta(t' - t)|s, t'\rangle + \theta(t' - t)\frac{\partial}{\partial t'}|s, t'\rangle = i\frac{\partial}{\partial t'}G^+(t', t)|s, t\rangle. \quad (11.13)$$

Using equations 11.1, 11.8 and the assumption that V , the potential, does not have a time derivative operator term, one obtains

$$\theta(t' - t)\frac{\partial}{\partial t'}|s, t'\rangle = \hbar^{-1}HG^+(t', t)|s, t\rangle. \quad (11.14)$$

Also, from a property of the delta function (see equation 1.72) it is seen that

$$\delta(t' - t)|s, t'\rangle = \delta(t' - t)|s, t\rangle. \quad (11.15)$$

Hence, from equations 11.13, 11.14 and 11.15 we get

$$\left[i\frac{\partial}{\partial t'} - \frac{1}{\hbar}H \right] G^+(t', t)|s, t\rangle = \delta(t' - t)|s, t\rangle. \quad (11.16)$$

As $|s, t\rangle$ is an arbitrary state, one concludes that the following operator relation is true in general.

$$\left[i\frac{\partial}{\partial t'} - \frac{1}{\hbar}H \right] G^+(t', t) = \delta(t' - t). \quad (11.17)$$

It can be shown that $G^-(t', t)$ also satisfies equation 11.17:

$$\left[i\frac{\partial}{\partial t'} - \frac{1}{\hbar}H \right] G^-(t', t) = \delta(t' - t). \quad (11.18)$$

For the special case of the free particle, the Green's functions will be called G_0^+ and G_0^- . They satisfy the following equations.

$$\left[i\frac{\partial}{\partial t'} - \frac{1}{\hbar}H_0 \right] G_0^+(t', t) = \delta(t' - t). \quad (11.19)$$

$$\left[i\frac{\partial}{\partial t'} - \frac{1}{\hbar}H_0 \right] G_0^-(t', t) = \delta(t' - t). \quad (11.20)$$

Now we shall write $G^\pm(t', t)$ (the \pm superscript denotes both the retarded and the advanced functions) in terms of $G_0^\pm(t', t)$ and the potential. In order to do this we need some formal mathematical definitions.

Definition 38 A general TIME DOMAIN LINEAR OPERATION Q on a time dependent state $|s, t\rangle$ is defined by an operator $Q(t, t')$ with the two time arguments t and t' as follows:

$$|s, t\rangle \xrightarrow{Q} |r, t\rangle$$

such that

$$|r, t\rangle = Q|s, t\rangle \equiv \int Q(t, t')|s, t'\rangle dt',$$

where the integration over t' is from $-\infty$ to $+\infty$. Note the compact notation chosen for the time domain operation by Q . Q , without the time arguments, denotes an integration over one of the time arguments as shown. This allows a natural matrix interpretation that will be discussed in the following.

$G^\pm(t', t)$ represent such time domain operators. The product of two time domain operators P and Q is defined through their successive operations:

$$|s, t\rangle \xrightarrow{Q} |r, t\rangle \xrightarrow{P} |u, t\rangle, \quad (11.21)$$

such that

$$\begin{aligned} |u, t\rangle &= P|r, t\rangle = PQ|s, t\rangle = \int P(t, t')|r, t'\rangle dt' \\ &= \int P(t, t') \left[\int Q(t', t'')|s, t''\rangle dt'' \right] dt' \\ &= \int \left[\int P(t, t')Q(t', t'')dt' \right] |s, t''\rangle dt''. \end{aligned} \quad (11.22)$$

Hence, the product of two time domain operators P and Q is found to be also a time domain operator that is defined by the following operator:

$$[PQ](t, t'') = \int P(t, t')Q(t', t'')dt'. \quad (11.23)$$

It is to be noted that these operators bear a strong resemblance to matrix operations. The two time arguments are continuous equivalents of the two discrete indices of a matrix. This matrix nature of time domain operators is over and above their matrix nature in the linear vector space of kets (see chapter 1). In fact, if these operators and the corresponding kets are written in the position representation, both the space and time coordinates behave as continuous indices for matrices. This process of putting space and time on similar footing is useful for the discussion of relativistic quantum mechanics (see chapter 13). For now, we need to identify the identity operator for such time domain operations. It is clearly seen to be the delta function:

$$I_t = \delta(t - t'). \quad (11.24)$$

The operator in the square brackets in equation 11.17 is also a time domain operator. It can be represented in the standard form with two time arguments as follows:

$$\left[i \frac{\partial}{\partial t} - \frac{1}{\hbar} H \right] \rightarrow i \delta'(t - t') - \hbar^{-1} H \delta(t - t'), \quad (11.25)$$

where $\delta'(t - t')$ represents the derivative of the delta function as defined in chapter 1. The form of the operator shown in equation 11.25 can be shown to be correct by operating on an arbitrary time dependent ket $|s, t\rangle$ i.e. one can verify that

$$\left[i \frac{\partial}{\partial t} - \frac{1}{\hbar} H \right] |s, t\rangle = \int [i \delta'(t - t') - \hbar^{-1} H \delta(t - t')] |s, t'\rangle dt'. \quad (11.26)$$

If the time domain operator of equation 11.25 is called K , then equations 11.17 and 11.18 can be written in the following compact form.

$$KG^\pm = I_t. \quad (11.27)$$

Hence, if K^{-1} is defined to have the standard meaning of an inverse, then

$$G^\pm = K^{-1}. \quad (11.28)$$

This gives the interesting result that the inverse of K is not unique! The free particle version of equation 11.28 would be

$$G_0^\pm = K_0^{-1}, \quad (11.29)$$

where the differential operator version of K_0 is

$$K_0 = \left[i \frac{\partial}{\partial t} - \frac{1}{\hbar} H_0 \right]. \quad (11.30)$$

Then, from equations 11.2 and 11.28, one may write

$$G^\pm = [K_0 - V/\hbar]^{-1}. \quad (11.31)$$

Now, for some small enough V , we may formally expand the right side of the above equation as a binomial expression. Such an expansion for operators, is not always valid. We shall justify it for the present case after considering the result as follows.

$$\begin{aligned} G^\pm &= [K_0(I_t - K_0^{-1}V/\hbar)]^{-1} = [I_t - K_0^{-1}V/\hbar]^{-1} K_0^{-1} \\ &= [I_t + K_0^{-1}V/\hbar + (K_0^{-1}V/\hbar)(K_0^{-1}V/\hbar) + \dots] K_0^{-1} \\ &= G_0^\pm + \hbar^{-1} G_0^\pm V G_0^\pm + \hbar^{-2} G_0^\pm V G_0^\pm V G_0^\pm + \dots \end{aligned} \quad (11.32)$$

Here it is seen that the nonuniqueness of the inverse of K on the left hand side shows up in the nonuniqueness of the inverse of K_0 on the right hand side. G_0^\pm has been used for K_0^{-1} with the understanding that the equation is correct for all '+' superscripts or all '-' superscripts. It can be seen that mixing '+' and '-' superscripts will not give the correct

limit for $V \rightarrow 0$. If we consider V to be time dependent, its time domain operator would be represented as

$$V \rightarrow V(t)\delta(t-t'). \quad (11.33)$$

Then it can be seen that each term in the infinite series of equation 11.32 has the meaning of a time domain operator. For example:

$$\begin{aligned} & [G_0^\pm V G_0^\pm V G_0^\pm](t_1, t_6) \\ &= \int \int \int \int G_0^\pm(t_1, t_2) V(t_2) \delta(t_2 - t_3) G_0^\pm(t_3, t_4) \times \\ & \quad \times V(t_4) \delta(t_4 - t_5) G_0^\pm(t_5, t_6) dt_2 dt_3 dt_4 dt_5 \\ &= \int \int G_0^\pm(t_1, t_2) V(t_2) G_0^\pm(t_2, t_4) V(t_4) G_0^\pm(t_4, t_6) dt_2 dt_4. \end{aligned} \quad (11.34)$$

With some changes in the time parameters the above result is rewritten as

$$\begin{aligned} & [G_0^\pm V G_0^\pm V G_0^\pm](t, t') \\ &= \int \int G_0^\pm(t, t_1) V(t_1) G_0^\pm(t_1, t_2) V(t_2) G_0^\pm(t_2, t') dt_1 dt_2. \end{aligned} \quad (11.35)$$

Similarly,

$$\begin{aligned} & [G_0^\pm V G_0^\pm](t, t') \\ &= \int G_0^\pm(t, t_1) V(t_1) G_0^\pm(t_1, t') dt_1. \end{aligned} \quad (11.36)$$

For some small enough V , if the infinite series of equation 11.32 converges¹, it can be shown to give solutions of equations 11.17 and 11.18. In doing this, we substitute the series solutions from equation 11.32 into the left hand sides of equations 11.17 and 11.18:

$$\begin{aligned} KG^\pm &= [K_0 - V/\hbar]G^\pm = K_0G^\pm - \hbar^{-1}VG^\pm \\ &= [I_t + \hbar^{-1}VG_0^\pm + \hbar^{-2}VG_0^\pm VG_0^\pm + \dots] \\ & \quad - [\hbar^{-1}VG_0^\pm + \hbar^{-2}VG_0^\pm VG_0^\pm + \dots] \\ &= I_t, \end{aligned} \quad (11.37)$$

where equation 11.29 is used to see that $K_0G_0^\pm = I_t$. This is the justification for the binomial expansion procedure used in equation 11.32. It can be readily verified that equation 11.32 could also be written as

$$G^\pm = G_0^\pm + \hbar^{-1}G_0^\pm VG^\pm. \quad (11.38)$$

Although we have derived the above equations for both the retarded and the advanced Green's functions, only the retarded function G^+ will be needed for most applications here. In the study of quantum field theories, an interesting linear combination of G^+ and G^- is used to represent the behavior of both particles and antiparticles. This combination is called the Feynman propagator.

¹Here convergence means the existence of a finite limiting value of the quantity $\langle r|G^\pm|s \rangle$ where $\langle r|$ and $|s \rangle$ are arbitrary.

11.2 The scattering matrix

Equipped with the Green's function, we can now return to the scattering problem. As stated earlier we visualize the scattering process to start with a beam of free particles that have a sufficiently sharply defined momentum such that each particle may be considered to be in a momentum eigenstate corresponding to that momentum. As before, the density of the beam is taken to be low enough to assume each particle to be a system by itself that does not interact with the others. *After* this momentum eigenstate is in place, at some time $-T/2$, the scattering potential is turned on. Let the initial state at some time t ($t < -T/2$) be denoted by $|\mathbf{p}, t\rangle$ where \mathbf{p} has the corresponding momentum eigenvalue components. The potential is turned off at a later time $T/2$. At a time t' ($t' > T/2$), the initial state would have transformed to some scattered state $|s, t'\rangle$ given by the Green's function to be

$$|s, t'\rangle = iG^+(t', t)|\mathbf{p}, t\rangle. \quad (11.39)$$

The scattering cross section computation involves the probability of detecting a particle in some given direction. This could be found from the probability of the final particle being in a momentum eigenstate $|\mathbf{p}', t'\rangle$ where \mathbf{p}' is in the given direction. The probability of such an event is given by postulate 4 of chapter 2 to be proportional to $|\langle\mathbf{p}', t'|s, t'\rangle|^2$. Hence, one must compute the following quantity which is known as the scattering matrix.

$$S(\mathbf{p}', \mathbf{p}) = \langle\mathbf{p}', t'|s, t'\rangle = i\langle\mathbf{p}', t'|G^+(t', t)|\mathbf{p}, t\rangle. \quad (11.40)$$

Using equation 11.38, one then gets

$$\begin{aligned} S(\mathbf{p}', \mathbf{p}) &= i\langle\mathbf{p}', t'|G_0^+(t', t)|\mathbf{p}, t\rangle \\ &+ i\hbar^{-1} \int \langle\mathbf{p}', t'|G_0^+(t', t_1)V(t_1)G^+(t_1, t)|\mathbf{p}, t\rangle dt_1. \end{aligned} \quad (11.41)$$

As the state $|\mathbf{p}, t\rangle$ is a momentum eigenstate, it must be an energy eigenstate for the free particle (see chapter 4). Hence, its time dependence is given by equation 4.4 to be

$$|\mathbf{p}, t\rangle = \exp(-iEt/\hbar)|\mathbf{p}\rangle, \quad (11.42)$$

where $E = p^2/(2m)$ and $|\mathbf{p}\rangle = |\mathbf{p}, 0\rangle$ is the momentum eigenstate at time $t = 0$. Thus the operation with the free particle propagator gives (from equation 11.8)

$$G_0^+(t', t)|\mathbf{p}, t\rangle = -i|\mathbf{p}, t'\rangle, \quad (11.43)$$

as it is known that $t' > t$. It can also be shown from the relation of G^- and the adjoint of G^+ (see problem 2) that in the integral in equation 11.41

$$\langle\mathbf{p}', t'|G_0^+(t', t_1) = -i\langle\mathbf{p}', t_1|, \quad (11.44)$$

where it is noted that $t' > t_1$ as the potential is defined to be zero for any time greater than $T/2$. From equations 11.41, 11.42, 11.43 and 11.44 one obtains

$$S(\mathbf{p}', \mathbf{p}) = \delta^3(\mathbf{p}' - \mathbf{p}) + \hbar^{-1} \int \langle\mathbf{p}', t_1|V(t_1)G^+(t_1, t)|\mathbf{p}, t\rangle dt_1, \quad (11.45)$$

where the three dimensional form of equation 3.24 is used to get

$$\langle \mathbf{p}' | \mathbf{p} \rangle = \delta^3(\mathbf{p}' - \mathbf{p}) \equiv \delta(p'_x - p_x) \delta(p'_y - p_y) \delta(p'_z - p_z),$$

which is the product of three delta functions in the three momentum directions. As the potential goes to zero for $t_1 < t$, one may replace the following in equation 11.45.

$$G^+(t_1, t) | \mathbf{p}, t \rangle = -i | s, t_1 \rangle, \quad (11.46)$$

where $| s, t_1 \rangle$ denotes the state of the system at the time t_1 . Then we have

$$S(\mathbf{p}', \mathbf{p}) = \delta^3(\mathbf{p}' - \mathbf{p}) - i\hbar^{-1} \int \langle \mathbf{p}', t_1 | V(t_1) | s, t_1 \rangle dt_1. \quad (11.47)$$

In particular, if the scattering potential is time independent, then one may write

$$V(t) = Vg(t), \quad (11.48)$$

where V depends on position alone and

$$g(t) = \begin{cases} 1 & \text{for } -T/2 < t < T/2 \\ 0 & \text{otherwise} \end{cases} \quad (11.49)$$

and $T \rightarrow \infty$. For practical purposes an actual measurement of scattering is made in a direction away from the incident beam direction² and hence the scattered particle momentum \mathbf{p}' is different from \mathbf{p} . Hence, scattering amplitude computations require only the following part of $S(\mathbf{p}', \mathbf{p})$.

$$S_s(\mathbf{p}', \mathbf{p}) = -i\hbar^{-1} \int \langle \mathbf{p}', t_1 | V(t_1) | s, t_1 \rangle dt_1. \quad (11.50)$$

The corresponding probability is proportional to $|S_s(\mathbf{p}', \mathbf{p})|^2$. As the eigenstates $|\mathbf{p}'\rangle$ are continuous, the transition is expected to occur to a group of states in the infinitesimal neighborhood of \mathbf{p}' . If the number of states in this neighborhood is dM , then the probability of transition is proportional to $|S_s(\mathbf{p}', \mathbf{p})|^2 dM$. Similarly, the probability of scattering in any specific direction (including the direction of the incident beam) can be seen to be $|S(\mathbf{p}', \mathbf{p})|^2 dM$ and hence, the total probability of scattering is $\int |S(\mathbf{p}', \mathbf{p})|^2 dM$. Then, the fractional probability of scattering, in the direction \mathbf{p}' , relative to the total probability of scattering is

$$W = \frac{\int_E |S_s(\mathbf{p}', \mathbf{p})|^2 dM}{\int |S(\mathbf{p}', \mathbf{p})|^2 dM}, \quad (11.51)$$

²Most of the incident beam goes through unscattered. So the off-axis scattering is small and requires high sensitivity particle detectors for its measurement. Such detectors would get overloaded (and probably destroyed) if the strong unscattered beam were to hit them. Besides, the unscattered beam has no interesting information anyway.

where \int_E denotes an integral over only those states $|\mathbf{p}'\rangle$ that are in the same scattering direction but have different energies E . As V is usually small, equation 11.47 allows us to approximate $S(\mathbf{p}', \mathbf{p})$ by $\delta^3(\mathbf{p}' - \mathbf{p})$. Thus we have

$$W = \frac{\int_E |S_s(\mathbf{p}', \mathbf{p})|^2 dM}{\int |\delta^3(\mathbf{p}', \mathbf{p})|^2 dM}. \quad (11.52)$$

We now need to write dM in terms of a d^3p' , the infinitesimal volume element in the \mathbf{p}' space. To do this we realize that for a discrete set of eigenstates $|i\rangle$, labelled by an integer i , the number of states ΔM between $i = M$ and $i = M + \Delta M$ can be written as

$$\Delta M = \sum_{i=M}^{M+\Delta M} \langle i|I|i\rangle, \quad (11.53)$$

where I is the identity operator. This can be generalized for the continuous states $|\mathbf{p}'\rangle$ to get

$$dM = \langle \mathbf{p}'|I|\mathbf{p}'\rangle d^3p'. \quad (11.54)$$

For the continuous eigenstates $|\mathbf{p}\rangle$, the identity may be written as

$$I = \int |\mathbf{p}\rangle \langle \mathbf{p}| d^3p. \quad (11.55)$$

Hence,

$$dM = \left[\int \langle \mathbf{p}'|\mathbf{p}\rangle \langle \mathbf{p}|\mathbf{p}'\rangle d^3p \right] d^3p'. \quad (11.56)$$

From normalization it is seen that $\langle \mathbf{p}'|\mathbf{p}\rangle = \delta^3(\mathbf{p}' - \mathbf{p})$. This gives

$$dM = \delta^3(\mathbf{p}' - \mathbf{p}) d^3p'. \quad (11.57)$$

The $\delta^3(\mathbf{p}' - \mathbf{p})$ in the above expression is infinite. However, we shall leave it as such expecting later cancellation³. Now the integral in the denominator of equation 11.52 can be evaluated to give the following.

$$W = \frac{\int_E |S_s(\mathbf{p}', \mathbf{p})|^2 d^3p'}{\delta^3(\mathbf{p} - \mathbf{p})}. \quad (11.58)$$

Physically, it can be seen that W depends on T , the duration for which the scattering potential is turned on. Hence, it would be convenient to define the transition rate per unit time as

$$w = W/T. \quad (11.59)$$

³If one feels uncomfortable carrying around such infinities, it is possible to use one of the limiting forms of the delta function (see for example equation 1.78) without taking the limit right away. The limit can be taken at the end of all computations at which point no infinities will remain!

In equation 5.16 the probability current was defined. Using that definition we can see the probability current of the incident beam is

$$\mathbf{S} = (\mathbf{p}/m)|\langle \mathbf{r}|\mathbf{p}\rangle|^2, \quad (11.60)$$

where $|\mathbf{r}\rangle$ is the position eigenstate and $\langle \mathbf{r}|\mathbf{p}\rangle$ is the position representation of the momentum eigenstate. From this one may compute the fractional probability current, by dividing by the total probability of the incident beam which is

$$\int |\langle \mathbf{r}|\mathbf{p}\rangle|^2 d^3r = \langle \mathbf{p}|\mathbf{p}\rangle = \delta^3(\mathbf{p} - \mathbf{p}). \quad (11.61)$$

Hence, the fractional probability current of the incident beam is

$$\mathbf{s} = \frac{\mathbf{p}|\langle \mathbf{r}|\mathbf{p}\rangle|^2}{m\delta^3(\mathbf{p} - \mathbf{p})}. \quad (11.62)$$

A three dimensional generalization of equation 3.20 gives

$$\langle \mathbf{r}|\mathbf{p}\rangle = (2\pi\hbar)^{-3/2} \exp(i\mathbf{r} \cdot \mathbf{p}/\hbar). \quad (11.63)$$

Hence,

$$\mathbf{s} = \frac{\mathbf{p}}{(2\pi\hbar)^3 m \delta^3(\mathbf{p} - \mathbf{p})}. \quad (11.64)$$

Using fractional probabilities in equation 5.21 would give the scattering cross section σ as follows.

$$\sigma d\omega = w/|\mathbf{s}|. \quad (11.65)$$

Using equations 11.58, 11.59 and 11.64 one then obtains

$$\sigma d\omega = \frac{(2\pi\hbar)^3 m}{pT} \int_E |S_s(\mathbf{p}', \mathbf{p})|^2 d^3p', \quad (11.66)$$

where $p = |\mathbf{p}|$. As the final energy of the scattered particle is given by $E' = p'^2/(2m)$, it can be seen that in the spherical polar coordinates in \mathbf{p}' space

$$d^3p' = p'^2 d\theta d\phi dp' = mp' dE' d\theta d\phi = mp' dE' d\omega. \quad (11.67)$$

Equations 11.66 and 11.67 would then lead to the following.

$$\sigma = \frac{(2\pi\hbar)^3 m^2}{pT} \int p' |S_s(\mathbf{p}', \mathbf{p})|^2 dE', \quad (11.68)$$

where the subscript E on the integral is dropped as the range of integration is now clear from the differential dE' .

11.3 The stationary case

The computation of σ from equation 11.68 is quite formidable in general. However, a useful special case (stationary case) can be handled with relative ease. This is the situation where the potential is time independent as given by equation 11.48, and the energy of a state is conserved during its propagation in time i.e. the state $|s, t_1\rangle$ defined in equation 11.46 can be written as

$$|s, t_1\rangle = \exp(-iEt_1/\hbar)|s\rangle, \quad (11.69)$$

where $E = p^2/(2m)$ is the energy of the incident particle and $|s\rangle$ is time independent. Then, from equations 11.42, 11.48 and 11.50 one obtains

$$S(\mathbf{p}', \mathbf{p}) = -i\hbar^{-1}\langle\mathbf{p}'|V|s\rangle \int g(t_1) \exp[i(E' - E)t_1/\hbar] dt_1. \quad (11.70)$$

Inserting this in equation 11.68 gives

$$\sigma = \frac{(2\pi)^3 \hbar m^2}{pT} \int p' |\langle\mathbf{p}'|V|s\rangle|^2 \left| \int g(t_1) \exp[i(E' - E)t_1/\hbar] dt_1 \right|^2 dE'. \quad (11.71)$$

The integral over t_1 can be seen to tend to a delta function as $T \rightarrow \infty$. This delta function peaks at $E = E'$. Hence, the term $p' |\langle\mathbf{p}'|V|s\rangle|^2$ can be replaced by its value at $E = E'$ and then taken out of the dE' integral. This gives $|\mathbf{p}'| = p' = p$ and hence

$$\sigma = \frac{(2\pi)^3 \hbar m^2}{T} |\langle\mathbf{p}'|V|s\rangle|^2 \int \left| \int g(t_1) \exp[i(E' - E)t_1/\hbar] dt_1 \right|^2 dE'. \quad (11.72)$$

The integral in equation 11.72 can be written as

$$\begin{aligned} & \int \left| \int g(t_1) \exp[i(E' - E)t_1/\hbar] dt_1 \right|^2 dE' \\ &= \int \int \int g(t) g^*(t') \exp[i(E' - E)(t - t')/\hbar] dt dt' dE' \\ &= 2\pi\hbar \int \int g(t) g^*(t') \delta(t - t') dt dt' \\ &= 2\pi\hbar \int |g(t)|^2 dt = 2\pi\hbar T. \end{aligned} \quad (11.73)$$

In the last step above, the definition of $g(t)$ from equation 11.49 is used. Now equation 11.72 reduces to

$$\sigma = (2\pi)^4 \hbar^2 m^2 |\langle\mathbf{p}'|V|s\rangle|^2. \quad (11.74)$$

This happens to be a special case of Fermi's golden rule. The state $|s\rangle$ can be computed to a desired order of accuracy using the Green's function and then inserted in equation 11.74.

11.4 The Born approximation

For the lowest order computation of σ , one writes equation 11.39 as

$$|s, t'\rangle = iG_0^+(t', t)|\mathbf{p}, t\rangle = |\mathbf{p}, t'\rangle = \exp(-iEt'/\hbar)|\mathbf{p}\rangle. \quad (11.75)$$

Hence, in the lowest order approximation (called the Born approximation), $|s\rangle$ is the initial momentum eigenstate $|\mathbf{p}\rangle$. Then equation 11.74 gives

$$\sigma = (2\pi)^4 \hbar^2 m^2 |\langle \mathbf{p}' | V | \mathbf{p} \rangle|^2. \quad (11.76)$$

As the potential is usually given as a function of position the following computation is best done in the position representation.

$$\begin{aligned} \langle \mathbf{p}' | V | \mathbf{p} \rangle &= \int \int \langle \mathbf{p}' | \mathbf{r}' \rangle \langle \mathbf{r}' | V | \mathbf{r} \rangle \langle \mathbf{r} | \mathbf{p} \rangle d^3 r' d^3 r \\ &= \int \int \langle \mathbf{p}' | \mathbf{r}' \rangle V(\mathbf{r}) \delta^3(\mathbf{r}' - \mathbf{r}) \langle \mathbf{r} | \mathbf{p} \rangle d^3 r' d^3 r \\ &= \int \langle \mathbf{p}' | \mathbf{r} \rangle V(\mathbf{r}) \langle \mathbf{r} | \mathbf{p} \rangle d^3 r \\ &= (2\pi\hbar)^{-3} \int V(\mathbf{r}) \exp[i\mathbf{r} \cdot (\mathbf{p} - \mathbf{p}')/\hbar] d^3 r, \end{aligned} \quad (11.77)$$

where equation 11.63 is used as the position representation of the momentum eigenstate. If

$$\mathbf{q} = (\mathbf{p}' - \mathbf{p})/\hbar, \quad (11.78)$$

then equations 11.76 and 11.77 give

$$\sigma = \frac{m^2}{4\pi^2 \hbar^4} \left| \int V(\mathbf{r}) \exp[-i\mathbf{q} \cdot \mathbf{r}] d^3 r \right|^2. \quad (11.79)$$

For a spherically symmetric potential, the angular part of the integration in equation 11.79 is quite straightforward. The result is

$$\sigma = \frac{4m^2}{\hbar^4 q^2} \left| \int r V(r) \sin(qr) dr \right|^2, \quad (11.80)$$

where $r = |\mathbf{r}|$ is the radial coordinate and $q = |\mathbf{q}|$. For the stationary case, being discussed here, $|\mathbf{p}'| = |\mathbf{p}| = p$. The angle θ between the directions of \mathbf{p}' and \mathbf{p} is then given by

$$\cos \theta = \mathbf{p}' \cdot \mathbf{p} / p^2. \quad (11.81)$$

Also

$$q = \hbar^{-1} |\mathbf{p}' - \mathbf{p}| = (p/\hbar) [2 - 2 \cos \theta]^{1/2} = (2p/\hbar) \sin(\theta/2). \quad (11.82)$$

As an example, we shall consider the case of electrons scattering from a neutral atom of atomic number Z . The potential for such a scatterer was given in chapter 9:

$$V(r) = -\frac{Ze^2}{4\pi\epsilon_0} \frac{\exp(-r/B)}{r}. \quad (11.83)$$

The resulting scattering cross section can be computed to be

$$\sigma = \frac{m^2 Z^2 e^4}{(2\pi\epsilon_0 \hbar^2)^2 (q^2 + B^{-2})^2}. \quad (11.84)$$

The angle dependence of σ is due to q . The total cross section can be computed by integrating over all angles. As there is no ϕ dependence, this leads to

$$\begin{aligned} \sigma_t &= 2\pi \int_0^\pi \sigma \sin\theta d\theta = \frac{2\pi\hbar^2}{p^2} \int_0^{2p/\hbar} \sigma q dq \\ &= \frac{m^2 Z^2 e^4 B^4}{\pi(\epsilon_0 \hbar)^2 (4p^2 B^2 + \hbar^2)}. \end{aligned} \quad (11.85)$$

Problems

- Using the definitions in equations 11.8 and 11.11 show the following.

$$G^+(t', t) = iG^+(t', t_1)G^+(t_1, t) \quad \text{if } t' > t_1 > t.$$

$$G^-(t', t) = -iG^-(t', t_1)G^-(t_1, t) \quad \text{if } t' < t_1 < t.$$

- Show that the adjoint of $G^-(t', t)$ is $G^+(t, t')$ i.e.

$$[G^-(t', t)]^\dagger = G^+(t, t').$$

- Prove equation 11.12 from the definition of the delta function given in chapter 1.
- Prove equation 11.18.
- Verify the relation in equation 11.26.
- Find the scattering cross section in the Born approximation for the following spherically symmetric potential.

$$V(r) = \begin{cases} V_0 & \text{for } r < a \\ 0 & \text{for } r > a \end{cases}.$$

where V_0 is a constant.

Chapter 12

Spin and Atomic Spectra

A comparison of the classical and quantum postulates of chapter 2 brings out an interesting distinction. The classical descriptor, the trajectory, is postulated to be directly observable. But the quantum descriptor, the state vector, is observed only indirectly through the probabilities of measurements. Also, the state vector is defined only through its properties and not as any specific mathematical object. This allows the theoretical freedom of choosing different types of mathematical objects as state vectors. Infinite dimensional wavefunctions (as defined in chapter 3) as well as finite dimensional vectors can represent state vectors in different situations. Until now we have not discussed the observational consequences of having a finite dimensional vector as the state vector. In fact, it is not self evident that there exist any physical systems described by such state vectors. A study of atomic spectra shows that they do.

The emission spectra of atoms show a “fine structure” of lines that cannot be explained by the usual representation of state vectors viz. single valued wavefunctions [10]. This is not a weakness of the principles of quantum mechanics but a result of a system dependent assumption made earlier (chapter 3). In fact, in chapter 3, we had made two significant system dependent assumptions.

Assumption 1 *All positions on a rectangular coordinate system are measurable (i.e. they are eigenvalues of the position operator).*

Assumption 2 *Position eigenstates are nondegenerate.*

The first of these two assumptions is seen to be observationally correct for most systems that need a quantum description. A counterexample would be the postulated structure of the universe in relativistic cosmology. In such a model (Friedmann), space is curved like the three dimensional surface of a sphere embedded in four dimensions. If a rectangular

coordinate system were to be set up in such a space, the positive coordinate direction would loop around and return to the origin! Thus, no coordinate would represent observed positions beyond a certain distance. Similar effects of curved space may also be observed near massive stars. Since most quantum observations are made in a scale significantly smaller than the universe (putting it mildly!) and far enough from massive stars, the approximation of a rectangular coordinate system existing, is a very good one. Hence, we shall not discuss this matter any further in this text.

A counterexample of the second assumption is the subject of discussion in this chapter. Once the position eigenstates are no longer assumed to be nondegenerate, the wavefunctions will not remain single valued. A systematic study of such systems can be seen to explain the “fine structure” in atomic spectra and other phenomena that have no classical analog. Such systems can also be seen to have observable angular momenta that are *not due to spatial rotation* [10]! The extra degrees of freedom from the degenerate position eigenstates are called the SPIN degrees of freedom due to the angular momentum they generate, although there is no accompanying spatial rotation. Spin angular momentum can have the half integer quantum numbers that were seen to be possible, in general, in chapter 7.

12.1 Degenerate position eigenstates

Due to existing experimental evidence, we shall assume the degree of degeneracy d , of position eigenstates, to be finite. It is also assumed that d is a constant for a given particle at all positions. For example, the electron is known to have $d = 2$. For a given position eigenvalue \mathbf{r} , the possible (orthonormal) position eigenstates will now be labelled as $|i, \mathbf{r}\rangle$, $i = 1, 2, \dots, d$. Then the position representation of any state $|s\rangle$ is defined as

$$\langle i, \mathbf{r} | s \rangle = \psi_{si}(\mathbf{r}), \quad \text{for } i = 1, \dots, d. \quad (12.1)$$

The different functions $\psi_{si}(\mathbf{r})$ for each value of i are called the *spinor components*. The operation of a rotation operator U_R , on $|s\rangle$ has the position representation $\langle i, \mathbf{r} | U_R | s \rangle$ that may be written as

$$\langle i, \mathbf{r} | U_R | s \rangle = \sum_{j=1}^d (U_S)_{ij} \psi_{sj}(a^{-1}\mathbf{r}), \quad (12.2)$$

where a is the rotation matrix defined in chapter 7. The U_S matrix operator is yet to be defined. It is to be kept in mind that U_R , a , and U_S depend on three parameters given by the direction $\hat{\mathbf{n}}$ and magnitude θ of the rotation i.e. the functional forms of these operators can be written as $U_R(\hat{\mathbf{n}}, \theta)$, $a(\hat{\mathbf{n}}, \theta)$ and $U_S(\hat{\mathbf{n}}, \theta)$. U_S is introduced to include the possibility that a rotation might transform a single spinor component to a linear combination of spinor components. It can be seen to be unitary as U_R is unitary:

$$\sum_{j=1}^d (U_S)_{ji}^* (U_S)_{jk} = \delta_{ik} \quad \text{or} \quad U_S^\dagger U_S = I. \quad (12.3)$$

We notice that to maintain the physical effect of a rotation, the U_S operators for different physical rotations must mimic the algebra of the corresponding U_R operators. That is if

$$U_R(\hat{\mathbf{n}}_1, \theta_1)U_R(\hat{\mathbf{n}}_2, \theta_2) = U_R(\hat{\mathbf{n}}_3, \theta_3), \quad (12.4)$$

then

$$U_S(\hat{\mathbf{n}}_1, \theta_1)U_S(\hat{\mathbf{n}}_2, \theta_2) = U_S(\hat{\mathbf{n}}_3, \theta_3). \quad (12.5)$$

Hence, the set of U_S operators for all possible rotations must be a finite dimensional unitary representation of the rotation group $SO(3)$. In chapter 7 we have already seen such finite dimensional representations, but have not identified them as such.

It can be seen that the generators of rotation (angular momentum components), on operating on the angular momentum eigenstates $|l, m\rangle$, produce a linear combination of angular momentum eigenstates that have the same value of l . Hence, a subspace

$$\mathcal{V}_l = \left\{ \sum_{m=-l}^l c_m |l, m\rangle, \quad c_m \in \mathcal{C} \right\}, \quad (12.6)$$

of linear combinations of angular momentum states with fixed l is closed under the operation of the angular momentum operators. For example, the operation of L_x on the state $|l, m\rangle$ can be written as

$$L_x |l, m\rangle = \sum_{m'=-l}^l (L_x)_{mm'} |l, m'\rangle, \quad (12.7)$$

where L_x is a finite dimensional matrix of dimensionality $2l + 1$. This matrix must be a representation of L_x within the subspace \mathcal{V}_l . L_y and L_z would similarly represent L_y and L_z . Exponentiating these angular momentum representations as in equation 7.34, gives U_{Rl} , the representation of the rotation group in the subspace \mathcal{V}_l .

$$U_{Rl}(\hat{\mathbf{n}}, \theta) = \exp(-i\mathbf{L}_l \cdot \hat{\mathbf{n}}\theta/\hbar), \quad (12.8)$$

which can be seen to be a representation of $SO(3)$ from the fact that if

$$U_R(\hat{\mathbf{n}}_1, \theta_1)U_R(\hat{\mathbf{n}}_2, \theta_2) = U_R(\hat{\mathbf{n}}_3, \theta_3), \quad (12.9)$$

then

$$U_{Rl}(\hat{\mathbf{n}}_1, \theta_1)U_{Rl}(\hat{\mathbf{n}}_2, \theta_2) = U_{Rl}(\hat{\mathbf{n}}_3, \theta_3). \quad (12.10)$$

This finite dimensional representation is said to be “carried” by the subspace \mathcal{V}_l . Two such representations, U_{Rl} and $U_{Rl'}$ placed in a block diagonal form in a matrix of dimensionality $2(l + l') + 2$, can also be seen to be a representation of $SO(3)$. In fact, any matrix of the following block diagonal form, with two or more blocks, can be seen to be a representation of $SO(3)$.

$$U_{RB}(\hat{\mathbf{n}}, \theta) = \begin{pmatrix} U_{Rl}(\hat{\mathbf{n}}, \theta) & & & \\ & U_{Rl'}(\hat{\mathbf{n}}, \theta) & & \\ & & U_{Rl''}(\hat{\mathbf{n}}, \theta) & \\ & & & \ddots \end{pmatrix}. \quad (12.11)$$

where l , l' , and l'' are the corresponding total angular momentum quantum numbers. U_{RB} is a *reducible* representation of $SO(3)$. A general definition of reducible and irreducible representations is as follows.

Definition 39 Any representation U_A , of a group, that can be reduced to a block diagonal form U_B of more than one block (for all elements), by some unitary transformation given by a constant unitary operator U as follows, is called a *REDUCIBLE REPRESENTATION*.

$$U_B = UU_AU^\dagger.$$

Definition 40 Any representation that is not a reducible representation is called an *IRREDUCIBLE REPRESENTATION* (IRR for short).

We shall now state two theorems without proof [11, 12].

Theorem 12.1 The finite dimensional representations U_{Rl} for each l are IRRs.

Theorem 12.2 Every unitary IRR of the rotation group can be transformed, by a unitary transformation, to a U_{Rl} for some l .

A constant unitary transformation of operators as in definition 39 accompanied by the transformation

$$|s\rangle \rightarrow U|s\rangle, \quad (12.12)$$

of all states $|s\rangle$, can be seen to be equivalent to a fixed coordinate transformation and hence, does not change any physical relationships. So, group representations that are unitarily related, as in definition 39, will be considered identical.

Returning to equation 12.5, it is now evident that U_S must be a finite dimensional matrix of the form U_{RB} . The simplest form of U_S would be the following d dimensional matrix which has the form of U_{RB} with $l = l' = l'' = \dots = 0$.

$$U_S = \begin{pmatrix} U_{R0} & & & \\ & U_{R0} & & \\ & & U_{R0} & \\ & & & \ddots \end{pmatrix}. \quad (12.13)$$

$U_{R0}(\hat{\mathbf{n}}, \theta)$ is seen to be one dimensional and for all $\hat{\mathbf{n}}$ and θ (see problem 2)

$$U_{R0}(\hat{\mathbf{n}}, \theta) = 1. \quad (12.14)$$

Hence, U_S is an identity matrix and does not mix spinor components in a rotation operation as given by equation 12.2. Physically this is equivalent to each spinor component representing a different *scalar particle*¹. Correspondingly, the U_{R0} representation is called the *scalar representation* as it does not change the components of the wavefunction. It is sometimes also called the trivial representation.

Similarly, any spin system that carries a reducible representation U_S , can be split into independent systems each of which would carry an IRR of $SO(3)$ that is included in the original U_S . Hence, one needs to study only those spin systems that carry IRR's of $SO(3)$. Accordingly, a spin system is named after the l value (angular momentum quantum number) of the IRR that it carries. For example, the $l = 0$ IRR is carried by systems with one spinor component (i.e. the kind we have been studying before this chapter) and they are called the spin-zero particles. The spin-half particles carry the $l = 1/2$ IRR and hence have $d = (2l + 1) = 2$. Similarly, spin-one particles have $d = 3$, and are sometimes called vector particles as usual three dimensional vectors also carry the $l = 1$ IRR.

12.2 Spin-half particles

The simplest nontrivial example of spin is spin-half, which has $d = 2$ and U_S as the $l = 1/2$ IRR of $SO(3)$:

$$U_S = U_{R1/2}. \quad (12.15)$$

Some examples of spin-half particles are the electron, the proton and the neutron. The $l = 1/2$ representation must have 2 dimensional (i.e. $2l + 1$) matrices for its generators (viz. L_{lx}, L_{ly}, L_{lz}). For the present case, these generators will be named S_x, S_y and S_z respectively and they will be considered to be the components of the vector operator \mathbf{S} . This is known as the spin operator. Using the standard definition of the angular momentum eigenstates $|l, m\rangle$, it can be seen that (see problem 1)

$$S_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad S_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (12.16)$$

Here, to conform with standard notation for spin operators, the spin-up state ($m = +1/2$) is chosen to have the matrix index 1 and the spin-down state ($m = -1/2$) is chosen to have the matrix index 2². The matrix parts of the above operators are called the Pauli spin matrices and they are written as σ_x, σ_y , and σ_z which are the components of the vector operator $\boldsymbol{\sigma}$. Hence,

$$\mathbf{S} = \frac{\hbar}{2} \boldsymbol{\sigma}. \quad (12.17)$$

¹These are particles that have nondegenerate position eigenstates.

²Mathematically, the more natural choice would have been the smaller m value for the smaller index.

where

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (12.18)$$

The exponential form in equation 12.8, for $l = 1/2$, will now give U_S for the present case.

$$U_S(\hat{\mathbf{n}}, \theta) = \exp(-i\mathbf{S} \cdot \hat{\mathbf{n}}\theta/\hbar). \quad (12.19)$$

As an example, a rotation of θ about the x axis is seen to be given by (see problem 3)

$$\begin{aligned} U_S(\hat{\mathbf{i}}, \theta) &= \exp(-iS_x\theta/\hbar) \\ &= \begin{pmatrix} \cos \theta/2 & -i \sin \theta/2 \\ -i \sin \theta/2 & \cos \theta/2 \end{pmatrix} \\ &= \cos \theta/2 - i\sigma_x \sin \theta/2, \end{aligned} \quad (12.20)$$

where a multiplication by the 2×2 identity matrix is implicit for the term that does not appear to be a matrix. The two spinor components of the wavefunction, as defined in equation 12.1, can be written as a column vector as follows.

$$\langle \mathbf{r} | s \rangle = \psi_s(\mathbf{r}) = \begin{pmatrix} \psi_{s1}(\mathbf{r}) \\ \psi_{s2}(\mathbf{r}) \end{pmatrix}. \quad (12.21)$$

These two-component column vectors are the *spinors* for the spin-half system. Then, from equation 12.2, the rotation operation on $\psi_s(\mathbf{r})$ can be written as the following matrix product.

$$\langle \mathbf{r} | U_R | s \rangle = U_S \langle a^{-1}\mathbf{r} | s \rangle = U_S \psi_s(a^{-1}\mathbf{r}) = U_S \begin{pmatrix} \psi_{s1}(a^{-1}\mathbf{r}) \\ \psi_{s2}(a^{-1}\mathbf{r}) \end{pmatrix}. \quad (12.22)$$

From the derivation of the rotation operation on a single component wavefunction (section 7.4), it can be seen that

$$\psi_{si}(a^{-1}\mathbf{r}) = \exp(-i\mathbf{L} \cdot \hat{\mathbf{n}}\theta/\hbar) \psi_{si}(\mathbf{r}), \quad (12.23)$$

where $\mathbf{L} = \mathbf{R} \times \mathbf{P}$, is the usual spatial angular momentum operator. Hence, from equations 12.19, 12.22 and 12.23 we obtain the rotation operation on a spinor to be given by

$$\langle \mathbf{r} | U_R | s \rangle = \exp[-i(\mathbf{L} + \mathbf{S}) \cdot \hat{\mathbf{n}}\theta/\hbar] \langle \mathbf{r} | s \rangle. \quad (12.24)$$

Thus the rotation operator for such systems can be written as

$$U_R(\hat{\mathbf{n}}, \theta) = \exp[-i\mathbf{J} \cdot \hat{\mathbf{n}}\theta/\hbar], \quad (12.25)$$

where

$$\mathbf{J} = \mathbf{L} + \mathbf{S}. \quad (12.26)$$

The components of the vector operator \mathbf{J} , are clearly seen to be the generators of rotation for spinors. Consequently, from corollary 7.1, one concludes that the components of \mathbf{J} are

conserved in spherically symmetric systems. From a physical point of view this means that, in the absence of external torques, it is \mathbf{J} , and not \mathbf{L} , that is conserved. Hence, it is \mathbf{J} , and not \mathbf{L} , that must be the observable angular momentum. \mathbf{L} , which is sometimes called the orbital angular momentum, has a classical analogue that is due to the rotational motion of the particle. \mathbf{S} , which is sometimes called the spin angular momentum, has no classical analogue and is *not related to spatial rotation*[10]! Nonetheless, \mathbf{S} is physically significant. A spin-half particle, with no orbital angular momentum, will give \mathbf{S} as its observed angular momentum. For example, a stationary electron will still have an angular momentum. This spin angular momentum, in fact, can never vanish because the only possible eigenvalue of $S^2 (= \mathbf{S} \cdot \mathbf{S})$ is $3\hbar^2/4$. Similarly, the measured values of S_x , S_y and S_z can only be their eigenvalues which are $\hbar/2$ and $-\hbar/2$.

Let us now consider a free spin-half particle (e.g. an electron). The hamiltonian

$$H = \frac{P^2}{2m}, \quad (12.27)$$

and all other spin independent operators are implicitly assumed to be multiplied by 2×2 identity matrices to define their operation on spinors. Thus we expect simultaneous eigenstates of the mutually commuting operator set $\{H, S^2, S_z, P_x, P_y, P_z\}$. Using the representation in the equations 12.16 for spin matrices, one finds the position representation of these simultaneous eigenstates to be as follows.

$$\psi_+(\mathbf{r}) = (2\pi\hbar)^{-3/2} \begin{pmatrix} \exp(i\mathbf{p} \cdot \mathbf{r}/\hbar) \\ 0 \end{pmatrix}, \quad (12.28)$$

$$\psi_-(\mathbf{r}) = (2\pi\hbar)^{-3/2} \begin{pmatrix} 0 \\ \exp(i\mathbf{p} \cdot \mathbf{r}/\hbar) \end{pmatrix}, \quad (12.29)$$

where the components of \mathbf{p} are the eigenvalues of the corresponding components of \mathbf{P} , the eigenvalue of H is $p^2/(2m)$ and the eigenvalue of S^2 is $3\hbar^2/4$. The eigenvalue of S_z corresponding to ψ_+ is $\hbar/2$ and corresponding to ψ_- is $-\hbar/2$.

12.3 Spin magnetic moment (Stern-Gerlach experiment)

A direct experimental verification of the electron spin is seen with a rather simple setup (the well-known *Stern-Gerlach* experiment[13]). As shown in fig. 12.1, a beam of electrons (silver atoms were used in the original experiment) passing through a nonuniform magnetic field in the z direction, is seen to split according to the z component of spin of the electrons. Such a setup can also be used to demonstrate some of the peculiar properties of quantum systems[14], for example, the collapse of a state (postulate 5). This is done by first splitting an electron beam according to its two possible z -components of spin. If the $+1/2$ spin beam is selected out (by blocking the other beam), the electrons in this beam will be in a $+1/2$

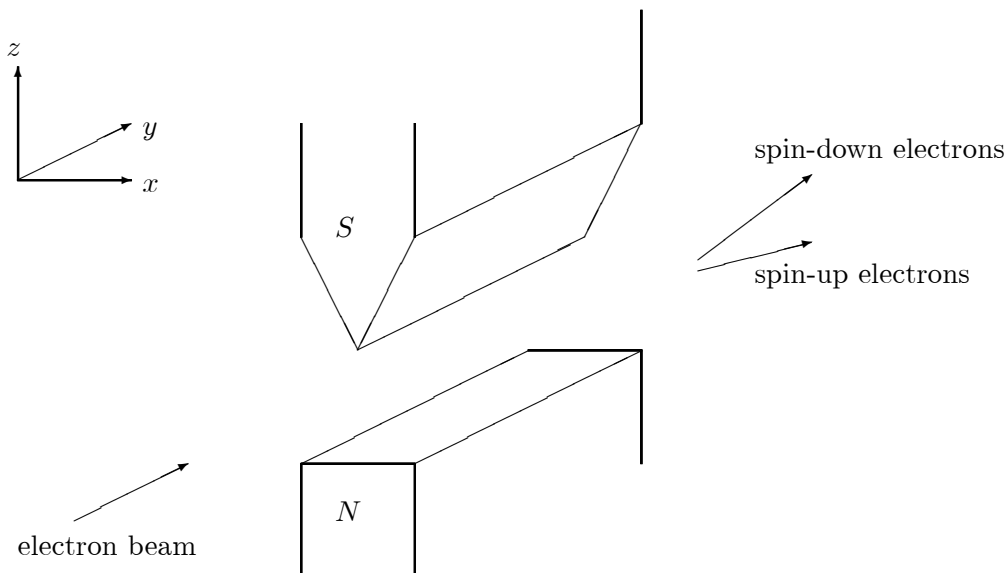


Figure 12.1: A Stern-Gerlach setup – an electron beam in the y direction is split in two along the z direction by a z direction magnetic field that has a strong uniform part and a weak nonuniform part. The nonuniformity is produced by shaping the pole pieces as shown.

spin collapsed state. Hence, a second Stern-Gerlach setup for the z -component of spin will not split this beam any further. However, if the second setup is for the x -component of spin (magnet rotated in the x direction), the beam would split in two once again corresponding to the two possible x -components of spin (see problem 7). Each of these beams can now be split into both a $+1/2$ and a $-1/2$ spin component in the z direction by using another z -component Stern-Gerlach setup!

In this section, the spin magnetic moment will be discussed and a theoretical basis for the Stern-Gerlach experiment will be presented.

A classical particle of mass m , charge q and angular momentum \mathbf{L} , has a magnetic dipole moment

$$\mathbf{M}_L = \frac{q}{2m} \mathbf{L}. \quad (12.30)$$

For the equivalent quantum system the same relation must be true in an operator sense if \mathbf{L} is an orbital angular momentum. For spin angular momenta such a relation cannot be expected to be true as spin has no classical analogue and is not due to a spatial rotation. However, a magnetic dipole moment \mathbf{M}_s , related to the spin \mathbf{S} , is experimentally observed. The relation is seen to be

$$\mathbf{M}_s = \frac{gq}{2m} \mathbf{S}, \quad (12.31)$$

where g is a dimensionless constant that is found to be different for different particles.

For an electron $g = 2$. The relativistic quantum theory of the electron (Dirac[1]) agrees with this experimental value for g (see chapter 13). The Dirac theory cannot explain the experimental values of the magnetic dipole moments of the proton or the neutron if they are assumed to be structureless particles. This was one of the original reasons for suspecting that the proton and the neutron have a substructure. Now it is known that this substructure is that of quarks. Quarks are spin-half particles and they have been found to have $g = 2$ within experimental error as expected from Dirac's theory.

From electrodynamics it is known that a magnetic dipole, of moment \mathbf{M} , placed in a magnetic field \mathbf{B} , has the energy

$$H_m = -\mathbf{M} \cdot \mathbf{B}. \quad (12.32)$$

This energy term was seen for the orbital angular momentum in our discussion of the Zeeman effect (equation 10.49). The electron in an atom has both \mathbf{M}_L and \mathbf{M}_s and hence, its magnetic dipole energy is

$$H_m = -(\mathbf{M}_L + \mathbf{M}_s) \cdot \mathbf{B}. \quad (12.33)$$

If the charge of the electron is given by $-e$, then from equations 12.30, 12.31, 12.33 and the fact that $g = 2$ for the electron, one obtains

$$H_m = \frac{e}{2m}(\mathbf{L} + 2\mathbf{S}) \cdot \mathbf{B}. \quad (12.34)$$

Thus we see that the Zeeman effect computation of chapter 10 is not complete. We shall do the complete computation later. At present, we notice that equation 12.34 can be applied to the Stern-Gerlach setup.

An electron influenced by no force other than a magnetic field, in a suitable coordinate system, has $\mathbf{L} = 0$. Hence, from equation 12.34, in a magnetic field \mathbf{B} , it has the magnetic energy

$$H_m = \frac{e}{m}\mathbf{S} \cdot \mathbf{B}. \quad (12.35)$$

If \mathbf{B} is a uniform magnetic field in the z direction with a magnitude B , then

$$H_m = (eB/m)S_z. \quad (12.36)$$

If this is added to the free particle hamiltonian of equation 12.27 it can be seen that ψ_+ and ψ_- are still the energy eigenstates but with different eigenvalues. The corresponding eigenvalues are

$$E_+ = \frac{p^2 + eB\hbar}{2m}, \quad E_- = \frac{p^2 - eB\hbar}{2m}. \quad (12.37)$$

This separates the two spin states by energy, but the splitting of the beam as required by the Stern-Gerlach setup is still not achieved. To spatially separate the two spin states, one can typically have the electron beam in the y direction and include a small *nonuniform* magnetic field in the z direction. The nonuniform part of the field is kept small compared to

the uniform part to keep the x and y components of the magnetic field small and ignorable³. The x and y components must be kept small to make sure that their contribution to H_m is small enough to consider ψ_+ and ψ_- to be the approximate eigenstates. Now, electrons of the two spin states can be seen to develop a z component of momentum in opposite directions. From equation 4.19 the time rate of change of the expectation value of P_z is seen to be

$$\frac{d}{dt}\langle P_z \rangle_s = \frac{\langle [P_z, H] \rangle_s}{i\hbar}. \quad (12.38)$$

The subscript s for the expectation value denotes the state $|s\rangle$ of the system. Each electron in the beam is a system in itself and could be in either the state ψ_+ or the state ψ_- ⁴. If the expectation values in these states are represented as $\langle \rangle_{\pm}$, then equations 12.36 and 12.38 give

$$\frac{d}{dt}\langle P_z \rangle_{\pm} = \mp \frac{e\hbar}{2m} \frac{dB}{dz}. \quad (12.39)$$

Hence, the two eigenstates develop small z components of momenta in opposite directions thus splitting the beam. It is interesting to observe that, although a nonuniform field must exist in the x or the y direction (in addition to the one in the z direction as $\nabla \cdot \mathbf{B} = 0$), there is no beam splitting in those direction. This is because the strong uniform field in the z direction forces the energy eigenstates to be eigenstates of S_z (the small nonuniform field will perturb this eigenstate only negligibly). If an electron trajectory were to bend in the x direction, it would have to be in an eigenstate of S_x which would not be an energy eigenstate and as energy is being measured inadvertently, this would not be possible (postulate 5).

12.4 Spin-orbit coupling

The inclusion of electron spin will change the hamiltonian for the hydrogen atom as discussed in chapter 8. In the hydrogen atom, from the point of view of the electron, it is the proton that orbits around it. If the velocity of the electron is \mathbf{v} then it “sees” the proton to be moving with respect to itself at a velocity $-\mathbf{v}$. From electrodynamics, it is seen that such a moving charge produces a magnetic field of

$$\mathbf{B} = \mathbf{E} \times \mathbf{v}/c^2, \quad (12.40)$$

where c is the speed of light and \mathbf{E} is the electric field due to the charge:

$$\mathbf{E} = \frac{k_e e}{r^3} \mathbf{r}. \quad (12.41)$$

³A nonuniform z component of the field will always produce x or y components due to the Maxwell equation $-\nabla \cdot \mathbf{B} = 0$.

⁴This statement needs a subtle explanation. Although ψ_{\pm} are the energy eigenstates (ignoring the small effects of the nonuniform field), the system is not expected to be in these states unless an energy measurement is made. Hence, it is important to realize that an “inadvertent” energy measurement occurs in this experiment – One can deduce the energy from the direction of bending of the beam! So, from postulate 5 the system must collapse to an energy eigenstate.

Here the eigenvalue \mathbf{r} is used instead of the operator \mathbf{R} as the position representation is the most convenient for these computations. Thus

$$\mathbf{B} = \frac{k_e c}{c^2 r^3} \mathbf{r} \times \mathbf{v} = \frac{k_e c}{c^2 m r^3} \mathbf{L}. \quad (12.42)$$

Using this relation in equation 12.35 gives the magnetic energy of the spin dipole for the hydrogen atom:

$$H'_{so} = \frac{k_e c^2}{c^2 m^2 r^3} \mathbf{L} \cdot \mathbf{S}. \quad (12.43)$$

However, this energy term is not complete. Due to a relativistic effect called *Thomas precession*, the correct energy of the spin dipole is $H'_{so}/2$. This result will be derived in chapter 13. For now, we shall use the following addition to the hydrogen atom hamiltonian.

$$H_{so} = f(r) \mathbf{L} \cdot \mathbf{S}, \quad (12.44)$$

where

$$f(r) = \frac{k_e e^2}{2c^2 m^2 r^3}. \quad (12.45)$$

This additional term changes the energy spectrum of the hydrogen atom only slightly. Hence, the term *fine structure* is used for the spectral lines shifted as a result of this extra term. H_{so} is sometimes called the *spin-orbit coupling* term due to its dependence on both the spin and the orbital angular momenta. This treatment of spin-orbit coupling can be generalized for other kinds of atoms – in particular alkali metal atoms. An alkali metal atom has one outer shell electron that is loosely bound to the rest of the atom and hence the rest of the atom can be approximated as a rigid spherical object with some charge distribution. This charge distribution produces an electric field that affects the outer electron in a fashion similar to the proton of the hydrogen atom. The actual electric field is, of course, different. But this is remedied by choosing a suitable function $f(r)$ for each alkali metal.

Now, for hydrogen and alkali metals the hamiltonian can be written as

$$H = H_0 + H_{so}, \quad (12.46)$$

where

$$H_0 = \frac{P^2}{2m} + V(r). \quad (12.47)$$

As the effects of spin-orbit coupling are known to be small, one can use the degenerate perturbation method to find the corrections to the energy eigenvalues due to H_{so} . To avoid complicated determinant computations it is desirable (if possible) to rewrite the degenerate eigenstates of H_0 as some linear combinations such that they are also eigenstates of H_{so} . This is seen to be possible by noticing that

$$\mathbf{L} \cdot \mathbf{S} = (J^2 - L^2 - S^2)/2. \quad (12.48)$$

Hence, J^2 , L^2 and S^2 commute with H_0 and H_{so} . So we write the angular part of the eigenstates of H_0 as the eigenstates $|l, j, m\rangle$ of the operators $\{L^2, S^2, J^2, J_z\}$. This is a special case of the addition of angular momenta as discussed in chapter 7. Here the two angular momenta being added are \mathbf{L} and \mathbf{S} and their sum is \mathbf{J} . The quantum number corresponding to S^2 is omitted among the labels of the eigenstate because it is always $1/2$. The eigenvalue equations for each of the operators define the standard meanings of the labels as follows.

$$L^2|l, j, m\rangle = l(l+1)\hbar^2|l, j, m\rangle, \quad (12.49)$$

$$S^2|l, j, m\rangle = (3\hbar^2/4)|l, j, m\rangle, \quad (12.50)$$

$$J^2|l, j, m\rangle = j(j+1)\hbar^2|l, j, m\rangle, \quad (12.51)$$

$$J_z|l, j, m\rangle = m\hbar|l, j, m\rangle. \quad (12.52)$$

Using the Clebsch-Gordan coefficients the states $|l, j, m\rangle$ can be written as linear combinations of the states $|l, m_l, +\rangle$ and $|l, m_l, -\rangle$ that are eigenstates of the operators $\{L^2, S^2, L_z, S_z\}$. The label for S^2 is once again omitted and the rest of the labels are defined by the following eigenvalue equations.

$$L^2|l, m_l, \pm\rangle = l(l+1)\hbar^2|l, m_l, \pm\rangle, \quad (12.53)$$

$$S^2|l, m_l, \pm\rangle = (3\hbar^2/4)|l, m_l, \pm\rangle, \quad (12.54)$$

$$L_z|l, m_l, \pm\rangle = m_l\hbar|l, m_l, \pm\rangle, \quad (12.55)$$

$$S_z|l, m_l, \pm\rangle = \pm(\hbar/2)|l, m_l, \pm\rangle. \quad (12.56)$$

Hence, it is seen that $j = l \pm 1/2$. Using the angular momentum eigenstates $|l, j, m\rangle$, for a degenerate perturbation computation using H_{so} as the perturbation, one obtains the correction terms in energy to be

$$\Delta E_{nlj} = F_{nl}[j(j+1) - l(l+1) - 3/4]\hbar^2/2, \quad (12.57)$$

where

$$F_{nl} = \int_0^\infty |R_{nl}(r)|^2 f(r) r^2 dr, \quad (12.58)$$

and R_{nl} are the radial parts of the corresponding energy eigenfunctions of H_0 .

12.5 Zeeman effect revisited

In chapter 10 we had discussed the Zeeman effect for spinless electrons. Such electrons are, of course, not realistic. To achieve agreement with experiment one needs to use equation 12.34 for the perturbing hamiltonian due to a uniform external magnetic field \mathbf{B} :

$$H_m = \frac{e}{2m}(\mathbf{L} + 2\mathbf{S}) \cdot \mathbf{B}. \quad (12.59)$$

Thus an alkali atom outer electron can be described by the complete hamiltonian

$$H = H_0 + H_{so} + H_m. \quad (12.60)$$

Now neither of the angular states $|l, j, m\rangle$ or $|l, m_l, \pm\rangle$ are eigenstates of both H_{so} and H_m . Hence, perturbation computations would require the diagonalization of matrices. However, for a weak external magnetic field $H_{so} \gg H_m$ and the $|l, j, m\rangle$ states are an appropriate choice. Similarly, for strong magnetic fields $H_{so} \ll H_m$ and the $|l, m_l, \pm\rangle$ are the better choice for computations. The Clebsch-Gordan coefficients are used in either case to find the effect of an operator on a state which is not its eigenstate. The detailed computations will not be shown here (see problem 6).

Problems

1. From the definition in equation 12.7, find the matrices L_{lx} , L_{ly} , and L_{lz} for $l = 0, 1/2$, and 1.
2. From the definition in equation 12.8 and the results of problem 1, find $U_{Rl}(\hat{\mathbf{n}}, \theta)$ for rotations about the x axis (i.e. $\hat{\mathbf{n}} = \hat{\mathbf{i}}$) for $l = 0, 1/2$, and 1.
3. Prove equation 12.20.
4. Prove the following identities:
 - (a) $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1$,
 - (b) $\sigma_x\sigma_y = -\sigma_y\sigma_x = i\sigma_z$,
 - (c) $\sigma_y\sigma_z = -\sigma_z\sigma_y = i\sigma_x$,
 - (d) $\sigma_z\sigma_x = -\sigma_x\sigma_z = i\sigma_y$.
5. Find the spin-orbit interaction energies for the first excited states ($n = 2$) of the hydrogen atom.
6. Find the weak field Zeeman effect corrections for the first excited state ($n = 2$) with $j = 3/2$ of the hydrogen atom.
7. A beam of electrons in the $+1/2$ state for the z -component of spin is travelling along the y direction. If a Stern-Gerlach setup for the x -component of spin is placed in its path, what fraction of the electrons will be seen to have $+1/2$ spin in the x direction?

Chapter 13

Relativistic Quantum Mechanics

At the end of the nineteenth century there were two major unexplained issues – the black-body radiation distribution and electromagnetic theory in the context of Galilean relativity. The explanation of the first led to quantum mechanics and that of the second led to special relativity. The independent development of the two subjects was quite satisfactory. However, when a meeting of the two subjects became necessary (for the description of particles that travelled at speeds close to that of light), there were serious problems.

One of the problems arose primarily due to the postulate of collapse of quantum states (postulate 5). Such a collapse presumably occurs instantaneously. But relativity does not allow any instantaneous movement. The speed of light is a speed limit. This contradiction is illustrated by the well known Einstein-Podolsky-Rosen (EPR) paradox. Originally this was proposed as a paradoxical gedanken experiment. But more recently, such experiments have actually been done[15]. The resolution of the paradox comes from the understanding that the special relativistic speed limit is restricted to the movement of information (or energy).

The second problem is more serious. The description of interaction of several relativistic particles is very tricky even in a classical (non-quantum) setup. In fact, at one time a theorem (no interaction theorem[16]) was proved which claimed that no interaction is possible between two relativistic particles unless they occupied the same space-time point! It was later shown that the conditions used for this proof were too stringent and physical reality did not require them. However, it still illustrates the difficulty of introducing interactions at a distance between relativistic particles. The source of this theoretical dilemma is the four-dimensional nature of relativistic position. For a relativistic description each particle position must include its individual time coordinate. For a physical measurement one knows that the time coordinate of all particles must be the same for a given system. So, constraints are required for the coordinates of a system of particles. These constraints themselves must be relativistically covariant to maintain general covariance. The choice of

such constraints is the sticking point.

A solution to the problem of interaction at a distance is to avoid all interactions at a distance. This is done in quantum field theory (QFT). In QFT every interaction between particles is considered to be mediated by other particles. For example, electromagnetic interactions between electrons are mediated by photons (light particles). An electron produces a photon (at its own space-time point) and transfers energy and momentum to it. The photon travels like a particle to another electron, deposits its energy and momentum to it and then disappears. QFT's have been proposed for every kind of quantum interaction – electromagnetic, weak, and strong. However, only the electromagnetic case has had the most experimental success. This is called quantum electrodynamics (QED).

A discussion of QED is beyond the scope of this book[18, 19]. Instead we shall discuss the simple case of the quantum mechanics of a single relativistic particle in a fixed background potential. This case does not have the problem of interaction at a distance as there is only one particle involved. The relativistic hydrogen atom can be approximated to be such a system as the proton can be approximated to be almost stationary in the frame of reference of the atom due to its significantly larger mass compared to the electron. The stationary proton produces the fixed background potential for the orbiting electron.

13.1 The Klein-Gordon equation

The intuitive approach to a relativistic quantum theory of a single particle would be to write a relativistic form the Schrödinger equation. The form of the Schrödinger equation presented in postulate 1 is general enough to include relativity. Just the hamiltonian H needs to be appropriately chosen. For the free particle, H could be chosen to be energy and formally related to momentum as in special relativity:

$$H = \pm\sqrt{P^2c^2 + m^2c^4}, \quad (13.1)$$

where m is the rest mass of the particle and c the speed of light. The immediate problem of this choice is the sign ambiguity. A rather cavalier decision to drop the possibility of the negative square root can be made. However, that causes serious mathematical problems for the completeness of the eigenstates of the hamiltonian. The other choice is to operate the Schrödinger equation (equation 2.1) by H and then use itself a second time to give a sign unambiguous equation:

$$-\hbar^2 \frac{\partial^2}{\partial t^2} |s\rangle = H^2 |s\rangle. \quad (13.2)$$

Using equations 13.1 and 13.2, one obtains the so-called Klein-Gordon equation for a free particle:

$$-\hbar^2 \frac{\partial^2}{\partial t^2} |s\rangle = (P^2c^2 + m^2c^4) |s\rangle. \quad (13.3)$$

Using the position representation, this equation takes the following form.

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}\right) \Psi_s = \frac{m^2 c^2}{\hbar^2} \Psi_s. \quad (13.4)$$

For $m = 0$, this is exactly the linear wave equation for a wave that travels at the speed of light c .

For the case of light, Ψ_s is the relativistic four component vector potential A^μ ($\mu = 0, 1, 2, 3$) for electromagnetic fields. The zeroth component $A^0 = \Phi/c$ where Φ is the standard scalar potential and the other three are the three components of the magnetic vector potential \mathbf{A} . The correspondence is sometimes written as $A^\mu = (\Phi/c, \mathbf{A})$. The notation here is obvious and can also be used to depict the four component forms of other vectors. For example, the position four-vector is $x^\mu = (ct, \mathbf{r})$ where t is time and \mathbf{r} is the usual position vector. A lower index form of these vectors is sometimes defined for convenience by changing the sign of the zeroth component¹. For example, for position $x_\mu = (-ct, \mathbf{r})$. Besides brevity, the four-vector notation has another benefit. It makes it easier to keep track of relativistic consistency of equations – and all physical equations must be consistent with relativity. Such consistency is sometimes called covariance. A covariant equation, written in the four-vector form, does not change in form in different frames of reference. For example, equation 13.4 can be written as

$$\frac{\partial^2}{\partial x^\mu \partial x_\mu} \Psi_s = \frac{m^2 c^2}{\hbar^2} \Psi_s, \quad (13.5)$$

where it is implicitly assumed that the left hand side is summed over all four values of μ . This assumption will be made whenever a contravariant index and a covariant index are the same and appear in a product of components as in the above equation. This is called the Einstein summation convention and is used extensively to avoid repeated use of the summation sign. For the case of light the corresponding equation is²

$$\frac{\partial^2}{\partial x^\mu \partial x_\mu} A^\nu = 0. \quad (13.6)$$

These are four equations for the four components of A^ν . It is interesting to note that, unlike the usual wavefunctions, the A^ν for light are actually measurable³!

It is also to be noted that A^ν has four components – somewhat like the two-component nature of spinors. This provides a hint about the spin of the light particle – the photon. It

¹The upper index vector is called a contravariant vector and the lower index vector is called a covariant vector. Different conventions for the definitions of such vectors may be found in other texts. A more general approach to such vectors is used in the general theory of relativity.

²The Lorentz gauge is chosen here[17].

³Of course, there is some ambiguity due to gauge choice and the true measurable quantities are the electric and magnetic field vectors.

can be shown that the photon is a spin-one particle[18]. It can also be shown that a particle with a scalar (relativistic) wavefunction has zero spin[18]. In fact, it is possible to write down Klein-Gordon equations for particles of any spin. However, it is seldom necessary to go beyond spin-one as all known fundamental particles have spins of one or less⁴. For spin-half particles, the wavefunction is a four component object, but it is not a four-vector. It is called a Dirac spinor. To understand the component structure of wavefunctions for different spins, one could look back at the two-component spinors (Pauli spinors) discussed in chapter 12. The Pauli spinors were seen to carry the two-dimensional IRR of the SO(3) group. In fact, it was seen that for every IRR of the SO(3) group, there exists a possible spin. An extension of the SO(3) group is the group of all possible special relativistic coordinate transformations (not considering translations). This is the so-called Lorentz group. Every IRR of the Lorentz group corresponds to a possible relativistic spin. The spin-zero IRR is one dimensional (the scalar wavefunction), the spin-half IRR is four-dimensional (the Dirac spinor) and the spin-one IRR is also four dimensional (relativistic four-vector). The Dirac equation (not the Klein-Gordon equation) happens to be the natural choice for spin-half particles and it will be discussed in the next section.

For the free Klein-Gordon particle, the hamiltonian commutes with the momentum operator and hence, as expected, momentum is conserved. The momentum eigenstates $|\mathbf{p}\rangle$ are also energy eigenstates and the energy-momentum relationship is found as follows.

$$E|\mathbf{p}\rangle = H|\mathbf{p}\rangle = \pm\sqrt{P^2c^2 + m^2c^4}|\mathbf{p}\rangle = \pm\sqrt{p^2c^2 + m^2c^4}|\mathbf{p}\rangle, \quad (13.7)$$

where, as before, \mathbf{p} is the eigenvalue of the momentum operator \mathbf{P} . Hence,

$$E = \pm\sqrt{p^2c^2 + m^2c^4}. \quad (13.8)$$

It is seen that the energy could be either positive or negative. It has a minimum positive value of mc^2 and a maximum negative value of $-mc^2$. For free particles, negative energy is physically meaningless. However, the negative energy eigenvalues cannot be mathematically ignored as they are needed for the completeness of the energy eigenstates. A similar dilemma for the Dirac equation had prompted Dirac to postulate what are now known as antiparticles. For every particle of energy E there exists an antiparticle with identical properties with energy $-E$. An antiparticle travels backward in time and hence, from the physical point of view of time translation (equation 7.23), it would appear to be a particle of positive energy travelling forward in time. This resolves the physical problem of negative energy, but introduces the need to experimentally detect the postulated antiparticles. Since the original postulate by Dirac, antiparticles have been amply detected (for example antiprotons, antineutrons, antielectrons (or positrons) etc.).

The energy-momentum relationship of equation 13.8, can also be written in a covariant four-vector form. The zeroth component of the momentum four-vector is related to energy.

⁴Gravitons, if they exist, must have spin-two. Sometimes some tightly bound composites like mesons, baryons or even atomic nuclei might be approximated to be single particles with higher spin values.

The four-vector momentum is $p^\mu = (E/c, \mathbf{p})$. Then the covariant form of equation 13.8 is

$$p^\mu p_\mu + m^2 c^2 = 0. \quad (13.9)$$

It is possible to find energy eigenvalues for Klein-Gordon particles in the presence of static background potentials. However, we shall not do it here as such problems are of little practical value. The hydrogen atom problem is that of an electron in a background potential. But the electron has a spin of half and hence, obeys the Dirac and not the Klein-Gordon equation. It is possible to have a pi meson (pion) orbit around a proton to simulate the Klein-Gordon situation. But this situation is further complicated by other possible interactions of a pion that can be dealt with only by a quantum field theory. The photon interacting with a static charge may be considered to be an interacting Klein-Gordon particle. But such a problem is nothing more than the case of electromagnetic fields in the presence of charge sources which is adequately dealt with in all electromagnetic theory texts.

13.2 The Dirac equation

The square root of an operator as shown in equation 13.1, is a mathematical complication that was avoided by the Klein-Gordon equation by squaring the operator. Dirac[1] avoided the problem by choosing a hamiltonian that is linear in the momentum \mathbf{P} . This choice turned out to be the correct one for spin-half particles like the electron. The Dirac hamiltonian for a free particle is as follows.

$$H = c\boldsymbol{\alpha} \cdot \mathbf{P} + \beta mc^2, \quad (13.10)$$

where $\boldsymbol{\alpha}$ and β are constants yet to be determined. This is the simplest linear relationship possible as mc^2 must be the rest energy of the free particle. Using H from equation 13.10 in equation 2.1 gives the so-called Dirac equation. Once again, H and P are seen to commute and hence, they have simultaneous eigenstates that can be labelled by the momentum eigenvalues \mathbf{p} and written as $|\mathbf{p}\rangle$. Then the energy eigenvalue equation will give:

$$(E - c\boldsymbol{\alpha} \cdot \mathbf{p} - \beta mc^2)|\mathbf{p}\rangle = 0. \quad (13.11)$$

This provides an energy-momentum relationship that does not resemble equation 13.8 in any way. But equation 13.8 is a relativistic kinematic equation for all free particles and hence, must be true in this case as well. The situation is saved by considering β and the components of $\boldsymbol{\alpha}$ to be matrices of dimensionality greater than one and the state vectors to have a column vector nature of the same dimensionality. With this assumption, if we multiply equation 13.11 from the left by $(E + c\boldsymbol{\alpha} \cdot \mathbf{p} + \beta mc^2)$, we get:

$$\begin{aligned} & [E^2 - c^2(\alpha_1^2 p_1^2 + \alpha_2^2 p_2^2 + \alpha_3^2 p_3^2 + (\alpha_1 \alpha_2 + \alpha_2 \alpha_1) p_1 p_2 \\ & + (\alpha_2 \alpha_3 + \alpha_3 \alpha_2) p_2 p_3 + (\alpha_3 \alpha_1 + \alpha_1 \alpha_3) p_3 p_1) - m^2 c^4 \beta^2 \\ & - mc^3((\alpha_1 \beta + \beta \alpha_1) p_1 + (\alpha_2 \beta + \beta \alpha_2) p_2 + (\alpha_3 \beta + \beta \alpha_3) p_3)] |\mathbf{p}\rangle = 0, \end{aligned} \quad (13.12)$$

where α_i , ($i = 1, 2, 3$) are the three components α_x , α_y and α_z of $\boldsymbol{\alpha}$. Now, if equation 13.8 were to be satisfied the following relations must be true.

$$\alpha_i^2 = \beta^2 = 1, \quad \text{for } i = 1, 2, 3, \quad (13.13)$$

$$\alpha_i \alpha_j + \alpha_j \alpha_i = 0, \quad \text{for } i \neq j \text{ and } i, j = 1, 2, 3, \quad (13.14)$$

$$\alpha_i \beta + \beta \alpha_i = 0, \quad \text{for } i = 1, 2, 3. \quad (13.15)$$

As the hamiltonian must be hermitian, the four matrices α_i and β must be hermitian. Then, from the above equations, it can be shown that these matrices must be traceless and their only possible eigenvalues are ± 1 (see problem 1). Hence, their dimensionality must be even. For the simplest possible theory, one chooses the lowest possible dimensionality. However, a dimensionality of 2 does not work. This is because two-dimensional hermitian traceless matrices can have only three independent parameters:

$$\alpha = \begin{pmatrix} c & a - ib \\ a + ib & -c \end{pmatrix}, \quad (13.16)$$

where a , b and c are real and α is an arbitrary two-dimensional hermitian traceless matrix. So, α can be written as a linear combination of the Pauli spin matrices defined in chapter 12:

$$\alpha = a\sigma_x + b\sigma_y + c\sigma_z. \quad (13.17)$$

Then, it can be shown that four such matrices (α_i and β) cannot be found such that they satisfy the conditions of equations 13.14 and 13.15 (see problem 2). Hence, we pick the next simplest choice for the dimensionality – namely 4. For four-dimensional matrices, there are an infinite number of possibilities that will satisfy the conditions of equations 13.14 and 13.15. However, all such possibilities can be seen to produce the same physical results (namely, the eigenvalues of observables). Hence, we shall choose one convenient form of these matrices:

$$\alpha_1 = \begin{pmatrix} 0 & \sigma_x \\ \sigma_x & 0 \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} 0 & \sigma_y \\ \sigma_y & 0 \end{pmatrix}, \quad \alpha_3 = \begin{pmatrix} 0 & \sigma_z \\ \sigma_z & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (13.18)$$

where each entry in the matrices is a 2×2 matrix – 0 represents the zero matrix, 1 represents the identity matrix and the others are the standard Pauli spin matrices (equation 12.18). A more compact form is given by the following vector notation:

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (13.19)$$

Now that the hamiltonian operator has a 4×4 matrix aspect to it, the state vector must be a corresponding 4-component object. Such a state vector is called a *Dirac spinor*.

Accordingly, $|\mathbf{p}\rangle$ must be a 4-component column vector. As this is still a momentum eigenstate, its position representation can be written as:

$$\langle \mathbf{r} | \mathbf{p} \rangle = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} \exp(i\mathbf{p} \cdot \mathbf{r}/\hbar). \quad (13.20)$$

Inserting this in the energy eigenvalue equation (equation 13.11) and using equation 13.18, one obtains the following matrix equation for the u_i .

$$\begin{pmatrix} (E - mc^2) & 0 & -cp_3 & -c(p_1 - ip_2) \\ 0 & (E - mc^2) & -c(p_1 + ip_2) & cp_3 \\ -cp_3 & -c(p_1 - ip_2) & (E + mc^2) & 0 \\ -c(p_1 + ip_2) & cp_3 & 0 & (E + mc^2) \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = 0. \quad (13.21)$$

These are a set of homogeneous equations in the u_i and hence, for a non-zero solution to exist, the determinant of the matrix must vanish. This condition leads to the following solutions for the energy eigenvalue:

$$E_+ = +\sqrt{c^2p^2 + m^2c^4}, \quad E_- = -\sqrt{c^2p^2 + m^2c^4}. \quad (13.22)$$

Equation 13.21 also gives two possible eigenstates for each of these eigenvalues. Each eigenstate can be written as the column vector formed by the four components u_i . For E_+ the eigenstates are:

$$u_{++} = \begin{pmatrix} 1 \\ 0 \\ \frac{cp_3}{E_+ + mc^2} \\ \frac{c(p_1 + ip_2)}{E_+ + mc^2} \end{pmatrix}, \quad u_{+-} = \begin{pmatrix} 0 \\ 1 \\ \frac{c(p_1 - ip_2)}{E_+ + mc^2} \\ \frac{-cp_3}{E_+ + mc^2} \end{pmatrix}. \quad (13.23)$$

For E_- the eigenstates are:

$$u_{-+} = \begin{pmatrix} \frac{cp_3}{E_- - mc^2} \\ \frac{c(p_1 + ip_2)}{E_- - mc^2} \\ 1 \\ 0 \end{pmatrix}, \quad u_{--} = \begin{pmatrix} \frac{c(p_1 - ip_2)}{E_- - mc^2} \\ \frac{-cp_3}{E_- - mc^2} \\ 0 \\ 1 \end{pmatrix}. \quad (13.24)$$

The first subscript for the eigenstate gives the sign of the energy and the second gives the sign of the z -component of its spin. The relation between particle spin and the Dirac spinor will be discussed later. For now, we notice that the negative energy states cannot be avoided even by a hamiltonian linear in momentum. So, once again, they are to be explained as antiparticles as discussed in the case of Klein-Gordon particles.

It is to be noted that the relativistic covariance of the Dirac equation is imposed through the conditions in equations 13.13, 13.14 and 13.15. A more manifestly covariant

form of presenting all relevant equations is possible. For example, equation 13.11 could be multiplied from the left by $-\beta/c$ to obtain

$$(\gamma_\mu p^\mu + mc)|\mathbf{p}\rangle = 0, \quad (13.25)$$

where

$$\gamma_\mu = (-\beta, \beta\boldsymbol{\alpha}), \quad (13.26)$$

and, as before, $p^\mu = (E/c, \mathbf{p})$. However, in this text, we shall not use this notation. It is too compact for an introductory discussion. Once the student becomes reasonably comfortable with manipulations of the standard Dirac matrices $\boldsymbol{\alpha}$ and β , he/she can use the more compact and manifestly covariant formulation.

13.3 Spin and the Dirac particle

If a Dirac particle is placed in a spherically symmetric potential V , its angular momentum must be conserved (see chapter 7). The relevant hamiltonian would be

$$H = c\boldsymbol{\alpha} \cdot \mathbf{P} + \beta mc^2 + V. \quad (13.27)$$

To be conserved, the angular momentum operator must commute with the hamiltonian. The orbital angular momentum $\mathbf{L} = \mathbf{R} \times \mathbf{P}$, by itself, does not commute with H . This is seen by first noticing that $[\mathbf{L}, V] = 0$ (see problem 3) and then computing $[\mathbf{L}, H]$. For example, for the x component

$$[L_x, H] = [L_x, c\boldsymbol{\alpha} \cdot \mathbf{P}] = i\hbar c(\alpha_y P_z - \alpha_z P_y). \quad (13.28)$$

So, in general, for all three components:

$$[\mathbf{L}, H] = i\hbar c\boldsymbol{\alpha} \times \mathbf{P}. \quad (13.29)$$

Hence, quite clearly, \mathbf{L} is not the complete angular momentum of the particle, although it must be part of it. The remaining part of the total angular momentum can be seen to be

$$\mathbf{S} = \frac{\hbar}{2}\boldsymbol{\sigma}', \quad (13.30)$$

where

$$\boldsymbol{\sigma}' = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}. \quad (13.31)$$

Now, if the total angular momentum is defined as

$$\mathbf{J} = \mathbf{L} + \mathbf{S}, \quad (13.32)$$

it is straightforward to show that (see problem 4)

$$[\mathbf{J}, H] = 0, \quad (13.33)$$

and hence, \mathbf{J} must be the conserved total angular momentum. \mathbf{S} is the spin angular momentum. The block diagonal form of \mathbf{S} shows that the first two and the last two components of the Dirac spinor see the effect of \mathbf{S} in pairs exactly the same way as the Pauli spinors see the effect of the two-dimensional spin operator (see chapter 12). Thus it is very satisfying to see spin angular momentum naturally built into the Dirac hamiltonian. In the following, it will be seen that the Dirac hamiltonian also includes the correct spin-orbit coupling term and the correct magnetic moment of the electron.

13.4 Spin-orbit coupling in the Dirac hamiltonian

As discussed in chapter 12, a nonrelativistic analysis misses a factor of half in spin-orbit coupling. We shall now see that this extra factor, due to Thomas precession, is built into the Dirac hamiltonian. To recognize the spin orbit term, as seen in chapter 12, we need to see a nonrelativistic approximation of the Dirac energy eigenvalue problem with a spherically symmetric potential. The eigenvalue equation is

$$E|E\rangle = (c\boldsymbol{\alpha} \cdot \mathbf{P} + \beta mc^2 + V)|E\rangle, \quad (13.34)$$

where the hamiltonian from equation 13.27 is used with the standard notation $|E\rangle$ for the energy eigenstate. Unlike in the free particle case, $|E\rangle$ is not the same as $|\mathbf{p}\rangle$. In a nonrelativistic limit, the first two and the last two components of the Dirac spinor $|E\rangle$ will be seen to decouple. So, for convenience, we shall write

$$|E\rangle = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \quad (13.35)$$

where v_1 and v_2 are two-component Pauli spinors. First, let us assume $|E\rangle$ to be a particle state (not antiparticle). Then E must be positive. In a nonrelativistic limit, the rest mass energy would constitute most of the energy. The quantity that was considered as energy in our earlier nonrelativistic work did not include the rest mass energy. That quantity will now be called E' :

$$E' = E - mc^2. \quad (13.36)$$

So, $E' \ll mc^2$. Now, equation 13.34 can be written as a pair of Pauli spinor equations using equations 13.35 and 13.36:

$$(E' - V)v_1 - c\boldsymbol{\sigma} \cdot \mathbf{P}v_2 = 0, \quad (13.37)$$

$$(E' + 2mc^2 - V)v_2 - c\boldsymbol{\sigma} \cdot \mathbf{P}v_1 = 0. \quad (13.38)$$

The second of this pair of equations shows that, in the nonrelativistic limit, v_2 is smaller than v_1 by a factor of the order of v/c where v is the speed of the particle⁵. For antiparticles, it is the reverse (see problem 7). So, for particles, v_2 is eliminated from equation 13.37 by using equation 13.38. The result is

$$E'v_1 = \frac{1}{2m}(\boldsymbol{\sigma} \cdot \mathbf{P}) \left(1 + \frac{E' - V}{2mc^2}\right)^{-1} (\boldsymbol{\sigma} \cdot \mathbf{P})v_1 + Vv_1. \quad (13.39)$$

The actual approximation is given by the following:

$$\left(1 + \frac{E' - V}{2mc^2}\right)^{-1} \simeq 1 - \frac{E' - V}{2mc^2}. \quad (13.40)$$

The approximation condition is $(E' - V) \ll 2mc^2$. To reduce equation 13.39 to a form similar to the standard nonrelativistic equation, the following identities are used (see problems 5 and 6).

$$\mathbf{P}V = V\mathbf{P} - i\hbar\nabla V, \quad (13.41)$$

$$(\boldsymbol{\sigma} \cdot \nabla V)(\boldsymbol{\sigma} \cdot \mathbf{P}) = (\nabla V) \cdot \mathbf{P} + i\boldsymbol{\sigma} \cdot [(\nabla V) \times \mathbf{P}]. \quad (13.42)$$

Now, using equations 13.40, 13.41 and 13.42 in equation 13.39, we obtain:

$$E'v_1 = \left[\left(1 - \frac{E' - V}{2mc^2}\right) \frac{P^2}{2m} + V \right] v_1 - \frac{i\hbar}{4m^2c^2}(\nabla V) \cdot \mathbf{P}v_1 + \frac{\hbar}{4m^2c^2}\boldsymbol{\sigma} \cdot [(\nabla V) \times \mathbf{P}]v_1. \quad (13.43)$$

As $P^2/(4m^2c^2)$ is already small of the order of v^2/c^2 , the $(E' - V)$ can be further approximated to be $P^2/(2m)$. Also, for a spherically symmetric potential

$$\nabla V = \frac{1}{R} \frac{dV}{dR} \mathbf{R}, \quad (13.44)$$

where $R = \sqrt{\mathbf{R} \cdot \mathbf{R}}$. So now, equation 13.43 can be written as

$$E'v_1 = \left(\frac{P^2}{2m} - \frac{P^4}{8m^3c^2} + V - \frac{i\hbar}{4m^2c^2}(\nabla V) \cdot \mathbf{P} + \frac{1}{2m^2c^2} \frac{1}{R} \frac{dV}{dR} \mathbf{S} \cdot \mathbf{L} \right) v_1, \quad (13.45)$$

where $\mathbf{S} = \hbar\boldsymbol{\sigma}/2$ is the Pauli spinor form of the spin operator and $\mathbf{L} = \mathbf{R} \times \mathbf{P}$ is the orbital angular momentum operator. The last term in equation 13.45 is seen to be the correct spin-orbit coupling term as discussed in chapter 12. The first and third terms are the standard nonrelativistic hamiltonian terms and the remaining two terms have no simple nonrelativistic interpretation.

⁵Consider the \mathbf{P} operator to produce a factor of the order of $m\mathbf{v}$ where \mathbf{v} is the velocity.

13.5 The Dirac hydrogen atom

The nonrelativistic approximation of the last section is useful to have for general spherically symmetric potentials. Such potentials are good approximations for alkali atoms where the single outer shell electron can be treated as a single particle in the spherically symmetric background potential of the nucleus and the other filled shells of electrons. However, for the specific case of the hydrogen atom, the potential is simple and the energy eigenvalue problem can be solved exactly[8].

In doing this, we shall first separate the angular and the radial parts of the Dirac hamiltonian as given in equation 13.27. The radial component of momentum in classical physics is written as $\mathbf{p} \cdot \hat{\mathbf{r}}$, where $\hat{\mathbf{r}}$ is the unit vector in the radial direction. For the quantum analog, it might be tempting to just replace \mathbf{p} and \mathbf{r} by their corresponding operators. However, such a representation of the radial momentum can be seen to be nonhermitian due to the noncommuting nature of position and momentum (see problem 8). In general, to obtain a quantum analog of a product of classical observables, one picks the hermitian part. For example, for two hermitian operators A and B , the hermitian part of AB is $(AB + BA)/2$ and the antihermitian part is $(AB - BA)/2$ (see problem 9). So the radial momentum would be

$$P_r = \frac{1}{2} \left(\mathbf{P} \cdot \frac{\mathbf{R}}{R} + \frac{\mathbf{R}}{R} \cdot \mathbf{P} \right), \quad (13.46)$$

where $R = \sqrt{\mathbf{R} \cdot \mathbf{R}}$. This can be simplified as (see problem 10)

$$P_r = \frac{1}{R} (\mathbf{R} \cdot \mathbf{P} - i\hbar). \quad (13.47)$$

The radial component of $\boldsymbol{\alpha}$ has a simpler form as it commutes with \mathbf{R} :

$$\alpha_r = \boldsymbol{\alpha} \cdot \mathbf{R}/R. \quad (13.48)$$

For a spherically symmetric potential the angular part of the Dirac hamiltonian must be contained in the $c\boldsymbol{\alpha} \cdot \mathbf{P}$ term. To isolate this angular part, we subtract out the radial part $c\alpha_r P_r$. So, the angular part of the hamiltonian is

$$H_a = c(\boldsymbol{\alpha} \cdot \mathbf{P} - \alpha_r P_r). \quad (13.49)$$

To find the relationship of H_a and P_r , we notice that $\alpha_r^2 = 1$ and hence,

$$\alpha_r H_a = c(R^{-1} \boldsymbol{\alpha} \cdot \mathbf{R} \boldsymbol{\alpha} \cdot \mathbf{P} - P_r). \quad (13.50)$$

The first term on the right hand side can be simplified by using the following identity (see problem 11):

$$\boldsymbol{\alpha} \cdot \mathbf{A} \boldsymbol{\alpha} \cdot \mathbf{B} = \mathbf{A} \cdot \mathbf{B} + i\boldsymbol{\sigma}' \cdot (\mathbf{A} \times \mathbf{B}), \quad (13.51)$$

where \mathbf{A} and \mathbf{B} are two arbitrary vectors with no matrix nature relating to $\boldsymbol{\alpha}$. Hence, equation 13.50 reduces to

$$\alpha_r H_a = cR^{-1} i(\boldsymbol{\sigma}' \cdot \mathbf{L} + \hbar). \quad (13.52)$$

where $\mathbf{L} = \mathbf{R} \times \mathbf{P}$. Multiplying both sides by α_r gives

$$H_a = cR^{-1}i\alpha_r(\boldsymbol{\sigma}' \cdot \mathbf{L} + \hbar). \quad (13.53)$$

The angular part of this is in the expression within the parenthesis. Let us call it $\hbar K'$:

$$\hbar K' = \boldsymbol{\sigma}' \cdot \mathbf{L} + \hbar. \quad (13.54)$$

If K' were to commute with H , we could find their simultaneous eigenstates and replace K' by its eigenvalue in equation 13.53. This will be seen to make the energy eigenvalue problem a differential equation in the radial coordinate alone. However, to form such a conserved quantity K' must be multiplied by β . Hence, we define the conserved quantity K (see problem 12):

$$K = \beta K' = \beta(\boldsymbol{\sigma}' \cdot \mathbf{L} / \hbar + 1). \quad (13.55)$$

As $\beta^2 = 1$, we can now write

$$H_a = cR^{-1}i\hbar\alpha_r\beta K. \quad (13.56)$$

Now, using equations 13.27 13.49, and 13.56, we obtain

$$H = c\alpha_r P_r + cR^{-1}i\hbar\alpha_r\beta K + \beta mc^2 + V. \quad (13.57)$$

Then, the energy eigenvalue problem can be written as

$$E|E\rangle = (c\alpha_r P_r + cR^{-1}i\hbar\alpha_r\beta k + \beta mc^2 + V)|E\rangle, \quad (13.58)$$

where k is the eigenvalue of K , and $|E\rangle$ represents simultaneous eigenstates of H and K . To reduce the equation to a differential equation we use the position representation and the spherical polar coordinates. The position representation of P_r in polar coordinates can be found to be (see problem 13):

$$P_r = -i\hbar \left(\frac{\partial}{\partial r} + \frac{1}{r} \right). \quad (13.59)$$

The only remaining matrix behavior is from α_r and β . These two matrices commute with everything other than each other. The necessary relationships of α_r and β are:

$$\alpha_r\beta + \beta\alpha_r = 0, \quad \alpha_r^2 = \beta^2 = 1. \quad (13.60)$$

As seen in section 13.2, these are the only relations that are necessary to maintain the physical correctness of the Dirac equation. The actual form of the matrices can be picked for computational convenience. In the present case such a form would be:

$$\beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \alpha_r = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad (13.61)$$

where each element is implicitly multiplied by a 2×2 identity matrix. With a similar understanding, the position representation of the Dirac spinor can be written as the following object:

$$\langle \mathbf{r} | E \rangle = \begin{pmatrix} F/r \\ G/r \end{pmatrix}, \quad (13.62)$$

where F and G are functions of the radial coordinate r alone. The $1/r$ term is separated out to make the differential equations a little more compact. Using equations 13.59, 13.61 and 13.62, the energy eigenvalue problem of equation 13.58 can be written in the position representation to be:

$$(E - mc^2 - V)F + \hbar c \frac{dG}{dr} + \frac{\hbar ck}{r} G = 0, \quad (13.63)$$

$$(E + mc^2 - V)G - \hbar c \frac{dF}{dr} + \frac{\hbar ck}{r} F = 0. \quad (13.64)$$

These are a pair of coupled first order ordinary differential equations. To find solutions, it is, once again, convenient to define a dimensionless independent variable:

$$\rho = \delta r. \quad (13.65)$$

Inserting this in equations 13.63 and 13.64, and requiring ρ to be dimensionless in the simplest possible way, gives

$$\delta = +\sqrt{\delta_1 \delta_2}, \quad \delta_1 = \frac{mc^2 + E}{\hbar c}, \quad \delta_2 = \frac{mc^2 - E}{\hbar c}. \quad (13.66)$$

Now, the equations 13.63 and 13.64 can be written as

$$\left(\frac{d}{d\rho} + \frac{k}{\rho} \right) G - \left(\frac{\delta_2}{\delta} + \frac{V}{\hbar c \delta} \right) F = 0, \quad (13.67)$$

$$\left(\frac{d}{d\rho} - \frac{k}{\rho} \right) F - \left(\frac{\delta_1}{\delta} - \frac{V}{\hbar c \delta} \right) G = 0. \quad (13.68)$$

For the specific case of the hydrogen atom the spherically symmetric potential is $V = -k_e e^2 / r$, where $-e$ is the electron charge and $k_e = 1/(4\pi\epsilon_0)$ in the usual SI units. Now, a convenient dimensionless parameter can be defined as:

$$\alpha = \frac{k_e e^2}{\hbar c}. \quad (13.69)$$

This is the so-called *fine structure constant* that appears in the fine structure splitting terms of atomic spectra. It is roughly equal to $1/137$. The smallness of this number is critical to many power series approximation methods used in quantum physics (in particular quantum electrodynamics). For the nonrelativistic hydrogen atom (and the harmonic oscillator) a standard method was used to separate the large distance behavior of the eigenfunctions. The same method can be used here as well to separate the two functions F and G as:

$$F(\rho) = f(\rho) \exp(-\rho), \quad G(\rho) = g(\rho) \exp(-\rho). \quad (13.70)$$

Using equations 13.69 and 13.70 in the two equations 13.67 and 13.68, we obtain:

$$\frac{dg}{d\rho} - g + \frac{kg}{\rho} - \left(\frac{\delta_2}{\delta} - \frac{\alpha}{\rho} \right) f = 0, \quad (13.71)$$

$$\frac{df}{d\rho} - f - \frac{kf}{\rho} - \left(\frac{\delta_1}{\delta} + \frac{\alpha}{\rho} \right) g = 0. \quad (13.72)$$

A standard power series solution may be assumed for f and g :

$$f = \rho^s \sum_{i=0}^{\infty} a_i \rho^i, \quad g = \rho^s \sum_{i=0}^{\infty} b_i \rho^i, \quad (13.73)$$

where $b_0 \neq 0$ and $a_0 \neq 0$. Inserting this into the two differential equations for f and g and collecting terms of the same powers of ρ gives the recursion relations:

$$(s + i + k)b_i - b_{i-1} + \alpha a_i - \frac{\delta_2}{\delta} a_{i-1} = 0, \quad (13.74)$$

$$(s + i - k)a_i - a_{i-1} - \alpha b_i - \frac{\delta_1}{\delta} b_{i-1} = 0, \quad (13.75)$$

for $i > 0$. The lowest power terms give the following equations.

$$(s + k)b_0 + \alpha a_0 = 0, \quad (13.76)$$

$$(s - k)a_0 - \alpha b_0 = 0. \quad (13.77)$$

A non-zero solution for a_0 and b_0 can exist only if

$$s = \pm \sqrt{k^2 - \alpha^2}. \quad (13.78)$$

It will soon be seen that $k^2 \geq 1$. Hence, to keep f and g from going to infinity at the origin, we must choose the positive value for s . With this choice, a relationship of a_0 and b_0 can be found. For the other coefficients, one can find from equations 13.74 and 13.75 that

$$b_i[\delta(s + i + k) + \delta_2\alpha] = a_i[\delta_2(s + i - k) - \delta\alpha]. \quad (13.79)$$

Using this back in the the same two equations, gives a decoupled pair of recursion relations. What we need to see from such equations is the behavior of a_i and b_i for large i :

$$a_i \simeq \frac{2}{i} a_{i-1}, \quad b_i \simeq \frac{2}{i} b_{i-1}. \quad (13.80)$$

This shows that both series behave as $\exp(2\rho)$ for large ρ . Hence, to keep the eigenfunctions from going to infinity at large distances, the series must terminate. Fortunately, this is seen to happen to both series with just one condition. If both series were to terminate at $i = n'$ such that $a_{n'+1} = b_{n'+1} = 0$, then it is seen that all subsequent terms in the series also vanish. For different values of n' we obtain different eigenvalues and eigenfunctions. By using either of the equations 13.74 and 13.75 for $n' = i - 1$ we get

$$\delta_2 a_{n'} = -\delta b_{n'}, \quad n' = 0, 1, 2, \dots \quad (13.81)$$

Using equation 13.79 for $i = n'$ along with this gives

$$2\delta(s + n') = \alpha(\delta_1 - \delta_2). \quad (13.82)$$

Writing δ , δ_1 and δ_2 in terms of the energy E (equation 13.66), and solving for E shows that it is positive and given by

$$E = mc^2 \left[1 + \frac{\alpha^2}{(s + n')^2} \right]^{-1/2}. \quad (13.83)$$

That E is positive shows that positrons (electron antiparticles) cannot form bound states with a proton potential⁶.

In equation 13.83, s depends on k . So, we need to find the possible values of k . From equation 13.55, it is seen that K^2 is related to the magnitude of the total angular momentum as follows.

$$\begin{aligned} K^2 &= \hbar^{-2}[(\boldsymbol{\sigma}' \cdot \mathbf{L})^2 + 2\hbar\boldsymbol{\sigma}' \cdot \mathbf{L} + \hbar^2] \\ &= \hbar^{-2}[L^2 + 2\mathbf{S} \cdot \mathbf{L} + \hbar^2] \\ &= \hbar^{-2}[(\mathbf{L} + \mathbf{S})^2 + \hbar^2/4] \\ &= \hbar^{-2}[J^2 + \hbar^2/4], \end{aligned} \quad (13.84)$$

where the result of problem 6 is used with the knowledge that $\mathbf{L} \times \mathbf{L} = i\hbar\mathbf{L}$. As the eigenvalues of J^2 are $j(j+1)$ ($j = 1/2, 1, 3/2, 2, \dots$), the eigenvalues K^2 are seen to be

$$k = \pm 1, \pm 2, \pm 3, \dots \quad (13.85)$$

Although we have found only the values of k^2 , both positive and negative roots for k are considered eigenvalues. This is because the form of the operator K shows that both signs are equally likely for its eigenvalues.

Equation 13.83 agrees very well with experiment including the fine structure splitting of energy levels. To see its connection with nonrelativistic results one may expand it in powers of α^2 (remember the α^2 dependence of s). Keeping terms of upto order α^4 , this gives

$$E = mc^2 \left[1 - \frac{\alpha^2}{2n^2} - \frac{\alpha^4}{2n^4} \left(\frac{n}{|k|} - \frac{3}{4} \right) \right], \quad (13.86)$$

where $n = n' + |k|$. The second term gives the usual nonrelativistic energy levels and the third term gives the fine structure splitting.

⁶Note that E includes the rest mass energy and hence cannot be negative for electrons, although $E < mc^2$.

13.6 The Dirac particle in a magnetic field

The effect of a magnetic field is introduced in the Dirac hamiltonian in a manner similar to classical mechanics. The momentum \mathbf{P} is replaced by $\mathbf{P} - q\mathbf{A}$ where q is the particle charge and \mathbf{A} is the magnetic vector potential. So, the hamiltonian becomes

$$H = c\boldsymbol{\alpha} \cdot (\mathbf{P} - q\mathbf{A}) + \beta mc^2. \quad (13.87)$$

To isolate and recognize the spin magnetic moment term, we need to find the nonrelativistic limit. This is done by squaring the hamiltonian to get

$$H^2 = c^2[\boldsymbol{\alpha} \cdot (\mathbf{P} - q\mathbf{A})]^2 + m^2c^4, \quad (13.88)$$

where the identities in equations 13.13 and 13.15 have been used. Equation 13.51 gives

$$[\boldsymbol{\alpha} \cdot (\mathbf{P} - q\mathbf{A})]^2 = (\mathbf{P} - q\mathbf{A})^2 + i\boldsymbol{\sigma}' \cdot [(\mathbf{P} - q\mathbf{A}) \times (\mathbf{P} - q\mathbf{A})]. \quad (13.89)$$

Using the result of problem 5 it is seen that

$$(\mathbf{P} - q\mathbf{A}) \times (\mathbf{P} - q\mathbf{A}) = -q(\mathbf{A} \times \mathbf{P} + \mathbf{P} \times \mathbf{A}) = i\hbar q \boldsymbol{\nabla} \times \mathbf{A} = i\hbar q \mathbf{B}, \quad (13.90)$$

where \mathbf{B} is the magnetic field. Using equations 13.89 and 13.90 in equation 13.88 gives

$$H^2 = c^2(\mathbf{P} - q\mathbf{A})^2 - \hbar qc^2 \boldsymbol{\sigma}' \cdot \mathbf{B} + m^2c^4. \quad (13.91)$$

Hence, the energy eigenvalue equation is

$$E^2|E\rangle = [c^2(\mathbf{P} - q\mathbf{A})^2 - \hbar qc^2 \boldsymbol{\sigma}' \cdot \mathbf{B} + m^2c^4]|E\rangle. \quad (13.92)$$

For positive particle energies, we can once again write $E = E' + mc^2$, where E' is the energy as defined in the nonrelativistic limit. As $E' \ll mc^2$ in the nonrelativistic limit, one may approximate equation 13.92 to be

$$m^2c^4 \left(1 + \frac{2E'}{mc^2}\right) |E\rangle = [c^2(\mathbf{P} - q\mathbf{A})^2 - \hbar qc^2 \boldsymbol{\sigma}' \cdot \mathbf{B} + m^2c^4]|E\rangle. \quad (13.93)$$

This leads to

$$E'|E\rangle = \left[\frac{1}{2m}(\mathbf{P} - q\mathbf{A})^2 - \frac{\hbar q}{m} \mathbf{S} \cdot \mathbf{B} \right] |E\rangle. \quad (13.94)$$

where \mathbf{S} is the spin operator. The interaction term of spin and magnetic field gives the experimentally correct expression for the magnetic dipole moment of an electron (see chapter 12).

Problems

1. Show that, in order to satisfy equations 13.13, 13.14 and 13.15, the matrices α_i and β must be traceless and their only possible eigenvalues must be ± 1 . [Hint: Show, for example, $\beta = \alpha_1 \alpha_1 \beta = -\alpha_1 \beta \alpha_1$ and use the cyclic property of the trace: $\text{Tr}(\alpha\beta\gamma) = \text{Tr}(\gamma\alpha\beta)$.]
2. Show that the matrices α_i and β cannot satisfy the equations 13.14 and 13.15 if they are two-dimensional. [Hint: Use the general form given by equation 13.17.]
3. A spherically symmetric scalar potential V is a function of $\mathbf{R} \cdot \mathbf{R}$ alone. Show that all components of the angular momentum \mathbf{L} commute with V , that is

$$[\mathbf{L}, V] = 0.$$

[Hint: Assume V to be a power series in $\mathbf{R} \cdot \mathbf{R}$. A Taylor series expansion about a suitable origin is used to avoid negative powers.]

4. Prove that the total angular momentum \mathbf{J} of the Dirac particle is conserved for a spherically symmetric potential.
5. If V is a function of position alone, show that

$$[\mathbf{P}, V] = -i\hbar \nabla V.$$

[Hint: Assume V to be a power series in each of the three position coordinates. A Taylor series expansion about a suitable origin is used to avoid negative powers.]

6. If \mathbf{A} and \mathbf{B} are two vectors with no matrix nature related to spin matrices, then show that

$$(\boldsymbol{\sigma} \cdot \mathbf{A})(\boldsymbol{\sigma} \cdot \mathbf{B}) = \mathbf{A} \cdot \mathbf{B} + i\boldsymbol{\sigma} \cdot (\mathbf{A} \times \mathbf{B}).$$

[Hint: Use the results of problem 4 of chapter 12.]

7. Find the nonrelativistic limit of equation 13.34 for antiparticle states.
8. Show that the following operator analog of the radial momentum is nonhermitian.

$$\mathbf{P} \cdot \frac{\mathbf{R}}{\sqrt{\mathbf{R} \cdot \mathbf{R}}}$$

9. Show that $(AB + BA)/2$ is hermitian and $(AB - BA)/2$ is antihermitian if A and B are hermitian operators. [Definition: An operator C is antihermitian if $C = -C^\dagger$.]
10. Using equation 13.46, show that

$$P_r = \frac{1}{R}(\mathbf{R} \cdot \mathbf{P} - i\hbar).$$

[Hint: Operate on an arbitrary state in the position representation.]

11. Prove the identity in equation 13.51.

12. Prove the following relations:

(a) $[\alpha_r, K] = 0,$

(b) $[\beta, K] = 0,$

(c) $[P_r, K] = 0,$

(d) $[\alpha_r, P_r] = 0.$

13. Prove that in spherical polar coordinates the position representation of P_r is

$$P_r = -i\hbar \left(\frac{\partial}{\partial r} + \frac{1}{r} \right).$$

Appendix A

'C' Programs for Assorted Problems

The following 'C' programs are not particularly "user-friendly" or "robust". They are presented in this form so that students can easily identify the essential components of the numerical methods involved. The code is meant to be used in conjunction with the material in the text. Definitions of parameters must be understood before using the programs. Comments are provided to help do this.

A.1 Program for the solution of energy eigenvalues for the rectangular potential well

```
#include <stdio.h>
#include <math.h>

void main()
{
float mid, lhs, rhs, inter, acc, gam, xi;
int n;

printf("\n Enter potential parameter gamma\n\n");
scanf("%f",&gam);
printf("\n enter accuracy \n\n");
scanf("%f",&acc);
```

```

n=0;
xi =0;
inter=PI/2;
while (xi <= gam)
{
while (inter > acc)
{
inter /= 2;
mid = xi + inter;
if (mid < gam)
{
lhs = mid*tan(mid);
rhs = sqrt(gam*gam - mid*mid);
if (lhs < rhs) xi = mid;
}
}
printf("\n The %d th root for xi is %f\n", n, xi);
n++;
xi = n*PI;
inter = PI/2;
}
}

```

A.2 General Program for one dimensional scattering off arbitrary barrier

```

#include <stdio.h>
#include <math.h>

double ul,u,uu,vl,v,vu;

void main()
{
double e,vt,del,ki,kt;
double din,ud,vd,ref,tran;
extern void scatter();

/* scatter() defines the specific potential to be used. See next
listing for the case of rectangular barrier. */

```

```

printf("enter data\n\n e,vt,del\n\n");

/* Choose units such that sqrt(2m)/hbar = 1. */
/* 'e' is energy of incoming particle in above units. */
/* 'vt' is potential energy in scattered region in above units. */
/* 'del' is interval in computation of dimensionless variable y of text. */

scanf("%lf %lf %lf",&e,&vt,&del);

ki = sqrt(e);
kt = sqrt(e-vt);

ul = 1; u = 1;
vl = 0; v = -del*kt/ki;

scatter(e,del);

/* scatter() defines the specific potential to be used. See next
listing for the case of rectangular barrier.*/

ud = (ul-u)/del; vd = (vl-v)/del;
din = (u+vd)*(u+vd) + (ud-v)*(ud-v);

printf("\n %lf %lf %lf %lf %lf %lf %lf %lf\n",ul,u,uu,ud,vl,v,vu,vd);

ref = ((u-vd)*(u-vd) + (ud+v)*(ud+v))/din;
tran = 4*kt/(ki*din);

printf ("\n\n reflection coeff. = %lf \n\n
transmission coeff. = %lf \n",ref,tran);
}

```

A.3 Function for rectangular barrier potential

```

#include <stdio.h>
#include <math.h>

extern double ul,u,uu,vl,v,vu;

```

```

void scatter(e,del)

double e,del;

{
double r,pot,rf;

printf("\n enter barrier height and width (dimensionless) \n");

/* Units are discussed in calling program */

scanf("%lf %lf",&pot,&rf);

for (r = -del; r > -rf; r += del)
{
uu = ((pot/e - 1)*del*del + 2)*u - ul;
vu = ((pot/e - 1)*del*del + 2)*v - vl;
ul = u; u = uu;
vl = v; v = vu;
}
}

```

A.4 General energy eigenvalue search program

```

#include <stdio.h>
#include <math.h>

void main()
{
double de,dem,be,del,e,e1,f2n1,f2n;
int ne,k,j;
extern double diff();

/* diff() defines the specific potential to be used. See next two
listings for examples. */

double zeroin();
FILE *fopen(), *fp;

printf("enter data\n\n de,dem,be,ne,k,del\n");

```



```

/* Data parameters are defined for equations written with a dimensionless
   position variable 'r' (different symbols used for different problems
   in text). 'de' is the energy interval for rough linear search.
   'dem' is the tolerable error for energy eigenvalues. 'be' is the lower
   starting point for energy search. 'ne' is the number of
   eigenvalues to be computed. 'k' is used with different meanings
   for different potentials. */

scanf("%lf %lf %lf %d %lf %d %lf %lf",
&de,&dem,&be,&ne,&k,&del);
if((fp=fopen("outdat","r+")) == NULL)
    if((fp=fopen("outdat","w")) == NULL)
        printf("\n error opening outdat\n");
if(fseek(fp,0,2))
    clrerr(fp);
fprintf(fp,"\n parameters de,dem,be,ne,k,del are \n");
fprintf(fp," %f %f %f %d %f %d %f %f",de,dem,be,ne,k,del);

e=be; f2n1=0;

for(j=0;j<ne;j++)
    {
        f2n=diff(e,k,del);

/* diff() defines the specific potential to be used. See next two
listings for examples. */

        if(f2n1*f2n<0)
            {
                e1=zeroin(e,de,dem,f2n,k,del);
                printf("\n %f %f %f \n",e1,f2n1,f2n);
                fprintf(fp,"\n %f %f %f \n",e1,f2n1,f2n);
            }
        f2n1=f2n; e=e+de;
    }
fclose(fp);
}

double
zeroin(e,de,dem,f2n,k,del)

double e,de,dem,f2n,del;

```

```

int k;

{
double de1,del1,em,ff;
extern double diff();

de1=de;
del1=del;
while(de1>dem)
  {
  de1=de1/2; em=e-de1;
  del1=del1/2;
  ff=diff(em,k,del1);
  if(ff*f2n>=0.0)
    {
    e=em;
    em=em-de1;
    ff=diff(em,k,del1);
    if(ff*f2n>=0.0) e=em;
    }
  else
    {
    ff=diff(em,k,del1);
    if(ff*f2n<0.0) e=e+de1;
    }
  }
return(e);
}

```

A.5 Function for the harmonic oscillator potential

```

#include <stdio.h>
#include <math.h>

double
diff(e,k,del)

```

```

/* 'k' is zero for even wavefunctions and non-zero for odd wavefunctions.
The odd wavefunction is handled somewhat differently from the text
for better accuracy. It is taken as an even function multiplied by 'r'
and the even function is then computed numerically. Energy search

```

```
    can be started at zero (the value for 'be'). */

double e,del;
int k;
{
double vl,vu,v,r;

vl = 1;
v = 1;
r = del;

if (k == 0)
    while(abs(v)<10)
        {
        vu = v*((r*r-e)*del*del + 2) - vl;
        r=r+del;
        vl = v; v = vu;
        }

else
    while(abs(v)<10)
        {
        vu = v*((r*r-e)*del*del + 2) + vl*(del/r-1);
        vu = vu/(1+del/r);
        r = r+del;
        vl = v; v = vu;
        }

return(v);
}
```

A.6 Function for the hydrogen atom potential

```
#include <stdio.h>
#include <math.h>
#define TRUE 1
#define FALSE 0

double
diff(e,k,del)
```

```

/* 'k' is the total angular momentum quantum number 'l'. Here 'e' is not
quite energy. But it is related to energy. It is the parameter defined
in the text as the greek letter 'beta'. One may use 'be' as zero for
this case too. Although, energy eigenvalues are negative and the
ground state cannot be easily estimated, 'beta' can still be found
to have a lower bound of zero. */

double e, del;
int k;
{
int tail=FALSE;
register double vl,v,r,vu,ab;
double dif;
register double small,large;
double verysmall;
small = del/10;
verysmall = small/10;
vl = 1;
v = 1 - e*del/(2*k+2);
r = del;
large = 5*vl;
ab = abs(v);

while(ab<large && !tail)
{
do
{
vu = ((2+(0.25-e/r)*del*del)*v + ((k+1)*del/r-1)*vl)/(1+(k+1)*del/r);
r += del;
vl = v; v = vu;
ab = abs(v);
}
while (ab>small && ab<large);
dif = abs(v-vl);
if (dif<verysmall) { large = del; tail = TRUE; }
}
if (tail)
while(abs(v)<large)
{
vu = ((2+(0.25-e/r)*del*del)*v + ((k+1)*del/r-1)*vl)/(1+(k+1)*del/r);
r += del;
}
}

```

```
    vl = v; v = vu;  
    }  
  
return(v);  
}
```

Appendix B

Uncertainties and wavepackets

In chapter 4, it was seen that both position and momentum cannot be measured precisely at the same time. It can also be seen that this would be true for any two noncommuting observables. A more precise statement of this fact will be made now in the form of the well-known Heisenberg uncertainty principle. First, we need to define the uncertainty in the measurement of an observable. The root-mean-squared error is a good measure of uncertainty. For the position operator X , this would be

$$\Delta x = \sqrt{\langle (X - \langle X \rangle_s)^2 \rangle_s}, \quad (\text{B.1})$$

where the expectation values are for some given state $|s\rangle$ as defined in equation 4.18. For simplicity, a Kronecker delta normalization is assumed for this state: $\langle s|s\rangle = 1$. Similarly, for the momentum operator P , the measure of uncertainty would be

$$\Delta p = \sqrt{\langle (P - \langle P \rangle_s)^2 \rangle_s}. \quad (\text{B.2})$$

Now it can be shown that the product of these uncertainties has a minimum possible value. In doing so, we consider the square of the uncertainty product:

$$(\Delta x)^2 (\Delta p)^2 = \langle s|A^2|s\rangle \langle s|B^2|s\rangle, \quad (\text{B.3})$$

where

$$A = X - \langle X \rangle_s, \quad B = P - \langle P \rangle_s. \quad (\text{B.4})$$

As A is hermitian

$$(A|s\rangle)^\dagger = \langle s|A^\dagger = \langle s|A. \quad (\text{B.5})$$

A similar relation is true for B as well. Hence, one may define

$$|u\rangle = A|s\rangle, \quad |v\rangle = B|s\rangle, \quad (\text{B.6})$$

such that equation B.3 can be written as

$$(\Delta x)^2(\Delta p)^2 = \langle u|u\rangle\langle v|v\rangle \geq |\langle u|v\rangle|^2. \quad (\text{B.7})$$

The above inequality can be proved as follows.

$$\begin{aligned} 0 &\leq \left| |u\rangle - \frac{\langle v|u\rangle}{\langle v|v\rangle} |v\rangle \right|^2 \\ &= \langle u|u\rangle - \frac{|\langle u|v\rangle|^2}{\langle v|v\rangle}. \end{aligned} \quad (\text{B.8})$$

Hence the result. Now, using the definition of the commutator bracket, equations B.6 and B.7 give

$$\begin{aligned} (\Delta x)^2(\Delta p)^2 &\geq \left| \langle s| \left[\frac{1}{2}[A, B] + \frac{1}{2}(AB + BA) \right] |s\rangle \right|^2 \\ &= \frac{1}{4} |\langle s|[A, B]|s\rangle|^2 + \frac{1}{4} |\langle s|(AB + BA)|s\rangle|^2. \end{aligned} \quad (\text{B.9})$$

As the commutator $[X, P] = i\hbar$, it is seen that

$$[A, B] = i\hbar. \quad (\text{B.10})$$

Hence, from equation B.9, we get

$$(\Delta x)^2(\Delta p)^2 \geq \hbar^2/4, \quad (\text{B.11})$$

where the term involving $(AB + BA)$ is dropped as it is seen to be non-negative and does not change the inequality. This gives the celebrated Heisenberg uncertainty relation to be

$$(\Delta x)(\Delta p) \geq \hbar/2. \quad (\text{B.12})$$

This result has been derived for a position operator and the corresponding momentum operator. But it must be true for any two operators with the same commutator. It is clearly true for position and momentum in each of the three spatial dimensions. In the above derivation, the only place where the commutator relation is actually used is equation B.9. Hence, an uncertainty relation can be found for any pair of arbitrary operators by inserting the appropriate commutator in equation B.9. If the commutator is zero, the minimum uncertainty is zero and hence, the corresponding observables can be measured simultaneously with indefinite accuracy.

To get a feel for the minimum uncertainty product, we shall find a one-dimensional single particle state which has the minimum possible uncertainty product of position and

momentum as allowed by equation B.12. To obtain the minimum, we must find the conditions for the equality options in the inequalities of equations B.7 and B.11. These conditions are quickly seen to be:

$$A|s\rangle = \lambda B|s\rangle, \quad (\text{B.13})$$

$$\langle s|(AB + BA)|s\rangle = 0, \quad (\text{B.14})$$

where λ is a constant yet to be determined. To find $|s\rangle$ in its position representation, we write equation B.13 in its position representation (see chapter 3). It reduces to the following first order differential equation (using the definitions in equation B.4).

$$\frac{d\psi}{dx} = \left[\frac{i(x - x_0)}{\lambda\hbar} + \frac{ip_0}{\hbar} \right] \psi, \quad (\text{B.15})$$

where ψ is the position representation of $|s\rangle$, $x_0 = \langle X \rangle_s$ and $p_0 = \langle P \rangle_s$. This differential equation has the solution:

$$\psi = N \exp \left[\frac{i(x - x_0)^2}{2\lambda\hbar} + \frac{ip_0x}{\hbar} \right], \quad (\text{B.16})$$

where N , the integration constant, is the normalization constant. Hence, it can be determined by the following normalization condition.

$$1 = \langle s|s\rangle = \int |\psi|^2 dx. \quad (\text{B.17})$$

To determine λ , we eliminate A in equation B.14 by using equation B.13 and its conjugate:

$$\langle s|A = \lambda^* \langle s|B. \quad (\text{B.18})$$

This gives

$$(\lambda + \lambda^*) \langle s|B^2|s\rangle = 0. \quad (\text{B.19})$$

As $\langle s|B^2|s\rangle = \langle v|v\rangle$ is the norm of a nonzero ket, it must be nonzero. Hence, to satisfy equation B.19, λ must be purely imaginary. λ must also be negative imaginary to prevent ψ from going to infinity at infinity (see equation B.16). We can also relate λ and N to the position uncertainty Δx by using the following condition on ψ which really is the definition given in equation B.1 written in the position representation.

$$\int (x - x_0)^2 |\psi|^2 dx = (\Delta x)^2. \quad (\text{B.20})$$

Now, the conditions in equation B.17 and B.20 will give

$$\psi = [2\pi(\Delta x)^2]^{-1/4} \exp \left[-\frac{(x - x_0)^2}{4(\Delta x)^2} + \frac{ip_0x}{\hbar} \right]. \quad (\text{B.21})$$

It can be seen that the ground state of the harmonic oscillator is exactly this state. As it is an eigenstate of the hamiltonian, it does not change with time. The same minimum uncertainty state is also possible for the free particle. However, for the free particle, it is not an eigenstate of the hamiltonian and hence, it does change with time. The change, with time, of the free particle minimum uncertainty state can be seen to be similar to that of the position eigenstates as discussed in chapter 4. The wavefunction spreads with time and does not remain a minimum uncertainty state.

The minimum uncertainty state is sometimes called the minimum uncertainty wavepacket. The term “wavepacket” is loosely used for any wavefunction that is localized in a small region of space. The minimum uncertainty wavepacket can be visualized as a gaussian (bell shaped) wavefunction that has a forward motion in time with momentum p_0 .

Bibliography

- [1] Dirac P. A. M., *The Principles of Quantum Mechanics*, (Oxford University Press).
- [2] Goldstein H., *Classical Mechanics*, (Addison-Wesley Publishing Company Inc.).
- [3] Mathews J. and Walker R. L., *Mathematical Methods of Physics*, (W. A. Benjamin Inc.).
- [4] von Klitzing K. et. al., Phys. Rev. Lett. **45**, 494 (1980).
- [5] Press W. H., Teukolsky S. A., Vetterling W. T. and Flannery B. P., *Numerical Recipes in C*, (Cambridge University Press).
- [6] Herzberg G., *Atomic Spectra and Atomic Structure*, (Dover Publications).
- [7] Watson G. N., *Theory of Bessel Functions*, (Macmillan).
- [8] Schiff L. I., *Quantum Mechanics*, (McGraw-Hill Book Company).
- [9] Merchant S. L., *Impurity States in a Quantum Well: A Numerical Approach*, Masters thesis, SUNY at New Paltz (1993).
- [10] Biswas T., Samuel Goudsmit and Electron Spin, *Models and Modellers of Hydrogen*, (ed. A. Lakhtakia, World Scientific).
- [11] Hamermesh M., *Group Theory*, (Addison-Wesley Publishing Company Inc.).
- [12] Tinkham M., *Group Theory and Quantum Mechanics*, (McGraw-Hill Book Company).
- [13] Stern O., Z. Physik **7**, 249 (1921); Gerlach W. and Stern O., Z. Physik **8**, 110 and **9**, 349 (1922); Ann. Physik **74**, 673 (1924).
- [14] Feynman R. P., Leighton R. B. and Sands M., *The Feynman Lectures on Physics (vol. 3)*. (Addison-Wesley Publishing Company).
- [15] Aspect A. and Grangier P., Experiments on Einstein-Podolsky-Rosen type correlations with pairs of visible photons, *Quantum Concepts in Space and Time* (ed. R. Penrose and C. J. Isham, Oxford University Press, 1986).

- [16] Currie D. G., Jordan T. F. and Sudarshan E. C. G., *Rev. Mod. Phys.* **35**, 350 (1963).
- [17] Jackson J. D., *Classical Electrodynamics*, (John Wiley and Sons Inc.).
- [18] Jauch J. M. and Rohrlich F., *The Theory of Photons and Electrons*, (Springer-Verlag, 1976).
- [19] Itzykson C. and Zuber J.-B., *Quantum Field Theory*. (McGraw-Hill Book Company)